DELPHI Collaboration

DELPHI 2002-0 XXXX 000
10 March, 2002

Optimal statistical treatment of parameterized systematic errors in a unified Bayesian-Frequentist Monte Carlo Framework.

Niels Jørgen Kjær
CERN, CH - 1211 Geneva 23, Switzerland
e-mail: Niels.Jorgen.Kjaer@cern.ch

Abstract

This paper describes a method to explicitly incorporate systematic errors in the statistical part of a likelihood analysis.

In Monte Carlo based likelihood analyses systematic errors can be interpreted as uncertainties in the underlying Monte Carlo Model. Often, these uncertainties can be reliable parameterized giving parameters, $\mathcal{S}$, in the simulation, which are either controlled by the information from the data itself or limited by external information from other data.

In this case it is shown that the optimal way to treat systematic errors is to simultaneously estimate the observables of interest and the systematic parameters themselves. In the k2-filter approach this is done using an event-by-event analysis with optimal observables for both the observables of interest and $\mathcal{S}$. Even though events contain no explicit information on the systematic $\mathcal{S}$, the measurement will often depend on a combination of $\mathcal{S}$ and other observables.

The practical implementation of the proposed procedure is simply to generate events according to the a priori known distribution of $\mathcal{S}$. If the measurement does not depend on $\mathcal{S}$ this causes no harm, while if it does depend on $\mathcal{S}$, the Monte Carlo sample can be used to extract the correlations with $\mathcal{S}$. This is then transformed into multi-dimensional likelihood curve for the observables and $\mathcal{S}$. The simultaneous measurements can hence be done using classical techniques. An a priori knowledge of $\mathcal{S}$ can be folded into the likelihood, and experiments can be combined directly taking the systematic correlations into account.

A few simple examples are given together with comparisons with classical approaches.

Contributed paper to "Advanced Statistical Techniques in Particle Physics",
Grey College, Durham, 18 - 22 March 2002

1 Introduction

Most analysis in modern particle physics are Monte Carlo (MC) based. In this framework, the MC can be seen as **the** model which describes everything related to the measurement. This leads to a change in paradigm for treating systematic errors. There are then no systematics in either the data or related to the analysis. The analysis method is actually not even "free" to choose, but directly forced to be the unique one with the best final sensitivity, including systematics. All systematics are then related to the imperfection of the MC model. There are several advantages on such an approach:

- One statistical component of the systematics vanishes. In many classical analyses, a significant part of the systematics is evaluated by varying the analysis, (cuts, fitting procedures, etc.), within "reasonable" limits, and then use the variation as a measure of belief in stability. When varying the MC model within "reasonable" limits this statistical component is no longer important.

- All systematics are treated on equal footing, which is very beneficial when results from different measurements are combined.

- Basically all systematics can be modelled inside the MC framework. When there is a near unlimited amount of MC capability available it is possible to simulate all the systematics.

In this paper such a MC framework is assumed. Section 2 is devoted to show how an optimal treatment of systematic errors can be developed and controlled. The systematic analysis is basically Bayesian, while it is controlled by a frequentist estimation of the errors in the MC frame work. Section 3 is devoted to examples and recommendations, while Section 4 deals with a related problem of including systematic errors in likelihood curves.

2 Optimal treatment of systematic errors

The individual systematics can be regarded as nuisance parameters, whose values are either totally unknown or are bounded from outside knowledge. Throughout this paper these parameters are denoted by $\mathcal{S}$. The goal of an optimal treatment of the systematic errors is to reduce the sensitivity of the measurement to them, and to minimize the total final error of the analysis. Depending on the $\mathcal{S}$, several different scenarios might occur:

- The $\mathcal{S}$ biases all events and measurements in the same way and with the same magnitude and the individual data events contain no direct information about the value of $\mathcal{S}$. Such systematics are typically related to unknown scales, like efficiencies and energy scales. They are also quite frequent when addressing theoretical errors.

- Other systematics bias the measurements mostly in the same way but with a magnitude that depends on the individual events. Typical examples of such systematics are event-by-event background estimations and non-symmetric geometrical uncertainties like alignment etc.

- The last group of systematics are those where the data events do contain direct sensitivity to $\mathcal{S}$.

First assume that the measurement does not contain any information of the systematics. Then the combined likelihood for the measurement of x and the systematics $\mathcal{S}$ for event i can be written as:

$$L_i(x, \mathcal{S}) = L_i(x + c\mathcal{S}), \tag{1}$$

where c is a constant. Figure 1 shows a toy example of how the 2-dimension likelihoods might look like and the combined likelihood. It is obvious that no information is gained

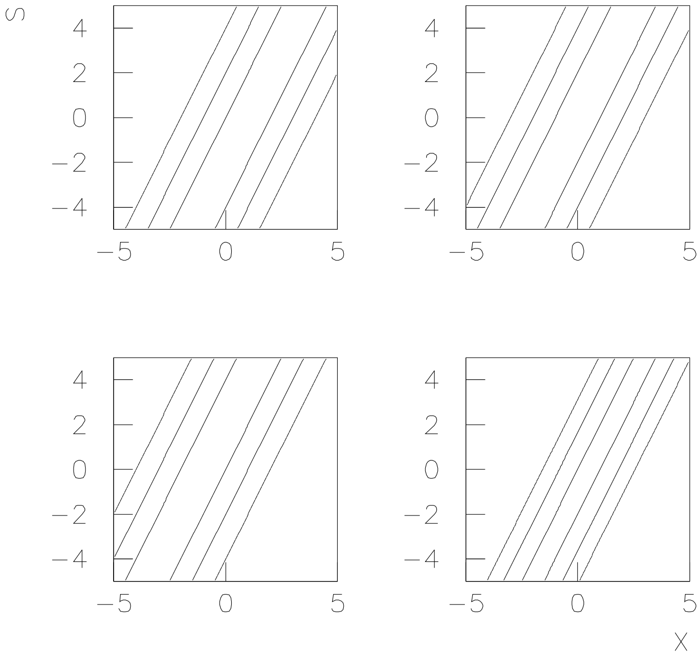

Figure 1: The 2-dimensional $(x, \mathcal{S})$ likelihood curves (1-3 sigma contours) are shown for three different events when the systematic is uncorrelated with the events. The lower right plot shows the combined likelihood contours.

on the systematics. The global error then becomes the measurement error added in quadrature to a global systematic error. The first error behaves like one over the square root of events, while the systematic error remains constant, and infinite without external information to bound it.

Then assume that the measurement only contains indirect information on the systematics:

$$L_i(x, \mathcal{S}) = L_i(x + c_i\mathcal{S}), \tag{2}$$

where c_i varies from event to event. Figure 2 shows a similar toy example for such systematic errors. Even though the individual events do not contain information about the systematic, the combined results does. If this information is included in the analysis also this systematic error will behave like one over the square root of events. In the case of Gaussian likelihood curves it is possible to do the combination analytically. If for example c_i is distributed uniformly from 0 to $c_{\max}$ and there is no external information about $\mathcal{S}$, the optimal measurement of x consists of weighting the information from each event, where the event weight becomes $1 - \frac{3c_i}{2c_{\max}}$. It is beyond the scope of this paper to go into details of such classical estimations.

Finally, assume that there is direct information about the systematics in the data. Figure 3 shows a similar toy example for such systematic errors. Here the information is

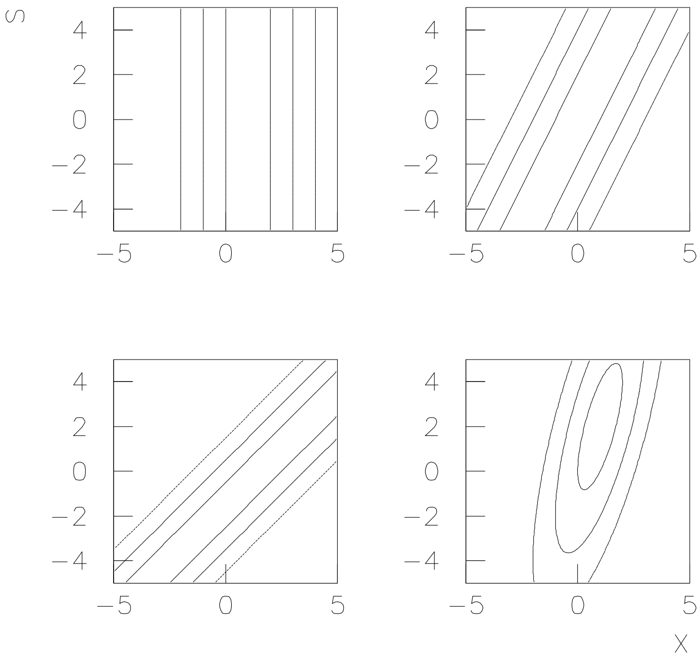

Figure 2: The 2-dimensional $(x, \mathcal{S})$ likelihood curves (1-3 sigma contours) are shown for three different events when the systematic is different for the different events, but the individual events contain no information on the systematic. The lower right plot shows the combined likelihood contours.

genuinely 2-dimensional, and the optimal way to include the systematics is obviously to make a full 2-dimensional simultaneous estimation of x and $\mathcal{S}$.

In [1] it was shown how to make optimized correlated measurements within a MC framework. By using reweighting techniques and an extension of Kalman filters it was shown that multi-dimensional measurements can be optimized independently locally providing derivatives of the log-likelihoods as optimal “actions”. The full likelihood curves are then simply constructed from integrating the optimized derivatives.

This method is fully applicable to the situations described above. In principle all the systematics can be included in a n-dimensional treatment together with all the observables that contain information on x and $\mathcal{S}$ on an event-by-event basis. In practice it is more efficient only to concentrate on those observables and systematics which are relevant, and often limit such investigations to one systematic and a few observables.

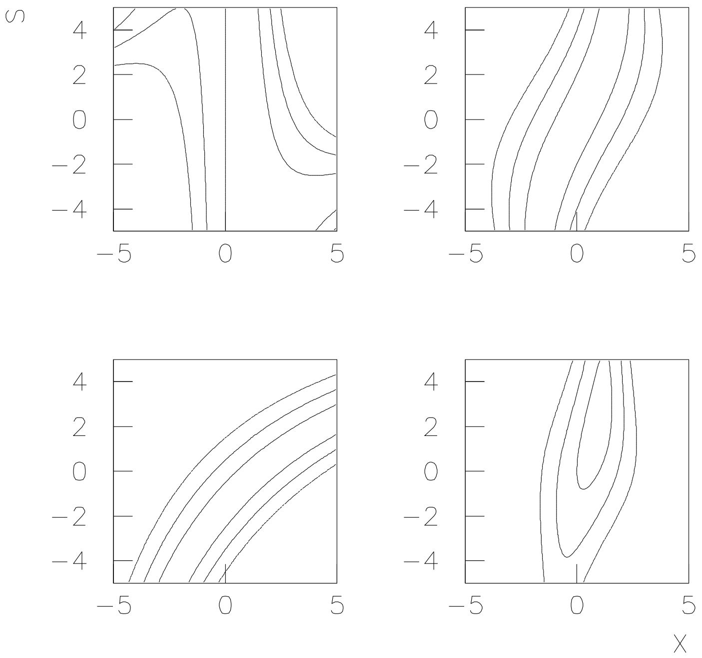

Figure 3: The 2-dimensional $(x, \mathcal{S})$ likelihood curves (1-3 sigma contours) are shown for three different events when the individual events contains information about the systematic. The lower right plot shows the combined likelihood contours.

When the full 2-dimensional $(x, \mathcal{S})$ information is available it is then straight forward also to do combinations of the measurement, since the log-likelihood simply add up. It then depends on the structure of the likelihood curves and the number of events whether Gaussian approximations can be applied or not. Finally, the results can be combined with any external information, like theoretical predictions on the systematics and so forth.

3 Examples of possible implementations

3.1 Luminosity measurement

There are many classical measurements where different correlations between observables and systematic are taken into account. A standard example is the luminosity measurements performed at LEP. There [2], advantage was taken of the fact that, the absolute distance between the opposite luminometers was precisely known. Then the effect of not knowing the absolute position w.r.t. the interaction point cancels to first order if the detectors are absolutely symmetric. This effectively leads to a precision measurement of the position of the interaction point relative to the luminometers.

3.2 SK1 measurement

In [1] a measurement of a model parameter κ was described using the k2-filter technique. Figure 4 shows a typical 2-d likelihood from such a measurement. The measurement of m_{W} is always very parabolic and it is only in κ the variation is not linear. The measurement is then performed as a grid with 12 points in κ and only two in m_{W}. Such results can afterwards easily be combined with other measurements and convoluted with any theoretical predictions on κ without loss of information.

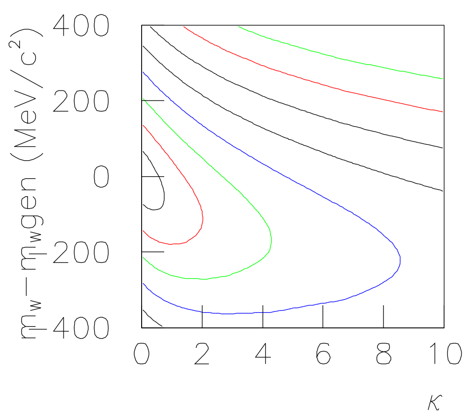

Figure 4: The 2-dimensional $(\kappa, m_{\mathrm{W}} - m_{\mathrm{W}}^{\mathrm{gen}})$ likelihood curve (1-8 sigma contours) is shown for MC events generated with $\kappa = 0$ and $m_{\mathrm{W}}^{\mathrm{gen}}$.

3.3 QCD errors

Many HEP measurements have significant systematic errors coming from the description of QCD and fragmentation processes. In fragmentation models, the process is steered by a finite number of so-called QCD parameters. Quite often systematics have been estimated by varying these parameters and comparing various different models.

For LEP measurements the value of the parameters have been extracted by tunings to somewhat arbitrary sets of event observables. Within the framework of this paper such an approach can be refined. Assume that events are generated not with fixed values of the QCD parameters, but with a variation which corresponds roughly to the current uncertainty on these parameters. Then the analysis strategies could be significantly altered for some scenarios:

- Measurements which have no sensitivity to the QCD parameters do not feel this alternative generation procedure and are therefore left unchanged.
- Measurements which do have a systematic sensitivity to the QCD parameters should be altered. They can either be used to perform a simultaneous measurement including the QCD parameters, or the analysis can be constrained to not depend on the parameters as was described in [1](Section 3.6).
- The estimation of the QCD parameters could be performed on the "standard" generation using the full detector simulation, and not be forced to use less reliable unfolding techniques.

3.4 The complex b-measurement sector

The heavy flavours provide some of the most interesting and involved analyses and measurements. The list of systematics is very long and the analyses typically provide information which can be used for several measurements. The working groups who are responsible for the combination [3] take into account all the different systematics and their potential correlations. It is not always completely clear how to estimate the correlations and to make sure that measurement results are not used to verify their own results. If Kalman filters were used to linearize the problem it would in principle become easy to control the correlations rigorously and even to filter out if there are regions of the measurements which are less prone to systematics and therefore should obtain more statistical weight in the combination. The author is of the opinion that such an approach is viable, but also acknowledges the limitations of such an approach, where basically everything is measured simultaneously on an event-by-event basis.

4 Likelihood curves and systematic errors

It is desirable to include systematic uncertainties in likelihood curves. Likelihood curves like the one in Figure 5 show the relative likelihood as a function of the measurable x. It is expected that additional systematics would widen the resulting likelihood curve. Several procedures have been proposed to make a practical implementation, but none of them have been shown to work fully consistently for any likelihood and any systematics.

The r.h.s. of Figure 5 shows a non-linear systematics. It should be understood in such a way that there is an unknown nuisance parameter distributed as a Gaussian around zero with unit width. For a given value of this parameter, the log-likelihood will change with the amount shown in the plot multiplied by the parameter. For certain "fix-points", like points A and B there will be no systematics, hence the log-likelihood difference should remain the same. If these fix-points lie close to a local minimum and a local maximum as is the case in Figure 5, there is no longer a way to solve the problem. The inclusion of systematics will the tend to pull the log-likelihood curve in opposite directions for the x values which lie in between the two points.

So in practice it is not possible to include systematics in likelihood curves since it leads to inconsistencies.

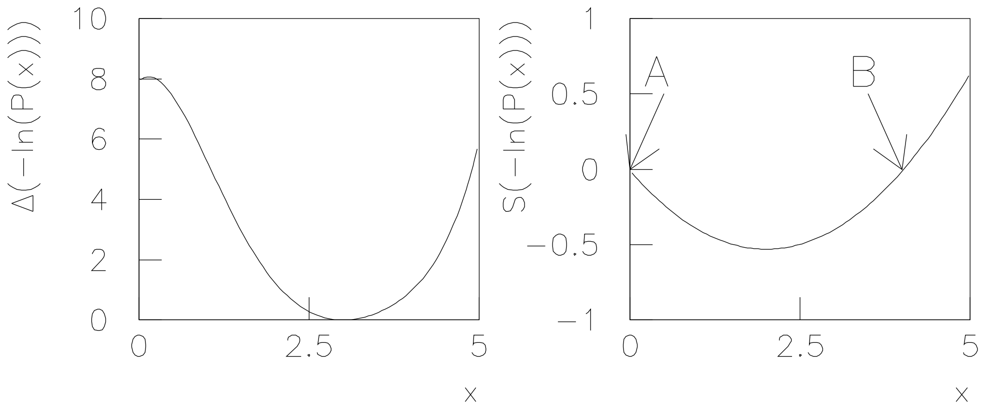

Figure 5: A hypothetical likelihood curve is shown in the left hand plot. The right hand plot shows the one-sided 1-sigma variation for a given hypothetical systematic, $\mathcal{S}$.

The solution to this problem is somewhat related to Bayes priors. Either the measurement is a precision measurement and the inclusion of systematics is trivial, or there is so little information that the measurement is meaningless without a prior. In the latter case the problem can be regularized by imposing that the measurement is a likelihood ratio measurement w.r.t. a reference value of the parameter. Such a reference does often exist and is usually related to a "Standard Model". When the problem is reduced to "two-point" likelihood ratios, the systematics can be included by simple integration over the nuisance parameters.

5 Conclusion

The optimal way to treat the systematics is to measure it. This paper was intended to address this rather trivial statement. It was shown that even though the data do not contain direct information of a systematic error, its correlations can be exploited to provide an indirect measurement, which can be quite as powerful. Using MC parameterizations of possible systematics provides the most powerful tool to extract the necessary information from the data events. When this information can be addressed via MC reweighting the k2-filter provides the optimal way to extract the simultaneous information about the measurement and the systematics. Finally, a few examples were mentioned and the problem of including systematic errors in likelihood curves addressed.

References

[1] N. J. Kjær, J. D'Hondt
Exact optimization of likelihood analyses in a unified Bayesian-Frequentist Monte Carlo Framework: The "k2-filter"
Contributed paper to "Advanced Statistical Techniques in Particle Physics",
Grey College, Durham, 18 - 22 March 2002

[2] As an example:
The OPAL Collaboration Eur. Phys. J. C14 (2000) 373-425

[3] Combined results on b-hadron production rates and decay properties, CERN-EP/2001-50