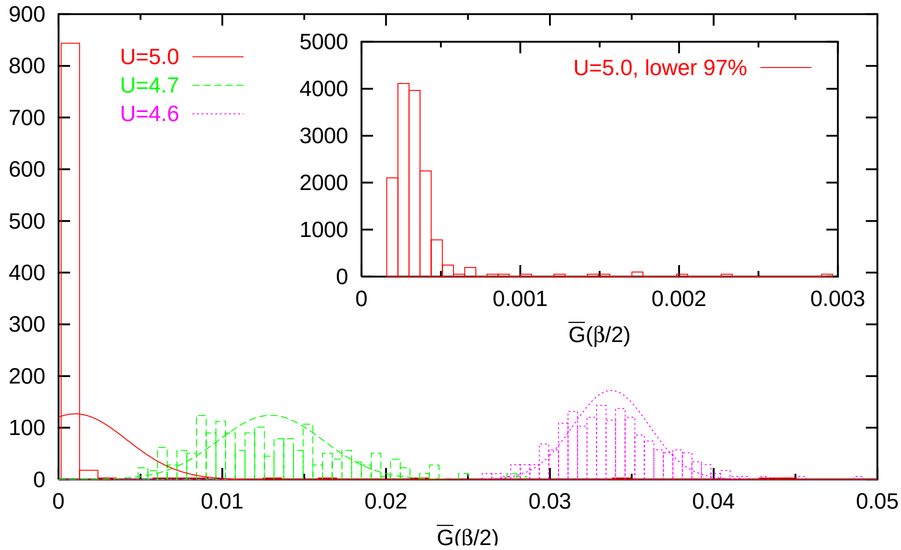

Figure 3.69: Histograms of $\bar{G}(\beta/2)$ for a Bethe DOS, $T = 1/20$, and $\Delta\tau = 0.1$ based on 300 measurements for each value of U (boxes). Smooth lines indicate Gaussian fits reproducing both the averages and the variances of each of the measured distributions. For $U = 5.0$, the variance is dominated by a small fraction of outliers; 97 percent of all measurements are below 0.003 (inset).

Fig. 3.69. Evidently, the distribution indicated by the histogram (boxes with solid lines) strongly differs from a Gaussian distribution corresponding to the same error and variance. In particular, the true distribution is strongly asymmetric while the Gaussian fit extends into the forbidden region $G(\beta/2) < 0$. In fact, the variance is primarily associated with a small fraction of outliers [i.e., measurements corresponding to very large values of $G(\beta/2)$]; eliminating the top 3% of outliers results in a much narrower (but still asymmetric) distribution as seen in the inset of Fig. 3.69. In comparison, the histograms for the insulating phase at $U = 4.7$ (dashed lines) and for the strongly correlated metallic phase at $U = 4.6$ (dotted lines) are much better approximated by Gaussians. As a consequence, the application of the MEM formalism deep inside the insulating phase implies the use of an inappropriate likelihood function[59] unless the number and accuracy of measurements is increased (by some orders of magnitude) so that Gaussian statistics is restored for the averages.[60] MEM algorithms have been formulated, however, for non-Gaussian error statistics in other contexts (Bryan and Skilling, 1980); we are therefore confident that a practical solution could be found leading to even more reliable spectra for insulators.[61]

[59] Our attempt of solving this problem by transforming G (to a new unbounded variable) lead to severe numerical instabilities so that the corresponding method could not be used in practical applications.

[60] At this point, the statistical error may be lower than systematic errors so that potentially the full error is never well approximated by a Gaussian distribution.

[61] Note that the presence of far outliers as seen for $U = 5.0$ in Fig. 3.69 poses an even more severe problem than skew and kurtosis since the characterization of the corresponding probability