THE ASTROPHYSICAL JOURNAL, 475:843–852, 1997 February 1

THE SHAPES OF CROSS-CORRELATION INTERFEROMETERS

ERIC KETO
Lawrence Livermore National Laboratory, University of California, and Smithsonian Astrophysical Observatory,
60 Garden Street, Cambridge, MA 02138
Received 1995 October 16; accepted 1996 July 30

ABSTRACT

Cross-correlation imaging interferometers designed in the shape of a curve of constant width offer better sensitivity and imaging characteristics than other designs because they sample the Fourier space of the image better than other shapes, for example, "T's" or "Y's." In a cross-correlation interferometer each pair of antennas measures one Fourier component with a spatial wavenumber proportional to the separation of the pair. Placing the individual antennas of the interferometer along a curve of constant width, a curve that has the same diameter in all directions, guarantees that the spatial resolution of the instrument will be independent of direction because the measured Fourier components will have the same maximum spatial wavenumber in all directions. The most uniform sampling within this circular region in Fourier space will be created by the particular symmetric curve of constant width that has the lowest degree of rotational symmetry or fewest number of sides, which is the Reuleaux triangle. The constant width curve with the highest symmetry, the circle is the least satisfactory although still considerably better than T's or Y's. In all cases, the sampling can be further improved by perturbing the antenna locations slightly off a perfect curve to break down symmetries in the antenna pattern which cause symmetries and hence nonuniformities in the sampling pattern in Fourier space. Appropriate patterns of perturbations can be determined numerically. As a numerical problem, optimizing the sampling in Fourier space can be thought of as a generalization of the traveling salesman problem to a continuous two-dimensional space. Self-organizing neural networks which are effective in solving the traveling salesman problem are also effective in generating optimal interferometer shapes. The Smithsonian Astrophysical Observatory's Submillimeter Array, a cross-correlation imaging interferometer for astronomy, will be constructed with a design based on the Reuleaux triangle.

Subject headings: instrumentation: interferometers

1. INTRODUCTION

This paper reports on the results of research conducted for the Smithsonian Astrophysical Observatory's Submillimeter Array project on the shapes of interferometers. The Submillimeter Array (SMA) is a multiple element or cross-correlation interferometer under construction at the Mauna Kea Observatory in Hawaii for astronomical imaging in the submillimeter region of the electromagnetic spectrum. Cross-correlation interferometers such as the SMA measure the brightness of a source in discrete spatial frequencies proportional to the separations of all pairs of antennas in the array and with relative orientations given by the relative orientations of the respective antenna pairs. This set of pairwise separations, mathematically equivalent to the cross-correlation function of the antenna locations, is the spectral response of the interferometer and the Fourier transform of the beam pattern. Thus, the shape of the interferometer determines the response function of the instrument.

In radio astronomy, an observation may follow a source from rise to set across the sky. At each instant, the spectral response will be the cross-correlation function of the interferometer projected by the angle between the source and the optical axis. In addition, the response function will rotate in following the source because the rotation of the Earth will cause the orientation of the interferometer, fixed on the Earth, to rotate with respect to the source. At the end of the observation, the total spectral coverage will be the integral sum of the time-varying instantaneous spectral response which is the rotated and projected cross-correlation function. This technique, called Earth rotation synthesis, allows for significantly better coverage of the Fourier plane than could be achieved by an instantaneous observation alone. However, the rotation and projection functions will depend on the location of the source in the sky and the latitude of the interferometer. Thus the instrument response is a function not only of the shape of the interferometer but also the mode of observation and may be different for each observation.

This report addresses two questions: what is the best response function for a general-purpose imaging interferometer such as the SMA, and how to determine the shape that best approximates that required response taking into consideration that the instrument will be used most of the time in Earth rotation synthesis. These two questions are difficult for two very different reasons. The first question is quite broad: for any particular scientific goal, the optimal response is likely to be unique. The second question is difficult because the spectral response must be the cross-correlation of the antenna locations, and the cross-correlation function is not invertible. A simple consideration suggests that numerical solutions are also difficult. If we imagine the antenna locations to be restricted to a grid, then the number of possible configurations increases exponentially with the number of antennas. This exponentially explosive character has previously prohibited numerical solutions for arrays with more than a few antennas.

In this report we will suggest that considering the most basic imaging requirements, the best response is the one that provides the most complete sampling of the Fourier space of the image out to the limit of some best spatial resolution. We will then offer two heuristic solutions, one analytic and one numerical, both suggesting why arrays

shaped in the form of a curve of constant width, in particular the Reuleaux triangle, provide the best approximation to this response. We will show how to estimate the reduction of the signal-to-noise of an image made with other interferometer shapes. For cases where some other spectral response is dictated by different concerns, we will show how to use the numerical procedure to find approximations to these other response functions.

2. CHOICE OF RESPONSE FUNCTION

In optimizing an imaging system for a particular task, we would first ask what are the properties of the expected images and how do we match the instrument response to those properties given the goals of the experiment. For a general-purpose instrument, we might rephrase the question and ask, given no prior assumptions about the type of images expected, what are the most basic or common properties that the imaging system ought to have. While there are many such properties, selecting just a few, for example, resolution, signal-to-noise, and sampling accuracy, will be sufficient to specify the response and thus the shape of the interferometer.

Starting with resolution, in the absence of any prior information about the image or the goals of an observation, we will prefer the interferometer to have equal resolution in all directions. Because the resolution in any direction is determined by the radial distribution of points in the Fourier space of the response function, equal resolution in all directions implies that the response function should be circularly symmetric. In particular, because the maximum resolution is set by the maximum wavenumber, the response function should be contained within a circular boundary in Fourier space.

The following two arguments concerning the signal-to-noise ratio and the accuracy of the imaging will show that the Fourier plane within this boundary ought to be sampled uniformly. First, in the absence of any prior information about the image, the highest overall signal-to-noise and the highest resolution will be achieved simultaneously if the Fourier domain within the boundary is sampled uniformly. Uneven sampling will always result in either poorer spatial resolution or poorer signal-to-noise. For example, a "T" or "Y" shaped cross-correlation interferometer is characterized by a centrally condensed sampling distribution with more sampling points at low spatial frequencies. In such a case, if the image is formed with equal weight given to all the measured Fourier components (so-called natural weighting), the spatial resolution will be less than if all areas of the Fourier plane are given equal weight. The latter can be achieved by regridding the data to a uniform grid (uniform weighting). However, this necessarily implies assigning nonuniform weights to the data and a corresponding loss of signal-to-noise. The decrease in the signal-to-noise ratio owing to nonuniform weighting is analogous to the inefficiency incurred in averaging any set of independent measurements with other than equal weights. That is, in the absence of prior information, we can do no better than performing a straight or unweighted average of independent measurements.

For a cross-correlation interferometer, we can examine the effect of weighting on the signal-to-noise of the image by considering an oversimplified case. If the interferometer is observing a point source at the phase center of the image, then the signal-to-noise can be written as a simple analytic expression (Thompson, Moran, & Swenson 1986, p. 162),

$$\frac{S}{N} = \frac{s}{n}\sqrt{n_d}\,\frac{\sum w_i}{\sqrt{\sum w_i^2}},$$

where s/n is the signal-to-noise ratio measured at a single sampling point, n_d, the number of data points in the sample, and w_i, the weights to the data. This equation illustrates that the highest signal-to-noise is obtained with weights identically equal to unity. Although the measurement of the flux of a point source is not an imaging problem, this equation is a good illustration because the detrimental effect of nonuniform weighting is still of the same character even when considering more complex sources. For more complex sources, the summations are simply more lengthy and complex. Thus the highest resolution with the highest signal-to-noise achievable is obtained with uniform sampling in the Fourier plane. Using the specialized terminology of radio astronomy, the response function of interferometer should be such as to make natural weighting and uniform weighting the same.

Second, uniform sampling will provide images least susceptible to errors arising from unmeasured Fourier components. Because a cross-correlation interferometer samples a discrete number of the Fourier components, nonuniform sampling will imply less complete sampling. That is, there will be larger gaps in some places in the Fourier plane and higher sampling densities elsewhere. Incomplete sampling implies unmeasured Fourier components which will lead to errors in the image which are quite separate from errors arising from noise. Consider an oversimplified example involving a digitized image with a finite number of discrete pixels. In this simple case, complete sampling is defined. Given a choice, would we prefer more measurements on fewer Fourier components to achieve a high signal-to-noise on these measured components, or would we prefer to measure all the Fourier components, although with lower signal-to-noise, to completely sample the Fourier plane. The answer is of course the latter because an image formed with complete sampling will be accurate to within the measured signal-to-noise whereas an image formed with incomplete sampling may have arbitrarily large errors.

In practice, real imaging systems never achieve complete sampling. In radio astronomy, incomplete sampling is addressed with image enhancement algorithms, such as CLEAN (Clark 1980) or maximum entropy (Gull 1989), which perform the inverse Fourier transform from the measured visibilities to the image, attempting to infer information about missing spatial frequencies by a combination of interpolation from nearby measured frequencies and use of simple prior information about the image. (A straightforward inverse Fourier transform implicitly makes the assumption that the unmeasured components are zero.) The larger the gaps in the sampling pattern, the greater the possibility that the interpolation, based on uncertain prior assumptions, will not be correct, and the greater the possibility for error in the final image. This notion of minimal gaps in the sampling function was also suggested as a design criterion for the Very Long Baseline Array (VLBA) (Walker 1984).

These considerations of resolution, signal-to-noise, and sampling accuracy precede and are not fundamentally changed by post-observational error correction or deconvolution procedures. For example, an algorithm such as

self-calibration (Schwab 1980) which attempts to reduce phase errors introduced by atmospheric fluctuations by enforcing phase closure may improve an image by reducing a particular contribution of noise, but this will not affect the resolution, the weighting of the Fourier components, nor the sampling. However, goals different from general imaging may require a different response function. We will discuss ways of designing interferometers for different response functions in § 6.

3. AN ANALYTIC APPRECIATION

The discussion of the previous section suggests that we look for interferometer shapes whose response functions are a uniform distribution of sampling points within a circular boundary. As mentioned in the Introduction, astronomical cross-correlation interferometers are almost always used in Earth rotation synthesis where the response depends on the duration of the observation and the locations of the source and interferometer. However, it is instructive to first consider the case where the problem can be exactly defined, where the spectral response is simply the cross-correlation of the antenna locations. This corresponds to instantaneous imaging, with the interferometer and the image plane parallel and their centers aligned. We will return to Earth rotation synthesis in § 5.

The criterion that the interferometer have the same resolution in all directions leads directly to a design around a curve of constant width by the following argument. For the case we are interested in, where the distribution of sampling points will be uniform inside a circular boundary, the resolution will be set by the radius, or wavenumber of the boundary. Placing the antennas along a curve of constant width ensures that the maximum separation of the antennas, and thus the resolution, will be the same in all directions. This defining property of a curve of constant width may be visualized by imaging such a curve placed between two parallel planes and rolled. The curve has constant width if the planes do not bump up and down. The circle is the simplest curve of constant width, but there are an infinite number of constant width curves, symmetrical and non symmetrical. Any regular polygon with an odd number of sides can be made into a curve of constant width by replacing each of the sides with a circular arc with the opposite vertex as the center. Of all the curves of constant width, the circle is the one with the highest degree of rotational symmetry, and of the curves formed from regular polygons, the curve formed from the equilateral triangle has lowest. Construction procedures and other interesting facts about these curves may be found in Gardner (1968) and Rademacher & Toeplitz (1957). Any of these shapes will create an interferometer with equal resolution in all directions. Thus the requirement for equal resolution in all directions defines the boundary of the spectral response function and specifies the family of curves of constant width as the correct shape. Interferometers built on a circular plan include the Culgoora Radioheliograph (Swenson & Mathur 1967).

The requirement that the sampling function be uniform within the previously defined circular boundary is sufficient to select the symmetric triangular curve of constant width, the Reuleaux triangle, as the best out of the infinite family of constant width curves by the following argument. Consider the sampling pattern produced by an interferometer with individual antennas evenly spaced around the perimeter of a circle. In this case, the antenna separations in real space will be evenly spaced in angle so that the samples in Fourier space must lie on radial spokes (Fig. 1). For any single angle, the separations in that direction, which are parallel chords across the circle, will be distributed in length as the cosine of the angle subtended by the chord from the center of the circle. At the shortest separations, the change in separation is rapid as is the change in cosine of a small angle. But at the largest separations, there is very little change in distance between pairs. The sampling points at the shortest spacings are then relatively far apart radially but tightly bunched in the perpendicular direction, in azimuth. At the longest spacings, the sampling points are very similar in radius, but far apart in azimuth around the perimeter of the outer boundary in Fourier space. In the midrange the sampling approximates a uniform grid (Fig. 1).

If the antennas are distributed uniformly around the perimeter of a Reuleaux triangle (Fig. 2), the spacings will be differently distributed in length and orientation. For example, consider the shortest spacings. In both the circle and the triangle, these are the spacings produced by the separations between adjacent pairs, every other pair, and so on. In the case of the circle, the smallest spacings are the same at all orientations as are the next smallest and so on. In the case of the Reuleaux triangle, the length of the spacings connecting antennas along the sides of the triangle is different than the lengths connecting antennas across the vertices. Thus the lengths change with different orientations. This difference breaks down some of the symmetry in the circular array resulting in a more uniform distribution (Fig. 2).

This notion of achieving more uniform sampling by reducing the symmetry of the sampling pattern makes sense

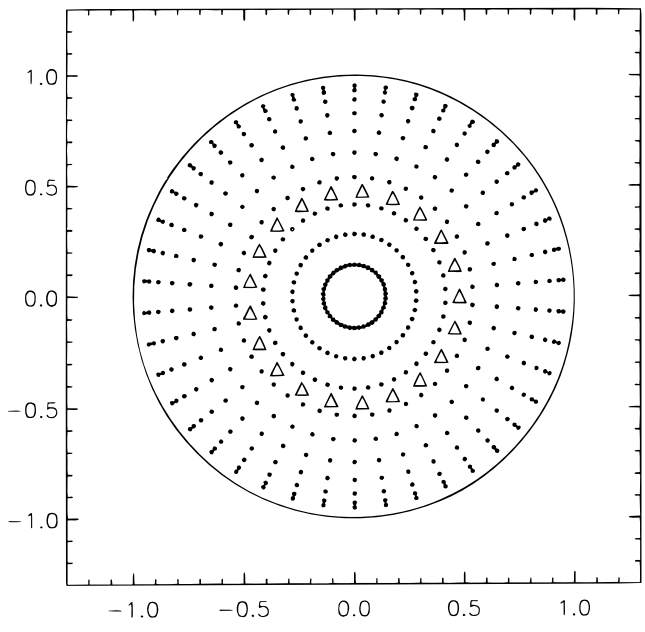


FIG. 1.—Sampling pattern for a circular array of 21 antennas. Triangles indicate the antenna locations and the dots indicate the separations of all pairs of antennas. These separations are proportional to the wavenumber of the sampling in Fourier space and are also equivalent to the cross-correlation function of the antenna locations. Thus the axes indicate physical space for both the antennas and their separations and are proportional to wavenumber in Fourier space. The circular array produces a sampling pattern that is too tightly packed in radius at large wavenumbers and too tight in azimuth at small. In between the sampling is fairly uniform.

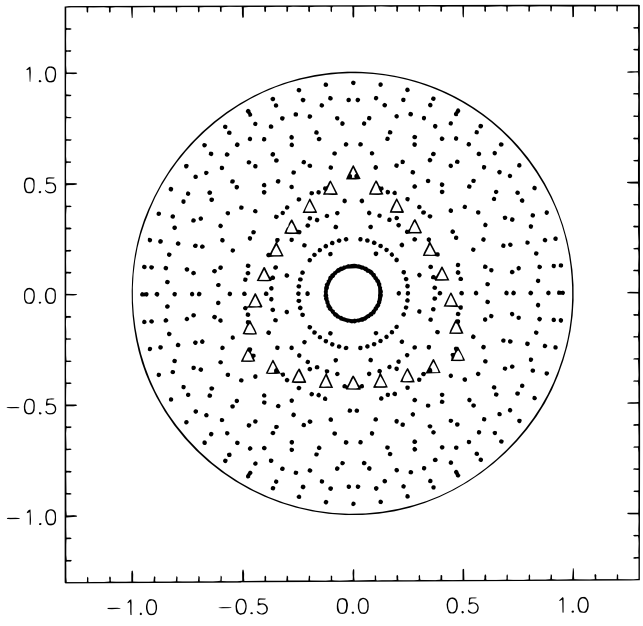


FIG. 2.—Sampling pattern for a symmetric array of 24 antennas in a Reuleaux triangle. Differences in separations between antennas along the sides of the triangle and across the vertices create a less symmetric and more uniform sampling pattern in Fourier space. See Fig. 1 for an explanation of the triangles and dots.

if we think of uniform sampling in the sense of an average equal density of the sampling points across the Fourier domain. Apart from the symmetry represented by a regular, periodic, tessalation of the plane—which is apparently not achievable as a cross-correlation function except over a limited subset of the sampling domain (Golay 1976)—other types of symmetries will in general reduce unformity. Therefore, it should be possible to improve the sampling of designs based on curves of constant width, by perturbing the symmetry of the antenna locations off of a perfect curve. Cornwell (1988) explored this notion of asymmetry in the design of circular arrays with small numbers of antennas. He used a numerical algorithm to place the antennas around the perimeter of a circle to achieve the most uniform sampling within a circular domain in Fourier space resulting in a quasi-random looking distribution around the circle. Continuing along this line of thought, we might suppose that the sampling could be further improved by perturbing the antenna locations in two dimensions, radius and azimuth, rather than just azimuth. A comparison of two designs for nine antennas, one generated by a computer algorithm that also allows variation in radius, shows that this is the case (Fig. 3).

This analysis suggests that to achieve the most uniform sampling within a circular boundary the antennas should be placed in a slightly perturbed pattern along the perimeter of a Reuleaux triangle. The perturbations will have a quasi-random character in that while there will be an infinite number of perturbed patterns with equally good sampling, random perturbations will not always produce good results especially for small numbers of antennas. Thus we need a procedure for determining a good pattern of perturbations. For this we will employ a numerical algorithm described in the next section.

4. A NUMERICAL APPROACH

While we know the forward mapping from a shape to the sampling pattern, we do not know the inverse mapping. But certainly, we should be able to find the shape which best approximates a desired sampling pattern with a numerical optimization program. Heuristic search algorithms such as simulated annealing, genetic algorithms, or neural networks are indicated by the exponentially explosive character of the

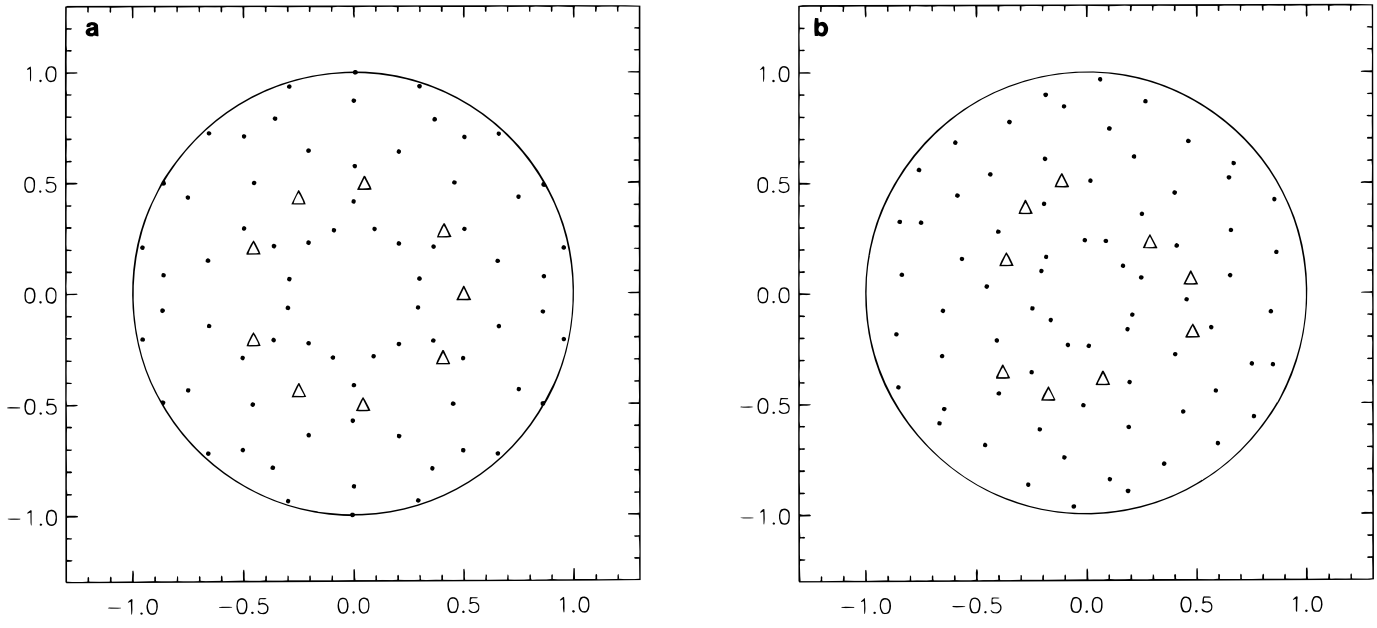


FIG. 3.—(*a*) Circular array of nine antennas where the antenna locations are confined to the perimeter of the circle, but optimized in one dimension, azimuth, to produce the most uniform sampling pattern within a circular region in Fourier space. (*b*) Triangular array of nine antennas where the antenna locations are optimized in two dimensions to produce the most uniform sampling within a circular region in Fourier space. Relative to Fig. 3*a*, perturbations in radius in addition to azimuth allow further breakdown of symmetry, and hence better uniformity in the sampling pattern. In particular the central hole is smaller as are the empty spaces between sampling points throughout the region.

problem. While all of these algorithms will work, for example, Cornwell (1988) used simulated annealing while Golay (1970) used a random guessing algorithm, the most efficient in this application is the neural or elastic net. Efficiency is key because not only does the search space expand exponentially with the number of antennas, but calculation of the cross-correlation function is itself an n^2 operation. This adverse scaling has previously prevented solution for arrays with more than a few antennas and even then on a limited solution space. For example, the solutions of Golay (1970) are restricted to a fixed grid while those of Cornwell (1988) are confined to the perimeter of a circle.

One might suppose that the difficulty with simulated annealing or stochastic search is that after some initial adjustments, it becomes progressively more difficult to blindly guess an adjustment which results in an improvement, and most of the trials, calculated at a cost of n^2 operations, are simply discarded. In contrast, gradient descent algorithms which calculate the direction of best improvement based on the current solution make better use of the available information and converge much more rapidly. However, as is well known, they are prone to trapping in local minima whereas a random guessing algorithm such as simulated annealing will eventually escape local minima to find the best solution at the global minimum. These considerations suggest that to improve on the performance of simulated annealing we might experiment with algorithms with a different mix of randomness and gradient descent yet which are still effective in avoiding local minima. Neural or elastic nets are suggested by their successful performance on the well-known traveling salesman problem which like our array design problem poses the difficulty of an exponentially explosive search space plagued by numerous local minima.

The traveling salesman's problem is to choose the shortest closed tour through a number of randomly distributed cities. To solve this problem by simulated annealing, we start from an initial guess and randomly swap pairs of cities in the order of tour. In the course of random guessing, we always accept solutions which shorten the tour, but we also accept solutions which increase the tour length according to a Boltzmann probability $e^{-E/kT}$, where E is the figure of merit, analogous to an energy, and kT can be thought of as a proportionality constant times a temperature which decreases slowly over time allowing the solution to converge. This process suggests crystallization by slow cooling or annealing. The occasional uphill steps in the algorithm allow the algorithm to escape local minima. In fact, it can be shown that if the temperature is cooled slowly enough, the algorithm will asymptotically converge to the correct solution (Geman & Geman 1984). However, in practice, the proof of asymptotic convergence is but meager consolation since there is no guarantee that any solution deemed sufficiently acceptable by whatever merit will be found in any finite period of time.

To solve the traveling salesman's problem using an elastic net, we imagine the tour to be represented by an elastic band. Initially the elastic band is placed on the plane containing the cities as a small closed loop which need not go through any cities. We then pick cities at random, find the closest spot on the elastic band, and stretch the band a small amount in the direction of the city. Over time the band will stretch to become closer and closer to a valid and short tour. Previous research shows that the elastic net is competitive with simulated annealing in the traveling salesman problem (Durbin & Wilshaw 1987).

To proceed toward an algorithm capable of solving the array design problem, we generalize the travel salesman problem to a two-dimensional continuous space. Instead of requiring the band to pass through discrete cities we require the band to cover as best as possible some two-dimensional region of arbitrary shape. In this case we discretize the elastic band so that it is made up of a fixed number of nodes connected by elastic. Following the same iterative procedure except randomly picking points on the plane rather than cities, we find the closest nodes on the band and move them closer to the randomly selected point. Over time, the nodes will distribute themselves over the plane so that no part of the plane is too far from one of the nodes. The result is a minimax solution; the algorithm minimizes the average maximum distance between the nodes and all points on the plane (see § 6).

Continuing the abstraction, we may imagine instead of a band, a two-dimensional elastic sheet or net of nodes connected by elastic bands. In this case our iterative procedure tugs at the sheet, stretching it into shape like pulling pizza dough to cover a pan. In the case of a two-dimensional sheet covering a two-dimensional surface, the problem is still interesting, that is to say nontrivial, if the sheet has a different topology than the shape it is to cover. For example, the sheet may be topologically simply connected, that is it has no holes, while the region to be covered is a torus. The elastic net algorithm in these topological mapping examples has been studied as a so-called self-organizing neural network (Kohonen 1984).

Finally, we may relate these algorithms to interferometer design by a further abstraction, imagining that our elastic net has the topology of the cross-correlation function of the antennas. If we index the n antennas, the net can be specified as an $n \times n - 1$ grid of nodes labeled by two antennas. The elastic may be strung between neighboring nodes defined as those having antenna indices differing by one. Each node then has four nearest neighbors, i, j connected to $i + 1, j$, $i - 1, j$, etc. Starting from an initial configuration of the antennas, we calculate the shape of the net as the cross-correlation function of the antenna locations. Then we randomly choose points in the circular region in Fourier space, stretching and pulling the net. In this case stretching the net means repositioning the respective antennas associated with the node in motion. Of course, it is not possible to move only a single node since each antenna is related to $n - 1$ nodes. But nevertheless averaging the required movements, over time, the net or cross-correlation function, will adapt to the specified topology, which is the required spectral response of the interferometer.

More specifically, the algorithm used in this study consists of two steps, a search step and a relaxation step. In the search step we find the point of the cross-correlation function $m_c(t)$ closest to a given random point $x(t)$ in the Fourier domain

$$\| x(t) - m_c(t) \| = \min_i \| x(t) - m_i(t) \| .$$

In the relaxation step we adjust the positions of those points in the cross-correlation function within some neighborhood, $N_c(t)$ around $m_c(t)$ which may be larger than the nearest neighbors as specified above. The adjustment is made so that on average—allowing that we cannot adjust one cross-correlation point without moving $n - 1$ others—

the points in the neighborhood lie closer to the randomly selected point in Fourier space $x(t)$:

$$m_i(t + \Delta t) = m_i(t) + \alpha(t)[x(t) - m_i(t)] \quad \text{for } i \in N_c(t) ,$$

$$m_i(t + \Delta t) = m_i(t) \quad \text{otherwise} .$$

The function $\alpha(t)$ is a monotonically decreasing function which allows the algorithm to converge, but α might also depend on the neighborhood function, $N_c(t)$. The algorithm carries the appellation "neural network" because we may imagine the neighborhood function to be the output of a similarity matching processor or neuron which is used adaptively in relaxing the synaptic weights of the neurons (the positions of our cross-correlation points). In the language of neural nets, the algorithm is said to have learned the required inverse mapping, albeit for a particular case, in that the information becomes encoded in the neural network in the synaptic weights connecting each neuron (Kohonen 1984).

We would expect, based on our analytic appreciation of the array design problem, that starting from any pattern of antennas, for example, random, the algorithm would always find a solution based on the Reuleaux triangle such as illustrated in Figures 3 or 4. Experience suggests that this is indeed the case. It is also possible to start from a circle or triangle and allow the algorithm to adjust the antenna locations a small amount simply to break down symmetries in the sampling pattern. Figure 5 is a circular pattern with much of the symmetry removed by the elastic net. However, in the case of nontriangular patterns the algorithm must be stopped at some arbitrary point, because allowed to converge, the solution is always triangular. Comparing Figures 4 and 5, the triangular solution has better sampling than the circular near the edge and center of the Fourier domain.

As opposed to the traveling salesman problem where simulated annealing and the elastic net are competitive, in this two-dimensional topological mapping problem, the elastic net is very much faster. For example, for 12 antennas, the difference on a RISC workstation is a few minutes versus a few hours. The concern that the algorithm is susceptible to trapping in local minima must be valid, but experience shows that, as with the traveling salesman problem, very good results are obtained very rapidly. Multiple trials with different initial configurations and with different randomly selected points in the Fourier plane will find different patterns of symmetry breaking perturbations on the Reuleaux design. But the various solutions are in various measures equally good. A particular solution cannot be guaranteed, but this is characteristic of heuristic search algorithms. For example, there is no guarantee that simulated annealing will not be trapped in a local minimum for a very long time, exceeding any allowed computational patience.

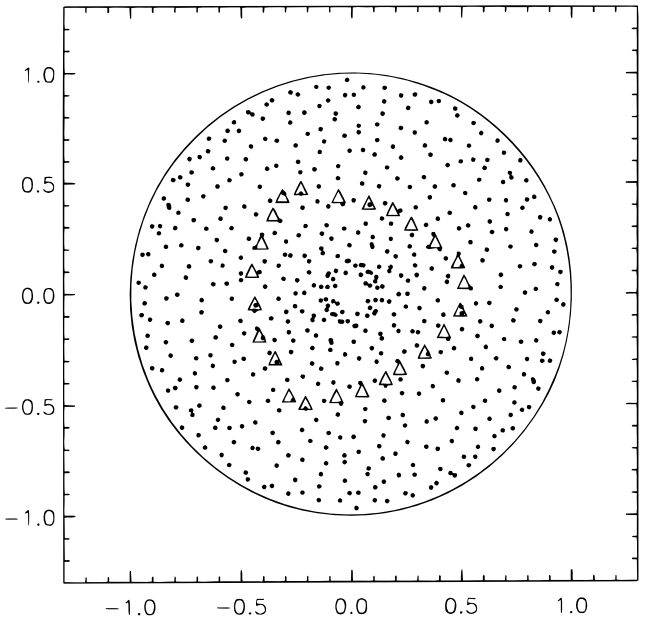


FIG. 4.—Triangular array of 24 antennas which has been optimized for the most uniform coverage by an elastic or neural net algorithm. Perturbations of the antenna locations away from a perfect Reuleaux triangle break down symmetry in the sampling pattern leading to more uniform coverage. Compare with Fig. 2.

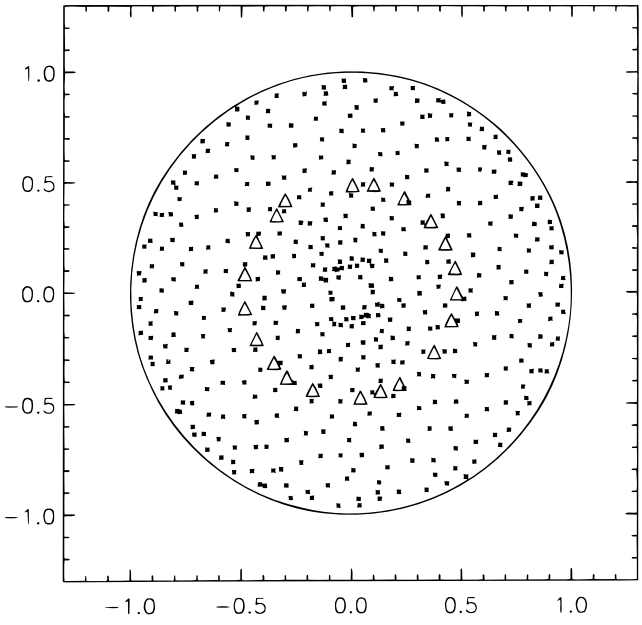


FIG. 5.—Circular array of 21 antennas partially optimized for the most uniform coverage by an elastic or neural net algorithm. In this case the solution was frozen before convergence to preserve the circular shape of the array. Allowed to run to convergence the solution would be triangular similar to the solution in Fig. 4. The perturbed circle here has much better coverage than the perfect circle in Fig. 1, but not as good as the triangular array in Fig. 4.

5. EARTH ROTATION SYNTHESIS

As mentioned in the Introduction, the response function of interferometers used in Earth rotation synthesis is not simply the cross-correlation function of the antenna locations, but rather the integral sum of a time-varying function which is the cross-correlation function rotated and projected by a constantly changing viewing angle as the interferometer tracks the source from rise to set across the sky. In Earth rotation synthesis an individual sampling point will trace out an elliptical arc in the Fourier plane over the course of an observation (Thompson et al. 1986, p. 86), and the interferometer design should consider the distribution of these arcs rather than the individual points in the cross-correlation function. This additional complexity does not change the original imaging requirements, and the choice of

a response function which provides for uniform sampling of the Fourier plane out to some maximum resolution is still the same. However, perhaps the biggest difficulty is that the optimization of the shape of the interferometer is no longer a well-defined problem because the same shape will generate different response functions depending on the particular observation, that is, the location of the source in the sky and the location of the interferometer on Earth.

Nonetheless, the problem when taken apart is not as intractable as it first appears. First, if we have a circularly symmetric response, the response will be invariant under rotation of the array, and we need consider only the effects of the projection of the response. A projection is a simple linear operation which will transform a uniform response within a circular boundary to a contracted uniform response within an elliptical boundary. During a tracking observation, the projection will be the least at transit, when the source passes the meridian, and greater on either side. Thus the instantaneous responses over the course of the observation will be a family of uniform distributions with elliptical boundaries. These ellipses, centered on the origin in the Fourier plane, will rotate and expand to a maximum area as the source rises to transit, rotating and decreasing as the source sets. Thus the final response function will be a sum of ellipses of varying sizes, all centered on the origin, and at the end of the observation we will always have a response which is more heavily weighted toward the center of Fourier plane. To compensate, we might ask for an array shape which provides a sampling function with lower density in the center so that this deficit in the center would be filled in by the smaller and denser projected sampling patterns early and late in the observation.

Such a request may be in vain in that there probably are no array patterns whose cross-correlation functions are centrally evacuated. One way to understand this statement is to think of the cross-correlation function as the overlap of the function with itself separated by some distance or lag in x and y. These lags are the independent variables in the two-dimensional cross-correlation. Suppose we have a function of value unity on a circle and zero elsewhere. Then the cross-correlation or overlap of this circle with itself has a value of two or zero almost everywhere, the exceptions being the points where the circles are tangent and the unique point of zero lag where the overlap is $2\pi r$. Now consider figures such as the "Y" or the "T." These have an overlap of unity or zero almost everywhere, but much higher values along certain directions where the arms of the two overlapping figures line up. Along these special directions, the overlap decreases linearly with increasing lag from the maximum at zero lag to the point where the overlapping figures separate. Along other directions, the lag falls immediately to unity. This is another way of understanding why the cross-correlation function of closed figures like the circle or Reuleaux triangle will be more uniform than the cross-correlation of open figures such as the "T" or "Y." Ignoring the point at zero lag, which is not relevant for radio interferometers built of discrete elements, it is difficult to imagine a two-dimensional figure whose overlap with itself decreases with smaller lag. Thus it will not be possible to have an instantaneous response function which is centrally evacuated. The best we can do in Earth rotation synthesis is start with a shape whose instantaneous response is flat so that over the course of the observation, although the response function will become denser in the center, the effect will be less so that if we had started with a centrally condensed instantaneous response.

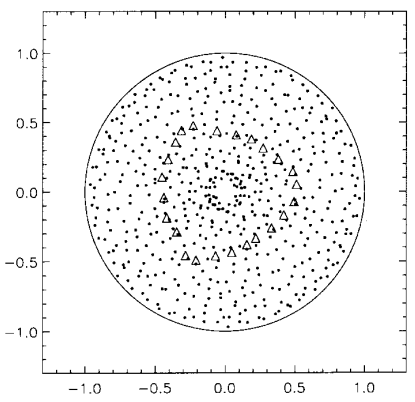
Figure 6 shows a comparison of the coverage obtained in Earth rotation synthesis with the 24-element perturbed Reuleaux triangle of Figure 4 and the 27-element "Y"-shaped Very Large Array. Although Earth rotation synthesis always improves the sampling over the instantaneous response, the deficiencies of the nonuniform instantaneous response generally remain as deficiencies in the time-integrated response. For example, the instantaneous response of a "Y" pattern has a higher density in the center and a boundary shaped like a six pointed star. After Earth rotation synthesis, the sampling of the "Y" is improved, but the outlying points of the star have simply become outlying tracks and the relative density in the center has increased further. Similarly, the Fourier coverage of the Reuleaux triangle is also much improved, and the response maintains its good properties of equal resolution in all directions and more nearly uniform coverage.

The dependence of the response function on the declination of the source and the duration of the observation means that to demonstrate a more formal proof, either numerical or deductive, we would have to specify the declination and duration of the observation in the definition of the problem. For example, we might choose the midpoint of the operating range of the instrument, or possibly the average response function over a range of declinations and durations about some midpoint. In this regard, our previous discussions of the response as the simple cross-correlation of the antenna locations is quite relevant because the midpoint of the operating range of an interferometer is generally the zenith. The array will likely observe as far north as south, and the observations are generally scheduled with equal time before and after transit. Thus the response function at zenith is arguably the correct single response to choose for the optimization of the response for Earth rotation synthesis.

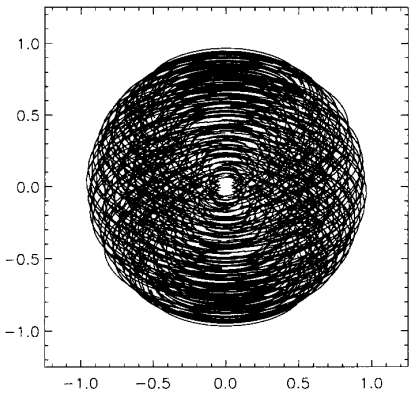
As this discussion illustrates, there are a number of ways to approach the problem. Yet all the ones tried lead back to the choice of the Reuleaux triangle. In summary, Figure 6 and the discussion above clearly demonstrate why the Reuleaux triangle will outperform "T's" and "Y's" in Earth rotation synthesis. The figure and the discussion also suggest very intuitively why there is probably no better shape for Earth rotation synthesis. However, a more definitive argument would require a more definitive statement of the problem, including the particulars of the operation of the array. Perhaps for a general imaging instrument, we have gone as far as we can.

6. THE TRADE-OFF BETWEEN SIDELOBES AND SIGNAL-TO-NOISE

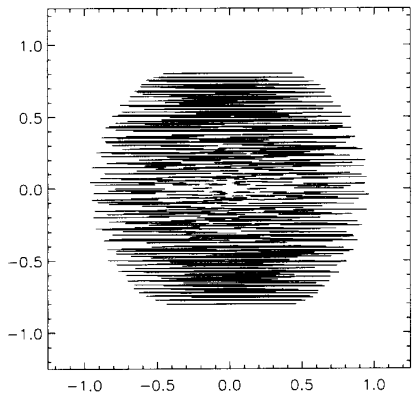
As mentioned in the discussion on image quality, in other circumstances we might wish to specify a nonuniform spectral sensitivity function. For example, in the design of single-dish antennas, there is a well-known trade-off between overall signal-to-noise and the sidelobe level of the antenna beam. In Fourier space, the maximum signal-to-noise at the highest resolution is achieved with a pillbox sampling distribution which will result in a beam shaped as the Bessel function $J_1(x)/x$ with sidelobes at the 19% level. In contrast, a truncated Gaussian distribution in Fourier space will result in an approximately Gaussian shaped beam. The main beam will be broader than that achieved by



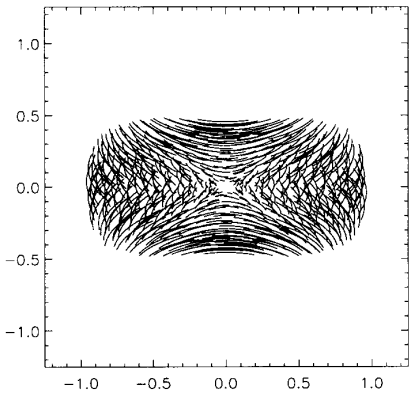
Snapshot at Zenith



$\delta = +30$

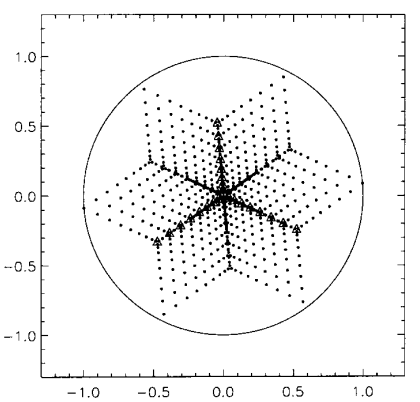


$\delta = 0$

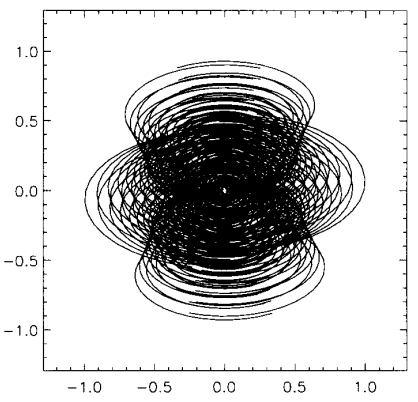


$\delta = -28$

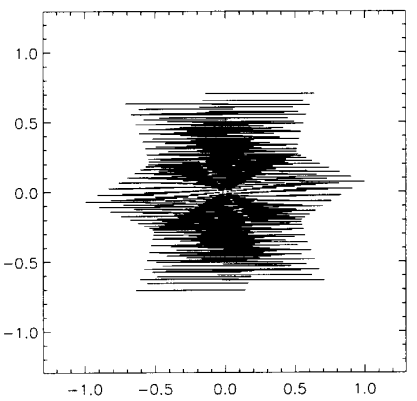
FIG. 6a



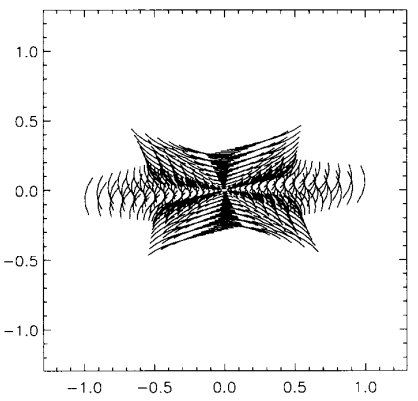
Snapshot at Zenith



$\delta = +30$



$\delta = 0$



$\delta = -28$

FIG. 6b

FIG. 6.—Comparison of the sampling pattern for a 24 antenna triangular array (*a*) and the 27 antenna VLA (*b*). The arrays are located at +34 latitude, the location of the VLA, and track sources between elevation limits of 25°, about 6 hr of tracking. Arrays built on triangular curves of constant width provide excellent sampling in Earth rotation synthesis as well as snapshot imaging. In particular, the circular boundary in Fourier space achieved by the triangular design is preserved in Earth rotation synthesis, providing equal angular resolution in all directions. In contrast, the outliers in the star-shaped sampling pattern of the "Y" remain outliers. Interior to the boundary, the sampling pattern of the triangular design is more uniform than the "Y," a characteristic which does not change under Earth rotation synthesis.

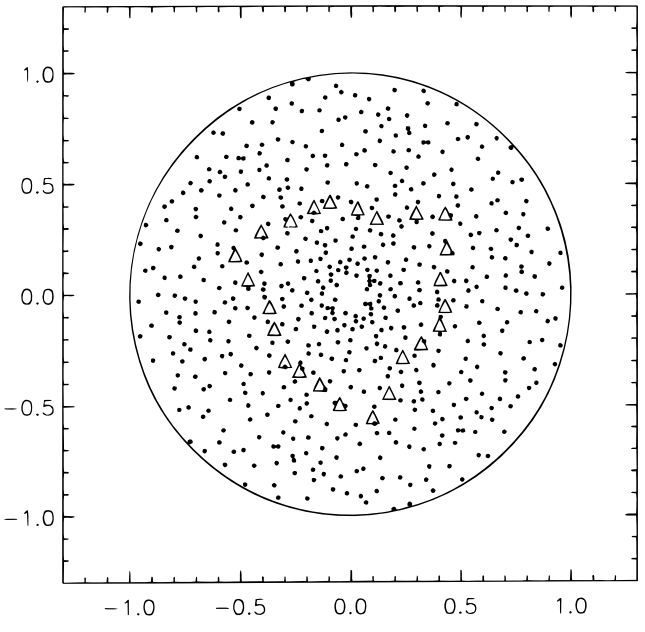


FIG. 7.—Array of 24 antennas which provides a tapered sampling function in Fourier space useful in applications requiring small sidelobes in the beam pattern. The array shape is optimized to provide the closest approximation to a Gaussian response within a circular boundary. Evidently, there are no patterns whose cross-correlation function closely meets these criteria, but the pattern here has a softer edge than, for example, the pattern in Fig. 4.

the pillbox sampling, implying a reduced resolution, but the sidelobes will be significantly less. For some applications, for example, to reduce ringing or concentrate the beam power, we might prefer an approximately Gaussian response. To achieve this, we might substitute a Gaussian distribution of randomly selected points in the Fourier plane in place of our uniform distribution to bias the stretching of the net. Previous research shows that this is nearly but not quite correct. The problem is analogous to early work on the optimal distribution of frequency channels in telephony where the task is to assign a limited number of channels to cover some bandwidth with the condition that the signals have a known but nonuniform probability distribution in frequency (Bennet 1948). In this case the result will be a weighted minimax solution. Extension to multiple dimensions has come to be known as vector quantization because the task may be stated as minimizing the average quantization error between the set of input signal vectors x, described by a probability distribution $p(x)$, and $m_{c(x)}$, the closest member in a set of measurement vectors,

$$\langle e \rangle = \int_{V_x} \| x - m_{c(x)} \| \, p(x) dV_x \, ,$$

where dV_x is a volume differential, and $\| \cdot \|$ is an arbitrary metric. Previous work on neural networks has shown that if the input signal vector x is a random variable with a probability density $p(x)$, then a self-organizing neural network will converge such that the distribution of the measurement vectors, in the asymptotic limit of a large number of such vectors, will be of the form $f[p(x)]$, where f is some continuous, monotonically increasing function (Kohonen 1984). The function f has been calculated analytically for the case where the neighborhood function N_c is restricted to the measurement vector m_c itself, that is, a zero width neighborhood, or equivalently if we let the self-organizing neural network converge with $N_c \to 0$ as $t \to \infty$. In the limit the elastic net algorithm is the same as the vector quantization algorithm proposed by Linde, Buzo, & Gray (1980), and we find that the measurement vectors will have a density $[p(x)]^q$, $q = n/(n + r)$, where n is the dimensionality of the measurement vectors and r is the power in the metric $\| f \| = [\int f^r dx]^{1/r}$ (Zador 1982). Figure 7 illustrates one array which approximates a Gaussian distribution in Fourier space with a circular boundary. The approximation is not particularly good presumably because there are no functions whose cross-correlation function is a Gaussian truncated with a circular boundary. However, because of the tapered sampling density this array would have a beam with a lower sidelobe level.

7. TRADE-OFFS TO SUBOPTIMAL DESIGN

How much better is the optimal configuration than another configuration? For example, the "Y" configuration offers certain advantages in construction requiring only half the path length to connect the antennas as required by the optimal triangle. Alternatively, if one is considering upgrading an existing "T"-shaped interferometer, would money be better spent on more antennas or on reconfiguring the antenna locations? Referring back to our discussion in § 2 we may ask how many antennas in a "Y" would be required to equal the image quality of the triangle. if the nonuniform cross-correlation function of the "Y" were regridded on the same scale as the cross-correlation function of the triangle. The best "Y" configuration for seven antennas has a point source signal-to-noise advantage including the penalty from regridding, which is 5% better than the six antenna triangle; however, on the grid scale of the triangular array, there are empty grid cells in the cross-correlation function of the "Y" which are covered by the cross-correlation function of the triangular array. The seven antenna "Y" is therefore a poorer design than the six antenna triangle since the 5% improvement in signal-to-noise is not worth the image degradation caused by the incomplete coverage. With eight antennas in a "Y," the poorer image quality owing to nonuniform coverage will be offset to some extent by a 24% improvement in signal-to-noise. A nine antenna "Y" performs better in all respects than the six antenna triangle, with an improvement in signal-to-noise of 46% and better overall coverage. In some rough sense we may say that six antennas in a triangle is worth approximately eight antennas in a "Y." The performance benefit or cost saving can be substantial if the antennas themselves are expensive, as is always the case with interferometers designed for astronomical observations at high frequency in the radio spectrum.

8. CONCLUSIONS

Interferometer shapes based on slightly perturbed curves of constant width, in particular the Reuleaux triangle, offer the most complete sampling in the Fourier space of the image. This holds not only in so-called snapshot imaging in which the spectral response is simply the cross-correlation function of the antenna locations, but is also generally true in Earth rotation synthesis in which the response is a time integration of the changing instantaneous responses obtained in tracking a source from rise to set across the sky.

Self-organizing neural networks are effective in generating the perturbed shapes.

The Smithsonian Astrophysical Observatory's Submillimeter Array, a cross-correlation interferometer for astronomical imaging at submillimeter wavelengths, will be built with a shape based on a triangular curve of constant width.

Work performed at Lawrence Livermore National Laboratory under the auspices of United States Department of Energy contract number W-7405-ENG-48.

REFERENCES

Bennet, W. R. 1948, Bell System Tech. J., 27, 446
Clark, B. G. 1980, A&A, 89, 377
Cornwell, T. J. 1988, IEEE Trans. Antennas Propagat., AP-36, 1165
Durbin, R., & Willshaw, D. 1987, Nature, 326, 689
Gardner, M. 1968, The Unexpected Hanging and Other Mathematical Diversions (Chicago: Univ. of Chicago Press)
Geman, S., & Geman, D. 1984, IEEE Trans. Pattern Analysis Machine Intelligence, PAMI-6, 721
Golay, M. J. E. 1970, J. Opt. Sci. Am., 61, 272
Gull, S. F. 1989, in Maximum Entropy and Bayesian Methods, ed. J. Skilling (Dordrecht: Kluwer), 53
Kohonen, T. 1984, Self-Organization and Associative Memory (New York: Springer)
Linde, Y., Buzo, A., & Gray, R. M. 1980, IEEE Trans. Comm., 28, 84
Rademacher, H., & Toeplitz, O. 1957, The Enjoyment of Mathematics (Princeton: Princeton Univ. Press)
Schwab, F. 1980, SPIE, 231, 18
Swenson, G. W., Jr., & Mathur, N. C. 1967, Proc. IREE Australia, 28, 370
Thompson, A. J., Moran, J. M., & Swensen G. W., Jr. 1986, Interferometry and Synthesis in Radio Astronomy (New York: Wiley Interscience)
Walker, R. C. 1984, in Indirect Imaging, ed. J. A. Roberts (Cambridge: Cambridge Univ. Press), 53
Zador, P. 1982, IEEE Trans. Inf. Th., IT-28, 139