# 1 Error Model for Mass Spectrometry

## 1.1 Context

In order to understand the Nf-$\kappa$B signaling pathway of the cell a good model of the process has to be obatined. The state of the art aproche is used which is to define a structure of the model with certain parameters left to be fitted such that the model agrees with the messured data. This structure is generated by pure thought and prior knowleg of the systems. A vital part of this pursuite is to determine the functional dependencies of values in the model and there degree of freedome represented by the number of parameters that can adjust these dependencies. It is of high interesst to limit this number of paramters to a minimum that just allows the model to fit all realistic systems of the kind it is suppose to model but not more, as it could get "overfitted" which results in unrealistic predictions.

Since messured data is not free of noise this noise has to be accounted for when trying to fit the model. Thus the model should not predict the messured data exactly but the messured data should be destributed arround the prediction just as the messurment is distributed arround the real values due to the noise of messurment. Since these destributions are not know they have to be fittet as part of the model. The right type of destribution has to be picked and the parameters fittet such that it represents the destribution of the messurements arround the prediction and therefor hopefully the real value.

With the error model obtained we can calculate the probability of a certain messurent to be taken and hence the porobability to messure a set of certain values like the messurments already taken. This liklyhood to messure the existing data will be used as the goodness of fit. Thus fitting the selected model means trying to find a set of parameters for the model that maximizes this likelyhood.

## 1.2 Goal

Here we want to analyze the mechanics of the procedure to messure the data in order to obtain an understanding of the noise which is produced relative to the real values. We hope to derive a general functional depedency of the messured values to their errors with only a few paramters to be fitted. Experirience in this field of science has shown that such mechanistic approches ussually do not work perfectly due to the lack of accurate knowlege about the underlying process. Henc alternative modles will be worked out and compared by means of ACI, $\text{AIC}_C$, BIC, etc. If sufficent data is provided we

will also compute Shapiro-Wilk tests and the like to analyse the distribution of error arround an established prediction, e.g. multiple messurments of the same dilution of a certain protein. If ther is more than one of such data sets we could also analyze the errors dependency to properties of the protein as size or lipophilicity.

## 1.3 Mechanic Description

The cells which ought to be analyzed conatins a descrete numer $n$ of a protein of interest.