Binned ML fit PWA and etc.

Beijiang LIU

IHEP, CUHK & HKU

(for PWA working group)

BESIII PWA tutorials, Beijing

May. 25-27, 2009

Outline

- Motivation
- Binned ML fit
- Fit the resonance parameters
  - Include the resolution
- Summary and outlook

Motivation

Partial Wave Analysis is a powerful tool of insight in the spectrum of hadrons, in particular the finding of exotic states.

Issues in PWA: (see Prof.Jin's talk)

- Parameterizations of intermediate resonances/processes
- Including the resolutions
- Background modeling
- Correction of data/MC inconsistencies in PWA
- Minimization strategy / Finding the global minimum
- ....

- **Time consumption (in front of high statistics data)**

PWA: a multi dimensional fit to massive data

$$\frac{d\sigma}{d\Phi_n} = \frac{1}{2}\sum_{\mu=1}^{2} A^{\mu}A^{*\mu} = \frac{1}{2}\sum_{i,j}\Lambda_i\Lambda_j^* \sum_{\mu=1}^{2} U_i^{\mu}U_j^{*\mu} \equiv \sum_{i,j} P_{ij}\cdot F_{ij}$$

$$A = \psi_{\mu}(m)A^{\mu} = \psi_{\mu}(m)\sum_i \Lambda_i U_i^{\mu}$$

$$P_{ij} = P_{ji}^* = \Lambda_i\Lambda_j^*,$$

$$F_{ij} = F_{ji}^* = \frac{1}{2}\sum_{\mu=1}^{2} U_i^{\mu}U_j^{*\mu}.$$

$$P(\xi : \alpha) = \frac{\omega(\xi,\alpha)\epsilon(\xi)}{\int d\xi\omega(\xi,\alpha)\epsilon(\xi)}$$

ω: differential cross section

$$\omega(\xi,\alpha) = \frac{d\sigma}{d\Phi}$$

ε: detection efficiency

ξ: a certain combination of four momentum of final state particles

α: parameters in amplitudes

$$\sigma = \int d\xi\omega(\xi)\epsilon(\xi)$$

$$\sigma = \frac{1}{N_{gen}}\sum_{k=1}^{N_{MC}}\left(\sum_{ij} PA_{ij}FU_{ij}(\xi_k)\right) = \sum_{ij} PA_{ij}\left[\frac{1}{N_{gen}}\sum_{k=1}^{N_{MC}} FU_{ij}(\xi_k)\right]$$

$$P(\xi_1,\xi_2,\ldots,\xi_n : \alpha) = \prod_{i=1}^{n} P(\xi_i : \alpha) = \prod_{i=1}^{n}\frac{\omega(\xi_i,\alpha)\epsilon(\xi_i)}{\int d\xi\omega(\xi,\alpha)\epsilon(\xi)}$$

$$\ln P(\xi_1,\xi_2,\ldots,\xi_n : \alpha) = \sum_{i=1}^{n}\ln\left(\frac{\omega(\xi_i,\alpha)}{\int d\xi\omega(\xi,\alpha)\epsilon(\xi)}\right) + \sum_{i=1}^{n}\ln\epsilon(\xi_i)$$

$$\ln\mathcal{L} = \sum_{i=1}^{n}\ln\left(\frac{d\sigma}{d\Phi}/\sigma\right)$$

PWA: a multi dimensional fit to massive data

Computation Demands:

the likelihood calculation

Most memory consuming

Coupling parameters (determined by fit)

Partial wave amplitudes (precalculated)

$$\ln L = \sum_{n=1}^{N_{data}} \ln\left(\frac{d\sigma}{d\Phi} / \sigma\right) = \sum_{n=1}^{N_{data}} \ln\left(\frac{\sum_{i,j} PA_{ij} \cdot FU_{ij}}{\sum_{i,j} PA_{ij} \left(\frac{1}{N_{gen}} \sum_{k=1}^{N_{MC}} FU_{ij}\right)}\right)$$

$N_{data} * N_{waves}^2$

Sum over data events

Independent of fit parameters (precalculated)

$N_{iter} * N_{data} * N_{waves}^2$

Most time consuming

$N_{MC} * N_{waves}^2$

More on Computation Demands:

At BESII era, the computation demands of a full PWA:

- ~O(1000) CPU days
- One fit takes several minutes. (@ a P4 PC)
- A lot of fits (~O(10^5)) are needed to complete an analysis.

 ('global' minimum finding; scan of resonance parameters; determination of significance, statistical error, systematic error)

At BESIII, ~O(100) higher statistics than BESII.

Be aware of that the jobs are strongly coupled.

 (you have to wait for the fitting results to decide your next move.)

- You need **a lot of** computers.
- You also need do analysis with **short turnover times**

How to make the PWA faster

Reduce the computation load

- **Precalculation (**store things will be used for next calculation**)**
- **Binned fit**
- **Fit the resonance parameters**

  } Will be discussed here
- **Improved 'global' minimum finding**
- **....**

Improve the speed

- **raw computation power**
- **I/O performance**
- **cache coherence**
- **memory hierarchy**
- **....**

Speeding up with parallelism

- (For an "unbinned" "mass-dependent" PWA) The differential cross sections of all the events are needed to build the likelihood.

  You can NOT split the job as MC production

- Events are independent.

  You can calculate the differential cross section of each event in parallel: "Data parallelism"

  GPUPWA: a PWA framework based on GPGPU, see Nik's talk

- Some of the fits can be done in the same time.

  You can run individual fits on a cluster (a group of PCs loosely coupled): "Task parallelism"

Maximum Likelihood fits

Unbinned extended ML

$$f(x_i;\mathbf{p})/A(\mathbf{p})$$

$$\mathcal{L}_u(\mathbf{p}) = \frac{A(\mathbf{p})^N e^{-A(\mathbf{p})}}{N!} \prod_{i=1}^{N} \frac{f(x_i;\mathbf{p})}{A(\mathbf{p})} .$$

$$\mathcal{F}_u(\mathbf{p}) = -\sum_{i=1}^{N} \ln f(x_i;\mathbf{p}) + A(\mathbf{p})$$

Unbinned ML

$$f'(x;\mathbf{p}) \equiv f(x;\mathbf{p})/A(\mathbf{p})$$

$$\mathcal{F}'_u(\mathbf{p}) = -\sum_{i=1}^{N} \ln f'(x_i;\mathbf{p}) .$$

Binned ML

$$\mathcal{L}_b(\mathbf{p}) = \prod_{b=1}^{M} \frac{f_b^{n_b} e^{-f_b}}{n_b!}$$

$$\mathcal{F}_b(\mathbf{p}) \equiv -\ln \mathcal{L}_b(\mathbf{p}) = -\sum_{b=1}^{M} (n_b \ln f_b - f_b - \ln n_b!) .$$

RECAP: the likelihood calculation

How many inputs are needed to retain all the information in data?

$$\ln L = \sum_{n=1}^{N_{data}} \ln\left(\frac{d\sigma}{d\Phi}/\sigma\right) = \sum_{n=1}^{N_{data}} \ln\left(\frac{\sum_{i,j} PA_{ij} \cdot FU_{ij}}{\sum_{i,j} PA_{ij}\left(\frac{1}{N_{gen}} \sum_{k=1}^{N_{MC}} FU_{ij}\right)}\right)$$

What we will care:

How many inputs are needed to extract the physics results?

Time consumption:

Unbinned ML fit, T $\propto$ O(N_{data})

Binned ML fit, T $\propto$ O(N_{bin})

Binned ML fit

$$L = \prod_{b=1}^{N_{bin}} \frac{\mu_b^{\,n_b} e^{-\mu_b}}{n_b\,!},$$

where n_b, observed bin contents

$$\mu_b = \frac{\int_{bin} \varepsilon(\xi)\varpi(\alpha;\xi)d\xi}{\int \varepsilon(\xi)\varpi(\alpha;\xi)d\xi} N_{data},$$ expected bin contents

Object function:

$$-0.5\chi^2 \equiv -\ln L(\alpha) = -\sum_{b=1}^{N_{bin}} (n_b \ln \mu_b - \mu_b - \ln n_b\,!), \text{ drop } \ln n_b\,!,$$

$$-\ln L(\alpha) = -\sum_{b=1}^{N_{bin}} (n_b \ln \mu_b - \mu_b)$$

if $\sum_{b=1}^{N_{bin}} \mu_b$ is independent of the fit parameters (in our case $\sum_{b=1}^{N_{bin}} \mu_b = \mathrm{N}_{\mathrm{data}}$),

$$-\ln L(\alpha) = -\sum_{b=1}^{N_{bin}} n_b \ln \mu_b$$

Binned ML fit

$$L = \prod_{b=1}^{N_{bin}} \frac{{\mu_b}^{n_b} e^{-\mu_b}}{n_b\,!},$$

where $n_b = \int_{bin} \sum_{i=1}^{N_{data}} \delta(\xi - \xi_i) d\xi,$ observed bin contents

$\mu_b = \frac{\int \varepsilon(\xi)\varpi(\alpha;\xi)d\xi}{\int \varepsilon(\xi)\varpi(\alpha;\xi)d\xi} N_{data},$ expected bin contents

Object function:

$$-\ln L(\alpha) = -\sum_{b=1}^{N_{bin}} (n_b \ln \mu_b - \mu_b - \ln n_b\,!)$$

$$-\ln L(\alpha) = -\sum_{b=1}^{N_{bin}} (n_b \ln \mu_b - \mu_b)$$

if $\sum_{b=1}^{N_{bin}} \mu_b$ is independent of the fit parameters (in our case $\sum_{b=1}^{N_{bin}} \mu_b = N_{data}$),

$$-\ln L(\alpha) = -\sum_{b=1}^{N_{bin}} n_b \ln \mu_b$$

With a Poisson distribution for the number of events in each bin,

the likelihood could be defined with the observed bin contents and expected bin contents.

Binned ML fit

$$\mu_b = \frac{\int_{bin} \varepsilon(\xi)\varpi(\alpha;\xi)d\xi}{\int \varepsilon(\xi)\varpi(\alpha;\xi)d\xi} N_{data}, \quad \varpi(\alpha;\xi) = \frac{d\sigma}{d\Phi} = \sum_{ij} PA_{ij} FU_{ij}$$

$$\int \varepsilon(\xi)\varpi(\alpha;\xi)d\xi \simeq \frac{1}{N_{gen}} \sum_{i}^{N_{gen}} N_{gen} \Delta\xi_i \varepsilon(\xi_i)\varpi(\alpha;\xi_i) = \frac{1}{N_{gen}} \sum_{k}^{N_{acc}} \varpi(\alpha;\xi_k)$$

$$= \frac{1}{N_{gen}} \sum_{k}^{N_{acc}} \sum_{ij} PA_{ij} FU_{ij} = \frac{1}{N_{gen}} \sum_{ij} PA_{ij} \sum_{k}^{N_{acc}} FU_{ij}$$

$$\mu_b = \frac{\frac{1}{N_{gen}} \sum_{ij} PA_{ij} \sum_{bin}^{N_{acc}} FU_{ij}}{\frac{1}{N_{gen}} \sum_{ij} PA_{ij} \sum_{phsp}^{N_{acc}} FU_{ij}} N_{data} = \frac{\sum_{ij} PA_{ij} \sum_{bin}^{N_{acc}} FU_{ij}}{\sum_{ij} PA_{ij} \sum_{phsp}^{N_{acc}} FU_{ij}} N_{data}$$

Binned ML fit

$$\mu_b = \frac{\int_{bin} \varepsilon(\xi)\varpi(\alpha;\xi)d\xi}{\int \varepsilon(\xi)\varpi(\alpha;\xi)d\xi} N_{data}, \quad \varpi(\alpha;\xi) = \frac{d\sigma}{d\Phi} = \sum_{ij} PA_{ij}FU_{ij}$$

The expected bin contents could be obtained by Monte-Carlo integrations.

The parts independent of fit parameters can be precalculated and cached.

$$\mu_b = \frac{\frac{1}{N_{gen}}\sum_{ij} PA_{ij} \sum_{bin}^{N_{acc}} FU_{ij}}{\frac{1}{N_{gen}}\sum_{ij} PA_{ij} \sum_{phsp}^{N_{acc}} FU_{ij}} N_{data} = \frac{\sum_{ij} PA_{ij} \sum_{bin}^{N_{acc}} FU_{ij}}{\sum_{ij} PA_{ij} \sum_{phsp}^{N_{acc}} FU_{ij}} N_{data}$$

Binned ML fit

individual lnL:

$\ln L_b = n_b \ln \mu_b$, the bins $n_b = 0$ or $\mu_b = 0$ are dropped

individual gradient:

$$\ln L_b' = n_b \frac{\mu_b'}{\mu_b},$$

individual hessian:

$$\ln L_b'' = -\frac{n_b(\mu_b')^2}{\mu_b^2} + \frac{n_b \mu_b''}{\mu_b} = -\frac{n_b(\mu_b')^2}{\mu_b^2}$$

μ_b' can be easily written down analytically.

$\mu_b'' \approx 0$, $\ln L_b''$ can be calculated with μ_b'

Binned ML fit

Note:

- We do NOT use $\ln L_b = n_b * \ln(p_{\text{bin center}})$, which assumed the approximation of "p.d.f is constant in each bin".
- Arbitrary binning is allowed.
- Detailed information (smaller than bin size) will be lost. Different binning methods (bin size, dimension) will give different sensitivities.
- But binned ML fit is a rigorous procedure in statistics. Binned ML fit should give correct fitting results. mean value, error, significance, etc.

A demo analysis of $J/\psi \to \gamma K^+K^-$

- **Data sample: 500000 'data' events**

The MC truth contains

a 2^+ resonance (M=2.00 GeV/c^2 Γ=0.13 GeV/c^2)

a 0^+ resonance (M=2.15 GeV/c^2 Γ=0.05 GeV/c^2)

and an uniform phase space.

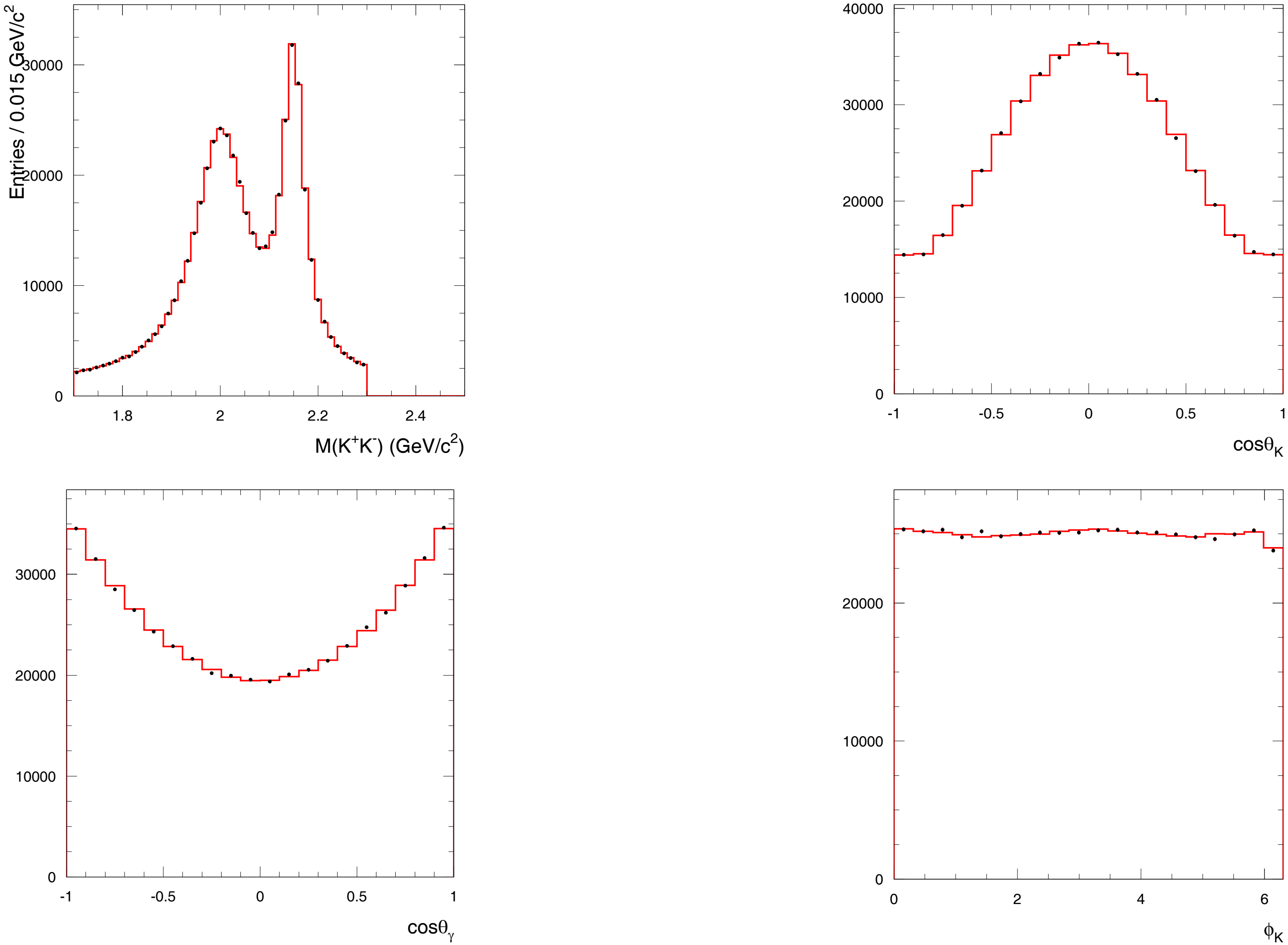
As well as 10^7 PS events for MC integrations

- **Binning: for this 3-body decay, using "square" Dalitz coordinates (4 dimensions)**

$M(K^+K^-)$, $\cos\theta_K$', ϕ_K', $\cos\theta_\gamma$ (40*20*10*10=80000 bins)

***Note: Arbitrary binning is allowed**

Projections



Fitting validation (1)

Fraction of components:

- 0^+ : 0.3099±0.0011 (input=0.310)
- 2^+ : 0.5839±0.0017 (input=0.585)

Resonance parameters:

- Mass of 0^+ : 2.1500±0.0001 (input=2.150)
- Width of 0^+ : 0.0501±0.0003 (input=0.050)
- Mass of 2^+ : 2.0000±0.0002 (input=2.000)
- Width of 2^+ : 0.1297±0.0006 (input=0.130)

Mean value and errors are obtained from 97x500000 events. (Thx to Ms. HuangYP)

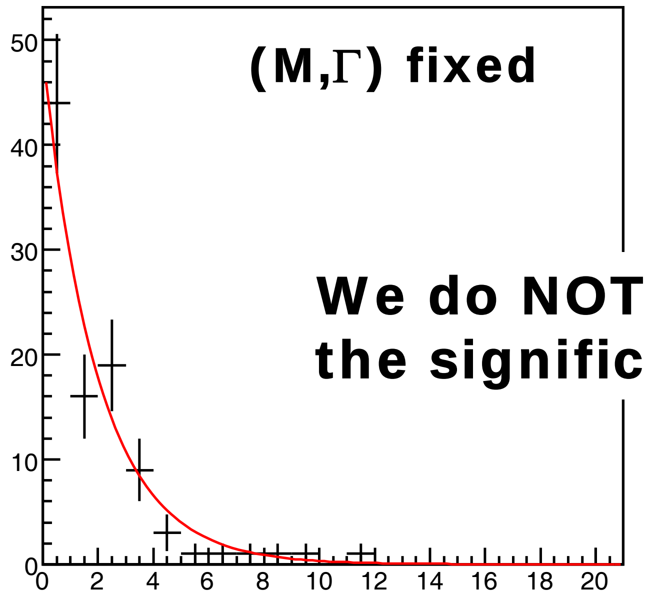
Fitting validation (2)

Hypothesis tests for significance:

- **include a fake 0^+ in the fit**
- **$\Delta\chi^2 \equiv -2(\Delta \ln L)$;**

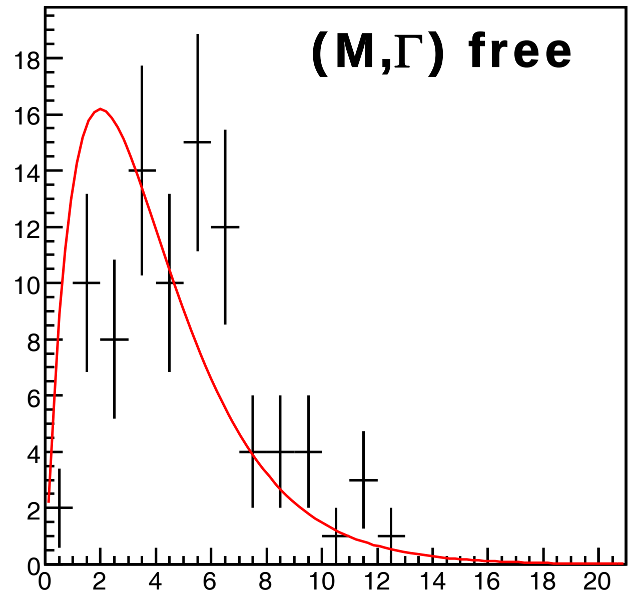
Confidence level: **$\alpha = PROB(\Delta\chi^2, \Delta\ n.o.f)$;**

Gaussian expression of significance: **$GAUSIN(1-\alpha/2)$.**



$\Delta\chi^2$ distribution (Δ n.o.f=2)

We do NOT over estimate the significance



$\Delta\chi^2$ distribution (Δ n.o.f=4)

Results yielded by unbinned fit

99x50000 events

Fraction of components:

- 0^+ : 0.3107±0.0029
  (scaled to 500000: ±0.0009)
- 2^+ : 0.5838±0.0046
  (scaled to 500000: ±0.0015)

Resonance parameters:

- Mass of 0^+ : 2.1498±0.0003
  (scaled to 500000: ±0.0001)
- Width of 0^+ : 0.0502±0.0009
  (scaled to 500000: ±0.0003)
- Mass of 2^+ : 1.9998±0.0006
  (scaled to 500000: ±0.0002)
- Width of 2^+ : 0.1297±0.0016
  (scaled to 500000: ±0.0005)

Binned ML fit

Fraction of components:

- 0^+ : 0.3099±0.0011
- 2^+ : 0.5839±0.0017

Resonance parameters:

- Mass of 0^+ : 2.1500±0.0001
- Width of 0^+ : 0.0501±0.0003
- Mass of 2^+ : 2.0000±0.0002
- Width of 2^+ : 0.1297±0.0006

We do NOT lose too much sensitivity

Timing

	BESII	Binned ML fit
Data set	Data: 73000 events MC: 10^6 events	Data: 72758 bins/500000 events MC: 10^7 events
MC Integration	42 s	350 s
LUT	4 s	
Fit	48 s /16 iter	55 s /13 iter

Recap

- The fitting results consist with inputs.

- The **results**, **errors** and **significance** yielded by **binned fit** consist with those of **unbinned fit**.

- The **time** consumption is reduced to $O(N_{bin})$, even though with large data sets.

Fit the resonance parameters

At BES2, the resonance parameters (mass and width) are fixed in the fitting procedure.

The resonance parameters are determined by scan.

- **One dimension scan**
  - **Vary one parameter to a certain value while keeping others fixed.**
  - **Repeat the fitting (optimization) procedure.**
- **Scan the resonance parameters one by one.**
- **do the above 2 steps iteratively, until the –lnL steady.**

*** One dimension scan loops ≠ Grid search**

Fit the resonance parameters

Motivation:

- **Less #fits than brutal scan.**
- **Mathematically better than one-dimension scan loops.**

$$\ln L = \sum_{n=1}^{N_{data}} \ln\left(\frac{d\sigma}{d\Phi} / \sigma\right) = \sum_{n=1}^{N_{data}} \ln\left(\frac{\sum_{i,j} PA_{ij} \cdot FU_{ij}}{\sum_{i,j} PA_{ij} \left(\frac{1}{N_{gen}} \sum_{k=1}^{N_{MC}} FU_{ij}\right)}\right)$$

$N_{iter} * N_{waves}^{2}$

To fit the resonance parameters,

MC integration have to be recalculate in every fit iteration

Can't afford this @BES2

$$\ln L = \sum_{n=1}^{N_{data}} \ln\left(\frac{d\sigma}{d\Phi} / \sigma\right) = \sum_{n=1}^{N_{data}} \ln\left(\frac{\sum_{i,j} PA_{ij} \cdot FU_{ij}}{\frac{1}{N_{gen}} \sum_{k=1}^{N_{MC}} \sum_{i,j} PA_{ij} \cdot FU_{ij}}\right)$$

$N_{iter} * N_{MC} * N_{waves}^{2}$

Fit the resonance parameters

A bonus of binned fit

- In Dalitz coordinates, with fine binning in Daltiz plane, for BW(s), s is approximately constant in a bin.

- Hence, for MC, we can cache the amplitudes (FU_{ij}) at the Dalitz plane bin level (~$O(10^2)$).

- The additional computation and I/O loads are trivial. We can fit the resonance parameter without a significant draw-back of speed.

Fit the resonance parameters

$$\frac{d\sigma}{d\Phi} = -\frac{1}{2}\sum_{\mu=1}^{2} A_{\mu\nu} g_{\nu\nu'}^{(\perp\perp)} A^{\mu\nu'} = \sum_{ij} P_{ij} F_{ij},$$

$$P_{ij} = P_{ji}^{*} = \Lambda_i \Lambda_j^{*}, \Lambda = ae^{ib} = a\cos b + i * a\sin b,$$

$$F_{ij} = -\frac{1}{2}\sum_{\mu=1}^{2} U_i^{\mu\nu} g_{\nu\nu'}^{(\perp\perp)} U_j^{*\mu\nu'}, U_i = BW_i(s) * T_i,$$

move BW from U into Λ, redefine $\Lambda_i \equiv \Lambda_i * BW_i(s)$, $U_i \equiv T_i$.

then, for each wave there're 4 parameters:

coupling coefficient a,b; resonance parameters m,w;

in the procedure of minimization,

F_{ij} is cached, P_{ij} will be yielded by fit.

calculate $\frac{\partial P_{ij}}{\partial a}, \frac{\partial P_{ij}}{\partial b}, \frac{\partial P_{ij}}{\partial m}, \frac{\partial P_{ij}}{\partial w}$ with $\left.\frac{\partial F(x)}{\partial x}\right|_{x_0} = \frac{F(x_0 + d) - F(x_0)}{d}$.

Fit the resonance parameters

$$\sigma = \int \varepsilon(\xi)\varpi(\alpha;\xi)d\xi$$

$$\simeq \frac{1}{N_{gen}} \sum_{i}^{N_{gen}} N_{gen}\Delta\xi_i \varepsilon(\xi_i)\varpi(\alpha;\xi_i) = \frac{1}{N_{gen}} \sum_{k}^{N_{acc}} \varpi(\alpha;\xi_k)$$

$$= \frac{1}{N_{gen}} \sum_{k}^{N_{acc}} \sum_{ij} PA_{ij} FU_{ij}$$

$$= \frac{1}{N_{gen}} \sum_{l}^{N_{dalitzbin}} \sum_{ij} PD_{ij} \sum_{dalitzbin}^{N_{acc}} FU_{ij}$$

$$\int_{bin} \varepsilon(\xi)\varpi(\alpha;\xi)d\xi = \frac{1}{N_{gen}} \sum_{bin}^{N_{acc}} \sum_{ij} PA_{ij} FU_{ij} = \frac{1}{N_{gen}} \sum_{ij} PB_{ij} \sum_{bin}^{N_{acc}} FU_{ij}$$

Fit the resonance parameters

$$\text{individual lnL: } \ln L_b = n_b \ln \mu_b = n_b \ln \frac{\sum_{ij} PB_{ij} \sum_{bin}^{N_{acc}} FU_{ij}}{\sum_{l}^{N_{dalitzbin}} \sum_{ij} PD_{ij} \sum_{dalitzbin}^{N_{acc}} FU_{ij}} N_{data}.$$

$$\text{individual gradient: } \ln L_b' = n_b \frac{\mu_b'}{\mu_b},$$

$$\mu_b' = \frac{(\sum_{ij} PB_{ij} \sum_{bin}^{N_{acc}} FU_{ij})'}{\sum_{l}^{N_{dalitzbin}} \sum_{ij} PD_{ij} \sum_{dalitzbin}^{N_{acc}} FU_{ij}} - \frac{(\sum_{ij} PB_{ij} \sum_{bin}^{N_{acc}} FU_{ij}) * (\sum_{l}^{N_{dalitzbin}} \sum_{ij} PD_{ij} \sum_{dalitzbin}^{N_{acc}} FU_{ij})'}{(\sum_{l}^{N_{dalitzbin}} \sum_{ij} PD_{ij} \sum_{dalitzbin}^{N_{acc}} FU_{ij})^2},$$

$$(\sum_{ij} PB_{ij} \sum_{bin}^{N_{acc}} FU_{ij})' = \sum_{ij} (PB_{ij})' \sum_{bin}^{N_{acc}} FU_{ij},$$

$$(\sum_{l}^{N_{dalitzbin}} \sum_{ij} PD_{ij} \sum_{dalitzbin}^{N_{acc}} FU_{ij})' = \sum_{l}^{N_{dalitzbin}} \sum_{ij} (PD_{ij})' \sum_{dalitzbin}^{N_{acc}} FU_{ij}$$

Fitting results

		input	fit	scan
0^+	fraction	0.310 $\pm$0.0009	0.310	0.3099
	mass	2.15 $\pm$0.0001	2.1500	2.1500
	width	0.05 $\pm$0.0003	0.0506	0.0501
2^+	fraction	0.585 $\pm$0.0015	0.584	0.5839
	mass	2.00 $\pm$0.0002	2.0002	2.0000
	width	0.13 $\pm$0.0005	0.1305	0.1297

Timing results

Fitting ranges are rather large

for 0^+: (2.0, 2.3), (0.01, 0.09)

for 2^+: (1.9, 2.2), (0.06, 0.2)

	Fit (s) (5+4 parameters)	Scan (s) (5 parameters)
LUT generation (MC Integration)	341*1	424 *(#scan points)
Fit x 50 times to find 'global minimum'	6240*1	1643*(104 scan points)=170872
Fit (with a fake 0^+) X 50 times	21174*1	3082*(156 scan points)=480792

Include the resolution (under working)

What is resolution?

$p4_{physics}$->Detector->…->Reconstruction

->$p4_{reconstructed}$

Why need you include the resolution?

Your amplitudes are written in $p4_{physics}$.

Your inputs are $p4_{reconstructed}$.

When the resolution is comparable to the size of structure, it is not negligible.

How to include the resolution in the fitting?

Convolution:

$pdf(p4_{physics}) \otimes resolution = pdf(p4_{reconstructed})$

$$f(u) = \int_{-\infty}^{\infty} f_x(x) f_y(u - x)\, dx = \int_{-\infty}^{\infty} f_y(y) f_x(u - y)\, dy \ .$$

Include the resolution (under working)

Hints from 1D fit:

To include the resolution when fitting a mass spectrum:

- **Analytical convolution, voigtian≡BW ⊗ Gausian**

- **Histogram pdf: the shape of simulated data, which already includes the resolution.**

  - **This is a sort of numerical convolution integral by MC sampling**

Include the resolution (under working)

As far as histogram (binned fit) is concerned, what is resolution?

- The events in a bin will be smeared to the neighborhood bins.
- The effective expected number events of a bin is the sum of number of events which fall in this bin

x_i is the collection of events in bin_i

pdf: the probability a event falls in bin_i

$$f(x_i;p) \equiv f(x_i;p)/A(p)$$

- smeared pdf: x_i is smeared to $r^j_i * x_j$ (j=i- cutoff, ..., i+ cutoff)

$$f(x_i;p) \otimes resolution = \Sigma\, r^j_i * f(x_i;p)$$

- the effective pdf for bin_i :

$$F(x_i;p) = \Sigma\, r^i_j * f(x_j;p)$$

Include the resolution (under working)

To-do:

- **We already have $f(x_i;p)$ (a.k.a PA*FU)**
- **Get the response matrix r^j_i of each bin.**
- **Calculate $F(x_i;p)=\Sigma\ r^i_j * f(x_j;p)$ to be used in the fitting.**

Include the resolution (under working)

How to get the r^j_i :

- **Use the MC Truth of reconstructed simulation events.**
- **Do the binning of the MC Truth PHSP and the binning of the reconstructed PHSP. Map it out for the r^j_i.**
- **r^j_i depends only on the detector properties, it does NOT depend on the expected result.**
- **This procedure is NOT a simple 'data correction'. It is to build the response matrix.**
- **(De)convolution is a complex mathematical operation and requires a good understanding of the detector.**

Summary and Outlook

- A binned ML fit approach of PWA is established.

- From timing results and preliminary results of fitting validation, the binned ML fit method looks promising.

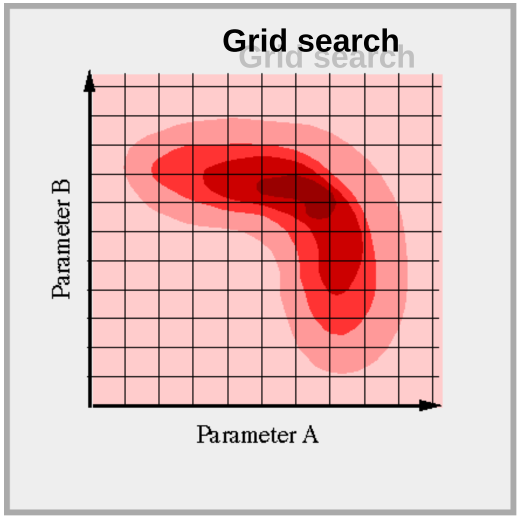
- Binned ML fit provides several bonus: making it easy to fit resonance parameters, to include resolutions (under working), ….

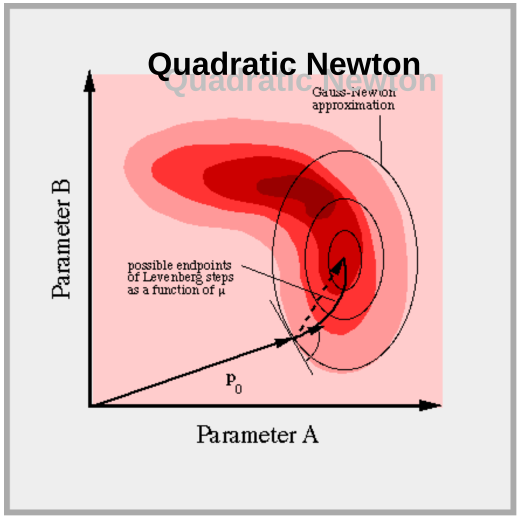
- Technically, the binned ML fit could be combined with other approaches (e.g. GPUPWA) smoothly.

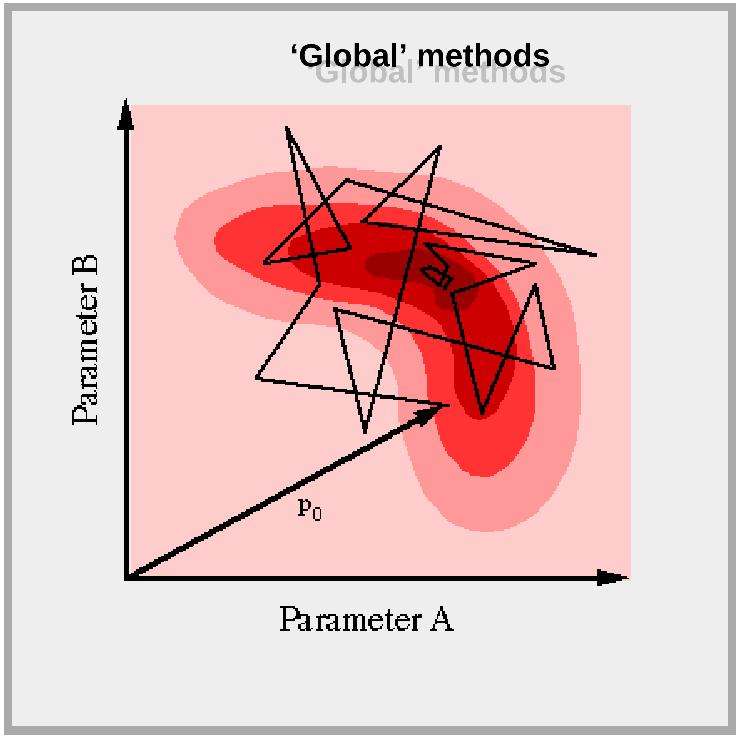
- Besides the time consumption, there're several issues (even not so urgent) still need to be solved.

Thank you

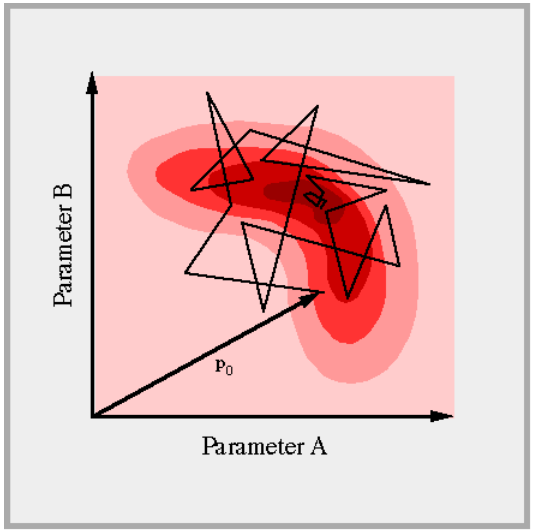
More thoughts about minimize

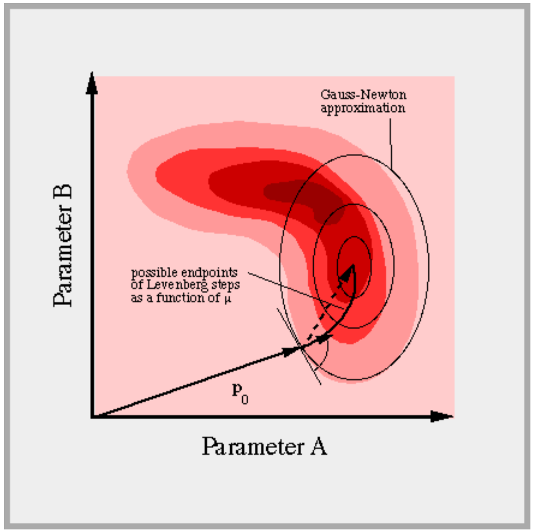






Finder + Converger





- **Global** optimisers :
  - **Grid search, Monte Carlo sampling, Random walk, Genetic Algorithm, Simulated Annealing, ...**
  - **Good global minimum finder**, but maybe **poor accuracy**

- Default solution in HEP:

  [How much longer do we need to suffer ?]
  - Gradient-driven search, using variable metric, can use quadratic Newton-type solution
  - **Poor global minimum finder**, gets quickly stuck in presence of **local minima**