IV. INFLUENCE OF DATA UNCERTAINTIES ON FIELD DISTRIBUTION AND RESPONSES

Let us assume that uncertainties are associated with the system parameters \( p_j \) \((j = 1, 2, \ldots, J)\). These uncertainties may be very well represented by the so-called variance-covariance (symmetric) matrix

\[
C_p = \begin{bmatrix}
c_{11} & c_{12} & \cdots & c_{1J} \\
c_{12} & c_{22} & \cdots & c_{2J} \\
\vdots & \vdots & \ddots & \vdots \\
c_{1J} & c_{2J} & \cdots & c_{JJ}
\end{bmatrix}
\]  

(31)

with elements \( c_{ij} \) represented by the expected values

\[
c_{ij} = E[(p_x^{\mu} - p_j)(p_y^{\mu} - p_i)] \quad (i, j = 1, 2, \ldots, J)
\]  

(32)

The quantity \( p_j \) here represents the true value of the \( j \)th parameter, whereas \( p_x^{\mu} \), that deduced from experiment (or, more generally, inferred from any source of coherent information).

Assuming that the sensitivities are calculated with negligible error, expanding the field vector \( \mathbf{f} \) (which describes the state of the system considered) around a reference solution \( \mathbf{f}_0 \) (corresponding to the available parameter values \( p_x^{\mu} \)), gives

\[
\mathbf{f} = \mathbf{f}_0 + \sum_{j=1}^{J} \mathbf{f}_{ij} \delta p_j + \frac{1}{2} \sum_{j,k=1}^{J} \mathbf{f}_{jk} \delta p_j \delta p_k + \ldots
\]

(33)

where by \( \mathbf{f}_{ij} \) and \( \mathbf{f}_{jk} \) we have denoted the (total) derivatives of \( \mathbf{f} \) with respect to \( p_j \) and \( p_i \), \( p_k \), respectively. Assuming that \( \delta p_j (j = 1, 2, \ldots, J) \) represent the errors associated with the parameters, the estimated average value \( \mathbf{f} \) of \( \mathbf{f} \) will then result (assuming that the average \( \delta p_j = 0 \))

\[
\bar{\mathbf{f}} = \mathbf{f}_0 + \frac{1}{2} \sum_{j,i=1}^{J} \mathbf{f}_{ijj} \delta p_j
\]

(34)

whereas its variance

\[
\text{var}(\bar{\mathbf{f}}) = \sum_{j,i=1}^{J} \mathbf{f}_{ijj} \delta p_j \delta p_i
\]

(35)

So, even if no correlation exists among parameters (i.e., \( c_{ij} = \delta_{ij} \) [Kronecker symbol]), we easily see that the average value \( \bar{\mathbf{f}} \) does not generally coincide with the most probable value \( \mathbf{f}_0 \) (obtained on the basis of the best available parameter values) if the second order derivatives are different from zero.\(^5\) Calculating the derivative functions \( \mathbf{f}_{ij} \) is often unnecessary for uncertainty analysis, i.e., in those cases in which we are interested in analyzing a limited number of quantities \( Q_{\ell} (\ell = 1, 2, \ldots, L) \), representing responses, or functionals, linear or nonlinear with respect to \( \mathbf{f} \). In these cases expressions analogous to those above, relevant to average values \( \bar{Q}_\ell \) and their variance, \( \text{var}(Q_\ell) \), can be written

\[
\bar{Q}_\ell = Q_{\ell,0} + \frac{1}{2} \sum_{j,i=1}^{J} s_{\ell,j} c_{ij}
\]

(36)

\[
\text{var}(\bar{Q}_\ell) = \sum_{j,i=1}^{J} s_{\ell,j} s_{\ell,i} c_{ij} \quad (= S_{\ell} C S_{\ell}^T)
\]

(37)
where $s_{e,j}$ and $s_{e,i}$ represent the sensitivity coefficients as defined by Equations 21 and 30, respectively, while $S_e = [s_{e,1}, s_{e,2}, \ldots, s_{e,l}]$.

V. ADJUSTMENT METHODS

We shall derive now the adjustment methodology to be adopted in order to incorporate new experimental information into the set of system parameters, via their adjustments. To this purpose, we shall use well-known concepts of statistics, in particular, the methods based on the maximum likelihood function, assuming that prior and posterior experimental (or other coherent) information available is normally distributed. As is well known, the parameter estimators obtained in this way result with optimal properties, i.e., unbiasedness and minimal variance, with respect to estimators obtained by any other method.

The adjustment methodology which we are going to describe can be subdivided into two alternative approaches to be selected on the basis of the specific problem at hand. Following the nomenclature used by Linnik, these approaches will be referred to as method of reduction by the Lagrange multipliers and method of reduction by elements.

A. Method of Reduction by the Lagrange Multipliers

Consider a number $L$ of integral quantities, or responses, $Q_e(\ell = 1, 2, \ldots, L)$ relevant to one or more physical systems which can be described by set of $J$ parameters $p_j (j = 1, 2, \ldots, J)$. Assume that to these quantities $Q_e$ and parameters $p_j$, corresponding experimental values $Q_e^{ex}(\ell = 1, 2, \ldots, L)$ and experimental (or otherwise coherently determined) values $p_j^{ex} (j = 1, 2, \ldots, J)$ are available, together with their error (variance-covariance) estimates.

If $Q_e (\ell = 1, 2, \ldots, L)$ and $p_j (j = 1, 2, \ldots, J)$ represent true values, assuming for the moment that the inaccuracy associated with the theoretical model adopted for the calculations is negligible, we can express the quantities $Q_e$ as functions of $p_j$, i.e.,

$$Q_e = Q_e(p_1, p_2, \ldots, p_j) \quad (\ell = 1, 2, \ldots, L) \quad (38)$$

If we assume a given set of values** $p_0,j$, close enough to the true ones $p_j$, we may expand Equation 38 disregarding second and higher order terms and obtain

$$Q_e = Q_e(p_0,1, \ldots, p_0,J) + \sum_{j=1}^{J} \frac{\partial Q_e}{\partial p_j} \bigg|_{p_0} (p_j - p_0,j) \quad (\ell = 1, 2, \ldots, L) \quad (39)$$

The integral quantities $Q_e$ are assumed different from each other so that Equation 39 may be considered linearly independent. The values $Q_e (p_0,1, \ldots, p_0,J)$, calculated on the basis of values $p_0,j$, will be called $Q_e^{cal}$, i.e.,

$$Q_e^{cal} = Q_e(p_0,1, \ldots, p_0,J) \quad (40)$$

* For a general theory approach see, for example, Linnik (1961). For its applications to nuclear reactor analysis and survey of methods, see Gandini (1973, 1974), Drat et al. (1977) and Weisbin et al. (1978).

** These values are independent from the experimental ones, $p_j^{ex}$, although they may be set numerically equal to them.
Let us define the quantities

\[ y_j = \frac{P_j - P_{0,j}}{P_{0,j}} \quad (j = 1, 2, \ldots, J) \]  

(41)

\[ y_{j+\ell} = \frac{Q_{\ell} - Q_{\ell}^{cal}}{Q_{\ell}^{cal}} \quad (\ell = 1, 2, \ldots, L) \]  

(42)

\[ s_{\ell,j} = \frac{P_{0,j}}{Q_{\ell}^{cal}} \frac{\partial Q_{\ell}}{\partial P_j} \bigg|_{p_0} \quad (\ell = 1, 2, \ldots, L; \ j = 1, 2, \ldots, J) \]  

(43)

where the sensitivity coefficients \( s_{\ell,j} \) can be calculated via the GPT sensitivity methods described in the previous section. Introducing vectors

\[
y = \begin{bmatrix} y_p \\ y_Q \end{bmatrix}
\]

(44)

\[
y_p = \begin{bmatrix} y_1 \\ \vdots \\ y_J \end{bmatrix}
\]

(45)

\[
y_Q = \begin{bmatrix} y_{J+1} \\ \vdots \\ y_{J+L} \end{bmatrix}
\]

(46)

and the matrix

\[
W = [S - U]
\]

(47)

\( U \) being an \( L \times L \) unit matrix, while

\[
S = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,J} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,J} \\ \vdots & \vdots & \ddots & \vdots \\ s_{L,1} & s_{L,2} & \cdots & s_{L,J} \end{bmatrix}
\]

(48)

Equation 39 may be written, in vector form,

\[
Wy = 0
\]

(49)

Since we assumed that Equation 39 is linearly independent, we may say that the rank of matrix \( W \) is equal to the number, \( L \), of the constraints. Matrix \( S \) (and, therefore, matrix
\( y_{ij}^{\text{ex}} = \frac{p_{ij}^{\text{ex}} - p_{0,j}}{p_{0,j}} \quad (j = 1, 2, \ldots, J) \)  \( (50) \)

\[ y_{j+\ell}^{\text{ex}} = \frac{Q_{\ell}^{\text{ex}} - Q_{\ell}^{\text{cal}}}{Q_{\ell}^{\text{cal}}} \quad (\ell = 1, 2, \ldots, L) \]  \( (51) \)

Along with the arguments discussed previously, we now assume that the measured quantities \( Q_{\ell}^{\text{ex}} \) and \( p_{ij}^{\text{ex}} \) (and, therefore, \( y_{ij}^{\text{ex}} \)) are normally distributed with a dispersion matrix \( C_y \), the \((i, k)\)th element of which is given by the expected value

\[ c_{ik} = E[(y_{ij}^{\text{ex}} - y_i)(y_{kj}^{\text{ex}} - y_k)] \]  \( (52) \)

where \( y_i \) represents a true value. The \((i, i)\)th element of matrix \( C \) represents the variance of the experimental value \( y_i^{\text{ex}} \). Since no correlation is assumed between the quantities \( Q_{\ell}^{\text{ex}} \), having dispersion matrix \( C_Q \), and the quantities \( p_{ij}^{\text{ex}} \), having dispersion matrix \( C_p \), the (symmetric) dispersion matrix \( C_y \) will be of the form

\[ C_y = \begin{bmatrix} C_p & 0 \\ 0 & C_Q \end{bmatrix} \]  \( (53) \)

The likelihood function of vector \( y \) then results

\[ L(y|y^{\text{ex}}) = \frac{1}{(2\pi)^{(L+J)/2} |\det B_j|^{1/2}} \exp\left[ -\frac{1}{2} (y^{\text{ex}} - y)^T C_y^{-1} (y^{\text{ex}} - y) \right] \]  \( (54) \)

This function is maximized if we choose an estimator \( \hat{y} \) of \( y \) such that

\[ (y^{\text{ex}} - \hat{y})^T C_y^{-1} (y^{\text{ex}} - \hat{y}) = \text{minimum} \]  \( (55) \)

with the constraints

\[ W\hat{y} = 0 \]  \( (56) \)

Vector \( \hat{y} \) may be evaluated by means of the Lagrange multipliers \( \lambda_1, \lambda_2, \ldots, \lambda_L \). Introducing the vector

---

* The maximum likelihood function is formally identical with the distribution function of \( y^{\text{ex}} \), with the difference that, while in this distribution \( y^{\text{ex}} \) represents a vector of variables and \( y \) a vector of constant values, in the likelihood function, \( y^{\text{ex}} \) represents a vector of fixed values and \( y \) a vector of parameters of which optimal estimators are to be determined. These are obtained by searching that vector \( \hat{y} \) of estimators which maximize the likelihood function itself, as may be expected on simple intuitive grounds. In fact, since the values \( y^{\text{ex}} \) have really occurred, it seems natural that the probability density around these values (which, together with the covariance matrix, represents all the information available on which to base the estimate) should be high, on average.
the Lagrange function takes the form

$$\psi = (y^{ex} - y)^T C^{-1}_y (y^{ex} - y) - \lambda^T W y$$

(58)

Differentiating, we obtain

$$2(y - y^{ex}) C^{-1}_y - \lambda^T W = 0$$

(59)

which, together with the constraints

$$W y = 0$$

(60)

allows determination of vector $\hat{y}$. From Equation 59 we obtain

$$\hat{y} = y^{ex} + \frac{1}{2} C_y W^T \lambda$$

(61)

and introducing this equation into Equation 60:

$$\lambda = -2(W C_y W^T)^{-1} W y^{ex}$$

(62)

Finally, from Equation 61 we obtain the estimators

$$\hat{y} = y^{ex} - C_y W^T (W C_y W^T)^{-1} W y^{ex}$$

(63)

In particular, for what concerns $\hat{y}_p$, recalling Equations 47 and 53, we have

$$\hat{y}_p = y^{ex}_p + C_p S^T (C_Q + SC_p S^T)^{-1} (y^{ex}_Q - S y^{ex}_p)$$

(64)

If the starting parameter vector $p_0$ is assumed numerically coincident with $p^{ex}$, then $y^{ex}_p = 0$, and then this equation reduces to

$$\hat{y}_p = C_p S^T (C_Q + SC_p S^T)^{-1} y^{ex}_Q$$

(65)

For determining the error dispersion matrix $\hat{C}_y$, relative to the random vector $\hat{y}$, let us write Equation 63 in the form

$$\hat{y} = (U - T) y^{ex}$$

(66)

$U$ being a unit matrix and where

$$T = C_y W^T (W C_y W^T)^{-1} W$$

(67)

Since $T^T = T$ and $T^T = C_y^{-1} T C_y$, we easily obtain
\[ \mathbf{C}_y = (\mathbf{U} - \mathbf{T}) \mathbf{C}_y (\mathbf{U} - \mathbf{T})^T = \mathbf{C}_y - \mathbf{T} \mathbf{C}_y \]
\[ = \mathbf{C}_y - \mathbf{C}_y \mathbf{W}^T (\mathbf{W} \mathbf{C}_y \mathbf{W}^T) \mathbf{W} \]  
(68)

and, in particular,
\[ \mathbf{C}_p = \mathbf{C}_p - \mathbf{C}_p \mathbf{S}^T (\mathbf{C}_p + \mathbf{S} \mathbf{C}_p \mathbf{S}^T) \mathbf{S} \mathbf{C}_p \]  
(69)

We notice that with the above methodology, matrix \((\mathbf{SC}_y \mathbf{S}^T)\) or \((\mathbf{C}_p + \mathbf{SC}_p \mathbf{S}^T)\), the inverse of which enters in the adjustment formulations, Equations 63 and 68, is a square matrix the order of which is given by the number \((L)\) of constraints. Therefore, this method appears preferable to that known as the method of reduction by elements (see Section V.B) in those cases in which the number of constraints results significantly smaller than that \((J)\) of the parameters. The dependence of the solutions \(\hat{\mathbf{y}}_p\) and \(\mathbf{C}_p\) from matrices \(\mathbf{S}, \mathbf{C}_p,\) and \(\mathbf{C}_p\) is quite clear from Equations 64 and 69. The availability of correct values is then crucial if reliable estimates are to be obtained.**

** Method of Reduction by Elements

In the previous section we have seen how the optimal estimators \(\hat{y}\) of \(\mathbf{y}\) satisfy the condition, Equation 55, and the fundamental equations (constraints), Equation 56. Assuming that rank \((\mathbf{W}) = L\), and recalling that \(\mathbf{W}\) is an \(L \times (L + J)\) matrix, there exist, then, \(J\) independent variables, which in our case coincide with those with indexes \(1, 2, \ldots, J\) (i.e., corresponding to the \(J\) parameters \(p_j\)).

The other variables, with indexes \(J + 1, J + 2, \ldots, J + L\), can then be expressed as functions of the former ones, i.e.,
\[ y_{r+1} = \sum_{i=1}^{J} s_i \hat{y}_i \quad (r = 1, 2, \ldots, L) \]  
(70)

To these equations we may add the identities
\[ y_i = \sum_{i=1}^{J} \delta_i y_i \quad (i = 1, 2, \ldots, J) \]  
(71)

* There can be other reasons, however, requiring the use of the reduction by elements rather than that of the Lagrange multipliers, i.e., when matrix \(\mathbf{C}_p\) is singular, due to the complete lack of information concerning the errors of one or more parameters. We shall further comment on this in Section V.H.

** To evidence in particular, the importance of adequately taking into account the correct structure of the covariance matrices, let us consider the very simple case of one response \(y\) and two parameters \(p_1\) and \(p_2\). For simplicity, let us set the variances \(c_{y,11} = c_{y,22} = c_y\). If no correlation between \(p_1^*\) and \(p_2^*\) is assumed, it results immediately in
\[ \hat{y}_r = \frac{1}{\Delta} \left[ \frac{1}{s_1} \begin{pmatrix} s_2^2 & s_1 s_2 \end{pmatrix} \right] \]
\[ \mathbf{C}_r = \mathbf{C}_r \left[ \mathbf{U} - \frac{1}{\Delta} \begin{pmatrix} s_1^2 & s_1 s_2 \end{pmatrix} \begin{pmatrix} s_1 & s_2 \end{pmatrix} \right] \]

where \(s_1\) and \(s_2\) are the sensitivities, \(\mathbf{U}\) is a unit matrix, and \(\Delta = \frac{c_y}{s_1^2} + s_1^2 + s_2^2\). If, instead, full correlation between \(p_1^*\) and \(p_2^*\) is assumed, (i.e., \(c_{y,12} = 1\)), the following different expressions result
\[ \hat{y}_r = \frac{1}{\Delta + 2 s_2 s_2} \begin{pmatrix} s_1 s_2 \end{pmatrix} \]
\[ \mathbf{C}_r = \mathbf{C}_r \left[ \mathbf{U} - \frac{1}{\Delta + 2 s_2 s_2} \begin{pmatrix} s_1^2 + s_1 s_2 & s_1 s_2 + s_2 s_2 \end{pmatrix} \begin{pmatrix} s_1 & s_2 \end{pmatrix} \right] \]
If we interpret the parameters $y_j$ ($j = 1, 2, \ldots, J$) at the right-hand side of Equations 70 and 71 as coefficients, or elements, $a_j$, and define matrix $P$ as

$$ P = \begin{bmatrix} U \mid S \end{bmatrix} $$

$U$ being a $J \times J$ unit matrix. Equations 71 and 70 may be represented, in vector form, by the equation

$$ y = Pa $$

and our problem is transformed into that of evaluating an estimate $\hat{a}$ of vector $a$, once the vector of experimental values $y^e$ is known, such that

$$ (y^e - Pa)^T C_\gamma^{-1} (y^e - Pa) = \text{minimum} $$

The necessary condition for a minimum is of the form

$$ P^T C_\gamma^{-1} (y^e - Pa) = 0 $$

which represents the so-called system of normal equations. From this we may write the solution

$$ \hat{a} = (P^T C_\gamma^{-1} P)^{-1} P^T C_\gamma^{-1} y^e $$

Substituting into Equation 73, the estimate $\hat{y}$ can be obtained. In particular,

$$ \hat{y}_p = (S^T C_Q^{-1} S + C_\gamma^{-1})^{-1} (C_\gamma^{-1} y^e + S^T C_Q^{-1} y^e) $$

Relationship 76 allows determination of the covariance matrix relevant to vector $\hat{a}$. Recalling Equation 37, it results in:

$$ C_\gamma = (P^T C_\gamma^{-1} P)^{-1} $$

and, consequently, recalling Relationship 73,

$$ \tilde{C}_\gamma = P C_\gamma P^T = P (P^T C_\gamma^{-1} P)^{-1} P^T $$

In particular, we can write

$$ C_\gamma^{-1} = C_p^{-1} + S^T C_Q^{-1} S $$

Since Conditions 74 and 73 result equivalent to Conditions 55 and 49, the above method and that described in the previous section, making use of the Lagrange multipliers, coincide.

With this method, $J \times J$ dimension matrices have to be inverted, and then it is advantageous with respect to the method using the Lagrange multipliers in those cases in which $J$ is significantly smaller than $L$.

**C. Added Experimental Information**

Let us consider that added experimental information has to be included, besides that which has been used to obtain the adjusted quantities $\tilde{p}$ and $\tilde{C}_\gamma$. This new experimental information
will be represented by the quantities $Q_\ell^\alpha$ ($\ell = L + 1, L + 2, \ldots, L$), to which the error covariance matrix $C_\omega$ is associated. Let us extend, then, vector $\hat{y}_p^\alpha$ and matrix $C_\omega$ to include also this new experimental information. We can then define the new vector

$$\hat{y}_p^\alpha = \begin{bmatrix} y_p^\alpha \\ y_Q^\alpha \\ \hat{y}_p^\alpha \\ \hat{y}_Q^\alpha \end{bmatrix} = \begin{bmatrix} y_p^\alpha \\ \hat{y}_p^\alpha \end{bmatrix}$$ (81)

with vector $\hat{y}_Q^\alpha$ formed with elements $y_Q^\alpha$, given by Equation 51, with indexes $\ell = L + 1, \ldots, L$, and the new error covariance matrix

$$C_\omega = \begin{bmatrix} C_\omega & 0 \\ 0 & \hat{C}_\omega \end{bmatrix}$$ (82)

where we have assumed that there is no correlation between this new information and the previous one. If we correspondingly define also the extended sensitivity matrix

$$S = \begin{bmatrix} S_p \\ S_Q \end{bmatrix}$$ (83)

$S$ representing the matrix of sensitivity coefficients relevant to the new quantities $Q_\ell$ ($\ell = L + 1, \ldots, L$) considered, for the new estimators $\tilde{y}_p$ and $\tilde{C}_p$, we could consider expressions analogous to those determined in the previous section. Considering the method of reduction by elements, substituting Expression 81 through 83 into Equations 77 and 80, gives

$$\tilde{y}_p = (C_p^{-1} + S^T C_Q^{-1} S + \hat{S}^T \hat{C}_Q^{-1} \hat{S})^{-1} (C_p^{-1} \hat{y}_p^\alpha + S^T C_Q^{-1} y_Q^\alpha + \hat{S}^T \hat{C}_Q^{-1} \hat{y}_Q^\alpha)$$

$$= (\tilde{C}_p^{-1} + \hat{S}^T \hat{C}_Q^{-1} \hat{S})^{-1} (C_p^{-1} \hat{y}_p^\alpha + S^T C_Q^{-1} y_Q^\alpha + \hat{S}^T \hat{C}_Q^{-1} \hat{y}_Q^\alpha)$$

$$= (\tilde{C}_p^{-1} + \hat{S}^T \hat{C}_Q^{-1} \hat{S})^{-1} (C_p^{-1} \hat{y}_p^\alpha + S^T C_Q^{-1} y_Q^\alpha + \hat{S}^T \hat{C}_Q^{-1} \hat{y}_Q^\alpha)$$ (84)

$$\tilde{C}_p^{-1} = (C_p^{-1} + S^T C_Q^{-1} S + \hat{S}^T \hat{C}_Q^{-1} \hat{S})$$

$$= (C_p^{-1} + \hat{S}^T \hat{C}_Q^{-1} \hat{S})$$ (85)

The corresponding, equivalent expressions relevant to the Lagrange multipliers method of reduction can be written

$$\bar{y}_p = \bar{y}_p + \bar{C}_p \bar{S}^T (\bar{C}_Q + \tilde{S} \bar{C}_p \bar{S}^T)^{-1} (\bar{y}_Q - \bar{S} \bar{y}_p)$$ (86)

$$\bar{C}_p = \bar{C}_p - \bar{C}_p \bar{S}^T (\bar{C}_Q + \tilde{S} \bar{C}_p \bar{S}^T)^{-1} \hat{S} \bar{C}_p$$ (87)

The above results indicate that the new information can be added without the need of rerunning the entire problem. Rather than from $\hat{y}_p^\alpha$ and $C_\omega$, we can then restart from the
adjusted parameters and corresponding covariance matrix $\hat{y}_p$ and $\hat{C}_p$, respectively, and proceed with the same methodology adopted for determining them. This, more generally, means that independent experimental information can be considered piecewise in a series of subsequent adjustments, this giving the same results obtainable with the one run calculation. This can be particularly important when using the method of the Lagrange multipliers, since in this case the size of the matrices to be inverted can be greatly reduced.

The above recurrent approach by which subsequent updated estimates can be obtained from prior (independent) knowledge can be interpreted as a recurrent Bayesian inference procedure (see Appendix IV) or as subsequent Kalman filtering runs.*

D. Lognormal Distribution

There may be circumstances in which, although physically known as positive, some parameters result negative after the adjustment procedure. This can be due, for example, to large inaccuracies associated with the starting, $a$ priori, parameters. In these circumstances, in order to make use of the information concerning the positiveness of the parameters, their logarithmic value

$$z_j = \ln \left( \frac{p_{j}}{p_{0j}} \right) \quad (j = 1, 2, ..., J)$$  \hspace{1cm} (88)

can be considered\textsuperscript{11-15} and the quantities

$$z_j^* = \ln \left( \frac{P_{j}^*}{p_{0j}} \right) \quad (j = 1, 2, ..., J)$$  \hspace{1cm} (89)

are assumed normally distributed. The likelihood function in this case could be assumed the same as that (Equation 54) considered previously, with vectors $z$ and $z^*$ in place of $y_p$ and $y_p^*$ entering vectors $y$ and $y^*$, respectively. In fact, we can easily see that the sensitivities of $Q_i$ with respect to $z_i$ ($j = 1, 2, ..., J$) result in:

$$\frac{1}{Q^*_j} \frac{\partial Q_i}{\partial z_j} = \left. \frac{\partial Q_i}{\partial p_j} \frac{\partial p_j}{\partial z_j} \right|_{p_0} = \left. \frac{\partial Q_i}{\partial p_j} \right|_{p_0}$$  \hspace{1cm} (90)

which corresponds to the same sensitivity coefficients $s_{ij}$ (Equation 43) previously considered. Quite analogously, recalling Relationship 88, and that covariance matrix $C_p$ is relevant to relative errors, it results in:

* The so-called Kalman filtering\textsuperscript{11} is generally applied to dynamic systems in which independent observations are made at subsequent times. This method has strong analogies with that described above. Considering Equations 86 and 87 (relevant to the method of the Lagrange multipliers) and interpreting $y_p^*$ and $y_p^*$ as vectors of quantities determined by measurements at times $t_{i-1}$ and $t_i$, respectively, we can rewrite them in the form

$$y_p = y_p^* - K_i(y_p^* - S_{i-1}y_{i-1})$$

$$C_p = C_p^* - S_{i-1}K_iS_{i-1}$$

where

$$K_i = C_p^{-1}S_{i-1}^T(S_{i-1}C_p^{-1}S_{i-1}^T + C_{i-1})^{-1}$$

Matrix $K_i$ is referred to as the "Kalman gain", whereas the difference $(y_p^* - S_{i-1}y_{i-1})$ is called the "filter innovation" since it represents the new information incorporated in the measurements at $t_i$, which are potentially of use in estimating $y_{i-1}$.
\[ C_c = C_y \] (91)

So, the adjustment methodology previously derived can be adopted also for this case, and the same formulations considered with vector \( z \) and \( \bar{z} \) in place of \( y_p \) and \( y_q \). The adjusted parameters \( \bar{p}_j \) will then be given by the expression

\[ \bar{p}_j = p_{0j} e^{\bar{z}_j} \quad (j = 1, 2, \ldots, J) \] (92)

For small relative corrections (for which negative values \( \bar{p}_j \) are not expected) we can write

\[ \bar{p}_j = p_{0j}(1 + \bar{z}_j) \quad (j = 1, 2, \ldots, J) \] (93)

i.e., in this case the adjustments with the normal and the lognormal distributions coincide.

For what concerns the covariance matrix, it is easy to see that in any case it is

\[ \bar{C}_p = C_y \] (94)

It should be noted that the lognormal distribution represents a relatively simple way to inglobe the prior knowledge that the parameters to be adjusted are nonnegative. A correct distribution along with information theory concepts would be that given by a truncated gaussian (see Appendix III).

E. The Consistent Method

It may happen that, in place of \( p_j \) \((j = 1, 2, \ldots, J)\), another set of \( p_j \) \((j = 1, 2, \ldots, J)\), parameters, e.g., \( q_k \) \((k = 1, 2, \ldots, K)\), is used, more adequately defining the system, or systems, considered. Parameters \( p_j \) are then considered as function of \( q_k \), i.e.,

\[ p_j = p_j(q_1, q_2, \ldots, q_k) \quad (j = 1, 2, \ldots, J) \] (95)

As an example, we may recall the definition of the energy multigroup cross-sections as functions of the basic nuclear parameters (resonance widths, neutron temperatures, etc.) which enter the nuclear model.\(^6\) The methods described above can be easily extended to this case. To this purpose, we need to define the \( J \times K \) sensitivity matrix \( Z \), with the \( j \), \( k \text{th} \) element

\[ z_{jk} = \frac{\partial p_j}{\partial q_k} \quad (j = 1, 2, \ldots, J; \ k = 1, 2, \ldots, K) \] (96)

where \( q_{0,k} \) as well as \( p_{0,j} \) represent given values assumed close to the expected (true) values. Here it is assumed that \( q_{0,k} \neq 0 \). Expanding Equation 95 we than obtain, to first order

\[ \frac{p_j - p_{0,j}}{p_{0,j}} = \sum_{k=1}^{K} z_{jk} \frac{q_{k} - q_{0,k}}{q_{0,k}} \] (97)

or, in vector notation,

\[ y_p = Z y_q \] (98)

where the \( j \text{th} \) component of \( y_p \) is given by Equation 41, while the \( k \text{th} \) one of vector \( y_q \) by the expression
\[ y_{q,k} = \frac{q_k - q_{0,k}}{\partial q_{0,k}} \quad (k = 1, 2, \ldots, K) \] (99)

Assuming the covariance matrix \( C_q \) is given, the adjustment formulations derived previously can still be used if we substitute the sensitivity matrix \( S \) with \( SZ \) and \( B_p \) with \( B_q \).

**F. Consistency Test**

After an adjustment procedure is made and new estimates of the parameters have been obtained, tests of consistency of the results with the starting hypotheses and with the adjustment procedure itself can be made. In this respect, special consideration should be given to the well-known \( \chi^2 \) test, so called since, if the quantities of \( \hat{y}_p \) and \( \hat{y}_q \), represented by vector \( \hat{y} \) given by Equation 63, are normally distributed, then the random quantity

\[ R = (\hat{y} - y^*)^T C_y^{-1} (\hat{y} - y^*) \] (100)

is distributed as \( \chi^2 \), i.e., with \( L \) degrees of freedom. Therefore, its expected value \( E(R) \) is equal to the number, \( L \), of constraints (assumed independent).\(^6\) According to this, we can test the assumptions made with the above adjustment procedure, i.e.,

1. The values \( P^j \) (\( j = 1, 2, \ldots, J \)) represent random quantities normally distributed.
2. The values \( Q^\ell \) (\( \ell = 1, 2, \ldots, L \)) represent random quantities normally distributed.
3. No systematic errors are present in the experimental information, and the dispersion matrices are adequate.
4. Quantities \( Q \) can be expanded linearly around \( Q^\ell \), and the accuracy of the sensitivity coefficients is adequate.

The procedure to be adopted according to the current methods of hypothesis testing is standard. Assumed an apriori probability, or confidence coefficient, \( r_0 \) (e.g., 0.9), a value \( \alpha_L \) is found such that the integral

\[ \int_0^{\alpha} \chi^2(x)dx = r_0 \] (101)

If the result obtained by this procedure shows that the value \( R \) falls within the interval \((0, \alpha_L)\), and, therefore, as is said in these cases, the probability falls in the \( r_0 \) level of significance, there are serious indications that the assumptions made are adequate. If, on the other hand, the value \( R \) falls out of the interval \((0, \alpha_L)\), and, therefore, the probability falls in the \((1 - r_0) \) level of significance, there are serious indications that one or more of the assumptions made are not satisfied. According to specific judgment, this method of testing may then be used effectively, in particular, for indicating the probable presence, or not, of systematic errors and helping to sort them out. It can as well be used for testing different system parameter sets of subsets available to help an optimal selection.

**G. Nonnormality of Errors**

A comment seems appropriate regarding the possible nonnormality of errors of the experimental data. In the above derivations normal distributions have been assumed in the likelihood function. As is well known, in this case the estimators \( \hat{y}_k \) for a wide class of unbiased estimators result optimal in a certain sense, i.e., the following probability relationship holds:\(^6\)

\[ P(\hat{y}_k - y_k \leq \epsilon) \geq P(y_k - y_k \leq \epsilon) \quad (k = 1, 2, \ldots, J + L) \] (102)
where $y_i^*$ represent a different unbiased estimator of $y_i$, and $\varepsilon$ is a small quantity. If the values $y_i^*$ are no more normally distributed, Equation 102 is no more generally valid, although the estimators $\tilde{y}_i$ still exhibit a few important properties, i.e.,

1. They maintain their unbiasedness; this may be easily verified from Equation 63, i.e.,
   $$E(\tilde{y}) = y.$$
2. They are characterized by minimal variance among all unbiased linear estimators in accordance with the theorem of Neyman and David (see Reference 6).
3. For a wide number of independent experimental data, they result asymptotically normally distributed, in accordance with the Liapunov theorem.\footnote{As previously mentioned, by $p_i^*$ are intended not merely measured quantities, but also parameter evaluations based on theoretical models and to which error estimates are associated (possibly by fitting with some experimental points).}

These properties are highly significant in the applications since they allow us, in general, to say that by the procedure followed above the estimators $\tilde{y}_i$ may be improved consistently with the improvement of the experimental information, whereas the quantities calculated using them may be assessed in terms of normal distribution and minimal variance. We can also say, inversely, that whatever the distribution of $y_i^*$ may be, the same above estimators would be obtained by imposing the properties of unbiasedness and that of minimal variance.

## H. Systematic Errors

As seen in Section V.F, the $\chi^2$ test provides a criterion by which we can verify the consistency of the adjustment procedure. If the $\chi^2$ value falls well out of the interval corresponding to a given level of significance, we may have serious suspicions that one or more assumptions implied with the adjustment procedure are not satisfied. The most stringent among these assumptions is certainly that concerning the absence of systematic errors in the quantities $p_i^*$ and/or $Q_i^*$. We shall show in the following how the same $\chi^2$ test can be used to help sort out these errors. Consider first that a systematic error affects one of the measurements $Q_i^*$. Its identification can be easily made by observing the large $\chi^2$ decrease when it is taken out of the adjustment procedure. The $\chi^2$ test could be routinely adopted to this purpose if we consider different (independent) pieces of information in subsequent adjustments, as mentioned in Section V.C. If, after this procedure, the presence of a systematic error is still suspected, the $\chi^2$ maintaining significantly out of the expected range, a close check on the parameter values $p_i^*$ should be done, to possibly identify those more likely to be affected by systematic errors. It may be, for instance, that a set of parameters is not correctly normalized, so that it is their common normalization which should be corrected. Or, likewise, it may be that a common parameter, to which they are correlated, has been badly estimated (e.g., the neutron temperature, from which the high-energy inelastic scattering cross-section depends\footnote{As previously mentioned, by $p_i^*$ are intended not merely measured quantities, but also parameter evaluations based on theoretical models and to which error estimates are associated (possibly by fitting with some experimental points).}). A $\chi^2$ test can then be made to decide on such hypothesis, if we can identify these parameters by a more or less accurate description of the systematic errors. A solution to this problem was proposed by Mitani and Kuroi\footnote{As previously mentioned, by $p_i^*$ are intended not merely measured quantities, but also parameter evaluations based on theoretical models and to which error estimates are associated (possibly by fitting with some experimental points).} in relation to neutron cross-section adjustments.

In order to find a most general procedure, let us assume that the measured quantities $p_i^*$ ($i = 1, 2, ..., J$) can be interpreted in terms of corresponding unbiased values $p_i^*$ ($i = 1, 2, ..., J$), to which the covariance matrix $C_p$ is associated, and of parameters $p_s^*$ ($s = 1, 2, ..., S \leq J$ and assumed also $\leq L$), accounting for systematic error deviations when different from an assigned set of constants $p_{s}$, i.e., in vector notation,

$$p^* = \psi(p^*, p^e)$$

(103)
where \( \psi \) is a given vector function such that, if \( \mathbf{p}^0 = \mathbf{p}_0^0 \),

\[
\psi(\mathbf{p}^0, \mathbf{p}_0^0) = \mathbf{p}'
\]  

(104)

Expanding around \( \mathbf{p}_0^0 \), we obtain, to first order,

\[
\mathbf{p}^{**} - \mathbf{p}^{*} = \mathbf{R}(\mathbf{p}^{0} - \mathbf{p}_0^0)
\]  

(105)

where

\[
\mathbf{R} = \begin{bmatrix}
\frac{\partial \psi_1}{\partial p_1^0} & \frac{\partial \psi_2}{\partial p_1^0} & \cdots & \frac{\partial \psi_L}{\partial p_1^0} \\
\frac{\partial \psi_1}{\partial p_2^0} & \frac{\partial \psi_2}{\partial p_2^0} & \cdots & \frac{\partial \psi_L}{\partial p_2^0} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial \psi_1}{\partial p_S^0} & \frac{\partial \psi_2}{\partial p_S^0} & \cdots & \frac{\partial \psi_L}{\partial p_S^0}
\end{bmatrix}
\]  

(106)

We define then the quantities

\[
y_j^{*} = \frac{\mathbf{p}_j^{**} - \mathbf{p}_j^{0}}{\mathbf{p}_j^{0}} \quad (j = 1, 2, \ldots, J)
\]  

(107)

\[
y_s^{0} = \mathbf{p}_s^{0} - \mathbf{p}_s^{0} \quad (s = 1, 2, \ldots, S)
\]  

(108)

so that we can write

\[
y^{**} = \mathbf{y}^{*} - \mathbf{Ry}^{0}
\]  

(109)

where

\[
\mathbf{R} = \left\{ \text{diag} \left[ \frac{1 \cdots 1}{\mathbf{p}_{0,1} \cdots \mathbf{p}_{0,J}} \right] \right\} \mathbf{R}
\]  

(110)

For example, in case there is the suspect that a normalization error affects, the first \( 1 \) \((<J)\) parameter measurements \( \mathbf{p}_j^{**} \) \((j = 1, 2, \ldots, J)\), we would obtain, having assumed \( \mathbf{p}^{0} = \mathbf{p}_0 \),

\[
\mathbf{p}_j^{**} = \mathbf{p}_j^{0} \mathbf{p}_j^{*} \quad (j = 1, 2, \ldots, J)
\]

\[
\mathbf{p}_0^{0} = 1
\]

\[
\mathbf{y}_j^{*} = \mathbf{y}_j^{**} - \begin{bmatrix}
1 & 1 \\
0 & 0
\end{bmatrix}
\]  

(111)
whereas if, for the first and second measurements $p_j^{**} \,(j = 1, 2)$, we suspect that $p_j^{**} = p_j^v + p_j^u = 0$, it would be

$$
p_{0,s} = 0 \quad (s = 1, 2)
$$

$$
\begin{bmatrix}
\frac{1}{p_{0,1}} & 0 \\
0 & \frac{1}{p_{0,2}}
\end{bmatrix}
$$

$$
y_p^v = y_p^{**} -
\begin{bmatrix}
y_1^v \\
y_2^v
\end{bmatrix}
\begin{bmatrix}
y_1^v \\
y_2^v
\end{bmatrix}
$$

Following the method suggested by Mitani and Kuroi,\textsuperscript{10} we shall first introduce $y^v$, as given by Expression 109, in place of $y^{**}$ in the likelihood function. The following expression is obtained:

$$
L = \alpha \exp \left\{ -\frac{1}{2} \left[ (y_Q^v - y_Q)^T C_Q^{-1} (y_Q^v - y_Q) + (y_p^{**} - y_p - R y^u)^T C_p^{-1} (y_p^{**} - y_p - R y^u) \right] \right\}
$$

(111)

where $\alpha$ represents a constant value. Since

$$
y_Q = S y_p
$$

(112)

setting, for simplicity, in Equation 41 $p_{0,j} = p_j^{**}$ (so that $y_p^{**} = 0$), substituting in Equation 111 and deriving with respect to $y_p$ and $y^v$, we obtain the normal equations

$$
S^T C_Q^{-1} (y_Q^v - S y) - C_p^{-1} (y_p + R y^u) = 0
$$

(113)

$$
R^T C_p^{-1} (y + R y^u) = 0
$$

(114)

From the second equation we determine the optimal estimate

$$
y^v = -(R^T C_p^{-1} R)^{-1} R^T C_p^{-1} y_p
$$

(115)

where, since it is assumed that rank $(R) = S$, the matrix in parenthesis may not be singular. Substituting this expression in Equation 113, gives

$$
[S^T C_Q^{-1} S + C_p^{-1} - C_p^{-1} R (R^T C_p^{-1} R)^{-1} R^T C_p^{-1}] y_p = S^T C_Q^{-1} y_Q^{**}
$$

(116)

which, apart from the third term in the sum in brackets, is formally equivalent to Expression 77 (after setting $y_p^{**} = 0$), i.e., to the solution in the case where no systematic errors are assumed. This means that the expressions which give the optimal estimates $y_p$ (and $y_Q$) and the dispersion matrix $C_p$ result equivalent to those already derived for the unbiased case, provided the dispersion matrix $C_p^{-1}$ is replaced by the matrix
\[(C_p^n)^{\dagger} = C_p^{-1} - C_p^{-1}R(R^T C_p^{-1}R)^{-1}R^T C_p^{-1} \quad (117)\]

As a consequence, once matrix \((C_p^n)^{\dagger}\) has been evaluated, an adjustment code already developed for unbiased problems may be used. On the other hand, it can be easily verified that, substituting Equation 115, the likelihood function, Equation 111, can be written

\[L = \alpha \exp \left\{ -\frac{1}{2} \left[ (y_Q^{ex} - y_Q)^T C_Q^{-1} (y_Q^{ex} - y_Q) + (y_p^{ex} - y_p)^T (C_p^n)^{\dagger} (y_p^{ex} - y_p) \right] \right\} \quad (118)\]

Since the number of degrees of freedom in now given by the difference

\[r = \text{rank}(S) - \text{rank}(R) \quad (119)\]

the scalar product

\[R = (\tilde{y} - y^{ex})^T C_p^{-1} (\tilde{y} - y^{ex}) \quad (120)\]

(with \((C_p^n)^{\dagger}\) in place of \(C_p^{-1}\) in the expression for \(C_p^{-1}\)) results having a \(\chi^2\) distribution with \(r\) degrees of freedom. A \(\chi^2\) test can then be made to evaluate the validity of the systematic error hypothesis considered.

\* It can be easily verified that matrix \((C_p^n)^{\dagger}\) is singular. In fact, multiplying Equation 117 on the left by \(R^T\) gives

\[R^T (C_p^n)^{\dagger} = R^T C_p^{-1} - R^T C_p^{-1} = 0\]

If \((C_p^n)^{\dagger}\) could be inverted, multiplying on the right by \(C_p^n\) would give \(R^T = 0\), which contradicts the hypothesis that \(R\) is not a null matrix. It follows that for these cases the method of reduction by elements (Section V.B) should be applied since with it the inverse of the error covariance matrix is used.