HORIZON 2020

HORIZON 2020 RESEARCH AND INNOVATION FRAMEWORK PROGRAMME OF THE EUROPEAN ATOMIC ENERGY COMMUNITY

Nuclear Fission and Radiation Protection 2018 (NFRP-2018-4)

Project acronym:	**SANDA**
Project full title:	**Solving Challenges in Nuclear Data for the Safety of European Nuclear facilities**
Grant Agreement no.:	**H2020 Grant Agreement number: 847552**

Workpackage N°:	***WP4.5***
Identification N°:	***D4.7***
Type of document:	***Deliverable***
Title:	Report on the possibility to generalize the high-energy model uncertainties methodology
Dissemination Level:	***PU***
Reference:	
Status:	***VERSION 1***
Comments:	

		Partner	Date	Signature
Prepared by:	J. Hirtz	3	17-07-2023	
WP leader:	J.-C. David	3	18-07-2023	
IP Co-ordinator:	E. González	1	xx-xx-2023	Signed by GONZALEZ ROMERO ENRIQUE MIGUEL - DNI ***3708** at 04/08/2023 with a certificate emitted by AC Sector Público

Contents

1 Introduction **2**

2 Objective **3**

3 Methodology **5**

3.1 Optimisation algorithm 6

3.2 CPU Optimisation 9

3.3 Limits of the approach 11

4 Experimental data treatment **12**

5 Toy models **14**

5.1 Toy model 1: Basic model 14

5.2 Toy model 2 : Stochasticity 15

6 Parameter optimisation **17**

6.1 Favourable case 17

6.2 Double differential neutron case 20

7 Summary and outlook **23**

A Derivation of the GLS method **24**

1 Introduction

As the Dutch physicist Walter Lewin wisely say it: "Any measurement that you make, without any knowledge of the uncertainty, is meaningless". It is true for experimental measurement as well as for theoretical models. As precise and reliable they can be, experimental data and models are always only an approximation of the reality and, therefore, the difference between data or models predictions on the one hand and the reality on the other hand have to be estimated to make them meaningful. However, while the estimation of uncertainties for experimental measurements became the norm a century ago, models uncertainties evaluation are much more recent and was a long time limited to statistical uncertainties.

The recent improvement of computer power in the last decades brought new tools for model uncertainty quantification, especially for Monte Carlo (MC) models. These models are of the highest importance in nuclear physics as measuring all required nuclear data is impossible for all the various fields of application (*e.g.*, fusion technology, medical hadron therapy, cosmogenic transmutation, etc.). Models able to predict the relevant data are needed to design instruments, radioprotection equipments, or to analyse data. The critical aspects in nuclear physics initiated a very early study of model uncertainties, which is commonly called *nuclear data evaluation*. Considering the wide range of applications of nuclear models and their relevance for societies, it is obvious that model calculations must be as precise and reliable as possible. This means the bias, the difference between the estimator and the true value of an observable, and the uncertainties of model must be estimated precisely for a proper use of these nuclear model.

The majority of methods proposed for the study of model uncertainties are based on Bayesian statistics. The Bayesian statistics is a general framework for inference of probabilities with a limited knowledge of the relevant information. It is normally used to attribute an *a posteriori* (density of) probability to different possibilities based on incomplete measures. This allows to estimate the likelihood of a result as well as its uncertainties.

In the 20th century, the nuclear data evaluation was mostly used in the energy domains which are relevant for nuclear power plant. Namely, nuclear data evaluations focussed on the energies below 20 MeV. This led to the creation of nuclear data libraries, which are basically tables of (double differential) cross sections. At present, new type of project are envisaged with operating energies much higher. As an example, the *Multi-purpose hYbrid*

Research Reactor for High-tech Applications (MYRRHA) [1] project will operate at energies up to 600 MeV. Therefore, the new (and large) energy range must be carefully studied too.

During the last Nuclear Data projects [2, 3, 4], an important effort has been devoted to the development, improvement and validation of the high energy nuclear models, in particular the INCL/ABLA combination of models that are now widely used for high energy applications.

In the CHANDA project [2], for the first time, a study has been conducted to investigate a possible methodology based on the Bayesian framework for quantifying the uncertainties linked to parameters in high energy models and propagating those uncertainties in MC transport codes. In this project, it is proposed to investigate if the methodology can be applied to a large number of parameters of INCL within a reasonable computational time.

In this study, we are interested in the model combination of the intranuclear cascade model INCL++6 and the de-excitation model ABLA++. In combination, they are able to simulate spallation reactions, which are high-energy nuclear reactions, in which a target nucleus is hit by an incident particle of energy greater than some tens of MeV.

Noteworthy, the objectives of this study (see section 2) are specific to nuclear evaluation with the combination INCL/ABLA. However, the methodology we developed to study our specific case is a general framework and can be apply to a large variety of model. In section 3, the limits of our approach will be discussed in the general case for a comprehensive description of cases where the method could be apply. Section 4 details the treatment of experimental data required before the use of our algorithm. Next, section 5 illustrates our methodology in some basic cases created on purpose. Then, the methodology is applied on INCL/ABLA with the use of real experimental data in section 6. Finally, we discuss the outlook of this work in section 7.

2 Objective

Task 4.5 of the European project SANDA (Supplying Accurate Nuclear Data for energy and non-energy Applications) called "High-energy model uncertainties" aims at estimating model bias as well as model parameters bias in the specific case of INCL/ABLA for energies above 20 MeV. The estimations of the model bias and of the parameters bias are actually orthogonal. One can estimate the model bias without optimising the model parameters

and vice versa.

In the CHANDA framework, we already validated the methodology based on the Bayesian approach, which estimates the bias of the model combination INCL/ABLA [5]. Doing so, we were able to demonstrate that we can estimate not only the bias of the model but also the uncertainties for specific observables. This has been done explicitly for inclusive double-differential neutron spectra from various proton induced reactions. Therefore, it has been decided to focus here on the parameters optimisation.

Model bias is, by definition, the expected difference between model predictions and the true values of these observables (*e.g.*, neutron multiplicity, angular distribution, mass distribution, etc.). Equally, model parameters bias is the expected difference between the parameters value provided to the model and their true values. However, the "true" values of parameters (when it is meaningful) are not accessible. Therefore, we have to consider that experimental data are the best estimations of the truth to characterise the model bias as well as the parameters bias. This is a reasonable consideration as there is not reason *a priori* for the experimental data to overestimate the reality nor to underestimated it. In other words, experimental data are the best estimation *a priori* of the reality we have and are *a priori* unbiased. With an equation, it can be written:

$$Bias = \mathbb{E}(\sigma_{exp} - \sigma_{true}) = 0, \tag{1}$$

with σ_{exp} experimental data and σ_{true} the true values of the corresponding observables.

One conceptual issue when we aim at measuring parameters bias is that the definition of the bias is meaningless when these parameters are not "physical" parameters. As an example, particles masses are "physical" parameters, while parameters used in INCL to determine when the model should stop running are model dependent parameters. Additionally, the estimation of the parameters bias will be carried out within the Bayesian framework, which assume the combination INCL/ABLA is a perfect model. In other words, the Bayesian procedure expects the model to perfectly reproduce the reality if the parameters of the model are the best ones. However, as says the famous quote attributed to the British statistician George E.P. Box: "All models are wrong, but some are useful". INCL/ABLA, as any model, cannot be perfect. This is why the procedure will not search for the true value of the model parameters but for the optimal parameters within the context of the model considered and of the observables studied.

On another hand, the uncertainties of the model parameters will also be evaluated. These uncertainties are useful as they provide information about the error propagation in the model. Small uncertainty for a given parameter would mean that a small modification significantly modify model predictions. Reciprocally, large uncertainties would mean that the parameter value does not play a major role on the model predictions.

In this study, our objectives are twofold. First, we aims at demonstrating the feasibility of our approach in real cases using the model combination INCL/ABLA. Second, we want to study the possibilities, the difficulties, and the limits of our procedure both for the evaluation of the optimal parameters of a model and for the evaluation of the corresponding uncertainties.

3 Methodology

As mentioned in section 2, the objective of this study is to estimate the optimal model parameters and the uncertainties associated to these parameters, which would provide information about the error propagation in the model. The related question of the model bias estimation is orthogonal to this study and has already been addressed in the previous study carried out by Schnabel within the CHANDA framework [5] and will not be further discussed.

The methodology we developed is based on the Generalised Least Squares (GLS) method [6], which is an important technique in nuclear data evaluation as it is used to estimate the unknown parameters in a linear regression model with a potential high degree of correlation between the observables. The GLS is a method of regression similar to the common χ^2 but the correlations are taken into account and the model parameters are treated as extra data, which therefore limits the risk of unphysical prediction for the parameters.

Our approach is divided in two main parts.

In a first step, we want to know what are the optimal parameters for the model. In other words, we want to estimate what are the parameters which will result in the best model prediction. Our approach being based on the Bayesian statistics, it takes both into account the reproduction of the experimental data and the respect of the *a priori* knowledge on the parameters. This is realised using the Expectation Maximisation (EM) method [7], which is an iterative algorithm which finds the parameter set maximising the likelihood of a model. In other words, the EM minimise the difference

between the model prediction and the experimental data taking into account the experimental uncertainties and the correlations.

In the second step, we want to know what are the uncertainties associated to each parameter as it provides information about the error propagation. In this respect we developed an approach based on the Gibbs sampling, which is suitable for stochastic model. With this approach, we sample parameters set in a multivariate normal distribution using the posterior covariance matrix of the model. The distribution of the parameters at the end of this second step allows to determine the uncertainties of these parameters.

It is important to mention that, if the model have difficulties to reproduce some of the experimental data with respect to their experimental error bars, the algorithm will focus on these data points and neglect the others. This is why the selection of experimental data to be included in the analysis as well as a careful study of their uncertainties must be carried out before trying to optimise the model parameters. In other words, the experimental data included in this approach have to be reasonably reproducible by the model. Otherwise, these toxic data points may make this approach highly inefficient.

3.1 Optimisation algorithm

Here, we will describe the use of our algorithm. As mentioned before, the algorithm we developed for the optimisation of parameters is divided in two steps: the Expectation Maximisation (EM) and the Gibbs sampling. The first step aims at evaluating the optimal parameter set for the model and the second step aims at determining the corresponding uncertainties.

The two steps are both a recursive algorithm. The number of iterations for both methods are free parameters, which can be fixed by the user. For the EM, it must be large enough to converge to the optimal parameters set. For the Gibbs sampling, it must be large enough to estimate the variance of the parameters with the distribution obtained. On the other hand, the computational time increases linearly with the number of iterations. Therefore, the minimum number of iterations required might range from a dozen to a hundred for the EM and from some hundreds to some thousands for the Gibbs Sampling.

The main idea of the EM is the following. We start with a model (here INCL/ABLA), experimental data $\vec{\sigma}_{exp}$ and, a set of parameters $\vec{p}_{ref}$, which represents the best estimation for these parameters *a priori* (*i.e.*, without knowledge about $\vec{\sigma}_{exp}$). Here, the model is seen as a function taking a vector

as input (the parameters) and with a vector as an output (the observables) corresponding to the experimental data. This means that the dimension of the model predictions, $\mathcal{M}(\vec{p})$, must fit the dimension of $\vec{\sigma}_{exp}$. In our specific case of INCL/ABLA, it is done with an additional layer above the standard version of the model which extracts the experimental setups (projectiles, targets, energies, angles, etc.) from the experimental data provided, runs the INCL/ABLA simulations with the same setups and, using the parameter set $\vec{p}$, extracts the observables corresponding to the experimental data from the ROOT files produced and, finally, produces a vector matching $\vec{\sigma}_{exp}$.

Then, we enter a loop to improve the set of parameters $\vec{p}_{ref}$. After the i-th loop, the improved set of parameters is called $\vec{p}_i$. With the knowledge of how the model varies locally, given by the Jacobian of the model in $\vec{p}_i$, and the difference between the model prediction $\mathcal{M}(\vec{p}_i)$ and the experimental data $\vec{\sigma}_{exp}$, one can determine what is the best set of parameter $\vec{p}_{i+1}$ to minimise the difference between the model and the experimental data, assuming the model is linear between $\vec{p}_i$ and $\vec{p}_{i+1}$, the optimal parameters. Since the model has no reason to be linear, the new set of parameters will likely not be the optimal parameters set. However, if the model is not completely erratic between $\vec{p}_i$ and $\vec{p}_{i+1}$, the linearisation of the model can be seen as an acceptable approximation of the model. Therefore, the new set of parameter $\vec{p}_{i+1}$ will certainly be an improvement with respect to $\vec{p}_i$. Then, we can reevaluate the local Jacobian and the real model prediction in $\vec{p}_{i+1}$ and restart the loop until convergence of $\vec{p}$.

Explicitly, the EM is executed as follows. At the beginning of each loop, we linearise the model with a Taylor approximation in $\vec{p}_i$:

$$\vec{T}_i(\vec{p}) = \mathcal{M}(\vec{p}_i) + J_{p_i} \otimes (\vec{p} - \vec{p}_i), \tag{2}$$

with the Jacobian matrix J_{p_i} of the model evaluated in $\vec{p}_i$:

$$J_{p_i} = \frac{d\mathcal{M}(\vec{p}_i)}{d\vec{p}_i}. \tag{3}$$

We introduce the matrix J_i:

$$J_i = \begin{pmatrix} \mathbb{I}_{n\times n} & J_{p_i} \\ \emptyset & \mathbb{I}_{m\times m} \end{pmatrix}, \tag{4}$$

with n the number of experimental data, m the number of parameter and, $\mathbb{I}$ the identity matrix.

The definition of J_i allows to define the matrix of regression as:

$$\tilde{\Sigma}_i = J_i \, \Sigma \, J_i^T = \begin{pmatrix} \tilde{\Sigma}_{DD_i} & \tilde{\Sigma}_{DI_i} \\ \tilde{\Sigma}_{ID_i} & \tilde{\Sigma}_{II_i} \end{pmatrix}, \tag{5}$$

with $\tilde{\Sigma}_{DD_i}$ of dimension $n \times n$, $\tilde{\Sigma}_{II_i}$ of dimension $m \times m$, and Σ the covariance matrix of the joint distribution of the experimental data and the input parameters:

$$\Sigma = \begin{pmatrix} \Sigma_{exp} & \emptyset \\ \emptyset & \Sigma_p \end{pmatrix}, \tag{6}$$

with Σ_{exp} and Σ_p the covariance matrix of the experimental data and of the parameters, respectively. The Σ matrices might be non-diagonal in case of correlation between the experimental data or between the parameters.

Then, we can evaluate an improved set of parameter using the formula from the GLS method:

$$\vec{p}' = \vec{p}_{ref} + \tilde{\Sigma}_{ID_i} \left(\tilde{\Sigma}_{DD_i} \right)^{-1} \left[\vec{\sigma}_{exp} - \vec{T}_i(\vec{p}_{ref}) \right]. \tag{7}$$

The demonstration of this formula is displayed in Appendix A.

This formula would provide directly the optimal parameters ($\vec{p}_{opti}$) for the model in case of a linear model. However, in the general case, $\vec{p}'$ is only an approximation of $\vec{p}_{opti}$. The quality of this approximation is directly correlated to the linearity of the model between $\vec{p}_i$ and $\vec{p}_{opti}$. Even if the model is not linear, $\vec{p}'$ is likely an improvement with respect to $\vec{p}_i$. Then, we can restart the procedure with $\vec{p}_{i+1} = \vec{p}'$. This will improve the quality of the GLS hypothesis of a linear model between $\vec{p}_i$ and $\vec{p}_{opti}$ and therefore, the precision of Equation 7. If the model is not completely erratic, we expect de difference $|\vec{p}_{opti} - \vec{p}_i|$ to decrease exponentially with the number of iterations.

In the case of a stochastic model, the hypothesis of a linear model between $\vec{p}_i$ and $\vec{p}_{opti}$ might be reasonable as long as the expected difference of the model predictions between $\vec{p}_i$ and $\vec{p}_{opti}$ dominates the stochasticity. However, as $\vec{p}_i$ approaches $\vec{p}_{opti}$, the formula becomes less and less valid. Therefore, we expect an initial quick convergence as $\vec{p}_i$ is far form $\vec{p}_{opti}$, then the $\vec{p}_i$ will start oscillating around $\vec{p}_{opti}$. In order to evaluate $\vec{p}_{opti}$, we then need to average the values of $\vec{p}_i$ along the oscillating phase. This will remove the effect of the stochasticity.

In the second phase of the algorithm, the Gibbs sampling, we want to sample $\vec{p}_i$ around $\vec{p}_{opti}$ using a multivariate normal distribution with the posterior

covariance matrix Σ_i in order to evaluate the distribution of $\vec{p}$ and therefore, to measure the uncertainties and the correlations of the parameters:

$$\vec{p}_{i+1} = \mathcal{N}(\vec{p}', \Sigma_{i+1}). \tag{8}$$

For each loop of the Gibbs sampling, $\vec{p}'$ is re-approximated using Equation 7, which means we process one loop of the EM within each iteration of the Gibbs sampling, and the covariance matrix Σ_{i+1} is an updated version of the initial parameters covariance matrix (see Appendix A for details), which includes the variance of the experimental data and the error propagation through the model:

$$\Sigma_{i+1} = \Sigma_p - \tilde{\Sigma}_{ID_i} \left(\tilde{\Sigma}_{DD_i}\right)^{-1} \tilde{\Sigma}_{DI_i}. \tag{9}$$

This Gibbs sampling is well suited in the case of stochastic models since the $\vec{p}'$ used is re-approximated in each loop and its dispersion over the loops allows to integrate the stochasticity of the model in the uncertainties of parameters. In addition to the dispersion of $\vec{p}'$, the sampling uses the updated covariance matrix Σ_{i+1}, which contains the information of the *a priori* knowledge on the parameters and the experimental data and also about the error propagation through the model. These uncertainties (and the correlations between the parameters) can be obtained with the covariance matrix of the multivariate distribution of $\vec{p}_i$ obtained with the Gibbs sampling. This allows to evaluate the parameters uncertainties in the Bayesian framework taking into account the *a priori* uncertainties of this parameters as well as the other relevant uncertainties, which are the error propagation through the model, the stochasticity of this model (which depends on the statistic used) and, the experimental data error bars. On the other hand, if we do not want to integrate the information about the stochasticity of the model in the parameters uncertainties, we can average the covariance matrix Σ_i over the iterations. In this case we would obtain parameters uncertainties with only the consideration of *a priori* uncertainties of this parameters, *a priori* uncertainties of experimental data, and the error propagation through the model.

3.2 CPU Optimisation

CPU Optimisation tricks has been used for the computation.

First, the inversion of the matrix $\tilde{\Sigma}_{DD_i}$ being very CPU consuming when using a large amount of experimental data, the Woodbury matrix identity is

used:

$$\left(\tilde{\Sigma}_{DD_i}\right)^{-1} = \left(\Sigma_{exp} + J_p \Sigma_p J_p^T\right)^{-1}$$
$$= \Sigma_{exp}^{-1} - \Sigma_{exp}^{-1}\ J_p\ \left(\Sigma_p^{-1} + J_p^T \Sigma_{exp}^{-1} J_p\right)^{-1} J_p^T \Sigma_{exp}^{-1}. \qquad (10)$$

When no correlation is taken into account, the Σ matrices being diagonal, the inversion is straightforward. When there are correlations, there are often limited to a small group of data and, inverting the matrix is therefore still very effective. Then, the problem of a large non-diagonal matrix inversion becomes a problem of large matrices multiplications, which is faster and can easily be parallelised.

Second, the Jacobian should theoretically be computed for each loop. However, it would be highly CPU inefficient as the Jacobian does not change drastically between two loops while it could requires a significant CPU time to be estimated. Therefore, at the end of each loop, we control if the Jacobian is still valid and, if not, it is revaluated. This is done with a comparison of $\mathcal{M}(\vec{p_i})$ (which is evaluated in each loop for the needs of the Taylor approximation) and $T_j(\vec{p_i})$, the Taylor approximation made the last time the Jacobian has been evaluated. If the two prediction in $\vec{p_i}$ differs by less than a predefined value, we consider the Jacobian is still valid. The condition for reevaluation is defined as a function of the number of experimental data, the number of parameters to be optimised, and the CPU time required by the model to evaluate the Jacobian itself.

Third, as we compile values express in different units in the Σ and, by extension, the $\tilde{\Sigma}_i$ matrices, it is not rare to have many orders of magnitude of difference between the different elements of these matrices. This can introduce errors due to the limited precision of computers while multiplying or inverting the matrices. In such a case, it might be useful to rescale the output of the model and the experimental data. In other words, we can choose to optimise the parameters for the model $\mathcal{M}' = A \times \mathcal{M}$ using the experimental data $\tau' = A \times \tau$ with A an arbitrary diagonal matrix. In such a case, the experimental error bars must also be updated but not the parameters. Proceeding this way is perfectly equivalent to optimise the parameters for the model $\mathcal{M}$ using the experimental data τ.

3.3 Limits of the approach

In our case with INCL/ABLA, there are three main limits for the use of our method.

First, one of the main challenge with nuclear data evaluation is the massive amount of observable to reproduce. One of the crucial assumption concerns the uncertainties of experimental data sets. Including automatically a large number of experimental data sets into the Bayesian procedure always brings the risk of some data sets having too optimistic uncertainty specifications. It is often the case with old experimental data when systematic errors were not evaluated or roughly fix to 10% per convention. As an example, some of the experimental data we included in our study were provided with relative uncertainties below 1%. In this case, the Bayesian procedure attributes a very high importance to this data while other data measured in experiment with a more rigorous evaluation of the uncertainties will under contribute to the final results. Therefore, a careful study of the experimental data included in the Bayesian procedure must be carried out in order to use realistic (or, at least, consistent) uncertainties for every set of data. This is further discussed in section 4.

Similarly, the over experimental measurement of some observables will induce an over fit of these data as each data point is considered individually and not as a set since they are almost never provided with their correlations. In other words, the more data points an experiment has, the more it will influence the final result of our study. For example, the neutron production cross section has been very well studied with respect to the proton production cross section. Therefore, if the two data sets are included in the same study, the approach will put a lot of effort to reproduce the experimental neutron cross sections as they are numerous while the proton cross sections will contributes marginally.

The second main issue is the stochasticity of the model used. The energy considered (above 20 MeV) is not described properly with deterministic model as the number of possibilities increase exponentially with the energy. MC models become necessary for these energies but it comes with the usual balance between precision and computation time. However, whatever the statistic, two simulations with the same initial state but different seeds will give different results. In order to avoid that the *a posteriori* probability associated to a parameter set varies too much from one run to another, the statistics must be carefully chosen in order to obtain a good balance between

CPU time and precision. This might become complex in a case with a large number of experimental data requiring very different statistics to be properly estimated by the model.

The last, but not least, issue: model deficiencies are not taken into account. Parameters can be optimised within the context of the model but the approach does not provide direct information about model deficiencies. This means two things: First, if the model deficiencies forbid to reproduce the experimental data whatever the parameters set used, the optimal set estimated will be unsatisfactory. As an example, if we try to optimise the parameter of a toy model where the data to be reproduced are distributed as a quadratic function and the toy model allows only linear function, the approach will optimised the parameter to minimise the bias but, despite the parameter will be optimal, the model will not be able to reproduce the quadratic shape of experimental data (see ref. [8], section 3.2). Second, as the approach minimises the variance, which evolves with the square of the difference between experimental data and the model predictions, a minor improvement in a region where the model is highly deficient will be seen as great improvement, while a large increase of the difference between experimental data and the model predictions in regions where the model reproduce properly the experimental data will be seen as a minor deterioration of the model. In other words, if we try to optimise the model with experimental data with part of them in deficient regions of the model and another part in efficient regions of the model, the algorithm will primarily improve the model prediction in the worst regions regardless of the effects on the model prediction in the good regions. In order to be complete, the quantification of the reliability of the model hypotheses should be done in a "global" study.

These limitation must always be kept in mind for the conclusion of the method.

4 Experimental data treatment

As mentioned in subsection 3.3, including a large amount of experimental data coming from a large number of experiences, teams, and decades, is very problematic as the quality of these data highly differs from each others. Actually, the main problem with experimental data is not their accuracy but the experimental error bars, which are crucial in our analysis as they define the Σ matrix. Sometimes, these error bars are not representative of the real

accuracy and precision of the experimental measurement. Additionally, the error bars were not evaluated the same way for all of the experimental data. Some can be pure statistical error bars while other includes systematics, themselves implying some arbitrary evaluation of the experimental setup. It appears clearly that some error bars provided are badly evaluated when two (or more) experiments exclude each others by several σ. This issue has been partly addressed by Schnabel within the CHANDA framework with the possibility to rescale automatically experimental error bars when several data set are available for the same observables [9]. However, most of the reaction studied have only one set of experimental data available. Therefore, we needed a more general approach for cases where only one set of data is available for an observable. Unfortunately, to our knowledge, there is no mathematical approach allowing to provide systematic error bars for a set of experimental data based only on the experimental data themselves.

The only way to provide systematics when only one set of data is available it to make an arbitrary estimation of them from the knowledge of the experimental setup. In our case, it is not reasonable to reprocess the systematic error bars of all sets of data included in our analysis. Therefore, we propose here an alternative approach taking the error bars provided with the experimental data and applying an arbitrary algorithm to normalise those error bars. This algorithm rules that experimental data with error bars too small to be realistic should be treated as experimental data with large uncertainties as there were badly evaluated. Additionally, the confidence we have in those data decreases with the increasing unlikelihood of the error bars provided. Therefore, the algorithm uses arbitrary thresholds under which uncertainties are rescaled up to predefined levels. On the other hand, we decide to trust the realistic uncertainties provided by other experiments regardless of the differences of uncertainties evaluation.

In practice, all the relative uncertainties below 1% are considered as very unrealistic and are rescaled to 30%. Those between 1% and 5% are considered as unrealistic and rescaled to 20%. Those between 5% and 10% are considered as realistic but likely underestimated and are rescaled to 10%. Finally, relative uncertainties above 10% are considered as properly estimated and are conserved as there are. Remark that this approach forbids relative uncertainties to go below 10% whatever the efforts put in the reduction of systematic errors.

This approach is needed for a proper execution of our algorithm but it has to be kept in mind that this treatment of experimental data is arbitrary.

However, when the amount of experimental data is large, our tests showed that modifying these thresholds and the rescaling do not significantly modify our results as long has the new error bars can be considered as realistic.

5 Toy models

In order to validate the methodology without the limitations on the experimental data mentioned above, we developed various toy models for which we control the "experimental data" we use. In this section, they are used to illustrate the parameter optimisation procedure. This will also highlight relevant considerations for the data analysis. Note that we always start our toy models with very bad values for the initial parameter in order to emphasis the expected behaviour of our algorithm.

5.1 Toy model 1: Basic model

In Figure 1, an example of parameters optimisation is displayed in the context of a non-stochastic model f. Here the model takes as input two parameters p_1 and p_2 and makes the product. We start with *a priori* values $p_1 = -3 \pm 1$, $p_2 = 4 \pm 1$, and the "experimental" observation for $f(p_1, p_2)$ is 15 ± 2. We do not consider *a priori* correlation between p_1 and p_2.

Applying our algorithm, one can see the quick convergence of the parameter set (p_1, p_2) in Figure 1 thanks to the EM. The convergence happens in $(2.00, 5.75)$. We remark that, after the convergence $f(p_1, p_2) \simeq 11.5 \neq 15$. This is due to the parameters variances, which forbid the parameters to become unreasonable with respect to their *a priori* estimation. In real cases, this feature will avoid model parameters to become unphysical.

In the second stage of the algorithm, one can see in Figure 1 the Gibbs sampling exploring the possible parameters sets $\vec{p}_i$ around the optimal set, which has been estimated with the EM. The x and y width of the zone explored depends on the variance of p_1 and p_2 while the correlation between the parameters depends, of course, on their real correlation on the results but also on the variance of the experimental datum. This dependence on the experimental error bar is due to the fact that the likelihood for a set of parameter decreases quickly when the model prediction drift apart of experimental data when the experimental error bars are smalls. Therefore, these small experimental error bars will reduce the size of the acceptable domain in

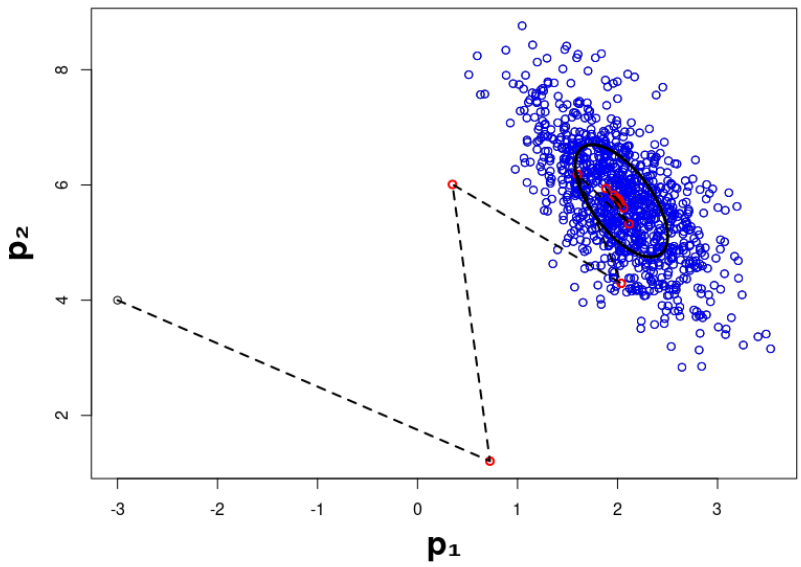

Figure 1: Example of parameters optimisation in a non-stochastic model. In red, the parameters sets along the Expectation Maximisation. In blue, the parameters sets along the Gibbs sampling. The black oval is the 1σ of the multivariate normal distribution used to fit the parameter set distribution obtained during the Gibbs sampling.

the parameters space and modify the dispersion of the parameter sets along the Gibbs sampling. The uncertainties for the parameters are provided by the covariance matrix of the distribution obtained with the Gibbs sampling.

5.2 Toy model 2 : Stochasticity

With the second toy model, we make the model slightly more complex with the addition of stochasticity in the results faking Monte Carlo processes. Here, the second toy model is a copy of the first toy model ($f(p_1, p_2) = p_1 \times p_2$) but with a Gaussian noise with a standard deviation of 1 above the output.

One of the main issue now is to evaluate the Jacobian since it is perturbed by the Gaussian noise. In the general case, the $d\vec{p}_i$ used to evaluate the Jacobian should not be too small in order not to be dominated by the noise

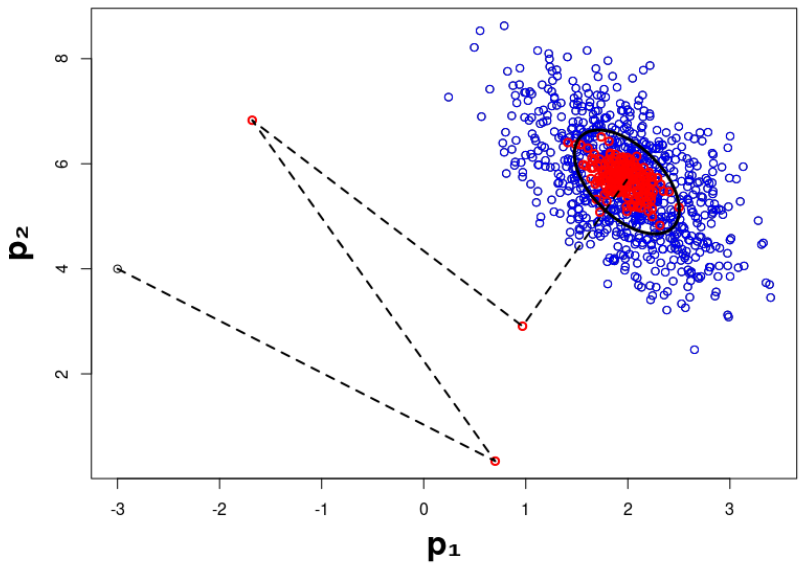

Figure 2: Same as Figure 1 with stochasticity.

but should also not be too large in order to get the local Jacobian and not a Jacobian averaged on a large scale.

The parameter set evolution is depicted in Figure 2. Note that running several times the algorithm would results in slightly different result due to the stochasticity of the model.

The second main problem is to know when the EM converged since the algorithm keeps modifying the parameters to compensate the stochasticity after the "convergence". Here, we have to draw a figure of merit along the optimisation and consider convergence is reached when the goodness of the model does not increase any more. Then, we can run several additional iterations in order to make the deconvolution between the noise and the optimal value for the parameters. This is actually optional as the Gibbs sampling also sample around the optimal parameter.

6 Parameter optimisation

The parameter optimisation might be very CPU intensive with modern model as INCL/ABLA. Especially in the case of "rare" observables. However, the principle remains the same as described in section 5.

In this study, we studied two different types of situation. First, a very favourable situation, but not fully physically meaningful, in order to demonstrate the feasibility and the capabilities of our method. Second, some cases representative of our long term objectives, in order to study the limits and difficulties. Our knowledge about the limits of our method and the difficulties we faced have been detailed in subsection 3.3.

6.1 Favourable case

For the first study, we chose the very favourable case of the subthreshold K^+ production as described by the experience carried out at LINP [10]. This case is very favourable for two main reasons. First, the subthreshold K^+ production is a very specific phenomenon, which implies just a few parameters. Additionally, there is a limited amount of experimental data (70 data points), all coming from the same experimental set up. This highly simplify both the mandatory analysis of the experimental data mentioned in subsection 3.3 and the results analysis. Second, the experimental data are very badly reproduced by INCL [11], which means there is a large room for improvement. Therefore, the figure of merit for this analysis should be very clear for interpretation.

On the other hand, this analysis have two limitations. First, the phenomenon studied is a very rare event with cross sections of the order of a few nanobarns. Additionally, each experimental data point corresponds to a different target and projectile energy, which requires individual calculations. Therefore, it is very CPU intensive to run INCL for this experiment. This forces us to limit the number of experimental data points used in our analysis to 24 representative points. Second, the parameters involved here have impact on other observables, which are not used in our analysis. Therefore, our approach completely neglect the very likely deterioration of these other observables when we will change the aforementioned parameters. Therefore, this first study is not physically complete. It will be a proof of concept showing that the approach we developed is functional for complex model like INCL.

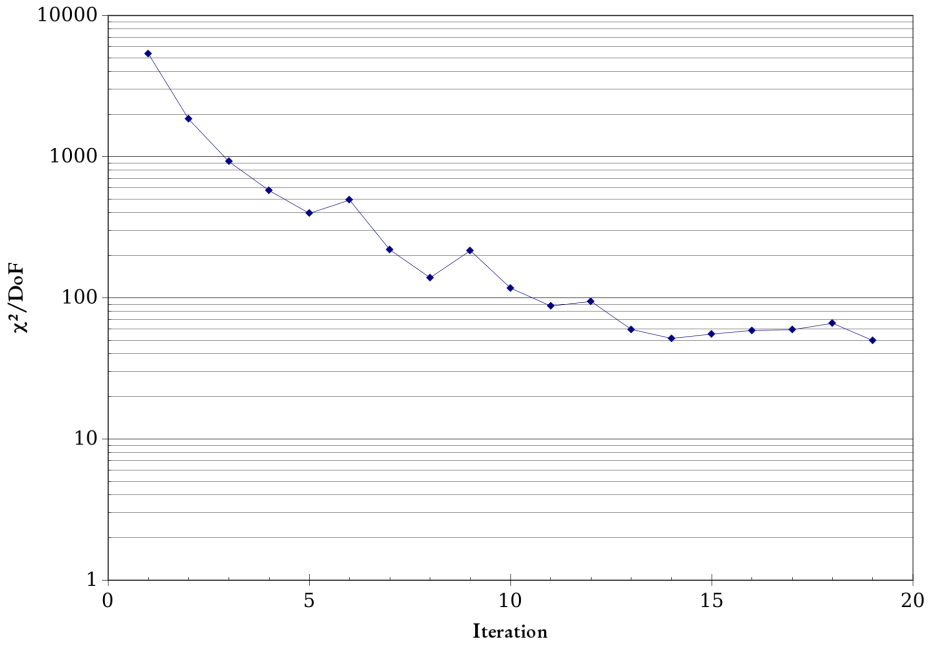

Figure 3: Figure of merit showing the evolution of the χ^2/DoF after each iteration. Iteration 1 corresponds to the initial version of INCL.

For this experience, we decided to consider four parameters to be optimised. Namely, we consider 3 scalars a_{NN}, $a_{\pi N}$, and $a_{\Delta N}$, which are multiplying factors applied to the strangeness production cross sections for NN, πN, and $\Delta N \rightarrow K + X$ reactions, respectively, and a fourth parameter, which is the Fermi momentum used in INCL.

Figure 3 depicts the evolution of the χ^2/DoF after each iteration of the algorithm as described in subsection 3.1. Here we only used the expectation maximisation due to CPU time restrictions. This calculation took 7 days using 20 cores. This means we will not be able to provide uncertainties on the parameters. After only a few iterations, one can see a huge improvement of the χ^2/DoF going from more than 5000 to roughly 50. The initial value of ~ 5300 is explained both by the poor initial description of the experimental data (factor 5 in average) and by the rather small experimental error bars (down to 3%). Regardless of the absolute value of the χ^2/DoF, the algorithm succeeds in improving the description of the experimental data by INCL as illustrated in Figure 4. On this figure, one can see that we started from

a model highly overestimating the experimental data and we ended with a pretty fair description of these data.

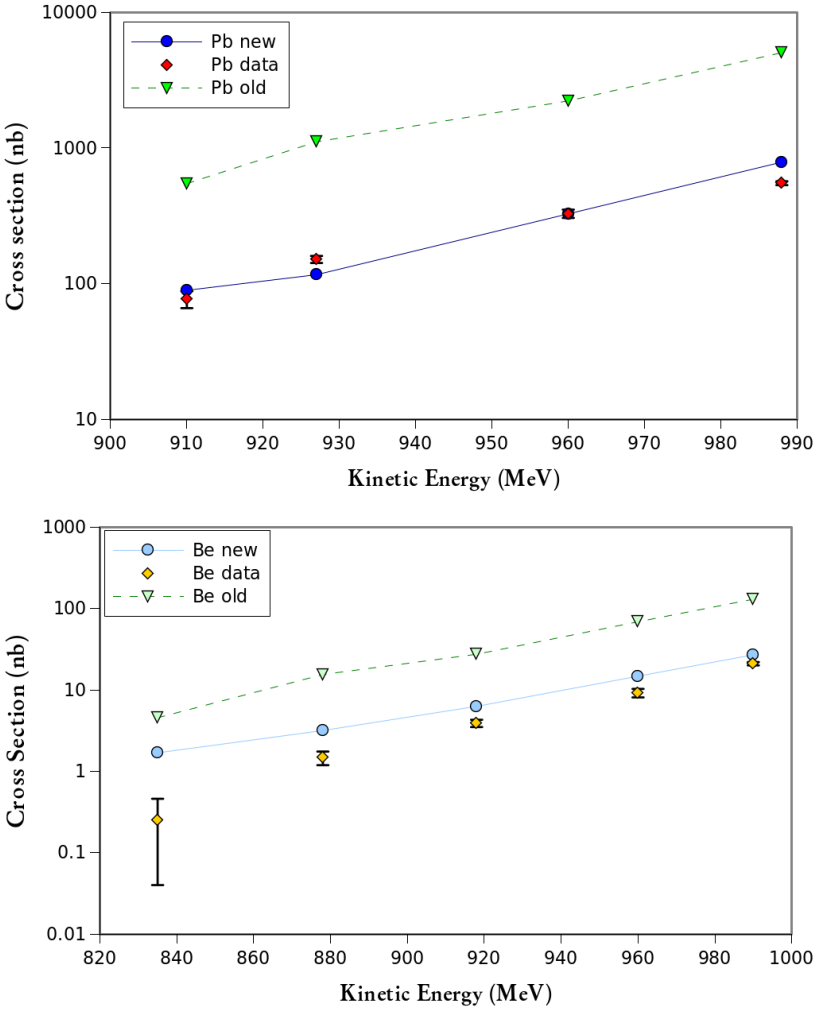

Figure 4: LINP experimental data [10] compared to INCL Before/After optimisation for (Top) Lead and (Bottom) Beryllium.

This clearly demonstrates the ability of our algorithm to optimise the parameters of a complex model like INCL.

6.2 Double differential neutron case

In a second step, we decided to apply our algorithm to the most important observables for the INCL/ABLA model applications: the double differential neutron cross section (DDNXS). Here, the parameters which can play a role are much more numerous than in the previous study. One can mention almost every single elementary double differential cross section (*e.g.*, $NN \to NN$, $\pi N \to N\pi\pi$, $\Delta N \to NN$, etc.), parameters dealing with the structure of the nucleus, the parameters ruling clustering, the freezing-out temperature in ABLA, etc. Almost everything matters for such a general feature. Here, it is not realistic in term of CPU to optimise every single parameter which might play a role in the DDNXS. It is therefore necessary to choose the parameters to be optimised. In our case, we chose to optimise the parameters for the $N\Delta \to NN$ cross section, the stopping time of the simulation, and the Fermi momentum. These parameters with enough leeway on their value are supposed to be those with the highest impact on the DDNXS and therefore the most interesting to study.

Once again, because of CPU time restriction, we limit the amount of experimental data to be taken into account. In this study, we work with the EXFOR data base [12]. Here, we decided to work with proton induced reactions with energies above 200 MeV and for target nuclei lighter than aluminium. This resulted in 7220 experimental data points taken into account. As mentioned in subsection 3.3, a careful study of the experimental data used and their possible correlations must be done beforehand, in order to obtain/use the best constraints. The most important in this preliminary study of the experimental data is to make sure the experimental error bars are consistent all together. If the error bars are globally over or underestimated, this will slightly modify the output of the optimisation, notably the parameter error bars, and the absolute value of the χ^2. However, it is of second order compare to the problem introduced by few unrealistically small error bars aside of much larger error bars as explained in section 4.

In the case studied here, we have experimental relative error bars down to 0.12% (EXFOR ID: E2387002, forward neutron emission at 117.5 MeV in the reaction p(137 MeV)+ C^{Nat} : 4.31 ± 0.00514 mb/MeV/sr). This kind of experimental data are toxic for our algorithm as they completely bias the value of the χ^2. Therefore, these problematic error bars need to be rescale. Otherwise, they can also be removed. We selected the first option. Our procedure to rescale experimental error bars is detailed in section 4. Our

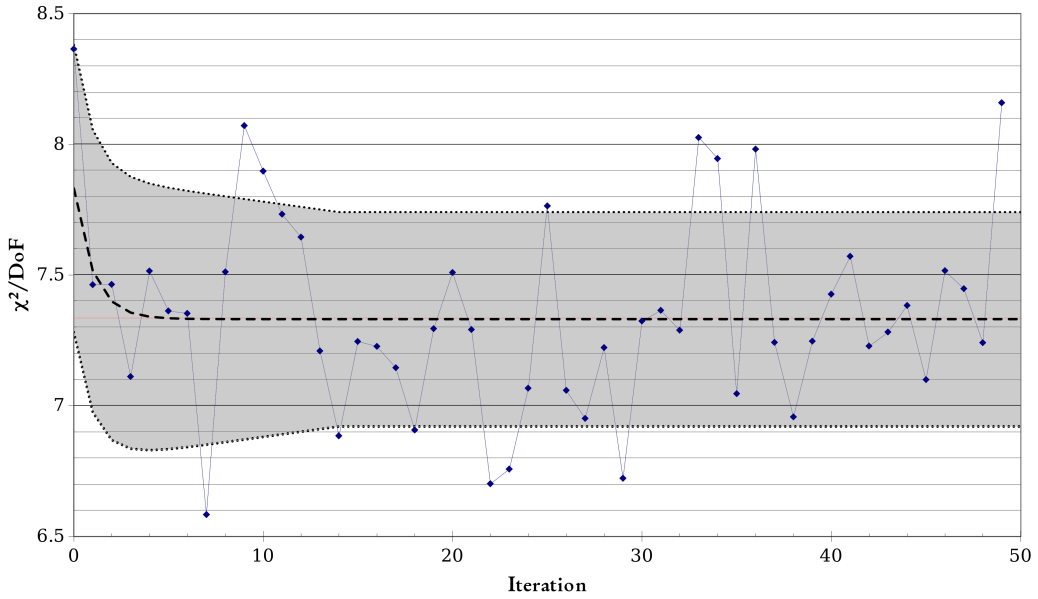

Figure 5: χ^2/Dof of INCL for the DDNXS evaluated for each iteration. The dashed line is a fit of the χ^2/Dof with the shape $a + e^{-bt}$, which is the shape expected with our algorithm. In gray, the 1 σ standard deviation evaluated for the statistics used along our study.

approach has not been pushed further forward as we are first interested in the feasibility of our method.

The execution of our algorithm on the CC-IN2P3 using 20 cores took roughly 60 hours.

First, we evaluated the model quality using the common reduced χ^2 throughout the algorithm. Noteworthy, the χ^2/Dof plotted in Figure 5 depends on the statistics used. A higher statistic reduces the statistical uncertainties and therefore the χ^2.

Using the standard values for the parameters, the χ^2/Dof is equal to 7.805 ± 0.125 with a standard deviation of 0.55. The uncertainty of the χ^2/Dof is due to the limited statistics. Using the optimal values as provided by our algorithm, the χ^2/Dof is now 7.34 ± 0.094 with a standard deviation of 0.41. This represents an improvement of 6% of the χ^2/Dof. The two χ^2/Dof provided have been estimated with the same statistic as in the algorithm to keep the coherence.

Second, the optimal parameters have been evaluated to 4.406 ± 0.131 (ini-

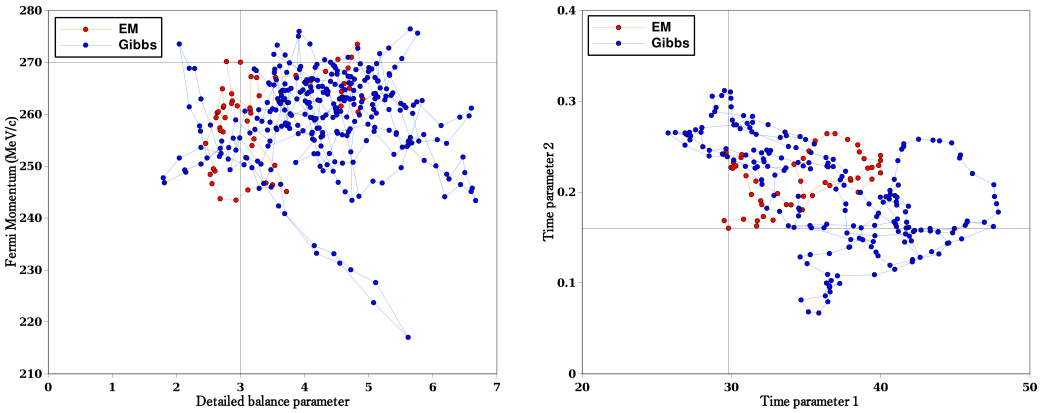

Figure 6: Parameters optimisation during the EM (red) and the Gibbs sampling (blue). The crosses indicates the initial values.

tially 3) for the detailed balance, to 266.4 ± 0.97 for the Fermi momentum (initially 270), and the stopping time parameters to $a = 37.13 \pm 0.59$ and $b = 0.226 \pm 0.005$ (initially $a = 29.8$ and $b = 0.16$). This is illustrated in the Figure 6 in red. Respectively, these new values mean that the cross section for the Δ-recombination ($\Delta N \rightarrow NN$) cross section has been increased of 50%, the maximal kinetic energy of nucleons has been slightly reduced and, the stopping time has been greatly increased ($t_{stop} = a \times A^b$, with A the mass number of the target nucleus). The uncertainties are due to the stochasticity of the model which is not fully compensated by a high number of iteration in the EM phase of the algorithm. In blue, we have the evolution of the parameters along the Gibbs sampling. This provide us the parameter set range within which the output of the model stays consistent with the experimental data. The *a posteriori* acceptability range for the parameters are provided by the standard deviation of the multivariate normal distribution obtained. Here, the 1 σ acceptability is: 0.986 (detailed balance), 8.822 (Fermi momentum), 6.061 (stopping time parameter 1), 0.0658 (stopping time parameter 2).

These uncertainties can be seen as a domain of validity being given the experimental data and the model, which is considered as a valid representation of the truth.

7 Summary and outlook

In this study, we developed an algorithm based on the Bayesian statistics able to optimise the parameters of a model using experimental data corresponding to the prediction of this model. This algorithm is able to fit the model prediction on the experimental data provided through the optimisation of the free parameters of the model. The objective of this algorithm is twofold: First, the algorithm determines optimal parameters which minimise the bias of the model and, by extension, the χ^2. Second, this algorithm aims at determining the uncertainties on the parameters.

This algorithm has been developed in a global framework, namely the Generalised Least Squares method, which allows an optimisation of the parameters for any model as long as these parameters are set free.

In our study, we first demonstrated the feasibility of such approach in a very suited case in which we reduced the χ^2 of the model by a factor 100. In a second stage, we study the case of the neutron double differential cross section with INCL in proton-induced reaction on light nuclei. It resulted in a reasonable improvement of the model prediction using thousands of experimental data with a reduction of the χ^2 by 6%. In this second case, we have been able to estimate the uncertainties of the free parameters used for the optimisation.

We also highlighted the limits of the approach with, first, high CPU requirements with several days of calculation with few tens of cores in the case of INCL. Another limit to the approach is the availability and the quality of the experimental data. Last but not least: the disparity of the quality of the experimental data is a toxic issue, which must be addressed before the use of our algorithm. This last point requires to exclude some toxic experimental data with unrealistically small error bars or to rescale these error bars in order to moderate the importance of the corresponding experimental data with respect to other data evaluated with more rigorous approach.

The specific case of INCL/ABLA has been studied with the optimisation of 4 parameters and more than 7000 experimental data from the EXFOR data base.

Once the parameters have been optimised, the model bias of the new version of INCL/ABLA can be estimated using the approach developed by Schnabel in the previous CHANDA framework.

A Derivation of the GLS method

The GLS method is the basis for inference in Bayesian networks of continuous variables with a multivariate normal prior distribution and linear relationships between variables. In this appendix, we will derivate from the Bayes theorem the Equation 7 and Equation 9, which are the base of our algorithm.

Lets assume a set of parameter of interest $\vec{p}$, a set of experimental data $\vec{\sigma}_{exp}$, and a model/function $\mathcal{M}$ that we assume perfect. Here, a perfect model means that $\mathcal{M}(\vec{p}_{true}) = \vec{\sigma}_{true}$.

The Bayes theorem gives the relation between, on the one hand, the posterior distribution of $\vec{p}$ called $\pi(\vec{p}|\vec{\sigma}_{exp})$, and, on the other hand, the prior distributions of $\vec{p}$, called $\pi_0(\vec{p})$, and $\vec{\sigma}_{exp}$, called $\pi(\vec{\sigma}_{exp})$, and the likelihood of $\vec{\sigma}_{exp}$ knowing $\vec{p}$, called $l(\vec{\sigma}_{exp}|\vec{p})$:

$$\pi(\vec{p}|\vec{\sigma}_{exp}) = \frac{l(\vec{\sigma}_{exp}|\vec{p}) \times \pi_0(\vec{p})}{\pi(\vec{\sigma}_{exp})} \tag{A.1}$$

Here, $\pi(\vec{\sigma}_{exp})$ is a scalar which guaranty the normalisation of $\pi(\vec{p}|\vec{\sigma}_{exp})$. Both the likelihood l and the prior distribution π_0 are supposed to be multivariate normal distributions. Therefore, we can write:

$$\pi_0(\vec{p}) \propto \exp\left(-\frac{1}{2}(\vec{p} - \vec{p}_{ref})^T \Sigma_p^{-1} (\vec{p} - \vec{p}_{ref})\right), \tag{A.2}$$

with $\vec{p}_{ref}$ the best estimation *a priori* of $\vec{p}$ and Σ_p the covariance matrix of $\vec{p}$, and:

$$l(\vec{\sigma}_{exp}|\vec{p}) \propto \exp\left(-\frac{1}{2}(\vec{\sigma}_{exp} - \mathcal{M}(\vec{p}))^T \Sigma_e^{-1} (\vec{\sigma}_{exp} - \mathcal{M}(\vec{p}))\right), \tag{A.3}$$

with Σ_e the covariance matrix of $\vec{\sigma}$ and $\mathcal{M}$ a perfect model.

Since the product of two (multivariate) normal distribution is also a (multivariate) normal distribution, we also have:

$$\pi(\vec{p}|\vec{\sigma}_{exp}) \propto \exp\left(-\frac{1}{2}(\vec{p} - \vec{p}_{opti})^T \Sigma_{opti}^{-1} (\vec{p} - \vec{p}_{opti})\right), \tag{A.4}$$

with $\vec{p}_{opti}$ and Σ_{opti} the optimal parameter set for the model knowing the experimental data set $\vec{\sigma}_{exp}$ and the corresponding covariance matrix, respectively.

Since the GLS method requires linear relationships between variables, we need to approximate the model $\mathcal{M}$ with a Taylor approximation:

$$\mathcal{M}(\vec{p}) = \mathcal{M}(\vec{p}_{ref}) + J_p(\vec{p} - \vec{p}_{ref}), \tag{A.5}$$

with J_p the Jacobian of the model.

Therefore, we can rewrite the likelihood as:

$$l(\vec{\sigma}_{exp}|\vec{p}) \propto \exp\left(-\frac{1}{2}(\vec{\sigma}_{exp} - \mathcal{M}(\vec{p}_{ref}) - J_p(\vec{p} - \vec{p}_{ref}))^T \right.$$
$$\left. \Sigma_e^{-1}(\vec{\sigma}_{exp} - \mathcal{M}(\vec{p}_{ref}) - J_p(\vec{p} - \vec{p}_{ref}))\right). \tag{A.6}$$

simplified as

$$l(\vec{\sigma}_{exp}|\vec{p}) \propto \exp\left(-\frac{1}{2}(H_{ref} - J_p\vec{p})^T\Sigma_e^{-1}(H_{ref} - J_p\vec{p})\right), \tag{A.7}$$

with the substitution of the constant term $H_{ref} = \vec{\sigma}_{exp} - \mathcal{M}(\vec{p}_{ref}) + J_p\vec{p}_{ref}$.

With a combination of equations A.1, A.2, A.4, and A.7 and knowing that $\pi(\vec{\sigma}_{exp})$ is a scalar, we have:

$$\exp\left(-\frac{1}{2}(\vec{p} - \vec{p}_{opti})^T\Sigma_{opti}^{-1}(\vec{p} - \vec{p}_{opti})\right) \propto \tag{A.8}$$
$$\exp\left(-\frac{1}{2}(\vec{p} - \vec{p}_{ref})^T\Sigma_p^{-1}(\vec{p} - \vec{p}_{ref})\right) \times \exp\left(-\frac{1}{2}(H_{ref} - J_p\vec{p})^T\Sigma_e^{-1}(H_{ref} - J_p\vec{p})\right)$$

It follows:

$$(\vec{p} - \vec{p}_{opti})^T\Sigma_{opti}^{-1}(\vec{p} - \vec{p}_{opti}) + C = \tag{A.9}$$
$$(\vec{p} - \vec{p}_{ref})^T\Sigma_p^{-1}(\vec{p} - \vec{p}_{ref}) + (H_{ref} - J_p\vec{p})^T\Sigma_e^{-1}(H_{ref} - J_p\vec{p}),$$

with C a constant of normalisation.

Since $\vec{p}$ is the only variable in Equation A.9, the coefficients must match for the terms with $(\vec{p})^T$ on the left side and those with $\vec{p}$ on the right side. We then have the four equations:

$$\Sigma_{opti}^{-1} = \Sigma_p^{-1} + J_p^T\Sigma_e^{-1}J_p \tag{A.10}$$
$$\Sigma_{opti}^{-1}\vec{p}_{opti} = \Sigma_p^{-1}\vec{p}_{ref} + J_p^T\Sigma_e^{-1}H_{ref} \tag{A.11}$$
$$(\vec{p}_{opti})^T\Sigma_{opti}^{-1} = (\vec{p}_{ref})^T\Sigma_p^{-1} + H_{ref}^T\Sigma_e^{-1}J_p \tag{A.12}$$
$$(\vec{p}_{opti})^T\Sigma_{opti}^{-1}\vec{p}_{opti} + C = (\vec{p}_{ref})^T\Sigma_p^{-1}\vec{p}_{ref} + H_{ref}^T\Sigma_e^{-1}H_{ref} \tag{A.13}$$

Using the Woodbury matrix identity with Equation A.10, we have:

$$\Sigma_{opti} = \Sigma_p - \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} J_p \Sigma_p \tag{A.14}$$

Multiplying Equation A.11 from left by Equation A.14, we get:

$$\begin{aligned}
\vec{p}_{opti} &= \left(\Sigma_p - \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} J_p \Sigma_p\right) \left(\Sigma_p^{-1} \vec{p}_{ref} + J_p^T \Sigma_e^{-1} H_{ref}\right) \\
&= \vec{p}_{ref} + \Sigma_p J_p^T \Sigma_e^{-1} H_{ref} - \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} J_p \vec{p}_{ref} \\
&\quad - \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} J_p \Sigma_p J_p^T \Sigma_e^{-1} H_{ref} \\
&= \vec{p}_{ref} - \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} J_p \vec{p}_{ref} \\
&\quad + \Sigma_p J_p^T \left(\Sigma_e^{-1} - (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} J_p \Sigma_p J_p^T \Sigma_e^{-1}\right) H_{ref} \\
&= \vec{p}_{ref} - \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} J_p \vec{p}_{ref} \\
&\quad + \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} \left((\Sigma_e + J_p \Sigma_p J_p^T) \Sigma_e^{-1} - J_p \Sigma_p J_p^T \Sigma_e^{-1}\right) H_{ref} \\
&= \vec{p}_{ref} - \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} J_p \vec{p}_{ref} + \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} H_{ref} \\
&= \vec{p}_{ref} + \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} (H_{ref} - J_p \vec{p}_{ref})
\end{aligned}$$

and, replacing H_{ref}, we finally obtain:

$$\vec{p}_{opti} = \vec{p}_{ref} + \Sigma_p J_p^T (\Sigma_e + J_p \Sigma_p J_p^T)^{-1} (\vec{\sigma}_{exp} - \mathcal{M}(\vec{p}_{ref})) \tag{A.15}$$

Remark that this equation is true only with the hypothesis of a linear model. However, most of realistic models can be approximated by a linear model only locally. Therefore, $\mathcal{M}(\vec{p}_{ref})$ must be estimated reversing the Equation A.5:

$$\mathcal{M}_{lin}(\vec{p}_{ref}) = \mathcal{M}(\vec{p}) + J_p(\vec{p}_{ref} - \vec{p}), \tag{A.16}$$

with J_p the Jacobian of the model in $\vec{p}$.

In order to simplify Equation A.15, we usually introduce the matrix of regression $\tilde{\Sigma}_i$ defined with Equation 5. Explicitly, the equation expands as:

$$\tilde{\Sigma}_i = J_i \, \Sigma \, J_i^T = \begin{pmatrix} \Sigma_e + J_p \Sigma_p J_p^T & J_p \Sigma_p \\ \Sigma_p J_p^T & \Sigma_p \end{pmatrix} = \begin{pmatrix} \tilde{\Sigma}_{DD_i} & \tilde{\Sigma}_{DI_i} \\ \tilde{\Sigma}_{ID_i} & \tilde{\Sigma}_{II_i} \end{pmatrix}. \tag{A.17}$$

Therefore, we obtain the Equation 7:

$$\vec{p}_{opti} = \vec{p}_{ref} + \tilde{\Sigma}_{ID_i} \left(\tilde{\Sigma}_{DD_i}\right)^{-1} \left[\vec{\sigma}_{exp} - \mathcal{M}_{lin}(\vec{p}_{ref})\right], \tag{A.18}$$

and the Equation 9 from Equation A.14:

$$\Sigma_{opti} = \tilde{\Sigma}_{II_i} - \tilde{\Sigma}_{ID_i} \left(\tilde{\Sigma}_{DD_i} \right)^{-1} \tilde{\Sigma}_{DI_i}. \tag{A.19}$$

As the difference $|\vec{p}_{opti} - \vec{p}|$ becomes smaller and smaller, the hypothesis of a linear model between $\vec{p}$ and $\vec{p}_{opti}$ becomes more and more correct, and therefore, the last two equations become more and more exact. This justifies the use of an iterative algorithm evaluating a linearisation of the model (Equation 2 : $\vec{T}_i$) and its Jacobian (Equation 4 : J_{p_i}) in $\vec{p}_i$, the best evaluation of the optimal parameters currently known and then, evaluating an improved $\vec{p}_{i+1}$ from $\vec{T}_i$ and J_{p_i} using Equation 7.

References

[1] MYRRHA project, Website.

[2] European commission: CHANDA, Website.

[3] European commission: ERINDA, Website.

[4] European commission: EUFRAT, Website.

[5] G. Schnabel, *"Estimating model bias over the complete nuclide chart with sparse Gaussian processes at the example of INCL/ABLA and double-differential neutron spectra"*, EPJ Nuclear Sci. Technol. 4, 33 (2018).

[6] C. Lawson, R. Hanson, *"Solving Least Squares Problems"*, Prentice-Hall, Englewood Cliffs, NJ, (1974).

[7] A. Dempster, N. Laird, D. Rubin, *"Maximum Likelihood from Incomplete Data Via the EM Algorithm"*, Journal of the Royal Statistical Society, Series B. 39 (1): 1-38, (1977)

[8] G. Schnabel, J.-C. David, D. Mancusi, and A. Boudard *"Report on the development of methodology for the uncertainty of high-energy models"*.

[9] G. Schnabel, *"Fitting and Analysis Technique for Inconsistent Nuclear Data"*, arXiv:1803.00960 [nucl-th]

[10] V. Koptev et al., *"Subthreshold K^+-meson production in proton-nucleus interactions"*, Zh. Eksp. Teor. Fiz. 94, 1-14 (1988).

[11] J. Hirtz et al., *"Strangeness production in the new version of the Liège Intra-Nuclear Cascade model"*, Phys. Rev. C 101, 014608 (2020).

[12] V.V. Zerkin and B. Pritychenko, *"The experimental nuclear reaction data (EXFOR): Extended computer database and Web retrieval system"*, NIM A 888, 31 (2018)