PWA Model Selection using a genetic algorithm

Stephan Schmeing,
Sebastian Neubert,
Karl Bicker

September 19th 2013
School on Concepts of Modern Amplitude Analysis Techniques
Flecken-Zechlin, Germany

Outline

Introduction: Partial-Wave Analysis at COMPASS

Genetic Algorithm for Model Selection

First Results

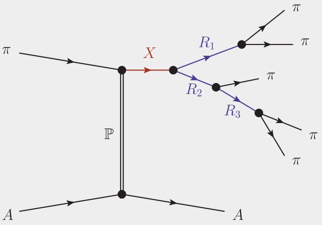
Conclusion

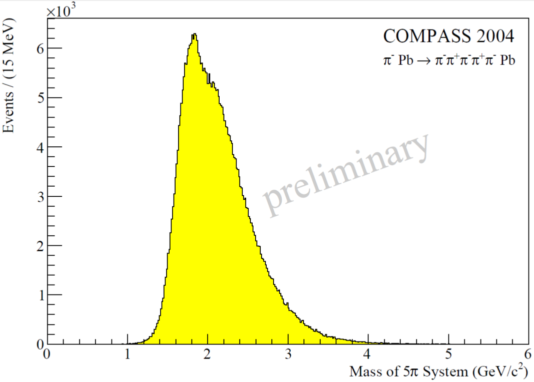
Outlook

Partial-Wave Analysis at COMPASS

Motivation

Diffractive Dissociation in 5 π





No "bump hunting" by eye possible $\Rightarrow$ Partial-Wave Analysis (PWA): J^P-Decomposition of mass spectrum using angular distribution

Partial-Wave Analysis at COMPASS

Partial-Wave Analysis

Partial-Wave Analysis

- Parametrisation of cross section (simplified):

$$\sigma(\tau) = \sigma_0 \sum_{i,j}^{Waves} \rho_{ij}(m_X)\phi_i(\tau)\phi_j(\tau)^*$$

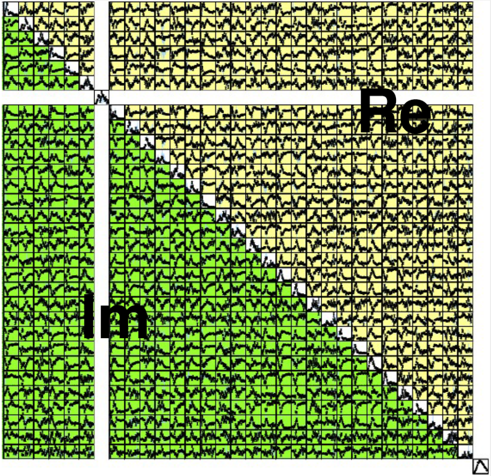
- 11-dimensional maximum likelihood fit to experimental kinematic distributions
- For the calculation the sum over all waves has to be truncated
- Truncation introduces systematic errors
- Optimal model for truncation has to be found

Partial-Wave Analysis at COMPASS

Partial-Wave Analysis

Technische Universität München

Spin Density Matrix



Partial-Wave Analysis at COMPASS

Partial-Wave Analysis

Model Requirements

1. The model should describe the data well
2. The number of parameters should be as small as possible
3. Correlations between parameters should be minimal

Partial-Wave Analysis at COMPASS

Partial-Wave Analysis

Technische Universität München

Model Selection

- Traditional way:
  - Compare log(likelihood) for different truncations
  - Use physical arguments and preexisting knowledge
  - Trial and error
- Introduces bias
- Has no methodical handle on systematic errors
- Too many possibilities for 5 π case

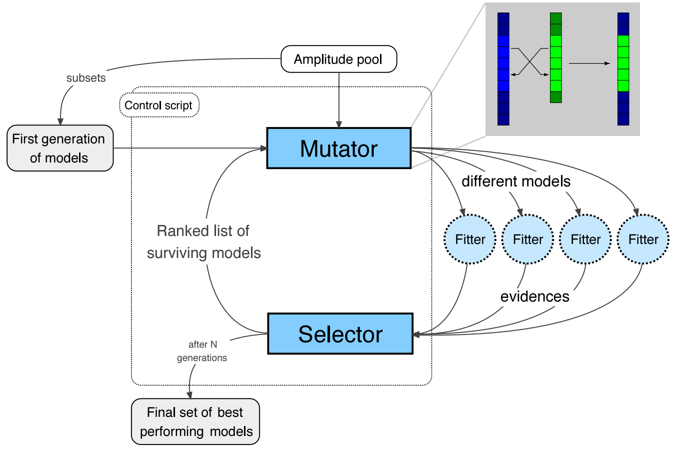
$\Rightarrow$ Use an algorithm for model selection

Genetic Algorithm for Model Selection

Principle

Technische Universität München

Working Principle



Genetic Algorithm for Model Selection

Ranking Criterion

Goodness-of-Fit Criterion

- Log(likelihood) alone cannot be used to quantify model quality, since more parameters tend to give better log(likelihood)
- Use Bayes' theorem to judge model quality

$$\begin{array}{ccccc} \text{Evidence} & \approx & \text{Best fit likelihood} & \cdot & \text{Occam factor} \\ P(Data|M_k) & \approx & P(Data|A^k_{ML}, M_k) & \cdot & P(A^k_{ML}|M_k)\,\sigma_{A^k|Data} \end{array}$$

- Additional factor to supress small waves with large errors is introduced
- A number of approximations are needed to calculate this

Genetic Algorithm for Model Selection

Optimization of Algorithm

Optimization Criteria

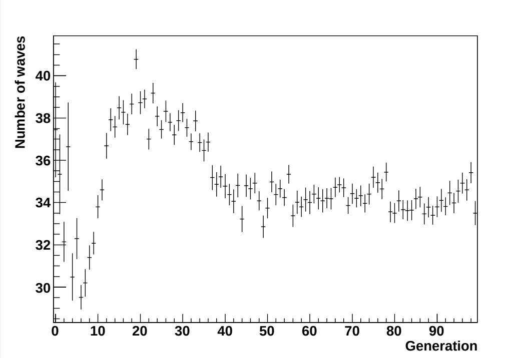
1. Full search space has to be explored: After a short starting phase the average evidence should fluctuate around constant well below maximum evidence
2. As few as possible created models should be invalid (for example due to not converging fits)
3. Final result should be (close to) optimal solution: Manually improvement should not be possible

First Results

Conditions

- Data from COMPASS 2004 hadron pilot run
- Use a pool of 284 Waves
- Run 100 generations with 50 models each

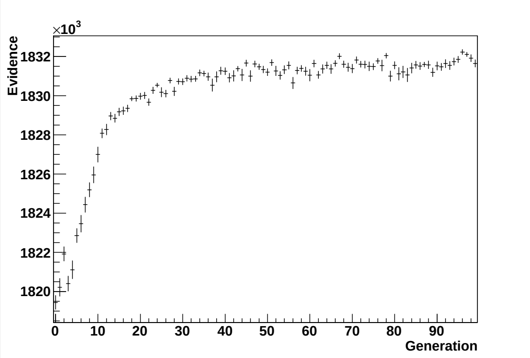
First Results



- Waveset size optimizes around 34 waves
- Finally chosen waveset contains 31 waves

First Results

First Results



- Average evidence varies slightly around constant not far from maximum($1834 \cdot 10^3$)
- Currently only between 16 and 32% of the created models are valid
- Simple manual breeding step can still increase final result

Conclusion

Conclusion

- A genetic algorithm for model selection has been implemented in the framework of the ROOTPWA toolkit: (http://sourceforge.net/projects/rootpwa/)
- A first partial-wave analysis using the algorithm has been performed
- The algorithm converged to a finite number of waves in the model
- High congruence between TOP 20 models

$\Rightarrow$ Goodness-of-Fit Criterion works

Outlook

Outlook

- Tuning of algorithm parameters and selection/mutation methods
- Tests of results with simulated dataset
- Transfer to other decay channels

Backup

PWA formula

$$\sigma(\tau) = \sigma_0 \sum_{\epsilon=-1}^{1} \sum_{i,j}^{Waves} \rho_{ij}^{\epsilon}(m_X)\phi_i^{\epsilon}(\tau)\phi_j^{\epsilon}(\tau)^*$$

Spin density matrix: $\rho_{ij}^{\epsilon}(m_X) = \sum_{r=1}^{Ranks} T_{ir}^{\epsilon} T_{jr}^{\epsilon *}$

Backup

Technische Universität München

Evidence

$$\ln P(Data|M_k) \approx \ln P(Data|A^k_{ML}, M_k) - \ln V^k_A + \ln \sqrt{(2\pi)^d |C_{A^k|Data}|} + \sum_a^{Waves} \ln S_a$$

$$\begin{aligned} \text{Dimension: } d &= \text{number of real parameters} \\ &= 2 \cdot \text{number of complex parameters} \end{aligned}$$

Significance: $S_a = \int\limits_{5\sigma_a}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x - |T_a|^2}{2\sigma_a^2}\right] dx$

Probability of the intensity of wave to be more than 5σ larger than zero

Parameter volume: $V^k_A = d \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)} r^{d-1}$

$(d-1)$-dimensional hypersphere with radius $r = \sqrt{N_{events}}$

(neglecting interference between waves)

Backup

Technische Universität München

Interpretation of Evidence

Evidence not normalised $\Rightarrow$ No absolut interpretation

Relative intepretation $\Rightarrow$ Bayes-Faktor: $B_{10} = \frac{P(Data|M_1)}{P(Data|M_0)}$

$\ln B_{10}$	B_{10}	Evidence
0 to 1	1 to 3	Not worth mentioning
1 to 3	3 to 20	Positive
3 to 5	20 to 150	Strong
≥ 5	≥ 150	Very strong