Practical Computation of FFT based Generalized Two-Dimensional Correlation Spectroscopy

Mengmeng HE,* Daisuke MIYATA,** Takakazu NAKABAYASHI,** and Shin-ichi MORITA*†

*Graduate School of Science, Tohoku University, 6-3 Azami-Aza-Aoba, Aoba, Sendai 980-8578, Japan
**Graduate School of Pharmaceutical Sciences, Tohoku University, 6-3 Azami-Aza-Aoba, Aoba, Sendai 980-8578, Japan

† To whom correspondence should be addressed.
E-mail: shinichi.morita.c6@tohoku.ac.jp

Abstract
A practical computation of fast Fourier transformation (FFT) based generalized two-dimensional (2D) correlation spectroscopy was described. Using simple sinusoids, we tested it and confirmed that the method served effectively and properly, invariant to the changes of the number of data points of the time-profiles. This computation is applicable to any type of waveforms in a versatile manner.

Keywords Correlation, spectroscopy, FT, FFT, computation, sinusoidal
Generalized two-dimensional (2D) correlation spectroscopy was proposed by Noda in 1993.\textsuperscript{1,2} Using two types of two-dimensionally spread maps with correlation peaks, that is, synchronous and asynchronous 2D spectra, researchers are capable of (i) detecting synchronicity and time-difference of signal intensity variations, (ii) resolution enhancement of overlapped bands, (iii) detection of slight but important signals, (iv) assignments of newly detected signals, and (v) determination of the sequential order determination of signal intensity changes. The versatile analytical tool has been widely applied in infrared, near-infrared, Raman, fluorescent spectroscopies, including any other optical measurements.\textsuperscript{3-10}

For the generalization of the correlation method, Fourier transform (FT) was incorporated to theoretical framework, by which the applicable range was drastically extended from the simple sinusoidal profiles\textsuperscript{11} to any type of non-periodic ‘general’ waveforms.\textsuperscript{1,2} Furthermore, using Hilbert transform (HT), Noda described another theoretical framework of generalized two-dimensional correlation spectroscopy,\textsuperscript{12} in which the generalized 2D correlation functions obtained through the FT and HT routes were confirmed identical to each other. Fig. 1 summarizes a comprehensive view of the generalized 2D correlation spectroscopy both for the HT and FT routes.

HT only uses the time domain, without passing through the frequency domain. This feature serves effectively for the practical computations using ‘discrete Hilbert transform’ (DHT) of generalized 2D correlation spectroscopy, which realized a user friendly analytical tool that produced many publications of application studies.\textsuperscript{2-10} On the other hand, the practical computation using discrete FT (usually researchers use the algorithm named ‘fast Fourier transform’ (FFT) because of the fast computation) was reported only in a limited manner, not widely used probably because of the computational complexity.

The FFT based generalized 2D correlation spectroscopy becomes important especially when researchers study the mechanism of the method. For instance, as summarized in Fig.1, it is possible to relate the generalized 2D correlation function $\Phi + i\Psi$ with cross correlation function $C(\tau)$ through $\phi_\omega - i\psi_\omega$ in the frequency domain ($\phi_\omega$ and $\psi_\omega$ are called co-spectrum and quad-spectrum, respectively).\textsuperscript{1,2} One of advantages of generalized 2D correlation spectroscopy is to detect ‘fast/slow’ changes in a monotonously changing system such as exponentially decaying curves, while cross correlation function is based on the operation of time shifting by $\tau$, difficult to estimate the fast/slow change in a directly meaningful manner. Numerical simulation study of FFT based generalized 2D correlation spectra will probably clarify the difference.
To date, practical computation of FFT based generalized 2D correlation spectroscopy was partially reported\textsuperscript{14,15} but has not been clearly described. For instance, the normalizing factor of the generalized 2D correlation function $\Phi + i \Psi$ in terms of the number of data points was not given or not properly given. This limitation has delayed the further mechanism study on generalized 2D correlation spectroscopy and related time-series correlation analysis.

The purpose of this research is to develop a practical computation of the FFT based generalized 2D correlation spectroscopy. The novelty and originality of this practical computation is to realize ‘versatility’ in terms of the number of data points and ‘compatibility’ between the DHT and FFT routes by selecting the normalizing factor of $\Phi + i \Psi$ properly.

We begin with the theoretical framework of the FT based generalized 2D correlation functions given by Noda.\textsuperscript{1,2} As functions of spectral variable $\nu$ and time $t$, spectra $y(\nu, t)$ are observed during the time range of $T_{\text{min}} \leq t \leq T_{\text{max}}$, where $T_{\text{min}}$ and $T_{\text{max}}$ are start and end times of the measurement. The spectra $y(\nu, t)$ are transformed to dynamic spectra $\tilde{y}(\nu, t)$ as

$$\tilde{y}(\nu, t) = \begin{cases} y(\nu, t) - \bar{y}(\nu) & \text{for } T_{\text{min}} \leq t \leq T_{\text{max}} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Here $\bar{y}(\nu)$ is named a reference spectrum. Usually, the mean spectrum is used for $\bar{y}(\nu)$, given by

$$\bar{y}(\nu) = \frac{1}{T_{\text{max}} - T_{\text{min}}} \int_{T_{\text{min}}}^{T_{\text{max}}} y(\nu, t) dt. \quad (2)$$

The dynamic spectra $\tilde{y}(\nu, t)$ are then subjected to FTs for yielding $\tilde{Y}_1(\omega)$ and $\tilde{Y}_2^*(\omega)$ as a function of frequency $\omega$ as

$$\tilde{Y}_1(\omega) = \int_{-\infty}^{\infty} \tilde{y}(\nu_1, t) e^{-i\omega t} dt$$

$$= \tilde{Y}_1^{\text{Re}}(\omega) + i \tilde{Y}_1^{\text{Im}}(\omega), \quad (3)$$

$$\tilde{Y}_2^*(\omega) = \int_{-\infty}^{\infty} \tilde{y}(\nu_2, t) e^{+i\omega t} dt$$

$$= \tilde{Y}_2^{\text{Re}}(\omega) - i \tilde{Y}_2^{\text{Im}}(\omega). \quad (4)$$
The symbol * of $\hat{Y}_2^*(\omega)$ indicates that $\hat{Y}_1(\omega)$ and $\hat{Y}_2^*(\omega)$ are conjugated with each other. The symbols Re and Im are the operations to take real and imaginary parts of the variable, respectively. The synchronous function $\Phi(v_1, v_2)$ and the asynchronous function $\Psi(v_1, v_2)$ are defined as

$$\Phi(v_1, v_2) + i \Psi(v_1, v_2) = \frac{1}{\pi (T_{\text{max}} - T_{\text{min}})} \int_0^\infty \hat{Y}_1(\omega) \hat{Y}_2^*(\omega) d\omega.$$  

(5)

Corresponding to the above theoretical framework, we here propose a practical computation of the FFT based generalized 2D correlation spectroscopy. As functions of spectral variable $v$ and time $t_j$, spectra $y(v, t_j)$ are observed during the time range of $T_{\text{min}} \leq t_j \leq T_{\text{max}}$, where $T_{\text{min}}$ and $T_{\text{max}}$ are start and end times of the measurement; for simplicity, we denote $y(v, t_j)$ as $y_j(v)$, that is, $y_j(v) \equiv y(v, t_j)$ where $j = 1, 2, 3 \ldots m$. $t_j$ is given as $t_j = T_{\text{min}} + (T_{\text{max}} - T_{\text{min}})(j - 1) / (m - 1)$. The spectra $y(v, t)$ are transformed to dynamic spectra $\tilde{y}(v, t)$ as

$$\tilde{y}_j(v) = y_j(v) - \bar{y}(v),$$  

(6)

where $\bar{y}(v)$ is named a reference spectrum. Usually, the mean spectrum is used for $\bar{y}(v)$, given by

$$\bar{y}(v) = \frac{1}{m} \sum_{j=1}^m y_j(v).$$  

(7)

If necessary, zero values are added to both sides (or one side) of $\tilde{y}_j(v)$ in Eq. (6), in order to reflect the statement of ‘$\tilde{y}(v, t) = 0$ otherwise’ in Eq. (1). Then, dynamic spectra $\tilde{y}_j(v)$ are subjected to discrete FT to give $\hat{\tilde{y}}_k(v_1)$ and $\hat{\tilde{y}}_k^*(v_2)$ as a function of $k = 1, 2, 3 \ldots m$ as

$$\hat{\tilde{y}}_k(v_1) = \text{conj} \left\{ \sum_{j=1}^m \tilde{y}_j(v_1) \exp \left[ -i 2\pi \frac{(k-1)(j-1)}{m} \right] \right\}$$  

(8)

$$\hat{\tilde{y}}_k^*(v_2) = \text{conj} \left\{ \sum_{j=1}^m \tilde{y}_j(v_2) \exp \left[ +i 2\pi \frac{(k-1)(j-1)}{m} \right] \right\}$$  

(9)
where the symbol * of $\tilde{Y}_k(\nu)$ indicates that $\tilde{Y}_k(\nu)$ and $\tilde{Y}_k^*(\nu)$ are conjugated with each other and $\text{conj}(x)$ is the operation to return the complex conjugate of $x$. Here we recall Euler’s formula that $e^{-i\omega t} = \cos \omega t - i \sin \omega t$. But for $e^{-i\omega t}$ at the first line of Eq. (3), the sign of ‘+’ used in the terms of $\tilde{Y}_1^{\text{Re}}(\omega) + i \tilde{Y}_1^{\text{Im}}(\omega)$ at the second line is opposite (the same relation appears also in Eq. (4)). This indicates that we practically need to take complex conjugation both for Eq. (8) and Eq. (9). The algorithm named ‘fast Fourier transform’ (FFT) is usually used for calculating of discrete FT in Eqs. (8) and (9) because of the fast computation.

The synchronous function $\Phi(\nu_1, \nu_2)$ and asynchronous function $\Psi(\nu_1, \nu_2)$ are defined as

$$\Phi(\nu_1, \nu_2) + i \Psi(\nu_1, \nu_2) = \frac{1}{\pi (m - 1)^2} \sum_{k=1}^{\text{ceil}(m^2/2)} \tilde{Y}_k(\nu_1) \tilde{Y}_k^*(\nu_2),$$

(10)

where $\text{ceil}(x)$ is the ceiling function to transform a real number $x$ to the least integer larger than the real number $x$ (we supposed odd $m$). The most important point of this practical computation to us is the versatility to $m$ by properly introducing the normalizing factor of $1/(m - 1)^2$. This realizes the compatibility between the DHT and FFT routes.

Using the proposed FFT based generalized 2D correlation functions, we practically calculated sinusoidal signals of $y(\nu_1, t) = \cos(t + \pi/6)$ and $y(\nu_2, t) = \cos(t)$. The phase difference between these is $\pi/6$ corresponding to $30^\circ$. Let us first exemplify discretization of sinusoidal signals in Fig. 2. Here we assume that the period $2\pi$ is divided into 9 data points, that is, $j' = 1, 2, 3 \cdots 9 \ (m' = 9)$, in which the data spacing $\Delta t = (T_{\text{max}} - T_{\text{min}})/(m' - 1) = \pi/4$. Please note that data points at the 1st ($t = 0$) and the 9th ($t = 2\pi$) positions repeat the same information. To avoid this double counting, the data point at $m' = 9$ was deleted with the data spacing $\Delta t = \pi/4$ kept, in order to reproduce $j = 1, 2, 3 \cdots 8 \ (m = 8)$. In other words, the time range is changed from $0 \leq t \leq 2\pi \ (j' = 1, 2, 3 \cdots m', m' = 9)$ to $0 \leq t < 2\pi \ (j = 1, 2, 3 \cdots m, m = 8)$. Please note that the symbol $\leq$ corresponds to the circle with the half white or black in Fig. 2, $\leq$ is used for the fully solid black circle, and $<$ is for the fully white open circle. This operation is important especially when we quantitatively estimate a target parameter such as the phase difference of $\pi/6$ initially set in this example. However, researchers usually do not have to operate the reduction of $m'$ to $m$ since generalized 2D correlation spectroscopy mainly aims to determine the sequential order of signal changes. The determination serves properly even without reducing the last data point.
Fig. 3 shows changes in $\Phi$, $\Psi$, and $\Theta$ values by increasing the number of data points $m$. $\Phi$ and $\Psi$ are synchronous and asynchronous values, respectively. $\Theta$ is global phase angle values defined by $\Theta = \text{atan} (\Psi/\Phi)$, here specifically corresponding to the phase angle between the simple sinusoids. $\Phi$, $\Psi$, and $\Theta$ values obtained by FFT (red solid circles) were compared with those by DHT (blue open circles). Starting from $m = 8$, $\Phi$ and $\Psi$ values for FFT decreased rapidly, from $m = 50$ to $100$, $\Phi$ and $\Psi$ became optimized at certain values. These indicate that the normalizing factor $(m - 1)^2$ in the denominator of Eq. (10) serves effectively and versatilely for the change in the number of data points of $m$. Also, the $\Theta$ values gave the stable value of $30^\circ$ invariant to the change in $m$, which agreed with the initial set value of $\pi/6$ in the simulation.

For $\Phi$ and $\Psi$ values for FFT as a function of $m$, (i) the signs of $\Phi$ and $\Psi$ are retained; (ii) the relation of $\Phi > \Psi$ are also retained. These assure that determination of the sequential order using the FFT based generalized 2D correlation spectroscopy serves in a versatile manner in sinusoidal system, invariant to the number of measured data points, $m$. As a matter of course, this computation of the FFT based generalized 2D correlation spectroscopy serves properly for any type of waveforms.

In conclusion, we discretized the theoretical framework of the FT based generalized 2D correlation spectroscopy. Tips and difficulties to realize the discretization are summarized in the followings: (i) researchers have to add zero values to both sides (or one side) of observed signals if necessary because FFT (or discrete FT) implicitly assumes repetition of signals; (ii) further complex conjugations were taken in Eq. (8) and Eq. (9) to be consistent with the Noda’s definition of the FT in Eqs. (3) and (4); (iii) considering the odd and even number of data points of $m$, ceiling function to $m/2$ was used approximately to return integers; (iv) to numerically estimate the phase difference of the sinusoids, we needed to reset the start and end points of the time range for avoiding the double counting of the both edges giving the same information due to the sinusoidal repetition. This practical computation properly selected the normalizing factor of $1/(m - 1)^2$ in the synchronous and asynchronous functions $\Phi + i\Psi$. This realizes versatility in terms of the number of data points $m$ and compatibility between the DHT and FFT routes.

References
1. I. Noda, Appl. Spectrosc., 1993, 47, 1329.
2. I. Noda and Y. Ozaki, “Two-Dimensional Correlation Spectroscopy Applications in
Vibrational and Optical Spectroscopy”, 2004, John Wiley and Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.
3. Y. Ozaki, Y. L. Liu, and I. Noda, Macromolecules, 1997, 30, 2391.
4. K. Nakashima, S. Yashuda, Y. Ozaki, and I. Noda, J. Phys. Chem. A, 2000, 104, 9113.
5. Y. Q. Wu, K. Murayama, and Y. Ozaki, J. Phys. Chem. B, 2001, 105, 6251.
6. W. Zhao, C. H. Song, B. Zheng, J. Liu, and T. Viswanathan, J. Phys. Chem. B, 2002, 106, 293.
7. I. Noda and C. Marcott, J. Phys. Chem. A, 2002, 106, 3371.
8. K. Izawa, T. Ogasawara, H. Masuda, H. Okabayashi, and I. Noda, Macromolecules, 2002, 35, 92.
9. H. C. Choi, S. B. Kim, and Y. M. Jung, J. Mol. Struct., 2006, 799, 91.
10. H. Fukuma, K. Nakashima, Y. Ozaki, and I. Noda, Spectrochim. Acta. A, 2006, 65, 517.
11. I. Noda, J. Am. Chem. Soc., 1989, 111, 8116.
12. I. Noda, Appl. Spectrosc., 2000, 54, 994.
13. T. Sugawara, T. Nakabayashi, and S. Morita, Anal. Sci., 2018, 34, 845.
14. R. Geitner, R. Fritzsch, J. Popp, T. W. Bocklitz J. Statistical Software, 2019, 90.
15. P. B. Harrington, A. Urbs, and P. J. Tandler, Chemom. Intell. Lab. Syst., 2000, 50, 149.
16. S. Morita, Y. Ozaki, and I. Noda, Appl. Spectrosc., 2001, 55, 1618.
Figure Captions

Fig. 1. A symbolic representation of the HT and FT routes to yield generalized 2D correlation spectroscopy. (FT): FT or inversed FT, (HT): HT or inversed HT, (τ): time-shifting, (·): multiplication, and (∫): integration. Dynamic spectra $\tilde{y}_1$ and $\tilde{y}_2$ are respectively transformed to $Y_1$ and $Y_2^*$ by FTs (* indicates conjugated relation). $\tilde{z}_2$ is HT of $\tilde{y}_2$. $\phi_\omega$ is co-spectrum and $\psi_\omega$ is quad-spectrum. $D_\tau$ is orthogonal correlation function. $\Phi$ is synchronous function and $\Psi$ is asynchronous function.

Fig. 2 Sinusoidal curves within the period of $2\pi$ and the discretization.

Fig. 3 $\Phi$, $\Psi$, and $\Theta$ values by increasing the number of data points from $m = 8$ to 100. $\Phi$, $\Psi$, and $\Theta$ values obtained by FFT (red solid circles) were compared with those by DHT (blue open circles).
Fig. 1. A symbolic representation of the HT and FT routes to yield generalized 2D correlation spectroscopy. (FT): FT or inversed FT, (HT): HT or inversed HT, (τ): time-shifting, (·): multiplication, and (∫): integration. Dynamic spectra \( \tilde{y}_1 \) and \( \tilde{y}_2 \) are respectively transformed to \( Y_1 \) and \( Y_2^* \) by FTs (\( * \) indicates conjugated relation). \( \tilde{z}_2 \) is HT of \( \tilde{y}_2 \). \( \phi_{\omega} \) is co-spectrum and \( \psi_{\omega} \) is quad-spectrum. \( D_{\tau} \) is orthogonal correlation function. \( \Phi \) is synchronous function and \( \Psi \) is asynchronous function.
Fig. 2: Sinusoidal curves within the period of $2\pi$ and the discretization.
Fig. 3 Φ, Ψ, and Θ values by increasing the number of data points from \( m = 8 \) to 100. Φ, Ψ, and Θ values obtained by FFT (red solid circles) were compared with those by DHT (blue open circles).
Graphical Index

\[ \Phi(\nu_1, \nu_2) + i \Psi(\nu_1, \nu_2) = \frac{1}{\pi (m - 1)^2} \sum_{k=1}^{\text{ceil}(m/2)} \tilde{Y}_k(\nu_1) \tilde{Y}_k^*(\nu_2) \]

FFT based generalized 2D correlation functions