GEOPHYSICS, VOL. 62, NO. 4 (JULY-AUGUST 1997); P. 1524–1532, 7 FIGS.

Reconditioning inverse problems using the genetic algorithm and revised parameterization

Andrew Curtis* and Roel Snieder*

ABSTRACT

The better conditioned an inverse problem is, the more independent pieces of information may be transferred from the data to the model solution, and the less independent prior information must be added to resolve trade offs. We present a practical measure of conditioning that may be calculated swiftly even for large inverse problems. By minimizing this measure, a genetic algorithm can be used to find a model parameterization that gives the best conditioned inverse problem. We illustrate the method by finding an optimal, irregular cell parameterization for a cross-borehole tomographic example with a given source-receiver geometry. Using the final parameterization, the inverse problem is almost a factor of three better conditioned than that using an average random parameterization. In addition, this method requires little additional programming when solving a linearized inverse problem. Hence, the improvement in conditioning and corresponding increase in independent information available for the model solution essentially come for free.

INTRODUCTION

Many geophysical inverse problems are ill conditioned because the model space contains more detail than can be resolved using the available data space (Bertero et al.,1985). The portion of the model space that cannot be resolved is called the null space. Such problems are usually dealt with in one of two ways: first, eigenvectors that span the null space may be removed from the model space explicitly using singular-value decomposition, creating a solution in a "reduced" model space (Michelena,1993). However, this method reduces the effective information content of the data set and resolution in the model space in a nontrivial manner. Alternatively, the inclusion of a priori information (smoothness constraints, reduction of the model space, compatibility with other data) may reduce the null space by directly restricting the range of possible solutions. First, however, prior information restricting the model space must be reliable and independent; second, smoothness constraints are often unrealistic, and finally, the solution depends on assumptions other than the data set alone. If our goal is to test some hypothesis about the Earth independently, then we must maximize the amount of information from the current data set and minimize the amount of additional information included in the solution.

Consider the problem of estimating average wave-propagation velocity structure across a region spanned by many earthquake source–receiver paths along which average wave velocities were measured. The structure of the data space (the source–receiver geometry) is often fixed and the relation between the data and model space is determined by the physical laws of nature (the propagation of seismic waves through the Earth). We may influence conditioning only through the model parameterization and hence structure of the model space.

Consider the path geometry in Figure 1. We suppose that an error-free measurement of average velocity is available for each source (circle)–receiver (square) path. The four cells shown in Figure 1a have raypath coverage exactly equal. From measurements along the two left-hand paths, the quantity $v_1 + v_2$ can be determined exactly. The quantity $v_1 - v_2$, however, remains entirely unresolved by the data, and hence velocities v_1 and v_2 cannot be determined. Similarly, the combination $v_3 + v_4$ may be determined exactly, but $v_3 - v_4$ and hence v_3 and v_4 cannot be determined.

Let $\mathbf{v} = [v_1, \ldots, v_4]^T$. Then the well-determined parameter combinations are $\mathbf{e}_1 \cdot \mathbf{v}$ and $\mathbf{e}_2 \cdot \mathbf{v}$, the undetermined combinations are $\mathbf{e}_3 \cdot \mathbf{v}$ and $\mathbf{e}_4 \cdot \mathbf{v}$, where

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \tag{1}$$

Manuscript received by the Editor October 18, 1995; revised manuscript received August 30, 1996.
*Department of Theoretical Geophysics, Faculteit Aardwetenschappen, Postbus 80.021, 3508 TA Utrecht, Netherlands.

and vectors $\mathbf{e}_1$ to $\mathbf{e}_4$ are the eigenvectors of the inverse problem of finding velocities v_1 to v_4. Parameter combinations parallel to eigenvectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are completely determined if, and only if (*iff*) the corresponding eigenvalues are large (and equal in this case). Similarly, combinations parallel to eigenvectors $\mathbf{e}_3$ and $\mathbf{e}_4$ are undetermined *iff* they correspond to zero eigenvalues. Hence, the eigenvalues in order of decreasing magnitude (and normalized by the largest) may be plotted as shown in Figure 1a. The two zero eigenvalues correspond to linear combinations of model parameters that are unconstrained by the data. All velocities v_1 to v_4 can be estimated only if we add two extra pieces of prior information that constrains the linear combinations $v_1 - v_2$ and $v_3 - v_4$. Otherwise we must remove altogether eigenvectors $\mathbf{e}_3$ and $\mathbf{e}_4$ that span the model null space and content ourselves with the "reduced" solution for $v_1 + v_2$ and $v_3 + v_4$.

If we had tried to find a cell parameterization based on our knowledge of the averaging properties of velocities along raypaths, and of the structure (geometry) of the data space, we might have chosen the cell parameterization shown in Figure 1b. Each cell has equal path coverage and the eigenvectors $[\mathbf{e}_1 \mid \mathbf{e}_2 \mid \mathbf{e}_3 \mid \mathbf{e}_4] = \mathbf{I}$, the identity matrix. The corresponding eigenvalues are all equal and may be normalized and plotted as shown in the Figure, reflecting the fact that in this case, velocities v_1 to v_4 are determined completely.

These examples show that raypath density, although necessary, is not nearly a sufficient criterion to design the model space structure, since it is a single summarizing statistic of the relationship between model and data space structures (the forward problem). The eigenvalues and eigenvectors on the other hand contain *all* relevant information about the forward problem (that path velocities are averages over traversed cells). The eigenvalues show exactly how well-determined each linear combination of model parameters is, and the more eigenvalues that are close to zero the more independent pieces of prior information must be added to resolve trade offs between the corresponding unconstrained parameter combinations. In general, when the number of data and model parameters is large we will find an eigenvalue spectrum similar to those shown in Figure 2 (for positive eigenvalues). However, from the argument above, the solid curve in Figure 2 corresponds to a model that is better constrained by the data than that of the dashed curve. Hence, a better measure of the constraint on the model space provided by the data space is the area under the eigenvalue curve, or the sum of normalized (positive) eigenvalues.

Below we present a method to estimate quickly and efficiently the area under the normalized eigenvalue curve for any inverse problem, and demonstrate how a genetic algorithm may be used to design the model parameterization to best suit the structure of the forward problem and data space using this new statistic. Usually the model parameterization is fixed (e.g., blocks of specific size, predefined basis functions), injecting prior information about the structure of the model space. The procedure shown here does not remove this prior

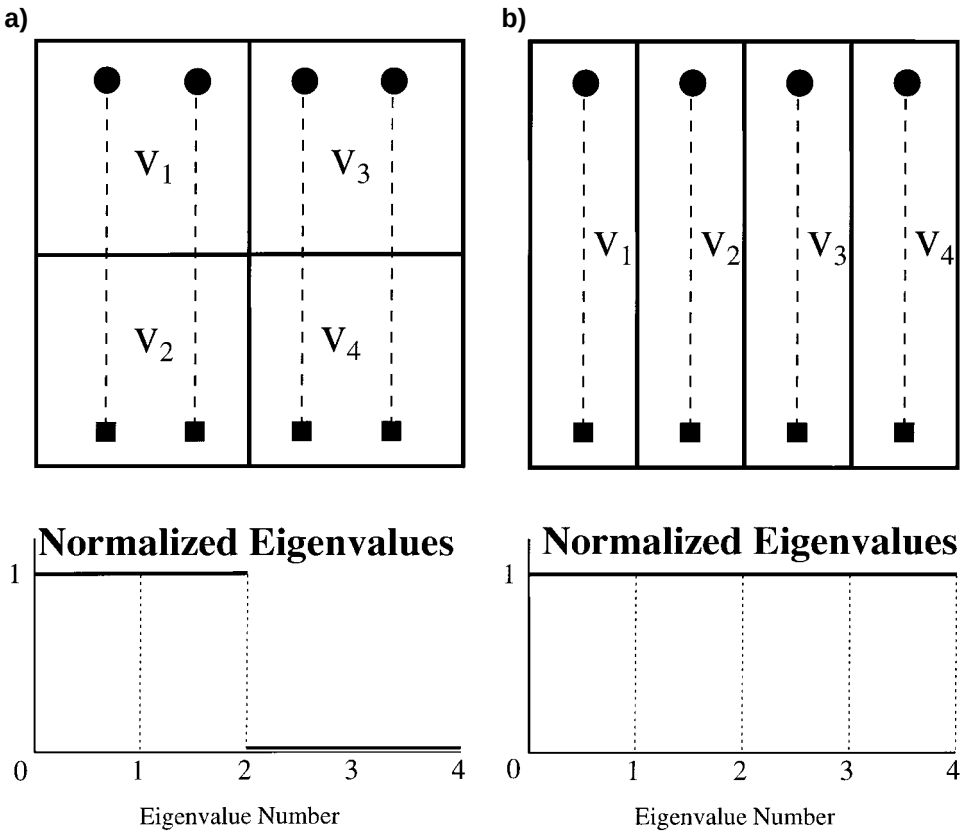


FIG. 1. Cell geometries for the event (circle)—station (square) paths shown, and corresponding eigenvalue spectrum normalized by the largest eigenvalue. In (a) the cells bisect each path whereas in (b) each cell contains exactly one path length. Both cell geometries have exactly the same homogeneous path coverage within each cell.

information completely, but it makes model parameterization a valuable and integral part of the inversion procedure that maximizes the number of independent pieces of information in any solution, rather than being an arbitrary addition of prior information. As long as the set of all possible parameterizations from which the genetic algorithm can choose are equally reasonably based on our prior knowledge, choosing any one of them should not bias our solution unreasonably. The final choice is optimum with respect to the particular data set available and to the forward problem. In effect we are gaining for free a better-conditioned inverse problem and hence more independent information in any final solution.

METHODOLOGY

For any given model vector $\mathbf{m}$ within model space M, let the forward problem of estimating the corresponding data vector $\mathbf{d}$ in data space D be solved by the matrix operator $\underset{\sim}{\mathbf{A}}$, i.e.,

$$\mathbf{d} = \underset{\sim}{\mathbf{A}}\mathbf{m}. \tag{2}$$

Then for a given data vector $\mathbf{d}_0$, we wish to find a model vector $\mathbf{m}_0 \in M$ such that $\|\mathbf{d}_0 - \underset{\sim}{\mathbf{A}}\mathbf{m}_0\|_2$ is minimized, where $\|\cdot\|_2$ is the L_2 norm. Theoretically, if the problem is purely over-determined this is accomplished by premultiplying equation (2) by $\underset{\sim}{\mathbf{A}}^T$ and taking a matrix inverse:

$$\mathbf{m} = (\underset{\sim}{\mathbf{A}}^T \underset{\sim}{\mathbf{A}})^{-1} \underset{\sim}{\mathbf{A}}^T \mathbf{d} \tag{3}$$

Instability in the solution arises because the $N \times N$ square matrix,

$$\underset{\sim}{\mathbf{L}} = \underset{\sim}{\mathbf{A}}^T \underset{\sim}{\mathbf{A}} \tag{4}$$

is often near singular, i.e., some of its eigenvectors, $\{\mathbf{e}_i : i = 1, \ldots, N\}$ say, have extremely small eigenvalues $\{\lambda_i : i = 1, \ldots, N\}$. Measurement errors in the data space D propagate into solution $\mathbf{m}$ parallel to each eigenvector $\mathbf{e}_i$ with an amplification $1/\lambda_i$; hence, the solution becomes unstable and unreliable (Menke, 1989).

As illustrated in the example above, if we have no prior information about preferred model parameterizations, we minimize the singularity of $\underset{\sim}{\mathbf{L}}$ and hence the amount of extra prior information necessary to stabilize the calculation of $\underset{\sim}{\mathbf{L}}^{-1}$ by choosing a parameterization that maximizes the area under the absolute, normalized eigenvalue curve. $\underset{\sim}{\mathbf{A}}$ is real so that $\underset{\sim}{\mathbf{L}}$ is positive definite (i.e., all eigenvalues of $\underset{\sim}{\mathbf{L}}$ are positive), and hence,

$$\text{Area under curve} = \frac{1}{\lambda_1} \sum_{i=1}^{N} \lambda_i, \tag{5}$$

where λ_1 is the largest eigenvalue. The complete eigenvalue spectrum is expensive to calculate. However, the trace of a matrix is invariant under a similarity transformation; since a diagonal matrix with elements equal to the eigenvalues of $\underset{\sim}{\mathbf{L}}$ can be found by a similarity transform of $\underset{\sim}{\mathbf{L}}$ alone, we have

$$\sum_{i=1}^{N} \lambda_i = \text{trace}(\underset{\sim}{\mathbf{L}}). \tag{6}$$

Also, the value of λ_1 may be estimated quickly using the power method as follows. Any random unit vector $\mathbf{U}$ can be expressed in terms of the eigenvector basis with coefficients α_i:

$$\mathbf{U} = \sum_{i=1}^{N} \alpha_i \mathbf{e}_i. \tag{7}$$

Then,

$$\underset{\sim}{\mathbf{L}}^n \mathbf{U} = \sum_{i=1}^{N} \alpha_i \lambda_i^n \mathbf{e}_i, \tag{8}$$

and if $\alpha_1 \neq 0$ then $\underset{\sim}{\mathbf{L}}^n \mathbf{U} \to \alpha_1 \lambda_1^n \mathbf{e}_1$ as $n \to \infty$. Hence

$$\frac{|\underset{\sim}{\mathbf{L}}^n \mathbf{U}|^2}{(\underset{\sim}{\mathbf{L}}^n \mathbf{U}) \cdot \mathbf{U}} \to \lambda_1^n \quad \text{as} \quad n \to \infty \tag{9}$$

with equality *iff* all eigenvalues are equal. We remove the problems associated with $\alpha_1 = 0$ by estimating λ_1 three times using equation (9). Each time a different source vector $\mathbf{U}$ is used, each rotated from the previous one by 73° in a random direction in each coordinate pair, where the first of the three source vectors has completely random direction. The resulting three estimates of λ_1 are averaged to obtain an estimate $\bar{\lambda}_1$.

The area in equation (5) has to be maximized, and this is achieved if Θ is minimized where

$$\Theta = \frac{N \cdot \bar{\lambda}_1}{\text{trace}(\underset{\sim}{\mathbf{L}})}. \tag{10}$$

The extra factor N in the numerator ensures that $\Theta \geq 1$, and $\Theta = 1$ *iff* all eigenvalues are equal.

The parameter Θ may be calculated efficiently for any matrix $\underset{\sim}{\mathbf{L}}$ and depends on model parameterization in a highly nonlinear manner. Hence, a genetic algorithm (or any other minimizing algorithm suitable for nonlinear problems) may be used to find a parameterization with reduced Θ. In general we will not find the global minimum value of Θ but *any* reduction in Θ improves the inverse problem conditioning.

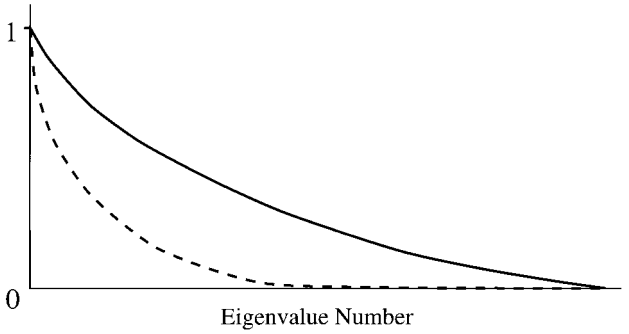


FIG. 2. Two possible normalized eigenvalue spectra in a many parameter inversion. The dashed curve contains many more near zero eigenvalues than the solid curve, hence the solid curve represents a better conditioned inverse problem.

A CROSS-BOREHOLE TOMOGRAPHIC EXAMPLE

We now illustrate the procedure by finding a triangular cell parameterization that best suits the raypath geometry shown in Figure 3a for a cross-borehole tomographic inversion. The exact depths of sources and receivers are unimportant, but they are irregularly spaced to reconstruct a situation where equipment failure or noise made data at certain regularly-spaced receivers unusable.

For irregular data geometries, we expect that the best-conditioned model parameterization will contain irregularly

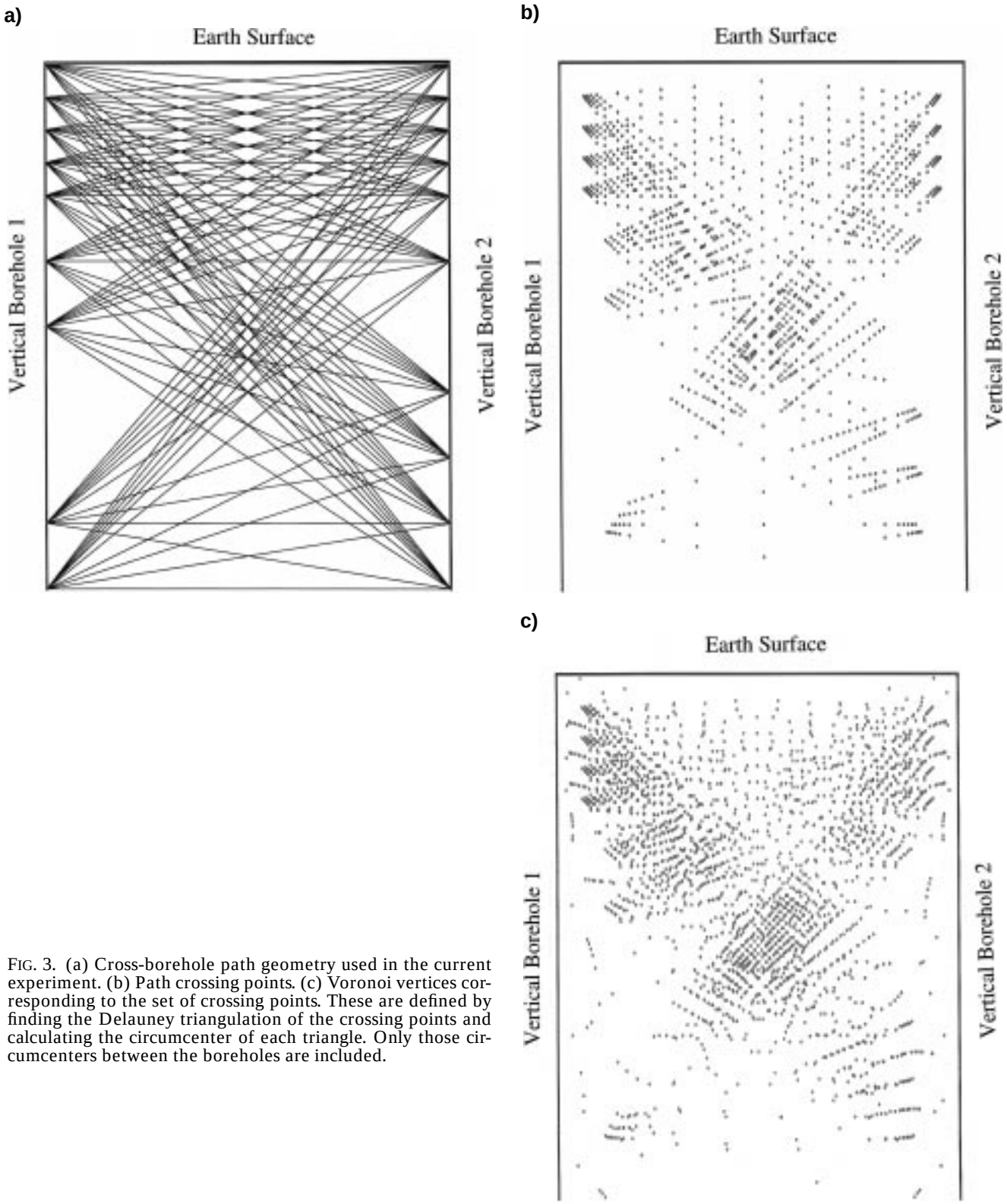


FIG. 3. (a) Cross-borehole path geometry used in the current experiment. (b) Path crossing points. (c) Voronoi vertices corresponding to the set of crossing points. These are defined by finding the Delauney triangulation of the crossing points and calculating the circumcenter of each triangle. Only those circumcenters between the boreholes are included.

sized cells, so we use Delauney triangular cells based on any irregular set of nodes (Delauney, 1934; Sibson, 1980; Sambridge et al., 1995). Using this method, pairs of nodes are joined by line elements to form triangles only if that would result in the smallest cumulative aspect ratio (the difference between the maximum and minimum internal angles of each triangle, summed over all triangles formed) (Fortune, 1992). Thus triangles are formed from (natural) neighboring nodes only, and an example is given in Figure 4a. The Figure shows a triangular cell geometry constructed automatically from a highly irregularly spaced set of nodes, and illustrates how triangle areas change based on the nodal density. Such triangulations are performed efficiently using standard algorithms (see references above).

Although homogeneous ray coverage in each cell is an insufficient condition for model design, it is a necessary criterion for a well-conditioned inverse system. Hence, it makes sense to select nodes to be triangulated based on the raypath density. Such a distribution of nodes is provided by the raypath crossing points shown in Figure 3b. However, triangulating the crossing point set is problematic since a triangle formed by three neighboring nodes contains no raypaths within its interior (Figure 4b). We construct a new set of points as follows:

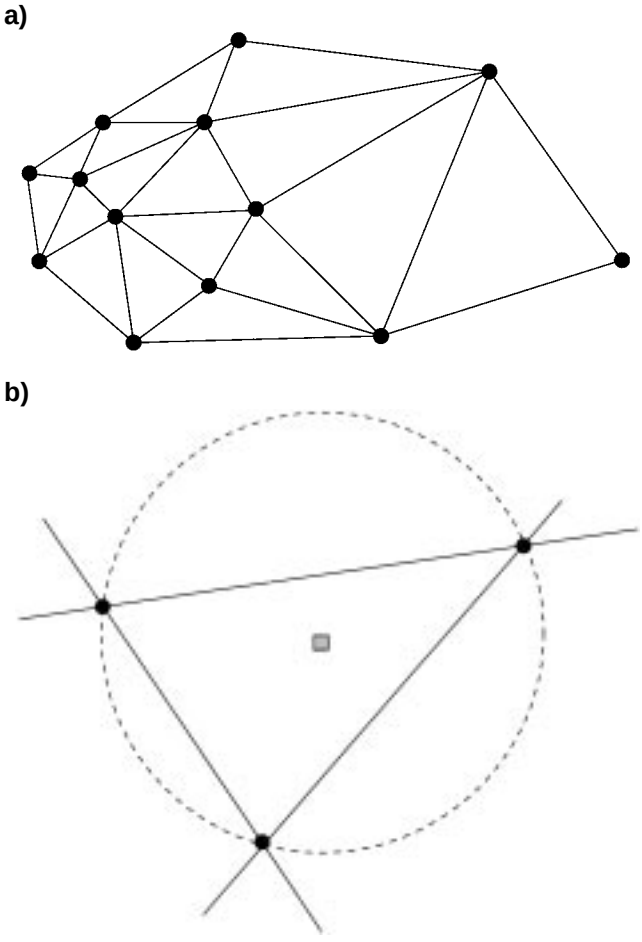


FIG. 4. (a) An example of Delauney triangulation of an irregular set of nodes (circles). (b) By definition, paths lie along the edges of any triangle formed from three neighboring nodes. To ensure that paths lie within cells the circumcenters of each triangle are used as cell vertices (square).

first, we triangulate the entire set of raypath crossing points. Then we calculate the circumcenter of each triangle and use those that lie between the two boreholes as our initial set of nodes (Figures 4b and 3c). Notice first that the raypath density criterion is approximately preserved for the node set. Second, in regions of approximately homogeneous path coverage (e.g., top center of the diagram) the node coverage is more homogeneous than in Figure 3b. Finally, in regions with no crossing points but where there may still be information there are now sparsely distributed nodes (e.g., in the neighborhood of each source/receiver node).

Triangulating all of these nodes would give a highly underdetermined model so we use the genetic algorithm (GA) to select a subset that gives the best-conditioned triangular cell parameterization based on the measure Θ in equation (10) [see Holland (1975) or Sambridge and Drijkoningen (1992) for a description of GA's]. In this application, a digital GA was used. That is, model nodes were stored digitally rather than in the more usual binary format. This makes the GA operation more intuitive: the "crossover" operation then just involves swapping some nodes between certain pairs of models in the model population, and "mutation" involves replacing an existing model with a new model comprising nodes selected randomly from the node space. The crossover and mutation steps were performed between each consecutive generation.

We wish to find a single, well-conditioned model (instead of a population of such models). To do this efficiently, the genetic algorithm was tuned for fast convergence and only the best models found were stored. After some experimentation, population sizes of 100 were used with probabilities of crossover equal to 0.7. Two different model sizes S (the number of nodes chosen) and corresponding probabilities of mutation P are illustrated here, $S = 10$ with $P = 0.2$, and $S = 200$ with $P = 0.05$. The improvement in conditioning of final models over initial models found in these runs was typical of that found in most experiments. The relatively high probability of mutation in the latter case ($10/S$, as opposed to $2/S$ in the former case) maintains a high diversity within each generation; mutating a point in the $S = 10$ model (with large cells) causes a much greater change in overall model structure than in the $S = 200$ case. For all node subsets selected, the source/receiver nodes were added since these define the convex hull of the model space (i.e., a set of nodes that form vertices of a convex polygon that contains the entire model space). In fact the number of source/receiver nodes to be included is arbitrary as long as the four corner nodes are included, but this does not affect the main results shown here.

RESULTS

Figure 5 shows result from the genetic algorithm search with $S = 10$. Plot 5a shows the first model constructed from nodes selected entirely at random from the initial node set. The condition measure Θ is 14.4, a value typical of most models in the first generation and which corresponds to a poorly conditioned eigenvalue spectrum as shown at the top of the figure. Many eigenvalues are close to zero corresponding to poorly constrained eigenvectors. Figure 5b shows the best model from the third generation—the 371st model tested. The condition measure of 5.8 has been reduced to 2/5 of its initial value corresponding to a far better conditioned system since the majority

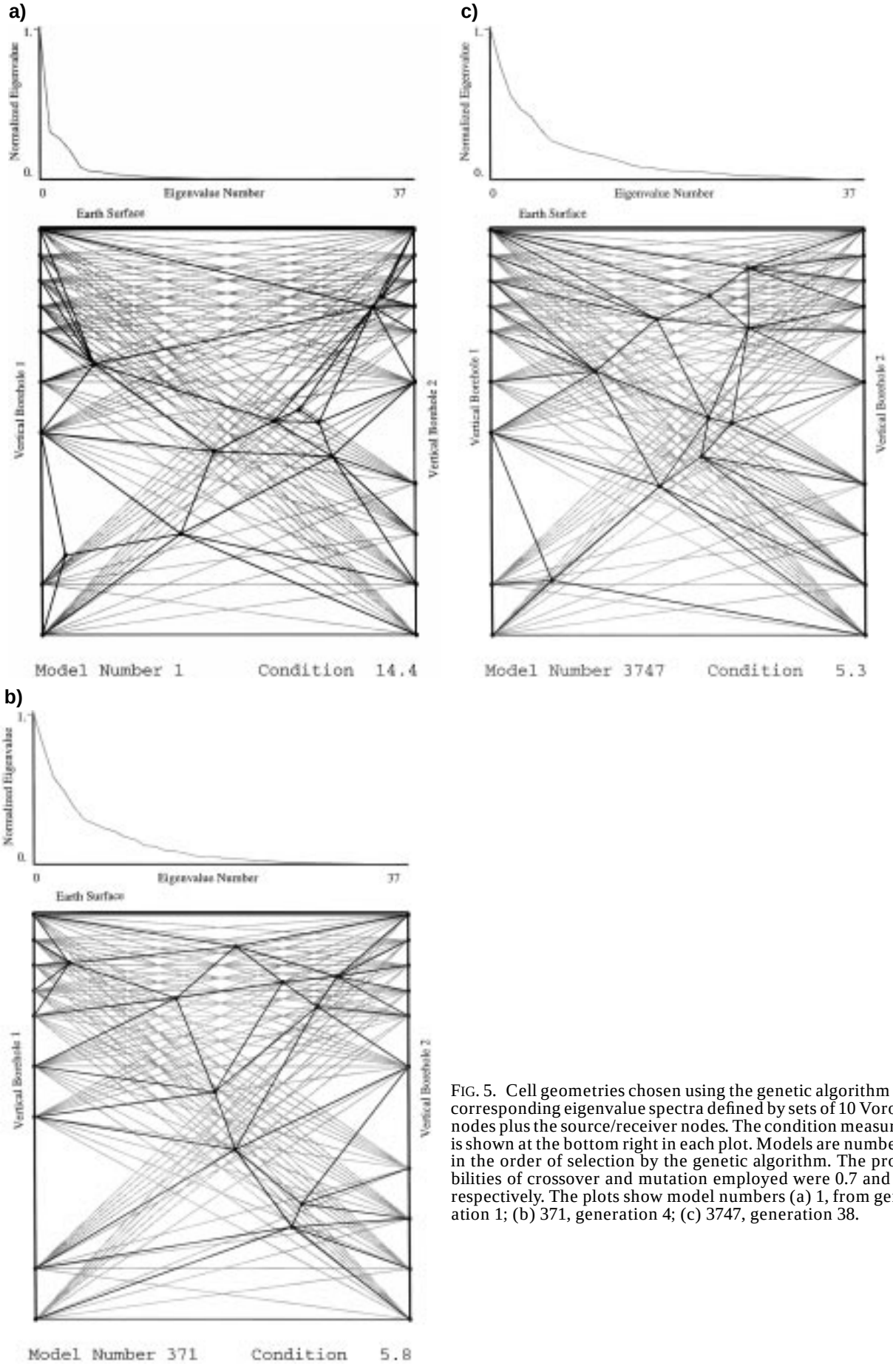


FIG. 5. Cell geometries chosen using the genetic algorithm and corresponding eigenvalue spectra defined by sets of 10 Voronoi nodes plus the source/receiver nodes. The condition measure Θ is shown at the bottom right in each plot. Models are numbered in the order of selection by the genetic algorithm. The probabilities of crossover and mutation employed were 0.7 and 0.2, respectively. The plots show model numbers (a) 1, from generation 1; (b) 371, generation 4; (c) 3747, generation 38.

of eigenvalues are visibly greater than zero. The best model from the first 8000 tested is shown in Figure 5c. The condition measure of 5.3 is almost 1/3 of its initial value, and almost all eigenvalues are significantly greater than zero. The conditioning in this system is such that we could almost invert for average model cell velocities with no prior smoothing or damping at all.

Figure 6 shows the model resulting from the genetic algorithm search with $S = 200$, while Figure 7 shows the corresponding eigenvalue spectra. Plot 6a shows the model constructed from the initial random selection of 200 nodes. Again the conditioning measure of 85.9 was typical of most models in the initial population and the eigenvalue spectrum in Figure 7 decays very sharply. However, as shown in Figure 6b, after only seven random attempts we were lucky enough to select a model that halved the conditioning measure and gave significantly less rapid decay in the eigenvalue spectrum. Notice first that both models have cell size distributions that approximately reflect the path density as expected. Also, both models have some extremely small triangles that cannot be resolved by the given ray geometry. These occur since the measure Θ does not penalize zero eigenvalues explicitly. When inverting for large numbers of cells, we will always require some regularization so these cells will not cause problems. (We could customize the measure Θ so that small cells are explicitly penalized, but this would not affect the main result shown here). Finally, from the overall cell geometry, there is no obvious reason why model 7 should be better conditioned than model 1; indeed, model 7 contains many more large cells that are almost entirely untraversed by rays than does model 1 (close to the left and right boreholes). Hence, choosing a model by eye might lead us to select the worst-conditioned model, number 1! Yet again this illustrates the inadequacy of data coverage as a criterion for model construction. The best model from the third generation is shown in Figure 6c, model number 379, with eigenvalues shown in Figure 7. Again, by the third generation we find a model with a conditioning measure 2/5 of its initial value. Finally, Figure 6d shows the best model found in 16 000 model evaluations, model number 8996 from the 90th generation, with a condition measure of 28.8, which is almost 1/3 of its initial value. There is still at least one large cell that is completely untraversed by rays and there is still no obvious criterion by which the naked eye could distinguish the improvement of models 8996 over model number 1. The improvement in the eigenvalue spectra in Figure 7, however, is clearly illustrates the power of the conditioning diagnostic Θ.

For the latter run with $S = 200$, model number 379 from the fourth generation (Figure 6c) was found in under 1 hour CPU time on a SUN Sparc 20 using nonoptimized code. The final model, model 8996 from the 90th generation (which only slightly further refined the conditioning) took 24 hours CPU time on the same machine.

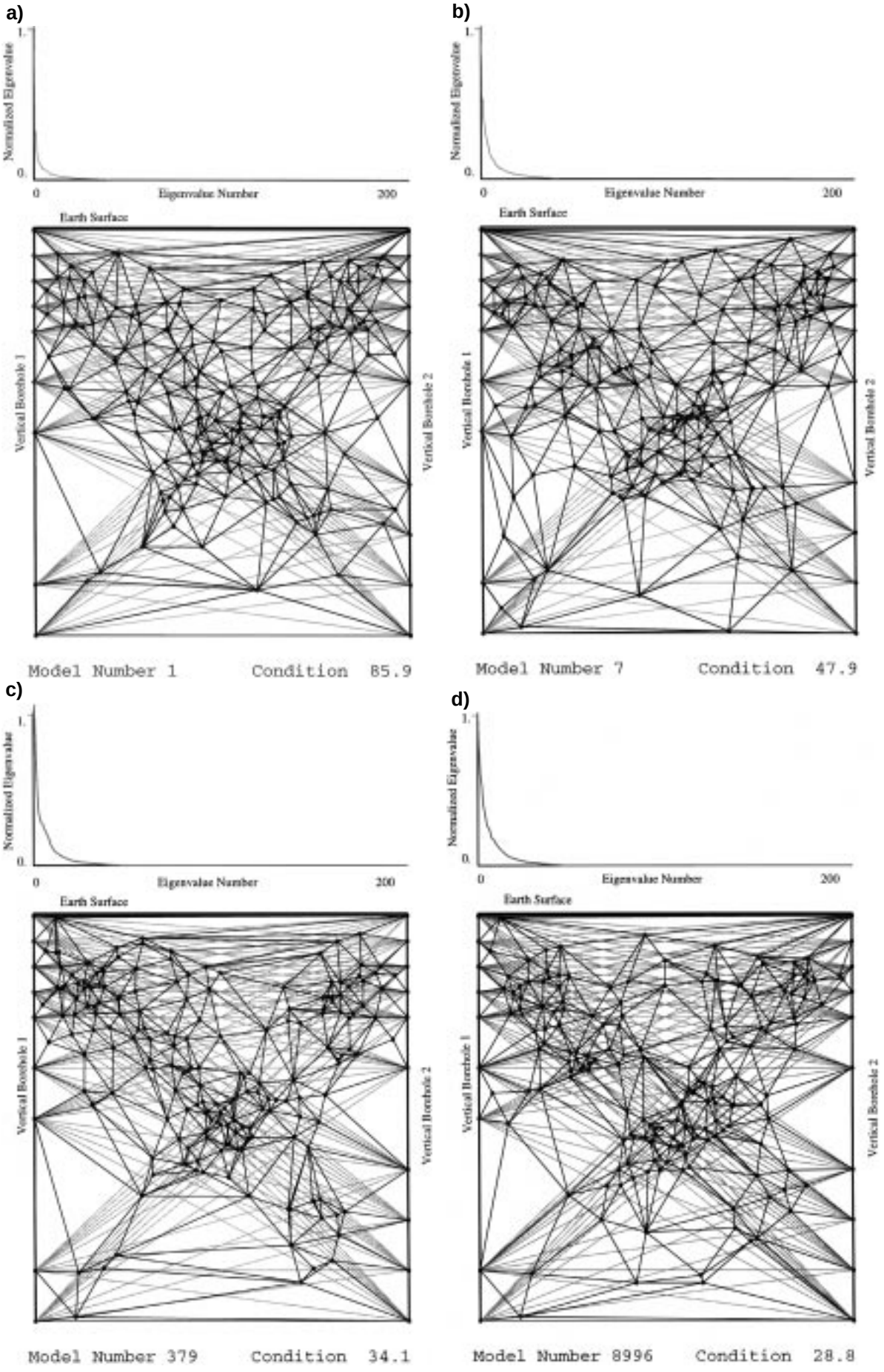
CONCLUSIONS

We have shown that data density, although necessary, may be an entirely insufficient criterion on which to select a model parameterization. As an alternative we have presented a measure Θ of inverse problem conditioning in equation (10), and a method with which it may be calculated efficiently for any model parameterization [equations (6) to (9)]. It was shown how a genetic algorithm can be used to find a particular model parameterization that gives a better-conditioned inverse problem than one where the model parameterization is chosen based on data coverage alone, and yet which does not look noticeably "better" nor physically less reasonable to the naked eye. The procedure achieved almost a factor of three improvement in the conditioning of a cross-borehole tomographic example using Delauney triangular cell parameterization. The conditioning of the final model could have been improved by either including only a subset of the source/receiver nodes in each model or by selecting nodes for each model using an algorithm that did not allow untraversed cells.

This process is carried out with minimal effort on the part of the operators: for any linearized inversion it is always necessary to program a routine that calculates the design matrix $\mathbf{A}$ and matrix $\mathbf{L} = \mathbf{A}^T\mathbf{A}$ for any particular model parameterization. The only extra programming required is the construction of a short subroutine that calculates the measure Θ given by equation (10) for any $\mathbf{L}$ and to place these in a standard genetic algorithm code to perform the minimization. The extra subroutine and genetic algorithm code can be used virtually unaltered for all subsequent inversions in any data and model spaces. The method makes model parameterization an integral and important part of the inversion process and allows a final model to be produced that includes more independent pieces of information from the data. Hence, there is no reason why this method should not become standard practice in future inversions.

The method is currently being applied to the much larger tomographic inversion problem of constructing a cell model consistent with over 14 000 surface wave source–receiver phase velocity measurements across the Eurasian continent. In exploration geophysics it might be applied similarly to parameterize models constrained by geomagnetic or gravity surveys. However, the conditioning measure Θ is not restricted to model design. In a future publication (manuscript in preparation), we will show that instead of varying the model space, we might fix a desired model parameterization and use Θ to assess any source–receiver network design. The latter technique might be used to design cross-borehole experiments, geomagnetic surveys or regional seismograph networks by varying locations of receivers (and sources if appropriate) in the genetic algorithm stage above.

At present, survey design [or in one case, designing irregular model parameterization grids, Vesnaver (1996)] is often guided by forward modeling experiments that calculate resolution or error propagation for some set of source–receiver array geometries and Earth structures (e.g., Macnae, 1984; Taylor, 1989; Sasaki, 1992), by maximizing the determinant of $\mathbf{L}$ (the D-criterion, e.g., Rabinowitz and Steinberg, 1990; Steinberg et al., 1995) or by "rules of thumb" gained from past experience (e.g., Rosencrans, 1992). Of these, the D-criterion is most similar to our method since it could be automated using a genetic algorithm, and indeed could be used to invert for an optimally conditioned model parameterization as done in the current work, rather than for survey design. However, repeated calculation of the determinant of $\mathbf{L}$ becomes prohibitively computer intensive as the dimension of $\mathbf{L}$ increases; our tests show that using our measure Θ carries out an equivalent operation in a fraction of the time, and hence is much more efficient. Finally, model resolution improves with the rank p of $\mathbf{L}$, so Barth and Wunsch (1990) suggested choosing array configurations that maximize the smallest nonzero eigenvalue λ_p. However, this


a)
Normalized Eigenvalue
1.
0.
0
Eigenvalue Number
200
Earth Surface
Vertical Borehole 1
Vertical Borehole 2
Model Number 1 Condition 85.9
b)
Normalized Eigenvalue
1.
0.
0
Eigenvalue Number
200
Earth Surface
Vertical Borehole 1
Vertical Borehole 2
Model Number 7 Condition 47.9
c)
Normalized Eigenvalue
1.
0.
0
Eigenvalue Number
200
Earth Surface
Vertical Borehole 1
Vertical Borehole 2
Model Number 379 Condition 34.1
d)
Normalized Eigenvalue
1.
0.
0
Eigenvalue Number
200
Earth Surface
Vertical Borehole 1
Vertical Borehole 2
Model Number 8996 Condition 28.8

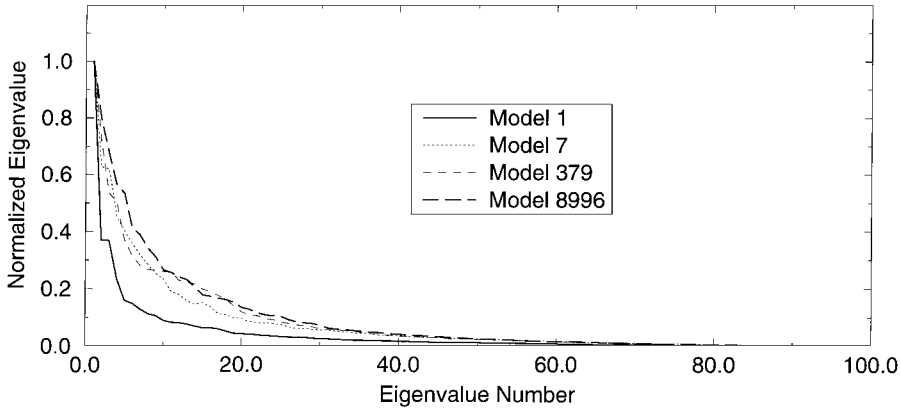


FIG. 7. Eigenvalue spectra corresponding to the models in Figures 6a to 6d.

process requires that we fix the desired rank p prior to array design. In the current method this is not necessary, and indeed we expect the rank to be maximized. Hence, the procedures outlined above enable array design to be ideally tailored to constrain any desired model parameterization (or vice versa) in a formal, objective manner that makes more independent information available to constrain the model space.

ACKNOWLEDGMENTS

We appreciated many helpful discussions with Malcolm Sambridge and Anthony Lomax. Thanks are extended to Kris Vasudevan, Subhashis Mallik, and Klaus Mosegaard for thoughtful reviews, and to Jonathan Shewchuk for allowing us to customize his 2-D Delauney triangulation routine. This research was supported by the Netherlands Organization for Scientific Research (NWO) through the Pionier project PGS 76-144. This is Geodynamics Research Institute (Utrecht University) publication 97.008.

REFERENCES

Barth, N., and Wunsch, C., 1990, Oceanographic experiment design by simulated annealing: J. Phys. Oceanography, **20**, 1249–1263.
Bertero, M., and De Mol, C., and Pike, E. R., 1985, Linear inverse problems with discrete data I: General formulation and singular systems analysis: Inverse Problems, **1**, 301–330.
Delaunay, B. N., 1934, Sur la sphere vide: Bull. Acad. Science USSR VII: Class. Sci. Math., 793-800.
Fortune, S., 1992, Voronoi diagrams and Delauney triangulations, *in* Du, D. Z., and Hwang, F., Eds., Computing in Euclidean geometry: World Scientific.
Holland, J. H., 1975, Adaptation in natural and artificial systems: Univ. of Michigan Press.
Mcnae, J. C., 1984, Survey design for multicomponent electromagnetic systems: Geophysics, **49**, 265–273.
Menke, W., 1989, Geophysical data analysis: Discrete inverse theory (Revised edition): International Geophysics Series, **45**, Academic Press Inc., Harcourt Brace Jovanovich, 119–125.
Michelena, R. J., 1993, Singular-value decomposition for crosswell tomography: Geophysics, **58**, 1655–1661.
Rabinowitz, N., and Steinberg, D. M., 1990, Optimal configuration of a seismographic network: A statistical approach: Bell. Seis. Soc. Am., **80**, 187–196.
Rosencrans, R., 1992, Cost-effective 3-D seismic survey design: The Leading Edge, **11**, No. 3, 17–24.
Sambridge, M., and Drijkoningen, G., 1992, Genetic algorithms in seismic wave-form inversion: Geophys. J. Int., **109**, 323–342.
Sambridge, M., and Braun, J., and McQueen, H., 1995, Geophysical parameterization and interpolation of irregular data using natural neighbors: Geophys. J. Int., **122**, 837–857.
Sasaki, Y. 1992, Resolution of resistivity tomography inferred from numerical simulation: Geophys. Prosp., **40**, No. 4, 453–464.
Sibson, R., 1980, A vector identity for the Dirichlet tessellation: Math. Proc. Camp. Phil. Soc., **87**, 151–155.
Steinberg, D. M., and Rabinowitz, N., and Shimshoni, Y., and Mizrachi, D., 1995, Configuring a seismic network for optimal monitoring of fault lines and multiple sources: Bull. Seis. Soc. Am., **85**, 1847–1857.
Taylor, G. G., 1989, Seismic resolution and field design: Success and failure at Taber, Alberta, Canada: Geophysics, **54**, 1101–1113.
Vesnaver, A. L., 1996, Irregular grids in seismic tomography and minimum-time ray tracing: Geophys. J. Int., **126**, 147–165.

←

FIG. 6. Similar to Figure 5, cell geometries chosen using the genetic algorithm defined by sets of 200 Voronoi nodes plus the source–receiver nodes are shown. The probabilities of crossover and mutation employed were 0.7 and 0.05, respectively. The plots show model numbers (a) 1, from generation 1; (b) 7, generation 1; (c) 379, generation 4; (d) 8996, generation 90.