

X.O. SAMTHN: THINNING DATA

Often a data file will contain more points than are needed to define the measurement (e.g., data appear to be very dense in energy regions where there is little structure in the data). SAMMY runs will therefore require far more computer time than is actually needed. For these reasons, it is often deemed reasonable to “thin” the data, at least for preliminary runs. A program SAMTHN has been written for this purpose; it has not, however, been studied extensively, so use with caution! Please note, also, that any results obtained with thinned data should ultimately be tested against the complete dense data set, to ensure that no important features were missed.

In running SAMTHN, the user provides two files, plus the name for the new thinned data file. The first user-provided file gives instructions on how the thinning should be done; the second is the dense data file. The first line of the thinning file contains three numbers (N, E_{min}, E_{max}) in I5, 2F10.1 format. The integer N tells how many data points are to be averaged together in the energy range from E_{min} to E_{max} . Second and subsequent lines contain two numbers (N, E_{max}); E_{min} is assumed to be equal to the previous E_{max} . Note that energies must be given from low to high, no matter what the ordering of data points in the original dense data file.

The fourth line of input (after the three file names) contains one character, which describes the type of input data in the dense file: O = Original (three energy-points per line, as in the original MULTI-style data format), C = CSISRS data, T = Twenty format, and D = Differential elastic data (ASCII format only). For differential elastic data, there is one additional line of input, an integer equal to the number of angles.

For O, C, or T, the format for the output thinned data file is always (3F20.10); in using this data file for SAMMY runs, the INPut file must contain the line “USE TWENTY SIGNIFICAnt digits for experimental data”. For D, the output format is (F10.2,7F10.5); if this format is inappropriate, for this simple code the author assumes the user can make the necessary changes him/herself.

The algorithms used by SAMTHN to average the data are as follows:

$$\bar{E} = \left(\frac{1}{N} \sum_{i=1}^N \frac{1}{\sqrt{E_i}} \right)^{-2} \quad (\text{X O.1})$$

and

$$\bar{D} = \frac{1}{\bar{V}} \sum_{i=1}^N \frac{D_i}{V_i} \quad (\text{X O.2})$$

with

$$\bar{V} = \left(\sum_{i=1}^N V_i^{-1} \right)^{-1} \quad (\text{X O.3})$$

in which E represents the energy, D the data, and V the variance (i.e., the square of the uncertainty). Quantities with subscripts are from the original dense data set, and quantities with bars represent the averaged values.

Test case tr072 has two examples for thinning data.