Applied and Computational Harmonic Analysis **9**, 99–119 (2000)
doi:10.1006/acha.2000.0302, available online at http://www.idealibrary.com on IDEAL®

Fast Functional Magnetic Resonance Imaging via Prolate Wavelets [1]

Larry Shepp and Cun-Hui Zhang

Rutgers University, New Brunswick, New Jersey

Communicated by Ronald R. Coifman

Received July 22, 1998

Our objective is to improve the *time-resolution* of functional magnetic resonance imaging by sampling only a small fraction of the Fourier transform of the spin density and using a *prolate wavelet (P-wavelet) filter* to approximately obtain not the usual susceptibility map, but instead the *integral*, $I(t)$, of the difference in susceptibility between task and pretask, over a prespecified region of interest in the brain at successive time-points, t. This space–time trade-off thus allows us to obtain, at high time-resolution, the *total activity*, $I(t)$, in a specified region, B, of the brain which processes the specific stimulus or task to learn or verify where the brain function takes place. We find that for a typical region, B, of the brain, say describing the hippocampus, believed to be involved in memory, consisting of say 100 points in a 64×64 brain image space, that our optimal choice of Fourier sampling region, A, has $a = 400$ points, which then gives a 10-fold speed-up compared to the usual method of sampling, since the usual sampling needs 64^2 points which exceeds 10 times 400. Of course we get this speed-up at the price of spatial resolution. Even faster sampling of the integral of the susceptibility difference ought to be possible for a set B of 100 points in this level of pixelization. Once A is fixed, the mathematical problem of choosing the optimal P-wavelet filter can be viewed as a natural generalization to 2 and 3 dimensions of the theory of prolate spheroidal wave functions, due to Landau, Pollak, and Slepian. The first prolate spheroidal wave function for fixed A, B is that function ϕ which is maximally concentrated on B and whose Fourier transform vanishes except on A. The problem of choosing the sampling region A, for fixed size $a = |A|$, to maximize the concentration on the given B remains open and is probably NP-complete, i.e., probably involves exponential search. However, we give a very simple *heuristic* for choosing a set A, which seems to give excellent results in the examples we have studied, and we give the optimal P-wavelet filter, $\phi = \phi_B$, for the choice of A, based on our heuristic.

The heuristic is to take as it A the *scaled polar set* of B, i.e.,

$$A = \Big\{\mathbf{k}\colon \max_{\mathbf{x},\mathbf{y}\in B} (\mathbf{k}, \mathbf{x}-\mathbf{y}) < c\Big\}, \tag{1}$$

[1] A prolate wavelet may also be called a prolate-spheroidal wave function. See Acknowledgments.

where c is chosen so that the size of A is a, as desired to achieve a given time-resolution for $I(t)$. We give evidence to support our claim that A gives good results.

1. TIME RESOLUTION IN fMRI

Ogawa [12–14] showed, in celebrated work, that cerebral blood flow, refreshing areas of the brain which are active during the execution of a mental task with oxygenated blood, changes the local magnetic susceptibility which can be measured with a magnetic resonance (MR) scanner. More particularly, hemoglobin exists in two different states, oxyfied and deoxyfied, and deoxyfied hemoglobin produces a different local magnetic field than does oxyfied hemoglobin, because of the presence of iron in hemoglobin and because of the fact that the magnetic field of iron is stronger in local combination without oxygen than it is with oxygen. Measuring susceptibility then indicates the flow of blood to those areas of the brain being utilized to perform the task.

A serious problem with Ogawa's widely used technique, and later related techniques, is slowness. Ogawa's method is usually implemented to give a *complete* image every 10 s (often even slower), while many mental processes, for example image recognition, take place on a *much faster* time scale.

PET (positron emission tomography) is even slower, and PET requires several subjects for a single experiment and consequent appropriate rescaling and renormalization in order to be used in physiology (function studies). Even so, it is still competitive with fMRI, despite the fact that it involves ionizing radio-activity. But maybe this will change with an improvement in the time-resolution of fMRI.

In all the standard methods, the Fourier transform of the spin density is sampled completely and repeatedly as time progresses. We propose speeding up the method by sampling only *a small fraction* of Fourier space when the physiologist wants to determine how much activity[2] there is in a *specific region* of the brain. This allows the speed-up because we avoid *complete* sampling, even though the Fourier transform is actually changing everywhere.

Our method is based on some old and mostly overlooked work on *prolate spheroidal wave functions* which later came to be called wavelets.

Sampling only a small region of Fourier transform ($\mathbf{k}$) space cannot allow an inverse Fourier transform with resolution equivalent to sampling all of $\mathbf{k}$ space, needless to say. But for the purposes of physiology, or functional understanding, high spatial resolution is not so important, since areas of the brain are not even *named*, or even *recognizable* in anatomy books, to high spatial resolution. So, it seems, a trade-off between spatial and temporal resolution is what is needed and this is what we are proposing.

If the Fourier transform, $\hat{f}(\mathbf{k}, t)$, of the spin density, $f(\mathbf{x}, t)$, at time t is sampled for $\mathbf{k}$ in a set A, and it is desired to learn how much activity there is in a set B in $\mathbf{x}$-space at time t, then, as we show, this can be done much more quickly using *higher dimensional* prolate-spheroidal wave functions. In one dimension, the elegant theory of these functions is due to Landau, Pollak, and Slepian [7, 8, 19, 20]. These functions are now more commonly known as wavelets. If the set A has Lebesgue measure a, and if a is decreased by a

[2] We use the term *activity* because of emission tomography [17].

factor of 2, then we obtain twice the time resolution, for example. We expect to speed up the procedure by a factor of 10 or more. Instead of trying to reconstruct *the entire* spin density at each t we only reconstruct the *change* in a specified region, B, say the visual cortex or some different area where physiologists believe that the "rolodex" of images is located.

This leads to the following mathematics problem, of interest in itself in wavelet theory. Given a set B, a positive number a, and a set A, contained in $\mathbf{k}$ space and having measure a, there is a function, $\hat{\phi}(\mathbf{k})$, which vanishes off A and whose Fourier transform has a maximal fraction of its energy in B. Call this maximal fraction, $F(A, B)$, and then choose A so that this maximum is as large as possible. Then sampling $\hat{f}$ only for $\mathbf{k} \in A$ will give a fast method (if a is small) for obtaining how much activity there is in B. Indeed, if $\hat{\phi}$ is the function having maximal concentration in B, for a best A, then by Parseval's theorem, with z^* denoting the complex conjugate of z,

$$\int f(\mathbf{x})\phi(\mathbf{x})^* \, d\mathbf{x} = \int \hat{f}(\mathbf{k})\hat{\phi}(\mathbf{k})^* \, d\mathbf{k}.$$

The right side is known because $\hat{\phi}(\mathbf{k}) = 0$ where $\hat{f}(\mathbf{k})$ is not measured, and the left side is (approximately) the integral of $f(\mathbf{x})$ over $\mathbf{x} \in B$, if the right normalizing constant is used for $\hat{f}$. Which set A would be optimal, for given a? We offer some heuristics and some evidence that if B is convex, then A ought to be taken as a scaled version of the dual convex set, $A = cB^{\perp}$, where

$$B^{\perp} = \{\mathbf{k} : (\mathbf{k}, \mathbf{x} - \mathbf{y}) \leq 1, \text{ for all } x, y \in B\}$$

and where c is chosen so that $A = cB^{\perp}$ has measure a. Of course the same would hold for the discrete Fourier transform setting, except that now A and B become discrete sets, and the polar set A can be also defined as

$$A = \{\mathbf{k} : \min_{\mathbf{x}\in B} \max_{\mathbf{y}\in B} |(\mathbf{k}, \mathbf{x} - \mathbf{y})| < c\}.$$

The two definitions coincide in the continuous case. It is amazing to us that in the examples we have tried, it seems that the optimal A as defined by maximizing the maximum eigenvalue, or concentration, in B is *nearly* the same as the *heuristically* obtained set A defined above in terms of the polar dual to B. Finding the true optimal A, i.e., the A of size a which maximizes the concentration in B of the most concentrated function in B whose Fourier transform is supposed by A, seems difficult and we expect it will be shown to be NP-complete, that is a computationally difficult problem involving an exponentially growth in N computer search. As in the traveling salesman problem, the true optimum is hard to find, but there is also a simple *heuristic* to find a route for a salesman's tour on an arbitrary graph that is *probably* very close to having minimal length. In our case, we have no such proof.

Of course, it is the uncertainty principle that keeps the maximum eigenvalue from being unity: a function and its Fourier transform cannot both have compact support.

Another way to argue that the choice of A and $\hat{\phi}(\mathbf{k})$ should be that prolate function, or wavelet, which is the main one in the Landau–Slepian–Pollak sense, rather than any old

wavelet concentrated in both $\mathbf{x}$ and $\mathbf{k}$ space, is that we need $\hat{\phi}(\mathbf{k}) \equiv \mathbf{0}$, off of A, since we do not measure $\hat{f}(\mathbf{k})$ there.

It might be argued that it would be more reasonable to sample $\hat{f}(\mathbf{k})$ for $\mathbf{k}$ in a very diversely distributed set, i.e., that the set A ought to be a fractal, rather than a scaled version, of the dual convex set to B, which is a convex set rather than a fractal. The reason we do not think that this will be the case is rather involved. Indeed, one could argue that one should merely take a lower sampling rate of the discrete Fourier transform of f and if $\phi(\mathbf{x})$ is an approximate indicator of a set B, then we can reconstruct the integral of f over B approximately, by using the Poisson summation formula (in $d = 2$ or $d = 3$ dimensions),

$$h^d \sum f(\mathbf{j}h)\phi(\mathbf{j}h)^* = \sum \hat{f}\left(\frac{\mathbf{k}}{h}\right)\hat{\phi}\left(\frac{\mathbf{k}}{h}\right)^*.$$

It might similarly be argued that it would be simpler to use the best approximation (in L^2) of the indicator function χ_B by a function ϕ whose Fourier transform has support of the Lebesgue measure a. This problem is indeed simple to solve. The answer is to use the function ψ whose Fourier transform has support of measure a, and which best approximates $\hat{\chi}_B$, *the Fourier transform of the indicator of* B. ψ is easily seen to be simply the function whose Fourier transform coincides with $\hat{\chi}_B$ on the set of measure a and of the form $\{\mathbf{k} : |\hat{\chi}_B(\mathbf{k})| > c\}$. However, we think choosing ψ will not work out as well as the proposed method of using the prolate function which goes with a scaled version of $B^{\perp}$ as described above, because the set B is not so precisely known and specifiable and because the second formulation is too rigid and nonrobust in that a point in B but near the boundary of B is given much more weight than a nearby point on the outside of B. The set $\{\mathbf{k} : |\hat{\chi}_B(\mathbf{k})| > c\}$ is not even connected unless its size a is small enough. The use of the prolate spheroidal function which goes with B and the best A will approximate the indicator, in a rough sense, and should provide a more robust estimate of the integral of f over B, deemphasizing the boundary of B in a reasonable way.

The intuition for *not* using a fractal A as in the Poisson sum formula is slightly different, but again heuristic and nonrigorous. If we knew nothing about f then the approximate integral over B would probably be best taken by the Poisson sum formula method above. So if f did not remain positive (or negative) within B, but instead could *change sign* within B, then we would probably not want to use the prolate method. But we are assuming that the mental process taking place inside of B will cause $f(\mathbf{x}, t) - f(\mathbf{x}, 0)$ to be of the same sign, because B is the area of increased deoxyhemoglobin. If our assumption is wrong, and in some parts of B there is less deoxyhemoglobin, then some other region should have been designated as B, perhaps a subregion of the B should have been chosen instead. We are not so much interested in getting the total integral of f over *some translate* of B (though we can obtain this as a control and at no additional sampling cost merely by using, instead of $\phi_B(\mathbf{j})$, a phase shift $\sum_{k \in A} e^{i(\mathbf{s}-\mathbf{j},\mathbf{k})}\hat{\phi}_B(\mathbf{k})$, which then corresponds to a translate of B to $B + \mathbf{s}$). Of course one should try all the above methods to verify the above reasoning in practice; we are not fully convinced ourselves by our intuitive reasoning. In principle one could study the choice of A and ϕ via the method of simulations, so effective in CT scanning, but simulations in MRI, not to mention fMRI, are not easily carried out due to the increased number of free variables present in MRI compared to CT.

For the background of the MRI and related fields, we refer to Monon *et al*. [11], Ogawa *et al*. [12–14], Buonocore *et al*. [1], and Cho *et al*. [2]. For more engineering details, see [18].

Prolate-spheroidal wave functions were considered by Slepian and Pollak [20], Landau and Pollak [7, 8], and Slepian [19]. Donoho and Stark [5, 6] considered applications of the prolate-spheroidal wave functions in a different context and proved that, in the one-dimensional case, for intervals B and $\mu(A)\mu(B) \leq 0.8$, where μ is the Lebesgue measure, $F(A, B)$ is maximized by an interval A. Wavelets are discussed in Chu [3], Daubechie [4], Meyer [9], and Meyer and Coifman [10].

2. CHOOSING A P-WAVELET

Because we are using a *quadratic* criterion for optimization, we have to make some assumptions on f to make the present approach sensible. Thus if f (which is the difference between the after-stimulus and the before-stimulus spin density signal in BOLD, which includes the effect of susceptibility) is as likely to be negative as positive inside of B, then there is not much sense in trying to find its integral over B, and there is no reason to want the filter $\phi(\mathbf{x})$ to be of one sign on B. But we believe [3] that $f(\mathbf{x})$ is of one sign (positive or negative) throughout the active region B and we expect that f mostly vanishes off B. Here we are relying on the physiologist to tell us that the stimulus is handled by the region, B, of the brain.

Now we want to argue that $\int_A \hat{f}\hat{\phi}_1^* \int_B \phi_1/\lambda_1$ is a reasonable proxy of $\int_B f$, where ϕ_1 is the main prolate-spheroidal wave function for A, B and λ_1 is the main eigenvalue. What is the reason for this?

Let us review the theory of prolate-spheroidal wave functions as wavelets. For a given set B in real space, and A, in Fourier transform space, consider the set of all functions, $\phi(\mathbf{x})$, for which the Fourier transform $\hat{\phi}(\mathbf{k}) = 0$ for $\mathbf{k}$ outside A. Let ϕ_1 be that function which maximizes $\int_B |\phi|^2 / \int_{\mathbb{R}^d} |\phi|^2$. Let ϕ_2 be defined similarly as the function maximizing the last ratio and also orthogonal to ϕ_1. We continue, in the same way, and define ϕ_j, $j = 1, 2, 3, \ldots$. This is exactly what was done in one dimension in [20], except we are dealing with general sets A and B in $\mathbb{R}^d$. In fact, for our purposes, space is discrete and we are using the discrete Fourier transform instead of the continuous Fourier transform. But this makes little difference. As in the one-dimensional case, the functions $\hat{\phi}_j$, the Fourier transformation of ϕ_j, are eigenfunctions of the prolate spheroidal operator (whence the unwieldy name), which in the discrete case is simply the $|A| \times |A|$ matrix,

$$K_{A,B}(\mathbf{k}, \mathbf{k}') = \sum_{\mathbf{j} \in B} \exp[(2\pi i/N)(\mathbf{j}, \mathbf{k} - \mathbf{k}')]/N^d, \qquad \mathbf{k}, \mathbf{k}' \in A.$$

This matrix can be written as $K_{A,B} = \chi_A F \chi_B F^* \chi_A$, where F is the $N^d \times N^d$ discrete Fourier transformation matrix with elements $\exp[(2\pi i/N)(\mathbf{k}, \mathbf{j})]/N^{d/2}$ and F^* its inversion. The eigenvalues, λ_j, $j \geq 1$, will all be nonnegative since $K_{A,B}$ is a positive-definite matrix, as is easy to see. We will have $1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|A| \wedge |B|} > 0$, since $\|F\| = \|F^*\| = 1$ and the rank of $K_{A,B}$ is $|A| \wedge |B|$. Let $\psi_j = \chi_B \psi_j$ be the

[3] This is because we expect that there is a dip in oxyhemoglobin in the active region B because the need for oxygen is supplied initially by the local blood pool.

orthonormal eigenvectors of $K_{B,A} = \chi_B F^* \chi_A F \chi_B$. Then $\chi_A F \psi_j = \chi_A F \chi_B \psi_j$ are orthogonal eigenvectors of $K_{A,B}$ and $\|\chi_A F \psi_j\|^2 = (\psi_j, K_{B,A}\psi_j) = \lambda_j$, so that $\hat{\phi}_j = \chi_A F \psi_j / \sqrt{\lambda_j}$, $\phi_j = F^* \hat{\phi}_j = F^* \chi_A F \psi_j / \sqrt{\lambda_j}$, and $\chi_B \phi_j = \sqrt{\lambda_j} \psi_j$. We will always choose A to be symmetrical about $\mathbf{k} = \mathbf{0}$, which means $A = -A \pmod N$, and then the eigenvectors ϕ_j can always be taken to be real and orthonormal in the entire space, as the symmetry of A guarantees that both matrices $K_{B,A}$ and $F^* \chi_A F$ are real and nonnegative definite. In addition, ψ_j ($= \chi_B \phi_j / \sqrt{\lambda_j}$ for $\lambda_j > 0$) form a real orthonormal and complete base in the space of all functions on B. Note that the polar of B is always symmetrical about $\mathbf{k} = \mathbf{0}$.

We are ready to provide justifications for the use of $\int_A \hat{f}\hat{\phi}_1 \int_B \phi_1/\lambda_1$ as a proxy of $\int_B f$. Since ψ_j form an orthonormal and complete base for functions on B, $\int_B f = \sum_j \int_B f \psi_j \int_B \psi_j$. We would like to replace the above sum by its first element, arguing that either $\int_B f \psi_j$ or $\int_B \psi_j$ are likely to be relatively small in absolute value for $j \geq 2$ and that there is likely to be a lot of cancellations in $\sum_{j \geq 2} \int_B f \psi_j \int_B \psi_j$. This is a weak point in our theory, but it seems reasonable in view of the fact that we are assuming that f is of one sign over B. Here we are certainly being led by the extreme elegance of the prolate spheroidal wavelet theory [20, 19, 7, 8]. Since ϕ_1 is essentially zero outside B ($\lambda_1 \approx 1$) and $\hat{\phi}_1$ is zero outside A, $\int_B f \sim \int_B f \psi_1 \int_B \psi_1 = \int_B f \phi_1 \int_B \phi_1/\lambda_1 \approx \int f \phi_1 \int_B \phi_1/\lambda_1 = \int_A \hat{f}\hat{\phi}_1^* \int_B \phi_1/\lambda_1$. The analogous formula in the discrete case is

$$\sum_B f \sim \sum_A \hat{f}\hat{\phi}_1^* \sum_B \phi_1/\lambda_1.$$

Now B is fixed by the physiologist, but how do we choose A, the region of $\mathbf{k}$-space where the Fourier transform of f is measured? If one wants to reconstruct f with a certain spatial resolution, one measures $\hat{f}(\mathbf{k})$, for $\mathbf{k}$ in a raster grid of N^d points, where N is the linear discrimination distance in each direction of d-space. Here, $d = 2, 3$. If A contains $|A| \ll N^d$ points, then time resolution is improved by a factor $N^d/|A|$, since each measured value of $\hat{f}(\mathbf{k})$ costs a fixed amount of time. It is a slight advantage if the set A consists of *nearby points* since the gradients do not have to be adjusted so much in this case to get each consecutive measurement. Connected A is preferred, and convex A is even more preferred. Fortunately, the proposed choice of A, the polar sets, are convex ones. We argue that it is roughly that A which maximizes the value of λ_1, the maximal eigenvalue of $K_{A,B}$, i.e., which maximizes the maximum ratio

$$\sum_B |\phi|^2 \Big/ \sum |\phi|^2$$

among all ϕ for which the size of the support of $\hat{\phi}$ is a. Since the set B is most probably only approximately known to the physiologist, an approximately optimal A is as good as an optimal A.

For convex sets B, the best A in the above criterion seems to be close to the scaled dual convex set of B. This heuristic is supported by our numerical results below in the two-dimensional case and Donoho and Stark [6] in the one-dimensional case. Actually, it is very easy to prove that the optimal A is a block of consecutive integers when B is such, for small a in the one-dimensional discrete case. For $B = \{1, \ldots, b\}$ and A' containing

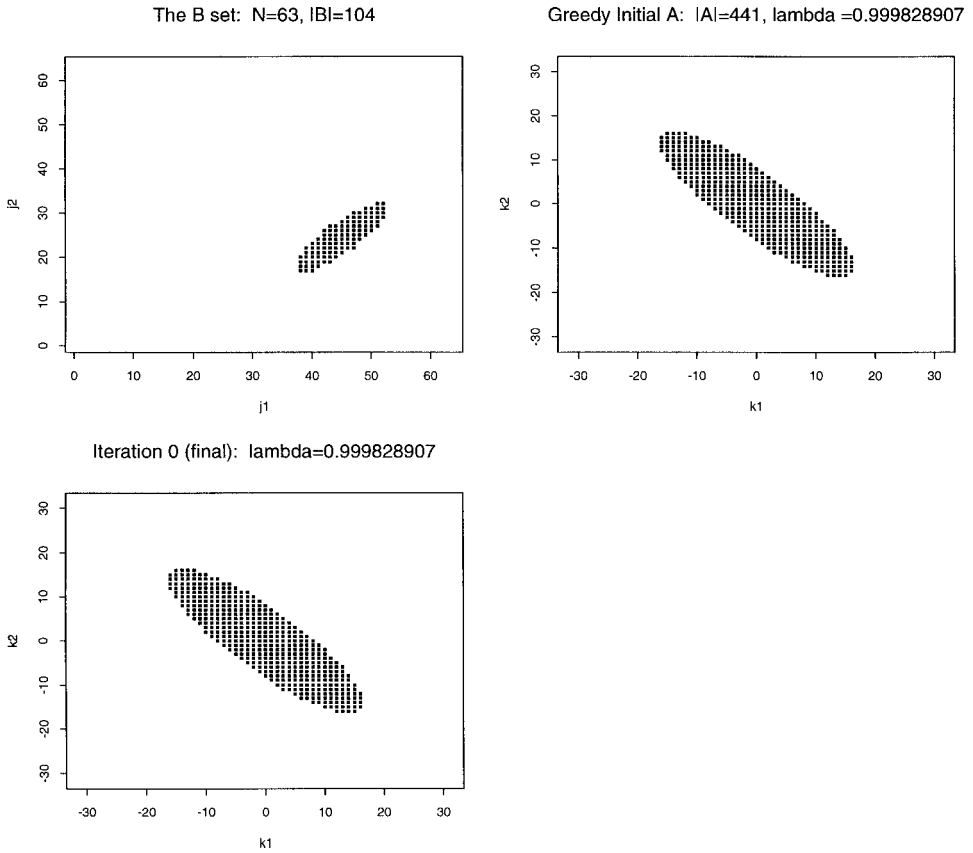

FIG. 1. MM-iterations with $a = 441$ and greedy set as initialization. We see that this greedy set in the upper right is stable.

$k_1 < \cdots < k_a$,

$$F(A', B) = \sup_{\|v\|=1} \sum_{\ell,\ell'} K(k_\ell - k_{\ell'}) v_\ell v_{\ell'}^*, \qquad K(k) = \frac{\sin(b\pi k/N)}{N \sin(\pi k/N)}.$$

For small a, $|K(k)|$ is decreasing in k for $0 \le |k| \le a$ and $|K(k)| \le |K(a)|$ for $|k| > a$. Thus,

$$F(A', B) \le \sup_{\|v\|=1} \sum_{\ell,\ell'} K(\ell - \ell') |v_\ell v_{\ell'}| = F(A, B),$$

where $A = \{1, \ldots, a\}$. We do not have a proof in the higher-dimensional case or for large a in the one-dimensional case. The optimization problem of finding A to maximize $F(A, B)$ for a given a can be solved by exhaustive search in the discrete case.

Suppose we have fixed A and B, and we compute the eigenvalues of $K_{A,B}$, and we find that $\lambda_1 = \lambda_2$. This is an indication that we have chosen $a = |A|$ too large and that a smaller a would suffice and would enable better time resolution. Indeed, in this case one could use *either* ϕ_1 or ϕ_2 so one could use a combination to *subdivide* B. The physiologist believes and tells us that the activity is happening in B. If it is happening almost everywhere

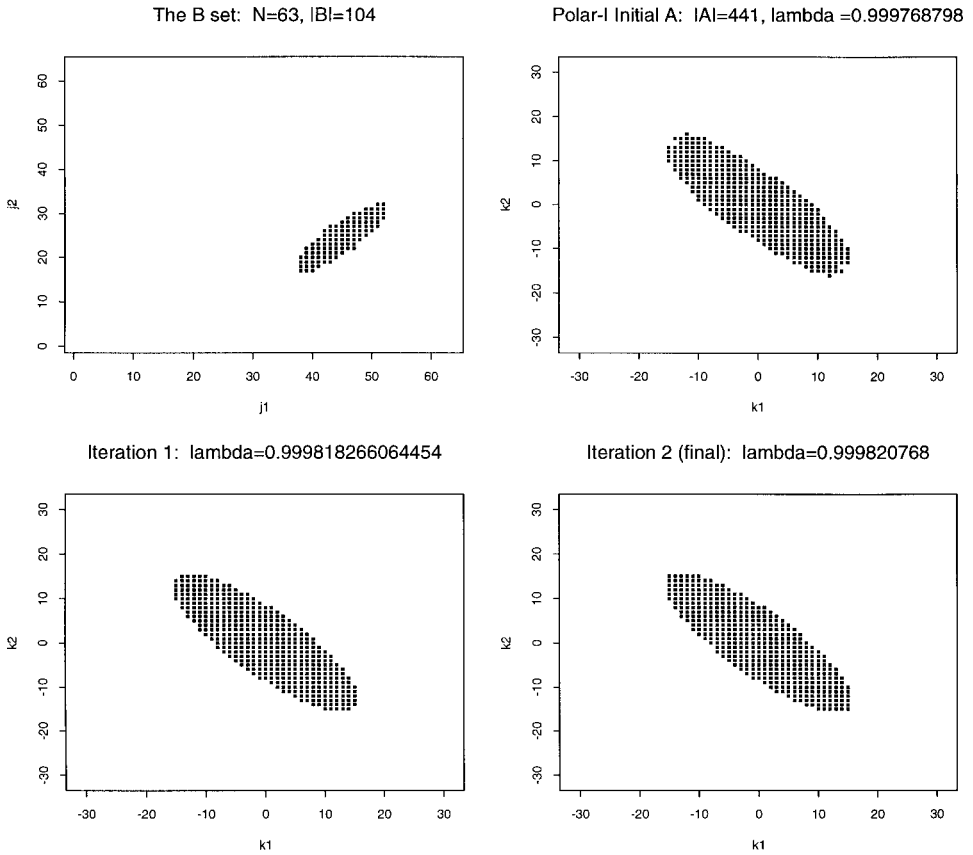

FIG. 2. MM-iterations with $a = 441$ and Polar-I as initialization. We see that this set in the upper right is unstable, but is stable after two iterations of MM.

in B, the optimal a should provide $\lambda_2 \ll \lambda_1 \approx 1$. If it is not known *where* in B it is happening, we may choose a such that $\lambda_{m+1} \ll \lambda_m \approx 1$ for some suitable $m > 1$, and partially reconstruct f on B by the formula

$$\chi_B f \sim \sum_{j=1}^{m} (f, \psi_j)\psi_j \approx \sum_{j=1}^{m} \frac{\sum_A \hat{f}\hat{\psi}_j^*}{\lambda_j} \chi_B \psi_j.$$

How large should $a = |A|$ be to assure $\lambda_2 \ll \lambda_1 \approx 1$? Since $\sum_j \lambda_j = \text{trace}(K_{A,B}) = |A||B|/N^d$, $\lambda_1 \approx 1$ requires $|A||B|/N^d \geq 1$. Some indication that $|A||B|/N^d$ should not be too large for the optimal A is that $\lambda_1 = 1$ and $\lambda_2 = 0$ for $|A||B| = N^d$ exactly when $|B| = 1$, whence $|A| = N^d$, and also when $|B| = N^d$, whence $A = \{0\}$ and $|A| = 1$. We have conducted experiments described next to see how good this relation is.

3. MATHEMATICAL EXPERIMENTATION

We chose for B several discrete subsets of an $N \times N$, $N = 63$, grid. We then considered the problem of optimally choosing a set A in discrete, Fourier space to maximize the

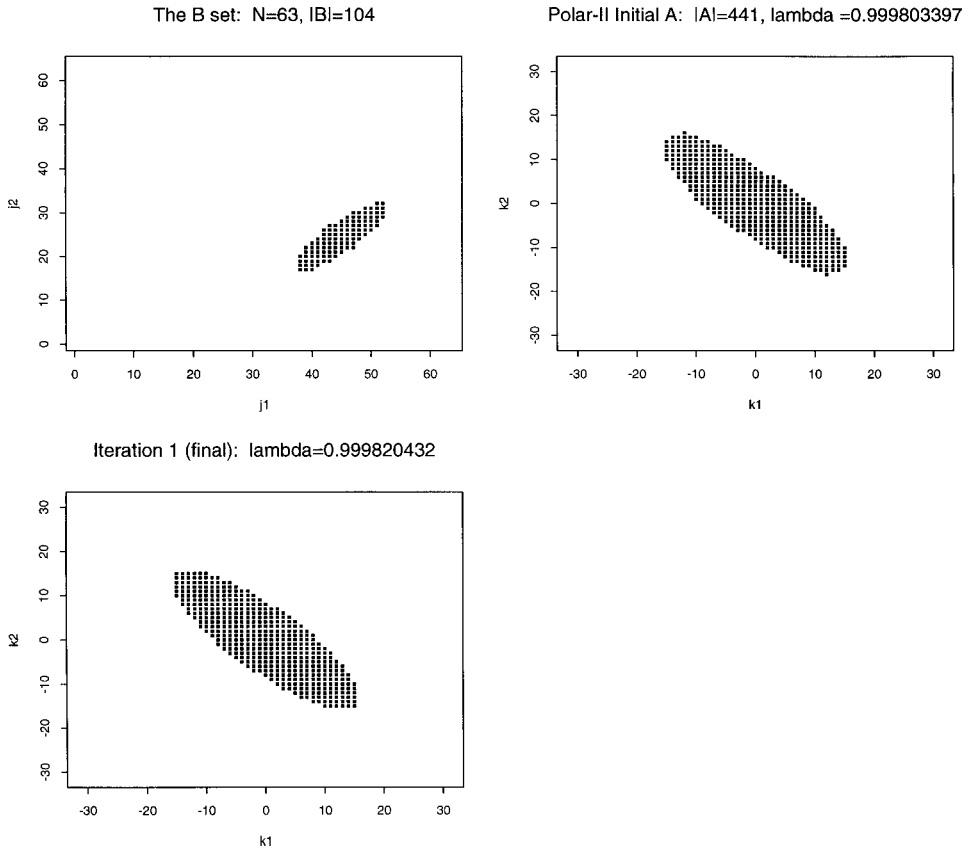

FIG. 3. MM-iterations with $a = 441$ and Polar-II as initialization. We see that this set is unstable, but is stable after one iteration of MM.

maximum eigenvalue of the kernel

$$K_{B,A}(\mathbf{j}_1, \mathbf{j}_2) = \sum_{\{\mathbf{k}\in A\}} \exp\{-2\pi i(\mathbf{k}, \mathbf{j}_1 - \mathbf{j}_2)/N\}/N^d,$$

or at least this is what we would like to do.

Four methods are considered to select a (nearly) optimal set A:

(i) Polar-I, $A = \{\mathbf{k} : \min_{\mathbf{x}\in B} \max_{\mathbf{y}\in B} |(\mathbf{k}, \mathbf{x} - \mathbf{y})| < c\}$ with $|A| = a$;
(ii) Polar-II, $A = \{\mathbf{k} : \max_{\mathbf{x},\mathbf{y}\in B}(\mathbf{k}, \mathbf{x} - \mathbf{y}) < c\}$ with $|A| = a$;
(iii) *Greedy search* to be described below;
(iv) MM-*algorithm* to be described below.

The greedy search begins with the single point set $\tilde{A}_1 = \{\mathbf{0}\}$ and at each step adds only two symmetric points to the previously optimal A; i.e., it sequentially searches for $\tilde{A}_1 \subset \tilde{A}_3 \subset \tilde{A}_5 \subset \cdots$ such that $\tilde{A}_{a+2}$ is the optimal symmetric $A \supset \tilde{A}_a$ with $|A| = a+2,\ a = 1, 3, 5, \ldots$. The result of the greedy search will be called the *greedy set*. Clearly, the greedy sets are optimal for $a = 1, \ldots, 2m+1$ iff $A_1 \subset \cdots \subset A_{2m+1}$, where A_a is the optimal symmetric A of size a. The MM-algorithm exploits the fact that our objective is

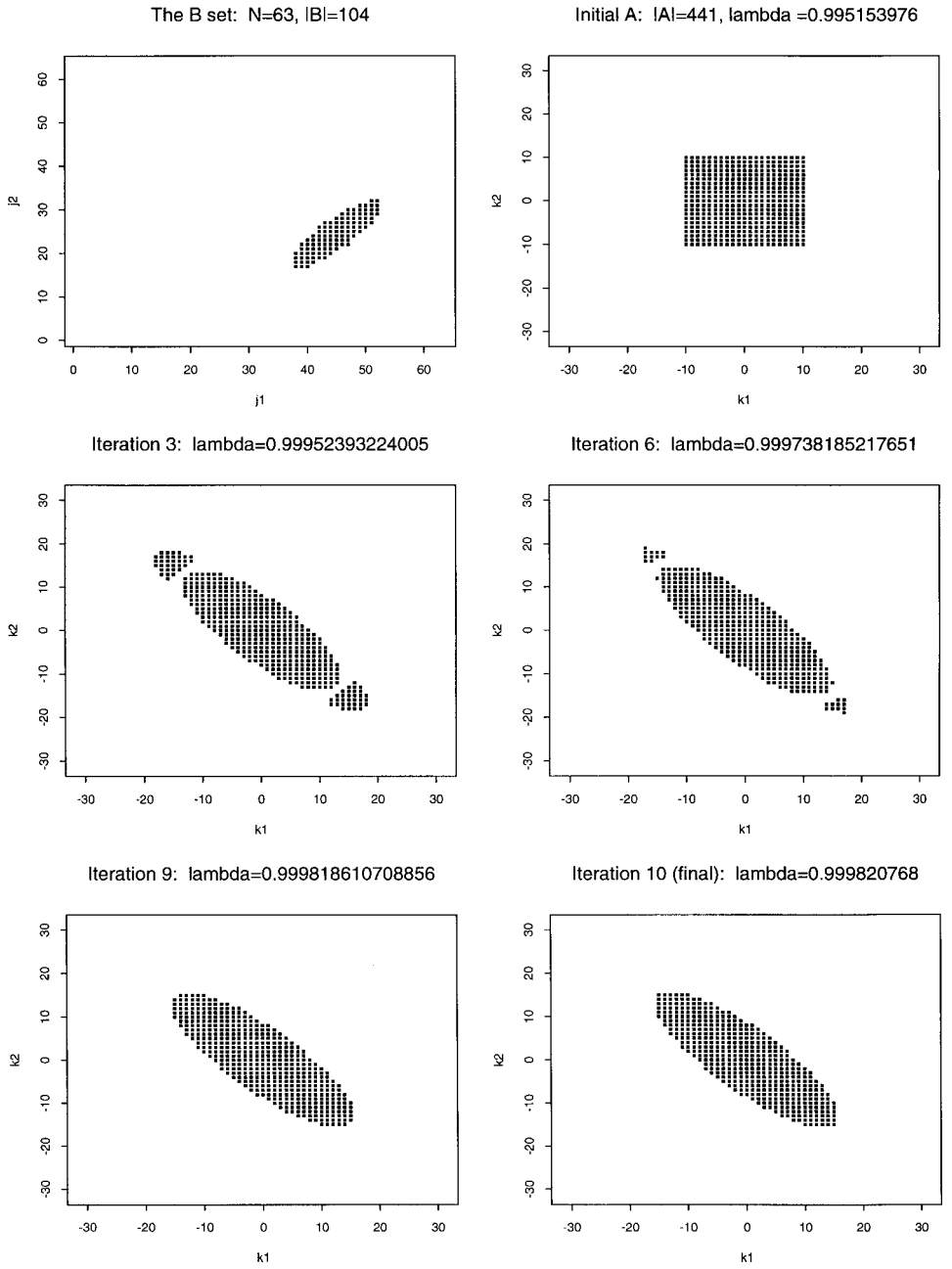

FIG. 4. MM-iterations with $a = 441$ and 21×21 square as initialization. We see that this set is unstable and needs 10 iterations of MM to become stable (at a different stable point than Fig. 3).

actually to find an A of size a at which

$$\max_{|A|=a} \max_{\|\psi\|=1} (\psi, K_{B,A}\psi) = \max_{|A|=a} \max_{\|\psi\|=1} \sum_{\mathbf{k}\in A} \left| \frac{1}{N^d} \sum_{\mathbf{j}\in B} \exp[(2\pi i)/N(\mathbf{j},\mathbf{k})]\psi(\mathbf{j}) \right|$$

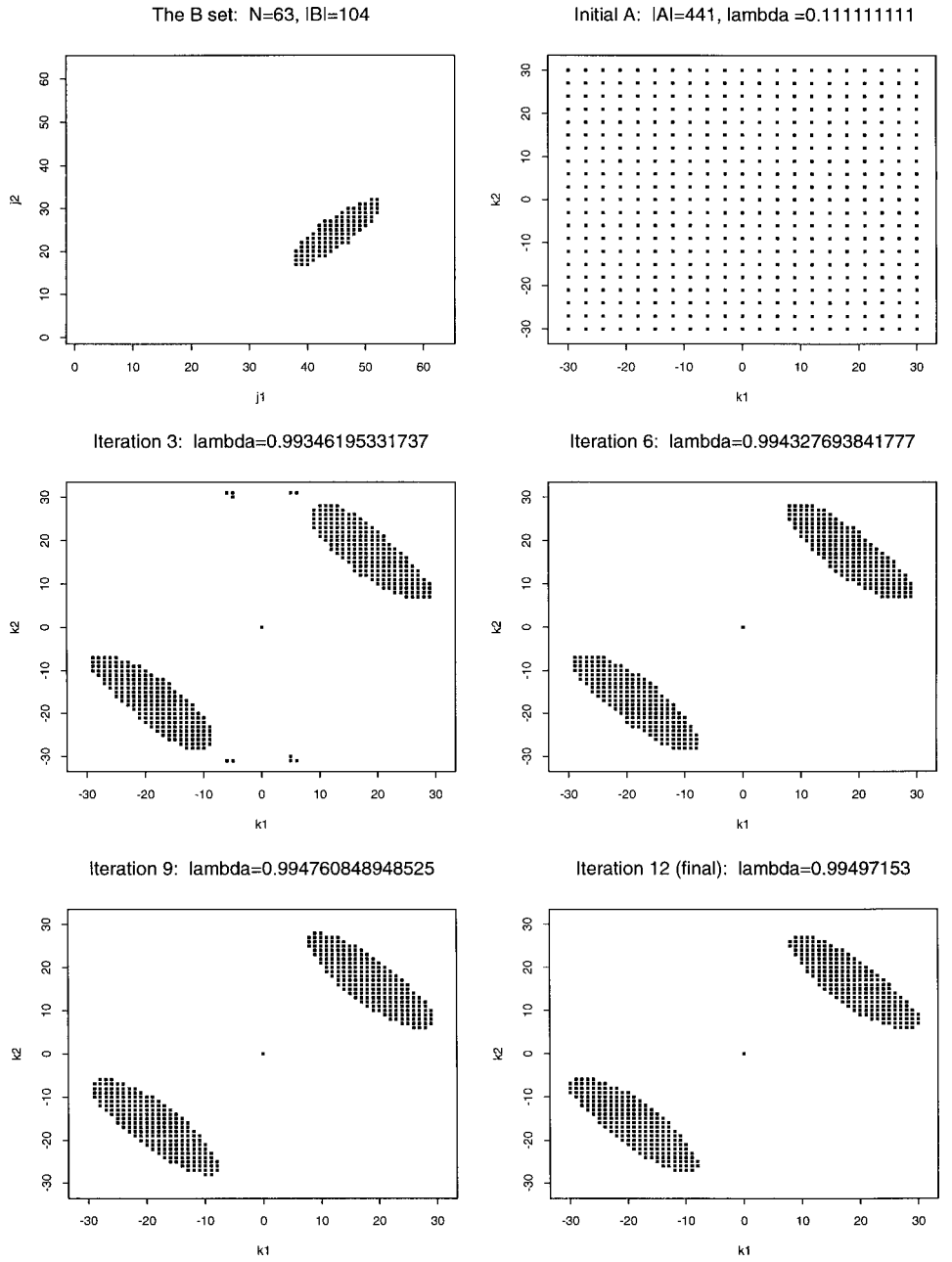

FIG. 5. MM-iterations with $a = 441$ and uniform A as initialization. We see that this set is unstable and needs 12 iterations of MM to become stable.

is reached, so that it alternately maximizes the sum above over ψ, and A; for a given initialization $A^{(0)}$, it generates iteratively

$$A^{(n+1)} = \left\{ \mathbf{k} : \left| \sum_B \exp[(2\pi i)/N(\mathbf{j}, \mathbf{k})] \psi^{(n)}(\mathbf{j}) \right| > c_n \right\}, \qquad \left| A^{(n+1)} \right| = a,$$

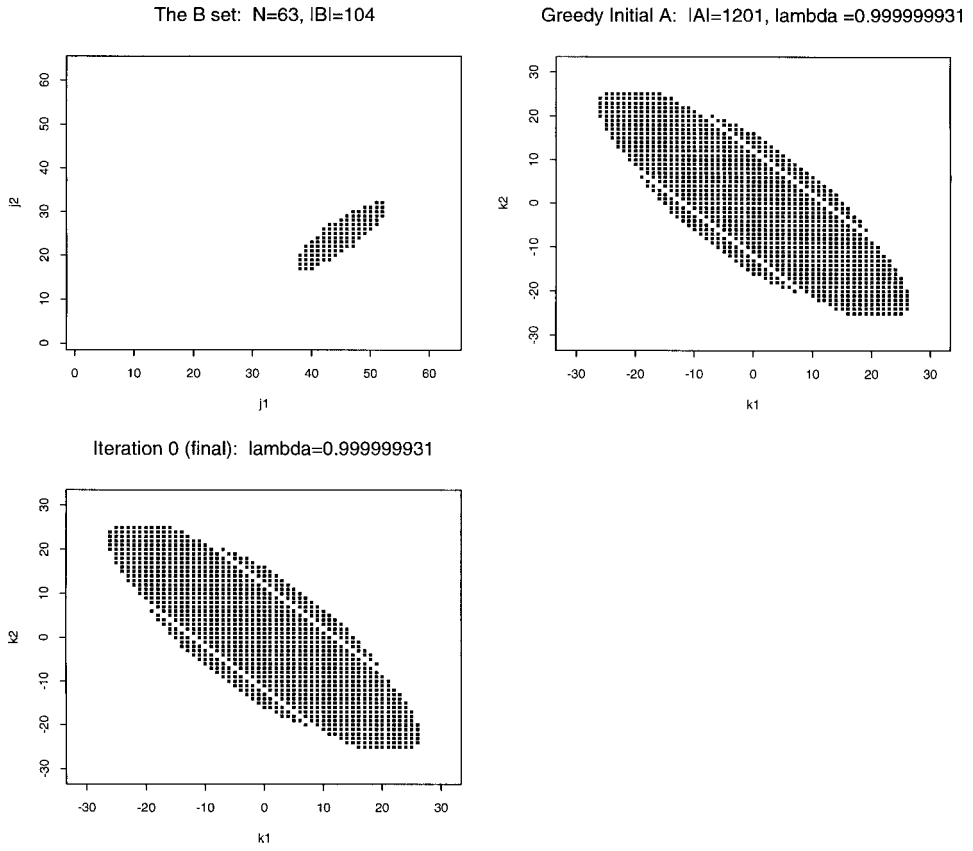

FIG. 6. MM-iterations with $a = 1201$ and greedy set as initialization. This greedy set is again stable.

where $\psi^{(n)}$ is the first eigenvector of $K_{B,A^{(n)}}$. For a given B, a set $A = A^{(n)}$ is stable if the first eigenvalue of $K_{B,A^{(n)}}$ is the same as that of $K_{B,A^{(n+1)}}$, and it is strongly stable if $A^{(n)} = A^{(n+1)}$. Clearly, the optimal A is a stable set.

What we have found strongly indicates that the polar sets are nearly optimal; we find:

(1) the best A at each step of the greedy search always remains very close to the polars of B, i.e., to Polar-I and Polar-II, with the same number of points;

(2) the polar sets and the greedy sets are *often* (strongly) stable;

(3) when they are not strongly stable the MM-algorithm does not provide much improvement in terms of the first eigenvalue;

Remark. It may seem that this is scanty evidence because of the fact that this is a greedy search, but we find it very surprising that the greedy search consistently nearly produced the polar set for the examples of B we used, as described below.

(4) in most cases with $|A| \le N^d/5$ the greedy sets are strongly stable and have larger first eigenvalue than the polars;

(5) for large a, polars have a larger first eigenvalue than the greedy sets.

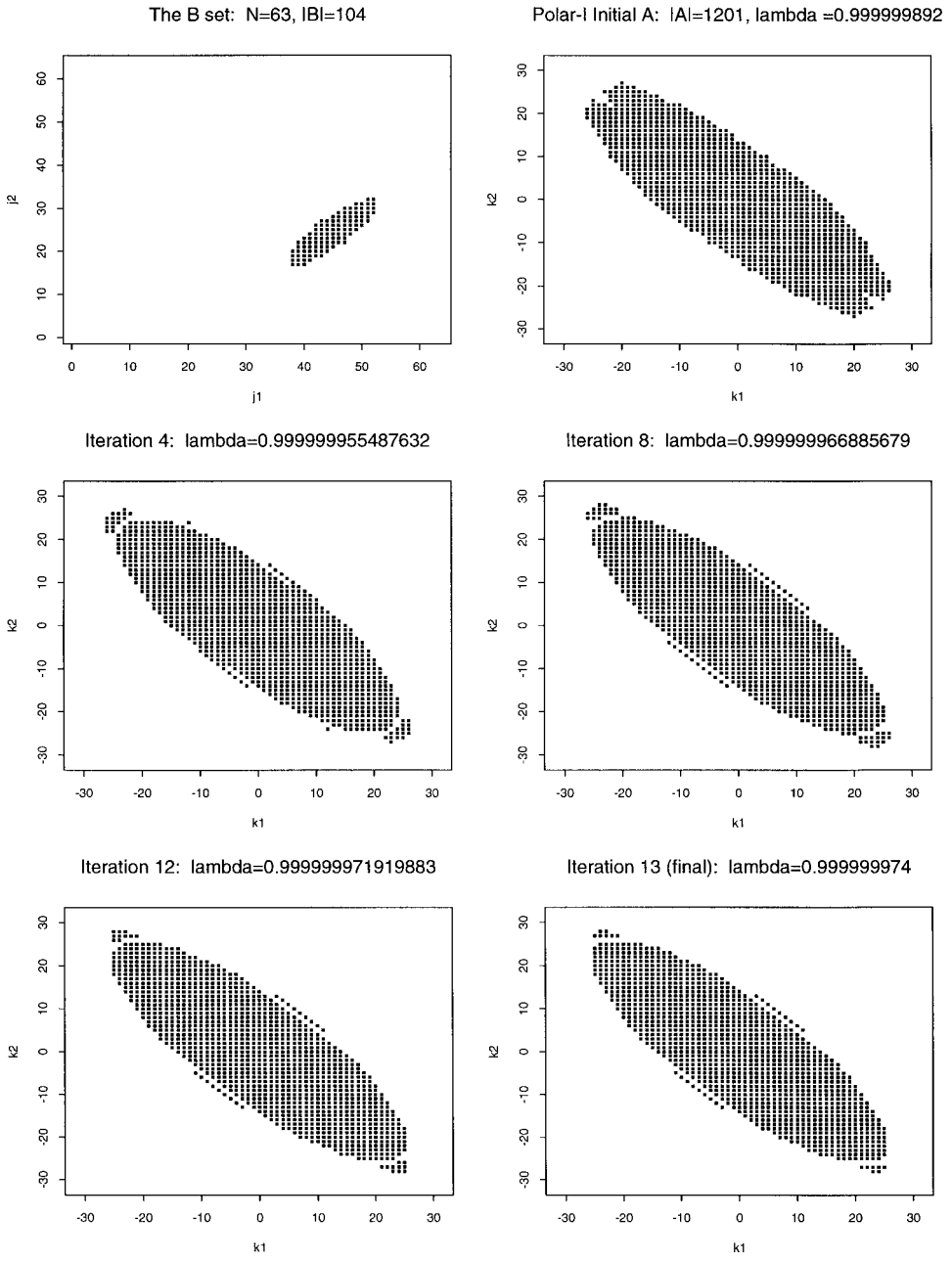

FIG. 7. MM-iterations with $a = 1201$ and Polar-I as initialization.

Remark. Since the greedy search is not always optimal, the optimal sets for different sizes a are not completely nested, but at least for small a, the greedy search seems to produce the optimal A.

The greedy sets are nearly convex in a discrete sense, but sometimes, they may have some holes. Of course, in the practical measurement situation the set of values where $\hat{f}(\mathbf{k})$

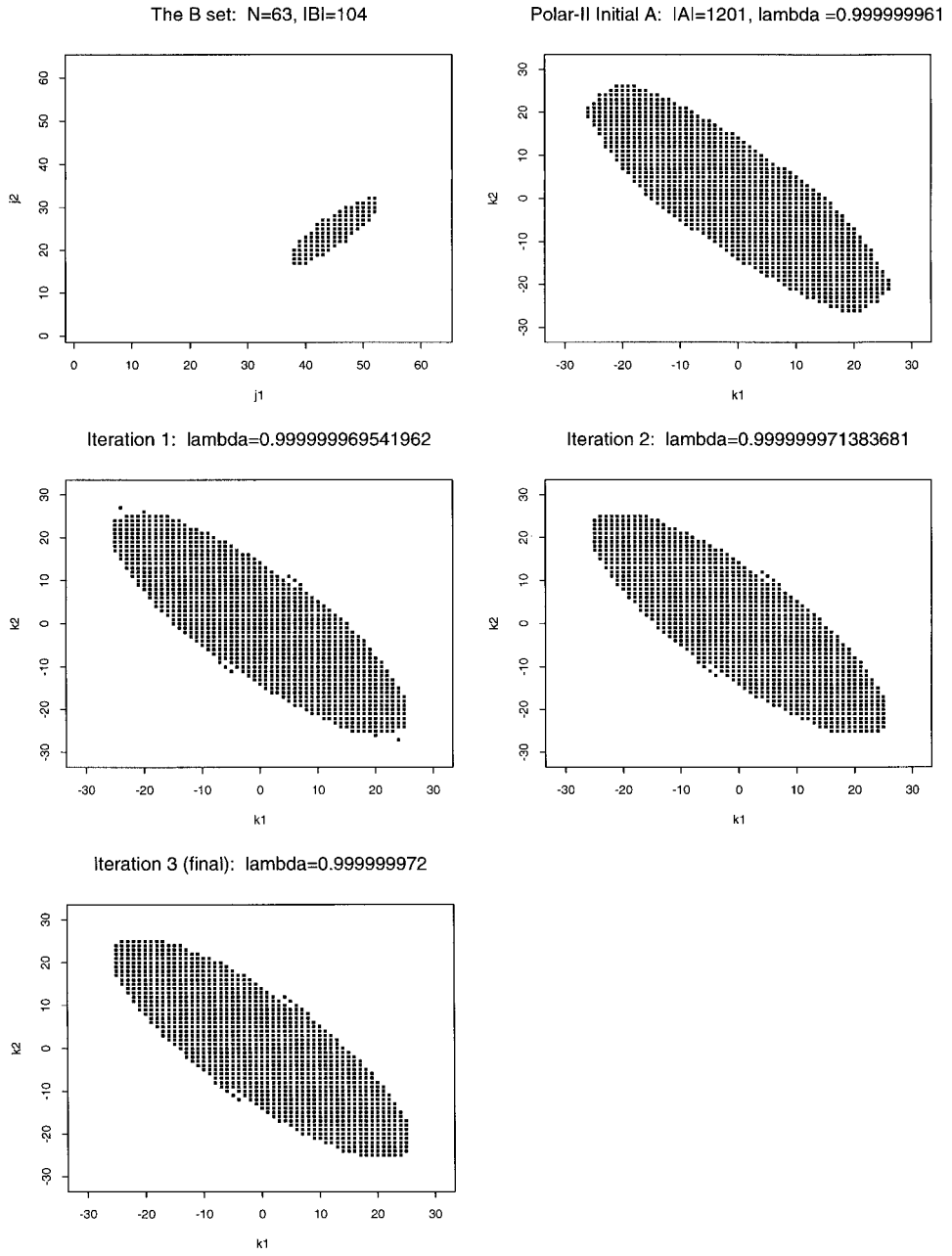

FIG. 8. MM-iterations with $a = 1201$ and Polar-II as initialization. In comparing Figs. 6–8, note that 1201 is too large for practice and the eigenvalues are nearly equal, but Polar-II does best.

is measured would certainly be taken to be convex, and the recommended choice of the scaled polar set should work well.

Figures 1–13 describe our experiments with a phantom compatible with one frequently used in CT scanning, in a 63×63 grid. For $a = 441$ ($\approx N^d/10$), Figs. 1–5 show the MM-algorithm with the initialization being the greedy set, Polar-I, Polar-II, square of

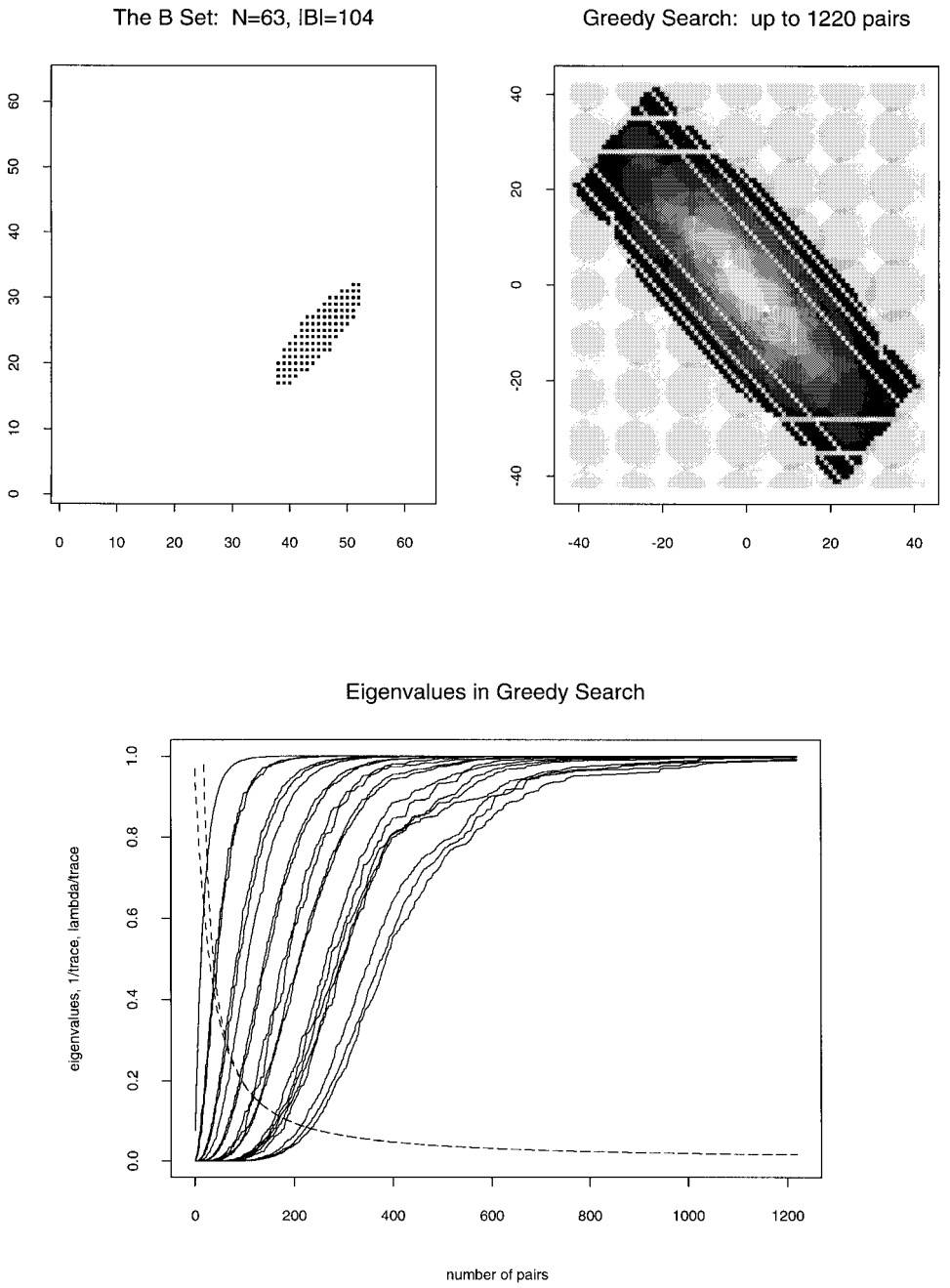

FIG. 9. The growth of greedy sets and the first 20 eigenvalues. Note the holes in the greedy sets for large a. Earlier pairs are lighter in gray level.

21×21, and the uniformly distributed A, respectively; the greedy set is the only stable one and provides the largest first eigenvalue. For $a = 1201$, Polar-II is the best as shown in Figs. 6–8. Figure 9 shows the growth of greedy sets as a varies and the first 20 eigenvalues

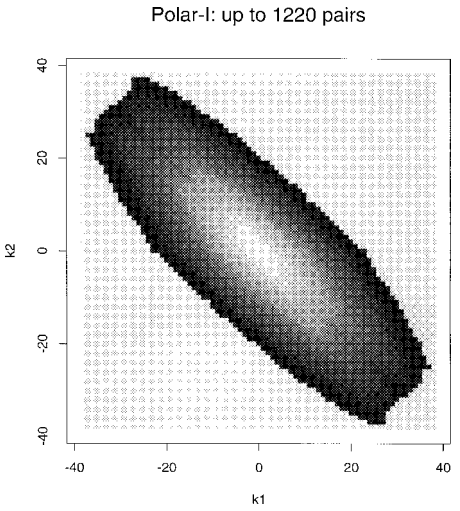

FIG. 10. The growth of Polar-I.

for the greedy sets as a varies; it can be seen that trace $= |A||B|/N^2 \approx 3$ $(1/\text{trace} \approx 1/3)$ provides $\lambda_2 \ll \lambda_1 \approx 1$. Figures 10 and 11 show the growth of polar sets. Figure 12 shows the optimal filters ϕ_1 for the greedy sets with various a; they are positive on the set B. Figure 13 shows the first six eigenvectors for the greedy set with $a = 225$; for $j \geq 2$, ϕ_j takes different signs on B. The greedy searches with two other B sets are described respectively in Figs. 14 and 15, which are self-explanatory.

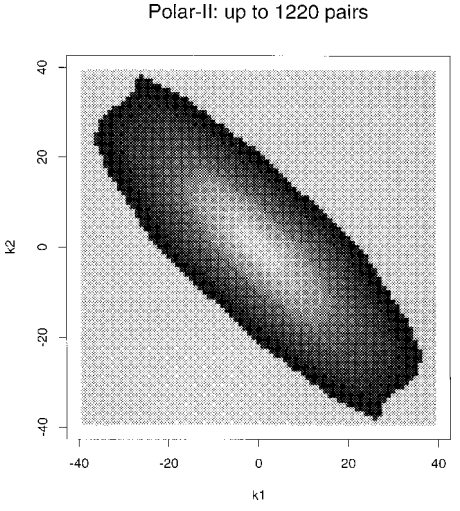

FIG. 11. The growth of Polar-II. We see that Polar-I and Polar-II are quite similar, but Polar-I has holes.

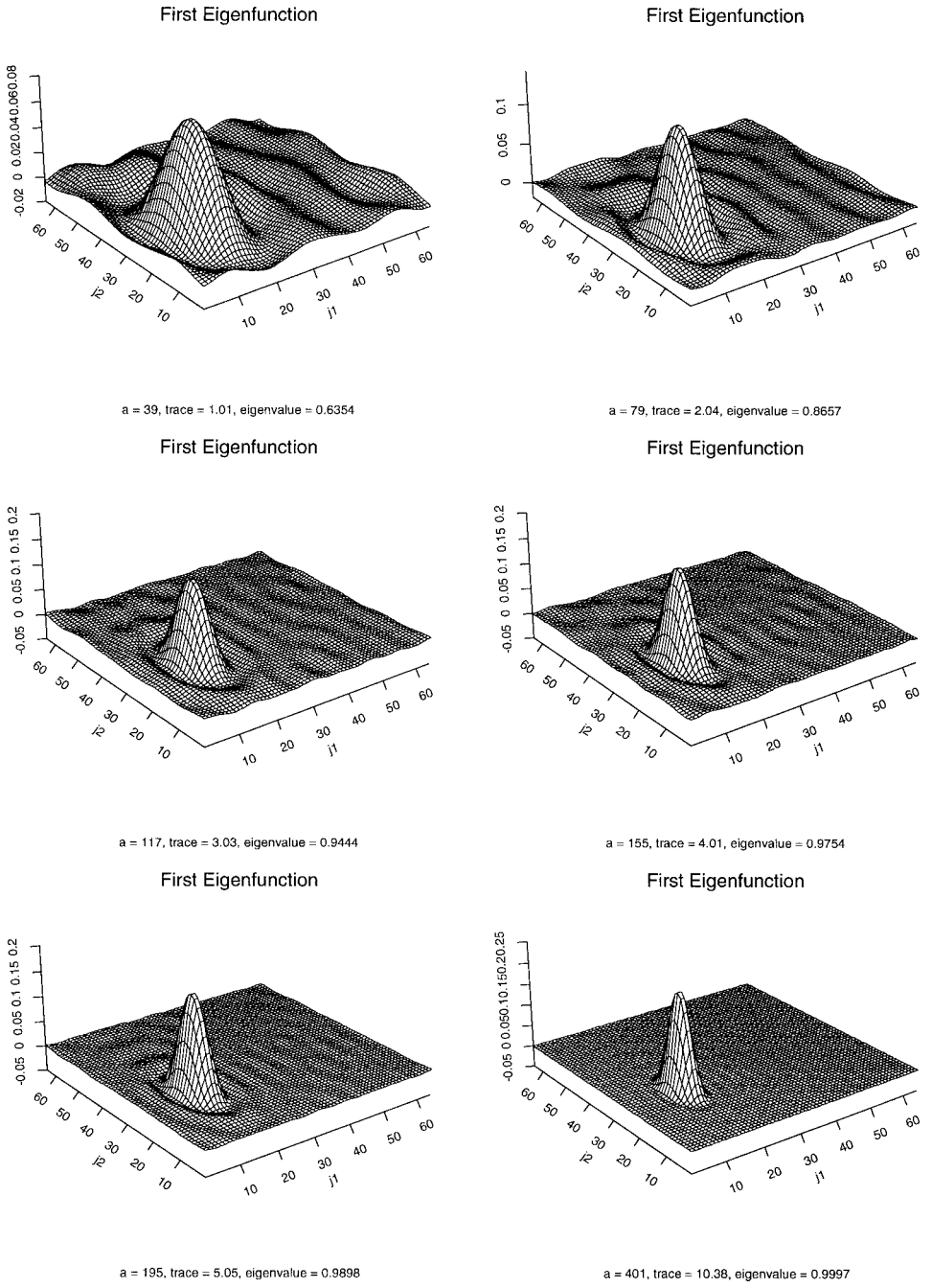

FIG. 12. Optimal filters for the greedy sets. One is getting a good integral of the activity even with 155 points (lower right).

4. CONCLUSIONS

We have specified a method to obtain a new measurement of the total mental activity in a specified region of the brain at an instant of time, which can be repeated

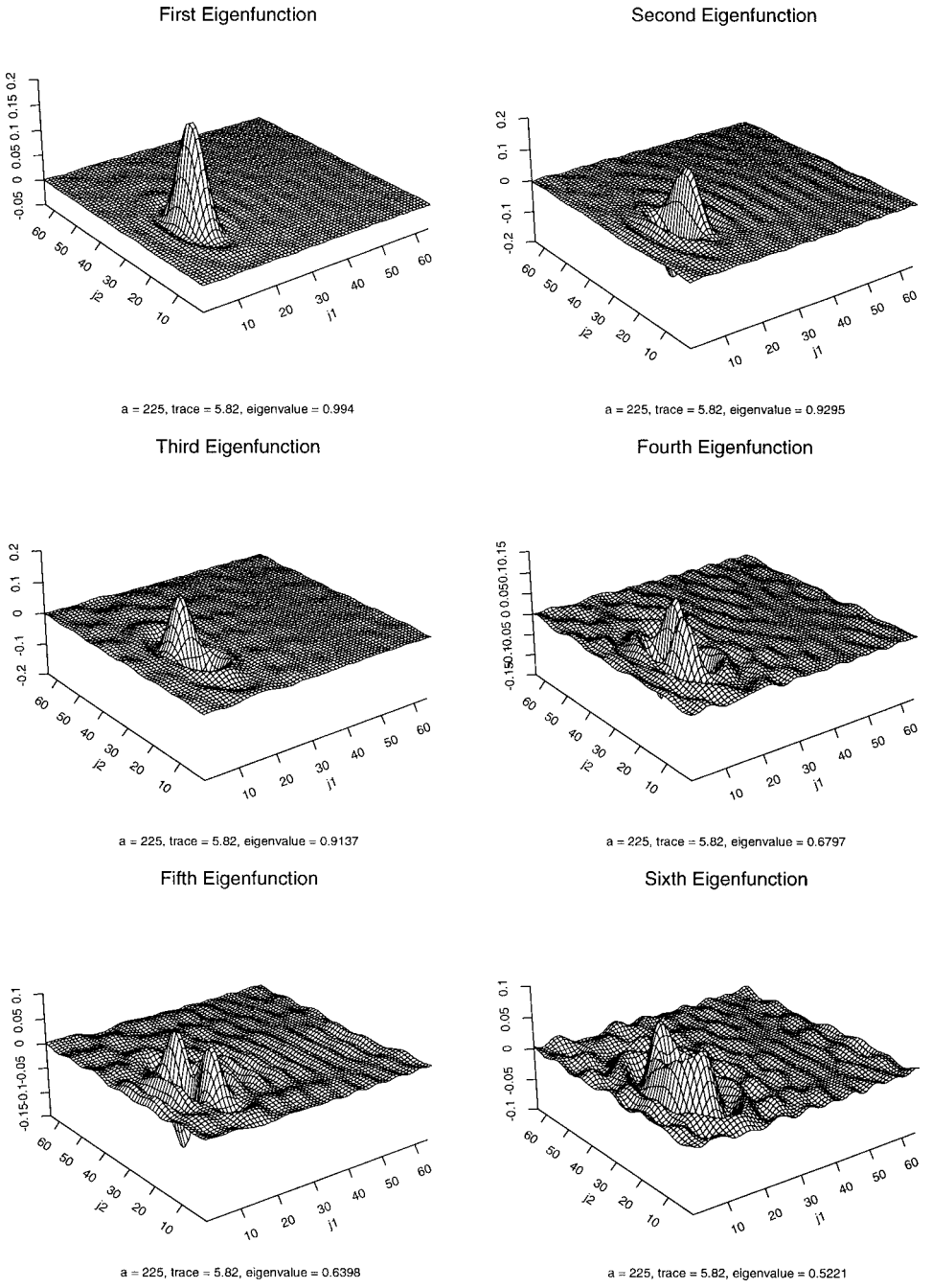

FIG. 13. First six eigenvectors for the greedy set. Some finer decomposition of the activity in B may be possible if $a = 225$.

at high time resolution. We hope our method will enable locating the place in the brain where more rapid mental processes and subprocesses take place. Our method poses an optimization problem in the theory of prolate-spheroidal functions, and we

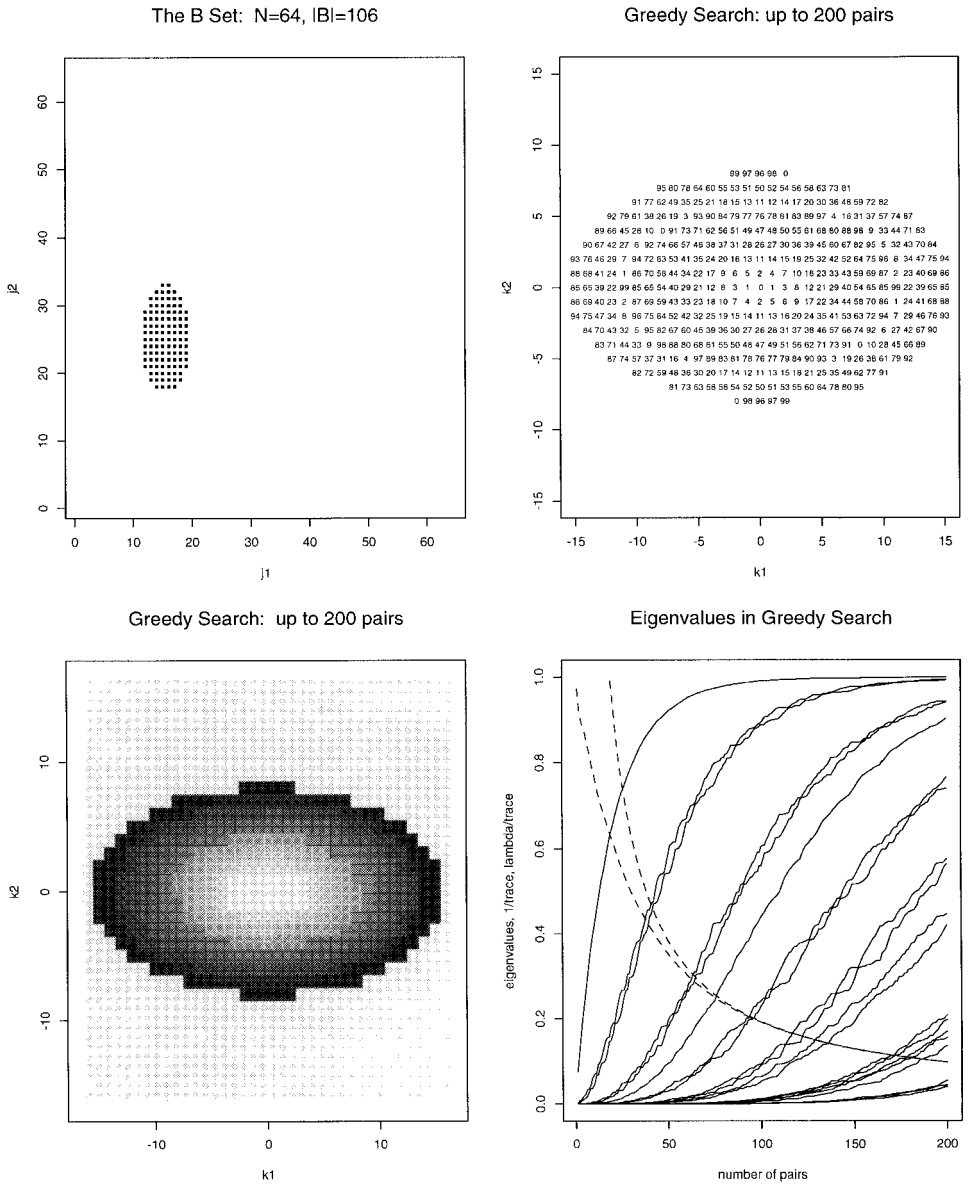

FIG. 14. The growth of greedy sets and the second eigenvalues for the second B. The upper right gives the time of entry (mod 100).

provide a heuristic approximate solution. Our computational experiments indicate that the optimization problem is likely to require exponential search in the discrete case, but the heuristic solutions are nearly optimal.

ACKNOWLEDGMENTS

We thank Gemaliel Isaacs for penetrating questions about the difference between disbursed regular sampling of the Fourier transform and the method suggested here, bordering on co-authorship. We thank Qing Yang, Hershey

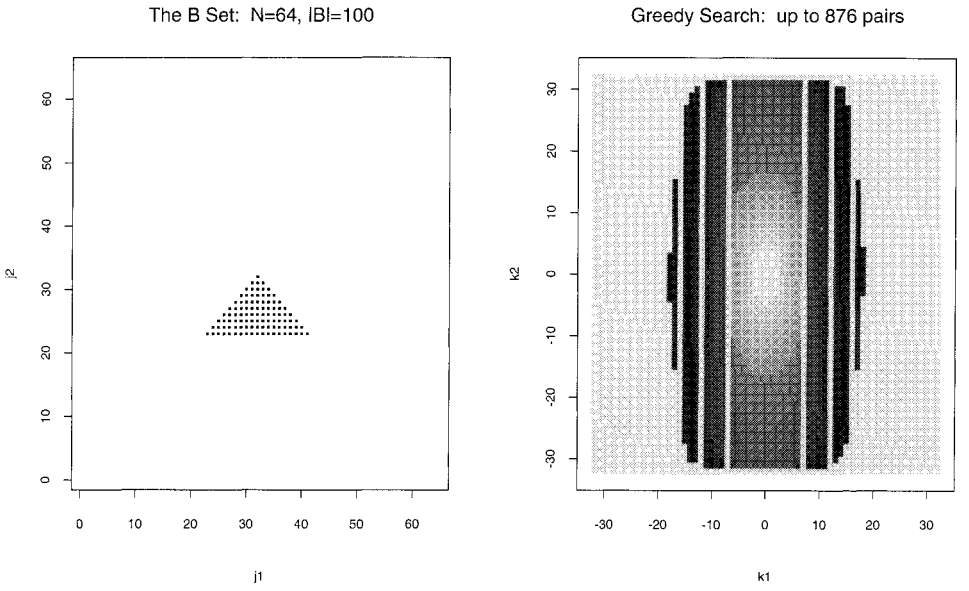

FIG. 15. The growth of greedy sets for the third B.

Medical Center, for introducing us to the problem and for stimulating discussions bordering on co-authorship. Finally, we thank Martin Farach for suggesting a way out of an impasse we had with the referee and editor of our paper. They felt strongly that the term "wavelet" already has a restricted meaning and can be used only for "functions where the support width scales inversely to the 'frequency' content" which is very different than the use we made here of this term. The term P-wavelet was coined by him to solve this problem.

REFERENCES

1. M. H. Buonocore, C. M. Wessinger, C. L. Kussmaul, G. R. Mangun, A. Jones, and M. S. Gazzaniga, MRI of auditory corex using pulsed tones, ISMRM, 1840 (1996).
2. Z. H. Cho, Y. M. Ro, and T. H. Lim, NMR venography using the susceptibility effect produced by deoxyhemoglobin, *Magn. Reson. Med.* **28** (1992), 25–38.
3. C. K. Chui, "An Introduction to Wavelets," Academic Press, New York, 1992.
4. I. Daubechies, "Wavelets," CBMS-NSF Series in Appl. Math., SIAM, Philadelphia, 1992.
5. D. L. Donoho and P. B. Stark, Uncertainty principles and signal recovery, *SIAM J. Appl. Math.* **49** (1989), 906–931.
6. D. L. Donoho and P. B. Stark, A note on rearrangements, spectral concentration, and the zero-order prolate spheroidal wavefunction, *Trans. Inform. Theory* **39** (1993), 257–260.
7. H. J. Landau and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty, II, *Bell System Tech. J.* **40** (1961), 65–84.
8. H. J. Landau and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty, III, *Bell System Tech. J.* **41** (1962), 1295–1336.
9. Y. Meyer, "Ondelettes et Opérateurs," Vols. I, II, Hermann, Paris, 1990.
10. Y. Meyer and R. R. Coifman, "Ondelettes et Opérateurs," Vols. III, Hermann, Paris, 1991.
11. R. S. Monon, S. Ogawa, D. W. Tank, and K. Ugurbil, 4 Tesla gradient recalled echo characteristics of photic stimulation-induced signal changes in the human primary visual cortex, *Magn. Reson. Med.* **30** (1993), 380–386.

12. S. Ogawa, T. M. Lee, and B. Barrere, The sensitivity of magnetic resonance image signals of a rat brain to changes in the cerebralvenous blood oxygenation, *Magn. Reson. Med.* **29** (1993), 205–210.
13. S. Ogawa, R. S. Menon, D. W. Tank, S. G. Kim, H. Merkle, J. M. Ellermann, and K. Ugurbil, Functional brain mapping 1 blood oxygenation level-dependent contrast magnetic resonance imaging: A comparison of signal characteristics with biophysic models, *Biophys. J.* **64** (1993), 803–812.
14. S. Ogawa, D. W. Tank, R. Menon, J. M. Ellermann, S. G. Kim, H. Merkle, and K. Ugurbil, Intrinsic signal changes accompanying sensory simulation: functional brain mapping with magnetic resonance imaging, *Proc. Natl. Acad. Sci. USA* **89** (1992), 5951–5955.
15. L. Shepp, Computerized tomography and nuclear magnetic resonance, *J. Comput. Tomogr.* **4** (1980), 94–107.
16. L. Shepp and B. F. Logan, The Fourier reconstruction of a head section, *IEEE Trans. Nucl. Sci.* **21** (1974), 21–43.
17. L. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography, *IEEE Trans. Med. Imaging* **1** (1982), 113–122.
18. L. Shepp and C. Zhang, Fast functional magnetic resonance imaging via space-time trade off, manuscript, 1998.
19. D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty, IV: Extensions to many dimensions; generalized prolate spheroidal functions, *Bell System Tech. J.* **43** (1964), 3009–3057.
20. D. Slepian and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty, I, *Bell System Tech. J.* **40** (1961), 43–64.
21. K. R. Thulborn, J. C. Waterton, P. M. Matthews, and G. K. Radda, Oxygenation dependence of the transverse relaxation of water protons in whole blood at high field, *Biochem. Biophys. Acta* **714** (1982), 265–270.
22. R. Turner, D. Le Bihan, C. T. Moonen, D. Despres, and J. Frank, Echo-planar time course MRI of cat brain oxygenation changes, *Magn. Reson. Med.* **22** (1991), 159–166.
23. G. A. Wright, B. S. Hu, and A. Macovski, Estimating oxygen saturation of blood *in vivo* with MR imaging at 1.5 T, *J. Magn. Reson. Imaging* **1** (1991), 275–283.