approximately half the interval between the two discrepant values.

We emphasize that our scaling procedure for *errors* in no way affects central values. And if you wish to recover the unscaled error $\delta\overline{x}$, simply divide the quoted error by S.

(b) If the number M of experiments with an error smaller than δ_0 is at least three, and if $\chi^2/(M-1)$ is greater than 1.25, we show in the Full Listings an ideogram of the data. Fig. 1 is an example. Sometimes one or two data points lie apart from the main body; other times the data split into two or more groups. We extract no numbers from these ideograms; they are simply visual aids, which the reader may use as he or she sees fit.

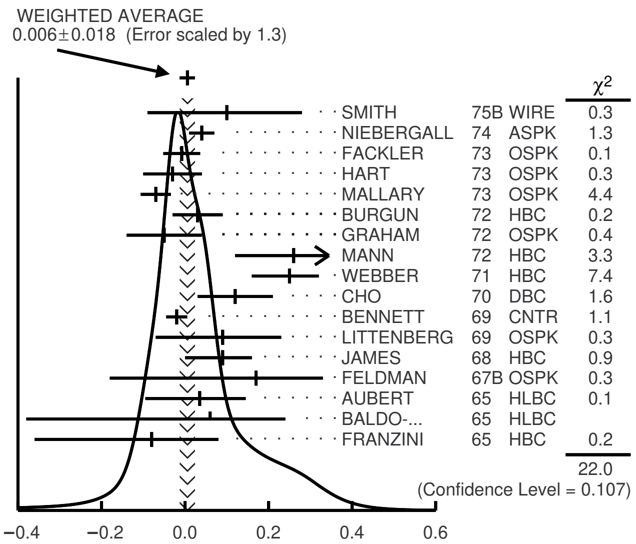

Figure 1: A typical ideogram. The "data point" at the top shows the position of the weighted average, while the width of the error bar (and the shaded pattern beneath it) shows the error in the average after scaling by the factor S. The column on the right gives the χ^2 contribution of each of the experiments. Note that the next-to-last experiment, denoted by the incomplete error flag ($\perp$), is not used in the calculation of S (see the text).

Each measurement in an ideogram is represented by a Gaussian with a central value x_i, error δx_i, and area proportional to $1/\delta x_i$. The choice of $1/\delta x_i$ for the area is somewhat arbitrary. With this choice, the center of gravity of the ideogram corresponds to an average that uses weights $1/\delta x_i$ rather than the $(1/\delta x_i)^2$ actually used in the averages. This may be appropriate when some of the experiments have seriously underestimated systematic errors. However, since for this choice of area the height of the Gaussian for each measurement is proportional to $(1/\delta\, x_i)^2$, the peak position of the ideogram will often favor the high-precision measurements at least as much as does the least-squares average. See our 1986 edition [2] for a detailed discussion of the use of ideograms.

4.2.3. *Constrained fits:* Except for trivial cases, all branching ratios and rate measurements are analyzed by making a simultaneous least-squares fit to all the data and extracting the partial decay fractions P_i, the partial widths Γ_i, the full width Γ (or mean life), and the associated error matrix.

Assume, for example, that a state has m partial decay fractions P_i, where $\sum P_i = 1$. These have been measured in N_r different ratios R_r, where, e.g., $R_1 = P_1/P_2$, $R_2 = P_1/P_3$, *etc.* [We can handle any ratio R of the form $\sum \alpha_i\, P_i / \sum \beta_i\, P_i$, where α_i and β_i are constants, usually 1 or 0. The forms $R = P_iP_j$ and $R = (P_iP_j)^{1/2}$ are also allowed.] Further assume that *each* ratio R has been measured by N_k experiments (we designate each experiment with a subscript k, e.g., R_{1k}). We then find the best values of the fractions P_i by minimizing the χ^2 as a function of the $m-1$ independent parameters:

$$\chi^2 = \sum_{r=1}^{N_r} \sum_{k=1}^{N_k} \left(\frac{R_{rk} - R_r}{\delta R_{rk}} \right)^2 , \tag{3}$$

where the R_{rk} are the measured values and R_r are the fitted values of the branching ratios.

In addition to the fitted values $\overline{P}_i$, we calculate an error matrix $\langle \delta\overline{P}_i\ \delta\overline{P}_j \rangle$. We tabulate the diagonal elements of $\delta\,\overline{P}_i = \langle \delta\,\overline{P}_i\ \delta\,\overline{P}_i \rangle^{1/2}$ (except that some errors are scaled as discussed below). In the Full Listings, we give the complete correlation matrix; we also calculate the fitted value of each ratio, for comparison with the input data, and list it above the relevant input, along with a simple unconstrained average of the same input.

Three comments on the example above:

(1) There was no connection assumed between measurements of the full width and the branching ratios. But often we also have information on partial widths Γ_i as well as the total width Γ. In this case we must introduce Γ as a parameter in the fit, along with the P_i, and we give correlation matrices for the widths in the Full Listings.

(2) We do *not* allow for correlations between input data. We *do* try to pick those ratios and widths that are as independent and as close to the original data as possible. When one experiment measures all the branching fractions and constrains their sum to be one, we leave one of them (usually the least well-determined one) out of the fit to make the set of input data more nearly independent.

(3) We calculate scale factors for both the R_r and P_i when the measurements for any R give a larger-than-expected contribution to the χ^2. According to Eq. (3), the double sum for χ^2 is first summed over experiments $k=1$ to N_k, leaving a single sum over ratios $\chi^2 = \sum \chi_r^2$. One is tempted to define a scale factor for the ratio r as $S_r^2 = \chi_r^2/\langle\chi_r^2\rangle$. However, since $\langle\chi_r^2\rangle$ is not a fixed quantity (it is somewhere between N_k and N_{k-1}), we do not know how to evaluate this expression. Instead we define

$$S_r^2 = \frac{1}{N_k} \sum_{k=1}^{N_k} \frac{\left(R_{rk} - \overline{R}_r\right)^2}{(\delta R_{rk})^2 - (\delta \overline{R}_r)^2} , \tag{4}$$

where $\delta\overline{R}_r$ is the fitted error for ratio r. With this definition the expected value of S_r^2 is one.

The fit is redone using errors for the branching ratios that are scaled by the larger of S_r and unity, from which new