The Astronomical Journal, 162:111 (14pp), 2021 September https://doi.org/10.3847/1538-3881/ac11f7

A Data-scientific Noise-removal Method for Efficient Submillimeter Spectroscopy With Single-dish Telescopes

Akio Taniguchi[1], Yoichi Tamura[1], Shiro Ikeda[2,3], Tatsuya Takekoshi[4,5], and Ryohei Kawabe[6,7]
[1] Division of Particle and Astrophysical Science, Graduate School of Science, Nagoya University, Furocho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan
taniguchi@a.phys.nagoya-u.ac.jp
[2] The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
[3] Department of Statistical Science, The Graduate University for Advanced Studies (SOKENDAI), 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
[4] Kitami Institute of Technology, 165 Koen-cho, Kitami, Hokkaido 090-8507, Japan
[5] Institute of Astronomy, Graduate School of Science, The University of Tokyo, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan
[6] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[7] Department of Astronomical Science, The Graduate University for Advanced Studies (SOKENDAI), 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan
Received 2021 February 1; revised 2021 June 30; accepted 2021 July 5; published 2021 August 19

Abstract

For submillimeter spectroscopy with ground-based single-dish telescopes, removing the noise contribution from the Earth's atmosphere and the instrument is essential. For this purpose, here we propose a new method based on a data-scientific approach. The key technique is statistical matrix decomposition that automatically separates the signals of astronomical emission lines from the drift noise components in the fast-sampled (1–10 Hz) time-series spectra obtained by a position-switching (PSW) observation. Because the proposed method does not apply subtraction between two sets of noisy data (i.e., on-source and off-source spectra), it improves the observation sensitivity by a factor of $\sqrt{2}$. It also reduces artificial signals such as baseline ripples on a spectrum, which may also help to improve the effective sensitivity. We demonstrate this improvement by using the spectroscopic data of emission lines toward a high-redshift galaxy observed with a 2 mm receiver on the 50 m Large Millimeter Telescope. Since the proposed method is carried out offline and no additional measurements are required, it offers an instant improvement on the spectra reduced so far with the conventional method. It also enables efficient deep spectroscopy driven by the future 50 m class large submillimeter single-dish telescopes, where fast PSW observations by mechanical antenna or mirror drive are difficult to achieve.

Unified Astronomy Thesaurus concepts: Atmospheric effects (113); Spectroscopy (1558); Astronomical methods (1043); Astronomy data reduction (1861); Millimeter astronomy (1061); Submillimeter astronomy (1647)

1. Introduction

Spectroscopy with large single-dish telescopes at the submillimeter wavelength is the key to understanding the dust-obscured cosmic star formation history of the universe. The wide-area ($>1\,\mathrm{deg}^2$) and sensitive ($\sigma \sim 0.1$ mJy) spectroscopic mapping of emission lines (e.g., [C II], [O III]) from galaxies enables us to measure the cosmic star formation rate density in the epoch of reionization and beyond (e.g., Kohno et al. 2019). To cover the huge three-dimensional volume of the universe (i.e., transverse area and time), large next-generation ground-based submillimeter telescopes ($D \sim 50$ m) have been proposed (Kawabe et al. 2016; Klaassen et al. 2019; Lou et al. 2020).

Several new spectroscopic instruments for "3D imagers" have also been developed. Wideband (~20 GHz) spectroscopy at a high spectral resolution ($\lambda/\Delta\lambda \gtrsim 10^5$) becomes promising by the recent development of heterodyne receivers with instantaneous wideband (Kojima et al. 2020) in radio frequency and digital spectrometers (Klein et al. 2012; Iwai et al. 2017). As ultrawideband ($\gtrsim$100 GHz) spectrometers at a medium spectral resolution ($\lambda/\Delta\lambda \sim 500$), integrated superconducting spectrometers (ISSs) based on an on-chip filterbank and microwave kinetic inductance detectors have been developed for blind redshift surveys (Wheeler et al. 2016; Endo et al. 2019). The deep spectroscopic high-redshift mapper (DESHIMA) demonstrates the detection of astronomical signals in the 345 GHz band with an instantaneous bandwidth of 45 GHz on the ASTE 10 m telescope (Endo et al. 2019) and will be upgraded to cover 220–440 GHz.

Alongside the new wideband instruments, observation strategies for them must be updated. Ground-based spectroscopy at submillimeter wavelengths is strongly affected by intense molecular emission from the Earth's atmosphere (e.g., H_2O, O_3), and removing such noise emission from observed spectra is essential to obtain astronomical signals. Calibration of the instrument (i.e., bandpass characteristics and absolute intensity scale) is also of great importance. For these purposes, the position-switching (PSW) method is widely used for heterodyne receivers (Wilson et al. 2012), where the atmospheric emission at the sky position with an astronomical target (on-source position, hereafter) is removed through subtraction at the sky position without the target (off-source position, hereafter). The two positions are alternatively observed generally by mechanical antenna driving at a switching interval of several seconds. The brightness temperature of the astronomical signals corrected for the atmospheric transmission, $T_\star$, is obtained by the following operation:

$$T_\star = T_{\rm sys}\,\frac{P_{\rm on} - P_{\rm off}}{P_{\rm off}}, \tag{1}$$

where $T_{\rm sys}$ is the system noise temperature, and $P_{\rm on}$ and $P_{\rm off}$ are the measured powers of the on-source and off-source positions, respectively. The standard-deviation noise level of the signals, $\sigma_\star$, is dependent on the frequency channel width of a

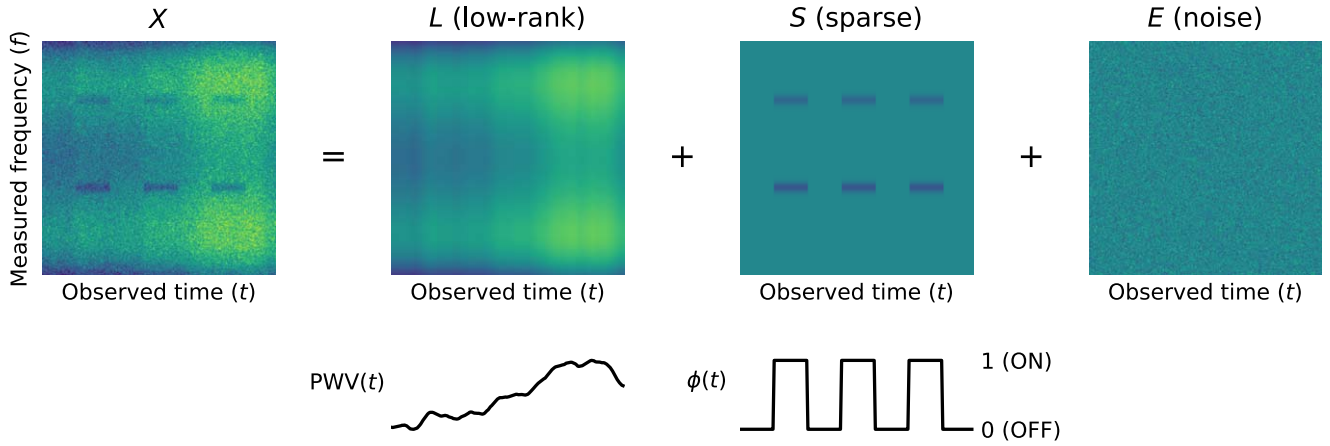

Figure 1. Schematic diagram of low-rank and sparse decomposition of a mock fast-sampled PSW observation. The top-left image represents a matrix of a dimensionless quantity, $X(f, t)$ (Equation (16)), where six rectangular features are two Gaussian-shaped emission lines observed by three on-source measurements. The other images represent decomposed low-rank (L), sparse (S), and noise (E) matrices. The bottom plots show the PWV and position indicator as a function of the observed time. Note that the values of the emission lines (nonzero values in S) are negative because of the definition of $X(f, t)$. Also note that emission lines here is set to be very bright for visualization: in real observations, they are often too faint and buried in the noise, the case of which will be discussed in Section 3.2.

spectrometer, $\Delta\nu$, and the total on-source time, $t_{\rm on}$:

$$\sigma_\star = \frac{\sqrt{2}\ T_{\rm sys}}{\sqrt{\Delta\nu\ t_{\rm on}}}. \tag{2}$$

Several observing methods have also been proposed. The frequency switching alternatively obtains spectra with two or more (Heiles 2007) different frequencies and subtracts each other. Because it does not need to point to the off-source position, efficient on-source integration is possible in emission-line observations with moderate line widths[8] ($\lesssim$100 km s^{-1}). As an improvement of the PSW method, wobbler switching is used in the IRAM 30 m telescope (Ungerechts et al. 2000), where faster switching interval (typically 0.5–1 s) is achieved by wobbling the secondary mirror.

The difficulty arises, however, for these "switching" observations with the new instruments. They apply subtraction of the observed noisy spectra, which results in the "addition" of noises to the on-source spectrum (we will refer to it as "the direct on–off subtraction"). This is why the factor $\sqrt{2}$ appears on the right-hand side of Equation (2). They also assume that the conditions of the atmosphere and the instrument (i.e., instrumental response and $T_{\rm sys}$) should be constant during each period of switching. In particular, the atmosphere is troublesome because its typical time-variation scale[9] is often close to the switching interval (10^{-1}–10^{0} Hz, e.g., Chapin et al. (2013)). When the switching interval between two positions or frequencies is insufficient, drift noise contribution from the atmosphere and the instrument may cause imbalance in subtraction (Schieder & Kramer 2001), which results in artificial signal in the spectrum (e.g., baseline ripples). This may even increase the effective noise level of the spectrum more than expected. Because fast position switching by mechanical antenna/mirror drive, or large frequency throw is not realistic for wideband spectroscopy with large single-dish telescopes, current observing methods are still the bottleneck.

To overcome these issues, it is essential to develop a new method that can separate astronomical signals from drift noise components. One of the promising solutions is to utilize a spectral correlation of the atmospheric emission: the atmospheric spectrum depends mostly on the amount of line-of-sight water vapor, which can be expressed as a simple function. If we continuously obtain the on-source and off-source spectra and view them as a matrix, it is well approximated with a low-rank matrix (see also Figure 1). By contrast, deep spectroscopic surveys are expected to detect only a few bright emission lines per target source, which suggests that emission-line components should be sparse (i.e., the fraction of nonzero elements is small) in a matrix composed of on-source and off-source spectra.

In data science, sparsity-based methods have revolutionized many applications. The innovative methods, Lasso (Tibshirani 1996) and compressive sensing (Candès et al. 2006; Donoho 2006) focused on the sparsity of vectors and have been successfully applied to astronomy (Uemura et al. 2015; The EHT Collaboration et al. 2019). The concept of sparsity has been generalized to matrices (Candès et al. 2011), where the sparsity of the singular values (i.e., low rank) and the component-wise sparsity have been utilized. This extension enabled us to decompose a matrix into low-rank and sparse components that made a variety of applications in computer vision. We have seen some successful applications in astronomy (Morii et al. 2017; Zuo et al. 2018) and our present approach also take the advantage of the effective method.

For the decomposition of a matrix into low-rank and sparse components, several computational algorithms have been proposed (Candès et al. 2011; Zhou & Tao 2011). This leads us to a possible solution: if we can (1) obtain fast-sampled time-series spectra of on-source and off-source measurements and (2) make a matrix that can be expressed as the sum of atmospheric and emission-line components, then we would achieve continuous and noiseless estimates of the atmospheric emission even during a switching observation.

In this paper, we propose a noise-removal method for submillimeter single-dish spectroscopy based on low-rank and sparse decomposition of a matrix of a fast-sampled time-series observation. In Section 2, we reformulate the PSW method and derive a time-series matrix that fulfills conditions (1) and (2).

[8] Maximum observable line width is limited by the maximum frequency throw (typically several tens of MHz).

[9] For example, assuming a typical wind speed of $v \sim 10$ m s^{-1} and $D \sim 10$–50 m, the transverse time of a cloud across the near field of the telescope's beam is $D/v \sim 1$–5 s.

Section 3 describes algorithms for low-rank and sparse decomposition for fast-sampled PSW observations. In Section 4, we apply the proposed method to real-observed fast-sampled PSW data and demonstrate that it can improve the observation sensitivity compared to the conventional direct on–off subtraction. Finally, we discuss the advantages, limitations, and potential applications of the proposed method in Section 5.

2. Reformulation

The main idea of the proposed method is summarized in Figure 1, which schematically describes low-rank and sparse decomposition applied for a mock fast-sampled PSW observation. We start with the general response functions (often referred to as observation equations) of a system that expresses the relation between input astronomical signals and output power spectrum of a spectrometer (Section 2.1). We then reformulate the PSW method to derive an observable expressed as a product of input signals and response terms (Section 2.2). Finally, we apply low-rank and sparse decomposition to the observable in a logarithmic space, derive the integrated spectrum of the mock observation, and compare it with that reduced by the direct on–off subtraction (Section 2.3).

2.1. General Response Functions of Submillimeter Spectroscopy

Let us start with the response functions of a system whose inputs are astronomical signals and outputs are the power spectrum of a spectrometer. In the following, we use a notation in which the input signal at an observed frequency, ν (radio frequency), is measured as the output at a measured frequency, f (i.e., intermediate frequency). The output power spectrum, $P(f, t)$, is a function of the measured frequency, f, and time, t. We assume that the response is linear, which applies to many cases of heterodyne receivers. This means that the system output is expressed as a single gain function, $G(\nu, f)$, which applies frequency conversion and amplification using the following equation:

$$P(f, t) = G(\nu, f)\, k_{\rm B}\, (T_{\rm in}(\nu, t) + T_{\rm noise}(\nu)), \tag{3}$$

where $k_{\rm B}$ is the Boltzmann constant, $T_{\rm in}$ is the brightness temperature input to a receiver, and $T_{\rm noise}$ is the equivalent noise temperature of the system. We define the frequency conversion from ν to f as

$$\nu = f + m, \tag{4}$$

where m is a frequency conversion term. In addition, $|m|$ is equivalent to the local-oscillator (LO) frequency in a heterodyne receiver. This is expressed as a function of time, $m(t)$, in frequency-modulation observations (Taniguchi et al. 2019); however, we do not consider the case hereafter (an application to this case is discussed in Section 5).

To correct for the gain, a blackbody at room temperature is measured as a calibrator, which is often referred to as a hot-load measurement or chopper-wheel calibration. Thus, there are two cases for the definition of the input temperature:

$$T_{\rm in}(\nu, t) = \begin{cases} \eta_{\rm fwd}(\nu)\, T_{\rm sky}(\nu, t) + (1 - \eta_{\rm fwd}(\nu))\, T_{\rm amb} & \text{(sky measurement)} \\ T_{\rm room} & \text{(calibrator measurement)}, \end{cases} \tag{5}$$

where $\eta_{\rm fwd}$ is the forward efficiency of a telescope feed, $T_{\rm sky}$ is the brightness temperature of signals from the sky, $T_{\rm amb}$ is the ambient temperature at a telescope site, and $T_{\rm room}$ is the room temperature around the receiver. We assume that the room temperature is constant during an observation.

There are also two cases for the definition of $T_{\rm sky}$:

$$T_{\rm sky}(\nu, t) = \begin{cases} \eta_{\rm atm}(\nu, t)\, T_\star(\nu, t) + (1 - \eta_{\rm atm}(\nu, t))\, T_{\rm atm} & \text{(on - source measurement)} \\ (1 - \eta_{\rm atm}(\nu, t))\, T_{\rm atm} & \text{(off - source measurement)}, \end{cases} \tag{6}$$

where $\eta_{\rm atm}$ is the line-of-sight atmospheric transmission and $T_{\rm atm}$ is the physical temperature of the atmosphere. In addition, $T_\star$ is the same as that defined in Equation (1) and is the term of interest. Because astronomical signals may change during an observation (e.g., on-the-fly mapping observations), it is expressed as a function of both time and frequency.

2.2. Reformulation of the PSW Method

We reformulate the PSW method (Equation (1)) to derive an observable as a product of input signals and response terms such as gain and efficiency. The keys are to unify the two cases of $T_{\rm sky}$ (Equation (6)) into a single expression and to take the difference between the sky and calibrator measurements, rather than between on-source and off-source measurements. Hereafter, we make the following assumptions:

$$T_{\rm atm} = T_{\rm amb} = T_{\rm room} \tag{7}$$

$$T_\star(\nu, t) < T_{\rm atm}. \tag{8}$$

The former is the same as what is assumed in Equation (1) and is used for the equation transformations in Section 2.2.2. Although we can still apply the proposed method in the case of $T_{\rm atm} \neq T_{\rm amb}$, this may affect the accuracy of absolute intensity scale (see also Section 5.2). We also assume that such physical temperatures should be estimated by another measurement, such as a thermometer or a weather monitor. The latter is a condition that ensures the matrix decomposition described in Section 2.3.

2.2.1. Unification of Sky Measurements

We start by introducing a position indicator, $\phi(t)$, which discriminates between on-source and off-source times during an observation:

$$\phi(t) = \begin{cases} 1 & \text{(when pointing to on - source position)} \\ 0 & \text{(when pointing to off - source position).} \end{cases} \tag{9}$$

Practically stated, it should be derived from an antenna log of the observation. Using the position indicator, we define unified astronomical signals, $T_{\rm ast}$, which express both on-source and off-source measurements at the same time:

$$T_{\rm ast}(\nu, t) = \begin{cases} T_\star(\nu, t) & (\phi(t) = 1) \\ 0 & (\phi(t) = 0), \end{cases} \tag{10}$$

and $T_{\rm sky}$ is then expressed as

$$T_{\rm sky}(\nu, t) = \eta_{\rm atm}(\nu, t)\, T_{\rm ast}(\nu, t) + (1 - \eta_{\rm atm}(\nu, t))\, T_{\rm atm}. \tag{11}$$

2.2.2. Difference Between Sky and Calibrator Measurements

We take the difference between the measurements of the sky and a calibrator to express a new observable for the following noise-removal method. As expressed in Equation (1), the PSW method takes the difference between on-source and off-source measurements. By contrast, the proposed sky-calibrator difference is expressed as a product of input signals and response terms, including the gain and efficiency. Using the equations above, we express the output power spectra measured for the sky and a calibrator as follows:

$$P_{\rm sky}(f,\,t) = G(\nu,f)\,k_{\rm B}\,[\eta_{\rm fwd}(\nu)\,T_{\rm sky}(\nu,\,t) + (1-\eta_{\rm fwd}(\nu))\,T_{\rm amb} + T_{\rm noise}(\nu)] \quad (12)$$

$$P_{\rm cal}(f,\,t) = G(\nu,f)\,k_{\rm B}\,[T_{\rm room} + T_{\rm noise}(\nu)]. \quad (13)$$

The sky-calibrator difference, $dP(f,\,t)$, is then expressed as follows:

$$\begin{aligned} dP(f,\,t) &\equiv P_{\rm sky}(f,\,t) - P_{\rm cal}(f,\,t) \quad (14)\\ &= G(\nu,f)\,k_{\rm B}\,[\eta_{\rm fwd}(\nu)\,(T_{\rm sky}(\nu,\,t) - T_{\rm amb}) + \underbrace{T_{\rm amb} - T_{\rm room}}_{=0}]\\ &= G(\nu,f)\,k_{\rm B}\,\eta_{\rm fwd}(\nu)\,[\eta_{\rm atm}(\nu,\,t)\,(T_{\rm ast}(\nu,\,t) - T_{\rm atm}) + \underbrace{T_{\rm atm} - T_{\rm amb}}_{=0}]\\ &= G(\nu,f)\,k_{\rm B}\,\eta_{\rm fwd}(\nu)\,\eta_{\rm atm}(\nu,\,t)\,(T_{\rm ast}(\nu,\,t) - T_{\rm atm}). \quad (15) \end{aligned}$$

Note that we ignore the cosmic microwave background (CMB) emission in the definition of $T_{\rm ast}$ (Equation (10)). If it is not negligible (e.g., observations at lower frequencies close to millimeter wavelength), $T_{\rm atm}$ in Equation (15) should be replaced with $T_{\rm atm} - T_{\rm cmb}(\nu)$, where $T_{\rm cmb}(\nu)$ is the brightness temperature of the CMB.

2.3. Signal Estimate by Low-rank and Sparse Matrix Decomposition

From Equation (15), $dP(f,\,t)$ is a product of the signal term $(T_{\rm ast} - T_{\rm atm})$ and the response terms $(Gk_{\rm B}\eta_{\rm fwd}\eta_{\rm atm})$. Here we introduce a dimensionless quantity, $X(f,\,t)$, as follows:

$$\begin{aligned} X(f,\,t) &\equiv \ln\left(-\frac{dP(f,\,t)}{k_{\rm B}\,T_{\rm atm}}\right)\\ &= \ln\{G(\nu,f)\,\eta_{\rm fwd}(\nu)\,\eta_{\rm atm}(\nu,\,t)\} + \ln\left(1-\frac{T_{\rm ast}(\nu,\,t)}{T_{\rm atm}}\right). \quad (16) \end{aligned}$$

This is a sum of two terms. The first term varies as the atmosphere fluctuates over time and always has nonzero values. As demonstrated in many submillimeter measurements (Dempsey et al. 2013; Cortés et al. 2016), the logarithm of $\eta_{\rm atm}(\nu,\,t)$ (i.e., the atmospheric opacity) is typically expressed as a linear function of line-of-sight precipitable water vapor, PWV(t):

$$\ln\,\eta_{\rm atm}(\nu,\,t) \simeq a(\nu)\,{\rm PWV}(t) + b(\nu), \quad (17)$$

where $a(\nu)$ and $b(\nu)$ are coefficients depending on ν. When $X(f,\,t)$ is discretely sampled in the time and frequency domains, it suggests that the first term of Equation (16) should be a low-rank matrix. By contrast, the second term is a sparse matrix because its elements only become nonzero when a telescope is pointing to the on-source position (i.e., $\phi(t) = 1$) and emission lines exist (i.e., $T_\star(\nu,\,t) \neq 0$). If the sparseness is ensured during an observation, $X(f,\,t)$ can be decomposed into low-rank (L) and sparse (S) components:

$$X = L + S + E, \quad (18)$$

where E is a noise term.[10] L and S correspond to the following:

$$L(\nu,f,\,t) \simeq \ln\,(G(\nu,f)\,\eta_{\rm fwd}(\nu)\,\eta_{\rm atm}(\nu,\,t)) \quad (19)$$

$$S(\nu,\,t) \simeq \ln\left(1-\frac{T_{\rm ast}(\nu,\,t)}{T_{\rm atm}}\right). \quad (20)$$

Although $T_{\rm ast}(\nu,\,t)$ is derived from $S(\nu,\,t)$, it is a noiseless estimate: if one prefers astronomical signals with errors as the final product, it is reasonable that the estimate of astronomical signals should be composed of both S and E. Finally, we define the estimate of astronomical signals, $\hat{T}_{\rm ast}$, using the low-rank component:

$$\hat{T}_{\rm ast}(\nu,\,t) = T_{\rm atm}\,[1 - \exp\,(X(f,\,t) - L(\nu,f,\,t))]. \quad (21)$$

For PSW observations, the integrated spectrum, $\bar{T}_\star(\nu)$, is derived as

$$\bar{T}_\star(\nu) = \frac{\int_0^{t_{\rm obs}} \hat{T}_{\rm ast}(\nu,\,t)\,\phi(t)\,dt}{\int_0^{t_{\rm obs}} \phi(t)\,dt}, \quad (22)$$

where $t_{\rm obs}$ is the total observation time of both on and off-source positions.

Figure 2 shows the integrated spectra of the mock observation (the same data in Figure 1) reduced by the proposed decomposition (Equation (21)) and the conventional direct on–off subtraction (Equation (1)). In this case, the noise level is improved by a factor of 1.44 ($\simeq\sqrt{2}$) while keeping the spectral shape of the astronomical signals.

The signal estimation works better when the astronomical signals of interest are sparse. The assurance of sparseness depends on the observing mode (e.g., fraction of on-source time) and spatial and spectral distribution of an astronomical target (e.g., number of channels the emission line enters and/or compactness of the target in the case of a mapping observation). In Sections 3 and 4, we demonstrate the case of fast-sampled PSW observations toward high-redshift galaxies where only a single emission line exists within a spectral band. Other cases such as frequency-modulation observations are discussed in Section 5.

3. Proposed Method

We describe the proposed method to achieve low-rank and sparse decomposition in this section. The method is based on the GoDec algorithm (Zhou & Tao 2011). We modify the sparse-identification step of the algorithm for the fast-sampled PSW observations. After introducing the mathematical notation in Section 3.1, we describe the algorithm for estimating astronomical signals from fast-sampled PSW observations in Section 3.2.

[10] Here, E appears because the noise terms (e.g., $T_{\rm noise}(\nu)$) are not constant but stochastic, which may follow the normal distribution. The sky-calibrator difference (Equations (14)–(16)) actually has an additional term related to E, but it is not explicitly described for simplicity.

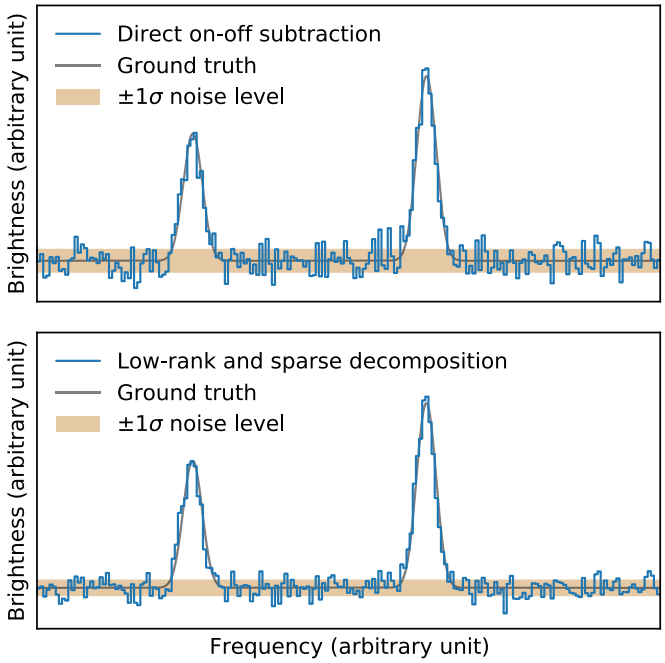

Figure 2. Integrated spectra of the mock observation (the same data in Figure 1) reduced by (top) the conventional direct on–off subtraction using Equation (1) and (bottom) low-rank and sparse decomposition using Equation (21). Gray lines are the ground truth of the astronomical signals. Dim orange spans indicate standard deviations of line-free channels of the spectra. Note that we independently create L, S, and E for Figure 1 and the spectra are made from them. This means that the spectrum of the bottom panel is not estimated from X using the GoDec algorithm shown in Section 3.

3.1. Mathematical Expression of Time-series Spectra

The time-series spectra of a fast-sampled PSW observation is represented as a discretely sampled matrix. For example, the dimensionless quantity, $X(f, t)$, is expressed as

$$X(f, t) \rightarrow x_{ij} \quad (1 \leqslant i \leqslant N_{\rm freq},\ 1 \leqslant j \leqslant N_{\rm time}), \tag{23}$$

where $N_{\rm freq}$ and $N_{\rm time}$ are the number of samples of the frequency and time domains, respectively. We use a bold upper-case letter to denote a matrix:

$$\boldsymbol{X} = (x_{ij}) \in \mathbb{R}^{N_{\rm freq} \times N_{\rm time}}. \tag{24}$$

The j-th column vector of a matrix corresponding to a sample spectrum is shown as a bold lower-case letter such as $\boldsymbol{x}_j$. The position indicator, $\phi(t)$, is also time-sampled and expressed as an $N_{\rm time}$-dimension column vector, $\boldsymbol{\phi}$. For other notations, $\mathbf{0}$ and $\mathbf{1}$ are matrices with all elements being zero and one, respectively. An element-wise product of two matrices (Hadamard product) is $\boldsymbol{X} \circ \boldsymbol{Y} \equiv (x_{ij} y_{ij})$. abs$(\cdot)$ and exp$(\cdot)$ are element-wise absolute-value and natural-exponential functions, respectively. The Frobenius norm of a matrix (the extension of a vector norm to a matrix) is expressed as $||\boldsymbol{X}||_F \equiv (\sum_i \sum_j x_{ij}^2)^{1/2}$.

Using the notations, low-rank and sparse decomposition (Equation (18)) becomes

$$\boldsymbol{X} = \boldsymbol{L} + \boldsymbol{S} + \boldsymbol{E}. \tag{25}$$

The estimates of astronomical signals and an integrated spectrum (Equations (21)–(22)), which we finally aim to derive, are expressed as follows:

$$\hat{\boldsymbol{T}}_{\rm ast} = T_{\rm atm}(\mathbf{1} - \exp(\boldsymbol{X} - \boldsymbol{L})), \tag{26}$$

$$\bar{\boldsymbol{t}}_\star = (\boldsymbol{\phi}^T \boldsymbol{\phi})^{-1}\, \hat{\boldsymbol{T}}_{\rm ast}\, \boldsymbol{\phi}. \tag{27}$$

Algorithm 1. GoDec algorithm

Input: $\boldsymbol{X} \in \mathbb{R}^{N_{\rm freq} \times N_{\rm time}}$, r, k, ε
Output: $\boldsymbol{L}$, $\boldsymbol{S} \in \mathbb{R}^{N_{\rm freq} \times N_{\rm time}}$
$n = 0$, $\boldsymbol{L}_0 = \boldsymbol{X}$, $\boldsymbol{S}_0 = \mathbf{0}$
While $||\boldsymbol{X} - \boldsymbol{L}_n - \boldsymbol{S}_n||_F^2 \,/\, ||\boldsymbol{X}||_F^2 > \varepsilon$
$\boldsymbol{L}_{n+1} = \sum_{j=1}^{r} \lambda_{jj}\, \boldsymbol{u}_j\, \boldsymbol{v}_j^T$ s.t. $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^T = \mathrm{SVD}(\boldsymbol{X} - \boldsymbol{S}_n)$
$\boldsymbol{\Omega}_{n+1} = \mathrm{SparseID}(\mathrm{abs}(\boldsymbol{X} - \boldsymbol{L}_{n+1});\, k)$
$\boldsymbol{S}_{n+1} = \boldsymbol{\Omega}_{n+1} \circ (\boldsymbol{X} - \boldsymbol{L}_{n+1})$
$n \coloneqq n + 1$
end While
return $\boldsymbol{L}_n$, $\boldsymbol{S}_n$

Algorithm 2. Sparse identification for fast-sampled PSW observations

Input: $\boldsymbol{Y} \in \mathbb{R}^{N_{\rm freq} \times N_{\rm time}}$, $\boldsymbol{\phi} \in \{0, 1\}^{N_{\rm time}}$, k
Output: $\boldsymbol{\Omega} \in \{0, 1\}^{N_{\rm freq} \times N_{\rm time}}$
function SparseID@PSW $\boldsymbol{Y}$; $\boldsymbol{\phi}$, k
$\boldsymbol{s} = (\boldsymbol{\phi}^T \boldsymbol{\phi})^{-1}\, \boldsymbol{Y}\boldsymbol{\phi}$
$\boldsymbol{\theta}_k = \mathrm{SparseID}(\mathrm{abs}(\boldsymbol{s});\, k)$
$\boldsymbol{\Omega} = \boldsymbol{\theta}_k\, \boldsymbol{\phi}^T$
return $\boldsymbol{\Omega}$
end function

3.2. GoDec Algorithm for Fast-sampled PSW Observations

We describe the GoDec algorithm and modify a part of it for estimating the astronomical signals of fast-sampled PSW observations. It is an iterative algorithm in which $\boldsymbol{L}$ and $\boldsymbol{S}$ are alternatively estimated, i.e., it assigns the low-rank approximation of $\boldsymbol{X} - \boldsymbol{S}$ to $\boldsymbol{L}$ and the sparse approximation of $\boldsymbol{X} - \boldsymbol{L}$ to $\boldsymbol{X}$. Algorithm 1 shows the GoDec algorithm (the notation is modified from Zhou & Tao (2011) to fit the present paper), where r is the rank of a low-rank matrix, k is the number of nonzero elements of a sparse matrix, and ε is a threshold for the convergence of the algorithm. In addition, SVD$(\cdot)$ conducts singular value decomposition[11], where $\boldsymbol{\Lambda}$ is a diagonal matrix of singular values, and $\boldsymbol{U}$ and $\boldsymbol{V}$ are left and right singular vectors, respectively. SPARSEID$(\cdot\,; k)$ is a sparse-identification step to compute an index matrix, $\boldsymbol{\Omega} = (\omega_{ij} \in \{0, 1\})$, whose elements only become unity if the corresponding elements of the input matrix are the first k-largest elements.

As the result of the GoDec algorithm, we expect the sparse matrix to extract the astronomical signals (Figure 1). In real observations, however, astronomical signals are often much weaker than the noise level in a spectrum at a sampling interval of $\sim$1 s: in high-redshift observations with a 50 m class telescope, the typical peak intensity of an emission line is $\sim$1 mK, and the signal-to-noise ratio (S/N) in a 1 s integrated spectrum is S/N $\sim 10^{-2}$. This indicates that the original sparse-identification step should be modified for the current purpose.

[11] Singular value decomposition (SVD) is one of the matrix decomposition methods, where it decompose a matrix into (left and right) singular vectors. Because the first r vectors represent r most dominant components in the matrix, it is sometimes used for low-rank approximation or dimensionality reduction of a matrix.

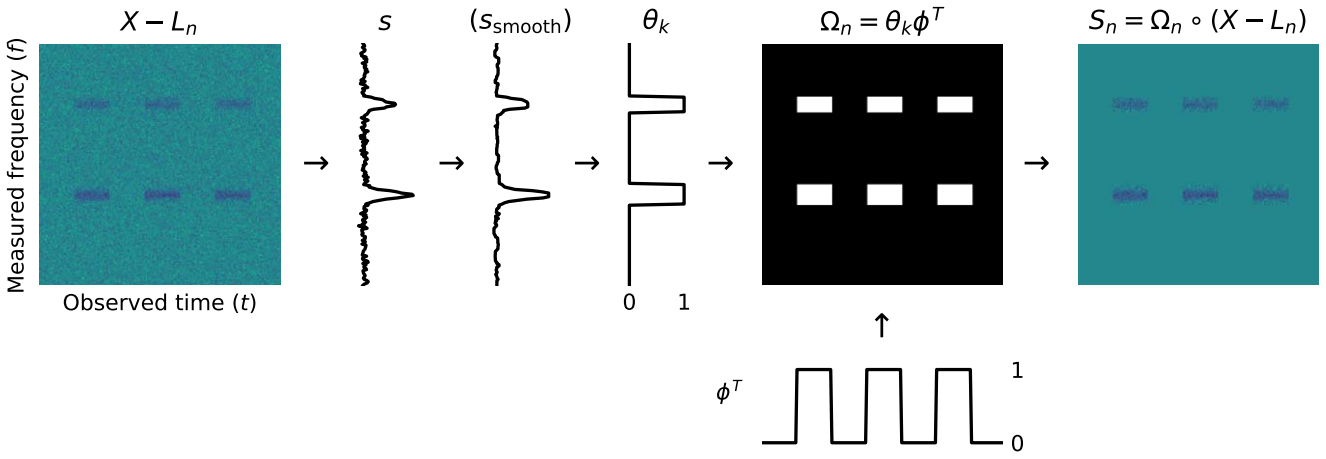

Figure 3. Schematic diagram of sparse identification for fast-sampled PSW observations (Algorithm 2). The top-left image represents a low-rank-subtracted matrix that should contain both astronomical signals and noise. In addition, $\boldsymbol{s}$ is a time-integrated on-source spectrum, and $\boldsymbol{s}_{\rm smooth}$ is a spectrum smoothed by a low-pass-like filter (optional step). Sparse identification (finding the first k-largest elements) is applied in the spectrum ($\boldsymbol{\theta}_k$), and the index matrix, $\boldsymbol{\Omega}$, is computed as an outer product of $\boldsymbol{\theta}_k$ and $\boldsymbol{\phi}$.

Herein, we propose a custom sparse-identification step, SPARSEID@PSW($\cdot$;$\boldsymbol{\phi}$, k), which employs information on the position indicator and the time-integrated spectrum, and replace SPARSEID with it. Figure 3 shows a schematic diagram of the step, where $\boldsymbol{s}$ is a column vector of the time-integrated on-source spectrum and $\boldsymbol{\theta}_k$ is a column vector of the index spectrum to indicate the first k-largest elements. The index matrix, $\boldsymbol{\Omega}$, is computed as the outer product of $\boldsymbol{\theta}_k$ and $\boldsymbol{\phi}$ (i.e., the intersection between on-source times and signal-detected frequencies). Algorithm 2 describes the step. Note that the number of nonzero elements in $\boldsymbol{S}$ is not k but $k(\boldsymbol{\phi}^T\boldsymbol{\phi})$.

As we will see in Section 4, the custom sparse-identification step enables us to identify astronomical signals even when $\mathrm{S/N} \ll 1$. The parameters (r, k) should be set a priori (i.e., hyperparameters) or optimized, for example, through a cross validation. The rank of the low-rank matrix, r, depends on the conditions of the atmosphere, the observed bandwidth, and the observed frequency. Taniguchi et al. (2019) demonstrates that $r \simeq 5$ is typically sufficient to remove correlated noise components in a frequency-modulation observation using the Nobeyama 45 m telescope, which ensures that r is much less than $\min(N_{\rm freq}, N_{\rm time}) \sim 10^3$. The number of nonzero elements in a time-integrated spectrum, k, depends on the number of channels at which emission lines are detected. In the case of a fast-sampled PSW observation toward a high-redshift emission line with a typical full width at half maximum of 1000 km s^{-1}, using a 4 GHz bandwidth in the 345 GHz band, the fraction of nonzero elements in $\boldsymbol{\Omega}$ (and thus $\boldsymbol{S}$) is $\sim$15%, which assures the sparseness.

As illustrated in Figure 3, smoothing a time-integrated spectrum before sparse identification is promising in the case of observations toward emission lines that have broader and fainter features (e.g., an outflow). In this case, the window length of a smoothing filer, w, is another parameter, which depends on the noise level of the spectrum or the spectral shape of the emission lines. We demonstrate the data reduction of actual fast-sampled PSW observations with spectral smoothing in Section 4.

4. Demonstration

We show an application of the proposed algorithms (Section 3) to fast-sampled PSW observations toward a high-redshift galaxy. We start by describing the conditions of the observations and data reduction using the conventional method (direct on–off subtraction and linear-baseline fitting) and the proposed method (Section 4.1). We then show the integrated spectra of the reduced data and demonstrate the improvement of the observation sensitivity achieved by the proposed method (Section 4.2). Finally, we investigate the noise characteristics of the reduced data to show the effect of low-rank noise removal (Section 4.3).

4.1. Data Description and Reduction

We use time-series spectra of fast-sampled PSW observations for an extremely luminous submillimeter galaxy PJ020941.3, which was originally found through submillimeter-to-far-infrared continuum imaging surveys of *Planck* (*Planck* Catalog of Compact Sources; Planck Collaboration et al. 2014) and *Herschel* (*Herschel* Stripe 82 Survey; Viero et al. 2014). Its spectroscopic redshift was determined to be $z_{\rm spec} = 2.5534 \pm 0.0002$ (Harrington et al. 2016) through a 3 mm observation of the redshifted CO ($J = 3$–2) line ($\nu_{\rm obs} = 97.314$ GHz) using the Redshift Search Receiver (RSR; Erickson et al. 2007) installed on the Large Millimeter Telescope (LMT; Schloerb 2008).

Our observations targeted on two redshifted emission lines of CO ($J = 4$–3) ($\nu_{\rm obs} = 129.739$ GHz) and CO ($J = 5$–4) ($\nu_{\rm obs} = 162.165$ GHz) using a 2 mm sideband-separating heterodyne receiver installed on the 50 m LMT (B4R[12]; Kawabe et al. in prep) connected to the XFFTS spectrometer (Klein et al. 2012). As summarized in Table 1, we carried out the fast-sampled PSW observations in late 2019 as part of a commissioning campaign of the B4R. For each observation, we took time-series spectra at a sampling rate of 1 Hz using two-sideband and dual-polarization settings. On-source and off-source spectra were alternatively observed for 10 s, and the sequence was repeated 30 times, which resulted in a total on-source time of 300 s. A 10 s measurement of a hot load preceded the observation. As a result, an observation yielded a matrix of continuous 600 s time-series spectra[13] whose

[12] http://lmtgtm.org/b4r/

[13] During the observation, we had a certain amount of antenna transition time between the two positions in which the spectrometer did not obtain data. This means that the spectra are not perfectly continuous and the system may have drifted.

Table 1
Observation Logs of the Target Taken by the B4R on the 50 m LMT

	CO ($J = 4$–3)	CO ($J = 5$–4)
Target name	PJ020941.3	
Target coordinates (J2000)	$\alpha = 02^{h}\,09^{m}\,41.3^{s}$, $\delta = +00°\,15'\,59''$	
Off-source relative coordinates (horizontal)	$d\mathrm{Az} = +00°\,01'\,00''$, $d\mathrm{El} = 00°\,00'\,00''$	
LMT project ID	2019S1B4RCommissioning	
LMT observation ID (hot load)	86889	86895
LMT observation ID (science target)	86890	86896
Observation date	2019 Nov 26	2019 Nov 26
Observation start time (UTC)	05:11:06	05:54:25
Opacity at 220 GHz	0.17	0.16
System noise temperature (K)	106	120
B4R first-LO frequency (GHz)	137.0	155.3
XFFTS frequency range (GHz)	128.9–131.4	160.9–163.4
256-ch binned channel width (GHz)	0.02 (40 km s^{-1})	0.02 (36 km s^{-1})
XFFTS data-dumping rate (sample s^{-1})	1.0	1.0
Integration time of hot load (s)	10.0	10.0
Integration time at on-source position (s)	300 (10 s × 30)	300 (10 s × 30)
Integration time at off-source position (s)	300 (10 s × 30)	300 (10 s × 30)
Transition time between two positions (s)	5–6	5–6

dimensions are $N_{\mathrm{time}} = 600$ and $N_{\mathrm{freq}} = 2^{15}$ (the number of channels of XFFTS).

After the observations, we carried out data reduction as follows: for common settings, we binned 256 frequency channels of a matrix to improve the signal-to-noise ratios ($N_{\mathrm{freq}} = 128$ after binning), which resulted in a velocity resolution of ∼36–40 km s^{-1} (see Table 1). We used a physical temperature of the atmosphere of $T_{\mathrm{atm}} = 273.0$ K. In the case of a conventional data reduction, we first time-integrated each 10-s observation of hot-load, on-source, and off-source measurements, and then applied chopper-wheel calibration to them (i.e., Equation (1)). This resulted in 30 spectra of calibrated astronomical signals. Linear baseline subtraction was applied to each spectrum. We finally integrated the spectra to obtain the final spectrum of the target.

In the case of the proposed data reduction, we integrated the 10-s hot-load measurement to obtain P_{cal}. We then obtained a dimensionless matrix, $\boldsymbol{X}$, according to Equations (14)–16 and applied Algorithm 1 to it using the custom sparse-identification step (Algorithm 2). As illustrated in Figure 3, we used a median filter to obtain a smoothed spectrum before sparse identification, $\boldsymbol{s}_{\mathrm{smooth}} = \mathrm{filter}(\boldsymbol{s}; w)$, where w is a parameter of the filter window length. We used the parameters $(r, k, w) = (5, 25, 5)$ for the CO (4–3) observation and $(r, k, w) = (7, 25, 7)$ for CO (5–4). The difference between the parameters assumed for CO (4–3) and CO (5–4) is attributed to a higher system noise temperature and a narrower velocity width per channel in the latter case. Finally, we estimated the time-series spectra of astronomical signals from $\boldsymbol{L}$ and $\boldsymbol{S}$ according to Equation (21) and time-integrated them to obtain an integrated spectrum of the target.

4.2. Improvement of Observation Sensitivity

We show the integrated spectra of the redshifted CO (4–3) emission line of PJ020941.3 in Figure 4 reduced by the conventional and proposed methods (those of CO (5–4) are shown in Figure 8). We detected the emission line at an observed frequency of 129.7 GHz in both methods. The spectral shapes and intensities of the emission lines reduced by both methods are consistent with each other. The full width at zero intensity of each line is ∼800 km s^{-1}, which is consistent with that of CO (3–2) obtained using RSR (Harrington et al. 2016).

The spectrum reduced by the proposed method in Figure 4 shows that the noise level is improved compared to that reduced by the conventional method. Because the proposed method does not apply on–off subtraction on noisy spectra, the factor $\sqrt{2}$ in Equation (2) is expected to decrease if the noise is ideally white. We summarize the achieved and expected noise levels of both cases in Table 2 (those of CO (5–4) are shown in Table 3). The achieved standard deviation of emission-free channels in the proposed case is lower than that of the conventional case by a factor of 1.67. This also means that a required noise level can be achieved in a 1.67^2-times shorter observation[14] with the proposed method.

The factor of 1.67 is slightly better than the expected improvement of $\sqrt{2}$. This extra improvement may be attributed to the subtraction of low-rank characteristics of atmospheric fluctuation and/or time variation of the instrumental response, which causes a baseline fluctuation in an integrated spectrum. As shown in Table 2, the expected noise level inferred from T_{sys} is close to the achieved value in the proposed case. By contrast, the achieved value in the conventional case is 1.2-times larger than the expected value, which suggests the existence of baseline fluctuations.

4.3. Characteristics of Noise on Time-series Spectra

We further investigate the characteristics of noise on the reduced time-series spectra. Figures 5 and 6 show the covariance matrices of time-series spectra of the CO (4–3) observation reduced by different methods and their histograms, respectively (those of CO (5–4) are shown in Figures 9 and 10). While the conventional case (center) indicates that nondiagonal covariance components (i.e., common-mode noises) remain in

[14] In the case of no overhead time. In an actual observation, constant overhead such as calibrator measurements may increase the total observation time.

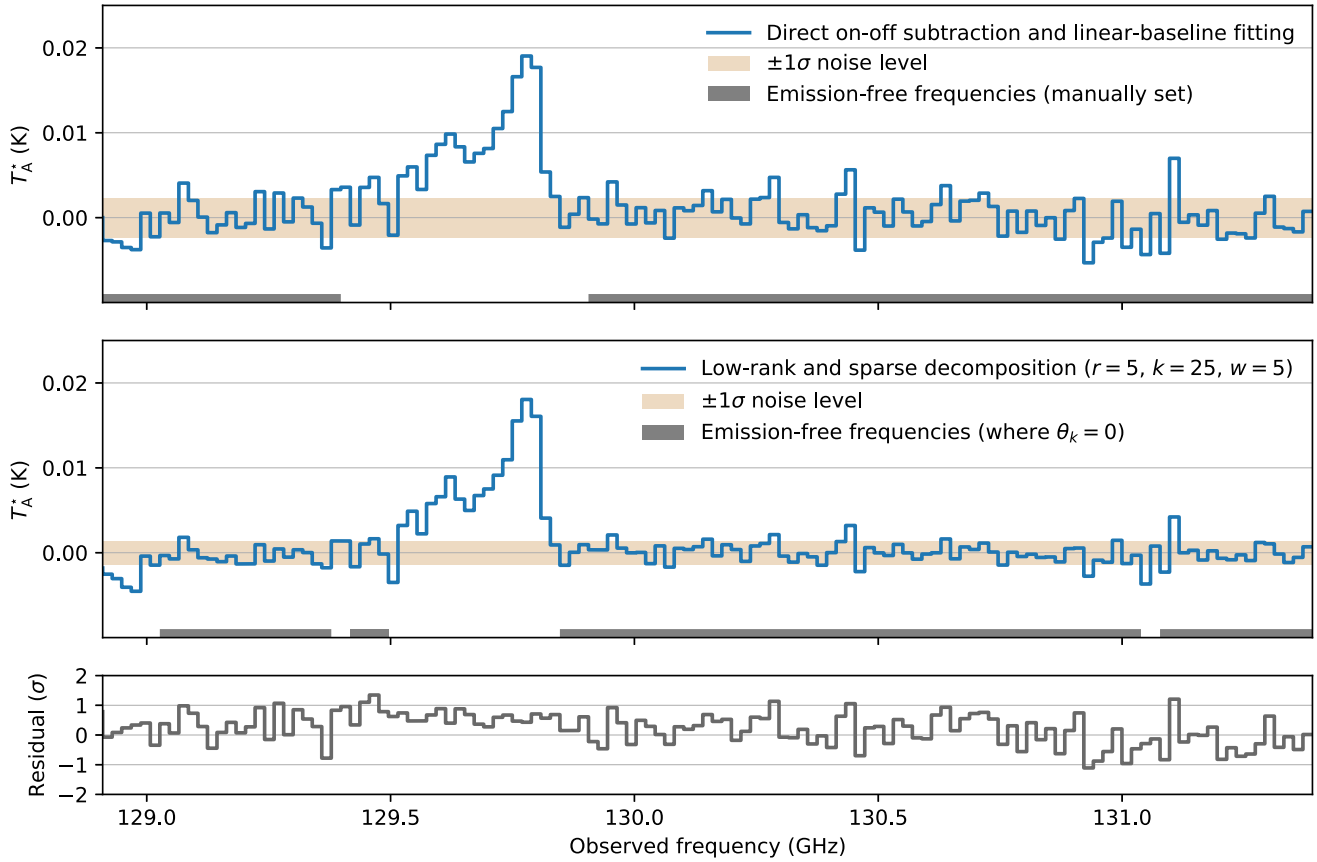

Figure 4. The integrated spectra of the redshifted CO (4–3) emission line of PJ020941.3 reduced by the conventional (top) and proposed (middle) methods. Dim orange spans indicate the achieved standard deviations of emission-free channels of the spectra (Table 2 left column). The unit of the vertical axis is $T_A^\star$, which corresponds to astronomical signals corrected for atmospheric attenuation. Other effects (e.g., main beam and/or beam dilution) are not corrected. In a conventional case, we excluded a frequency range of 129.4–129.9 GHz in a linear-baseline fitting. Emission-free frequencies are indicated as gray strips in the plot. In the proposed case, the parameters for the Algorithms 1–2 are $(r, k, w) = (5, 25, 5)$, where w is the window length of the median filter before sparse identification. Estimated emission-free frequencies (i.e., where $\theta_k = 0$) are indicated as gray strips in the plot. Other parameters are listed in Table 1. The bottom panel shows the difference spectrum between the two methods divided by the achieved standard deviation of the conventional method.

Table 2
Noise Levels of the CO (4–3) Spectra Reduced by Different Methods

	Achieved Standard Deviation of Emission-free Channels (mK)	Expected Noise Level Inferred from Equation (2) (mK)
Conventional method	2.30	1.92
Proposed method	1.38	1.35 (=1.92/$\sqrt{2}$)

Table 3
Noise Levels of the CO (5–4) Spectra Reduced by Different Methods

	Achieved Standard Deviation of Emission-free Channels (mK)	Expected Noise Level Inferred from Equation (2) (mK)
Conventional method	2.60	2.18
Proposed method	1.39	1.54 (=2.18/$\sqrt{2}$)

the integrated spectrum, most of them are subtracted using the proposed method (mean absolute covariance of <5%). This improvement can also be seen in the time-versus-noise plot in Figure 7 (that of CO (5–4) is shown in Figure 11). Although the achieved standard deviations of both cases are consistent at an integration time less than the PSW interval (∼10 s), the standard deviation of the conventional case becomes worse than the expected value at a longer integration time. This suggests the existence of atmospheric fluctuations and/or time variation of an instrumental response of slower than 0.1 Hz.

5. Discussion

We discuss the advantages, limitations, and possible applications of the proposed reformulation and algorithms in submillimeter spectroscopy.

5.1. Observation Sensitivity and Efficiency

One of the major advantages of the proposed method is the improvement in the observation sensitivity through a postprocess: If fast-sampled time-series spectra are available, the

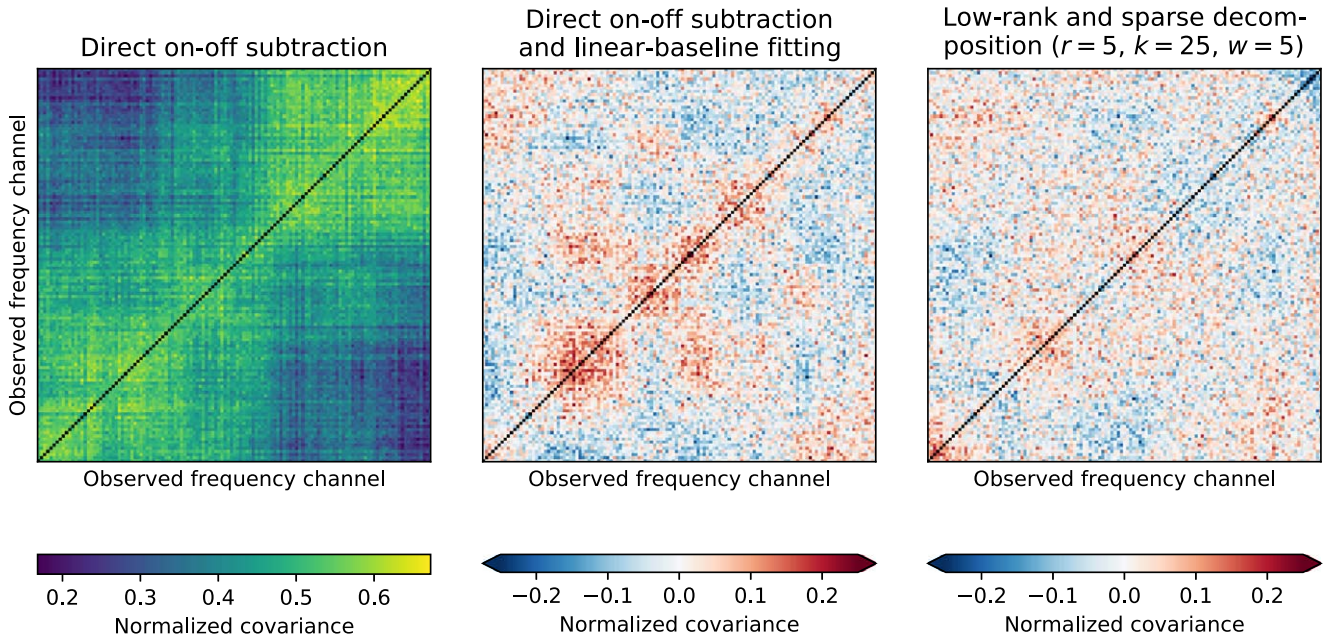

Figure 5. Covariance matrices of reduced time-series spectra of the redshifted CO (4–3) observation (note that they are normalized such that the diagonal elements are unity). Left: reduced only through a direct on–off subtraction. Center: reduced using a direct on–off subtraction and a linear-baseline fitting (the conventional method). Right: reduced through the GoDec algorithm for fast-sampled PSW observations (the proposed method). The parameters for these methods are listed in Figure 4.

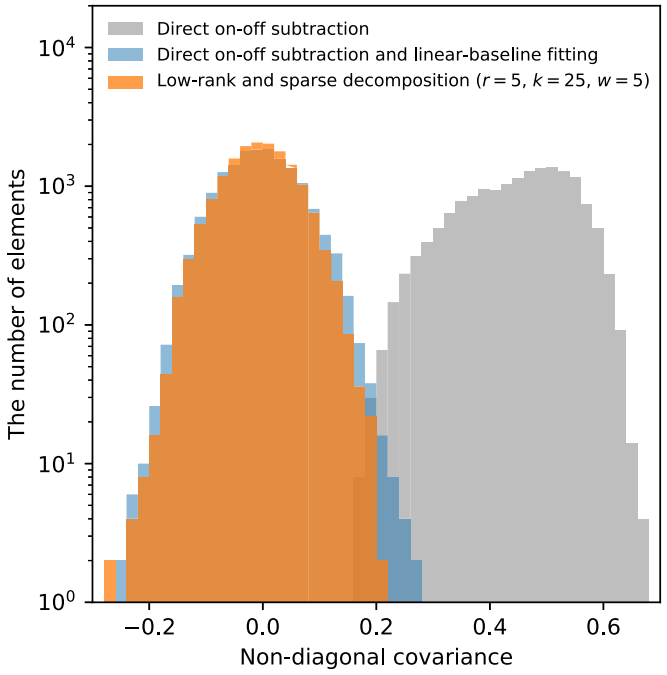

Figure 6. Histogram of nondiagonal elements of the covariance matrices in Figure 5.

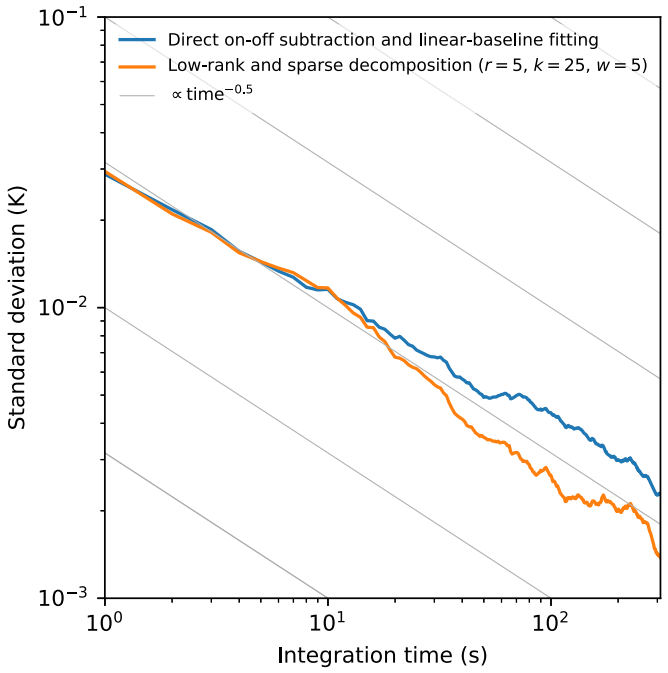

Figure 7. Integration time vs. achieved standard deviation of emission-free channels in the time-series spectra of the redshifted CO (4–3) observation reduced through the conventional (blue line) and proposed (orange line) methods. To obtain time-series spectra in the former case, we first time-integrated each 10 s off-source measurement and then subtract it from each 1 s on-source measurement. Gray sloped lines are proportional to the square root of the integration time (the standard deviation decreases in parallel in the case of white noise).

observation sensitivity can be improved by at least a factor of $\sqrt{2}$ even if they are obtained in a past observation. Because many single-dish telescopes already have the capability of on-the-fly mapping observations that obtain time-series spectra at >1 Hz, the fast-sampled PSW mode should be feasible. Thus, the proposed method can immediately improve their observation sensitivities.

The proposed method may also improve the observation efficiency (the fraction of on-source time over the total observation time). Because it does not require off-source measurements at a fixed sky position, the antenna transition time between the on and off-source positions can be partially used for the measurements, which was not available before. This means that the integration time at the off-source position can be reduced, while keeping the total off-source (sky) measurement time.

As another advantage, the proposed method can improve the baseline stability because the GoDec algorithm estimates common modes (singular vectors) in a matrix of time-series spectra. In addition to an atmospheric fluctuation, periodic baseline ripples caused by standing waves in an instrument not only worsen the observation sensitivity but they also make it difficult to distinguish emission lines from such artifacts. The proposed method may also be a powerful baseline fitting tool where fitting functions can be automatically determined from the data themselves.

As we demonstrated using the B4R implemented on the 50 m LMT, the observation sensitivities and baseline stabilities are improved by the proposed method. Although the targets are limited to luminous high-redshift galaxies with shorter observation times ($\sim$10 minutes), we show that it can be applied for single emission line observations. In a future study, the proposed method needs to be verified for observations of longer integration times (>1–10 hr) to establish it for deep spectroscopic science cases, as introduced in Section 1.

5.2. Calibration Accuracy

The proposed method does not improve the accuracy of absolute intensity scales compared to the conventional method because it continues to use a single hot-load calibration. The major assumptions of the method (Equation (7)) are sometimes unrealistic. For example, $T_{\rm atm}$ is expected to be less than $T_{\rm amb}$ under a good weather condition. We can still apply the method and expect the improvement; however, it may cause a systematic error in $\hat{T}_{\rm ast}$ in the same way as the conventional method. This is because an additional term, $G\,k_{\rm B}\,\eta_{\rm fwd}\,(T_{\rm atm}-T_{\rm amb})$, appears in the sky-calibrator difference (dP; Equation (15)), which is difficult to assume or estimate.

The systematic error can be reduced with better calibration. For example, one can directly derive the following quantity without the assumptions of Equation (7) using two-load calibration (i.e., measurements of hot and cold loads):

$$\frac{T_{\rm hot}-T_{\rm cold}}{T_{\rm atm}}\cdot\frac{P_{\rm hot}-P_{\rm sky}}{P_{\rm hot}-P_{\rm cold}} = \eta_{\rm fwd}\,\eta_{\rm atm}\left(1-\frac{T_{\rm ast}}{T_{\rm atm}}\right)-\eta_{\rm fwd}\left(1-\frac{T_{\rm amb}}{T_{\rm atm}}\right), \quad (28)$$

where $T_{\rm hot}$ and $T_{\rm cold}$ are the temperature of the hot and cold loads, respectively. $P_{\rm hot}$ and $P_{\rm cold}$ are given by Equation (13) whose $T_{\rm room}$ is replaced with $T_{\rm hot}$ and $T_{\rm cold}$, respectively. Unlike the additional term above, the second term in Equation (28) is close to zero in many cases. This means that the logarithm of Equation (28) is the sum of low-rank and sparse components and thus can be used as X. In the case of a more accurate calibration, where $T_{\rm sky}(\nu,\,t)$ can be directly derived (using sky-dip measurements, for example), the proposed algorithms can be applied to $T_{\rm sky}(\nu,\,t)-T_{\rm atm}$ instead of $dP(f,\,t)$. This strategy would be useful if an instrument has a nonlinear response to the signal (e.g., DESHIMA).

5.3. Observation Targets

Like other correlated noise-removal methods, the proposed method has certain limitations in terms of the spectral distribution of the emission. Because it estimates and removes common-mode spectra, it cannot be used for continuum observations where the emission uniformly enters all spectral channels. It also assumes the sparseness of the signal in a matrix. As we demonstrate in Section 4, the signal of $\sim$15% sparseness can be estimated properly. This suggests that observations of a single emission line (e.g., blind redshift surveys of distant galaxies) should be promising for most broadband spectrometers. It would be challenging, however, in the cases of multiple emission lines that occupy a substantial fraction of the spectrometer bandwidth (e.g., line surveys of nearby galaxies). In such cases, it would still be useful if the on-source fraction was less than 50% (e.g., mapping observations of compact sources). The feasibility of the proposed method for different target types should be further investigated using a large data set of either past or future observations.

5.4. Potential Applications

The proposed method has potential applications to many types of observations and/or instruments other than PSW observations by heterodyne receivers. It can be applied to observations with DESHIMA and other where the other efficient noise-removal method based on frequency modulation (Taniguchi et al. 2019) cannot be used. Moreover, because it does not depend on spatial scan patterns of antennas, it can be applied to on-the-fly mapping observations of compact objects. Taniguchi et al. (2019) also demonstrate that the correlated noise-removal method effectively reduces scanning effects (artificial stripes seen in an image along the scan direction) of such observations. As mapping observations become crucial to investigate the "3D property" of galaxies, the proposed method with 3D imagers would be expected to accelerate the survey speed of the investigation. An initial application of the proposed method to DESHIMA will be demonstrated in a future study.

6. Conclusions

In this paper, we propose a new noise-removal method based on low-rank and sparse decomposition to achieve continuous and noiseless estimates of the atmospheric emission for submillimeter single-dish spectroscopy. The conclusions are as follows:

1. We propose a new noise-removal method that achieves the continuous and noiseless estimate of atmospheric emission by obtaining fast-sampled time-series spectra of on and off-source measurements in a PSW observation and applying low-rank and sparse decomposition to them (Section 1).
2. We show that the time-series astronomical spectra of a single-dish telescope with a hot-load calibration can be expressed as a sum of low-rank atmospheric emission and sparse astronomical signals by reformulating the PSW method (Section 2).
3. We show that GoDec, one of the low-rank and sparse decomposition algorithms, can be applied to a matrix of time-series spectra with a custom sparse-identification step for fast-sampled PSW observations (Section 3).
4. We demonstrate that the proposed method improves the observation sensitivity by a factor of 1.67 using the data of fast-sampled PSW observations obtained by the B4R on the 50 m LMT (Section 4). We find that the improvement is better than the expected value of $\sqrt{2}$,

which suggests that the proposed method can also reduce the baseline fluctuation of an integrated spectrum.

5. We discuss the advantages, limitations, and potential applications of the proposed method (Section 5), and propose the application of it to future ultrawideband spectrometers such as DESHIMA. By contrast, application to long-integrated observations has yet to be investigated, and will be demonstrated in future studies.

We thank the anonymous referee for fruitful comments. This work is supported by KAKENHI (Nos. 15H02073, 17H06130, and 20H01951). This paper makes use of data taken by the Large Millimeter Telescope Alfonso Serrano (LMT) in Mexico. The LMT project is a joint effort of the Instituto Nacional de Astrófisica, Óptica, y Electrónica (INAOE) and the University of Massachusetts at Amherst (UMASS). We also appreciate the support of the technical staff and the support scientists of the LMT during the commissioning campaign of the B4R.

Facility: LMT (B4R).

Software: Astropy (The Astropy Collaboration et al. 2013, 2018), NumPy (Harris et al. 2020), matplotlib (Hunter 2007), scikit-learn (Pedregosa et al. 2011), SciPy (Virtanen et al. 2020), pandas (McKinney 2010), xarray (Hoyer & Hamman 2017).

Appendix A
Results and Analyses of the CO (5–4) Observation

Figures 8–11 show the results and analyses of the CO (5-4) observation (corresponding to those of the CO (4-3) observation in Figures 4–7, respectively).

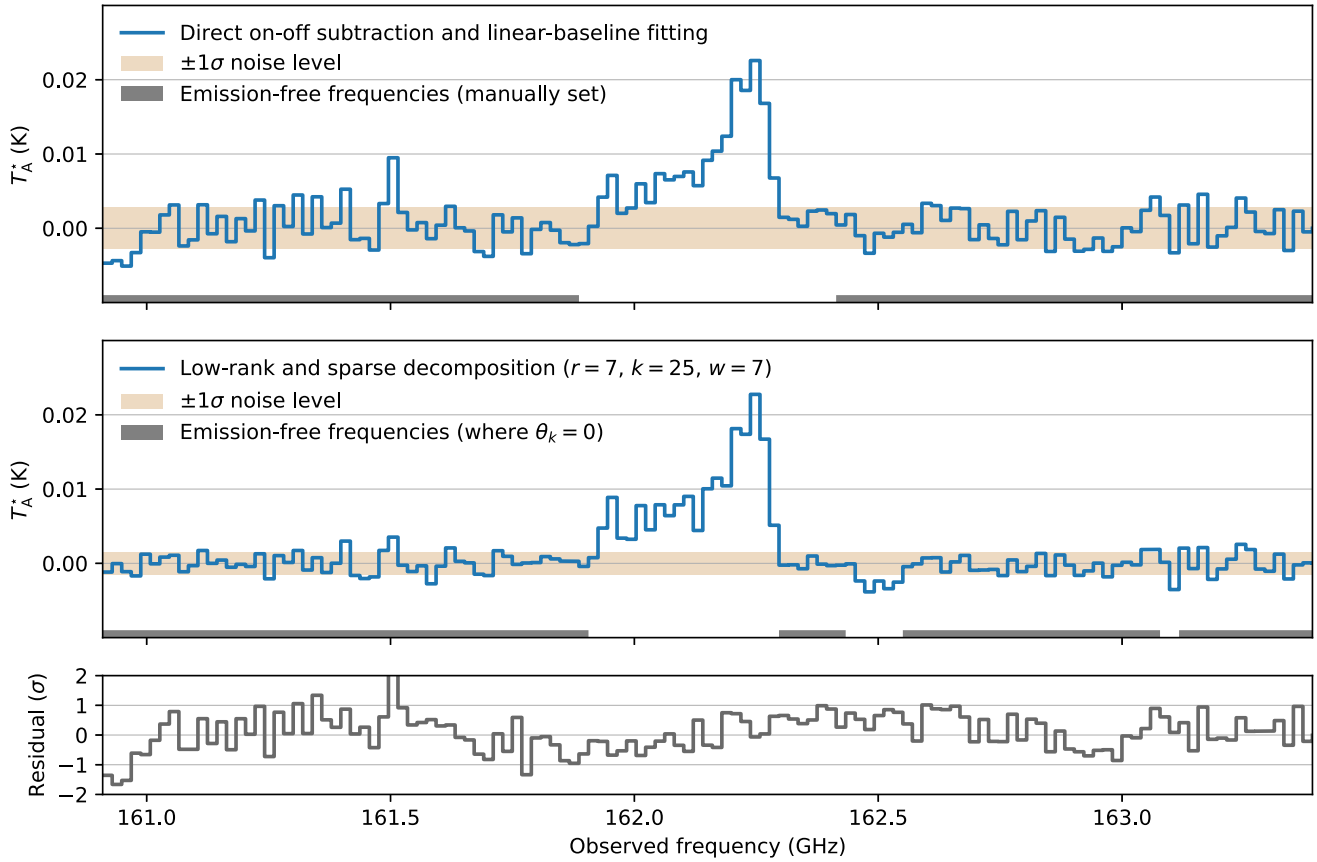

Figure 8. The integrated spectra of the redshifted CO (5–4) emission line of PJ020941.3 reduced by the conventional (top) and proposed (middle) methods. Dim orange spans indicate the achieved standard deviations of emission-free channels of the spectra (Table 3 left column). The unit of the vertical axis is $T_A^\star$, which corresponds to astronomical signals corrected for atmospheric attenuation. Other effects (e.g., main beam and/or beam dilution) are not corrected. In the conventional case, we excluded a frequency range of 161.9–162.4 GHz in a linear-baseline fitting. Emission-free frequencies are indicated as gray strips in the plot. In the proposed case, the parameters for the Algorithms 1–2 are $(r, k, w) = (7, 25, 7)$, where w is the window length of the median filter before sparse identification. Estimated emission-free frequencies (i.e., where $\theta_k = 0$) are indicated as gray strips in the plot. Other parameters are listed in Table 1. The bottom panel shows the difference spectrum between the two methods divided by the achieved standard deviation of the conventional method.

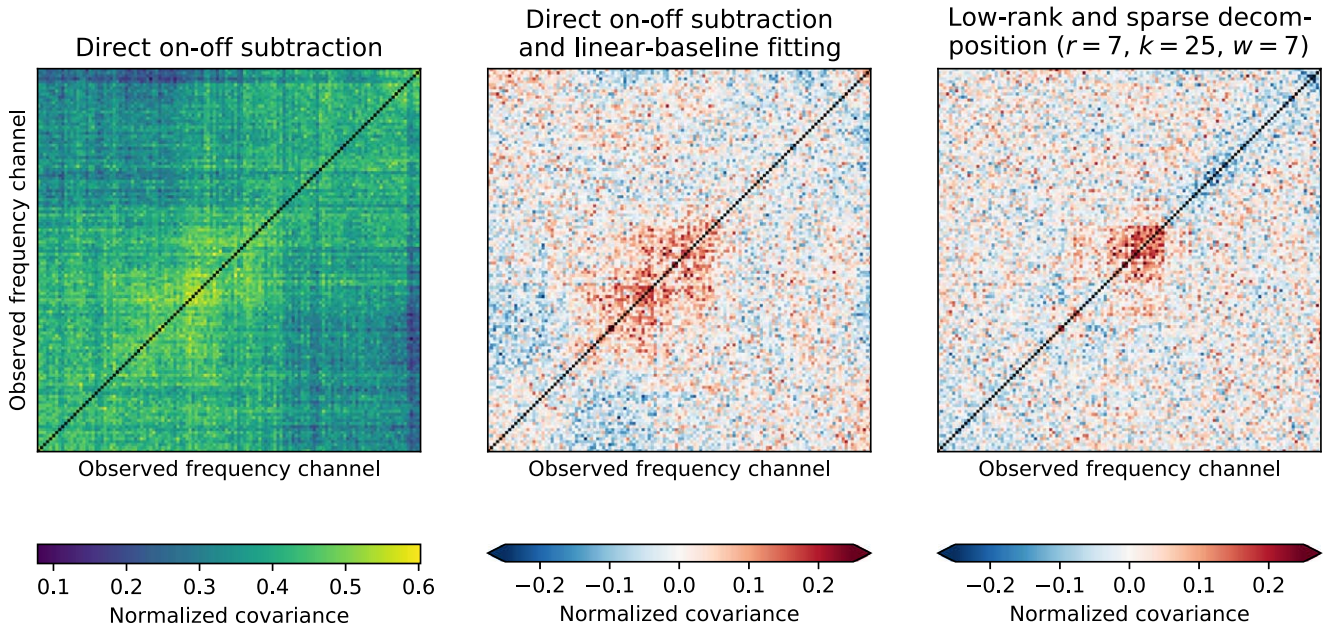

Figure 9. Covariance matrices of reduced time-series spectra of the redshifted CO (5–4) observation (note that they are normalized such that the diagonal elements are unity). Left: reduced using only a direct on–off subtraction. Center: reduced through a direct on–off subtraction and a linear-baseline fitting (the conventional method). Right: reduced using the GoDec algorithm for fast-sampled PSW observations (the proposed method). The parameters for these methods are listed in Figure 8.

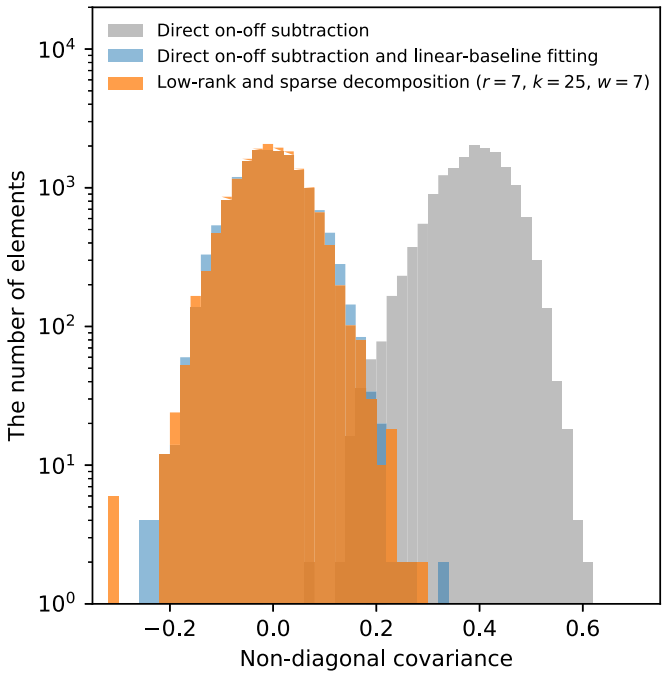

Figure 10. Histogram of nondiagonal elements of the covariance matrices in Figure 9.

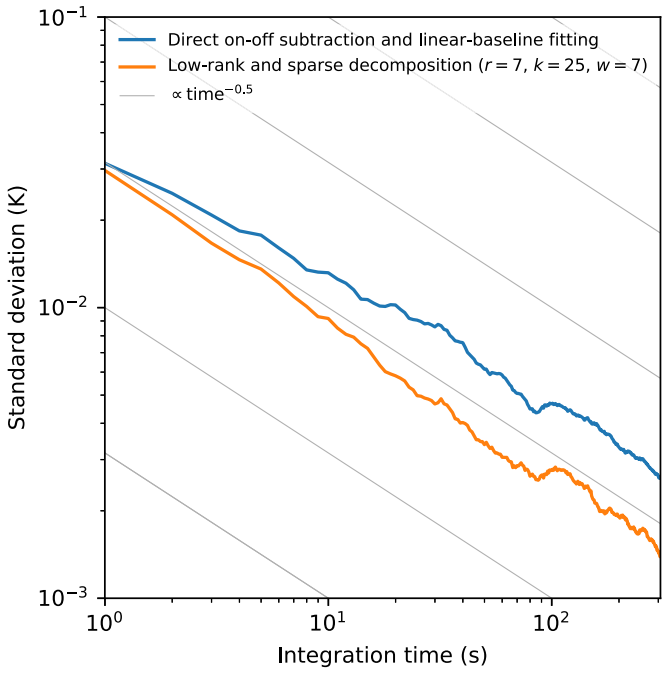

Figure 11. Integration time vs. achieved standard deviation of emission-free channels in the time-series spectra of the redshifted CO (5–4) observation reduced by the conventional (blue line) and proposed (orange line) methods. To obtain time-series spectra in the former case, we first time-integrate each 10-s off-source measurement and then subtract it from each 1-s on-source measurement. Gray sloped lines are proportional to the square root of the integration time (the standard deviation decreases in parallel in the case of white noise).

Table 4
Glossary of Section 1

Notation	First Appearance	Description
P	Equation (1)	Output power spectrum of a receiver-spectrometer system.
$T_\star$	Equation (1)	Brightness temperature of the astronomical signals corrected for the atmospheric transmission. Both signals from a target and CMB are included. Beam dilution may be corrected to obtain the intrinsic brightness temperature of the target.
$T_{\rm sys}$	Equation (2)	System noise temperature.
$\sigma_\star$	Equation (2)	Standard-deviation noise level of $T_\star$.
$\Delta\nu$	Equation (2)	Frequency channel width of a spectrometer.
$t_{\rm on}$	Equation (2)	Total on-source time of an observation.

Table 5
Glossary of Section 2

Notation	First Appearance	Description
$k_{\rm B}$	Equation (3)	The Boltzmann constant.
t	Equation (3)	Measured time or elapsed time from some origin.
ν	Equation (3)	Observed frequency or radio frequency (RF). Dependent on time, $\nu(t)$, in a frequency modulation observation.
f	Equation (3)	Measured frequency in a spectrometer. Equivalent to an intermediate frequency (IF) in a heterodyne receiver. $f=\nu$ for a direct detector.
m	Equation (4)	Frequency conversion term between the measured and observed frequencies. Here, $\lvert m\rvert$ is equivalent to the LO frequency in a heterodyne receiver. $m=0$ in a direct detector. Dependent on time, $m(t)$, in a frequency modulation observation.
ϕ	Equation (9)	Position indicator, which discriminates between on ($\phi(t)=1$) and off-source ($\phi(t)=0$) times during an observation.
G	Equation (3)	Gain between an RF input and an IF output.
$T_{\rm in}$	Equation (3)	Input brightness temperature to a receiver.
$T_{\rm noise}$	Equation (3)	Equivalent noise temperature of a receiver-spectrometer system.
$T_{\rm sky}$	Equation (5)	Brightness temperature of the signals from the sky.
$T_{\rm ast}$	Equation (11)	Brightness temperature of astronomical signals corrected for the atmospheric transmission, which expresses both on-source ($T_{\rm ast}=T_\star$) and off-source ($T_{\rm ast}=0$; virtual) measurements.
$T_{\rm amb}$	Equation (5)	Ambient temperature around a telescope.
$T_{\rm room}$	Equation (5)	Room temperature around the receiver.
$T_{\rm atm}$	Equation (6)	Physical temperature of the atmosphere.
$\eta_{\rm fwd}$	Equation (5)	Forward efficiency of telescope feed.
$\eta_{\rm atm}$	Equation (6)	Line-of-sight atmospheric transmission.
X	Equation (16)	Element of a two-dimensional matrix calculated from the powers of the sky and a hot load and considered to be decomposed into low-rank and sparse matrices.
L	Equation (18)	Element of a two-dimensional matrix whose rank is expected to be much lower than full rank.
S	Equation (18)	Element of a two-dimensional matrix whose elements are mostly expected to be zero.
E	Equation (18)	Noise arose from the atmosphere and a receiver-spectrometer system.
$t_{\rm obs}$	Equation (22)	Total observation time of both on and off-source positions (overheads are not included).

Table 6
Glossary of Section 3

Notation	First appearance	Description
$N_{\rm freq}$	Equation (23)	The number of frequency sample of a matrix.
$N_{\rm time}$	Equation (23)	The number of time sample of a matrix.
k	Algorithm 1	The number of nonzero elements of a vector or a matrix.
r	Algorithm 1	The rank of a low-rank matrix.
SVD	Algorithm 1	Singular value decomposition ($\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^T=\mathrm{SVD}(\boldsymbol{X})$).
$\boldsymbol{\Omega}$	Algorithm 1	Index matrix for sparse identification in a matrix.
$\boldsymbol{\theta}$	Algorithm 2	Index vector for sparse identification in a vector.
$\boldsymbol{s}$	Algorithm 2	Time-integrated on-source spectrum.
w	Figure 3	Filter length of a smoothed spectrum.

Appendix B
Glossary Lists

Tables 4–6 summarize the mathematical symbols that appeared throughout the paper.

ORCID iDs

Akio Taniguchi https://orcid.org/0000-0002-9695-6183
Yoichi Tamura https://orcid.org/0000-0003-4807-8117
Shiro Ikeda https://orcid.org/0000-0002-2462-1448
Tatsuya Takekoshi https://orcid.org/0000-0002-4124-797X
Ryohei Kawabe https://orcid.org/0000-0002-8049-7525

References

Candès, E. J., Li, X., Ma, Y., & Wright, J. 2011, J. ACM, 58, 11
Candès, E. J., Romberg, J., & Tao, T. 2006, ITIT, 52, 489
Chapin, E. L., Berry, D. S., Gibb, A. G., et al. 2013, MNRAS, 430, 2545

Cortés, F., Reeves, R., & Bustos, R. 2016, RaSc, 51, 1166
Dempsey, J. T., Friberg, P., Jenness, T., et al. 2013, MNRAS, 430, 2534
Donoho, D. L. 2006, ITIT, 52, 1289
Endo, A., Karatsu, K., Laguna, A. P., et al. 2019, JATIS, 5, 1
Endo, A., Karatsu, K., Tamura, Y., et al. 2019, NatAs, 3, 989
Erickson, N., Narayanan, G., Goeller, R., & Grosslein, R. 2007, in ASP Conf. Ser. 375, From Z-Machines to ALMA: (Sub)Millimeter Spectroscopy of Galaxies, ed. A. J. Baker et al. (San Francisco, CA: ASP), 71
Harrington, K. C., Yun, M. S., Cybulski, R., et al. 2016, MNRAS, 458, 4383
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Heiles, C. 2007, PASP, 119, 643
Hoyer, S., & Hamman, J. J. 2017, JORS, 5, 10
Hunter, J. D. 2007, CSE, 9, 90
Iwai, K., Kubo, Y., Ishibashi, H., et al. 2017, EP&S, 69, 95
Kawabe, R., Kohno, K., Tamura, Y., et al. 2016, Proc. SPIE, 9906, 779
Klaassen, P., Mroczkowski, T., Bryan, S., et al. 2019, BAAS, 51, 7
Klein, B., Hochgürtel, S., Krämer, I., et al. 2012, A&A, 542, L3
Kohno, K., Tamura, Y., Inoue, A., et al. 2019, Astro2020: Decadal Survey on Astronomy and Astrophysics, Vol. 2020 (Washington, DC: National Academies), 402
Kojima, T., Kiuchi, H., Uemizu, K., et al. 2020, A&A, 640, L9
Lou, Z., xi Zuo, Y., jun Yao, Q., et al. 2020, ApOpt, 59, 3353
McKinney, W. 2010, Proc. 9th Python in Science Conf., 445, 56
Morii, M., Ikeda, S., Sako, S., & Ohsawa, R. 2017, ApJ, 835, 1
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, J. Mach. Learn. Res., 12, 2825
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A28
Schieder, R., & Kramer, C. 2001, A&A, 373, 746
Schloerb, F. P. 2008, Proc. SPIE, 7012, 299
Taniguchi, A., Tamura, Y., Kohno, K., et al. 2019, PASJ, 72, 2
The Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, T. A.
The Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
The EHT Collaboration, et al. 2019, ApJL, 875, 4
Tibshirani, R. 1996, J. R. Stat. Soc. Series B Stat. Methodol., 58, 267
Uemura, M., Kawabata, K. S., Ikeda, S., & Maeda, K. 2015, PASJ, 67, 55
Ungerechts, H., Brunswig, W., Penalver, J., Perrigouard, A., & Sievers, A. 2000, Proc. SPIE, 4009, 327
Viero, M. P., Asboth, V., Roseboom, I. G., et al. 2014, ApJS, 210, 22
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wheeler, J., Hailey-Dunsheath, S., Shirokoff, E., et al. 2016, Proc. SPIE, 9914, 904
Wilson, T. L., Rohlfs, K., & Huttemeister, S. 2012, Tools of Radio Astronomy (Berlin: Springer)
Zhou, T., & Tao, D. 2011, in Proc. 28th Int. Conf. on Machine Learning (ICML-11) (Bellevue, WA: ICML), 33
Zuo, S., Chen, X., Ansari, R., & Lu, Y. 2018, AJ, 157, 4