International Journal of Mass Spectrometry 373 (2014) 1–14

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

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

The filter diagonalization method and its assessment for Fourier transform mass spectrometry

Beau R. Martini[a], Konstantin Aizikov[b], Vladimir A. Mandelshtam[a,*]

[a] Chemistry Department, University of California at Irvine, Irvine, CA 92697, USA
[b] Life Science Mass Spectrometry Scientific Instrumentation, Thermo Fisher Scientific (Bremen), GmbH Hanna-Kunath-Str. 11, 28199 Bremen, Germany

ARTICLE INFO

Article history:
Received 3 May 2014
Received in revised form 5 August 2014
Accepted 5 August 2014
Available online 23 August 2014

Keywords:
Fourier transform mass spectrometry
FTMS
Filter diagonalization method
FDM
Super resolution

ABSTRACT

Application of the filter diagonalization method (FDM) to Fourier transform mass spectrometry (FTMS) data is not new. Under certain conditions FDM provides resolution superior to Fourier transform (FT) and was proved to be useful in investigation of space charge phenomena in an ion cyclotron resonance cell (ICR) by O'Connor and Amster research groups. Kozhinov and Tsybin have reported substantial increase in resolution and/or acquisition speed of high-resolution molecular and macromolecular MS data. In light of fundamental difficulty in providing theoretical evaluation of the FDM performance under various spectral and noise conditions, this paper is an empirical investigation aimed at establishing the method's true potentials and areas where it may perform better than currently used technologies. The study was conducted on both synthetic transients and experimental Orbitrap transients. Unlike FT, resolution of FDM depends strongly on noise levels. Consequently, we identify the regimes at which FDM can provide a superior resolution even at moderate signal to noise ratios. Moreover, when individual peaks fail to be resolved either because of the small peak separation or high noise conditions, the FDM solution seems to preserve their cumulative intensity. This preservation of the true intensity seems to be very consistent across rather wide ranges of noise conditions and almost impervious to the peak separation.

1. Introduction

In the field of biological mass spectrometry, Fourier transform mass spectrometry (FTMS) [1] has established itself as a leader in mass accuracy and resolving power. This performance is made possible by means of observation of the motion of ionic species contained by an electric field in the case of Orbitrap, or by a combination of electric and magnetic fields in a Fourier Transform Ion Cyclotron Mass Spectrometer (FTICRMS), in which case their flight path may reach kilometers in length for a typical FTMS experiment. In a typical FTICR [2–5] experiment, trapped ions are excited to their respective cyclotron orbits. The coherent motion of ions in Orbitrap [6–10] is achieved by pulsed injection into an electrostatic trap. The image current induced by the moving ions is detected, digitized, and recorded for further processing. Typically, transient signals are zero padded, apodized, and converted to frequency domain via Fourier transform [11]. Even though FT is a powerful and well understood technique, which is both numerically stable and fast, it has drawbacks. Aside from other artifacts [12], inability to resolve ionic species in the frequency domain beyond the FT uncertainty [13] is of particular concern, especially when interference causes errors in calculation of the abundance of the resulting peaks. Given the intrinsic smoothness of the phase function of FTMS signals, a noticeable improvement in resolution is achieved by representing mass spectra in absorption mode [14–30]. Although improved, resolution of absorption spectra is still limited by the FT uncertainty and noticeable errors in abundance are caused by interference. A number of super-resolution (i.e., those which bypass the FT uncertainty) approaches [31–38] have been reported to be applied to FTMS transient data with varying degree of success. A unique combination of computational speed and resolving power seems to make the filter diagonalization method (FDM) so far the most promising technique. The original idea of the method was formulated by Wall and Neuhauser [39]. Later, Mandelshtam and Taylor [40] turned this idea into a numerically efficient algorithm. Some further developments of FDM can be found in Refs. [41,42]. The method as formulated in Ref. [40] demonstrated superb results on both simulated [43–45] and experimental [45–49] FTMS transient signals. Hypothetically, under ideal (particularly, noiseless) conditions, FDM can provide infinite spectral resolution. However, the super resolution ability of FDM appears to be a very sensitive to various imperfections of the data, i.e., its deviation from the sum of exponentially dumped sinusoids. Those imperfections include noise, non-stationary effects, non-Lorentzian line shapes, etc.

* Corresponding author. Tel.: +1 949 824 5509; fax: +1 949 824-8571.

http://dx.doi.org/10.1016/j.ijms.2014.08.010
1387-3806/

This article introduces a robust and computationally efficient variant of a FDM-FT hybrid spectral approach in the field of FTMS. The core of the method is a computationally stable and fast version of the FDM algorithm.

Given the difficulty in estimating the method's performance subject to noise [50] and other conditions, results of an empirical investigation are presented to demonstrate the regions of applicability, and potential advantages of the method. These findings are compared with the performance of the method when applied to the experimental Orbitrap data.

2. Theory

Most results of this section can be found in Refs. [40–42].

2.1. The harmonic inversion problem and spectral estimation

Consider a time signal $\{c_n := c(n\tau)\}$ $(n = 0, \ldots, N-1)$, sampled on an evenly-spaced time grid. With the assumption that the data is composed of a series of damped sinusoidal oscillations, the harmonic inversion problem corresponds to the parametric fit

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

with the set of $2K$ complex-valued unknowns $\{u_k, d_k\}$, $(k = 1, \ldots, K)$, which we refer to as the line list. Each spectral feature is then characterized by an amplitude d_k and parameter u_k, which is in turn related to the "complex frequency" ω_k as

$$u_k := e^{-i\tau\omega_k}; \quad \omega_k := \nu_k - i\gamma_k \tag{2}$$

While the actual oscillation frequencies are given by the real parts ν_k, the imaginary parts γ_k correspond to the decay constants.

Define the hypothetical infinite-time FT of c_n:

$$I(\omega) := \tau \sum_{n=0}^{\infty} c_n z^{-n}; \quad z := e^{-i\tau\omega} \tag{3}$$

The spectral estimation problem challenges to estimate $I(\omega)$ from the finite sequence $\{c_n\}$.

The two problems are closely related to one another. The parametric fit (i.e., the harmonic inversion problem) is more demanding, when the individual entries $\{\omega_k, d_k\}$ are assumed to correspond to physical observables. However, even when the line list is not exact, a meaningful spectrum can often still be produced. Consequently, in what follows, we will also focus on the algorithm's capabilities with respect to spectral estimation, with the hope that FDM can produce spectra that are better than a standard finite FT spectrum:

$$I^{(\mathrm{FT})}(\omega) := \tau \sum_{n=0}^{N-1} c_n g_n z^{-n}, \tag{4}$$

where g_n is a suitable apodization function. As is well known, the main drawback of Eq. (4) is its slow convergence with respect to the data size N, manifested by the "FT uncertainty principle"

$$\Delta\omega = \frac{2\pi}{N\tau} \tag{5}$$

where $\Delta\omega$ stands for the spectral resolution as defined by the fast Fourier Transform (FFT) grid spacing, or the FFT bin.

Starting with Eq. (1) and assuming the FDM line list is available (regardless of the accuracy of its individual entries), we define the "ersatz" FDM spectrum:

$$I^{(\mathrm{FDM})}(\omega) := \tau \sum_{n=0}^{\infty} \sum_{k} d_k u_k^n z^{-n} \tag{6}$$

$$= \tau \sum_{k} \frac{d_k}{1 - u_k/z} \tag{7}$$

As already well documented (see, e.g., Refs. [40,41]), when the data c_n satisfies Lorentzian and sparseness assumption (1), which also implies a high signal-to-noise ratio (SNR), the resolution of the FDM ersatz spectrum becomes superior to that of the finite FT spectrum (4). However, for severely truncated data, the FDM resolution is a nontrivial function of several factors, such as spectral line shapes, multiplet structure, and SNR, rather than the data size (N) alone. When the Lorentzian assumption (1) is a poor representation of the data, the harmonic inversion problem is ill-posed and its solution may be unstable and not useful for spectral estimation. In this latter case it is desired to have a way to correct the ersatz spectral estimate, such that it becomes at least as "good" as the finite FT spectrum (4). One-way to accomplish this is to use a hybrid FDM-FT spectrum (see, e.g., [40,41]), which combines the strengths of both FT and FDM. In other words, if for some reason a genuine peak happens to be resolved in the FT spectrum, then by construction, the hybrid spectrum will recover this peak even if it is missing in the ersatz spectrum (false negative). On the other hand, a spurious peak (false positive) appearing in the ersatz spectrum will be suppressed in the hybrid spectrum.

First, note that a numerical solution of parameter estimation problem (1) obtained by FDM (or even by any other method) may not be exact, with the residual signal

$$c_n^{(res)} = c_n - \sum_k d_k u_k^n \tag{8}$$

This information can then be utilized in the postprocessing steps associated, in particular, with the hybrid spectral estimate. Here, the formulation of the latter is somewhat different from the previous proposals:

$$\begin{aligned} I^{(\mathrm{hybrid})}(\omega) &= \tau \sum_k \frac{d_k}{1 - u_k/z} + \tau \sum_{n=0}^{N-1} \left[c_n - \sum_k d_k u_k^n \right] g_n z^{-n} \\ &= I^{(\mathrm{FDM})}(\omega) + I^{(\mathrm{FT})}(\omega) - I^{(\mathrm{FDMf})}(\omega) \end{aligned} \tag{9}$$

where the finite FT estimate of the FDM ersatz spectrum is given by

$$I^{(\mathrm{FDMf})}(\omega) = \tau \sum_k d_k \sum_{n=0}^{N-1} (u_k/z)^n g_n \tag{10}$$

In practice the last equation is better to be evaluated analytically. For this, consider the cosine apodization function

$$g_n = \cos(n\alpha) = \frac{1}{2}(\delta^n + \delta^{-n}) \tag{11}$$

with $\alpha = \pi/2N$ and $\delta = e^{i\alpha}$, so that $\cos(N\alpha) = 0$ and $\delta^N = i$. Substituting this into Eq. (10) we obtain

$$\begin{aligned} I^{(\mathrm{FDMf})}(\omega) &= \frac{\tau}{2} \sum_k d_k \sum_{n=0}^{N-1} [(u_k/z\delta)^n + (u_k\delta/z)^n] \\ &= \frac{\tau}{2} \sum_k d_k \left[\frac{1 - i(u_k/z)^N}{1 - (u_k\delta/z)} + \frac{1 + i(u_k/z)^N}{1 - (u_k/z\delta)} \right] \\ &= \tau \sum_k d_k \frac{1 + (u_k/z)^{N+1} \sin\alpha - (u_k/z)\cos\alpha}{1 + (u_k/z)^2 - 2(u_k/z)\cos\alpha} \end{aligned} \tag{12}$$

which is an expression that is easy to evaluate.

Given the FDM line list, regardless of whether or not it is meaningful physically, Eqs. (9)–(12) provide working expressions for computing the hybrid FDM-FT spectral estimate.