

ERRATUM

VLA SCIENTIFIC MEMORANDUM NO. 136

Robust Solution for Antenna Gains

Frederic R. Schwab

February, 1982

On page 5 of the memorandum, the definitions of L and M are in error.

The correct expressions are: $L \equiv e^{i(\psi_i - \psi_k)} v_{ik} \bar{v}_{ik}$ and $M \equiv e^{i(\psi_k - \psi_j)} v_{kj} \bar{v}_{kj}$.

VLA SCIENTIFIC MEMORANDUM NO. 136

Robust Solution for Antenna Gains

Frederic R. Schwab

September, 1981

VLA visibility data bases frequently are contaminated by "wild" points, or, in statistical jargon, by outliers. The outliers may be due, for example, to transient problems in the receiver/i.f. chains, to sampler problems or correlator problems, or to faults in system error detection. Because that part of the standard calibration procedure which consists in solution for antenna gains is quite sensitive to discordant data points, the calibrator data must be edited carefully before calibration. I will outline a new method for the solution for antenna gains which is much less sensitive to outliers than is the standard method and which, moreover, in the presence of well-behaved random errors, performs nearly as well as the standard least-squares method. Therefore, much of the burden of data editing can be shifted to the post-calibration stage of data reduction. Because the usual map/CLEAN procedure is quite sensitive to wild points, editing is still required -- but it may be accomplished somewhat more easily once the data have been corrected for antenna gains. I shall pay no attention here to the problem of identifying outliers, but only to the task of describing a calibration scheme which is not too severely affected by them.

Both the standard calibration program (ANTSOL) and the self-calibration program - until its most recent implementation - have been based on the method of least-squares. A popular (at least in linear regression), but non-classical, approach to parameter estimation in the presence of outliers is, instead of solving for parameters which minimize the sum of

squared residuals, to solve for parameters which minimize the sum of absolute values of the residuals. The latter approach is known as ℓ_1 minimization, and least-squares as ℓ_2 , since in both cases one minimizes a discrete ℓ_p norm of the residuals, $p = 1$ or 2 . To gain an intuitive feeling for the difference, consider that in the case of linear regression, if the constant term, or intercept, is included in the regression model, then the ℓ_2 solution has the property that the mean residual is equal to zero, whereas an ℓ_1 solution (the ℓ_1 solution to an overdetermined linear system is not necessarily unique) has the property that the median residual is zero. ℓ_1 solutions to linear problems can be computed by standard linear programming methods. Nonlinear ℓ_1 minimization, as is required for our problem, has not received much attention in the literature, but effective algorithms for this purpose recently have been published [1,2].

The basic idea of the approach to nonlinear ℓ_1 minimization that is outlined in [1] is this:

Replace the problem

find a vector x which minimizes $S(x) = \sum_i |f_i(x)|$

by a family of problems

find a vector x which minimizes $S_\epsilon(x) = \sum_i [f_i^2(x) + \epsilon]^{1/2}$

and, for a sequence $\epsilon_1 > \epsilon_2 > \dots$, tending to 0, apply an extrapolation procedure to the corresponding sequence of solution vectors, x_1, x_2, \dots .

(The extrapolation procedure converts a convergent sequence to a sequence which converges no less rapidly -- in practice, for a variety of extrapolation methods in common use, the convergence usually is accelerated.)

Such an indirect method is required because, for well-behaved f_i , $S(x)$

may be nondifferentiable at points where some f_i is equal to 0. Extrapolation usually is required because the problems get more ill-conditioned as ϵ tends to 0. The extrapolation method recommended in [1] is a standard one due to Fiacco and McCormick. In the present implementation of the method in the self-calibration program I haven't employed extrapolation because, in a variety of tests, the solutions have tended to stabilize for small ϵ before numerical instability has set in. I do, however, solve a sequence of problems (in a typical case, say, with $\epsilon = 5, .5, \text{ and } .05 \text{ mJy}$) since my minimization algorithms tend to require better initial guesses as ϵ decreases.

For general purpose ℓ_1 minimization the method outlined in [2] ought to be a more reliable method than that in [1]. For our specialized problem, I believe that the above approach is adequate.

For the case of complex antenna gains, we wish, given instantaneous observations \tilde{V}_{ij} obtained on $n(n-1)/2$ baselines $i-j$ ($i < j$), given corresponding model visibilities V_{ij} , and given weights w_{ij} , to solve for $g = (g_i)_{i=1}^n$ by minimizing the summation

$$S_\epsilon(g) = \sum_{i < j} w_{ij} [|\tilde{V}_{ij} - g_i \bar{g}_j V_{ij}|^2 + \epsilon]^{1/2}.$$

A necessary condition for S_ϵ to be minimized is that its gradient, $\nabla S_\epsilon(g)$, be equal to 0. We can rewrite the equation $\nabla S_\epsilon(g) = 0$ so that the k^{th} antenna gain, in terms of the others, is given by

$$\begin{aligned} g_k &= \frac{\sum_{i < k} w_{ik} [|\mathbf{R}_{ik}|^2 + \epsilon]^{-1/2} g_i v_{ik} \bar{\tilde{V}}_{ik} + \sum_{k < j} w_{kj} [|\mathbf{R}_{kj}|^2 + \epsilon]^{-1/2} g_j \bar{v}_{kj} \tilde{V}_{kj}}{\sum_{i < k} w_{ik} [|\mathbf{R}_{ik}|^2 + \epsilon]^{-1/2} |g_i v_{ik}|^2 + \sum_{k < j} w_{kj} [|\mathbf{R}_{kj}|^2 + \epsilon]^{-1/2} |g_j v_{kj}|^2} \\ &\equiv f_k(g_1, g_2, \dots, g_{k-1}, g_{k+1}, \dots, g_n), \text{ where } \mathbf{R}_{ij} \equiv \tilde{V}_{ij} - g_i \bar{g}_j V_{ij}. \end{aligned}$$

Choosing an initial guess, $g^{(0)}$, we can use an iterative relaxation method where, at the $(m+1)$ st iteration, the k^{th} antenna gain (for fixed ϵ) is approximated by

$$g_k^{(m+1)} = g_k^{(m)} + \omega [f_k(g_1^{(m+1)}, g_2^{(m+1)}, \dots, g_{k-1}^{(m+1)}, g_{k+1}^{(m)}, \dots, g_n^{(m)}) - g_k^{(m)}].$$

In practice, I have found successive over-relaxation (with $\omega = 1.5$) to converge reliably and in about the same number of iterations as the corresponding method for ℓ_2 solutions.* Typically, for each ϵ , 5-10 iterations are required in order to achieve a relative error of 10^{-4} in the solutions, starting, for ϵ_1 , with an initial guess $g_1^{(0)} = 1$.

Phase errors generally have a more deleterious effect on maps than do amplitude errors, thus, especially in self-calibration, one is sometimes driven by signal-to-noise considerations to wish to reduce the number of parameters and so to solve only for antenna-based phase errors, $\psi = (\psi_i)_{i=1}^n$. We then have

$$S_\epsilon(\psi) = \sum_{i < j} w_{ij} [|\tilde{v}_{ij} - e^{i(\psi_i - \psi_j)} v_{ij}|^2 + \epsilon]^{1/2}.$$

S_ϵ is periodic of period 2π in each of the ψ_k , and we may use this fact to drive an iterative scheme to a local minimum of S_ϵ rather than to a

*for ℓ_2 solutions,

$$g_k = \frac{\sum_{i < k} w_{ik} g_i \bar{v}_{ik} \tilde{v}_{ik} + \sum_{k < j} w_{kj} g_j \bar{v}_{kj} \tilde{v}_{kj}}{\sum_{i < k} w_{ik} |g_i \bar{v}_{ik}|^2 + \sum_{k < j} w_{kj} |g_j \bar{v}_{kj}|^2}.$$

local maximum:

$$\psi_k^{(m+1)} = \psi_k^{(m)} - \omega \tan^{-1} \left(\frac{\partial S_\epsilon / \partial \psi_k}{\partial^2 S_\epsilon / \partial \psi_k^2} \right).$$

$$\frac{\partial S_\epsilon}{\partial \psi_k} = - \sum_{i < k} w_{ik} [|R_{ik}|^2 + \epsilon]^{-1/2} \operatorname{Im}(L) + \sum_{k < j} w_{kj} [|R_{kj}|^2 + \epsilon]^{-1/2} \operatorname{Im}(M)$$

and

$$\begin{aligned} \frac{\partial^2 S_\epsilon}{\partial \psi_k^2} &= \sum_{i < k} w_{ik} \left\{ [|R_{ik}|^2 + \epsilon]^{-1/2} \operatorname{Re}(L) - [|R_{ik}|^2 + \epsilon]^{-3/2} \operatorname{Im}^2(L) \right\} \\ &\quad + \sum_{k < j} w_{kj} \left\{ [|R_{kj}|^2 + \epsilon]^{-1/2} \operatorname{Re}(M) - [|R_{kj}|^2 + \epsilon]^{-3/2} \operatorname{Im}^2(M) \right\}, \end{aligned}$$

where $L \equiv e^{i(\psi_i - \psi_k)}$, $M \equiv e^{i(\psi_k - \psi_j)}$, and $R_{ij} \equiv \tilde{V}_{ij} - e^{i(\psi_i - \psi_j)} V_{ij}$. *

In the circumstance in which the real parts of the visibility observations have an error distribution different from that of the imaginary parts, and perhaps in the circumstance in which the noise in the real part of an observation is independent of the noise in the imaginary part of that observation, it might be preferable to minimize the sum of the

*for ℓ_2 solutions,

$$2 \frac{\partial S}{\partial \psi_k} = - \sum_{i < k} w_{ik} \operatorname{Im}(L) + \sum_{k < j} w_{kj} \operatorname{Im}(M) \quad \text{and}$$

$$2 \frac{\partial^2 S}{\partial \psi_k^2} = \sum_{i < k} w_{ik} \operatorname{Re}(L) + \sum_{k < j} w_{kj} \operatorname{Re}(M).$$

absolute values of the real parts of the residuals plus the sum of the absolute values of the imaginary parts of the residuals, rather than the sum of the moduli of the complex residuals. However, I have tested this approach for the case of well-behaved random errors, with identical error distributions for real and imaginary, but independent errors in real part and imaginary part, and I found, in that case, very little difference in the computed solutions. In my implementation, this alternative solution method is much slower because more arithmetic is required. For brevity, the algebraic machinery for the alternative method has been omitted from this report.

The results of a series of tests comparing the ℓ_1 gain solution method with the ℓ_2 method are presented below. The data used in each test run simulate the observations of a one Jansky point source (located at the visibility phase reference position) by a 27 element interferometer. The visibility model exactly represents the source; i.e., $v_{ij} \equiv 1$. The observations \tilde{v}_{ij} ($1 \leq i < j \leq 27$) for each run are these model visibilities, modified by 27 multiplicative complex gains $g = (g_i)_{i=1}^{27}$, possibly with random noise added in. Weights $w_{ij} \equiv 1$ were used in all cases. In order to simulate wild observations, in some cases a number of the data points have been chosen at random and have no relation to the visibility model or to g .

Any solution to one of the minimization problems described above is unique only up to a multiplicative constant of unit modulus; that is, if \hat{g} is one local minimizer, then $c\hat{g}$ is another, whenever $|c|=1$. The solution algorithms are such that, with no special attention to the fact that the problems are underdetermined, they always converge upon some solution^{*} (barring unusual circumstances). In order most fairly to compare a com-

* Interestingly, if the argument of some gain is held fixed, then the rate of convergence is diminished.

puted solution, \hat{g} , with the true solution, g , each computed solution has been multiplied by that factor, c , of unit modulus, appropriate to minimize $\sum |g_i - c\hat{g}_i|^2$. In terms of the true solution, c is given by $\Sigma g_i \bar{\hat{g}}_i / |\Sigma g_i \bar{\hat{g}}_i|$.

Behavior when the data are contaminated only by well-behaved random errors

The ℓ_1 method clearly would be unattractive if, given data contaminated only by well-behaved random errors, it were to behave much worse than the least-squares method. Two series of tests were performed in which the model observations were contaminated by adding varying levels of zero mean (approximately-) Gaussian distributed random noise (independently) to the real parts and to the imaginary parts of the visibility observations; i.e., $\tilde{v}_{ij} = g_i \bar{g}_j v_{ij} + \text{noise}$ ($= g_i \bar{g}_j + \text{noise}$, for our 1 Jy point model). To test the two phase solution methods, the antenna gains were chosen to be of unit modulus, with phases drawn at random from the uniform distribution on $[0^\circ, 360^\circ]$. To test the complex gain solution methods, the phases of the antenna gains were chosen in the same manner as before, and the moduli were drawn from the uniform distribution on [.5, 1.5]. Ten trials were performed for each of eleven choices of standard deviation of the random noise. A different choice of gains and a different choice of noise were made for each trial (except that the same data were used for both the ℓ_1 and ℓ_2 solution tests; hence there will be some correlation in the mean solution errors, between ℓ_1 and ℓ_2). The rms error in each gain solution was computed as $\sqrt{n^{-1} \sum |g_i - \hat{g}_i|^2}$. The mean of these errors, for each set of ten trials, is shown in Table 1 and in the plot of Figure 1. For S/N $\gtrsim 3$ (resp., $\gtrsim 2$) the rms gain solution error remains $\lesssim .1$ for the phase solutions (resp., for the gain

solutions). Excluding the points at 2.0 Jy random noise, the ℓ_1 phase solution error (resp., gain solution error) is typically about 13% (resp., 14%) greater than the ℓ_2 error. The results of two representative trials (which were not included in the test sample) are shown in Figure 2.

Another test which might be of interest is the case $\hat{V}_{ij} = g_i \bar{g}_j (V_{ij} + \text{noise})$. Trials in which the observations were contaminated by noise of this form gave results which were very similar to those described above.

Sensitivity of the solution methods to observations chosen at random to be outliers

To test the sensitivity of the solution methods to randomly chosen outliers, some probability was assigned according to which an observation was chosen to be made a wild point. The wild points were assigned phases drawn at random from the uniform distribution on $[0^\circ, 360^\circ]$ and amplitudes drawn at random from the uniform distribution on $[0, 2]$ Jy. The "tame" observations were contaminated by 0.2 Jy Gaussian noise, as in the random noise tests, above. The results for complex gain solutions and for phase solutions are shown in detail in Figures 3 and 4, respectively. For data made wild with probabilities .05, .10, .20, and .50, the mean observed rms gain solution errors in ten trials were 6.3, 6.6, 8.1, and $23 (\times 10^{-2})$ (9.7, 12, 19, and $40 (\times 10^{-2})$) for the ℓ_1 (resp., for the ℓ_2) solution method. In the corresponding test of the phase solution methods, assigning the same probabilities, the mean observed rms errors were 4.1, 4.8, 4.9, and $7.7 (\times 10^{-2})$ (5.3, 7.3, 10.0, and $27 (\times 10^{-2})$) for the ℓ_1 (resp., for the ℓ_2) solution method.

To test the sensitivity to the presence of extreme outliers, two

trials were run in which each observation, with probability .10, was assigned a random phase and a random amplitude in the range $[0, 10^{10}]$ Jy. The results are shown in Figures 5a and 5b. The rms ℓ_1 (resp., ℓ_2) gain solution errors observed were 7.2 and 4.5 ($\times 10^{-2}$) (resp., 4.2×10^7 and 1.3×10^0) for the complex gain solution and for the phase solution, respectively. This test, with ten trials in each case, was repeated for the case of wild amplitudes in the range $[0, .01]$ Jy. The results are shown in Figures 5c and 5d. The mean observed rms gain solution errors for ℓ_1 (resp., for ℓ_2) were 7.2 and 4.6 ($\times 10^{-2}$) (10.9 and 4.1 ($\times 10^{-2}$)) in the complex gain solutions and in the phase solutions, respectively. In the presence of extreme outliers, the ℓ_1 method appears in all cases, except that of phase solutions with very small amplitude outliers, to outperform the ℓ_2 method.

Sensitivity of the solution methods to bad i.f.'s

An occasional problem with VLA data is that the data associated with a single i.f. may be bad; sometimes the system error detection facility fails to recognize the problem. In these tests, 0.2 Jy Gaussian distributed random noise was added to the observations, as above, except that all the visibilities associated with antenna 5 were assigned a given amplitude and a random phase uniformly distributed on $[0^\circ, 360^\circ]$. Representative results, along with the rms gain solution errors that were observed in ten trials with identical error distributions, are displayed in Figures 6 and 7. For amplitudes associated with antenna 5 of 0.1, 1.0, 2.0, and 5.0 Jy, mean rms gain solution errors (excluding antenna 5) of 5.8, 5.9, 6.6, and 8.2 ($\times 10^{-2}$) (5.2 , 5.3 , 6.6 , and 80 ($\times 10^{-2}$)) were observed for the ℓ_1 (resp., the ℓ_2) solutions when solving for full complex gains. In

ten trials with 26 antennas and 0.2 Jy random noise, with no outliers, the mean observed rms gain solution error was $5.9 \pm .5$ and $5.3 \pm .4$ ($\times 10^{-2}$) for the ℓ_1 and ℓ_2 solutions, respectively.

When solving only for phase errors, for visibility amplitudes associated with antenna 5 of 0.5, 1.0, 2.0, and 100.0 Jy, mean rms gain solution errors (excluding antenna 5) of 5.1, 4.6, 5.3, and 5.4 ($\times 10^{-2}$) (4.8, 5.2, 7.3, and 126 ($\times 10^{-2}$)) were observed with the ℓ_1 (resp., the ℓ_2) solution method. In ten trials with 26 antennas, 0.2 Jy random noise, and no outliers, the mean observed rms gain solution error was $4.6 \pm .6$ and $4.1 \pm .5$ ($\times 10^{-2}$) for the ℓ_1 and ℓ_2 solutions, respectively.

Clearly it is desirable to detect the presence of a bad i.f. before solving for antenna gains, no matter which solution method is used. In the case of full complex gain solutions, if there is a bad i.f., and if the associated amplitudes are very large, then the ℓ_1 method is preferable. For phase solutions, the ℓ_1 method is extremely insensitive to a bad i.f. In both cases, the ℓ_1 method is strikingly superior to the least-squares method only when the errors associated with the bad i.f. are very large.

A few additional tests were run: If the observations associated with antenna 5 are assigned not a random phase, but each is assigned, instead, a phase of 45 degrees, then the conclusions above remain unaltered. In the case of not a single bad i.f., but of five bad i.f.'s, it is still preferable to detect the presence of the bad i.f.'s before solving for the gains (the solutions are quite inferior to those for 22 good i.f.'s), and the ℓ_2 solutions still are not much inferior to the ℓ_1 solutions except when the amplitudes of the bad points are very large.

Conclusions

The ℓ_1 method for gain solutions appears to be an attractive alternative to the least-squares method, both in self-calibration and, perhaps more so, in standard calibration. In the case of standard calibration in the pipeline data reduction system, the ℓ_1 method has appeal because in the pipeline it might not be possible to correct data editing errors before calibration. The appeal in self-calibration lies in the fact that, in combination with a scheme for the identification of outliers, one would be able first to solve for the gains and then to identify and flag the outliers, resting assured that the gain solutions probably were not unduly influenced by the outliers (otherwise it would be necessary to recompute the gains after flagging data). Since the ℓ_1 method performs nearly as well as ℓ_2 in the presence of well-behaved errors, any decision not to use ℓ_1 in preference to ℓ_2 probably would be based either upon faith that there were not many harmful outliers or upon considerations of computational speed. In a careful implementation, especially if the gains are computed in an array processor, the penalty due to the added computational complexity ought not to be too great.

Cornwell's recent error analysis of calibration [3] is a useful complement to the test results reported here. I plan to do a related empirical study of the sensitivity of simultaneous solution for antenna gains and for constant or slowly time-varying additive or multiplicative correlator based errors, to errors in the observations and to perturbations in the source model.

The problem of identifying outliers merits careful attention. [4] ought to be a useful reference on this subject.

TABLE 1.

rms error in computed gain solution x100

rms noise (Jy)	phase solutions		complex gain solutions	
	ℓ_1	ℓ_2	ℓ_1	ℓ_2
.05	1.0 ± .1	1.0 ± .1	1.4 ± .2	1.3 ± .2
.10	2.0 ± .3	1.9 ± .3	2.9 ± .4	2.6 ± .3
.15	3.0 ± .3	2.9 ± .3	4.4 ± .5	3.8 ± .4
.20	4.7 ± .6	4.0 ± .5	6.0 ± .6	5.3 ± .5
.25	5.4 ± .6	4.7 ± .7	7.1 ± .8	6.4 ± .7
.35	6.9 ± .8	6.0 ± .4	11 ± 1	9.2 ± 1
.50	10 ± 1	8.7 ± 1	16 ± 2	13 ± 1
.75	16 ± 3	14 ± 2	24 ± 3	20 ± 2
1.00	22 ± 3	19 ± 3	30 ± 4	26 ± 4
1.50	35 ± 6	31 ± 4	49 ± 7	43 ± 5
2.00	69 ± 23	49 ± 12	74 ± 17	61 ± 10

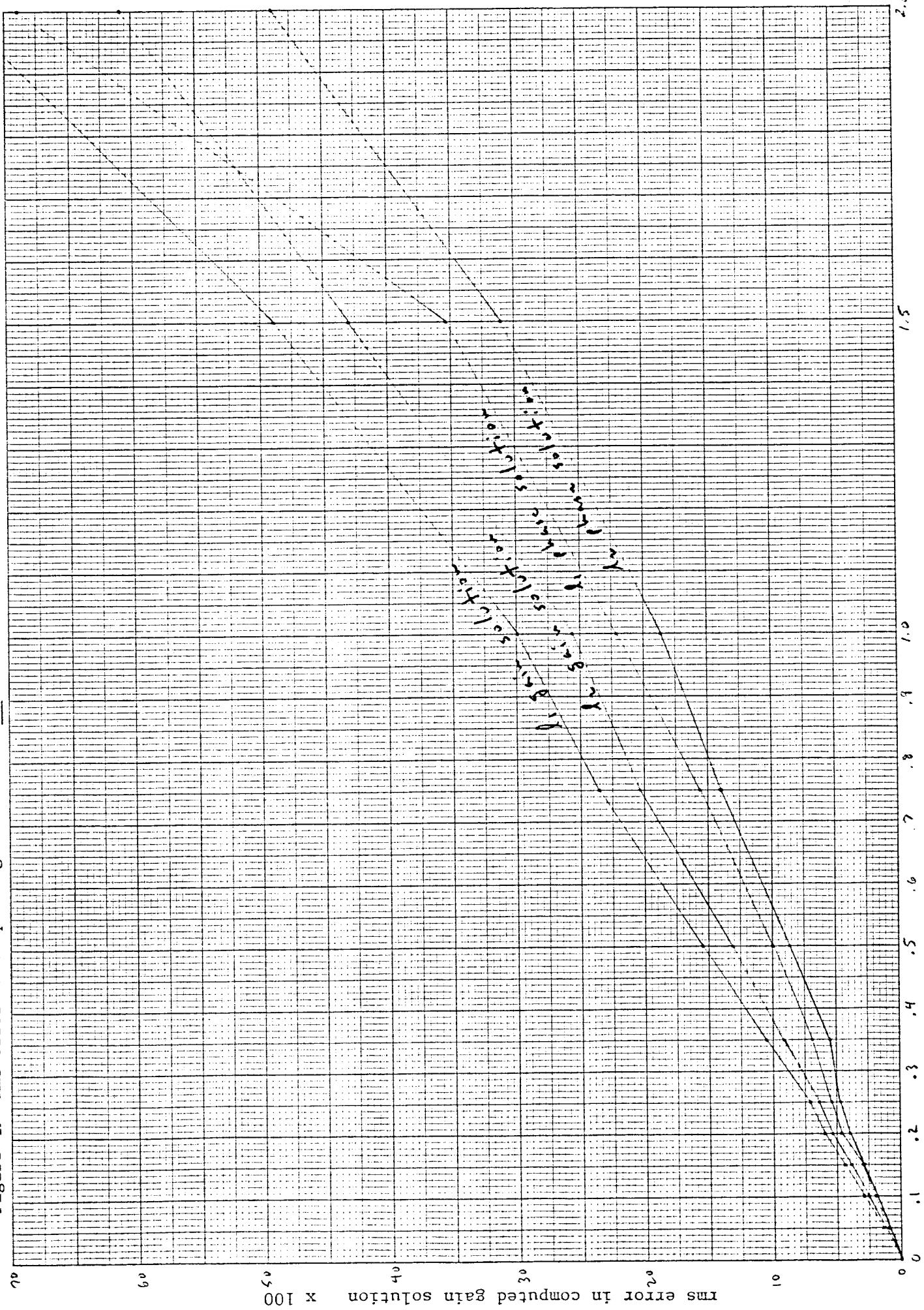
10 trials/datum

simulated observations of 1 Jy point source

REFERENCES

1. El-Attar, R. A., M. Vidyasagar, and S. R. K. Dutta, "An algorithm for ℓ_1 -norm minimization with application to nonlinear ℓ_1 -approximation", SIAM J. Num. Anal., 16 (1979), 70-86.
2. Murray, W. and M. L. Overton, "A projected Lagrangian algorithm for nonlinear ℓ_1 approximation", SIAM J. Sci. Stat. Comp., 2 (1981), 207-224.
3. Cornwell, T., "An error analysis of calibration", VLA Scientific Memo. No. 135, September 1981.
4. Barnett, V., and T. Lewis, Outliers in Statistical Data, Wiley, Chichester, 1978.

Figure 1. Rms error in computed gain solution vs. rms noise added to observations. (10 trials/datum)



-15-
L1 SOLUTION

TRUE SOLUTION

ANT	AMP	PHASE	AMP	PHASE	AMP	PHASE
1	0.74019	18.96	0.68422	16.14	0.69212	16.55
2	0.76571	-142.61	0.76475	-147.41	0.76112	-146.78
3	1.46513	135.12	1.43573	134.16	1.42796	134.83
4	0.69369	139.73	0.78972	140.11	0.78699	137.89
5	1.29463	-112.25	1.19058	-112.99	1.17391	-113.42
6	0.99279	114.12	0.98797	114.17	0.99179	112.62
7	1.25279	73.59	1.29354	71.34	1.29637	72.74
8	1.10175	-132.58	1.14248	-134.57	1.12624	-133.98
9	0.63555	-166.74	0.63576	-161.29	0.68157	-161.84
10	1.29148	64.92	1.29272	65.87	1.29738	65.12
11	0.59614	86.63	0.56424	84.91	0.57376	85.18
12	1.31672	161.59	1.34419	163.76	1.35684	159.88
13	1.21291	33.59	1.11781	33.09	1.13274	33.16
14	1.21615	115.15	1.26026	116.42	1.27599	117.71
15	0.68938	-77.35	0.73036	-79.32	0.69313	-73.14
16	0.71566	40.16	0.70078	42.26	0.63445	43.15
17	0.87368	155.69	0.85458	155.59	0.85356	156.47
18	0.74752	66.06	0.82046	64.84	0.79710	55.27
19	0.73871	41.33	0.72071	44.43	0.73614	43.78
20	0.90312	61.31	0.92673	58.27	0.94906	58.03
21	0.74135	111.30	0.73756	113.17	0.74264	111.79
22	1.10917	-134.93	1.02903	-137.61	1.03897	-137.41
23	0.55409	-21.38	0.53543	-25.68	0.58013	-24.61
24	0.61669	-132.47	0.53754	-129.21	0.62701	-129.51
25	1.22916	163.25	1.23000	167.84	1.23703	165.14
26	1.37211	-116.55	1.48111	-115.99	1.40623	-116.18
27	1.17198	8.24	1.20096	7.63	1.21970	7.94

RMS SOLUTION ERROR:

6.45E-02

5.76E-02

complex gain solution
0.2 Jy random noise

BADIF=F
NOISEAMP= 0.20 JY
WILDPCT= 0.00
WILDAMP= 0.000E+00 JY

TRUE SOLUTION

L1 SOLUTION

L2 SOLUTION

ANT	AMP	PHASE	AMP	PHASE	AMP	PHASE
1	1.00000	21.48	1.00000	19.56	1.00000	23.72
2	1.00000	-159.77	1.00000	-154.83	1.00000	-158.31
3	1.00000	119.24	1.00000	119.42	1.00000	118.35
4	1.00000	86.25	1.00000	91.90	1.00000	87.84
5	1.00000	129.27	1.00000	136.07	1.00000	135.03
6	1.00000	156.29	1.00000	161.20	1.00000	159.96
7	1.00000	-58.31	1.00000	-64.59	1.00000	-65.39
8	1.00000	121.55	1.00000	113.52	1.00000	117.62
9	1.00000	97.42	1.00000	102.99	1.00000	98.68
10	1.00000	167.08	1.00000	165.80	1.00000	168.37
11	1.00000	104.05	1.00000	109.47	1.00000	103.14
12	1.00000	-93.02	1.00000	-91.30	1.00000	-93.20
13	1.00000	81.05	1.00000	75.48	1.00000	76.08
14	1.00000	-78.73	1.00000	-77.62	1.00000	-77.98
15	1.00000	-154.90	1.00000	-150.37	1.00000	-151.60
16	1.00000	-11.00	1.00000	-16.54	1.00000	-15.28
17	1.00000	-81.36	1.00000	-81.81	1.00000	-76.75
18	1.00000	-149.09	1.00000	-152.95	1.00000	-151.72
19	1.00000	-156.57	1.00000	-153.85	1.00000	-155.18
20	1.00000	32.21	1.00000	32.30	1.00000	29.34
21	1.00000	-120.13	1.00000	-125.01	1.00000	-124.30
22	1.00000	98.61	1.00000	96.11	1.00000	95.71
23	1.00000	-134.91	1.00000	-131.66	1.00000	-129.81
24	1.00000	-28.66	1.00000	-27.06	1.00000	-25.22
25	1.00000	56.97	1.00000	50.29	1.00000	56.09
26	1.00000	71.30	1.00000	82.93	1.00000	77.83
27	1.00000	66.68	1.00000	64.44	1.00000	65.79

RMS SOLUTION ERROR:

8.35E-02

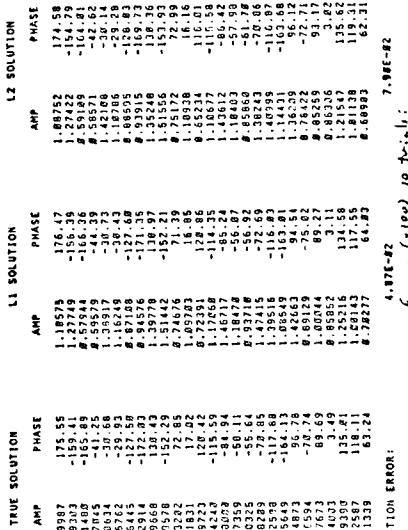
6.03E-02

phase solution
0.35 Jy random noise

BADIF=F
NOISEAMP= 0.35 JY
WILDPCT= 0.00
WILDAMP= 0.000E+00 JY

Figure 2. Two representative trials illustrating behavior of the solution methods in the presence of random noise.

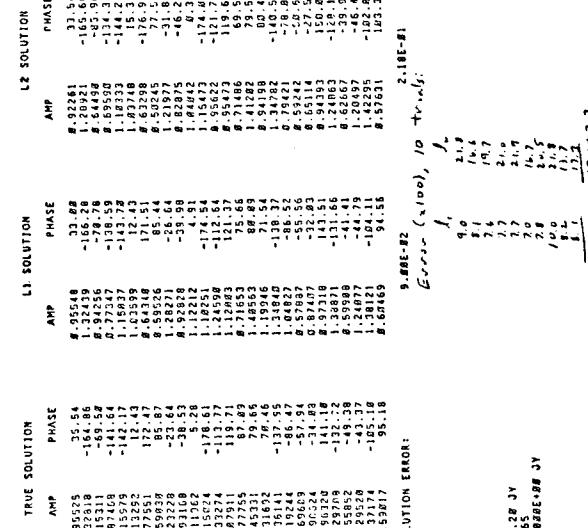
a) $\rho = .05$



$$\begin{aligned} & \text{L2 SOLUTION: } \mathcal{E}_{\text{true}}(f_1, f_2) = 7.98 \times 10^{-12} \\ & \text{RMS SOLUTION ERROR: } \mathcal{E}_{\text{rms}}(f_1, f_2) = 4.37 \times 10^{-12} \end{aligned}$$

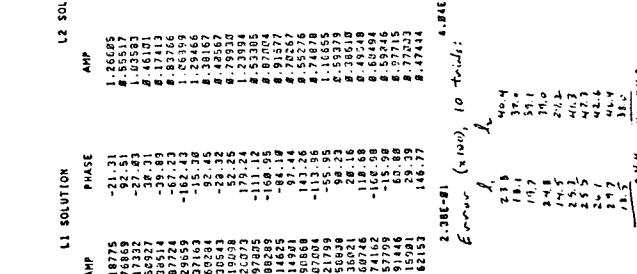
$\rho = .05$

c)



$\rho = .50$

d)



b)

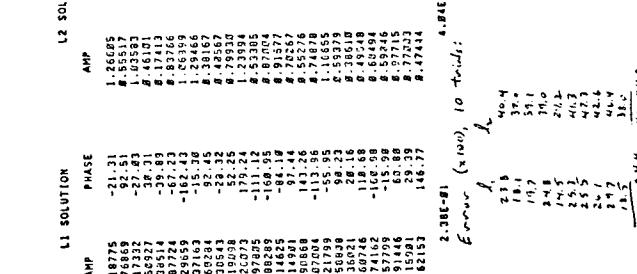


Figure 3. Sensitivity of complex gain solutions to outliers. The "tame" observations had random noise of 0.2 Jy added. Each observation, with a given probability, P, was made a wild point by assigning it a random phase and an amplitude in the range 0-2 Jy.

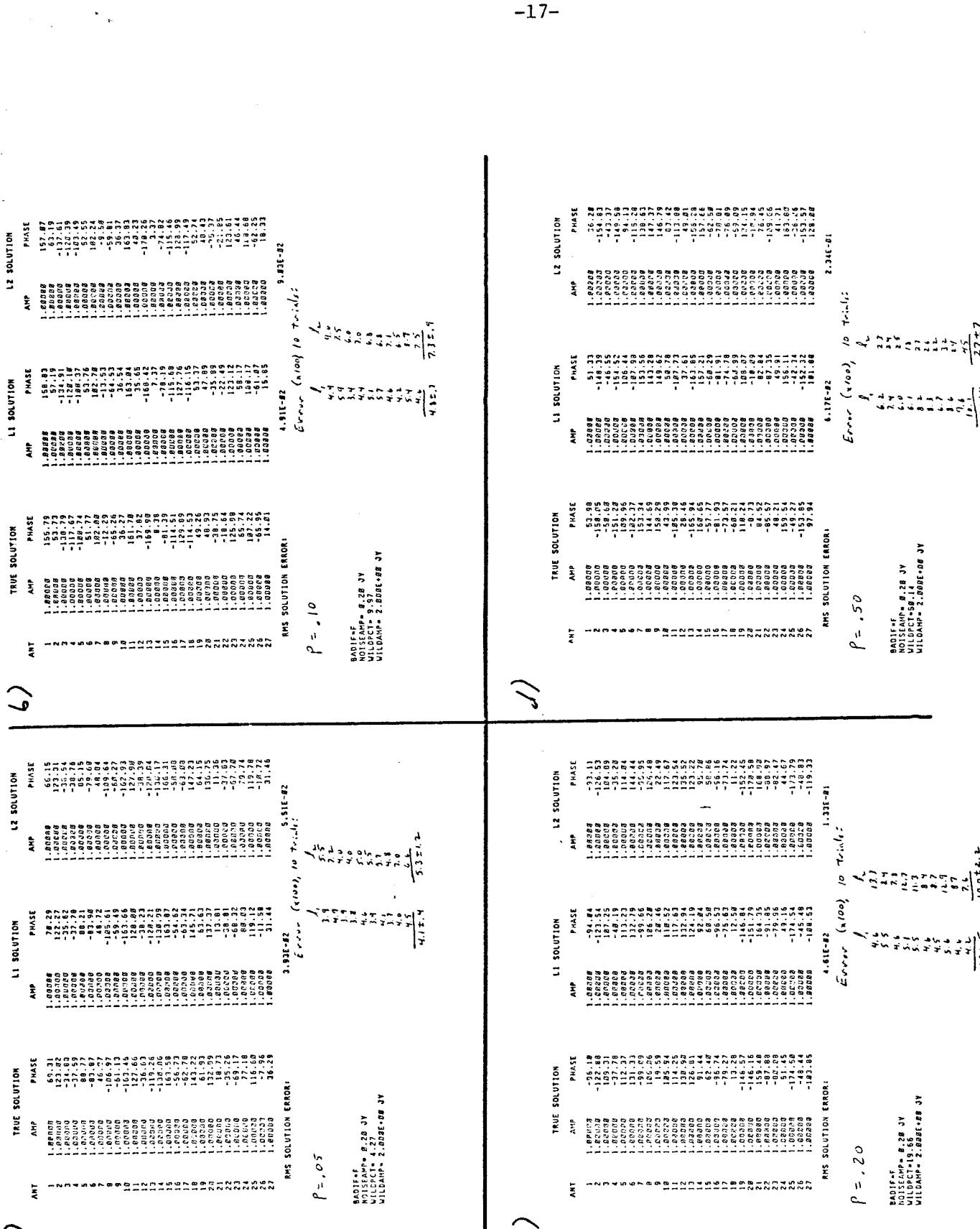


Figure 4. Sensitivity of phase solutions to outliers. The "tame" observations had random noise of 0.2 Jy added. Each observation, with probability P , was made a wild point by assigning it a random phase and an amplitude in the range $0\text{--}2$ Jy.

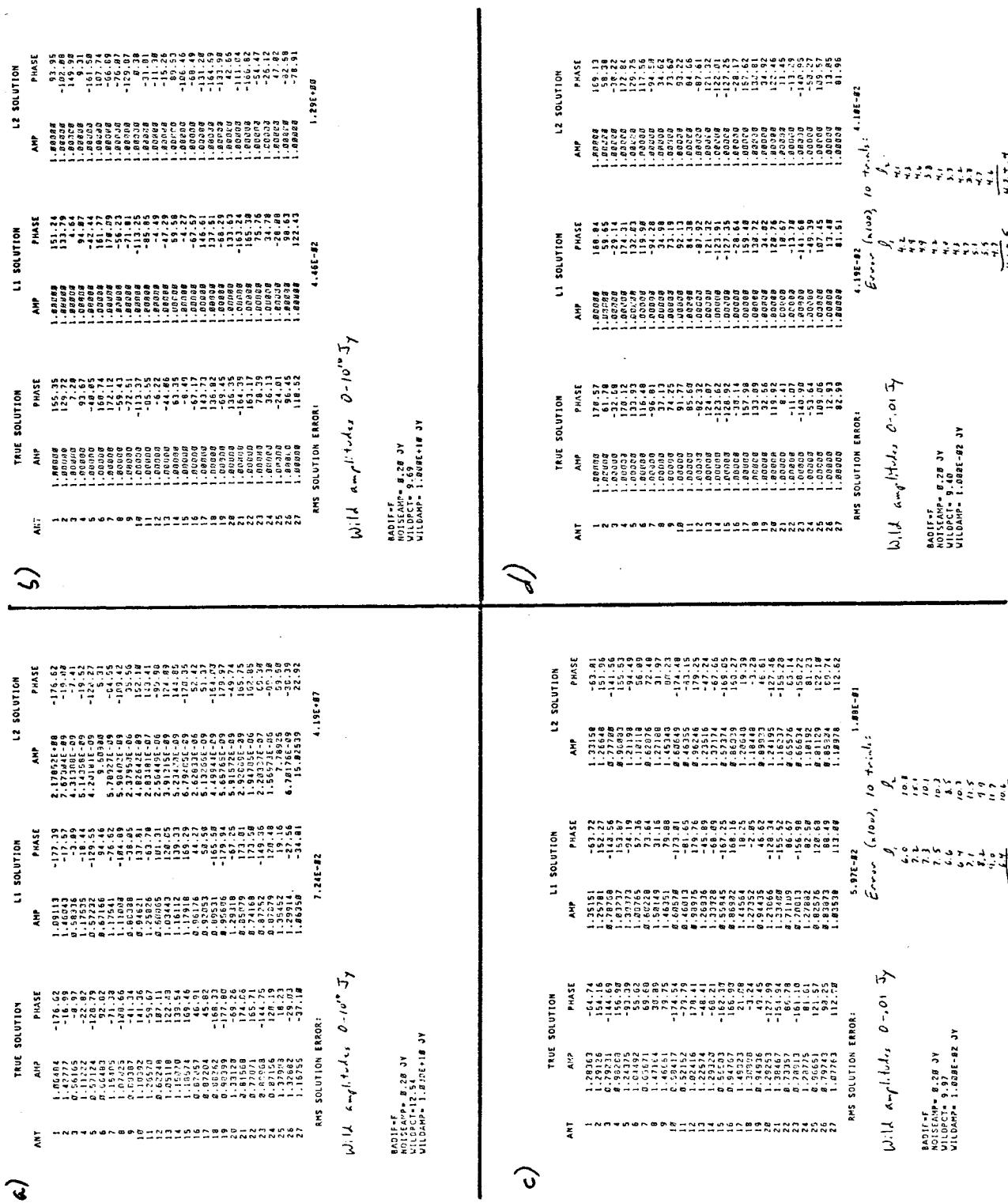


Figure 5. Sensitivity to the presence of extreme outliers. "Tame" observations have 0.2 Jy random noise added. Observations were made wild with probability .10 . a) and b) illustrate

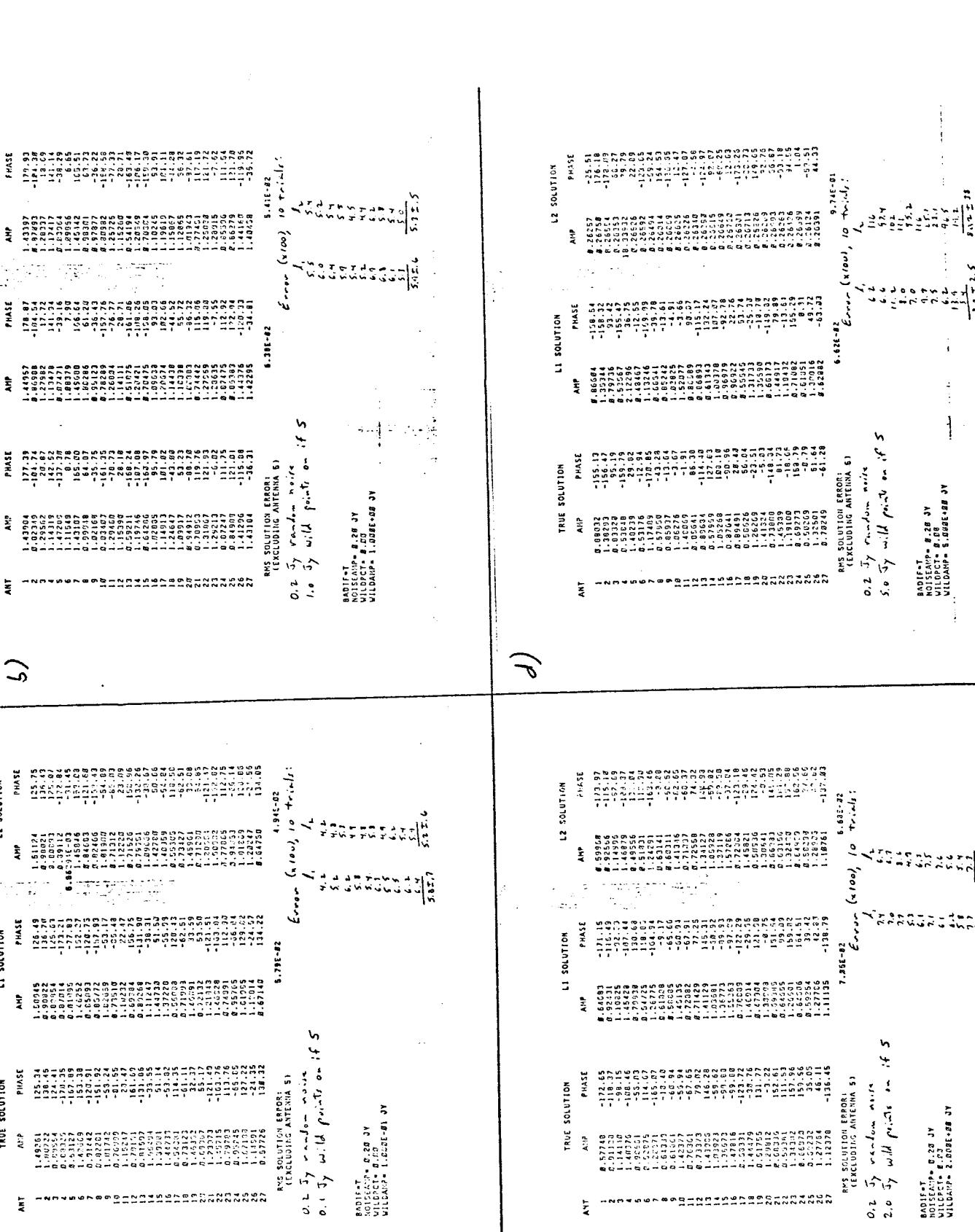


Figure 6. Sensitivity of complex gain solutions to a bad i.f.

