Journal of Magnetic Resonance 206 (2010) 120–126

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Sensitivity analysis of solutions of the harmonic inversion problem: Are all data points created equal?

Hasan Celik, A.J. Shaka, V.A. Mandelshtam *

Department of Chemistry, University of California, Irvine, CA 92697-2025, USA

ARTICLE INFO

Article history:
Received 20 April 2010
Revised 11 June 2010
Available online 30 June 2010

Keywords:
Spectral analysis
Super-resolution methods
Filter Diagonalization Method
FDM
Harmonic inversion problem
Spectral estimation problem
Sensitivity analysis
Non-uniform sampling

ABSTRACT

We consider the harmonic inversion problem, and the associated spectral estimation problem, both of which are key numerical problems in NMR data analysis. Under certain conditions (in particular, in exact arithmetic) these problems have unique solutions. Therefore, these solutions must not depend on the inversion algorithm, as long as it is exact in principle. In this paper, we are not concerned with the algorithmic aspects of harmonic inversion, but rather with the sensitivity of the solutions of the problem to perturbations of the time-domain data. A sensitivity analysis was performed and the counterintuitive results call into question the common assumption made in "super-resolution" methods using non-uniform data sampling, namely, that the data should be sampled more often where the time signal has the largest signal-to-noise ratio. The numerical analysis herein demonstrates that the spectral parameters (such as the peak positions and amplitudes) resulting from the solution of the harmonic inversion problem are least susceptible to the perturbations in the values of data points at the edges of the time interval and most susceptible to the perturbations in the values at intermediate times.

1. Introduction

With the growing interest in so-called "super-resolution" methods for spectral analysis of NMR data [1–10] it is important to understand the general limitations imposed by the *harmonic inversion problem* (HIP) itself, aside from specific algorithmic aspects of its numerical solution. The HIP is to fit the time-domain data as a superposition of complex sinusoids or, equivalently, to cast the NMR spectrum as a superposition of Lorentzian peaks. The Fourier transform, a linear method, treats each time point of a free induction decay (FID) $C(n\tau) \equiv c_n (n = 0,\ldots,N-1)$ on an equal footing. A small perturbation of any particular time point leads to a small perturbation in the baseline of the FT spectrum, and a predictable small change in peak positions and integrals. As such, no point in the time domain is especially important. However, in the HIP, where the spectrum is estimated from a parametric form, there is no guarantee of such uniform behavior with respect to small perturbations of particular points. The intrinsic sensitivity of the spectral estimate to errors in the input, which could be either noise or quantization errors from the analog-to-digital conversion, is important to ascertain in this case.

2. The harmonic inversion problem

Given an N-point time signal $C(n\tau) \equiv c_n (n = 0,\ldots,N-1)$, defined on an equidistant grid, consider the harmonic inversion problem

$$c_n = \sum_{k=1}^{K} d_k u_k^n, \tag{1}$$

where the two sets of spectral parameters, $u_k \equiv e^{i\tau\omega_k}$ and d_k $(k = 1,\ldots,K)$ identify each peak frequency and width, and integral and phase, respectively.

The optimization problem (1) may seem unremarkable and somewhat arbitrary. However, it is by far the most popular among all possible parametric forms. The reason for the ubiquity of the HIP lies in its unique property of having a linear algebraic solution (see for example Ref. [9]) and so allowing the consideration of a much larger parameter space than is feasible in generic non-linear optimization problems, and without the problems of local minima that can hinder other forms. That is, the HIP has a unique solution and a numerically efficient means to obtain it.

There is, however, a hidden subtlety in the formulation of the HIP that arises if the total number of terms K is *fixed* and so independent of N. This seemingly natural assumption, motivated by the underlying idea that the number of *signal* peaks is fixed once the sample for analysis is chosen, is generally flawed. The problem becomes ill-defined, making the solution extremely sensitive to the errors (noise) in the input data. Even starting with a very special case of an N-point FID consisting of exactly K damped sinusoids,

* Corresponding author. Fax: +1 949 824 8571.
E-mail addresses: hcelik@uci.edu (H. Celik), ajshaka@uci.edu (A.J. Shaka), mandelsh@uci.edu (V.A. Mandelshtam).

1090-7807/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2010.06.015

where $K < N/2$, a general infinitesimal perturbation of the input signal c_n will promptly destroy this special property, and can immediately require $K = N/2$ peaks to obtain an exact fit. A well-defined numerically stable formulation of the harmonic inversion problem is thus always to set $K = N/2$ (considering, without serious sacrifice, only even N), corresponding to having the number of unknowns consistent with the number of equations.

Assuming $N = 2K$, the fact that the HIP Eq. (1) has a unique solution is quite remarkable. This is a general result, independent of the algorithm used to numerically solve the harmonic inversion problem. In practice, however, obtaining the solution may be quite difficult. For example, the well known variant of the harmonic inversion problem, namely the multiexponential fit,

$$c_n = \sum_{k=1}^{K} d_k e^{-n\tau\lambda_k}, \tag{2}$$

where c_n, d_k and λ_k are all real, is a famous example of an appealing and easily formulated problem that typically has no meaningful numerical solution. That is, using finite arithmetic, it is generally impossible to uniquely fit a purely decaying sequence c_n by more than $K = 2$ exponentials. As more exponentials are included, the extracted parameters change abruptly, bearing little relation to those determined with fewer exponentials. These difficulties aside, the focus here is on the universal properties of the exact solution of Eq. (1) and its *sensitivity to the input data* c_n rather than the method used to obtain the exact solution of Eq. (1).

3. Solving the harmonic inversion problem

One of the simplest algorithms to solve Eq. (1) is the *Matrix-Pencil Method* (MPM); it is known in the literature under various names (see e.g. Refs. [1,2]) and is closely related to many other linear algebraic methods. Define the $M \times M$ data matrices

$$U^{(p)}_{mn} = c_{m+n+p} \tag{3}$$

with $n = 0,\ldots,M-1$, $m = 0,\ldots,M-1$, $M = N/2$ and $p = 0$ or 1. The solution of Eq. (1) is given by the generalized eigenvalue problem

$$(U^{(1)} - u_k U^{(0)})B_k = 0 \tag{4}$$

where the (column) eigenvectors B_k satisfy the orthonormality condition

$$B_k^{\mathrm{T}} U^{(0)} B_{k'} = \delta_{kk'}, \tag{5}$$

which follows from the symmetry of the data matrices, $[U^{(p)}]^{\mathrm{T}} = U^{(p)}$. The amplitudes are obtained by

$$d_k = \left(C^{\mathrm{T}} B_k\right)^2 \tag{6}$$

with $C^T = (c_0,\ldots,c_{M-1})$. It is Eq. (4) that identifies the MPM. The compact formula for the amplitudes, Eq. (6) was first given in Ref. [11]. Although with exact arithmetic Eqs. (3)–(6) give the exact solution to the HIP (1), such an approach is never practical whenever N is large. A much more practical algorithm, the Filter Diagonalization Method (FDM), is derived by using a transformation to a Fourier basis

$$\widetilde{U}^{(p)} = Z^{\mathrm{T}} U^{(p)} Z, \tag{7}$$

where Z is a rectangular $M \times K_{\mathrm{win}}$ matrix given by

$$Z_{n_j} = w_n z_j^{-n} \quad (j = 1,\ldots,K_{\mathrm{win}}, n = 0,\ldots,M-1), \tag{8}$$

and $z_j \equiv e^{2\pi i f_j}$ with real values f_j distributed uniformly within a generally small frequency interval $f_{\min} < f_j < f_{\max}$; and w_n defines a weighting function. The unusual form of the basis transformation (7) is chosen so that the transformed $K_{\mathrm{win}} \times K_{\mathrm{win}}$ matrices are also symmetric, $[\widetilde{U}^{(p)}]^{\mathrm{T}} = \widetilde{U}^{(p)}$. The resulting generalized eigenvalue problem reads

$$\left(\widetilde{U}^{(1)} - u_k \widetilde{U}^{(0)}\right)\widetilde{B}_k = 0; \quad \widetilde{B}_k^{\mathrm{T}} \widetilde{U}^{(0)} \widetilde{B}_{k'} = \delta_{kk'}, \tag{9}$$

with the new eigenvectors related to those in Eq. (4) by $B_k \approx Z\widetilde{B}_k$. The approximation holds assuming that the eigenvectors B_k can be expanded in the narrow-band Fourier basis for eigenvalues which have their real part of the frequency within the frequency window. For these frequencies, the amplitudes d_k can then be computed using

$$d_k = (C^{\mathrm{T}} Z \widetilde{B}_k)^2. \tag{10}$$

For the special case of $w_n = 1$ [11], the $\widetilde{U}^{(p)}$ matrices can be computed efficiently using

$$\widetilde{U}^{(p)}_{jj'} = (z_j - z_{j'})^{-1}\left[z_j \sum_{n=0}^{M-1} z_{j'}^{-n} c_{n+p} - z_{j'} \sum_{n=0}^{M-1} z_j^{-n} c_{n+p} - z_j^{1-M} \sum_{n=M}^{2M-2} z_{j'}^{M-n} c_{n+p} + z_{j'}^{1-M} \sum_{n=M}^{2M-2} z_j^{M-n} c_{n+p}\right]. \tag{11}$$

For $z_j = z_{j'}$ a numerically practical expression can be obtained after taking the $z_j \to z_{j'}$ limit:

$$\widetilde{U}^{(p)}_{jj} = \sum_{n=0}^{2M-2} (M - |M - n - 1|) c_{n+p} z_j^{-n}. \tag{12}$$

Generally the spectral parameters computed by FDM in a small frequency window using K_{win} basis functions are as accurate as those computed by MPM using $K = N/2$ basis functions over the entire spectral width. The tremendous practical advantage of FDM is that it scales quasi-linearly with N versus the $\sim N^3$ scaling of MPM. One reason for this property of FDM is that the Fourier-transformed matrices $\widetilde{U}^{(p)}$ are diagonally dominant, with $\widetilde{U}^{(1)}_{jj}/\widetilde{U}^{(0)}_{jj}$ being already a good approximation for the eigenvalues u_j. Given this property of FDM, the possibly large spectral width $[-1/2\tau, 1/2\tau]$ can be covered with small windows, and the whole "line-list" (d_k, ω_k) can be constructed by collecting the results from all the small-window calculations. However, one can always take $K_{\mathrm{win}} = K = N/2$ and $[f_{\min}, f_{\max}] = [-1/2\tau, 1/2\tau]$, in other words a complete Fourier basis, in which case the transformation matrix Z is unitary and with exact arithmetic the resulting spectral parameters using either Eq. (4) or Eq. (9) would be indistinguishable. It is also important to note that all other "correct" linear algebraic algorithms, such as the Prony method, *Linear Prediction* (LP), etc., should yield the same solution in exact arithmetic.

4. Sensitivity analysis

The parameters extracted from the linear algebraic fit can be in error when the data is noisy. The sensitivity of the output of the method to perturbations in the input can be investigated by the classical error, or sensitivity, analysis. In this section, we discuss how the solution of the HIP (1) depends on perturbation of the input values c_t. For example, the spectrum $S(f)$ computed from the parameters extracted from the solution of the HIP or estimated directly (see Appendix B) is a functional of $\{c_t\}$ ($t = 0,\ldots,N-1$). Given a finite variation, δc_t, of a single data point, c_t, the spectral sensitivity can be defined

$$E_t[S(f)] = \frac{1}{S(f)}\left|\frac{\delta S(f)}{\delta c_t}\right| \quad (t = 0,\ldots,N-1). \tag{13}$$

For small enough variations, δc_t, we can approximate the spectral sensitivity by the spectral derivative

$$E_t[S(f)] \approx \left|\frac{1}{S(f)}\frac{\partial S(f)}{\partial c_t}\right|. \tag{14}$$

(For large variations the sensitivity is a function of δc_t.) The spectral sensitivity itself is of limited utility, however, as small errors in the positions of narrow peaks would register as large spectral deviations, even though they might not be considered serious errors from a spectroscopic viewpoint. Therefore, even when considering the spectrum, a more convenient sensitivity measure must be associated with the "line-list" corresponding to the spectrum. For example, consider a narrow peak located at frequency f_k with amplitude $S_k = S(f_k)$, and assume that $S(f)$ has a maximum at $f = f_k$ (the peak center and peak amplitude may be defined differently though). Both f_k and S_k are then functions of $c_t (t = 0,\ldots,N-1)$, and we can define the peak sensitivities by

$$E_t[f_k] := \left|\frac{\delta f_k}{\delta c_t}\right|; \quad E_t[S_k] := \left|\frac{1}{S_k}\frac{\delta S_k}{\delta c_t}\right|. \tag{15}$$

The particularly interesting case corresponds to a spectrum with dominant features that are Lorentzian lines, i.e., when the HIP (1) is well-posed and the line-list is given by the set of complex frequencies ω_k and amplitudes d_k. The corresponding sensitivity measures can then be written

$$E_t[\omega_k] := \left|\frac{\delta\omega_k}{\delta c_t}\right|; \qquad E_t[d_k] := \left|\frac{1}{d_k}\frac{\delta d_k}{\delta c_t}\right| \quad (t = 0,\ldots,N-1). \tag{16}$$

Again, for small variations, δc_t, we have

$$E_t[\omega_k] \approx \left|\frac{\partial\omega_k}{\partial c_t}\right| = \left|\frac{1}{\tau u_k}\frac{\partial u_k}{\partial c_t}\right|; \quad E_t[d_k] \approx \left|\frac{1}{d_k}\frac{\partial d_k}{\partial c_t}\right|. \tag{17}$$

While for given variations δc_t calculation of the exact values of $E_t[\omega_k]$ may be time consuming, calculation of the derivatives, $\partial u_k/\partial c_t$ and $\partial d_k/\partial c_t$, is inexpensive within MPM (or FDM). The eigenvalue derivatives can be expressed as a variant of the Hellman–Feynman theorem

$$\frac{\partial u_k}{\partial c_t} = B_k^{\mathrm{T}}\left(\frac{\partial U^{(1)}}{\partial c_t} - u_k\frac{\partial U^{(0)}}{\partial c_t}\right)B_k, \tag{18}$$

and the eigenvector derivatives as

$$\frac{\partial B_k}{\partial c_t} = \sum_{m\neq k}\frac{B_m}{u_k-u_m}B_m^{\mathrm{T}}\left(\frac{\partial U^{(1)}}{\partial c_t} - u_k\frac{\partial U^{(0)}}{\partial c_t}\right)B_k - \frac{B_k}{2}B_k^{\mathrm{T}}\frac{\partial U^{(0)}}{\partial c_t}B_k. \tag{19}$$

(See Appendix A for derivations.) In the MPM the derivatives of the data matrices are

$$\left[\frac{\partial U^{(p)}}{\partial c_t}\right]_{mn} = \delta_{t,(m+n+p)}, \tag{20}$$

and therefore

$$\frac{\partial u_k}{\partial c_t} = \sum_{n=\max(0,t-M)}^{\min(M-1,t-1)} B_{kn}B_{k,(t-n-1)} - u_k\sum_{n=\max(0,t-M-1)}^{\min(M-1,t)} B_{kn}B_{k,(t-n)}, \tag{21}$$

where whenever $m < 0$ or $m > M-1$ we assume that $B_{km} = 0$. Using the result of Eq. (19) the amplitude derivative in the MPM can be given as

$$\begin{aligned}\frac{\partial d_k}{\partial c_t} &= \frac{C^{\mathrm{T}}B_k}{2}\left[C^{\mathrm{T}}\frac{\partial B_k}{\partial c_t} + \frac{\partial C^{\mathrm{T}}}{\partial c_t}B_k\right]\\ &= \frac{\sqrt{d_k}}{2}\left[\sum_{m\neq k}\frac{\sqrt{d_m}}{u_k-u_m}B_m^{\mathrm{T}}\left(\frac{\partial U^{(1)}}{\partial c_t} - u_k\frac{\partial U^{(0)}}{\partial c_t}\right)B_k - \frac{\sqrt{d_k}}{2}B_k^{\mathrm{T}}\frac{\partial U^{(0)}}{\partial c_t}B_k + B_{kt}\right]\\ &= \frac{\sqrt{d_k}}{2}\left\{\sum_{m\neq k}\frac{\sqrt{d_m}}{u_k-u_m}\left[\sum_{n=\max(0,t-M)}^{\min(M-1,t-1)} B_{mn}B_{k,(t-n-1)} - u_k\sum_{n=\max(0,t-M-1)}^{\min(M-1,t)} B_{mn}B_{k,(t-n)}\right]\right.\\ &\quad\left.\times\frac{\sqrt{d_k}}{2}\sum_{n=\max(0,t-M)}^{\min(M-1,t)} B_{kn}B_{k,(t-n)} + B_{kt}\right\}.\end{aligned} \tag{22}$$

Now consider a rectangular Fourier basis as defined in Eq. (8) but with $w_n = 1$. We have

$$\begin{aligned}\frac{\partial\widetilde{U}_{jj'}^{(p)}}{\partial c_t} &= \begin{cases}\left(z_j z_{j'}^{p-t} - z_{j'}z_j^{p-t}\right)(z_j - z_{j'})^{-1}, & p\leqslant t\leqslant M+p-1,\\ \left(z_{j'}^{1-M}z_j^{M+p-t} - z_j^{1-M}z_{j'}^{M+p-t}\right)(z_j - z')^{-1}, & M+p\leqslant t\leqslant 2M+p-2,\\ 0, \quad \text{otherwise},\end{cases}\\ \frac{\partial\widetilde{U}_{jj}^{(p)}}{\partial c_t} &= (M - |M+p-t-1|)z_j^{p-t}, \quad p\leqslant t\leqslant 2M+p-2.\end{aligned} \tag{23}$$

The diagonal elements of the derivative matrix scale with t as $\sim(M - |M+p-t-1|)$. This suggests that on average the eigenvalues are most sensitive to elements c_t with $t \sim M = N/2$, the sensitivity decaying linearly towards the beginning ($t \to 0$) and the tail ($t \to 2M-1 = N-1$) of the signal array. This sensitivity to errors in individual data entries is the central result of this paper. Data points in the center of the acquisition time window are more important than those at the head or tail of the data. In the two-dimensional case there is a direct product dependence of the sensitivity in each dimension, resulting in an extraordinary weighting of points in the center of both time domains. This result does not depend on the particular method used to obtain the numerical answer, but rather is an intrinsic feature of the HIP itself. Mathematically, as shown in Eq. (A3), the eigenvalue derivative is invariant under any basis transformation, making the non-uniform dependence of $\partial u_k/\partial c_t$ on t a general result that should show up in any inversion algorithm that solves the HIP.

5. A numerical example

A model time signal similar to that introduced earlier [7] was used to investigate the predicted non-uniform sensitivity. The spectrum (see Fig. 1) consists of triplets with couplings, line widths, chemical shift differences, and integrals all decreasing geometrically, to supply a full range of resolution tests in one go. The time-domain data was generated using the formula

$$\begin{aligned}c_n = 10\sum_{m=0}^{9}\{&\exp\left[in2\pi\tau((450+i)\cdot 0.75^m - 250)\right]\\ &+\exp\left[in2\pi\tau((460+i)\cdot 0.75^m - 250)\right]\\ &+\exp\left[in2\pi\tau((470+i)\cdot 0.75^m - 250)\right]\}\cdot 0.75^m,\end{aligned} \tag{24}$$

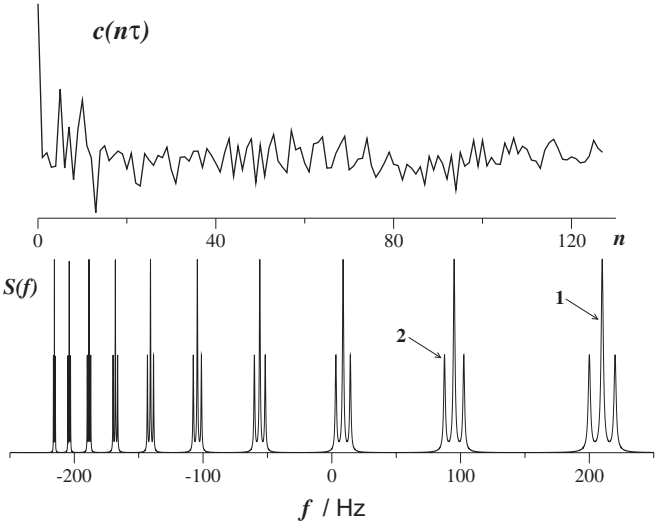

Fig. 1. The real part of the synthesized time signal $c_n = c(n\tau)$ (cf. Eq. (24)) used here for numerical demonstration and its exact infinite-time FT spectrum $S(f)$.

where $\tau = 0.002$ s, a 500 Hz spectral width. The sensitivity analysis focused on two peaks located at $f_1 = 210$ Hz and $f_2 = 87.5$ Hz (peaks 1 and 2 in Fig. 1).

We note that there are several possible ways to create the spectral representation from the parameters. It can either be created *exactly* using *Regularized Resolvent Transform* (RRT), or the FDM spectrum using (B8) (see Appendix B for further details). The differences between the RRT (B5) and FDM (B8) spectra were insignificant, so all the reported results are obtained by FDM. Also, to streamline the analysis we only present results corresponding to $N = 128$ and $K_{\text{win}} = 15$. These results are representative of many calculations that were tried. Two types of calculations were performed, labeled *Type A* and *Type B*.

Type A. Gaussian noise ξ_n,

$$c_n \to c_n + \xi_n, \tag{25}$$

was added to the model signal such that

$$\sum_n |c_n|^2 = 10^3 \cdot \sum_n |\xi_n|^2.$$

The noisy signal was then processed yielding the spectral parameters ω_k and d_k and the spectrum $S(f)$, together with the derivatives of the spectral parameters with respect to c_t, such as $\partial\omega_k/\partial c_t$. Next, for each value of $t = 0,\ldots,N-1$ each data point, c_t, is modified one at a time

$$c_t \to c_t + \delta c_t,$$

using $\delta c_t = |\delta c_t| \cdot e^{i\phi_t}$ with the random phase $\phi_t \in [0; 2\pi]$. The sensitivities, $E_t[f_k]$, $E_t[S_k]$, $E_t[\omega_k]$ and $E_t[d_k]$, as defined by Eqs. (15) and (16), were then computed using the new parameter list and the new spectrum.

Type B. The *Type A* procedure was repeated 1000 times using different realizations of noise ξ_n and random phases ϕ_t, and the results for the sensitivities were averaged.

Our tests showed that as long as the single-point perturbations $|\delta c_t|$ were of the order of noise, ξ_n, or less, that there was no difference between the actual sensitivities and the corresponding derivatives, that is, Eqs. (14) and (17) hold. Only when $|\delta c_t| \gg 0.1$ were differences noticeable, in any event they do not lead to qualitatively different conclusions. Thus, for simplicity, we only report results using $|\delta c_t| = 0.1$.

Fig. 2 shows the sensitivity of ω_1 (peak 1) for two different realizations of noise ξ_n. As can be seen the t-dependencies of $E_t[\omega_1]$ are quite erratic and are different for the two cases. However, the sensitivity obtained by averaging over 1000 different realizations of noise has a remarkable shape with the maximum

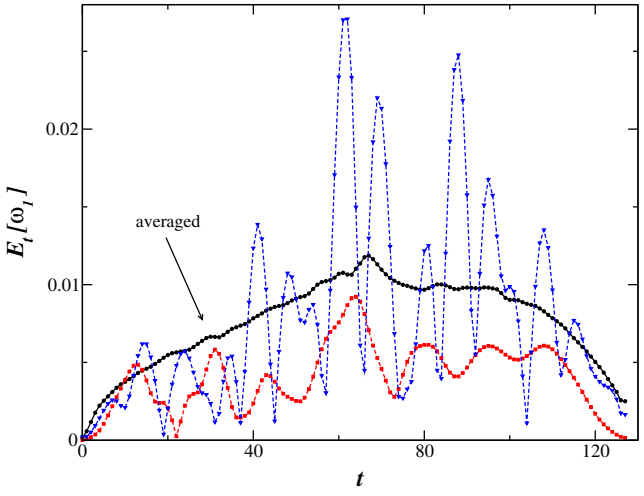

Fig. 2. Sensitivity of the complex frequency ω_1 of peak 1 (as in Fig. 1) for two realizations of noise (see text) and that averaged over 1000 different noise realizations.

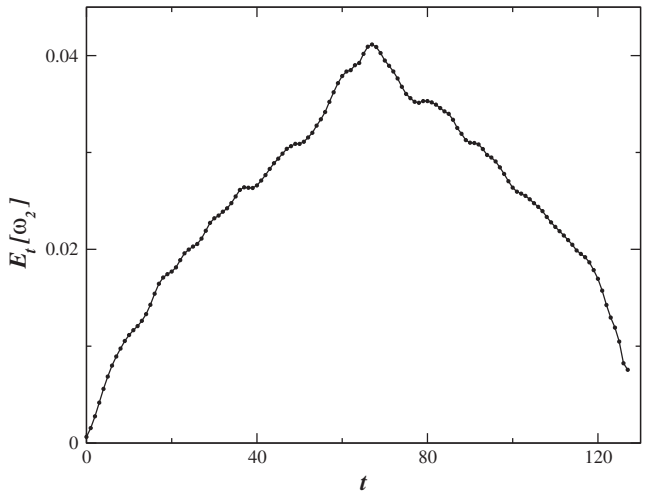

Fig. 3. Sensitivity of the complex frequency ω_2 of peak 2 averaged over 1000 different noise realizations.

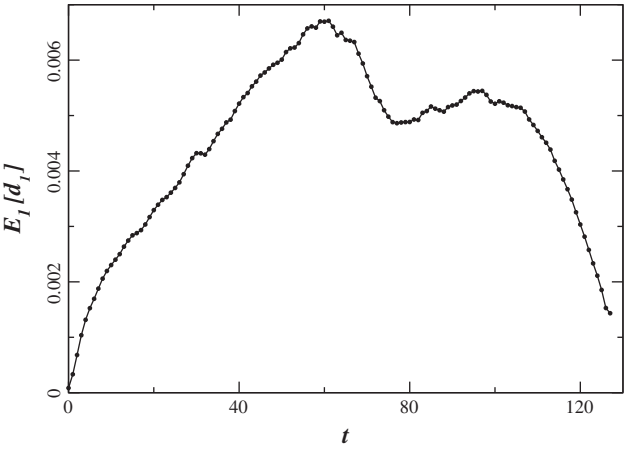

Fig. 4. Averaged sensitivity of the complex amplitude d_1 of peak 1.

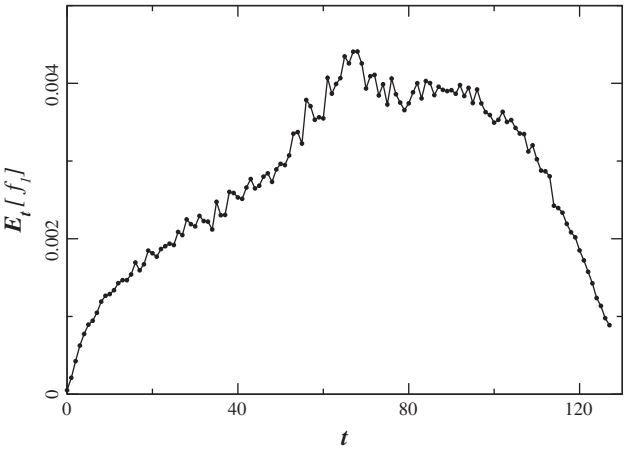

Fig. 5. Averaged sensitivity for the position, f_1, of the maximum of peak 1.

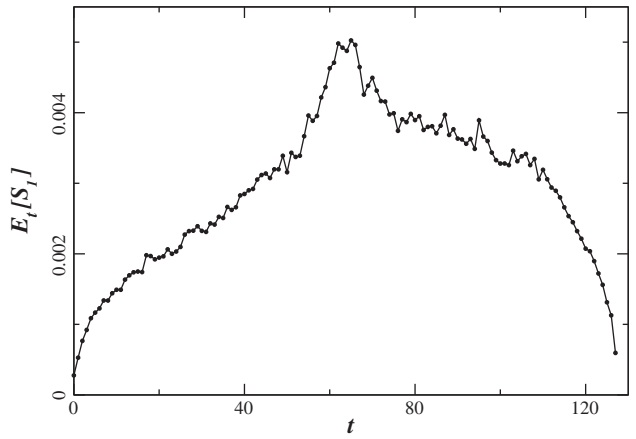

Fig. 6. Averaged sensitivity of the maximum value S_1 of peak 1.

at $t \sim N/2$ decreasing towards the edges of the time interval. For peak 2, the sensitivity function $E_t[\omega_2]$ shown in Fig. 3 has a similar shape, although multiplied by a factor of $\sim$4. The spectral parameters of peak 2 are thus less accurate and more sensitive. This is to be expected, as peak 2 is smaller and belongs to the triplet with closer lines.

Figs. 4–6 show averaged sensitivities of other spectral parameters for peak 1. The t-dependencies of the corresponding functions, $E_t(d_1)$, $E_t(f_1)$ and $E_t(S_1)$, have shapes similar to that of $E_t(\omega_1)$.

6. Discussion

This investigation focused on the HIP and the spectral estimation problems, which are key numerical problems in the super-resolution FT spectroscopy (such as NMR), and appear in many other related fields in science and engineering. A closer investigation of the sensitivity of the spectral estimate to perturbations of individual time-domain data points, clearly showed that the sensitivity follows a rather unusual trend. The spectral estimate was more severely affected by perturbations in the center of the specified acquisition time than similar perturbations at either the beginning or end of the data. Also, the number of times each data point c_t appears in the data matrices mirror a similar trend $(M - |M + p - t - 1|)$, namely a triangular function with an apex in the middle of the time interval $t = M$. Rather naively, one can associate this distribution of the data points (3) in the Hankel matrices with the bias in the sensitivity in order to understand the unusual results.

Clearly, these results may cause one to modify acquisition strategies, as obtaining superior signal-to-noise for the central points may have more effect on the accuracy of the spectrum than a corresponding improvement at the beginning or end of the data. This aspect would probably only be relevant in multidimensional NMR, where different numbers of transients could be acquired (and each increment then appropriately scaled). However, perhaps other interferometric methods like FT-IR could benefit by non-uniform signal averaging. In addition, the sensitivity analysis shows that very long acquisition times, in which most the signal intensity may have decayed, could pose problems for any method that tries to solve the harmonic inversion problem. In this case the fidelity of the central time points may not be high enough, due to noise contamination, to allow accurate spectral estimation. In other words, the "super-resolution" methods may not offer any improvement in resolution, but could instead become "super-sensitive" to small errors.

While FDM requires a uniform time grid to take advantage of the Fourier basis, so-called non-uniform sampling (NUS) strategies can operate by deleting acquisition of certain data points, by skipping increments or taking random time intervals within some sampling schedule. Some schedules, like exponential sampling, seem to have been justified by persuasive but *ad hoc* arguments, such as those based on the assumption that more data points should be taken at times where the signal is generally big, and fewer points should be taken at longer times where the signal is smaller from its decay. While the present paper does not completely exclude this argument, it presents results that are somewhat at odds with it. The true situation for NUS would be clarified by a proper sensitivity analysis, such as that undertaken here for the HIP.

7. Conclusions

The HIP is a key numerical problem across a wide range of disciplines, and is widely used in different contexts to try to obtain more information from a given data set than allowed by the Fourier transform uncertainty principle, or, equivalently, to speed up data acquisition by allowing shorter acquisition times. Surprisingly, the weight of a data point, as measured by the number of times it appears in the data matrices, translates into different sensitivity (in the mean) for parameters of the method. The present work introduced a straightforward approach to analyze the sensitivity of the solutions of the harmonic inversion (or spectral estimation) problems. This approach was implemented within FDM, which is an efficient, but particular method of solving these problems. It was argued that as long as the spectral analysis method is associated with the HIP, the sensitivity of the solution will behave universally as a function of time. In particular, the solution will be most sensitive to the data points in the middle of the time interval and least sensitive at its edges. This counterintuitive fact may be exploited, for example, by optimizing the sampling schedule accordingly to minimize the total experimental time for a given resolution requirement. However, additional studies are certainly needed to analyze other spectral inversion techniques. The new sensitivity theory proposed in this paper is well suited for such analyses.

Acknowledgments

VAM acknowledges NSF support, Grant CHE-0809108. AJS acknowledges NSF Grant CHE-0703182.

Appendix A. Eigenvalue and eigenvector derivatives

By multiplying the generalized eigenvalue problem Eq. (4) from the left by B_k^{T} and differentiating it with respect to c_t we have

$$\frac{\partial}{\partial c_t} B_k^{\mathrm{T}} (U^{(1)} - u_k U^{(0)}) B_k = 0.$$

Then using

$$0 = \frac{\partial}{\partial c_t} B_k^{\mathrm{T}} U^{(0)} B_k = \frac{\partial B_k^{\mathrm{T}}}{\partial c_t} U^{(0)} B_k + B_k^{\mathrm{T}} U^{(0)} \frac{\partial B_k}{\partial c_t} + B_k^{\mathrm{T}} \frac{\partial U^{(0)}}{\partial c_t} B_k,$$

we obtain a variant of the Hellman–Feynman theorem:

$$\frac{\partial u_k}{\partial c_t} = B_k^{\mathrm{T}} \left(\frac{\partial U^{(1)}}{\partial c_t} - u_k \frac{\partial U^{(0)}}{\partial c_t} \right) B_k, \tag{A1}$$

which provides the eigenvalue derivatives essentially without any extra numerical effort.

Besides its numerical convenience, the result given by Eq. (A1) is quite remarkable, in particular, because in infinite precision arithmetic the eigenvalue derivative is invariant under the basis

transformation, as long as the latter is not rank-deficient. Namely, consider square non-singular matrix Z, with $\det(Z) \neq 0$, so that the inverse Z^{-1} exists.

If Z does not depend on c_t we have

$$\frac{\partial \widetilde{U}^{(p)}}{\partial c_t} = Z^{\mathrm{T}} \frac{\partial U^{(p)}}{\partial c_t} Z. \tag{A2}$$

Eq. (A1) is then rewritten as

$$\begin{aligned}\frac{\partial u_k}{\partial c_t} &= \widetilde{B}_k^{\mathrm{T}} \left(\frac{\partial \widetilde{U}^{(1)}}{\partial c_t} - u_k \frac{\partial \widetilde{U}^{(0)}}{\partial c_t} \right) \widetilde{B}_k \\ &= \left(Z^{-1} B_k \right)^{\mathrm{T}} \left(Z^{\mathrm{T}} \frac{\partial U^{(1)}}{\partial c_t} Z - u_k Z^{\mathrm{T}} \frac{\partial U^{(0)}}{\partial c_t} Z \right) \\ &\quad \times Z^{-1} B_k B_k^{\mathrm{T}} \left(\frac{\partial U^{(1)}}{\partial c_t} - u_k \frac{\partial U^{(0)}}{\partial c_t} \right) B_k, \end{aligned} \tag{A3}$$

which proves that $\partial u_k / \partial c_t$ is invariant under the basis transformation.

To obtain the derivative of the eigenvector B_k we first expand it into the eigenbasis by writing

$$\frac{\partial B_k}{\partial c_t} = \sum_m x_m B_m, \tag{A4}$$

with unknown coefficients x_m, which can be obtained by differentiating the orthonormality condition (5):

$$\begin{aligned} 0 &= 2 B_k^{\mathrm{T}} U^{(0)} \frac{\partial B_k}{\partial c_t} + B_k^{\mathrm{T}} \frac{\partial U^{(0)}}{\partial c_t} B_k = 2 B_k^{\mathrm{T}} U^{(0)} \sum_m x_m B_m + B_k^{\mathrm{T}} \frac{\partial U^{(0)}}{\partial c_t} B_k \\ &= 2 x_k + B_k^{\mathrm{T}} \frac{\partial U^{(0)}}{\partial c_t} B_k. \end{aligned}$$

Differentiating Eq. (4),

$$\begin{aligned} \left(\frac{\partial U^{(1)}}{\partial c_t} - \frac{\partial u_k}{\partial c_t} U^{(0)} - u_k \frac{\partial U^{(0)}}{\partial c_t} \right) B_k &= \left(u_k U^{(0)} - U^{(1)} \right) \frac{\partial B_k}{\partial c_t} \\ &= \left(u_k U^{(0)} - U^{(1)} \right) \sum_m x_m B_m, \end{aligned}$$

and multiplying both sides by B_n^{T} $(n \neq k)$ and also using Eqs. (A1) and (5) we get

$$B_n^{\mathrm{T}} \times \ \mathrm{RHS} = B_n^{\mathrm{T}} (u_k - u_n) \sum_m x_m B_m = (u_k - u_n) x_n,$$

$$B_n^{\mathrm{T}} \times \ \mathrm{LHS} = B_n^{\mathrm{T}} \left(\frac{\partial U^{(1)}}{\partial c_t} - u_k \frac{\partial U^{(0)}}{\partial c_t} \right) B_k.$$

Combining the above results we obtain

$$\frac{\partial B_k}{\partial c_t} = \sum_{m \neq k} \frac{B_m}{u_k - u_m} B_m^{\mathrm{T}} \left(\frac{\partial U^{(1)}}{\partial c_t} - u_k \frac{\partial U^{(0)}}{\partial c_t} \right) B_k - \frac{B_k}{2} B_k^{\mathrm{T}} \frac{\partial U^{(0)}}{\partial c_t} B_k. \tag{A5}$$

By analogy with Eq. (A3) it can be shown that for a square non-singular matrix Z the eigenvector derivative is unchanged:

$$\frac{\partial B_k}{\partial c_t} = \frac{\partial \widetilde{B}_k}{\partial c_t}. \tag{A6}$$

Appendix B. The spectral estimation problem and Regularized Resolvent Transform

In this appendix we consider a problem related to the harmonic inversion problem, namely, given a finite time signal $c(n\tau) \equiv c_n (n = 0, \ldots, N-1)$, estimate its infinite-time Discrete Fourier Transform (DFT) spectrum $S(f)$

$$S(f) = -\frac{\tau c_0}{2} + \tau \sum_{n=0}^{\infty} e^{-i2\pi n \tau f} c_n \equiv -\frac{\tau c_0}{2} + \tau \sum_{n=0}^{\infty} z^{-n} c_n, \tag{B1}$$

where $z := e^{i2\pi\tau f}$.

If the signal c_n satisfies the harmonic form (1) with $2K \leqslant N$ (in exact arithmetic) the spectrum $S(f)$ can be computed *exactly* by the following expression (*Resolvent Transform*) [8]

$$S(f) = -\frac{\tau c_0}{2} + \tau C^{\mathrm{T}} (U^{(0)} - U^{(1)}/z)^{-1} C, \tag{B2}$$

or using the Fourier basis

$$S(f) = -\frac{\tau c_0}{2} + \tau \widetilde{C}^{\mathrm{T}} \left(\widetilde{U}^{(0)} - \widetilde{U}^{(1)}/z \right)^{-1} \widetilde{C}. \tag{B3}$$

Both Eqs. (B3) and (B7) generally require solution of an ill-conditioned linear system,

$$R(f) X(f) = C, \tag{B4}$$

with

$$R(f) = U^{(0)} - U^{(1)}/z,$$

and

$$S(f) = C^{\mathrm{T}} X(f). \tag{B5}$$

The ill-conditioned linear system can be solved using the regularization:

$$(R^{\dagger}(f) R(f) + q^2) X(f) = R^{\dagger}(f) C, \tag{B6}$$

$$I(f) = C^{\mathrm{T}} \left(R^{\dagger}(f) R(f) + q^2 \right)^{-1} R^{\dagger}(f) C, \tag{B7}$$

$$I(f) = \widetilde{C}^{\mathrm{T}} \left(\widetilde{R}^{\dagger}(f) \widetilde{R}(f) + q^2 \right)^{-1} \widetilde{R}^{\dagger}(f) \widetilde{C},$$

where q is the regularization parameter that controls the condition number of the initially ill-conditioned matrix $R^{\dagger}R$.

We note here that unless heavy regularization (in the case of highly ill-posed problem) is applied to solve Eq. (B4), the RRT spectrum $S(f)$ is indistinguishable from the FDM spectrum computed by

$$S(f) = -\frac{\tau c_0}{2} + \tau \sum_k d_k (1 - u_k/z)^{-1}. \tag{B8}$$

An expression for the spectral sensitivity Eq. (13) in terms of the spectral derivative, as defined by Eq. (14), can be obtained by differentiating Eq. (B7):

$$\frac{\partial}{\partial c_t} S(f) = C^{\mathrm{T}} \frac{\partial X}{\partial c_t} + X^{\mathrm{T}} \frac{\partial C}{\partial c_t},$$

where $X^{\mathrm{T}} \partial C / \partial c_t = X_t$ and

$$\frac{\partial X}{\partial c_t} = R^{-1} \frac{\partial R}{\partial c_t} R^{-1} C,$$

which after some algebra gives

$$\frac{\partial}{\partial c_t} S(f) = X_t + \sum_{n=\max(0,t-M)}^{\min(M-1,t)} X_n \left[X_{t-n} - z^{-1} X_{t-n-1} \right]. \tag{B9}$$

The first term in Eq. (B9) has a simple dependence on t, but the second term depends on t nonuniformly, in a fashion similar to that of the eigenvalue derivatives Eq. (21).

References

[1] R. Roy, A. Paulraj, T. Kailath, IEEE Trans. Acoust. Speech Signal Process. ASSP-34 (1986) 1340.
[2] Y. Hua, T.K. Sarkar, IEEE Trans. Acoust. Speech Signal Process. ASSP-38 (1990) 814;
Y. Hua, T.K. Sarkar, IEEE Trans. Signal Process. 39 (1991) 892.
[3] G.L. Bretthorst, J. Magn. Reson. 88 (1990) 533.
[4] R.A. Chylla, J.L. Markley, J. Biomol. NMR 5 (1995) 245.

[5] J.C. Hoch, A.S. Stern, in: D.M. Grant, R.K. Harris (Eds.), Encyclopedia of Nuclear Magnetic Resonance, John Wiley, London, 1996, pp. 2980–2988.
[6] P. Koehl, Prog. NMR Spectrosc. 34 (1999) 257.
[7] H. Hu, Q.N. Van, V.A. Mandelshtam, A.J. Shaka, J. Magn. Reson. 134 (1998) 76.
[8] J. Chen, A.J. Shaka, V.A. Mandelshtam, J. Magn. Reson. 147 (2000) 129.
[9] V.A. Mandelshtam, Progress in NMR Spectroscopy 38 (2001) 159.
[10] V.Y. Orekhov, I.V. Ibraghimov, M. Billeter, J. Biomol. NMR 20 (2001) 49.
[11] V.A. Mandelshtam, H.S. Taylor, J. Chem. Phys. 107 (1997) 6756.