VLA SCIENTIFIC MEMORANDUM NO. 147

A Method of Optimal Data Smoothing Applied to Analysis of Spectral Line Profiles

Fred Schwab

March, 1983

1. Introduction. The topics considered in this memorandum are the optimal smoothing of noisy one-dimensional spectral line velocity profiles, the estimation of moments and other parameters characterizing these profiles, the estimation of the uncertainties of these estimates, and an investigation of their biases. Smoothing is an essential element of the parameter estimation procedure, for it yields an estimate of the mean profile shape. Usually, some of the parameters of interest depend, either directly or indirectly, on the estimation of the location of a particular profile ordinate—perhaps on the location of the profile peak or of the half-maximum. For example, this dependence is direct in an an estimate of the 50% width (FWHM), and it is indirect in estimates of the profile moments, if they are calculated over a data window centered about the abscissa of the profile peak. Smoothing is desirable in that it causes the estimates of the locations of particular profile ordinates to become continuous functions of the measurement errors, and thereby allows the calculation of meaningful error estimates. It is undesirable in that it introduces bias, especially in the estimation of characteristic widths (see e.g., Lewis [L]).

In § 2 is outlined a method of variable data smoothing based on a real parameter α. Appropriate choice of α depends on the magnitude of the measurement errors. If an a priori estimate of the measurement error is available, then one can make a reasonable guess at an appropriate choice for α. On the other hand, if it can be assumed that the errors are independent and identically distributed (in particular, that there has been no prior smoothing of the data), then a choice can be made which is based on the data alone: For a particular choice of α, and successively ignoring one data point at a time, the data are repeatedly smoothed; and the optimal choice of the smoothing parameter is that which yields the least mean square data prediction error—i.e., it is that α which yields, on average, the best prediction of each missing datum. This is the method of cross validation, described in § 3.

In § 4 are defined some parameters—moments, etc.—which would be appropriate for the characterization of a unimodal, or singly peaked, emission line profile. This is so that an error estimation procedure can be outlined in § 5 and some test results presented in § 6. An attractive point to emphasize is that these nonparametric error estimates can be retrieved as by-products of the optimal smoothing method of § 3. For the analysis of multimodal profiles or absorption line profiles, similar ideas apply—but a useful characterization of a complicated profile necessarily requires the estimation of a larger number of parameters than the number needed to characterize a simple profile. As this number increases, it becomes more useful, perhaps, simply to look at a plot of the graph of the mean profile estimate that has been derived through some technique of optimal data smoothing, and to "let the data speak for themselves." In § 7, brief attention is given to the application of smoothing techniques to the problem of testing for multimodality.

To summarize, the method of analysis described below consists of three parts: smoothing of the data, by the calculation of a spline curve which passes near each of the data points, with particular attention to choosing the proper degree of smoothing; computation of parameters (moments, etc.) which characterize this curve; and estimation of the errors in these derived parameters.

2. Smoothing. The reasons for smoothing the data are threefold: 1) to produce an aesthetically pleasing estimate of the mean profile shape lacking (obvious) noise bumps; 2) to reduce the variance in the parameter estimates (at the expense of increased bias); and 3) to make the derived parameters continuous functions of the data, so that useful error estimates can be obtained.

Let the observed profile—i.e., the measured intensity as a function of velocity, at a particular position in the sky—be given by $y_i = f(v_i) + \epsilon_i$ for $i = 1, \ldots, n$, where $f(v)$ is the true intensity as a function of velocity and where the ϵ_i are independent zero-mean random variables of finite variance σ_i^2, representing the measurement error at velocity v_i.* Given a smoothness parameter $\alpha > 0$, and given estimates $\tilde{\sigma}_i$ of the σ_i, we shall choose to minimize the quantity

$$S = \sum_{i=1}^{n} \left(\frac{y_i - s(v_i)}{\tilde{\sigma}_i} \right)^2$$

over all twice continuously differentiable functions $s(v)$—subject to the smoothness constraint

$$\alpha = \int_{v_1}^{v_n} (s''(v))^2 \, dv \, .$$

The solution $s(\alpha; v)$ is a cubic spline function, comprised of cubic pieces pieced together at $v_2, \ldots, v_{n-1}$, with matching first and second derivatives at these points, and satisfying $s''(v_1) = s''(v_n) = 0$. For $v \in$

*In § 3 we shall assume, in addition, that the ϵ_i are identically distributed.

$[v_i, v_{i+1})$, we shall represent s by $s(v) = c_{i0} + c_{i1}(v - v_i) + c_{i2}(v - v_i)^2 + c_{i3}(v - v_i)^3$, where $c_{ij} \equiv s^{(j)}(v_i)/j!$. The c_{ij} are called the ***spline coefficients.*** The larger that α is chosen, the nearer $s(\alpha; v_i)$ stays to y_i; conversely, the smaller α, the smoother s becomes. In other words, α measures the roughness, or "flamboyancy", of s;* and S measures the degree of "infidelity" of s, in its relation to the data.

A reasonable choice for α might be one such that the corresponding value of S lies in the confidence interval $n - \sqrt{2n} \leq S \leq n + \sqrt{2n}$. An Algol subroutine for numerical computation of α and $s(\alpha; v)$, given a desired value of S, appears in Reinsch [**R**]; a Fortran version of Reinsch's algorithm is given by de Boor [**deB**, Chapt. XIV]; and another Fortran version is included in the IMSL Library—this is the subroutine (ICSSCU) which was used for the tests reported in § 6. de Boor's Fortran version would have to be used in AIPS, because of the proprietary nature of the IMSL library.

3. Optimal smoothing. In § 6 it is shown that the estimation of moments and related parameters of f, derived by calculating the corresponding parameters characterizing the smoothed velocity profile, is very sensitive to the *amount* of smoothing that is chosen. In particular, if one employs an automated procedure to pick out an interesting looking "hump" in the smoothed profile, positions a data window in the neighborhood of the hump, and then calculates moments over that window, then the velocity dispersion estimate is biased downward if too little smoothing is used and may be biased severely upward if too much smoothing is used. One might not expect this downward bias in the case of too little smoothing; it evidently arises from errors in calculating the appropriate placement, and width, of the data window.

For a particular choice of the smoothness parameter α, let $s_{(i)}(\alpha; v)$ denote the smoothing spline that is obtained if the ith data point is excluded from consideration. Then $y_i - s_{(i)}(\alpha; v_i)$, the ***data prediction error***, tends to be small (in absolute value) if the other data points, and the chosen degree of smoothing, do a good job of predicting y_i; and it tends to be large otherwise. Selecting that value of α which minimizes the mean square data prediction error

$$Q(\alpha) = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - s_{(i)}(\alpha; v_i)\right)^2$$

is called the ***method of cross validation.*** Assuming now that the measurement errors ϵ_i are identically distributed, we have a method of optimal data smoothing in which the degree of smoothing is determined from the data alone.

*This terminology is due to Silverman [S].

Spline smoothing by cross validation is described in detail by Craven and Wahba [C–W]. This method is implemented in the IMSL Library Fortran subroutine ICSSCV, which I have used for the test results reported in § 6. Obviously the idea of cross validation can be used in many other applications; it is described in a broader context by Efron [E1], [E2], and by Grace Wahba [W]. The latter reference describes the application to ill-posed problems given noisy data (as in maximum entropy deconvolution, choosing the trade-off between the entropy term and the error term). An application in periodogram analysis, a problem very similar to our own, is described in [W–W].

4. Moments, and related parameters. The first three (normalized) moments of the smoothed velocity profile are given by

$$\theta_1 = \int_{v_1}^{v_n} s(v)\,dv\,,$$
$$\theta_2 = \frac{1}{\theta_1}\int_{v_1}^{v_n} v s(v)\,dv\,,$$
$$\text{and}\quad \theta_3 = \frac{1}{\theta_1}\int_{v_1}^{v_n} (v-\theta_2)^2 s(v)\,dv\,.$$

Note that each of the integrals simply is equal to a weighted sum of the spline coefficients. For example, $\theta_1 = \sum_{i=1}^{n-1}\sum_{j=0}^{3} c_{ij}\Delta_i^{j+1}/(j+1)$, where $\Delta_i \equiv v_{i+1}-v_i$. Three additional parameters, of particular interest in the case of a unimodal profile, are

$$\theta_4 = \text{the abscissa, } v_0\text{, of the peak of } s\,, \qquad \theta_5 = s(v_0)\,,$$
$$\text{and}\quad \theta_6 = \frac{\theta_1}{\theta_5} = \Delta v_{\text{eq}},\ \text{the equivalent width.}$$

θ_6 is the width of that rectangular profile whose integrated intensity is the same as that of s, and whose peak intensity is the same as the peak intensity of the smoothed velocity profile. Additional parameters, likely to be more useful than θ_1–θ_3 for characterization of a noisy profile, are the moments of the smoothed profile restricted to a narrow window W centered about the abscissa, $v_0 = \theta_4$, of the peak of s (say, the window $W = [v_0 - \Delta v_{\text{eq}}, v_0 + \Delta v_{\text{eq}}]$):

$$\theta_7 = \int_W s(v)\,dv\,,$$
$$\theta_8 = \frac{1}{\theta_7}\int_W v s(v)\,dv\,,$$
$$\text{and}\quad \theta_9 = \sqrt{\frac{1}{\theta_7}\int_W (v-\theta_8)^2 s(v)\,dv}\,.$$

(One wants to restrict the computation of moments to a narrow data window in order to reduce the variance in the parameter estimates; but the selection of the window can be the main source of of bias). θ_9 is the velocity dispersion, over the window W, about the mean velocity θ_8, calculated over the same window. (I've taken the square root here, in hopes that the computed second moment will be positive). Often other parameters than these—say, the 50% width (the FWHM), the 20% width, etc.—also are of interest.

5. Estimation of errors. Useful non-parametric error estimates can be derived by a method which resembles cross validation, and, as a by-product of cross validation. Of course, errors in the moments of the unsmoothed velocity profile, computed over a *fixed* data window, follow from the standard normal distribution theory (see e.g., Kendall and Stuart [K–S, pp. 228–231]). But, more properly, the error estimates *ought* to reflect the uncertainty in the choice of an appropriate data window—so the matter becomes more complicated in the case of the parameters θ_7–θ_9 defined above. And, too, the matter is more complicated in the case of any other parameters (say, the 50% width) which depend on an estimate of the location of a particular ordinate or on an estimate of the height of the profile peak.

The method described below, known as the *jackknife*, first was described by Quenouille in 1949, and it was popularized in the 1950's by J. W. Tukey (of Cooley–Tukey FFT fame). The name *jackknife* was coined because the technique is a tool, like the pocket knife, which can be handy in diverse situations. For more thorough discussions of the jackknife and related methods than that given here, see Efron [E1], [E2].

Denote the (column) vector of parameters by Θ, $\Theta = (\theta_1, \ldots, \theta_9)^T$, and the vector of parameter estimates derived from cross validation by $\hat{\Theta}$. Also, let $\hat{\Theta}_{(i)}$ denote the vector of parameter estimates derived by ignoring the ith data point y_i. (To do things completely correctly, each $\hat{\Theta}_{(i)}$ ought to be obtained by a separate cross validation estimation, but we shall settle for the $\hat{\Theta}_{(i)}$'s that correspond to the single optimal α appropriate to the full set of observations).* Now, if we let $\hat{\Theta}_{J,i} = n\hat{\Theta} - (n-1)\hat{\Theta}_{(i)}$ and $\hat{\Theta}_J = \frac{1}{n}\sum_{i=1}^{n} \hat{\Theta}_{J,i}$, then the jackknife estimate of the covariance C of $\hat{\Theta}$ is given by

*The IMSL Library subroutine for computation of smoothing splines by cross validation, based on an algorithm given in [U1], requires $O(n)$ arithmetic operations when the abscissae are evenly spaced, versus $O(n^2)$ operations when they are not. (Normally the velocity abscissae are evenly spaced, but, of course, when a point is ignored, the remaining abscissae no longer are equispaced—unless one of the endpoints has been dropped). (Recently though, Utréras [U2] has published an $O(n)$ algorithm for the case of non-equispaced data).

$$\hat{C} = \frac{1}{n(n-1)} \sum_{i=1}^{n} (\hat{\Theta}_{J,i} - \hat{\Theta}_J)(\hat{\Theta}_{J,i} - \hat{\Theta}_J)^T .$$

$\hat{\Theta}_J$ is called the ***jackknifed estimate*** of Θ; it includes a bias correction term which, in our application, often is large compared to the standard error—so we shall ignore it. The estimate of the standard error of $\hat{\Theta}$ is the vector of square roots of the diagonal elements of $\hat{C}$. The matrix of estimated normalized correlation coefficients is obtained by dividing each element of $\hat{C}$ by the corresponding pair of standard error estimates.

Matloff [M] gives a versatile Fortran subroutine for jackknifing—given a user-supplied subroutine that is called repeatedly to compute the base estimator (Θ above). Note that the jackknife is relatively expensive to apply, requiring, in the analysis of a set of n data points, about n times as much work as usual. For large data sets, when it is impractical to apply the standard jackknife procedure, the data can be divided into groups, and groups deleted one at a time.

6. Test results and examples. Several examples of the application of cross validation and the error estimation procedure of § 5 are given below. More comprehensive tests, including comparisons with the traditional methods of profile analysis, may await practical experience with the method.

Examples of cross validation, with varying S/N. In Figures 1a–1f, a few sample results obtained by the cross validation technique are shown. Here, $n = 101$, $f(v)$ is a Gaussian, the ϵ_i are identical pseudo-random normal deviates of zero-mean, with standard deviation σ equal to the maximum of $f(v)$ divided by the indicated value of S/N, and the $\tilde{\sigma}_i \equiv \sigma$ in the definition of S. $\hat{\Theta}$ is shown in the Figures (except that θ_{10} is equal to the θ_0 of § 4, and θ_0 here is the velocity dispersion about v_0). Figure 1f, in which the signal to noise ratio (S/N) is equal to zero, is particularly interesting in that it illustrates the tendency for the cross validation smoothing spline to approach a straight line coincident with the v-axis whenever the data are well-behaved random noise.

Effect of varying the smoothing parameter. The effect of varying the smoothing parameter is illustrated in Figure 2. Here f is as in the previous example, $S/N = 4$, and α is chosen such that S_{opt} is equal to $n, n \pm \sqrt{2n}$. The cross validation curve also is shown. The cross validation curve looks best, in that the width of the largest bump is closest to the width of the true profile (similarly for the height). On the other hand, it has extraneous bumps, one negative and the other positive, at $v \approx .19, .67$. Since these bumps do not appear when α is chosen such that S_{opt} is close to n, they mightn't be considered real if this were a "real-life" situation (see § 7).

Bias of the width estimates. In Figures 3a–3c are shown histograms of width estimates (θ_0 of § 4), each based on 1000 trials, for the cases

$S/N = 3.333, 5, 10$, respectively. Here again, $f(v)$ is a Gaussian density with standard deviation 0.05. The distribution of widths of the smoothing splines derived by cross validation is shown in the top portion of each Figure. The middle histogram shows the distribution of width estimates derived naively from the raw data—that is, the integrals were approximated by rectangular sums (i.e., sums of the form $\theta_1 \simeq \sum y_i \Delta v_i$), and the height and the location of the profile peak were estimated by finding $\max y_i$; the equivalent width was calculated in this naive manner, as well. The bottom histogram represents widths that were estimated by using the cross validation estimates of θ_4–θ_6 to calculate the data window W, but then using rectangular sums and the raw data to calculate θ_7, θ_8, and the second moment over W.

In Figure 3a, corresponding to the case $S/N = 3.333$, we see that the cross validation width estimates are biased upward by 14% and the naive estimates downward by 17%, but that the third estimates are biased only by +0.7%. On the other hand, the variance of the width estimates derived by the third method is significantly greater than the variance of the width estimates derived by either of the other two methods. The bias is significantly less for the higher signal to noise cases shown in Figures 3b and 3c: for $S/N = 5$ we have $(+8.8\%, -9.5\%, +0.6\%)$; and for $S/N = 10$ we have $(+4.3\%, -3.7\%, +0.5\%)$.

The reason for the upward bias in the cross validation case is clear—it is simply due to the tendency of any smoothing technique to fatten a curve. However, as is suggested by inspection of Figure 2, the bias in the width of the smoothing spline would be even greater if the curve were derived not by cross validation, but rather by setting $S = n$, on the basis of an a priori estimate of the noise, and using the smoothing method of § 2. The downward bias exhibited in the case of the rectangular sums is due to the severe underestimation of the equivalent width, which in turn is due to the overestimation of the height of the profile peak; and probably contributing too is a greater uncertainty in the location of the abscissa of the profile peak. The third method appears somewhat more attractive than the other two, except for the greater variance in the estimated width.

Overall, cross validation generally seems not to broaden a profile much more than is necessary in order to get a pleasing appearing representation of the data. In addition to the bias in the width estimate derived by cross validation, there is a downward bias in the estimate of the height of the profile peak, and, hence, an upward bias in the estimate of the equivalent width. It is this effect which probably causes the larger variance of the width estimates derived by the third method; that is, the better estimate of the center position, and the tendency to broaden the data window, effectively remove the bias, but the noise added into the rectangular sums at the edges of the window increases the variance.

Example of the jackknife estimation of error. An example of parameter estimates and error estimates, derived by the method of § 5, is shown in Figure 4. Here $S/N = 5$. $\hat{\Theta}$ and $\hat{\Theta}_J$ both are shown in the Figure, along with the matrix $\hat{C}$ of estimated normalized correlation coefficients, and the estimated standard deviations of $\hat{\Theta}$. The matrix of correlation coefficients is arranged in the straightforward manner; for example, the observed correlation of the width estimate θ_{10} with the estimated height of the profile peak is -0.63. I have not studied systematically the distribution of the jackknifed estimates $\hat{\Theta}_J$, but their behavior has appeared somewhat erratic in a number of trials similar to this one.

7. Testing for multimodality. Following an idea proposed by Silverman (see [S], and references cited therein), suppose that one wishes, in the analysis of a measured profile, to test the hypothesis that the true profile has at least k modes, or "bumps", for some given $k \geq 1$. One may search for the critical value α_{crit}, which is defined as the smallest value of the smoothing parameter α for which k modes appear in $s(\alpha; v)$. The smaller α_{crit}, the greater the confidence level for the hypothesis that there are indeed at least k modes.

Generally, too, one is interested in specific profile features, or particular bumps. Analogous tests can be constructed for these situations—for example, one may search for the value α_{crit} at which a persistent bump first appears in the neighborhood of a given profile abscissa.

The problem of constructing significance tables appropriate to these situations merits investigation. A priori estimates of the $\tilde{\sigma}_i$ appearing in the definition of S in § 2 perhaps would be helpful here.

8. Discussion. Utréras' work is quite important in our application, for, by reducing the problem of selection of the optimal value of the smoothing parameter to an equivalent eigenvalue problem, he decreases the required computational effort roughly by a factor of n. Listings of his computer programs are available in a technical report [U3]. [U2] gives timing information for computation of smoothing splines by cross validation: on an IBM 360/67 computer the quoted run times for $n =$ 50, 100, 150, 200, 250, 300, 350, and 450 are .80, 1.14, 1.76, 2.47, 3.18, 3.79, 4.45, and 4.90 seconds, respectively. It is apparent that this method is practical for single-dish radio astronomical applications, and, in a limited sense, for aperture synthesis spectral line data reduction—but probably not for large-scale aperture synthesis reduction. Furthermore, the reduction in computational effort by a factor of n applies only to the computation of the cross validation smoothing spline, and not to the jackknife estimation of errors outlined in § 5; hence, my assertion in the Introduction that these error estimates may be retrieved as by-products of cross validatory data smoothing requires the obvious qualification that this only is so provided that the smoothing is done by the naive, computationally intensive algorithm. Some of the computations profitably could be performed in a high-speed array processor; but array processors are not particularly well-suited to the task.

An obvious objection to the cross validation smoothing technique, and to the error estimation scheme described here, is that, in actual practice, the measurement errors ϵ_i are not independent. Generally the data y_i are *averages* of spectra, each of which is obtained by computing the discrete Fourier transform of a discrete correlation function $r(\tau_j) = \sum_k x_1(t_k)x_2(t_k + \tau_j)$, where x_1 and x_2 are stationary time series. That the sampled values, $x_1(t_k)$, $x_2(t_k)$, usually may be considered independent is not of much direct utility, because of the possibly large volume of such data. But, if one were willing to combine the averaging operation with the operations of smoothing and parameter estimation, then in cross validation and jackknifing one could delete individual spectra rather than individual data points. If, for example, there were N spectra of n points each, given by y_{ij}, $i = 1, \ldots, N$, $j = 1, \ldots, n$, one would minimize $S = \sum_{i=1}^{N} \sum_{j=1}^{n} \left(y_{ij} - s(v_j)\right)^2$, subject to the smoothness constraint $\alpha = \int_{v_1}^{v_n} (s''(v))^2\, dv$; and select α by minimizing a mean square data prediction error given by $Q(\alpha) = \frac{1}{Nn} \sum_{i=1}^{N} \sum_{j=1}^{n} \left(y_{ij} - s_{(i)}(\alpha; v_j)\right)^2$, where now $s_{(i)}$ denotes the smoothing spline obtained if the ith *spectrum* is excluded from consideration. The jackknifed error estimates would be obtained analogously. If N were too large for this method to be practical, one could divide the spectra into groups, compute group averages, and then apply the cross validation technique and the jackknife to these averages. One thousand independent spectra, for example, might be divided into 25 groups of 40, and cross validation applied to the resultant 25 independent mean spectra. Such a scheme as this would be practical in single-dish radio astronomy, but probably not (at present) in aperture synthesis work (though, in fact, the technique could be applied with as few as two groupings of data).

References

[C–W] Craven, P. and G. Wahba, *Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross validation*, Numerische Mathematik **31** (1979), 377–403.

[deB] de Boor, C., "A Practical Guide to Splines", Springer-Verlag, New York, 1978.

[E1] Efron, B., "The Jackknife, the Bootstrap and Other Resampling Plans", Society for Industrial and Applied Mathematics, Philadelphia, 1982.

[E2] ———, *Computers and the theory of statistics: thinking the unthinkable*, SIAM Review **21** (1979), 460–480.

[K–S] Kendall, M. G. and A. Stuart, "The Advanced Theory of Statistics (Volume I)", Hafner, New York, 1969.

[L] Lewis, B. M., *The precision of velocity measurements from HI profiles*, preprint, March 1983.

[M] Matloff, N. S., *Algorithm AS 148: the jackknife*, Applied Statistics **29** (1980), 115–117.

[R] Reinsch, C. H., *Smoothing by spline functions*, Numerische Mathematik **10** (1967), 177–183.

[S] Silverman, B. W., *On a test for multimodality based on kernel density estimates*, MRC Technical Report No. 2181, Mathematics Research Center, University of Wisconsin—Madison, Madison, February 1981.

[U1] Utréras Díaz, F., *Cross-validation techniques for smoothing spline functions in one or two dimensions*; In Gasser, Th. and M. Rosenblatt (eds.), "Smoothing Techniques for Curve Estimation", Lecture Notes in Mathematics 757, Springer-Verlag, Berlin, 1979, pp. 196–232.

[U2] ——————, *Optimal smoothing of noisy data using spline functions*, SIAM Journal on Scientific and Statistical Computing **2** (1981), 349–362.

[U3] ——————, *A package for smoothing noisy data with splines*, Technical Report MA-80-B-209, Department of Mathematics, University of Chile, Santiago, 1980.

[W] Wahba, G., *Practical approximate solutions to linear operator equations when the data are noisy*, SIAM Journal on Numerical Analysis **14** (1977), 651–667.

[W–W] Wahba, G. and S. Wold, *Periodic splines for spectral density estimation: the use of cross-validation for determining the degree of smoothing*, Communications on Statistics **4** (1975), 125–141.

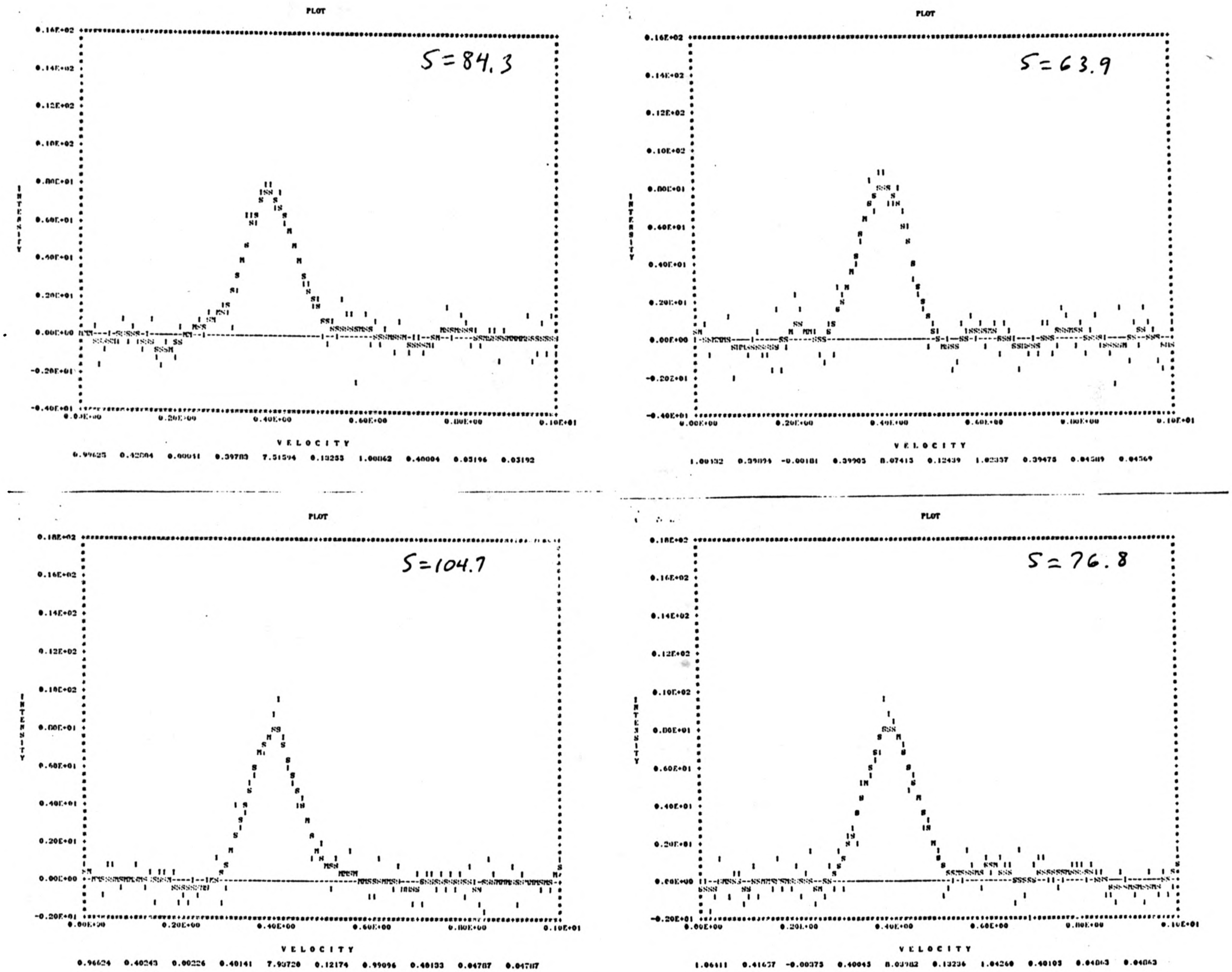


FIG. 1A. Cross validatory spline smoothing of a noisy Gaussian profile centered at $v = 0.4$, and of dispersion equal to 0.05. The pseudo-random normal noise has mean zero and a standard deviation of 0.1 times the true profile peak (i.e., $S/N = 10$).

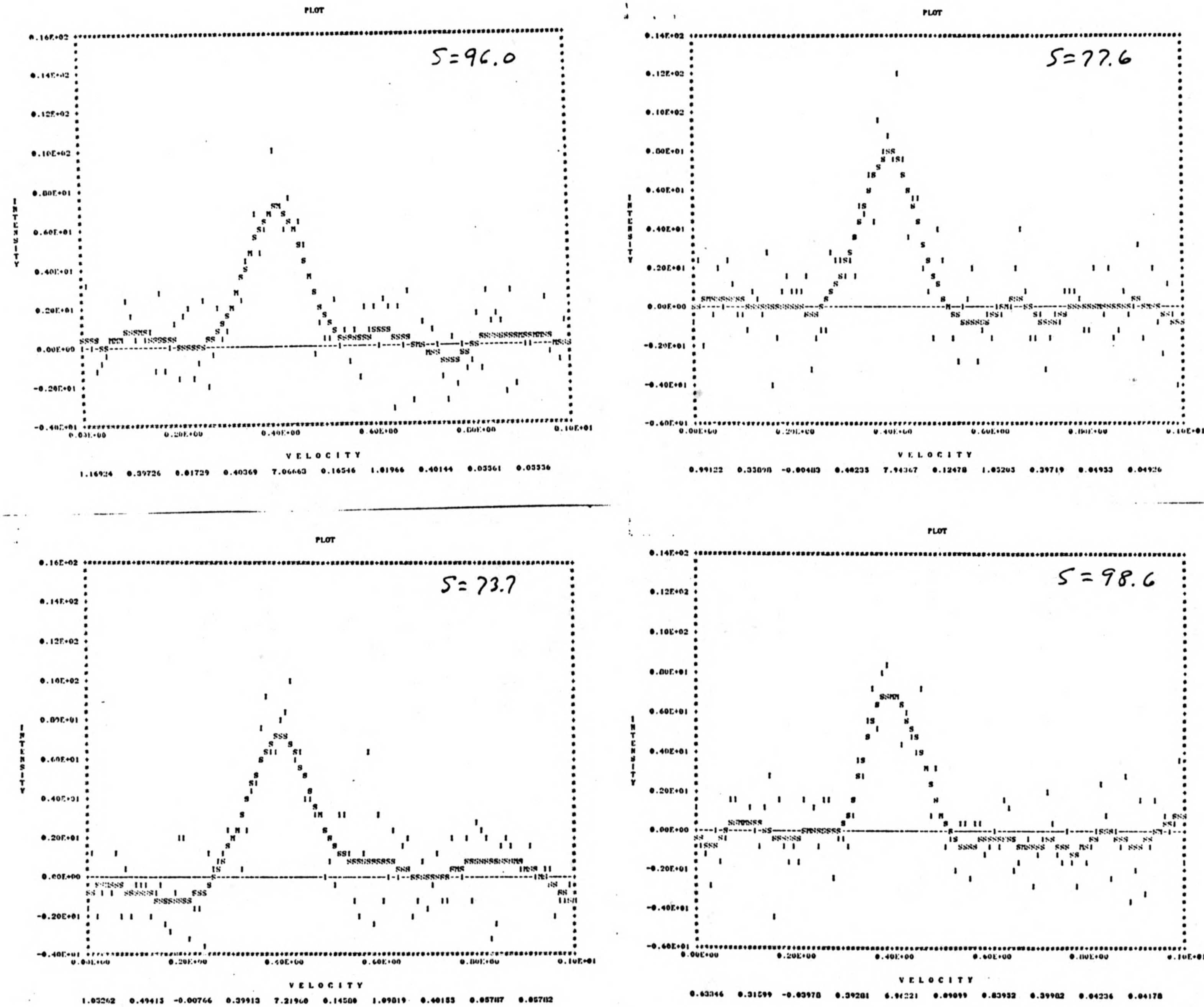


FIG. 1B. Same as FIG. 1A, except $S/N = 5$.

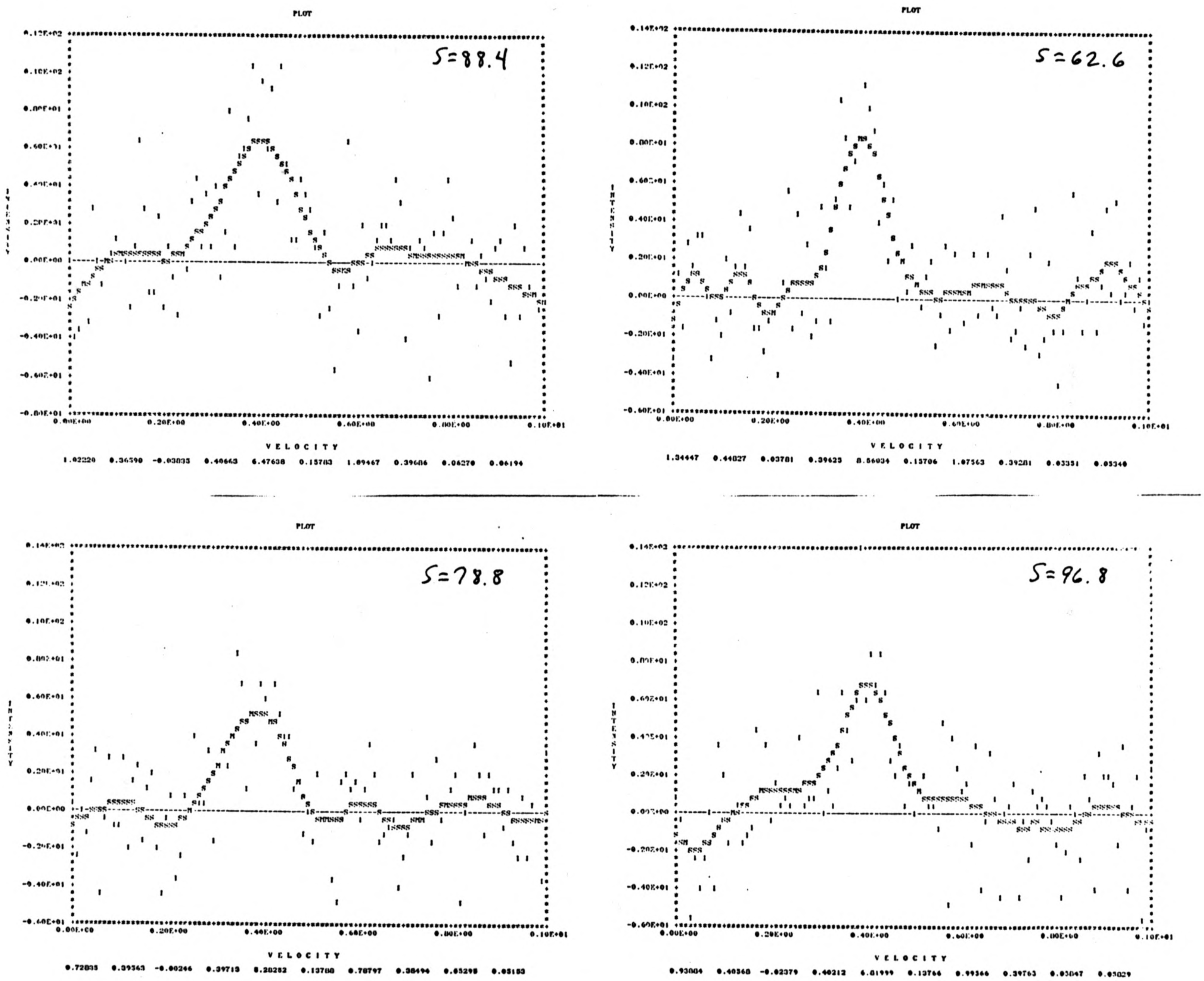


FIG. 1C. Same as FIG. 1A, except $S/N = 3.333$.

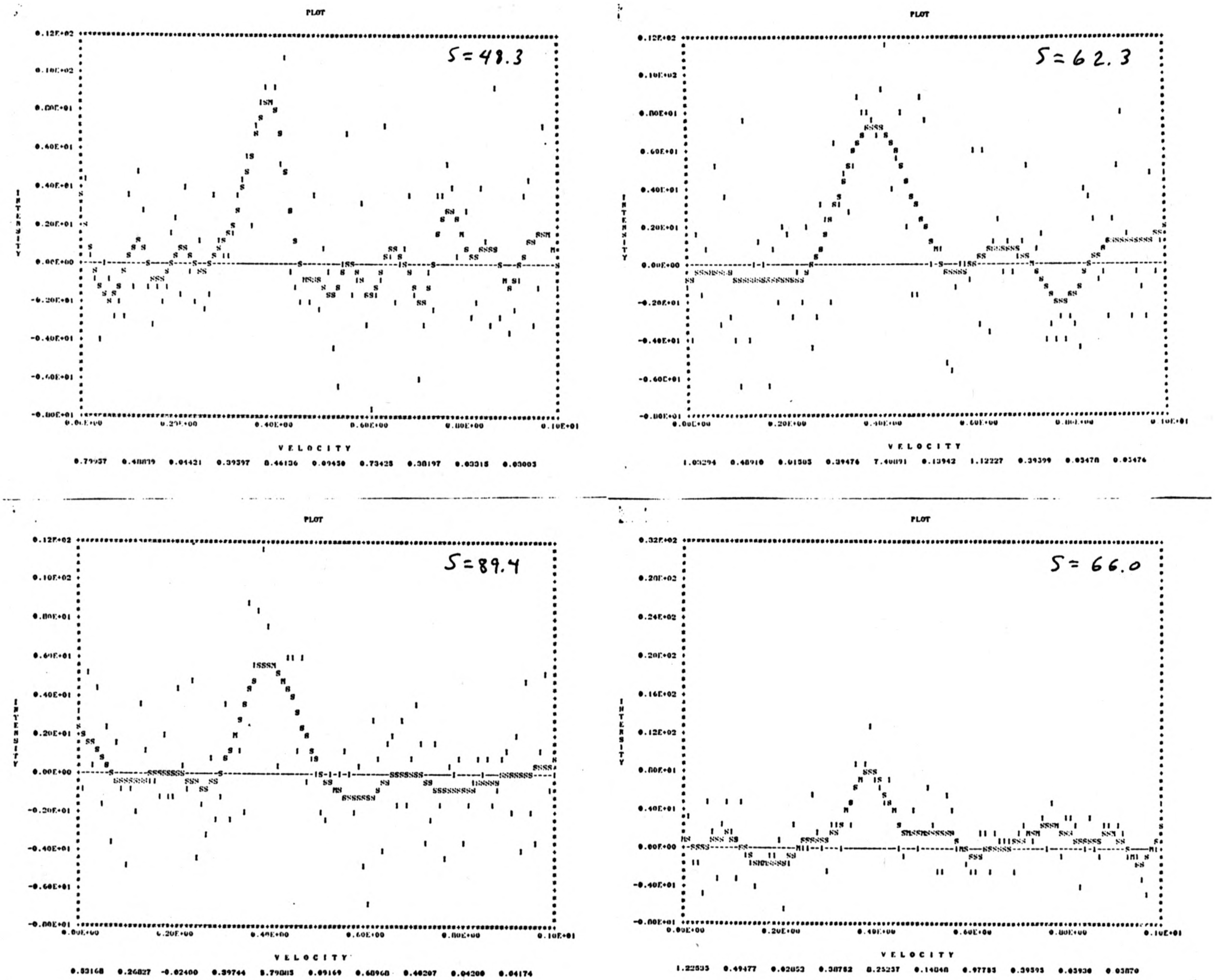


FIG. 1D. Same as FIG. 1A, except $S/N = 2.5$.

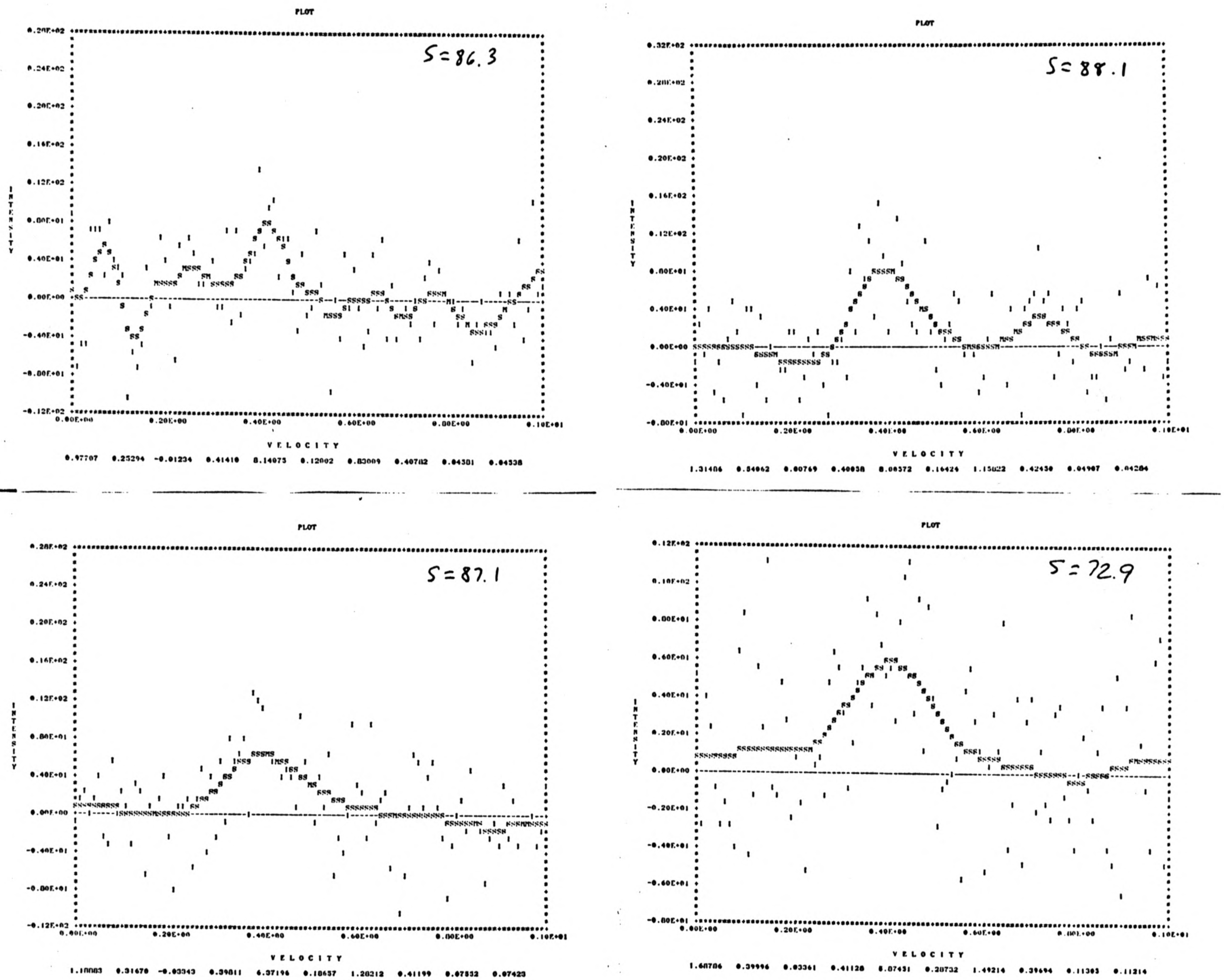


FIG. 1E. Same as FIG. 1A, except $S/N = 2$.

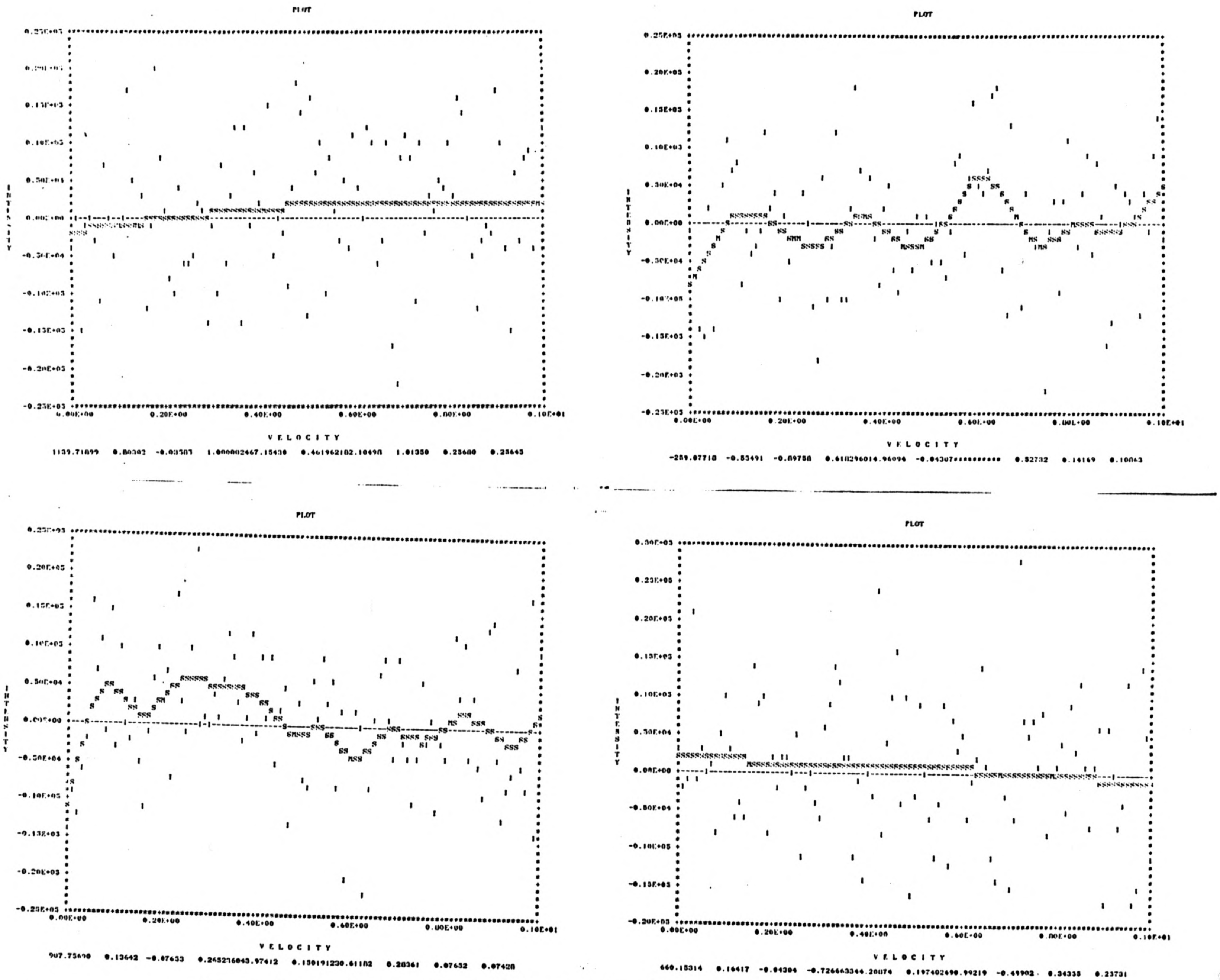


FIG. 1F. Same as FIG. 1A, except $S/N = 0$ — i.e., the data are pure noise. This behavior is typical; with $n = 101$, about half the time the smoothing curve obtained by cross validation is a straight line of nearly zero slope. This desirable tendency is accentuated when n is increased.

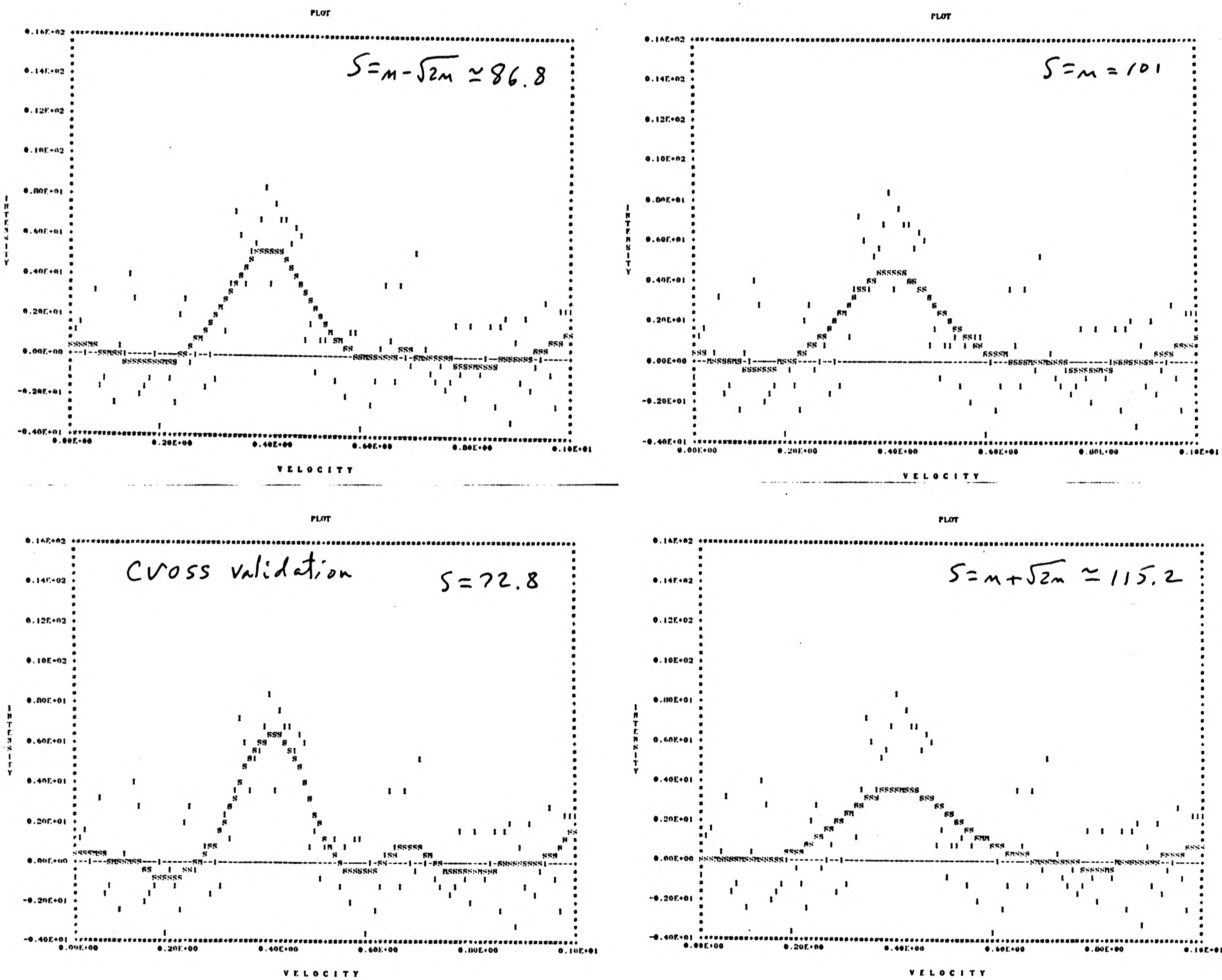


FIG. 2. The effect of varying the smoothing parameter. Here α is chosen such that $S_{\text{opt}} = n, n \pm \sqrt{2n}$. The cross validation smoothing spline also is shown.

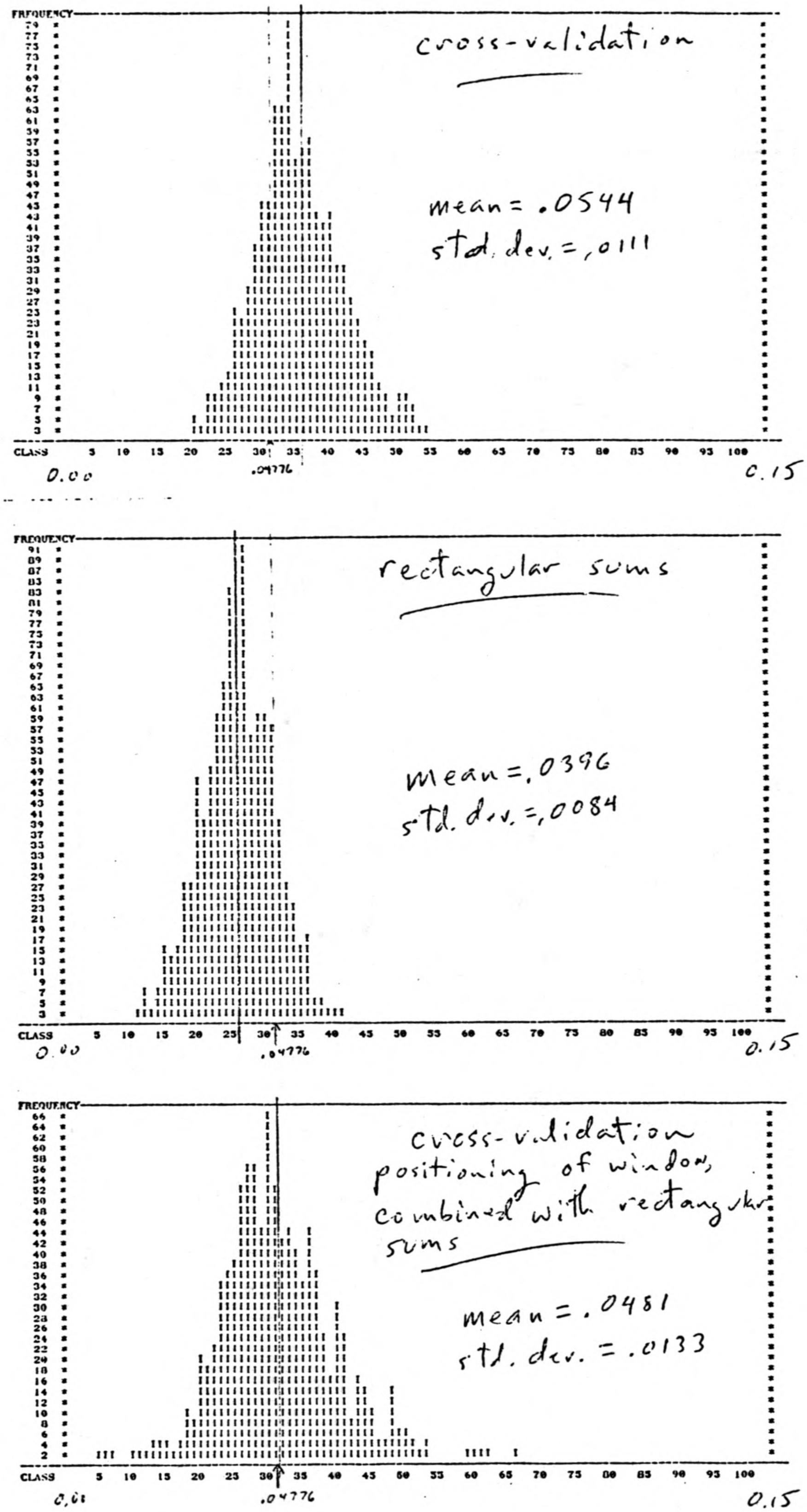


FIG. 3A. Histograms of width estimates, based on 1000 trials. The true width is 0.05 (0.04776 when calculated by integrating over a window of twice the equivalent width), and $S/N = 3.333$.

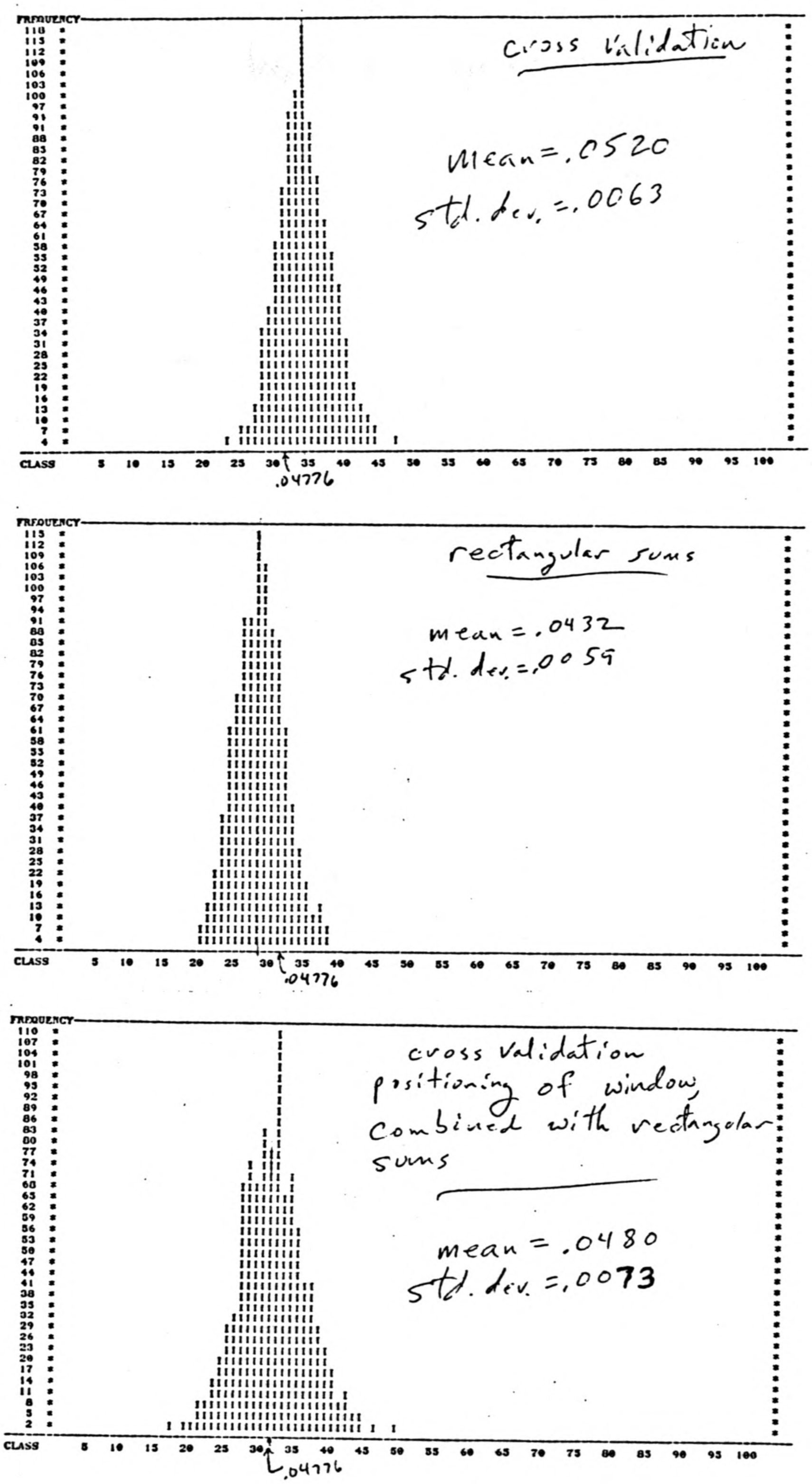


FIG. 3B. Same as FIG. 3A, except $S/N = 5$.

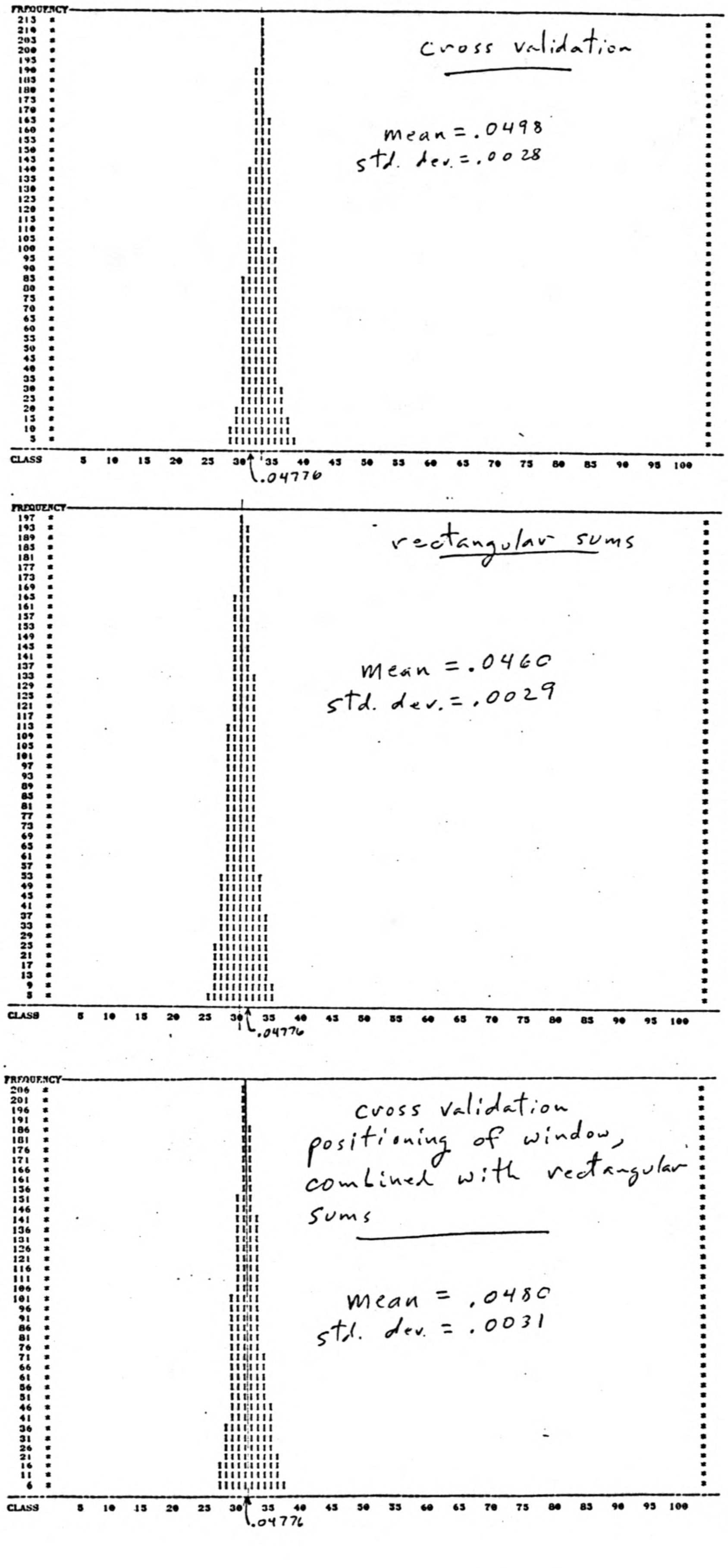
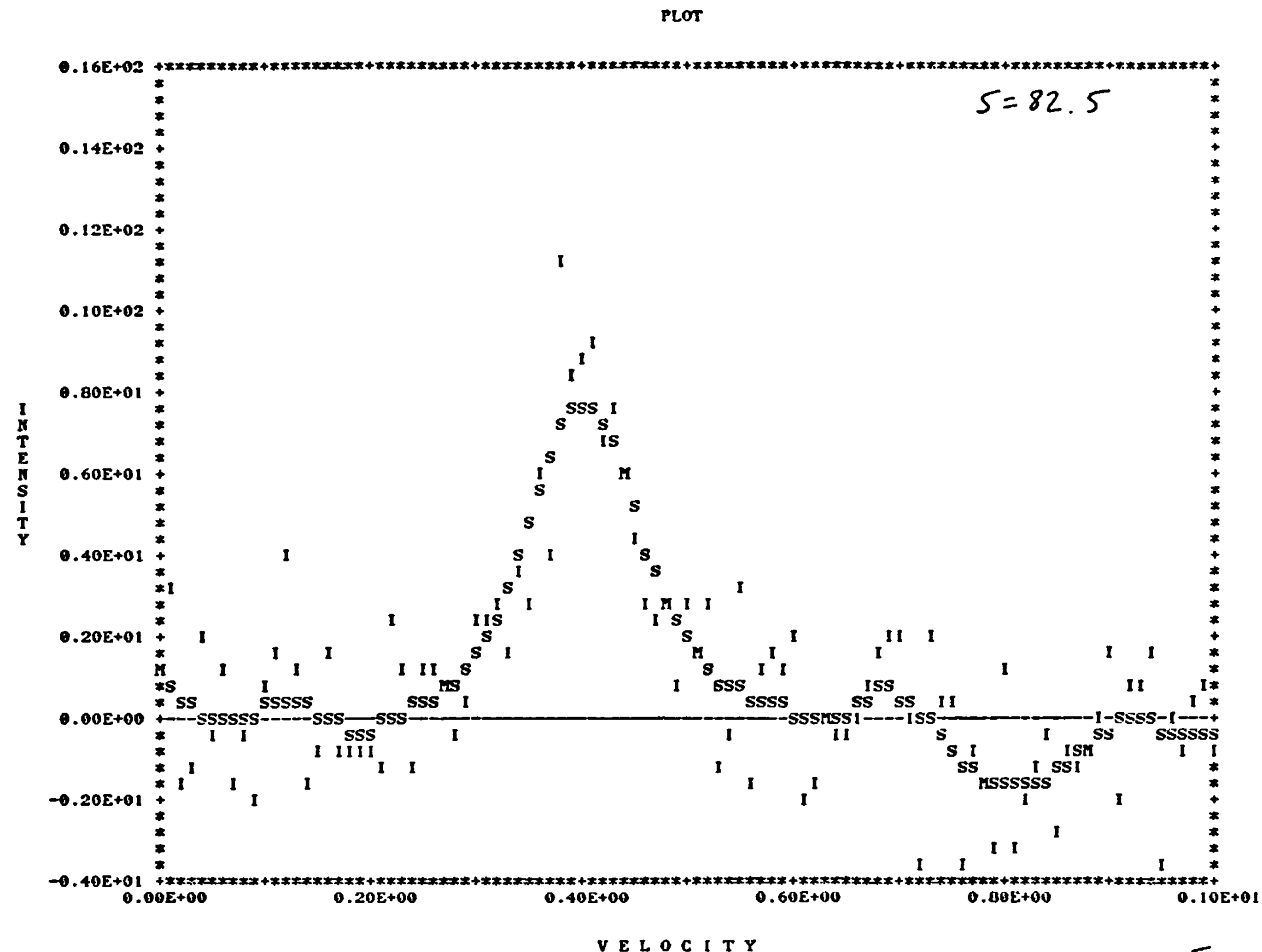


FIG. 3C. Same as FIG. 3A, except $S/N = 10$.



$$\hat{\theta} = (0.95764,\ 0.30832,\ -0.03729,\ 0.40023,\ 7.78205,\ 0.12306,\ 1.03836,\ 0.40193,\ 0.05207,\ 0.05204)^T$$

$$\hat{\theta}_J = (1.02833,\ 0.31753,\ -0.00015,\ 0.35756,\ 12.17877,\ 0.08164,\ 0.94442,\ 0.37532,\ 0.04477,\ 0.04399)^T$$

est'd standard deviations

0.24926	0.09778	0.05866	0.02526	2.35577	0.02482	0.14620	0.01273	0.00801	0.00868

1.00000	0.59764	0.78089	-0.24910	0.30015	0.65962	0.82659	-0.23471	0.59100	0.51963
0.59764	1.00000	0.87678	-0.04599	0.05766	0.48987	0.40309	0.00268	0.46664	0.42005
0.78089	0.87678	1.00000	-0.04018	0.05342	0.65750	0.52829	0.01520	0.63104	0.57506
-0.24910	-0.04599	-0.04018	1.00000	-0.76254	0.41849	-0.10083	0.96126	0.31545	0.46852
0.30015	0.05766	0.05342	-0.76254	1.00000	-0.51605	0.11514	-0.74974	-0.55940	-0.63227
0.65962	0.48987	0.65750	0.41849	-0.51605	1.00000	0.65278	0.41829	0.95714	0.95994
0.82659	0.40309	0.52829	-0.10083	0.11514	0.65278	1.00000	-0.18827	0.54423	0.52537
-0.23471	0.00268	0.01520	0.96126	-0.74974	0.41829	-0.18827	1.00000	0.33017	0.44472
0.59100	0.46664	0.63104	0.31545	-0.55940	0.95714	0.54423	0.33017	1.00000	0.97787
0.51963	0.42005	0.57506	0.46852	-0.63227	0.95994	0.52537	0.44472	0.97787	1.00000

est'd normalized correlation coefficients

FIG. 4. An example of error estimates derived by the procedure outlined in §5. Here $S/N = 5$, and the true parameters are (1.0, 0.4, 0.0025, 0.4, 7.9788, 0.1253, 0.9878, 0.4, 0.04776, 0.04776).