ESTIMATION OF THE ERRORS INVOLVED IN THE INTENSITY MEASUREMENT OF LOW S/N RATIO EMISSION LINES

Claudia Rola and Didier Pelat

DAEC, Observatoire de Meudon

We present a detailed analysis on the extent of the errors in the intensity measurements of low signal-to-noise narrow emission lines. Our first goal is to determine a model for the probability distribution function (p.d.f.) associated with the measured intensities of a line characterized by its signal-to-noise ratio. Our final purpose is to provide an error domain: - for the measured signal-to-noise ratio of an observed line; - for the ratio of two lines in terms of their signal-to-noise ratios, and eventually (with the knowledge of the noise energy) to get errors on the corresponding intensities and intensity ratios.

To reproduce a real emission line intensity measurement process, we have designed a program based on a Monte-Carlo simulation procedure. For the purposes of this simulation, the knowledge of the true line parameters (the Intensity - imposed by the true signal-to-noise $(S/N)_{true}$, the position, and width) and the noise characteristics were necessary. Each spectrum was modeled by the sum of an emission line plus a continuum of a known constant level, this sum being considered as the mean of a certain stochastic process. The line was a Gaussian profile. To this model was added a non-correlated noise ("white noise") following the Student distribution.

Our simulation was divided in two main parts: the first, called detection, tells if a line is suspected to be present in a specific segment of the spectra; the second tries to measure the detected line, that is, to give estimates of the intensity, position and width of the suspected line. Once a line is suspected in one of the samples, the program proceeds to a fitting algorithm. The line model is also a Gaussian, the parameters of which are known only to lie within reasonable bounds. The adjustment is made through a standard X^2 minimizing routine. Once the minimum is attained a X^2 rejection test is performed on the residuals of the fit. If the test is satisfied, a Gaussian line is detected.

We made at least 1000 simulations for each $(S/N)_{true}$. For each signal-to-noise $(S/N)_{true}$, we calculated the intensity of each line in the detected lines sample. We call it "observed" intensity and denote it by S_{obs}. These intensities are then normalized to the known true intensity, S_{true}, and we call this normalised intensity, i, i.e., $i = S_{obs}/S_{true}$. Fits made for each $(S/N)_{true}$ led us to consider the log-normal distribution as a possible model for S_{obs}/S_{true}, which proved to give very reasonable results. We computed, for each normalised sample, a few statistics, the "true" signal-to-noise ratio, the mean of the sample (normalised to the true intensity), its variance and the square-root of the mean quadratic error calculated relative to one (the true normalised intensity value). Our results show clearly that for $(S/N)_{true}$ up to three, there is a strong bias ($> 50\%$) towards overestimation of the values of the "observed" intensities.

In practice, the observer's interest is to determine from the knowledge of $(S/N)_{obs}$ a plausible range where $(S/N)_{true}$ may be found. In other words, one would like to know what are the limits which contain S_{true} with probability of γ. These error bars can be easily calculated since we know for each value of $(S/N)_{true}$ the distribution followed by $(S/N)_{obs}$. By a change of variable it is easy to derive the distribution followed by $(S/N)_{obs}$, which is also a log-normal. This parameter has the advantage of being a non-dimensional universal one. By integrating the $(S/N)_{obs}$ p.d.f. corresponding to each $(S/N)_{true}$ up to a confidence level γ we obtain the chart presented in Fig. 1, where the continuous thick line delimits the confidence intervals for $\gamma = 0.683$ (or one σ), and the thin line, the one for $\gamma = 0.954$ (or two σ). This chart was constructed horizontally, but must be used vertically with the following reasoning: an observer measures a line of $(S/N)_{obs}$ and obtains vertically the range on the $(S/N)_{true}$ values determined with confidence level γ. This chart shows clearly that, for example, at the one σ confidence level (or $\gamma = 0.683$), a line measured with $(S/N)_{obs} = 4$ is compatible with $0 \leq (S/N)_{obs} \leq 5$. One cannot then rule out the possibility that a noise fluctuation was mistaken for a line.
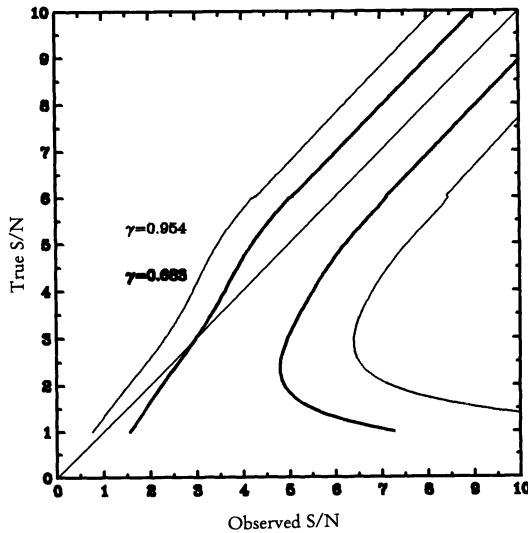

Fig. 1. Chart for the interval estimation of the $(S/N)_{true}$, given an observed value $(S/N)_{obs}$.

Furthermore, we calculated that a ratio of measured intensities or signal-to-noise ratios also follows a log-normal distribution and as such is subject to strong biases. Additionally, we determined the error domain on the ratio of a signal-to-noise ratios, which can be converted in an intensities ratio with the knowledge of the rrespondent noise energy ratio.