Kaon LT analysis discussion

Ioana & Gabriel Niculescu,
James Madison University

Introduction

As the JMU analysis progressed toward getting Σ^0/Λ ratios...

- We looked at the next logical analysis steps
- Realized that we might have an incomplete grasp of what these steps entail.
- As we dug a little deeper, we convinced ourselves that, OK, we do “get” the workflow (kinda!)
- However, in the process, (think that) we uncovered some potential “issues” that might affect the analysis and would like to share these with you

Information Sources

We’ve looked (in no particular order) at:

- G.H: Summary of L/T/LT/TT-separation... (Aug, 2022)
- H.B. it et al: PRC 78 (2008)
- **hcana**, **simc** code
- PhD theses: R. Mohring (1999), J. Volmer (2000), T. Horn (2006); V.Klimenko (2024) - Hall B! others...
- several statistics/data analysis books & papers (geared for our field):
  - F. James: “Statistical Methods In Experimental Physics (2nd ed)”
  - G. Cowan: “Statistical Data Analysis”
  - L. Lista: “Statistical Methods for Data Analysis: With Applications in Particle Physics (Lecture Notes in Physics)”
  - L. Lyons: “Statistics for nuclear and particle physicists”
  - W.J.Metzger: “Statistical Methods in Data Analysis”

Here is a brief list w/ the topics that might require attention

Due to time constraints, we will not be able to cover all today...

- origin and evolution of the "Monte Carlo ratio" (mcr) method.
- $\bar{f}(x)$ vs $f(\bar{x})$ (among other things) (I)
- bin migration. do we monitor/account for/care? (II)
- the pitfalls of bin–by–bin correction factors (III)
- (check the) handling statistical uncertainties (IV)
- recommended/proper way of handling of systematic uncertainties (bootstrapping?) (V)
- is the final result a cross–section or a comparison w/ a/some model(s)? (VI)

I:

afaikt...

- the crux of the **mcr** method* is this equation:

$$\sigma_{exp}(\bar{W}, \bar{Q}^2, t, \phi; \bar{\theta}, \bar{\epsilon}) = \frac{< Y_{exp} >}{< Y_{sim} >} \sigma_{MC}(\bar{W}, \bar{Q}^2, t, \phi; \bar{\theta}, \bar{\epsilon}) \quad (1)$$

- not 100% sure, but it seems this first appeared in JV's thesis (see next page):

J. Volmer PhD thesis (2000), page 81:

Determination of the model cross section

For the purpose of extracting cross sections, the data are binned in t, but integrated over W, Q^2 and θ^*. This integration is complicated because of the range covered in these quantities (see for instance Fig. 5.1) and the correlations between them. To make matters more complicated, only three of the four quantities W, Q^2, t and θ^* are independent. If the dependence of the cross section on these variables is poorly understood, problems will arise in determining the (average) cross section and also in the L/T separation. The problems are that the cross section has to be averaged over the phase space of the t bin, and that in order to do the L/T separation, the same values for W, Q^2 and t have to be used in the high and low ϵ data. If the average values of these quantities in a given t bin differ for both data sets, the cross sections must be scaled towards common values of W and Q^2. When using an imperfect model, the scaling error will influence the high and low ϵ data differently and thus influence the extraction of σ_L and σ_T. Also, the interference terms have to be modelled well, because the ϕ acceptance is not homogeneous (see right hand plot in Fig. 5.1 and leftmost plot in Fig. 5.7).

Both problems are alleviated by creating a model such that the ratio of experimental to simulated yields does not depend (or hardly depends) on W, Q^2, t, θ^* and ϕ. If the model meets this criterium, it is safe to use

$$\left(\frac{d\sigma}{dt}\right)_{\rm exp} = \frac{Y_{\rm exp}}{Y_{\rm sim}} \left(\frac{d\sigma}{dt}\right)_{\rm MC} \qquad (5.7)$$

for the extraction of unseparated experimental cross sections at any value of W, Q^2 and t. In this way, common values of W and Q^2 can be chosen freely for the high and low ϵ bins in order to make an L/T separation possible. The iterative fitting procedure used to create the model for the cross section is described below.

??

- ...seems to be in "response"/follow–up from Rick's thesis (see next page)
- contrast between "imperfect" model vs "model"
- if the better model exists, why not use it for scaling?
- **any** point?

Rick Mohring's PhD thesis (1999), page 140:

$$Y_{\rm MC} = L_H \times \int \left[\Gamma(Q^2, W) \left(\frac{d^2\sigma}{d\Omega^*_K} \right) \right] A(d^5V)\, R(d^5V)\, dQ^2 dW\, d\phi_e d\Omega^*_K . \quad (4.6)$$

Note that the general cross section can be written in terms of the cross section at a given point (called the "scaling point") $(\langle Q^2 \rangle, \langle W \rangle)$ and $\theta_{CM} = 0^\circ$ as

$$\left(\frac{d^2\sigma}{d\Omega^*_K} \right) = \left(\frac{\overline{d^2\sigma}}{d\Omega^*_K} \right) \Bigg|_{(\langle Q^2 \rangle, \langle W \rangle, \theta_{CM}=0^\circ)} \times \left(\frac{f(Q^2, W, \theta_{CM})}{f(\langle Q^2 \rangle, \langle W \rangle, \theta_{CM} = 0^\circ)} \right) \quad (4.7)$$

Notes:

- highlights the need for having a "scaling function"
- integrated over the bin and compared w/ its value at the "scaling point"
- also (page 141) Rick attempts a genuine cross-section extraction (see next page)
- better*

Rick Mohring's PhD thesis, page 141:

$$Y_{\rm MC} = L_H \times \left(\overline{\frac{d^2\sigma}{d\Omega_K^*}}\right)\Bigg|_{(\langle Q^2\rangle,\langle W\rangle,\theta_{CM}=0^\circ)} \times \left[\left(\frac{(\text{Number of MC successes})_{weighted}}{\text{Number of MC tries}}\right) \times \Delta^5 V_{\rm gen}\right]. \quad (4.9)$$

Finally, by setting the equivalent yields of MC and data to be equal (i.e., $Y_{\rm MC} = Y_{\rm DATA}$), the cross section at the scaling point can be extracted using Equation 4.9 as:

$$\left(\overline{\frac{d^2\sigma}{d\Omega_K^*}}\right)\Bigg|_{(\langle Q^2\rangle,\langle W\rangle,\theta_{CM}=0^\circ)} = Y_{\rm DATA} \times \frac{1}{L_H} \times \left[\left(\frac{\text{Number of MC tries}}{(\text{Number of MC successes})_{weighted}}\right) \times \frac{1}{\Delta^5 V_{\rm gen}}\right]. \quad (4.10)$$

Notes:

- "acceptance" is explicitly stated/employed
- eq. 4.9 and 4.10
- **NOTE:** That capability is partially lost in present day **simc**!
- **NOTE:** (still) a bin–by–bin correction*

Eq. 1...

...

- Justification of eq. 1 is not entirely clear. (additional insight would be welcomed here!)
- Sure, if the cross–section implemented in **simc** IS the cross–section we get in the real world AND we know this fact *a priori*, then eq. 1 is true everywhere.
- ...but then we do not need an experiment (as we know the true answer!).
- eq. 1 looks a lot like a variant of Cauchy's **Mean Value Theorem (MVT)**.
- see next page for a *wiki* snapshot!
- **NOTE:** There is an even better way of framing this!
  Hopefully we'll get to it today!

Mean Value Theorem

Cauchy's mean value theorem [edit]

Cauchy's mean value theorem, also known as the **extended mean value theorem**, is a generalization of the mean value theorem. [6][7] It states: if the functions f and g are both continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), then there exists some $c \in (a, b)$, such that

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

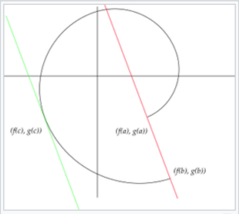
Of course, if $g(a) \neq g(b)$ and $g'(c) \neq 0$, this is equivalent to:

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Geometrically, this means that there is some tangent to the graph of the curve[8]

$$\begin{cases} [a, b] \to \mathbb{R}^2 \\ t \mapsto (f(t), g(t)) \end{cases}$$

which is parallel to the line defined by the points $(f(a), g(a))$ and $(f(b), g(b))$. However,



a couple of subtle but essential differences between MVT and mcr:

- The yield is not the integral of (just) the 2–fold x-sect (one can, maybe, argue it is the integral of the 5–fold)
- The **(e)MVT** is an "existence" theorem! So eq. 1 is ($\exists$), not ($\forall$)!
- It's possible that the statement was OK for the era (precision, size of bins,...)

Further eq. 1 studies

Looking at the present day analysis...

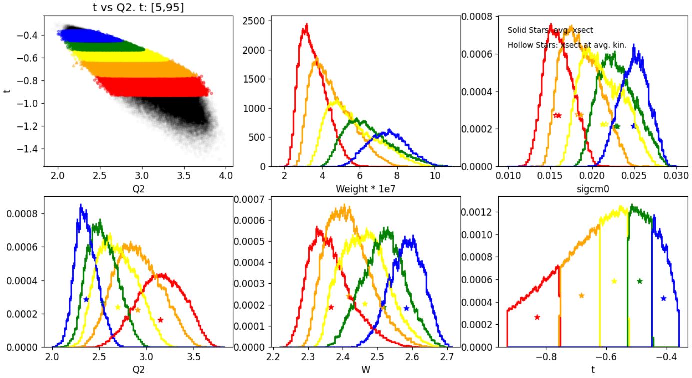
- We evaluate the **simc** model cross-section at the mean Q^2, etc. values
- We sought to gauge the “cost” of the approximation implied by eq. 1.
- Used the “stock” cross–section from simc.
- Unless we expect the model cross–section to be linear in all its variables (which would be very unusual!) subsequent iterations will not alter these, as they have more to do with the bin spans than the particularities of the model.
- We used 10 GeV kaon data (results shown are for Λ simc; Σ^0 available too).

Further eq. 1 studies (II)

...

- In t the data was split in five bins, spanning the 5–95% coverage range. Into similar (as in "close to equal") statistics.
- for each of these bins the average value was obtained for Q^2, W, t, sig_{cm}
- **Note:** All of these are "expectation values" with respect to the **Weight** kernel (which presumably contains the cross–section, all its prefactors, RC effects, etc.)
- The sig_{cm} was then evaluated at the avg. kin. and compared with the average cross–section (in absolute as well as relative terms).
- Here are some results of this study.

Kinematic Setting 0 (we did all of them, just randomly picking some):



Bin Centering Study Summary for Setting 0:

t bin: (-0.929,-0.754) mean xsect: 0.016 xsect at mean: 0.0157 Diff: **0.019**

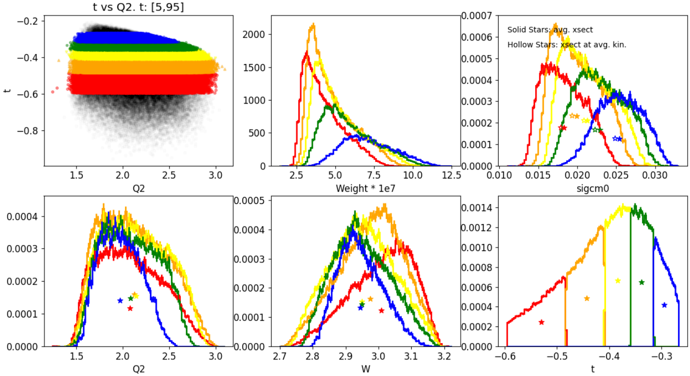
t bin: (-0.754,-0.621) mean xsect: 0.0187 xsect at mean: 0.0183 Diff: **0.021**

t bin: (-0.621,-0.528) mean xsect: 0.0212 xsect at mean: 0.0216 Diff: **−0.021**

t bin: (-0.528,-0.449) mean xsect: 0.023 xsect at mean: 0.0228 Diff: **0.007**

t bin: (-0.449,-0.358) mean xsect: 0.0249 xsect at mean: 0.0249 Diff: **0.001**

Kinematic Setting 3:



Bin Centering Study Summary for Setting 3:

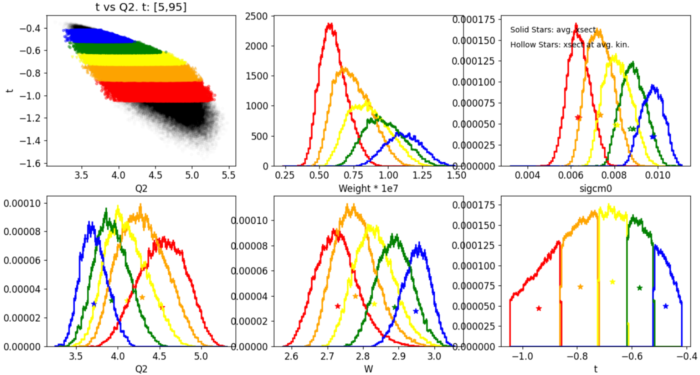
t bin: (-0.596,-0.483) mean xsect: 0.0183 xsect at mean: 0.0179 Diff: **0.023**

t bin: (-0.483,-0.409) mean xsect: 0.0199 xsect at mean: 0.0193 Diff: **0.031**

t bin: (-0.409,-0.359) mean xsect: 0.0213 xsect at mean: 0.0209 Diff: **0.021**

t bin: (-0.359,-0.316) mean xsect: 0.0229 xsect at mean: 0.0223 Diff: **0.029**

t bin: (-0.316,-0.267) mean xsect: 0.0254 xsect at mean: 0.0249 Diff: **0.021**



Bin Centering Study Summary for Setting 8:

t bin: (-1.047,-0.861) mean xsect: 0.00636 xsect at mean: 0.00631 Diff: **0.007**

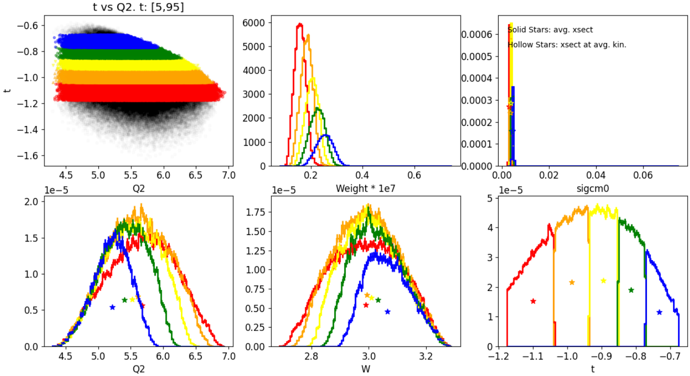
t bin: (-0.861,-0.722) mean xsect: 0.00734 xsect at mean: 0.00736 Diff: **−0.003**

t bin: (-0.722,-0.617) mean xsect: 0.00813 xsect at mean: 0.00805 Diff: **0.010**

t bin: (-0.617,-0.521) mean xsect: 0.00886 xsect at mean: 0.00874 Diff: **0.014**

t bin: (-0.521,-0.417) mean xsect: 0.00974 xsect at mean: 0.00981 Diff: **−0.007**

Kinematic Setting 16:



Bin Centering Study Summary for Setting 16:

t bin: (-1.176,-1.040) mean xsect: 0.00327 xsect at mean: 0.00315 Diff: **0.039**

t bin: (-1.040,-0.937) mean xsect: 0.00364 xsect at mean: 0.00374 Diff: **−0.027**

t bin: (-0.937,-0.853) mean xsect: 0.00395 xsect at mean: 0.00388 Diff: **0.019**

t bin: (-0.853,-0.771) mean xsect: 0.00426 xsect at mean: 0.00418 Diff: **0.018**

t bin: (-0.771,-0.675) mean xsect: 0.00461 xsect at mean: 0.00461 Diff: **−0.002**

(partial) sig_{cm} study

Looking at avg. xsect vs xsect at avg. kin...

- One incurs differences in the ± 2-3% for the t bins of the same "diamond".
- Accounting for the sign differences, one can have (easily) a 5% across the t distribution (w/ iterations, one can extend study in the ϕ direction as well).
- Seemingly at random (therefore hard to correct for with an overall eq.) unless specifically provided for via a "scaling function" (i.e. model).
- It could be that 5% random swings do not make a difference for the L–T separation or meson FF extraction
- **GN's $0.02:** Have a BC correction "at the ready" for direct use (preferably!) or to check if things get out of control would.

There is an even better way of framing...

The (proper!) way of looking at a particle/nuclear physics exp:

- When carrying out a measurement (or a simulation thereof!), limitations of the apparatus a/o known/understood physical effects (res, eff, RC, bkg...) distort the "physics" of interest.
- Essentially we have a **convolution** (see next page!) between the result we want to obtain and the "detector response".
- **IF** one wants to remove these exp. effects to recover the original "physics" (i.e. we want to quote a/some cross–section(s), polarization obs., etc. at a number of kinematic points)...
- this process is known as **unfolding**. **Way** harder than the former!!
- **NOTE1:** Unfolding is not needed for **current(!)** theory comparisons.
- **NOTE2:** However, for any subsequent comparisons, the whole machinery used for convolution needs to be provided.
- **NOTE3:** Otherwise results from exp. **A** cannot be compared w/ exp. **B**...

There is an even better way of framing (II)...

Convolution:

$$f(x) = \int_a^b g(y) r(x;y) dy \tag{2}$$

- $f(x)$ - the measured distribution
- $g(y)$ - **THE** physics we want to extract/publish
- $r(x;y)$ - the detector “response” (aka **kernel**)
- **NOTE:** Eq. 2 is known as the **Fredholm** eq. of the first kind*.
- Formally, we are dealing w/ the inverse of a diff. eq.: $Lf(x) = g(x)$
- Notorious (according to many sources!) difficult to tackle!
- Let’s see if **mcr** (iterated!) “solves” our little unfolding problem.

Ex: 2 step functions + Gaussian kernel

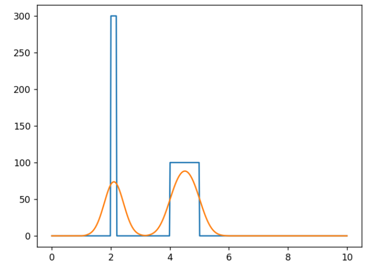
File Edit Format Run Options Window Help

```python
# testing convolution
from scipy import signal
import numpy as np
import matplotlib.pyplot as plt

x = np.linspace(0,10,500)
y = np.zeros_like(x)
y[200:250] = 100.
y[100:110] = 300.
plt.ion()
plt.plot(x,y)

kernel = signal.windows.gaussian(100, 16)
##
blurred = signal.fftconvolve(y, kernel,
                             mode='same')
blurred = blurred/blurred.sum()*y.sum()
plt.plot(x,blurred)
print(y.sum(), blurred.sum())
plt.show()
```

Figure 1



(x, y) = (0.99, 31.4)

IDLE Shell 3.13.2

File Edit Shell Debug Options Window Help

```
Python 3.13.2 (tags/v3.13.2:4f8bb39, Feb  4 2025, 15:23:48) [MSC v.1942 64 bit (
AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: E:\gabby_python_doodles\SamplesAndTests\conv_and_unfolding\test_convo
lution.py
8000.0 8000.000000000001
```

Just 20 (or so) python lines, courtesy of *Stackoverflow*

mcr & convolution/unfolding

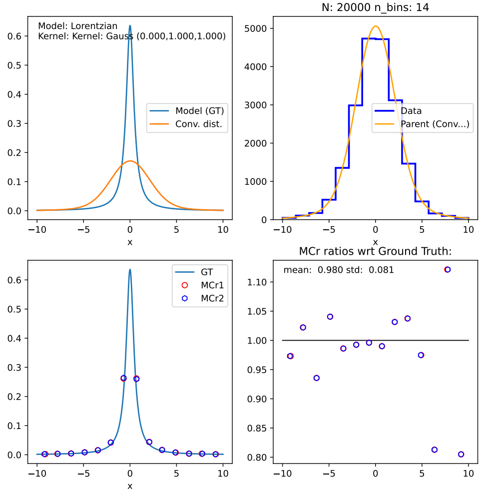
Q: Does the Monte Carlo Ratio method solve the unfolding problem?

- If **YES**: then we are good to go!
- If **NO**: then we can try a few **mcr** iterations.
  - if iterations **work**: then we are good to go!
  - if iterations **do NOT work**: we are still good to go, but where? (*Quo Vadis?*)
- we did this exercise w/ a couple of distributions that might be relevant for the **KaonLT** analysis.

Ex 3: Lorentzian + Gaussian kernel

$$f(x) = \int_a^b g(y) r(x; y) dy$$

$$\sigma_{exp}(\bar{W}, \bar{Q^2}, t, \phi; \bar{\theta}, \bar{\epsilon}) = \frac{< Y_{exp} >}{< Y_{sim} >} \sigma_{MC}(\bar{W}, \bar{Q^2}, t, \phi; \bar{\theta}, \bar{\epsilon})$$

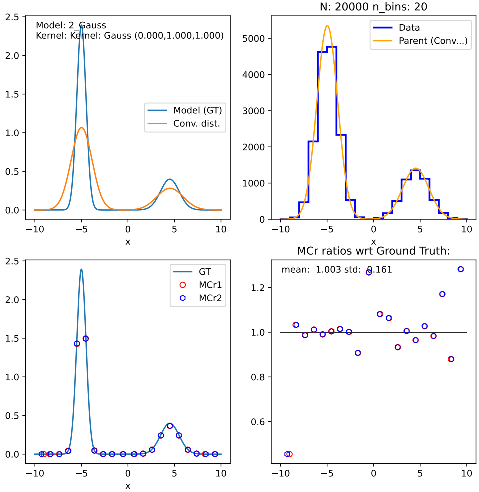


Steps:

1. Get **model**, **kernel**, Conv.
2. Use conv. dist. to generate **N** "data" ev. *ditto* 4 MC (10x).
3. Bin. Get data & MC yields.

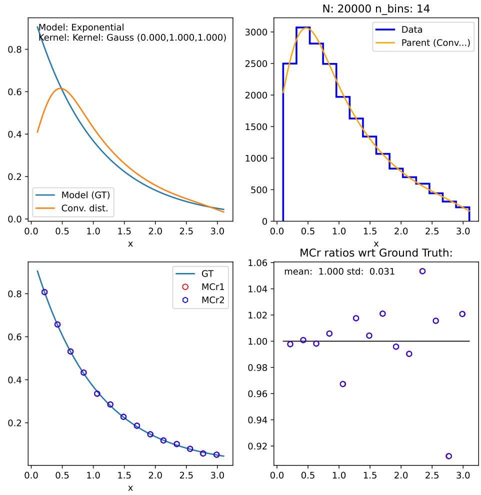
   NOTE: data & MC use **SAME** model!
4. Get $\bar{x}_i$ & $g(\bar{x}_i)$.
5. Plot **mcr** result & $g(x)$
6. Compare.

- No problem at the 5% lvl
- Beyond that, however...



- ~5% lvl? OK
- Beyond that, however...
- ...and we occasionally (in the tails) get large outliers.
- **NOTE:** this functional form not really relevant for **KaonLT** analysis.

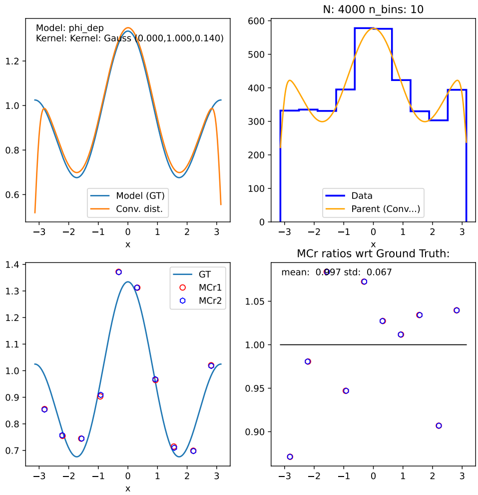
Ex 5: Exponential + Gaussian kernel



- this is not unlike the **t**-dependence...
- Again, no problem at the 5% level or so
- If we want to go beyond that, however...
- (from the audience): But t fitting is the last step of the analysis. There is ϕ fitting first...
- You are, of course, **right**.

So, let's look at some ϕ distributions!

phi distribution + Gaussian kernel



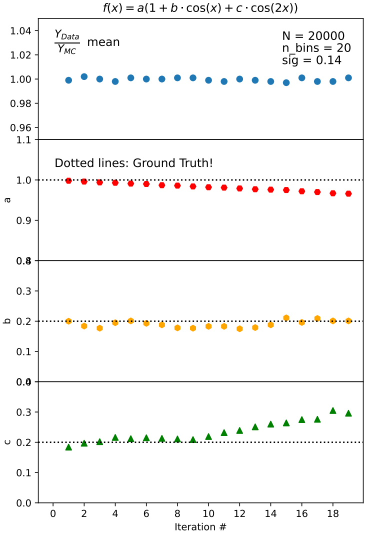
...

- realistic σ # of bins
- plausible statistics
- $\pm$8-10 excursions from unity

Steps

- use the **mcr** result points to fit our functional form
- $f(x) = a(1+b\cos(x)+c\cos(2x))$
- get a, b, and c
- use the new values to evaluate the MC side of mcr

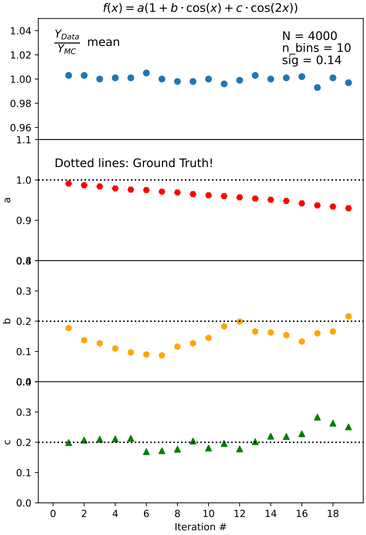
phi distribution mcr method iterations



Notes:

- the top panel **IS NOT** the stopping condition for the iterative loop (maybe it ought to be?)
- statistics are pretty high! (too high?)
- two of the three parameters **drift** monotonously. third one **dances** around as well.
- the model was already at its **GT** values on iteration 1!!!

phi distribution mcr method iterations (II)



Notes:

- the situation does not improve when we drop to "realistic" ("optimistic" if Σ^0 is your goal!) statistics.
- parameters go further and further from the ground truth...
- even though the result is self–consistent
- see difference between **precision** & **accuracy**

Quo Vadis?

We touched on one (I) of the assumptions/techniques essential to klt analysis

- I do not have a conclusion.
- Just an encouragement: **try this for yourselves**! (simple–enough in 1D)
- Perhaps there is a bug in our code.
- ...as well as in **ALL** the published, non–thesis resources cited (or similar).

- Thank you!