The Maximum Entropy Method for Optical Spectrum Analysis of Real-Time TDDFT

M. Toogoshi, S. S. Kano and Y. Zempo
Computer and Information Sciences, Hosei University, 3-7-2 Kajino, Koganei 184-8584 Japan
E-mail: 14t0007@cis.k.hosei.ac.jp

Abstract. The maximum entropy method (MEM) is one of the key techniques for spectral analysis. The major feature is that spectra in the low frequency part can be described by the short time-series data. Thus, we applied MEM to analyse the spectrum from the time dependent dipole moment obtained from the time-dependent density functional theory (TDDFT) calculation in real time. It is intensively studied for computing optical properties. In the MEM analysis, however, the maximum lag of the autocorrelation is restricted by the total number of time-series data. We proposed that, as an improved MEM analysis, we use the concatenated data set made from the several-times repeated raw data. We have applied this technique to the spectral analysis of the TDDFT dipole moment of ethylene and oligo-fluorene with \( n = 8 \). As a result, the higher resolution can be obtained, which is closer to that of FT with practically time-evoluted data as the same total number of time steps. The efficiency and the characteristic feature of this technique are presented in this paper.

1. Introduction

Time-dependent density-functional theory (TDDFT) is widely used to calculate optical properties of medium-to-large molecules, and it is recognized as a powerful tool for analyzing electronic behaviors of excited states to a certain level [1]. In our calculations, we employed the real-time and real-space technique to solve the time-dependent equation of motion [2, 3]. Within the framework of this approach, the wave functions are calculated by the finite difference method on real spatial grids [4] without using explicit bases such as plane waves or Gaussian. In the time domain, we evolve the perturbed initial wave function until we obtain enough dynamic motions. In the traditional method, the optical strength function is calculated from the Fourier transform (FT) of the dynamic dipole moment.

In the frequency domain, however, the maximum entropy method (MEM) is frequently applied to the spectrum analysis with a relatively small number of time-series data to obtain a fairly high resolution and accuracy [5]. The MEM is based on the relationship that the known autocorrelation are connected with the power spectrum by the Fourier transform. We apply this method to the spectrum analysis of the dynamic dipole moment, which is calculated by real-time TDDFT. In the comparison of MEM and FT spectrum for several molecules, we realize that MEM needs less time steps than FT to obtain the optical spectrum with same resolution [6].

We propose that, for MEM, we use the several-times repeated time-series data in stead of single time-series data, in order to reduce the total number of time steps significantly. Since the spectral resolution of MEM depends on the maximum lag \( M \) of the autocorrelation where \( M \) is limited by the total number of time steps \( N \), we need a large enough \( N \) for a reasonable
resolution. With the repeated data, we can select the maximum lag \( M \) large enough without an additional computational cost. We have found that the MEM of the repeated data shows higher resolution than the FT spectra with the double size data. In this paper, we present the effectiveness and characteristic features of our new MEM, comparing with the FT and the conventional MEM.

The paper is organized as follows. Our new technique is presented in the next section, and the observed effectiveness and characteristics of our MEM technique follow along with our conclusions in the last section.

2. Method
In this section, we briefly describe our new technique. We propose that, as an improved MEM analysis, we use the concatenated data set made from the several-times repeated raw data. The resolution of the MEM spectra depends on not only the total number of time steps \( N \) but also the maximum lag \( M \) of autocorrelation. The larger value of \( M \) provide the well-developed peaks. Generally, the high frequency components of time-series data are taken into account, even for smaller \( M \). But for the analysis of lower frequency components, \( M \) must be large. Since the maximum lag \( M \) limits the low frequency resolution where \( M \) is limited by \( N \), we need a large enough \( N \) for a reasonable spectral resolution, and the total machine time could not be shortened than we hope. To solve this difficulty, we duplicate the data and repeat it many times as shown in Fig. 1 (a).

A framework of the present study is based on Burg’s method [5], where the samples are given by the real-time TDDFT. Given that the time-series data obeys Gaussian distribution, the entropy \( h \) of a Gaussian process is given by

\[
h \propto \int_{-\pi/\Delta t}^{\pi/\Delta t} \log P(\omega) d\omega,
\]

where \( \Delta t \) is a sampling interval, and \( P(\omega) \) is the power spectrum. We find \( \hat{P}(\omega) \) which Eq.(1) is maximized under the constraints of the Wiener-Khintchin’s theorem. The solution is

\[
\hat{P}(\omega)/\Delta t = \left| \frac{\beta}{1 + \sum_{k=1}^{M} a_k \exp(-ik\omega\Delta t)} \right|^2,
\]

where \( M \) is the maximum length of autocorrelation \( C_M \). The parameters \( a_k \) and \( \beta \) are the Lagrange multipliers and they are the solution of Yule-Walker equation [7]

\[
\begin{pmatrix}
C_0 & C_1 & \cdots & C_M \\
C_1 & C_0 & \cdots & C_{M-1} \\
\vdots & \vdots & \ddots & \vdots \\
C_M & C_{M-1} & \cdots & C_0
\end{pmatrix}
\begin{pmatrix}
a_1 \\
a_2 \\
\vdots \\
a_M
\end{pmatrix} =
\begin{pmatrix}
1 \\
0 \\
\vdots \\
0
\end{pmatrix}.
\]

We adopt the Levinson-Durbin algorithm to solve Eq.(3) efficiently.

3. Comparison of our new MEM and conventional methods
3.1. Ethylene
As a simple example to show the advantage of our proposed method, we applied our MEM to a small molecule ethylene. The molecular structure is based on the ground state. The time evolution is carried out according to the real-time TDDFT [2, 3]. The spectrum \( S(E) \) is calculated from the dynamic dipole moment \( \mu(t) \) in both FT and our MEM technique. The calculation procedure as shown Figure 1 (a) is that, after the real calculation up to the total
Figure 1. Ethylene absorption spectra. (a) the dynamic dipole moment of the twice repeated data, (b) the spectra of FT with the real time steps \(N = 20000\), and the spectra calculated by our MEM with (c) \(N = 10000\) and (d) the repeated data \(N = 10000 + 10000\). For both MEM calculations, the same maximum lag \(M = 6000\) are adopted. The MEM data are scaled to allow comparison with those of FT.

time steps 10000, we used the twice repeated data of the dipole moment, i.e. \(N = 10000 + 10000\), for our MEM calculation. The time step is \(\Delta t = 0.002\) eV\(^{-1}\). The absorption spectra by FT of \(\mu(t)\) is shown in Fig. 1 (b) with the number of real time steps \(N = 20000\). Figure 1 (c) shows the result of MEM with the real time steps \(N = 10000\) and the maximum lag \(M = 6000\). Although the traditional MEM provides the higher resolution compared to that of FT [6], there are some peak splits around 15 eV. In the result of the repeated data as shown in Fig. 1 (d), on the other hand, there is no this kind of peak splits observed. In addition to it, our MEM makes the spectrum similar resolution to that of the FT with the real data \(N = 20000\).

3.2. Fluorene
Poly(9,9-dialkyl-fluorene) and their substituted derivatives are expected to be basic materials for a blue emitting polymer diodes. Their electronic structures have been extensively studied [8, 9, 10, 11, 12]. We employed the oligomer of fluorene (FL) with \(n = 8\) and carried out the MEM calculation. The calculation for this kind size molecule is quite expensive to realize enough time evolution. If we can save the calculation cost to figure out the lower energy spectrum, it is quite effective. As similar as the calculation of ethylene, we used the concatenated repeated dipole moment data. Thus, we compared the efficiency of MEM. The calculation was done by MEM with \(N = 5000\) and \(M = 2500\), and the four times repeated data \(N = 5000 \times 4\) with \(M = 2500\). The results were compared to that by FT with for the real times steps \(N = 20000\).

The results are shown in Fig. 2. We can clearly see that MEM has a significant advantage to figure out the lower energy part of the spectrum. In particular, the concatenated four times repeated data of the dipole moment can be used for MEM, which is much more effective to the conventional MEM.

The reason why the repeated data is efficient for the lower energy part of the spectrum analysis can be explained as follows. The lower energy resolution is attributed to the maximum lag \(M\) of autocorrelation. The lager number of \(M\) provides the information of the lower energy
Figure 2. Concatenated four times repeated data of the dipole moment. (a) Real time evolution up to $N = 20000$ for FT. (b) Four times repeated data of the dipole moment with $N = 5000 \times 4$ and $M = 2500$ for MEM.

Figure 3. Comparison of absorption spectrum of oligo-FL$_8$ obtained by FT and MEM. The solid line is the result of FT with $N = 20000$, the dashed line is that of MEM with $N = 5000$ and $M = 2500$, and the dotted line is that of MEM with $N = 5000 \times 4$ and $M = 2500$. Spectrum are normalized at the second low energy peak.

part of the spectrum. However, the unphysical results such as peak splits and false peaks are appeared by selecting too large value of $M$ from our experience. Also, $M$ is restricted by the total number of time-series data $N$. These are caused by the few contributions of the higher-order autocorrelations compared to the small-order ones.

Figure 4 graphically shows the elements of the Toepliz matrices of ethylene in the Yule-Walker equation as stated in Eq. (3) with the total time steps (a) $N = 10000$ and (b) $N = 10000 + 10000$, respectively. The maximum lag $M$ is used as the same value $M = 9999$ for both cases. The color is meant for the magnitude of $C_M$. All the values of the elements are normalized so as to make the diagonal value unity, which is corresponding to the zero-lag autocorrelation, colored red. In the case of the real data use, the values of high-order autocorrelations are almost close to zero for $M > 5000$ as shown in Fig. 4 (a). On the other hands, the result of repeated data
provides the effective values of $M$ from about $M = 5000$ to $M = 7500$ as shown in Fig. 4 (b). With the repeated data, the effective autocorrelations expressed by the maximum lag $M$ can be increased. In other words, the method enhances the information of the low energy part for time-series data without additional calculations, which save our calculation cost.

![Figure 4. Magnitude of the matrix elements of Eq.(3) in the cases of (a) $N = 10000$ and $M = 9999$, and (b) $N = 10000 + 10000$ and $M = 9999$.](image)

From these results above, MEM has several advantages. First, there is no the periodicity appeared to the signal in MEM, which is different from FT. When we applied FT to the repeated data in Fig. 1 (a), the periodicity was appeared to the spectrum as artificial frequency components, which disturbed the signal detection. Second, there is a phase jump for each frequency component at the connection point. However, MEM is originally used of the method for analysing the natural phenomena, which involve a lot of noises. This phase jump can be perceived as one of the noises. MEM is quite robust to detect signals from the system with noises. Thus, we observed the signal clearly in Fig. 1 (d).

4. Concluding remarks
By the use of MEM, the spectrum is generally calculated with the small number of time-series data, in comparison with the conventional FT [6]. We proposed the new MEM technique to obtain the spectrum with much smaller number of time-series data especially in the lower frequency part. In the new technique, we make a quasi-long size data, which is made by several times repeating the small size of time-series data. In this calculation, there are many higher-order autocorrelations involved, which are relevant to the lower frequency part of the spectrum. The resolution of the spectrum is dependent on the total number of time steps $N$ and the autocorrelation lag $M$. In our calculation, we use the quasi-longer $N$. Thus, we can apparently increase the higher-order autocorrelation lag $M$. As a result, the spectrum is efficiently enhanced in the lower frequency part.

The new technique is applied to the spectrum analysis of the dynamic dipole moment of several molecules such as ethylene and oligo-fluorene with $n = 8$. We confirm that it is quite efficient to obtain spectra from TDDFT. With a fairly small number of time-series data, we obtained much better resolutions of the low frequency part of the spectrum. This means that we can save our calculation cost efficiently.
Acknowledgments
Authors would like to express their acknowledgements to Professor H. Hanaizumi for useful discussions and encouragements. This work was partially supported by JSPS Grants-in-Aid for Scientific Research (C) Grant number 25390158, Sumitomo Chemical, Co. Ltd. and Simulation Corporation.

References
[1] Runge E and Gross E K U 1984 Phys. Rev. Lett. 52, 997
[2] Yabana K and Bertisch G F 1996 Phys. Rev. B54, 4484
[3] Zempo Y, Akino, Ishida M, Ishitobi M and Kurita Y 2008 J. Phys. Condens. Matter 20, 064231
[4] Chelikowsky J, Troullier N, Wu K and Saad Y 1994 Phys. Rev. B50, 11355
[5] Burg J P 1968 Advanced Study Institute on Signal Processing, 15-0, NATO, Erschede, Netherlands
[6] Toogoshi M, Kano S S, Zempo Y 2014 J. Phys. Conf. Ser. 510, 012027
[7] Haykin S and Kesler S 1979 "Non linear Methods of Spectral Analysis", ed S. Haykin, Springer(Berlin)
[8] e.g. Hummer K, Puschnig P, Sagmeister S and Ambrosch-Draxl C 2006 Mod. Phys. Lett. B20, 261 , and the references therein.
[9] Yu W L, Cao Y, Pei J, Huang W and Heeger A J 1990 Appl. Phys. Lett. 75 3270
[10] Wang J F, Feng J K, Ren A M, Liu X D, Ma Y G, Lu P and Zhang H X 2004 Macromolecules 37, 3451
[11] Brière J F and Coté M 2004 J. Chem. Phys. B108, 3123
[12] Grice A W, Bradley D D C, Bernius, M T, Inbasekaran M, Woo EP and Wu W W 1998 Appl. Phys. Lett. 73, 629