

# A Modified Cross-correlation Beamforming (MCBF) for Extracting Multimode Surface Wave Dispersion Curves from Ambient Seismic Noise

Tongwei Qin<sup>1</sup> and Laiyu Lu<sup>2</sup>

<sup>1</sup>Institute of Geophysics

<sup>2</sup>Institute of Geophysics, China Earthquake Administration

November 22, 2022

## Abstract

Beamforming (BF) and Frequency-Bessel transform (F-J) have been demonstrated to extract multimode surface wave dispersion curves from ambient seismic noise. F-J method implicitly assumes the structure under the array is laterally isotropic. As for the conventional BF method, although the azimuth-dependence phase velocity can be measured, the fictitious azimuth anisotropy created by array geometry would be projected into the result. In this paper, the weighted and modified cross-correlation beamforming (WCBF and MCBF) schemes are proposed to extract the multimode surface wave dispersion curves with sufficient resolution using quite short noise recordings. Compared with the conventional BF, only the plane waves with the azimuth consistent with the interstation orientations are considered in MCBF and the search over the incident plane waves from different azimuth is omitted. The azimuth-dependence velocity can therefore be extracted by MCBF, independent of the array geometry. As far as the measurement of azimuth-averaged velocity is concerned, we show that BF is equivalent with F-J. The explicit relationship between BF and F-J methods is derived. For the finite sampling in practical applications, the theoretical representations of the dispersion image generated by BF technique under different imaging conditions are given. These representations can be used to investigate analytically the features of the dispersion images in frequency-velocity domain and how the aliasing is eliminated by improved imaging condition. The proposed methods are validated for the synthetic data as well as the real data from the dense array at different scales.

1 **A Modified Cross-correlation Beamforming (MCBF) for Extracting**  
2 **Multimode Surface Wave Dispersion Curves from Ambient Seismic**  
3 **Noise**

4 **Tongwei Qin and Laiyu Lu\***

5 Institute of Geophysics, China Earthquake Administration, Beijing, 100081, China

6 \*Corresponding author: Laiyu Lu ([laiyulu@cea-igp.ac.cn](mailto:laiyulu@cea-igp.ac.cn))

7 **Key Points:**

- 8 • A modified beamforming is proposed to extract multimode surface wave dispersion curves  
9 using short noise recordings
- 10 • The explicit relationship between cross-correlation beamforming and Frequency-Bessel  
11 transform is given
- 12 • An improved scheme is proposed to remove the aliasing artifacts, and its validity is  
13 demonstrated based on the synthetic and real data

14

## 15 **Abstract**

16 Beamforming (BF) and Frequency-Bessel transform (F-J) have been demonstrated to extract  
17 multimode surface wave dispersion curves from ambient seismic noise. F-J method implicitly  
18 assumes the structure under the array is laterally isotropic. As for the conventional BF method,  
19 although the azimuth-dependence phase velocity can be measured, the fictitious azimuth  
20 anisotropy created by array geometry would be projected into the result. In this paper, the  
21 weighted and modified cross-correlation beamforming (WCBF and MCBF) schemes are proposed  
22 to extract the multimode surface wave dispersion curves with sufficient resolution using quite  
23 short noise recordings. Compared with the conventional BF, only the plane waves with the azimuth  
24 consistent with the interstation orientations are considered in MCBF and the search over the  
25 incident plane waves from different azimuth is omitted. The azimuth-dependence velocity can  
26 therefore be extracted by MCBF, independent of the array geometry. As far as the measurement  
27 of azimuth-averaged velocity is concerned, we show that BF is equivalent with F-J. The explicit  
28 relationship between BF and F-J methods is derived. For the finite sampling in practical applications,  
29 the theoretical representations of the dispersion image generated by BF technique under different  
30 imaging conditions are given. These representations can be used to investigate analytically the  
31 features of the dispersion images in frequency-velocity domain and how the aliasing is eliminated  
32 by improved imaging condition. The proposed methods are validated for the synthetic data as well  
33 as the real data from the dense array at different scales.

## 34 **Plain Language Summary**

35 Benefiting from the advance of seismic interferometry technology, traditional array-based methods  
36 developed to process the data from the event or active source can be directly or redesigned to  
37 process the virtual source records. Beamforming (BF) and Frequency-Bessel transform (F-J) have  
38 been demonstrated to extract multimode surface wave dispersion curves from ambient seismic  
39 noise. Both methods assume that the structure beneath the array is laterally uniform. The weighted  
40 and modified cross-correlation beamforming (WCBF and MCBF) schemes are proposed in this  
41 paper. It was proved that multimode dispersion curves of surface wave can be measured by WCBF  
42 and MCBF with sufficient resolution using quite short noise recordings. Moreover, the azimuth-

43 dependence velocity can be extracted by MCBF, independent of the array geometry. The explicit  
44 relationships and equivalence between BF and F-J are proved as far as the application in extracting  
45 multimode dispersion curves is concerned. An improved imaging condition is suggested to remove  
46 the artificial aliasing based on the theoretical representations of MCBF for the finite sampling on  
47 spatial wavefield in practical applications. The proposed methods are validated for the synthetic  
48 data as well as the real data from the dense array at different scales.

## 49 **1. Introduction**

50 The theory of seismic interferometry (SI) suggests the Green's function can be retrieved by cross-  
51 correlating the seismic ambient noise recorded at two stations ([Lobkis and Weaver, 2001](#); [Campillo  
52 and Paul, 2003](#)). That is, the records of a virtual source can be constructed by calculating the noise  
53 cross-correlation function (NCF) of the interstation. Benefiting from the advance in SI, as well as  
54 the deployment of large and dense arrays, array-based schemes such as spatial autocorrelation  
55 (SPAC) ([Aki, 1957](#); [Yamaya et al., 2021](#)), Frequency-Bessel transform (F-J) ([J. Wang et al., 2019](#)) and  
56 beamforming (BF, or called CBF, cross-correlation beamforming) ([Harmon et al., 2008](#); [Roux and  
57 Ben-Zion, 2017](#); [K. Wang et al., 2020](#)), are proposed or redesigned to extract the multimode surface  
58 wave dispersion curves using ambient seismic noise. Compared with traditional noise-based two-  
59 station method, where the pure-path inversion is required after extracting the interstation  
60 dispersion curves, array-based method provides an opportunity to measure directly the lateral  
61 variation of the velocity using the subsets of the array ([Roux and Ben-Zion, 2017](#); [K. Wang et al.,  
62 2020](#)). Another advantage of array-based methods over two-station surface wave method is their  
63 ability for extracting the multimode dispersion curves. The join of higher modes will increase the  
64 stability of the surface wave inversion and provide constraints on the deeper structure ([Xia et al.,  
65 2003](#)). Therefore, the array technology for extracting multimode surface wave dispersion with high  
66 accuracy using seismic ambient noise has attracted the attention of seismic community ([Roux and  
67 Ben-Zion, 2017](#); [J. Wang et al., 2019](#); [K. Wang et al., 2020](#); [Yamaya et al., 2021](#); [Qin et al., 2022](#)).  
68 Precise measurement of dispersion curves is the basis for high-resolution 3D S-wave velocity  
69 imaging using surface waves.

70 F-J method is rooted in the theoretical representation of the wavefield at the free surface of a  
71 layered model. For an isotropic layered model, the wavefield in the frequency domain at distance  
72  $r$  can be written as the Fourier Bessel integral over wavenumber  $k$  (Harkrider, 1964; Ben-  
73 Menahem & Singh, 1968; Chen, 1999). The integral kernel is expressed as a fractional form related  
74 to the structure and source parameters. The surface wave is given by the integral contribution of  
75 the residues determined by the roots that make the denominator of the kernel vanish. The kernel  
76 can thereby be obtained by taking the inverse Fourier-Bessel transform over the propagation  
77 distance  $r$ . As a result, in frequency-velocity (f-v) or frequency-wavenumber (f-k) domain the  
78 peaks of the kernel would be associated with the eigenvalues of surface wave. Forbriger (2003)  
79 has applied Fourier-Bessel transform to extract dispersion curves of multimode Rayleigh wave  
80 using the seismogram gather at the surface excited by a hammer source. J. Wang et al. (2019)  
81 apply similar Fourier-Bessel transform to the records of the virtual source, and call it Frequency-  
82 Bessel (F-J) transform.

83 When F-J transform is used for virtual recordings, the fact that the eigenvalues of the surface waves  
84 are associated with the peaks in f-v or f-k domain can also be explained via the orthogonality of  
85 the Bessel functions. Taking the vertical component of Rayleigh waves as an example, the NCF in  
86 the time domain corresponds to the SPAC coefficient in the frequency domain (Chávez-García and  
87 Luzón, 2005; Tsai and Moschetti, 2010; Lu, 2021), i.e., the zero-order Bessel function  $J_0(k_n r)$  of  
88 the first kind with argument  $k_n r$ , where  $k_n$  is the eigen-wavenumber of the Rayleigh wave and  $r$   
89 is the interstation distance. F-J transform implies to compute the integral of  $J_0(k_n r)J_0(kr)r$  over  
90 the distance  $r$  from 0 to infinite. Due to the orthogonality of the Bessel function, the F-J  
91 spectrogram would achieve a maximum at  $k = k_n$ . Considering this property, Hu et al. (2020)  
92 extends F-J to NCFs of cross components and to estimate dispersion curves of Love wave using the  
93 orthogonality of Bessel function with different orders. The artifacts caused by aliasing can be  
94 eliminated by considering only the waves propagating in one direction, replacing Bessel function  
95 with Hankel function, as done in Forbriger (2003), or by the modified F-J (Xi et al., 2021; Zhou and  
96 Chen, 2021). In practical application, all NCFs involved in the array at different azimuth are sorted  
97 according to their interstation distance. Treating NCFs as the virtual record propagated along a

98 supposed linear array, the F-J spectrogram is then obtained by implementing the F-J transform  
99 numerically. This means an azimuthally isotropic model is assumed. The velocity given by F-J  
100 transform is the azimuth-averaged result.

101 BF is another array-based method to estimate the phase velocity under the array (Harmon et al.,  
102 2008) using the ambient seismic noise. In terms of the azimuth-averaged phase velocity, BF is  
103 almost independent of the noise source and array configuration, and has already been successfully  
104 used in ChinArray with moving subarrays (K. Wang et al., 2020). Using the data from a dense array  
105 with interstation interval of about 1 km, Qin et al. (2022) proved that the multimode dispersion  
106 curves can also be extracted by BF. For the extraction of azimuth-averaged phase velocity, the  
107 assumption on lateral isotropy is made in BF, as done in F-J method. If the azimuthal anisotropy is  
108 an issue, BF can also give the phase velocity at different azimuth by picking the maximum of the  
109 beampower at that azimuth by omitting the summation over the azimuth (Löer et al., 2018).  
110 However, the azimuthal anisotropy obtained by conventional BF would be affected by noise  
111 sources and array configuration. The dominant orientation of the array distribution would be  
112 projected into the azimuthal anisotropy of the structure since the azimuth dependence of the  
113 artificial anisotropy caused by the station-pair orientation is consistent with the azimuth  
114 dependence of the surface wave velocity of the structure (Lu et al., 2018). Although the artificial  
115 anisotropy caused by the array geometry can be removed from the BF results (Lu et al., 2018), the  
116 additional process for correction may produce the uncertainty of the results. Moreover, the  
117 dispersion image given by conventional BF lacks clarity, especially at higher frequency range.

118 In this paper, the modified BF methods are proposed to extract the multimode dispersion curves  
119 of the surface wave from seismic ambient noise. Start from the conventional BF, we first correct  
120 the wavefield by multiplying  $\sqrt{kr}$  to reduce the effect of geometric spread of the wave in high  
121 frequency range. We call this method as weighted cross-correlation beamforming (WCBF).  
122 Moreover, the modified cross-correlation beamforming (MCBF) is proposed to reduce the effect  
123 of the array geometry on the azimuthal anisotropy. An improved imaging condition of MCBF is  
124 suggested to remove the artificial aliasing. Another purpose of the paper is to present the  
125 relationship between BF and F-J. As the array-based technique, both methods deal with NCFs. In

126 BF method, the delay-and-sum of the time domain NCFs are calculated, although this is usually  
127 done in the frequency domain. In F-J method, the Fourier transform of NCFs, or the spatial  
128 wavefield represented by SPAC, is integrated over the interstation distance. Seismic interferometry  
129 theory shows that NCFs in the time domain and the SPAC expression in the frequency domain are  
130 two descriptions for the same physics (Chavez-Garcia and Luzon, 2005; Yokoi and Margaryan, 2008;  
131 Tsai and Moschetti, 2010; Lu, 2021). Therefore, we speculate there might be some connection  
132 between BF and F-J. An explicit expression of such a relationship is investigated in this paper.

133 The structure of the paper is as follows: the theory of the conventional BF is introduced and  
134 revisited in section 2. The proposed WCBF and MCBF are investigated in section 3 and section 4,  
135 respectively. Their relationships with F-J are also examined in these two sections. In section 5, the  
136 theoretical representations of MCBF are derived for a finite sampling in practical applications and  
137 an improved imaging condition is suggested to remove the artificial aliasing. The proposed  
138 methods are validated in section 6 based on the synthetic data as well as the field data from the  
139 dense array at different scales. The discussion and conclusions are given in section 7.

## 140 **2. Revisiting the Theory on Beamforming**

141 The beamforming was originally designed to estimate the predominant seismic wave, which would  
142 be body or surface wave. Initial application mainly focused on the detection of nuclear explosions  
143 (Lacoss et al., 1969; Picozzi et al., 2010), and were later applied to the waveform from earthquake  
144 and ambient seismic noise, usually used to investigate the back azimuth and slowness of the  
145 dominant signal or to obtain the velocity structure under the array. A brief review of BF is provided  
146 in this section, mainly focusing on the beamforming of seismic noise. Conventions used in this  
147 paper for Fourier transform, Hilbert transform and cross-correlation are given in Appendix A.

### 148 **2.1. Conventional (Cross-correlation) Beamforming**

149 The basic idea of beamforming is delay-and-sum. For the seismic surface wave or ambient noise,  
150 it is usually assumed a horizontal plane wave travelling over the array. The signal received at each  
151 station in the array is thought as the summation of the time-shifted plane wave. The beam trace  
152 can be expressed in the time domain as

153 
$$b(\mathbf{s}, t, \theta) = \frac{1}{N} \sum_{i=1}^N W_i d(\mathbf{x}_i, t + \tau_i) \quad (1)$$

154 where  $d(\mathbf{x}_i, t)$  is the time series recorded at station  $\mathbf{x}_i$ .  $W_i$  is the weight.  $N$  is the number of  
 155 stations involved in the array.  $\mathbf{s}$  is the slowness vector of the incident plane wave with azimuth  $\theta$ .  
 156  $\tau_i = (\mathbf{x}_i - \mathbf{x}_0) \cdot \mathbf{s}$  is the time delay between station  $\mathbf{x}_i$  and the reference position  $\mathbf{x}_0$ . Equation 1 can  
 157 be written in frequency domain as,

158 
$$b(\mathbf{k}, \omega, \theta) = \frac{1}{N} \sum_{i=1}^N W_i d(\mathbf{x}_i, \omega) e^{-ik \cdot (\mathbf{x}_i - \mathbf{x}_0)} \quad (2)$$

159 where  $d(\mathbf{x}_i, \omega)$  is the Fourier transform of  $d(\mathbf{x}_i, t)$ .

160 The beampower  $B(\mathbf{k}, \omega, \theta)$  is usually taken as the beamforming output, which can be written as

161 
$$B(\mathbf{k}, \omega, \theta) = |b(\mathbf{k}, \omega, \theta)|^2 = \frac{1}{N^2} \left| \sum_{i=1}^N W_i D(\mathbf{x}_i, \omega) e^{-ik \cdot (\mathbf{x}_i - \mathbf{x}_0)} \right|^2 \quad (3)$$

162 with

163 
$$D(\mathbf{x}_i, \omega) = \frac{1}{L} d(\mathbf{x}_i, \omega) \quad (4)$$

164 where  $L$  is the signal length. Equation 3 can be recast into

165 
$$\begin{aligned} B(\mathbf{k}, \omega, \theta) &= \frac{1}{N^2} \left( \sum_{i=1}^N W_i D(\mathbf{x}_i, \omega) e^{-ik \cdot (\mathbf{x}_i - \mathbf{x}_0)} \right)^* \left( \sum_{j=1}^N W_j D(\mathbf{x}_j, \omega) e^{-ik \cdot (\mathbf{x}_j - \mathbf{x}_0)} \right) \\ &= \frac{1}{N^2} \left( \sum_{i=1}^N \sum_{j=1}^N W_{ij} e^{ik \cdot \mathbf{x}_i} C_{ij}(r_{ij}, \omega) e^{-ik \cdot \mathbf{x}_j} \right) \\ &= \frac{1}{N^2} \left( \sum_{i=1}^N \sum_{j=1}^N W_{ij} C_{ij}(r_{ij}, \omega) e^{-ik \cdot (\mathbf{x}_j - \mathbf{x}_i)} \right) \\ &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N W_{ij} C_{ij}(r_{ij}, \omega) e^{-ikr_{ij} \cos(\theta_j - \theta)} \end{aligned} \quad (5)$$

166 where

167 
$$C_{ij}(r_{ij}, \omega) = D^*(\mathbf{x}_i, \omega) D(\mathbf{x}_j, \omega) = \frac{1}{L^2} d^*(\mathbf{x}_i, \omega) d(\mathbf{x}_j, \omega) \quad (6)$$

168 is the cross spectral density matrix (CSDM) (e.g. [Gerstoft and Tanimoto, 2007](#); [Riahi et al., 2014](#)) or  
 169 cross-covariance matrix (e.g. [Capon, 1969](#); [Seydoux et al., 2017](#)), in which the phase shift between

170 stations  $\mathbf{x}_i$  and  $\mathbf{x}_j$  is contained.  $r_{ij}$  and  $\theta_{ij}$  denote the interstation distance and azimuth of the  
171 station pair, respectively. It can be found that Equation 3, which is derived based on the standard  
172 delay-and-sum beamforming, is identical to Equation 5, which can be obtained directly from the  
173 cross-spectral density matrix (Löer et al., 2018).

174 Löer et al. (2018) refers to Equations 3 or 5 as standard beamforming. Gal et al. (2019) call them  
175 conventional beamforming (or Bartlett beamforming). Ruigrok et al. (2017) call the beamforming  
176 shown in Equation 2 as conventional beamforming and the beamforming shown in Equation 5 as  
177 correlated beamforming (CBF). Meanwhile, they call the beamforming only considering the CSDM  
178 with  $i \neq j$  as the cross-correlation beamforming (CCBF) since the autocorrelation is discarded. The  
179 definitions of CCBF and CBF are not distinguished in this paper. We refer to the beamforming shown  
180 in Equation 5 as the correlation beamforming (CBF) regardless of the autocorrelation is contained.  
181 Which correlation components are considered in calculation is determined by choosing the  
182 elements presented in the matrix CSDM, such as an upper triangular matrix, a lower triangular  
183 matrix or a matrix without diagonal elements.

## 184 **2.2. Multimode Dispersion Image Obtained by CBF**

185 To estimate the vector wavenumber  $\mathbf{k}$  across the array, a search over all possible wavenumber and  
186 azimuth are performed in conventional BF. The beampower is calculated by fitting the CSDM  
187 between station  $\mathbf{x}_i$  and  $\mathbf{x}_j$  with a synthetic plane wave described by the vector wavenumber  $\mathbf{k}$ .  
188 Once the phase delay denoted by CSDM match the one required by the incident plane wave model,  
189 the constructive interference would occur and hence the summation shown in Equation 5 would  
190 reach a maximum. The slowness associated with the maximum gives the velocity of the structure  
191 under the array, and the corresponding azimuth associated with the maximum gives the direction  
192 of arrival (DOA) of the wave.

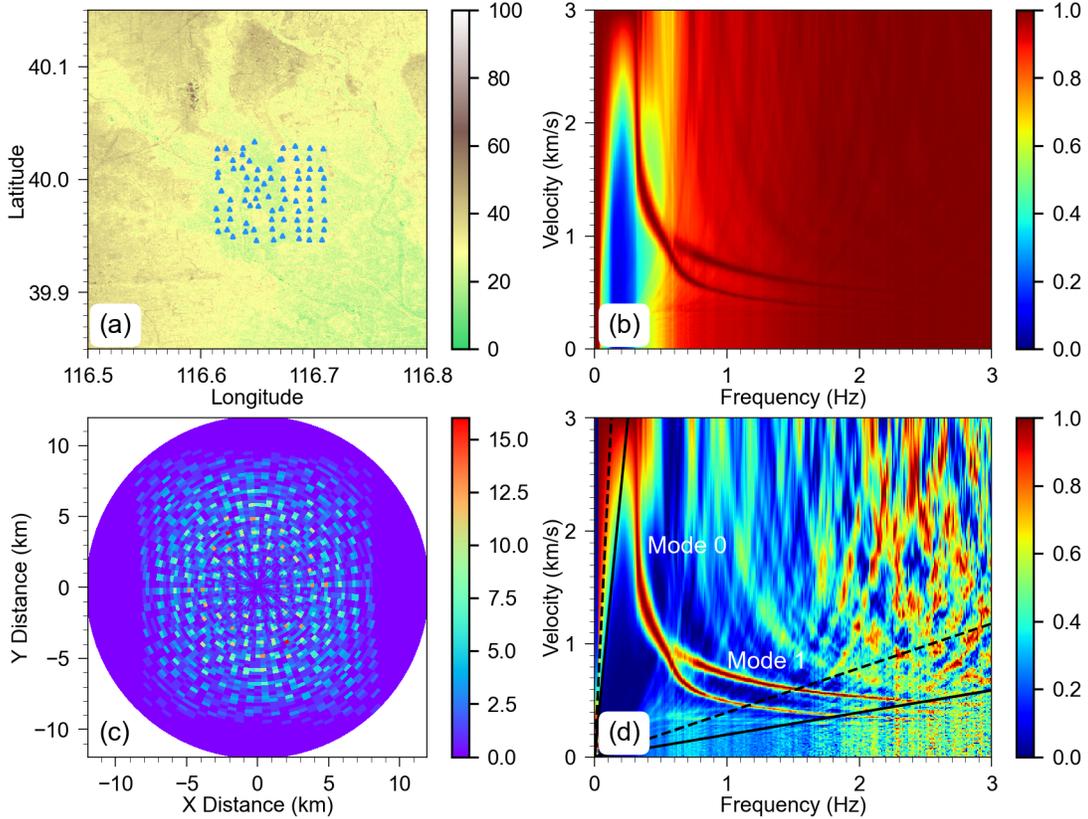
193 For surface waves, if more than one mode is incident as a plane wave at velocities with much  
194 difference, the multimode phase velocity can in principle be estimated by beamforming. This is  
195 more common at local scales, especially in sedimentary basins where the energy of higher modes  
196 is usually pronounced.

197 To generate the dispersion image in f-v domain, the beampower shown in Equation 5 for each  
198 frequency is calculated by

$$199 \quad CBF(\mathbf{k}, \omega, \theta) = \frac{1}{N^2} \left| \sum_{i=1}^N \sum_{j=1}^N C_{ij}(\mathbf{r}_{ij}, \omega) e^{-i\mathbf{k} \cdot \mathbf{r}_{ij}} \right|. \quad (7)$$

200 Equation 7 and Equation 5 are theoretically equivalent except that the weight  $W_{ij}$  is ignored in  
201 Equation 7. In practical application, NCFs are usually asymmetric due to the complex source  
202 distribution.  $C_{ij}(\mathbf{r}_{ij}, \omega)$ , the Fourier transform of NCFs, is not real. The modulus is therefore to be  
203 taken in Equation 7 as the beamforming output. The beampower is then averaged over  $\theta$ . The  
204 dispersion image can thereby be obtained by combining the azimuth-averaged beampower of each  
205 frequency. We call Equation 7 as the conventional CBF for extraction of dispersion curves. [K. Wang  
206 et al. \(2020\)](#) have used CBF to estimate the azimuth-averaged phase velocity of the fundamental  
207 mode Rayleigh wave.

208 As an example, Figure 1 gives an illustration for extraction of multimode Rayleigh waves using CBF.  
209 Figure 1a shows the configuration of the array consisting of seventy-nine stations, which are  
210 sampled from the dense array located in Tongzhou with intervals of 1-2 km ([Qin et al., 2022](#)). Figure  
211 1c shows the distribution of interstation distances as a function of azimuth. Figure 1b and 1d show  
212 the dispersion images. In Figure 1b, the beampower of each frequency is normalized by the  
213 maximum at that frequency. To highlight the energy along the dispersion branches at high  
214 frequencies, we normalize the difference between the beampower and the minimum at that  
215 frequency in Figure 1d. The fundamental mode (mode 0) and the first higher mode (mode 1) are  
216 observed in Figure 1d. Based on this CBF scheme, [Qin et al. \(2022\)](#) has successfully obtained the  
217 lateral variation of the phase velocity of these two modes. Note that Figure 1d is slightly different  
218 from Figure 5a in [Qin et al. \(2022\)](#) since different matrix elements of CSDM is used here to reduce  
219 the artifacts.



220  
221

222 **Figure 1.** An illustration for the measurement of multimode surface wave by CBF. (a) The configuration of the  
 223 array consisting of seventy-nine stations. (b) The azimuth-averaged dispersion image obtained by conventional  
 224 CBF. For each frequency, the beampower is normalized by the maximum at that frequency. (c) The distribution  
 225 of the number of station pairs as a function of the azimuth and interstation distance. (d) The same as (b) but with  
 226 different normalized strategy. The dashed and solid black lines in (d) give the resolved wavenumber ranges  
 227 estimated by  $[k_{\min}, k_{\max}]$  and  $[2k_{\min}, 2k_{\max}]$  using Equation 15, respectively.

### 228 2.3. Resolution and Aliasing of CBF

229 The resolution of BF depends on the array configuration and the characteristics of the wavefield  
 230 across the array. The wavefield, for example the energy of the surface wave modes carried by the  
 231 ambient noise, depends on the structure under the array and source characteristics which are  
 232 often what we are trying to figure out. Therefore, the resolution we are concerned with and can  
 233 be improved usually refers to the resolution determined by the array configuration.

234 The quantities appearing in Equation 7 are only related to the array configuration, except for the

235 CSDM which describes the wavefield. After removing CSDM, Equation 7 can be expressed as  
 236 ([Wathelet et al, 2008](#); [Ruigork et al, 2017](#))

$$237 \quad ARF(\mathbf{k}) = \frac{1}{N^2} \left| \sum_{i=1}^N \sum_{j=1}^N e^{-i\mathbf{k} \cdot \mathbf{r}_{ij}} \right| \quad (8)$$

238 Since

$$239 \quad e^{-i\mathbf{k} \cdot (\mathbf{x}_j - \mathbf{x}_i)} = \left[ e^{-i\mathbf{k} \cdot (\mathbf{x}_i - \mathbf{x}_j)} \right]^* \quad (9)$$

240 Equation 8 can also be written as ([Horike 1985](#), [Picozzi et al,2010](#))

$$241 \quad ARF(\mathbf{k}) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N e^{-i\mathbf{k} \cdot (\mathbf{x}_j - \mathbf{x}_i)} \quad (10)$$

242  $ARF(\mathbf{k})$  is termed spatial window function ([Lacoss et al, 1969](#); [Horike 1985](#)) or array response  
 243 function (ARF) ([Capon,1969](#); [Rost and Thomas,2002](#); [Ruigork et al, 2017](#)).

244 For an incident monochromatic plane wave with given angular frequency  $\omega_0$  and wavenumber  $\mathbf{k}_0$ ,  
 245 by ignoring the attenuation, the CSDM between station  $\mathbf{x}_i$  and  $\mathbf{x}_j$  can be expressed as

$$246 \quad C_{ij}(\omega) = \delta(\omega - \omega_0) e^{i\mathbf{k}_0 \cdot (\mathbf{x}_j - \mathbf{x}_i)} \quad (11)$$

247 where  $\delta(\omega - \omega_0)$  is the Dirac delta function. Substitute Equation 11 into Equation 5, the  
 248 beamforming output for a monochromatic plane wave can be written as ([Asten and Henstridge,](#)  
 249 [1984](#))

$$250 \quad B(\omega, \mathbf{k}) = \left| W_f(\omega - \omega_0) \right|^2 ARF(\mathbf{k} - \mathbf{k}_0) \quad (12)$$

251 where  $W_f(\omega - \omega_0)$  is the Fourier spectra of the time series with given limited length. For more  
 252 general case, the estimated wavenumber spectrum is the 2D convolution of the true spectrum  
 253 with ARF ([Lacoss et al., 1969](#); [Asten and Henstridge, 1984](#)). The beamforming resolution is thereby  
 254 controlled by ARF which depends on the array configuration.

255 At present, there is no global agreement about the capabilities of an array ([Wathelet et al, 2008](#)).  
 256 For the array with simple and regular geometries, for instance, a linear array with equal spacing,  
 257 the aliasing and resolution limit can be estimated by the maximum and minimum interstation  
 258 spacing in the array using the Nyquist sampling theorem. For the array with irregular geometry,

259 some empirical rules are proposed to estimate the reasonable results achieved by the array.  
 260 Tokimatsu (1997) use the minimum  $r_{\min}$  and maximum  $r_{\max}$  of the interstation distances inside the  
 261 array to determine the range of the resolved wavelength. As a rule of thumb, the resolved  
 262 minimum wavelength  $\lambda_{\min}$  and maximum one  $\lambda_{\max}$  are respectively  $\lambda_{\min} = 2r_{\min}$  and  $\lambda_{\max} = 3r_{\max}$ .

263 A more rigorous definition for resolved wavenumber is based on ARF. For two plane waves  
 264 travelling the array with wavenumbers  $\mathbf{k}_1$  and  $\mathbf{k}_2$ , the CSDM can be expressed as

$$265 \quad C_{ij}(r_{ij}, \omega) = A_1(\omega)e^{i\mathbf{k}_1 \cdot \mathbf{r}_{ij}} + A_2(\omega)e^{i\mathbf{k}_2 \cdot \mathbf{r}_{ij}} \quad (13)$$

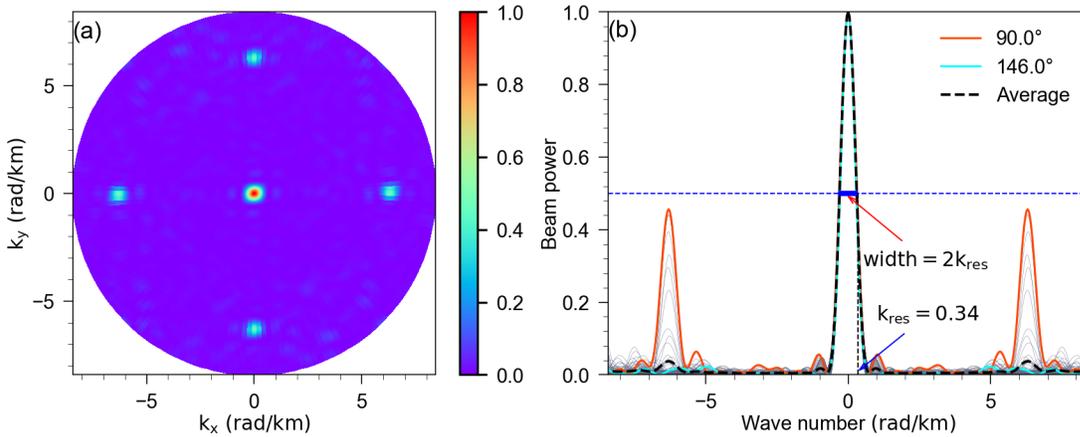
266 The f-k spectrum estimated from Equation 7 can be written as

$$267 \quad \begin{aligned} CBF(\mathbf{k}, \omega, \theta) &= \frac{1}{N^2} \left| \sum_{i=1}^N \sum_{j=1}^N \left( A_1(\omega)e^{i\mathbf{k}_1 \cdot \mathbf{r}_{ij}} + A_2(\omega)e^{i\mathbf{k}_2 \cdot \mathbf{r}_{ij}} \right) e^{-i\mathbf{k} \cdot \mathbf{r}_{ij}} \right| \\ &\leq \frac{1}{N^2} \left( \left| \sum_{i=1}^N \sum_{j=1}^N A_1(\omega)e^{i\mathbf{k}_1 \cdot \mathbf{r}_{ij}} e^{-i\mathbf{k} \cdot \mathbf{r}_{ij}} \right| + \left| \sum_{i=1}^N \sum_{j=1}^N A_2(\omega)e^{i\mathbf{k}_2 \cdot \mathbf{r}_{ij}} e^{-i\mathbf{k} \cdot \mathbf{r}_{ij}} \right| \right) \\ &= |W_1(\omega - \omega_1)|^2 ARF(\mathbf{k} - \mathbf{k}_1) + |W_2(\omega - \omega_2)|^2 ARF(\mathbf{k} - \mathbf{k}_2) \end{aligned} \quad (14)$$

268 This means the beampower for two plane waves is always lower than the summation of individual  
 269 plane waves (Wathelet et al., 2008). The aliasing and resolution of the array can be defined based  
 270 on the ARF by considering the summation of two shifted ARFs. If two wavenumbers are close to  
 271 each other, the summation of two shifted ARF would generate a wider main lobe rather than two  
 272 narrow main lobes. The width of the main lobe of ARF which makes two wavenumbers cannot be  
 273 distinguished is used to define the resolved wavenumber. The side lobes of two shifted ARFs may  
 274 overlap and resulted the beampower with the same magnitude as the main lobe. The wavenumber  
 275 associated with such overlapping sidelobes is defined as the aliasing wavenumber.

276 Figure 2a shows the ARF for the array shown in Figure 1a. The maximum occurred in the center at  
 277  $\mathbf{k} = \mathbf{0}$ . The aliasing is likely to occur at the wavenumbers where, for example, the four secondary  
 278 peaks are observed, which appear in four typical directions around 0, 90, 180, 270 degrees. Figure  
 279 2b shows the cross sections of ARF every two degrees. Concerning on the resolution, the thinner  
 280 is the central peak of ARF, the more capable is the array to distinguish two waves travelling at close  
 281 wavenumbers (Wathelet et al., 2008). How to evaluate quantitatively the beamforming resolution

282 and the confidence interval that aliasing does not appear, different definitions are proposed by  
 283 researchers. The wavenumber  $k_{\text{Res}}$  associated with the half-width of the main lobe, at the edge of  
 284 which the beampower is reduced to a given threshold relative to the maximum, is usually used to  
 285 define the resolving power of the wavenumber. As shown in Figure 2b, two wavenumbers cannot  
 286 be resolved if their difference is less than  $2k_{\text{Res}}$  since two main lobes would overlap. If the  
 287 maximum of the main lobe is 1, the threshold is generally selected as 0.5, i.e., the beampower  
 288 reduces to the half of the maximum or -3dB in the logarithmic coordinates (Woods and Lintz,1973;  
 289 Asten and Henstridge, 1984; Wathelet et al., 2008). For this definition,  $k_{\text{Res}}$  is equal to 0.34 Rad/km  
 290 for the array shown in Figure 1a.



291  
 292 **Figure 2.** The array response function (ARF) for the array shown in Figure 1a. The gray lines in (b) denotes  
 293 the cross sections of the ARF every two degrees. The black dashed line denotes the azimuth-averaged result.  
 294 The blue dashed line denotes the location of 0.5 beampower.

295 On the other hand, based on the sampling theorem, the maximum ( $k_{\text{max}}$ ) and minimum ( $k_{\text{min}}$ )  
 296 wavenumbers that can be resolved by the array are

$$\begin{aligned}
 k_{\text{max}} &= \pi/r_{\text{min}} \\
 k_{\text{min}} &= \pi/r_{\text{max}}
 \end{aligned}
 \tag{15}$$

298 where  $r_{\text{max}}$  and  $r_{\text{min}}$  are the largest and smallest interstation distances, respectively.  $k_{\text{max}}$  is also  
 299 called Nyquist wavenumber, the wavenumber exceeds which repetition would occur (e.g., see  
 300 Figure 10). Ruigrok et al. (2017) also approximate the resolution using  $k_{\text{min}}$ . For the array shown in  
 301 Figure 1a,  $k_{\text{max}} = 15.98$  Rad/km,  $k_{\text{min}} = 0.26$  Rad/km. By removing the aliasing in negative

302 frequency, the resolved wavenumber range would reach to  $[2k_{\min}, 2k_{\max}]$ . The dashed and solid  
 303 black lines in Figure 1d gives the resolved wavenumber ranges estimated by  $[k_{\min}, k_{\max}]$  and  
 304  $[2k_{\min}, 2k_{\max}]$ , respectively.

### 305 **3. Weighted Cross-correlation Beamforming (WCBF)**

306 It can be found from Figure 1b that for the conventional CBF, the energy difference around the  
 307 eigenvalues from the ambient noise is not manifest at high frequencies. As a result, at the first  
 308 glance in Figure 1b, the color scheme is different above and below 0.5 Hz, resulting the energy  
 309 along the dispersion branches are not prominent for the frequencies above 0.5 Hz. This is mainly  
 310 due to the attenuation and geometric spread of the wavefield at high frequencies. Of course, we  
 311 can highlight the energy along the dispersive branches by changing the normalization strategy, as  
 312 shown in Figure 1d. However, to correct the wavefield by multiplying  $\sqrt{kr}$  in CBF is a natural  
 313 alternative.

#### 314 **3.1. Correction of Wavenumber $k$ and Propagation Distance $r$**

315 Although representations are different literally, the conventional BF shown in Equation 3 and  
 316 Equation 5 are actually equivalent. However, seismic interferometry endows Equation 5 with more  
 317 physical explanations, which is not so obvious in Equation 3. The theory of seismic interferometry  
 318 states that the Green's function can be retrieved by cross-correlating the seismic ambient noise.  
 319 In the frequency domain, that is, the CSDM  $C(\mathbf{x}_i, \mathbf{x}_j, \omega)$  is equivalent to the Green's function and  
 320 can be represented as

$$321 \quad C(\mathbf{x}_i, \mathbf{x}_j, \omega) = C_{ij}(\omega) \propto S(\omega) \text{Im}[G(\mathbf{x}_i, \mathbf{x}_j, \omega)] \quad (16)$$

322 where  $G(\mathbf{x}_i, \mathbf{x}_j, \omega)$  is the Green's function between  $\mathbf{x}_i$  and  $\mathbf{x}_j$ . Due to the source distribution, the  
 323 surface wave usually dominates the seismic noise. The Green's function in Equation 16 is therefore  
 324 considered to be the surface wave Green's function. For the vertical component of a homogeneous  
 325 layered medium, Equation 16 can be written as

$$\begin{aligned}
C_{ij}(\omega) &\propto S(\omega) \operatorname{Im} \left[ G_{zz}(\mathbf{x}_i, \mathbf{x}_j, \omega) \right] \\
&= S(\omega) \operatorname{Re} \left[ \sum_n \frac{1}{8cUI_1} r_2(n, \omega) r_1(n, \omega) H_0^{(2)}(k_n r_{ij}) \right]
\end{aligned} \tag{17}$$

where  $c$  is the phase velocity,  $U$  the group velocity.  $r_1(z)$  and  $r_2(z)$  are the radial and vertical eigen-function of the Rayleigh wave, respectively.  $I_1 = \frac{1}{2} \int_0^\infty \rho(r_1^2 + r_2^2) dz$ .  $n$  denotes the mode number.  $k_n$  is the horizontal wavenumber in the direction of  $r_{ij}$ . By expressing the Hankel function as the sum of the first and second kind of Bessel functions, Equation 17 can be recast as

$$C_{ij}(\omega) \propto S(\omega) \sum_n \frac{1}{8cUI_1} r_2(n, \omega) r_1(n, \omega) J_0(k_n r_{ij}) \tag{18}$$

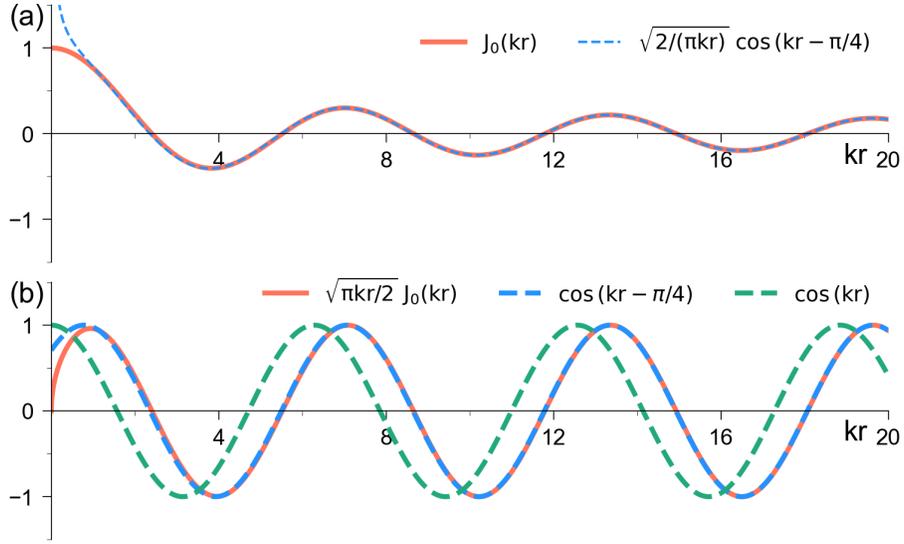
This means, under the assumption of laterally homogeneous model and uniform distribution of noise source, the monochromatic wavefield across the array can be represented as  $J_0(k_n r)$ . This is consistent with the result on SPAC of the microtremor in the frequency domain, originally given by [Aki \(1957\)](#). Therefore, as a function of the interstation distance  $r$ , the Fourier transform of NCFs can be thought as the spatial sampling on a cylindrical wavefield described by the Bessel function  $J_0(k_n r)$ .

In the far field,  $J_0(k_n r)$  can be expressed as

$$J_0(k_n r) \sim \sqrt{\frac{2}{\pi k_n r}} \cos\left(k_n r - \frac{\pi}{4}\right), k_n r \gg 0 \tag{19}$$

Figure 3a shows the variation of  $J_0(k_n r)$  and its far field approximation as a function of the distance. If the weight  $W$  in Equation 5 is the same for all NCFs, the contribution of the NCFs with larger interstation distance or the wave with high frequency would be much smaller due to the spread factor  $\sqrt{1/kr}$ . This makes the mode recognition along dispersion branches lack clarity in high frequency range in Figure 1b. To improve the CBF results for larger  $k_n r$ , we correct the wavefield by considering the following weighted cross-correlation beamforming.

$$WCBF(\mathbf{k}, \omega, \theta) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \sqrt{|\mathbf{k}| |\mathbf{r}_{ij}|} C_{ij}(\mathbf{r}_{ij}, \omega) e^{i\mathbf{k} \cdot \mathbf{r}_{ij}} \tag{20}$$



347

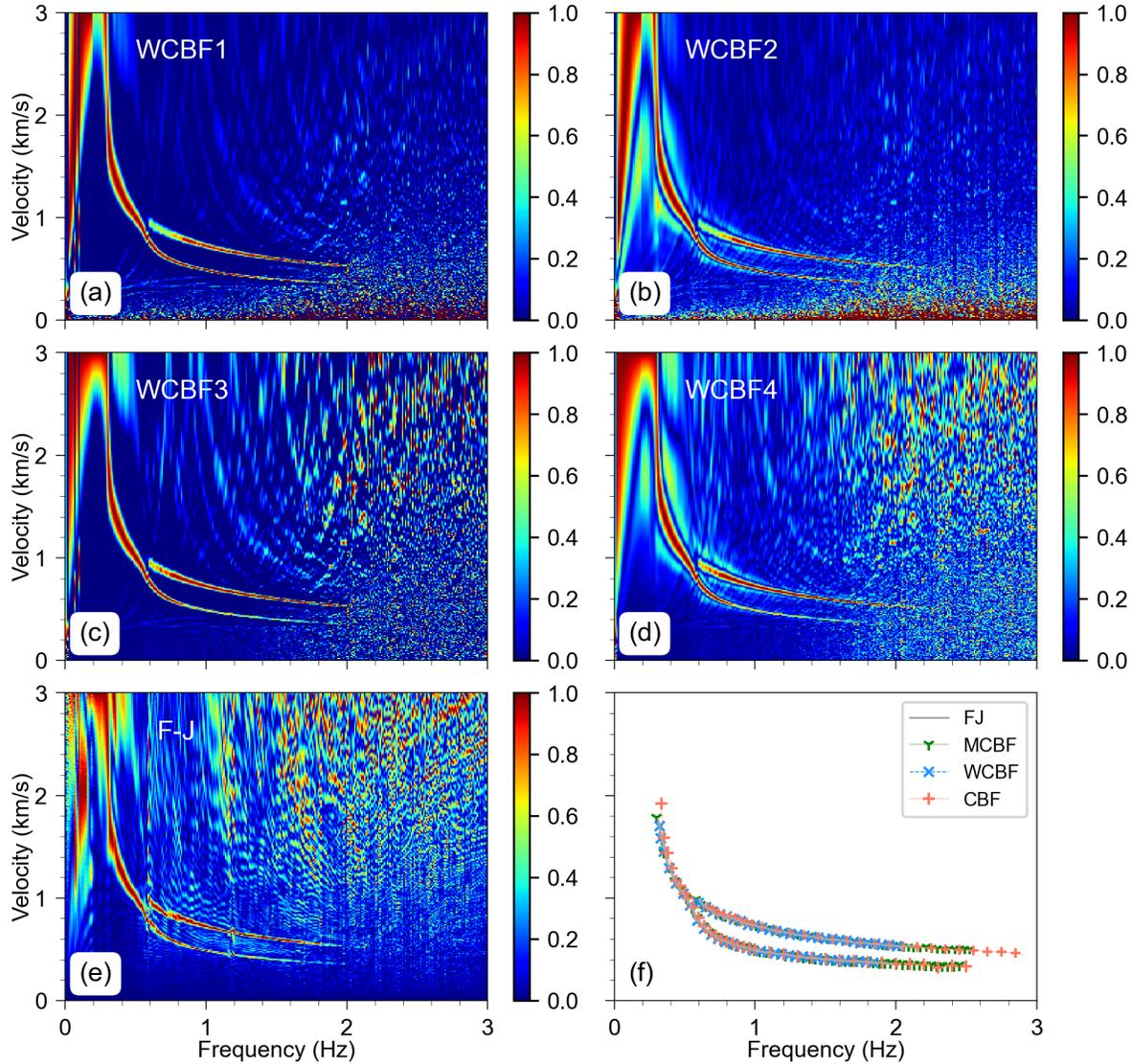
348 **Figure 3.** (a) The Bessel function  $J_0(k_n r)$  and its far field approximation.  
 349  $J_0(k_n r)$  and cosine functions.

350 The variation of the corrected wavefield with the distance is shown in Figure 3b. The azimuth-  
 351 averaged velocity is considered currently. The average over the azimuth implies to conduct the  
 352 summation of Equation 20 over the angle  $\theta$ , the azimuth of the wavenumber  $\mathbf{k}$ . This azimuth-  
 353 averaged operation would again introduce a decay with a factor of  $\sqrt{1/kr}$ . Therefore,  $\sqrt{kr}$  is again  
 354 multiplied for the azimuth-averaged of Equation 20. It reads

355 
$$\overline{WCBF}(k, \omega) = \frac{1}{N^2} \sum_{\theta=0}^{2\pi} \sum_{i=1}^N \sum_{j=1}^N kr_{ij} C_{ij}(\mathbf{r}_{ij}, \omega) e^{ikr_{ij}\theta} \quad (21)$$

356 The overbar is used to denote the average over the azimuth. We call Equations 20 and 21 as the  
 357 weighted cross-correlation beamforming (WCBF) since varying weight is adopted. Different from  
 358 Equation 7, we take the real part of Equation 21 or its' absolute value as the imaging conditions for  
 359 plotting, which are expressed as

360 
$$\begin{aligned} WCBF1(k, \omega) &= \text{Re}(\overline{WCBF}(k, \omega)) \\ WCBF2(k, \omega) &= \text{ABS}[\text{Re}(\overline{WCBF}(k, \omega))] \end{aligned} \quad (22)$$



361  
 362 **Figure 4.** The dispersion images obtained by WCBF (a-d) and F-J (e). The comparison of the picked values from  
 363 different methods are shown in (f). The gray lines in (f) show the F-J results picked from (e). The result of WCBF  
 364 in (f) is picked from (a), which is the same as that from (b-d). The result of CBF in (f) is picked from Figure 1d.

365 Figures 4a and 4b show dispersion images in f-v domain calculated from Equation 22. The  
 366 normalization strategy is the same as that used in Figure 1b. Figures 4a show the energy along two  
 367 dispersion branches is more prominent than that in Figure 1b. The efficiency of the normalization  
 368 strategy used in Figure 1d is achieved and exceeded by WCBF. Moreover, the dispersion images  
 369 are much clearer in Figure 4a and 4b, and the artificial images occurred in the upper region with  
 370 high velocities and frequencies in Figure 1d disappear. As a trade-off, at the lower of the panel,

371 some disturbing pixels appear around the area close to zero velocity. This would be caused by the  
 372 overcorrection on the large wavenumber around zero velocity. These interfering pixels have little  
 373 effect on mode recognition.

### 374 3.2. Relation Between WCBF and F-J

375 F-J is also an array-based method to extract the dispersion curves using ambient noise. In this  
 376 section, the relation between WCBF and F-J is investigated. We start from the 2D Fourier transform  
 377 of an arbitrary function. In polar coordinates, it can be expressed as (Baddour, 2011)

$$378 \quad 2\text{DFFT}(k, \omega) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{+\infty} C_{ij}(r, \omega, \theta) e^{ikr \cos \theta} r dr d\theta \quad (23)$$

379 where  $C_{ij}(r, \omega, \theta)$  can be an arbitrary function which depends on the azimuth  $\theta$  and the radial  
 380 distance  $r$ . It is assumed  $C_{ij}(r, \omega, \theta)$  is independent of the azimuth. For instance, assuming it be  
 381 the spectrum of the surface wave vertical component recorded at the surface of the laterally  
 382 isotropic layered model,  $C_{ij}(r, \omega, \theta)$  can be written as  $C_{ij}(r, \omega)$ . Since the Bessel function  $J_0(kr)$   
 383 can be expressed as

$$384 \quad J_0(kr) = \frac{1}{2\pi} \int_0^{2\pi} e^{ikr \cos \theta} d\theta \quad (24)$$

385 Equation 23 can then be rewritten as

$$386 \quad 2\text{DFFT}(k, \omega) = FJ(k, \omega) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{+\infty} C_{ij}(r, \omega) e^{ikr \cos \theta} r dr d\theta = \int_0^{+\infty} C_{ij}(r, \omega) J_0(kr) r dr \quad (25)$$

387 Equation 25 is the F-J transform proposed by J. Wang et al (2019), which can be thought as the 2D  
 388 Fourier transform of the radially symmetric function  $C_{ij}(r, \omega)$ . This implies the lateral isotropy is  
 389 assumed for F-J method. Under this assumption, the 2D Fourier transform of the radially symmetric  
 390 function is degenerated into the Fourier Bessel transform or Hankel transform.

391 The surface wave usually dominates the NCFs. As shown in Equation 18, for the vertical component,  
 392  $C_{ij}(r, \omega)$  is related to the spatial autocorrelation coefficient  $J_0(kr)$ , i.e.,

$$393 \quad C_{ij}(r, \omega) \sim S(\omega) J_0(k_n r) \quad (26)$$

394 Substituting Equation 26 into Equation 25 and considering the orthogonality of the Bessel function  
 395 (Morse and Feshbach, 1953, P943), we have

396 
$$FJ(k, \omega) = S(\omega) \int_0^{+\infty} J_0(k_n r) J_0(kr) r dr = S(\omega) \frac{\delta(k_n - k)}{k_n} \quad (27)$$

397 Consequently, the maxima in  $\omega - k$  domain are associated with the eigenvalues  $k_n$  of the Rayleigh  
 398 waves. The dispersion curves can then be measured by picking the velocities associated with the  
 399 maxima in f-v domain.

400 Comparing Equations 25 and 21, it can be found that azimuth-averaged  $\overline{WCBF}$  is the discrete form  
 401 of F-J except for an extra wavenumber factor occurred in Equation 21. We define the new imaging  
 402 conditions WCBF3 and WCBF4 by dividing Equation 22 by the wavenumber  $k$ . We have

403 
$$WCBF3(k, \omega) = \text{Re} \left( \frac{1}{k} \overline{WCBF}(k, \omega) \right) = FJ(k, \omega) \quad (28)$$

404 
$$WCBF4(k, \omega) = \text{ABS} \left[ \text{Re} \left( \frac{1}{k} \overline{WCBF}(k, \omega) \right) \right]$$

404 The WCBF3 shown in Equation 28 is now equivalent to the F-J method. Compared with WCBF1 in  
 405 Equation 22, only the propagation distance  $r$  is corrected in F-J method. Choosing  $\sqrt{k}$  or  $\sqrt{kr}$  to  
 406 correct the wavefield is a trade-off that determines whether artifacts appear in the upper region  
 407 with high velocities and frequencies, or in the lower region close to zero velocity. Figure 4c and 4d  
 408 show the results calculated from Equation 28. As opposite to Figure 4a and 4b, the artifacts in  
 409 Figure 4c and 4d appear in the upper region with higher velocities and frequencies. The disturbing  
 410 pixels around the area close to zero velocity are significantly reduced. In addition, for the results  
 411 of WCBF2 and WCBF4, where the absolute value of the real part is taken as the imaging condition,  
 412 the apparent side lobes around the dispersion branches are observed in Figure 4b and 4d. Because  
 413 the sampling of the array on the wavefield is always finite, Dirac delta function shown in Equation  
 414 27 would behave as a sinc function. The side lobes in Figure 4b and 4d originate from the negative  
 415 values adjacent to the main lobe of the sinc function. They disappear in Figure 4a and 4c since only  
 416 positive values are color-coded.

417 Figure 4e shows the result of F-J method. As expected, the dispersion image obtained by F-J  
 418 method has the same characteristics as that of WCBF3 (Figure 4c). The slight difference possibly  
 419 comes from the approximation of the numerical integration. An integral scheme based on the  
 420 trapezoidal integration is used in F-J (J. Wang et al, 2019), while the discrete summation shown in

421 Equation 21, similar as the discrete Fourier summation, is used to directly compute the result of  
422 WCBF.

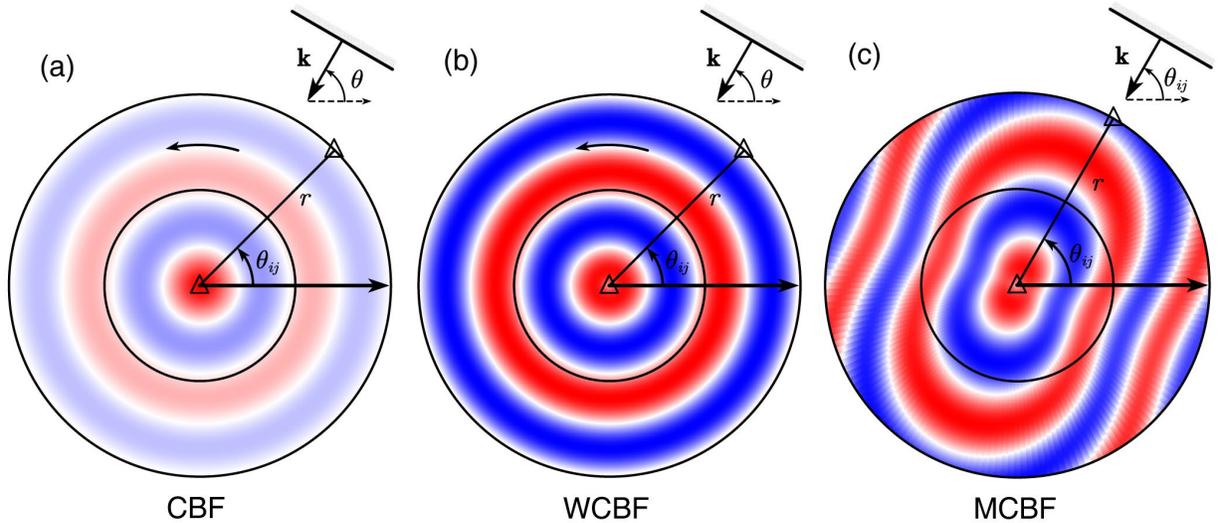
#### 423 **4. Modified Cross-correlation Beamforming (MCBF)**

424 Different from the estimation on the azimuth-averaged velocity, the phase velocity as well as the  
425 azimuth associated with the maximum beampower are measured by CBF in Roux and Ben-Zion  
426 (2017). Admittedly, the azimuth-dependent velocity can be estimated by picking the values  
427 associated with the maximum beampower at that azimuth. The azimuthal anisotropy can then be  
428 estimated by fitting the model proposed by Smith and Dahlen (1973), as done in L er et al. (2018).  
429 However, the artificial anisotropy introduced by the array geometry will be projected into the  
430 estimated results since the maximum beampower also depends on the orientation of station pairs  
431 determined by the array geometry. In this section, a modified CBF scheme is proposed to overcome  
432 the effect of the array geometry on the estimation of the azimuth-dependence phase velocity.

##### 433 **4.1. Algorithm for MCBF**

434 As mentioned in section 3.1, due to the equivalence between SPAC in the frequency domain and  
435 NCFs in the time domain, NCFs can be thought as the sampling on the cylindrical wavefield  
436 described by  $J_0(kr)$ . We can do the following thought experiment. Assuming the array is dense  
437 enough and infinite, the Fourier transformed NCFs of each azimuth are arranged from  $r = 0$  to  
438  $r = \infty$  according to their interstation intervals. The spatial wavefield would be a cylindrical wave  
439 with  $r = 0$  as the center, as shown in Figure 5a. Note that the central point  $r = 0$  is a reference  
440 point and does not actually correspond to any physical position of the array.

441 The beamforming shown in Equation 7 is designed to track the phase by fitting the cylindrical wave  
442 shown in Figure 5a with the plane wave incident from different azimuth. As a function of  $\theta$ , the  
443 beampower is thus obtained by summing over  $\theta_{ij}$  which is implemented by projecting the plane  
444 wave into all the station-pairs. The averaged beampower over the azimuth  $\theta$  is then used to  
445 estimate the azimuth-averaged velocity. The algorithm of WCBF shown in Equation 20 is the same  
446 as this process for Equation 7, but the geometric spread is corrected. The corrected cylindrical  
447 wavefield is shown in Figure 5b.



448  
 449 **Figure 5.** The schematic diagram of cylindrical wave field processed by different strategies of BF technique. (a)  
 450 The conventional CBF handle the cylindrical wavefield represented by  $J_0(kr)$ . The plane wave with a given  
 451 azimuth  $\theta$  is projected into all the interstation pairs and the beamforming is output as a function of  $\theta$ . (b) WCBF  
 452 handle the cylindrical wavefield with the correction factor  $\sqrt{kr}$ . The beamforming is also output as a function  
 453 of  $\theta$ . (c) Only the plane waves with the same azimuth as the orientation  $\theta_{ij}$  of the interstation pairs is  
 454 considered in MCBF. The beamforming is output as a function of  $\theta_{ij}$ . The wavefield can therefore be azimuthally  
 455 anisotropic for MCBF.

456 The above algorithm for CBF and WCBF suffers from two deficiencies. First, more computations  
 457 are spent on projecting a plane wave at a given azimuth into the orientation of all station-pairs,  
 458 but the improvement of the dispersion image is trivial. Second and more importantly, the  
 459 implement of projection implies the structure beneath the array is azimuthally isotropic. As shown  
 460 in Figure 5c, for the wavefield with an azimuth-dependence velocity, the azimuthal anisotropy will  
 461 be an issue. The isotropic assumptions required by the projection algorithm in Equations 7 and 20  
 462 would fail. Although the azimuth-dependence velocity can be estimated by omitting the average  
 463 over the azimuth  $\theta$ , the effect of the array geometry will be merged in the result due to the  
 464 projection algorithm. To this end, we modify Equation 20 by omitting the projection operation. For  
 465 the interstation with a given orientation, beamforming is conducted only for the plane wave

466 incident from the azimuth consistent with the interstation orientation. The modified beamforming  
 467 can be expressed as

$$468 \quad MCBF(k, \omega, \theta_{ij}) = \frac{1}{N^2} \sum_i^{N(\theta_{ij})} \sum_j^{N(\theta_{ij})} \sqrt{\pi k r_{ij}} C_{ij}(r_{ij}, \omega) e^{i\left(kr_{ij} - \frac{\pi}{4}\right)} \quad (29)$$

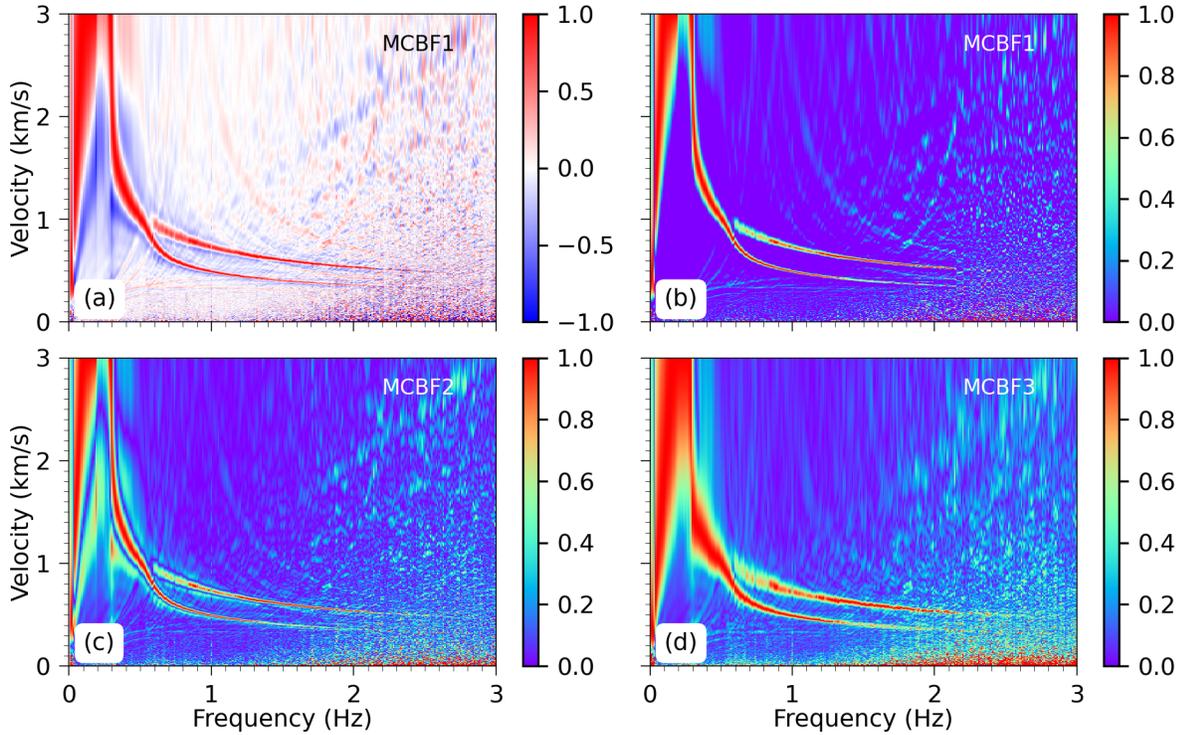
469 We call Equation 29 the modified cross-correlation beamforming (MCBF). The  $N(\theta_{ij})$  over the  
 470 summation symbol means only the station pair with orientation  $\theta_{ij}$  is taken, i.e., only the incident  
 471 plane wave along the direction of the station-pair orientation is considered. Correspondingly, the  
 472 vectors representing the wavenumber  $\mathbf{k}$  and distance  $\mathbf{r}_{ij}$  in Equation 20 degenerate into scalars  
 473 in Equation 29. The vector dot product  $\mathbf{k} \cdot \mathbf{r}_{ij}$  is replaced by the scalar product  $kr_{ij}$ . It must be  
 474 pointed out that  $-\pi/4$  phase shift is introduced in Equation 29 to balance the  $-\pi/4$  phase shift  
 475 appeared in CSDM caused by the stacking over noise sources in seismic interferometry. This  $-\pi/4$   
 476 phase shift is absent in Equation 20 because the projection and summation over  $\theta_{ij}$  would  
 477 automatically introduce  $-\pi/4$  phase shift, while the operation on the projection and summation  
 478 is neglected in Equation 29.

479 Using the same procedure as that for CBF and WCBF, the azimuth-averaged phase velocity can be  
 480 estimated by MCBF. Note that, the azimuth average of WCBF in Equation 21 is conducted over  $\theta$ ,  
 481 the azimuth of the incident plane wave. For MCBF, the azimuth average is performed over  $\theta_{ij}$ , the  
 482 orientation of the station-pairs. This average is nothing but the summation of all station-pairs and  
 483 thereby can be written as

$$484 \quad \overline{MCBF}(k, \omega) = \frac{1}{N^2} \sum_i^N \sum_j^N \sqrt{\pi k r_{ij}} C_{ij}(r_{ij}, \omega) e^{i\left(kr_{ij} - \frac{\pi}{4}\right)} \quad (30)$$

485 The imaging conditions can be written as

$$486 \quad \begin{aligned} MCBF1 &= \text{Re} \left[ \overline{MCBF}(k, \omega) \right] \\ MCBF2 &= \text{ABS} \left[ \text{Re} \left( \overline{MCBF}(k, \omega) \right) \right] \\ MCBF3 &= \text{ABS} \left[ \overline{MCBF}(k, \omega) \right] \end{aligned} \quad (31)$$



487  
 488 **Figure 6.** The dispersion images obtained by MCBF using the data from the array shown in Figure 1a. (a), (c) and  
 489 (e) are the results for the imaging conditions shown in Equation 31. (b) is the same as (a) but only the positive  
 490 value is color-coded.

491 Figure 6 shows the results obtained by MCBF using the data from the array shown in Figure 1a.  
 492 Figure 6a shows the normalized result of MCBF1, the real part of the azimuth-averaged  
 493  $\overline{MCBF}(k, \omega)$ . The positive maxima are observed along two dispersion branches. The negative  
 494 values, which are denoted by blue, appear adjacent to the maxima or spread in the image. These  
 495 negative values originate from the product of the sinc and trigonometric functions (See Equation  
 496 B13 in Appendix B for the theoretical representation of MCBF). If the absolute value of the real  
 497 part of  $\overline{MCBF}(k, \omega)$  is taken, i.e., the MCBF2 in Equation 31, side lobes would be observed on  
 498 both side of the maxima, as shown in Figure 6c. These side lobes would disappear if the modulus  
 499 of the of  $\overline{MCBF}(k, \omega)$  is taken as the imaging condition, but the width of the main lobe would  
 500 broaden as expected, as shown in Figure 6d. Therefore, a suitable display for the dispersion image  
 501 is to color-code only the positive values of the real part of  $\overline{MCBF}(k, \omega)$ , as shown in Figure 6b.  
 502 This display makes the dispersion image look clearer and easier to identify especially for the case

503 that the velocity of the modes is approaching.

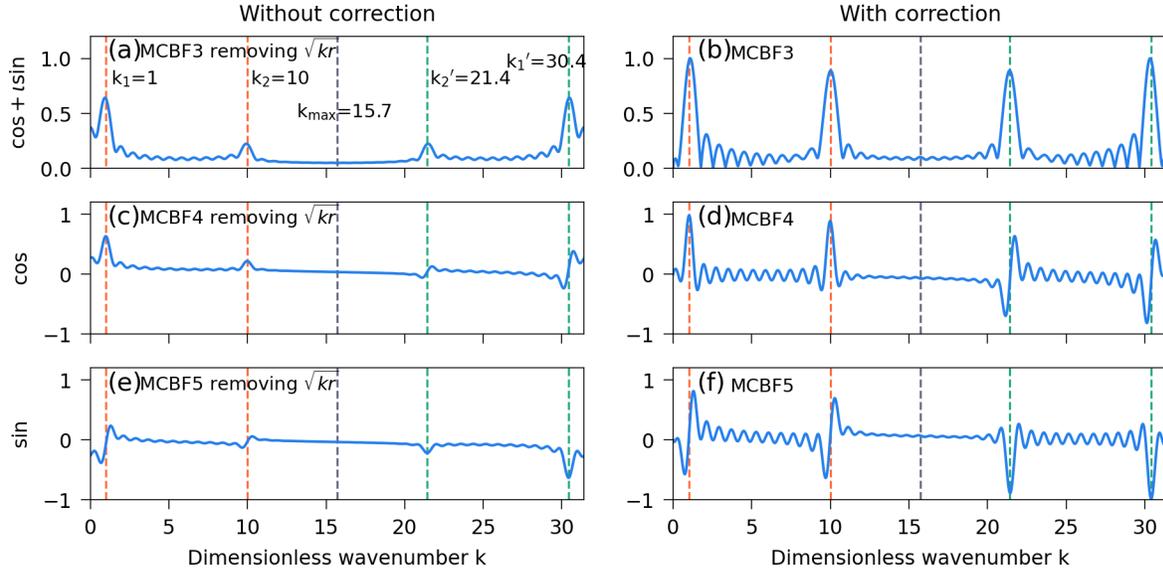
504 Moreover, the assumption on the laterally isotropy is not required in MCBF. For the case of  
 505 azimuthal anisotropy, the azimuth-dependence velocity can be extracted using MCBF by  
 506 processing the beamforming output at each azimuth  $\theta_{ij}$ , which is exactly the azimuth of the  
 507 incident plane wave. Different from the estimation of azimuth-dependence velocity using CBF or  
 508 WCBF, the spurious azimuthal anisotropy introduced by the array geometry would not be merged  
 509 into the results of MCBF since the projection to the orientation of the station pair in CBF or WCBF  
 510 is omitted in MCBF. The correction on the azimuth anisotropy introduced by array geometry, as  
 511 done in [Lu et al. \(2018\)](#), is not required.

#### 512 **4.2. The Other Imaging Conditions**

513 Theoretically, the cross spectra  $C_{ij}(r_{ij}, \omega)$  of the vertical component in Equation 30 is real and  
 514 equivalent to the Bessel function  $J_0(kr)$  if the noise field is perfectly isotropic. However, it is  
 515 usually complex in practice due to the complicated source feature. Different from Equation 30, we  
 516 can use the real part of  $C_{ij}(r_{ij}, \omega)$  to approximate  $J_0(kr)$ . At the same time, instead of the complex  
 517 exponential function, the sine or cosine functions are used to present the plane wave. The  
 518 corresponding imaging conditions can be expressed as

$$\begin{aligned}
 \text{519} \quad MCBF4(k, \omega) &= \frac{1}{N^2} \sum_i^N \sum_j^N \sqrt{kr_{ij}} \operatorname{Re} [C_{ij}(r_{ij}, \omega)] \cos \left( kr_{ij} - \frac{\pi}{4} \right) \\
 MCBF5(k, \omega) &= \frac{1}{N^2} \sum_i^N \sum_j^N \sqrt{kr_{ij}} \operatorname{Re} [C_{ij}(r_{ij}, \omega)] \sin \left( kr_{ij} - \frac{\pi}{4} \right)
 \end{aligned} \tag{32}$$

520 It is assumed that the cylindrical wavefield described by  $J_0(kr)$  contains two surface wave  
 521 eigenvalues  $k_1 = 1$  and  $k_2 = 10$ . We investigate the effect of different imaging conditions on the  
 522 result by comparing the beamforming output of such an ideal cylindrical wavefield. The results are  
 523 given in Figure 7. In simulation, the spatial wavefield is uniformly sampled with  $\Delta r = 0.2$  and the  
 524 maximum distance  $r_{\max} = 6.0$ . The wavenumber and distance are dimensionless.



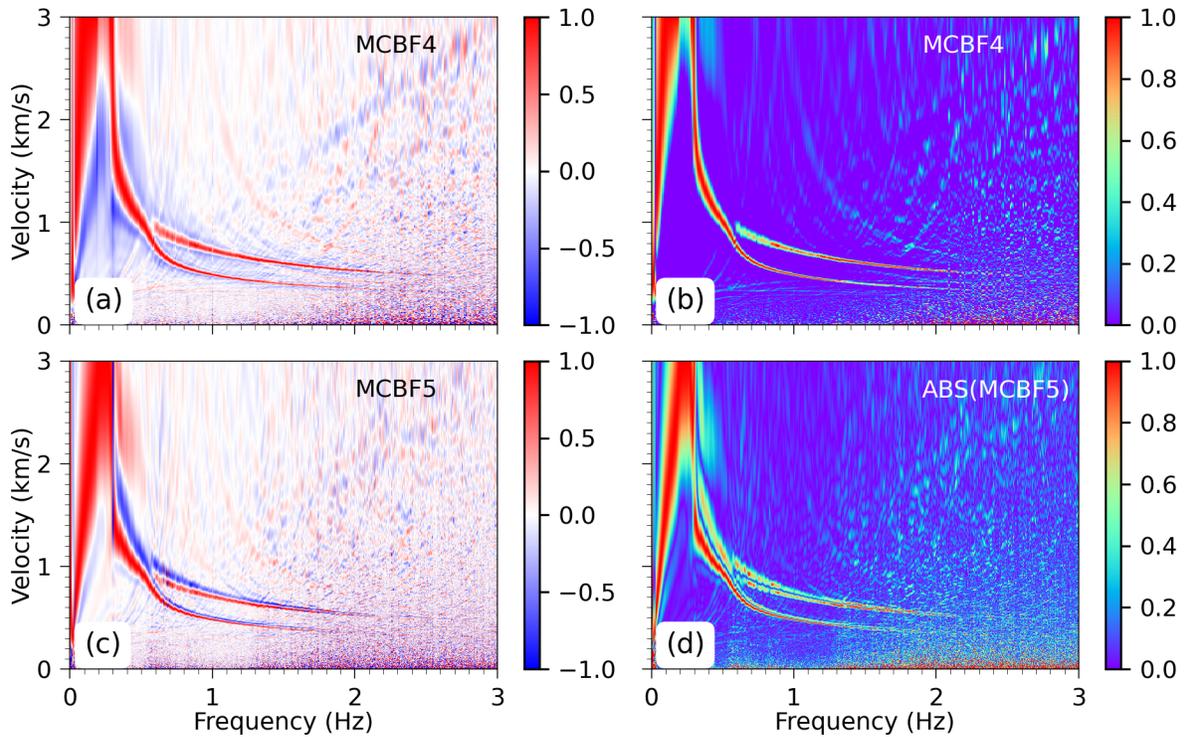
525

526 **Figure 7.** The beamforming output by applying MCBF with different basis functions to a cylindrical wavefield  
 527 containing two eigen-wavenumbers ( $k=1$  and  $k=10$ ). (b) The result of MCBF3, where the complex exponential  
 528 function is taken as the basis function. (d) and (f) are the results of MCBF4 and MCBF5, where the cosine and  
 529 sine functions are taken as the basis function, respectively. (a), (c) and (e) show the corresponding results by  
 530 removing the term  $\sqrt{kr}$  in Equations 30 and 32.

531 In Figure 7, right panels show the results of MCBF3, MCBF4 and MCBF5. The left panels show the  
 532 corresponding results without the correction for geometric spread, i.e., the results of MCBF3,  
 533 MCBF4 and MCBF5 but removing the correction term  $\sqrt{kr}$ . The vertical red dashed line denotes  
 534 the location of the eigen-wavenumbers. The vertical gray dashed line denotes the location of  
 535 resolving maximum wavenumber  $k_{\max} = \pi/\Delta r = 15.7$  estimated by Equation 15. Aliasing occurs at  
 536 the wavenumbers greater than  $k_{\max}$  ( $k > k_{\max}$ ). The aliasing wavenumbers are  
 537  $k'_1 = 2k_{\max} - k_1 = 30.4$  and  $k'_2 = 2k_{\max} - k_2 = 21.4$ , which are denoted by vertical green dashed lines.

538 It can be found from Figure 7: 1) The result of MCBF5 is the antisymmetry of that of MCBF4, as  
 539 shown in Figures 7d and 7f. The target wavenumbers  $k_1$  and  $k_2$  are associated with the maxima  
 540 of the beamforming output if cosine function is selected as the basis, while they are associated  
 541 with the zero-crossing points for sine basis function. On the contrary, the aliasing wavenumbers  
 542  $k'_1$  and  $k'_2$  are associated with the (negative) maxima for sine basis function, while they are

543 associated with the zero-crossing points for cosine basis function. If the modulus is taken as the  
 544 imaging conditions, the zero-crossing points associated with the target (or aliasing) wavenumbers  
 545 would behave as a trough between two extremes (See Figure S1 in the supporting information  
 546 where the modulus of MCBF4 and MCBF5 is plotted). 2) Compared with the results of MCBF4  
 547 based on the cosine function, the results of MCBF3 based on the exponential function have a wider  
 548 main lobe but fewer side lobes. This is more evident if the modulus of MCBF4 is taken (See Figure  
 549 S1 in the supporting information). 3) Both the target wavenumbers  $k_1, k_2$  and the aliasing  
 550 wavenumbers  $k'_1, k'_2$  are associated with the maximum for the result of MCBF3. 4) The actual  
 551 amplitudes, which are supposed to be 1 for two eigen-wavenumbers, can not be recovered without  
 552 the correction term  $\sqrt{kr}$ . This results in a dispersion image that lacks clarity at high frequencies,  
 553 as shown in Figure 1b.



554  
 555 **Figure 8.** The dispersion images obtained by MCBF (Equation 32) using the data from the array shown in Figure  
 556 1a. (b) is the same as (a) but only the positive value is color-coded.  
 557 Figure 8 shows the results obtained by MCBF4 and MCBF5 using the data from the array shown in  
 558 Figure 1a. As discussed above, the eigenvalues along the dispersion branches corresponds to the

559 maxima for the results of MCBF4. For the results of MCBF5 based on sine function, the eigenvalues  
 560 along the dispersion branches corresponds to the zero-crossing points. Since there are many zero  
 561 points spread in the image, it brings the handicap for the picking of the eigenvalues along dispersion  
 562 branches. Color-coding only the positive value of MCBF4 is proved to be a better imaging condition,  
 563 as shown in Figure 8b, where the dispersion image has a high resolution without side lobes such as  
 564 occurred in Figure 8a, 8c and 8d. Furthermore, the result of MCBF4 is almost the same as those of  
 565 MCBF1 shown in Figure 6. This means the distribution of the noise source is fairly uniform and  
 566 hence the SPAC coefficient is supposed to be close to  $J_0(kr)$  with a negligible imaginary part.

### 567 **4.3. Relation Between MCBF and F-J Method**

568 In practice, the sampling on the wavefield in space is always finite and usually non-uniform due to  
 569 the finite NCFs and irregular array. The integral over the distance  $r$  in Equation 25 from 0 to infinite  
 570 is often replaced by a finite interval. Since the relative rather than the absolute magnitude of the  
 571 spectrum in f-v domain is of interested, the integral is therefore normalized by the integral length.  
 572 We have

$$573 \quad FJ(k, \omega) = \frac{1}{R} \int_0^R C_{ij}(r, \omega) J_0(kr) r dr \quad (33)$$

574 For an array with N stations, the number of the station-pairs is  $M = N(N-1)/2$ . The interstation  
 575 distances are arranged in order from smallest to largest as  $r_1, r_2, \dots, r_j, \dots, r_M$ . The integral of  
 576 Equation 33 can be approximated numerically by

$$577 \quad \begin{aligned} FJ(k, \omega) = & \frac{1}{r_2 - r_1} \int_{r_1}^{r_2} C(r) J_0(kr) r dr + \dots \\ & + \frac{1}{r_{j+1} - r_j} \int_{r_j}^{r_{j+1}} C(r) J_0(kr) r dr + \dots \\ & + \frac{1}{r_M - r_{M-1}} \int_{r_{M-1}}^{r_M} C(r) J_0(kr) r dr \end{aligned} \quad (34)$$

578 Using the trapezoidal integral formula

$$579 \quad \int_a^b f(x) dx = (b-a) \frac{f(a) + f(b)}{2} \quad (35)$$

580 Equation 34 can then be expressed as

$$\begin{aligned}
FJ(k, \omega) &\approx \frac{1}{2} C(r_1) J_0(kr_1) r_1 + \sum_{j=2}^{M-1} C(r_j) J_0(kr_j) r_j + \frac{1}{2} C(r_M) J_0(kr_M) r_M \\
581 \quad &\approx \sum_{j=1}^M C(r_j) J_0(kr_j) r_j \approx \sum_{j=1}^M C(r_j) r_j \sqrt{\frac{2}{\pi k r_j}} \cos(kr_j - \frac{\pi}{4}) \\
&\approx \sqrt{\frac{2}{\pi k}} \sum_{j=1}^M C(r_j) \sqrt{r_j} \cos(kr_j - \frac{\pi}{4})
\end{aligned} \tag{36}$$

582 The far approximation of  $J_0(kr)$  shown in Equation 19 are applied in Equation 36.

583 On the other hand, the azimuth-averaged MCBF shown in Equation 30 can be recast into

$$584 \quad \overline{MCBF}(k, r) = \frac{1}{N^2} \sum_i^N \sum_j^N \sqrt{\pi k r_{ij}} C_{ij}(r_{ij}, \omega) e^{i(kr_{ij} - \frac{\pi}{4})} = \frac{\sqrt{2\pi k}}{N^2} \left( \frac{\sqrt{2}}{\sqrt{\pi k}} \sum_{j=1}^M C(r_j) \sqrt{r_j} e^{i(kr_j - \frac{\pi}{4})} \right) \tag{37}$$

585  
586 In practical applications, the real part of  $C(r_{ij})$  is usually taken in F-J method. Ignoring the  
587 difference caused by tiny imaginary part and comparing Equations 36 and 37, it can be found that  
588 the relationship between MCBF and F-J is

$$589 \quad \text{Re}[\overline{MCBF}(k, \omega)] \approx MCBF4 = \frac{\sqrt{2\pi k}}{N^2} FJ(k, \omega) \tag{38}$$

590 That is to say, as far as the measurement of the azimuth-averaged phase velocity, MCBF is also  
591 equivalent to F-J, differing by a factor of  $1/k$ . If we dividing MCBF4 by  $k$  to remove the correction  
592 on the wavenumber in Figure 8b, the disturbing pixels around the area near the zero velocity would  
593 be reduced and the resulted dispersion image would comparable to those shown in Figure 4c and  
594 4e, the results of WCBF3 and F-J. Although it is the cosine rather than Bessel function appears in  
595 MCBF, the correction on the wavenumber in MCBF reduces the near field effect. The extracted  
596 velocity values are quite consistent with that given by WCBF and F-J, even at low frequencies, as  
597 shown in Figure 4f.

598 As discussed before, similar as the extraction of azimuth-averaged velocity by CBF and WCBF, the  
599 velocity obtained by F-J is also the azimuth-averaged. Although F-J and MCBF are also equivalent  
600 for the estimation on azimuth-averaged velocity, the azimuth-dependence velocity can be  
601 estimated using MCBF by picking the velocities associated with the maxima of the beampower at  
602 a given azimuth. This cannot be achieved in principle via F-J by considering only the NCFs at a given  
603 azimuth. This is because the Bessel function appears in the F-J integration is the result of integrating

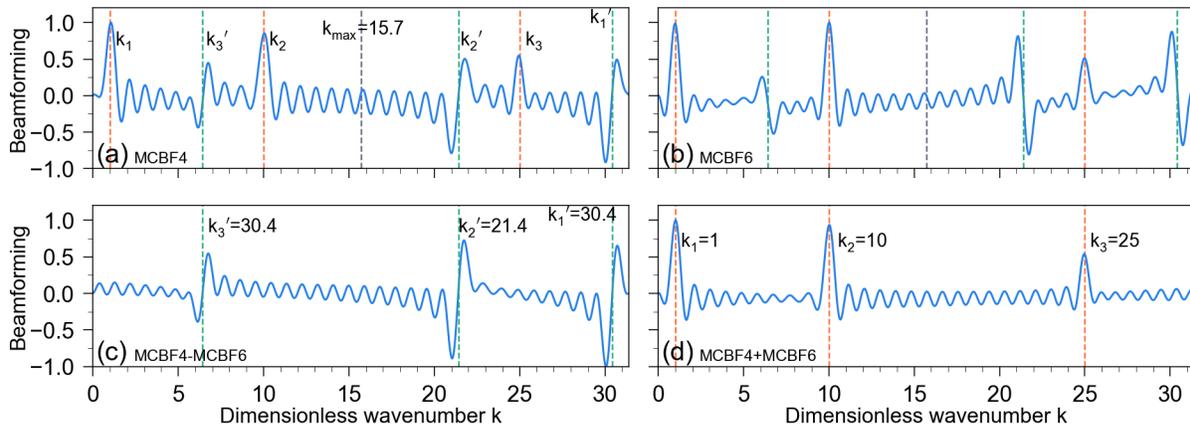
604 the exponential function over the omnidirectional azimuth. If only the NCFs at a given azimuth are  
 605 considered in F-J, the Bessel function should be replaced by the exponential function, i.e., the MCBF  
 606 scheme.

## 607 5. Aliasing and Its Reduction

608 The aliasing is inevitable in practical applications due to the finite sampling. In this section, we  
 609 demonstrate the aliasing features in terms of a simple cylindrical wavefield containing three eigen-  
 610 wavenumbers. Based on the theoretical representation of MCBF for the array with finite stations,  
 611 a scheme is then proposed to eliminate the aliasing introduced by negative wavenumber.

### 612 5.1. Aliasing and Its Features

613 The CSDM of a cylindrical wavefield containing three dimensionless eigen-wavenumbers  $k_1 = 1$ ,  
 614  $k_2 = 10$  and  $k_3 = 25$  can be expressed as  $C(r, \omega) = J_0(k_1 r) + J_0(k_2 r) + J_0(k_3 r)$ . For the same  
 615 sampling as that used in Figure 7, i.e.,  $\Delta r = 0.2$  and  $r_{\max} = 6.0$ , the resolved maximum and  
 616 minimum wavenumber are respectively  $k_{\max} = \pi / r_{\min} = 15.7$  and  $k_{\min} = \pi / r_{\max} = 0.52$ .



617 **Figure 9.** An illustration for aliasing wavenumbers and its reduction. The results are calculated by different imaging  
 618 conditions for a cylindrical wavefield containing three eigenvalues of horizontal wavenumbers  $k_1 = 1$ ,  $k_2 = 10$   
 619 and  $k_3 = 25$ , which are denoted by vertical red dashed lines. The corresponding aliasing wavenumbers  $k_1'$ ,  $k_2'$   
 620 and  $k_3'$  are denoted by vertical green dashed lines. The resolved maximum wavenumber is denoted by gray  
 621 dashed line. (a) The result of MCBF4. (b) The result of MCBF6. (c) The result of MCBF4-MCBF6. (d) The result of  
 622 MCBF4+MCBF6.

624 Figure 9a shows the beamforming output of MCBF4. Three target eigen-wavenumbers are denoted  
625 by vertical red dashed lines.  $k_{\max}$  estimated by Equation 15 is denoted by gray dashed line. Two  
626 aliasing wavenumbers  $k'_1 = 2k_{\max} - k_1 = 30.4$  and  $k'_2 = 2k_{\max} - k_2 = 21.4$  caused by the symmetry of  
627  $k_1$  and  $k_2$  are observed at  $k > k_{\max}$ , as shown by the vertical green lines. Similarly, due to the  
628 symmetry of  $k_3$ , aliasing wavenumber  $k'_3 = 2k_{\max} - k_3 = 6.4$  is observed at  $k < k_{\max}$ . The target  
629 wavenumbers correspond to the positive peaks of the beamforming output, while the aliasing  
630 wavenumbers are related to the zero-crossing points between two positive and negative peaks. If  
631 the modulus is taken as the imaging condition, the aliasing wavenumbers would behave as a trough  
632 between two extremes (See Figure S2 in the supporting information where the modulus is taken  
633 as the value for plotting). The presence of  $k'_3$  could affect the identification of the eigen-  
634 wavenumbers  $k_1$  and  $k_2$ , even if only the resolved target wavenumbers range  $[k_{\min}, k_{\max}]$   
635 determined by the sampling theorem is considered.

## 636 5.2. Theoretical Representation for Aliasing and Its Reduction

637 Aliasing arises from the sampling on the traveling waves at two directions (Forbriger, 2003). The  
638 spatial wavefield described by  $J_0(kr)$  can be thought as a cylindrical standing wave, which is the  
639 superposition of two cylindrical traveling waves propagating inward and outward. The  $J_0(k_n r)$  can  
640 be expressed as

$$641 \quad J_0(k_n r) = 1/2 \left[ H_0^{(1)}(k_n r) + H_0^{(2)}(k_n r) \right] \quad (39)$$

642 where  $H_0^{(1)}$  and  $H_0^{(2)}$  are respectively the first and second kind of Hankel function with zero-order.  
643 According to the convention of Fourier transform given in Appendix A,  $H_0^{(1)}$  represents the  
644 cylindrical wave propagating inwards while  $H_0^{(2)}$  represents the cylindrical wave propagating  
645 outwards. For the imaging condition MCBF4 in Equation 32, a plane wave represented by cosine  
646 function is used to fit the cylindrical standing wave via the operation of delay-and-sum. This means  
647 that two cylindrical waves propagating in opposite directions are fitted simultaneously. The eigen-  
648 wavenumbers could be also estimated for the propagating wave inwards, i.e., propagating along  
649 direction  $-\mathbf{k}$ . In the range  $[0, 2k_{\max}]$ , the aliasing wavenumber  $k'_n$  can be observed at

650  $k'_n = 2k_{\max} - k_n$  due to the symmetry with the target wavenumber for  $+\mathbf{k}$  direction.

651 Similar to the discussion in [Xi et al. \(2021\)](#) and [Zhou and Chen \(2021\)](#), we define MCBF6 as

$$652 \quad MCBF6(k, \omega) = \frac{1}{N^2} \sum_i^N \sum_j^N \sqrt{\pi r_{ij} k} \left\{ \mathcal{H} \left( \text{Re} [C_{ij}(r_{ij}, \omega)] \right) \sin \left( kr_{ij} - \frac{\pi}{4} \right) \right\} \quad (40)$$

653 The symbol  $\mathcal{H}$  represents the Hilbert transform, the convention of which is given in appendix A.

654 For the ideal cylindrical wavefield,  $C(r, \omega) = J_0(k_n r)$ , and its Hilbert transform is  $Y_0(k_n r)$ , the zero

655 order Bessel function of the second kind. Figure 9b shows the output of MCBF6, where both the

656 actual and aliasing wavenumbers are observed. For the beampower around the positive peaks

657 associated with the target wavenumbers, the result of MCBF6 is the same as that of MCBF4.

658 However, for the beampower around the zero-crossing points associated with aliasing

659 wavenumbers, the result of MCBF6 is the Centro symmetry of MCBF4. This means that the aliasing

660 wavenumbers can be removed by summing the results of MCBF4 and MCBF6. In fact, the

661 elimination of aliasing wavenumbers can be demonstrated analytically by investigating the

662 theoretical representations of MCBF4 and MCBF6 for the finite sampling.

663 Substituting  $C(r, \omega) = J_0(k_n r)$  into Equations 32 and 40, for the sampling with small equal interval

664  $\Delta r$ , we have ([See appendix B for details](#))

$$665 \quad \begin{aligned} MCBF4 &= \frac{\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left[ \cos \frac{R(k_n - k)}{2} \text{sinc} \frac{R(k_n - k)}{2} + \sin \frac{R(k_n + k)}{2} \text{sinc} \frac{R(k_n + k)}{2} \right] \\ MCBF6 &= \frac{\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left[ \cos \frac{R(k_n - k)}{2} \text{sinc} \frac{R(k_n - k)}{2} - \sin \frac{R(k_n + k)}{2} \text{sinc} \frac{R(k_n + k)}{2} \right] \end{aligned} \quad (41)$$

666 where  $\text{sinc}(x) = \sin x/x$  is the sinc function, and  $M = N(N-1)/2$ . It can be found from Equation

667 41 that the terms containing  $k_n - k$  and  $k_n + k$  are linearly separated. The linear combinations of

668 MCBF4 and MCBF6 can be used to separate the terms containing only the argument  $k_n - k$  from

669 the terms containing only argument  $k_n + k$ . The corresponding imaging conditions can be

670 expressed as

$$\begin{aligned}
MCBF4 + MCBF6 &= \frac{2\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \cos \frac{R(k_n - k)}{2} \operatorname{sinc} \frac{R(k_n - k)}{2} \\
MCBF4 - MCBF6 &= \frac{2\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \sin \frac{R(k_n + k)}{2} \operatorname{sinc} \frac{R(k_n + k)}{2}
\end{aligned} \tag{42}$$

671  
672 Figure 9c shows the result of  $MCBF4 - MCBF6$ . Only the aliasing wavenumbers are observed at  
673 the zero-crossing points which are associated with  $k'_n = 2k_{\max} - k_n$  ( $n = 1, 2, 3$ ). Figure 9d shows the  
674 result of  $MCBF4 + MCBF6$  where only the actual wavenumbers propagating in  $\mathbf{k}$  direction are  
675 kept, which are associated with the positive peaks at  $k = k_n$ . Therefore, the linear combination  
676  $MCBF4 + MCBF6$  can be used to remove the aliasing caused by the propagating wave in  $-\mathbf{k}$   
677 direction.

### 678 5.3. Numerical Simulation and the Example of the Field Data

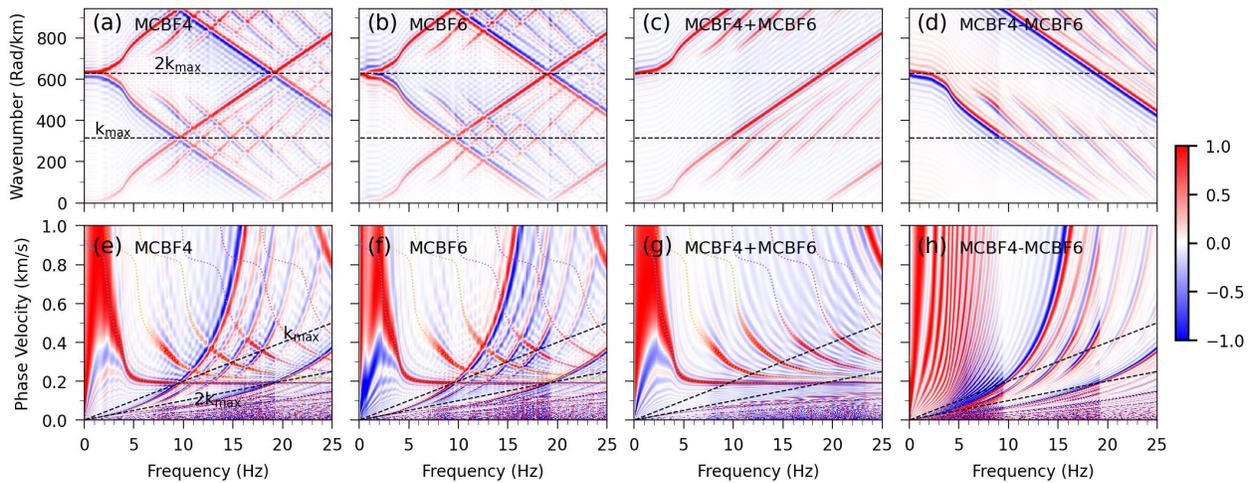
679 **Table 1.** The layered model (Zhou and Chen, 2021)

Layer thickness	P wave velocity	S wave velocity	Density
(km)	(km/s)	(km/s)	( $g / cm^3$ )
0.025	1.35	0.2	1.9
$\infty$	2.0	1	2.5

680  
681 In this section, the aliasing and its reduction are investigated by numerical simulation. The two-  
682 layered model shown in Table 1 is considered. This model has been discussed in Wathelet et al.  
683 (2008), Xi et al. (2021) and Zhou and Chen (2021). The quality factor Q is neglected here since it  
684 does not affect current discussion. The vertical component of Rayleigh wave is considered in the  
685 numerical simulation. Substituting the CSDM represented by Equation 18 into Equations 32 and 40,  
686 the dispersion images in f-v and f-k domain can be obtained by the imaging conditions MCBF4 and  
687 MCBF6 and their linear combinations.

688 In the simulation, we take the maximum radial distance  $r_{\max} = 200m$  and the interval  $\Delta r = 10m$ .  
689 Figure 10 shows the final results for different imaging conditions. Figure 10a and 10e show the  
690 results of MCBF4 in f-k and f-v domain, respectively. The corresponding results of MCBF6 are

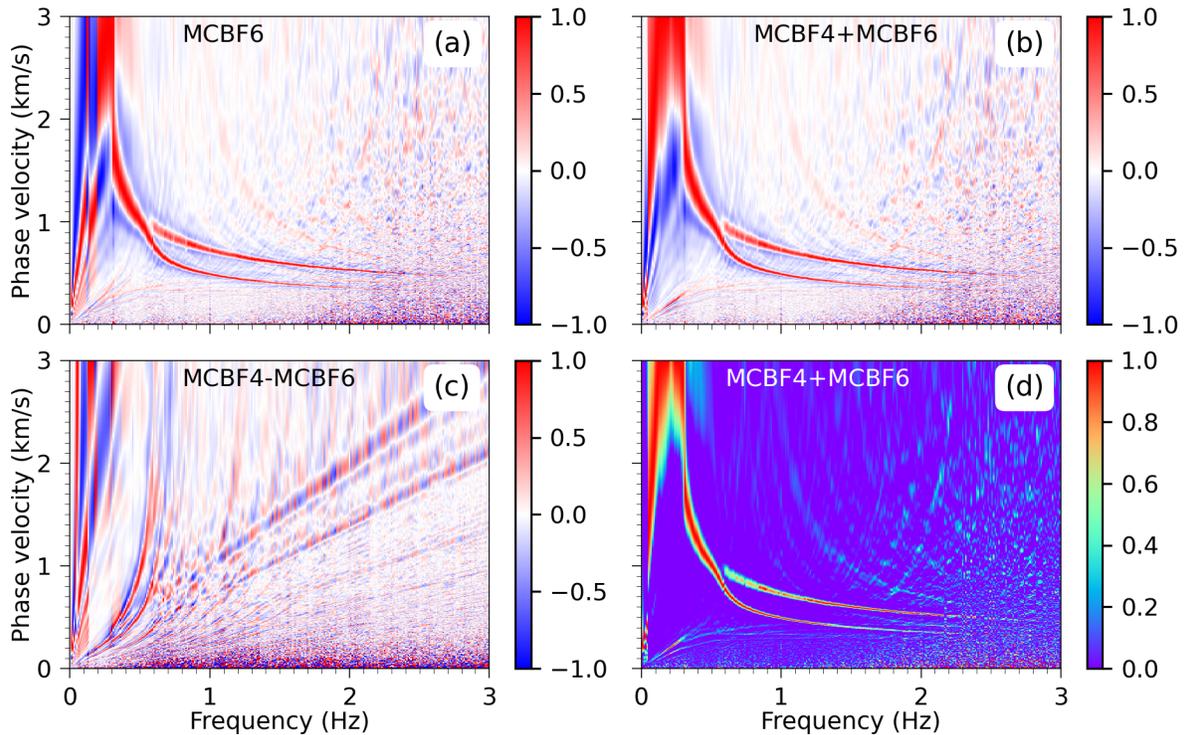
691 presented in Figure 10b and 10f. The aliasing wavenumbers caused by the waves propagating in  
 692  $-\mathbf{k}$  direction can be observed in Figure 10a and 10b, which appears as the maxima energy belt  
 693 with negative slope. In f-v domain, as shown by Figure 10e and 10f, this aliasing appears as a  
 694 hyperbolic shape with vertex close to the origin. As expected, the sign of the result on either side  
 695 of the zero-crossing points associated with the aliasing wavenumber is reversed for MCBF4 and  
 696 MCBF6. The target wavenumbers are associated with the maxima both for MCBF4 and MCBF6. As  
 697 a result, this aliasing can be separated and removed using the linear combinations of MCBF4 and  
 698 MCBF6. The separated aliasing in f-k and f-v domain are respectively shown in Figure 10d and 10h,  
 699 which are the results of  $MCBF4 - MCBF6$ . Figure 10c and 10g show the corresponding results of  
 700  $MCBF4 + MCBF6$ , where the aliasing caused by waves propagating in  $-\mathbf{k}$  direction is removed.



701  
 702 **Figure 10.** An illustration of the aliasing and its reduction. The panels at top row are the results in f-k domain  
 703 while the panels at lower row are the corresponding results in f-v domain. (a) The dispersion image in f-k domain  
 704 obtained MCBF4. (b) The result of MCBF6. (c)The result of MCBF4+MCBF6. The aliasing caused by the waves  
 705 propagating in  $-\mathbf{k}$  direction is removed. (d)The separated aliasing obtained by MCBF4-MCBF6. (e)-(h) are the  
 706 corresponding results in f-v domain of (a)-(d). The wavenumber of  $k_{\max}$  and  $2k_{\max}$  are denoted by black dashed  
 707 lines.

708 In Figure 10a, there also exist another periodic aliasing caused by limited sampling. Opposite to the  
 709 aliasing shown in Figure 10d, the energy belts presented this periodic aliasing has a positive slope  
 710 in f-k domain, as shown in the area above the dashed line of  $k = 2k_{\max}$  in Figure 10c. This periodic  
 711 aliasing repeats with a period of  $2k_{\max}$  and cannot be removed using  $MCBF4 + MCBF6$ . As an

712 artifact, this periodic aliasing appears more pronounced in f-k domain. In f-v domain, as shown in  
 713 Figure 10g, it appears in the area under the black dashed line of  $k = 2k_{\max}$ , a region with high  
 714 frequency but low velocity. Therefore, in the resolved frequency range determined by sampling,  
 715 this aliasing generally has little impact on the identification of dispersion curves in f-v domain.



716  
 717 **Figure 11.** An illustration for the separation and elimination of the aliasing wavenumbers for the array shown in  
 718 Figure 1a. (a) The result of MCBF6. (b) The result of MCBF4+MCBF6. (c) The result of MCBF4-MCBF6. (d) The  
 719 same as (b) but only the positive values are color-coded.

720 Figure 11 presents the results of MCBF6 and the linear combination of MCBF4 and MCBF6 for the  
 721 data from the array shown in Figure 1a. It shows the improvement of the dispersion image after  
 722 removing aliasing by the imaging conditions MCBF4+MCBF6 is not significant. In other words, the  
 723 aliasing of the dispersion image given by MCBF4 itself is not obvious, as shown in Figure 8b. We  
 724 speculate that this is related to the irregular station distribution of the real array. The summation  
 725 over the NCFs inside the array reduces the aliasing. This is different from the numerical simulation  
 726 shown in Figure 10e, where the regular sampling with equal intervals is adopted. The fact that the

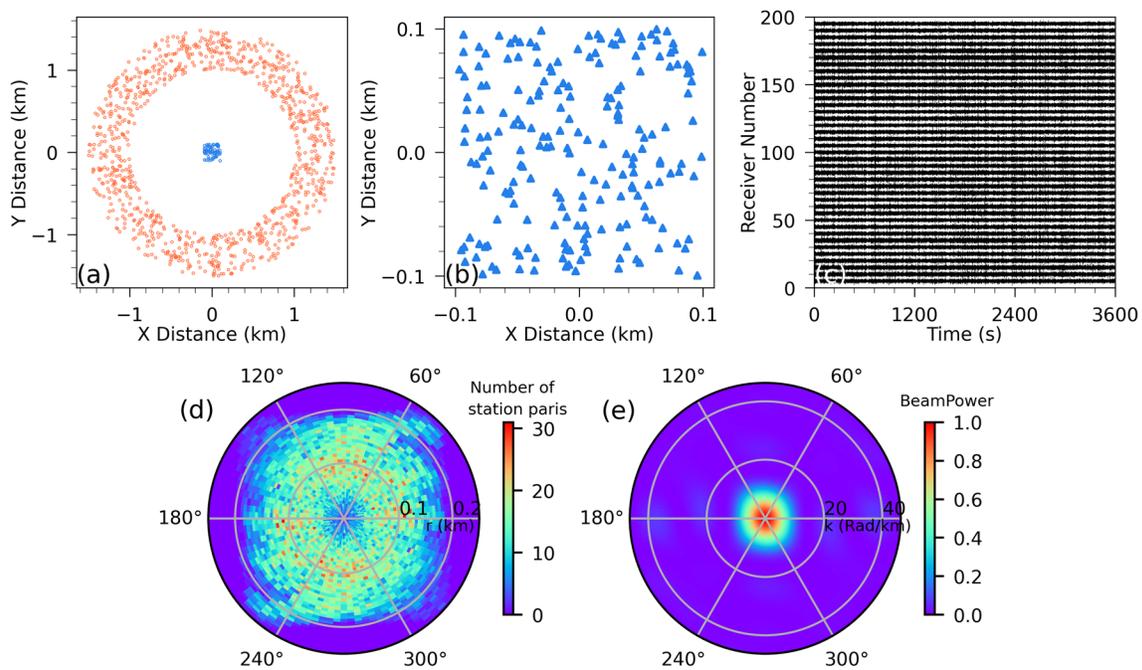
727 aliasing can be reduced by random station distribution can also be observed for the synthetic data  
728 in section 6.1.

## 729 6. Synthetic Data and More Examples of the Real World

### 730 6.1. Example of the Synthetic Data

#### 731 6.1.1. Model and the Synthesis of Ambient Seismic Noise

732 In this section, we investigate the extraction of multimode dispersion curves of Rayleigh wave using  
733 WCBF and MCBF based on the synthetic data. The same layered model shown in Table 1 is  
734 considered.

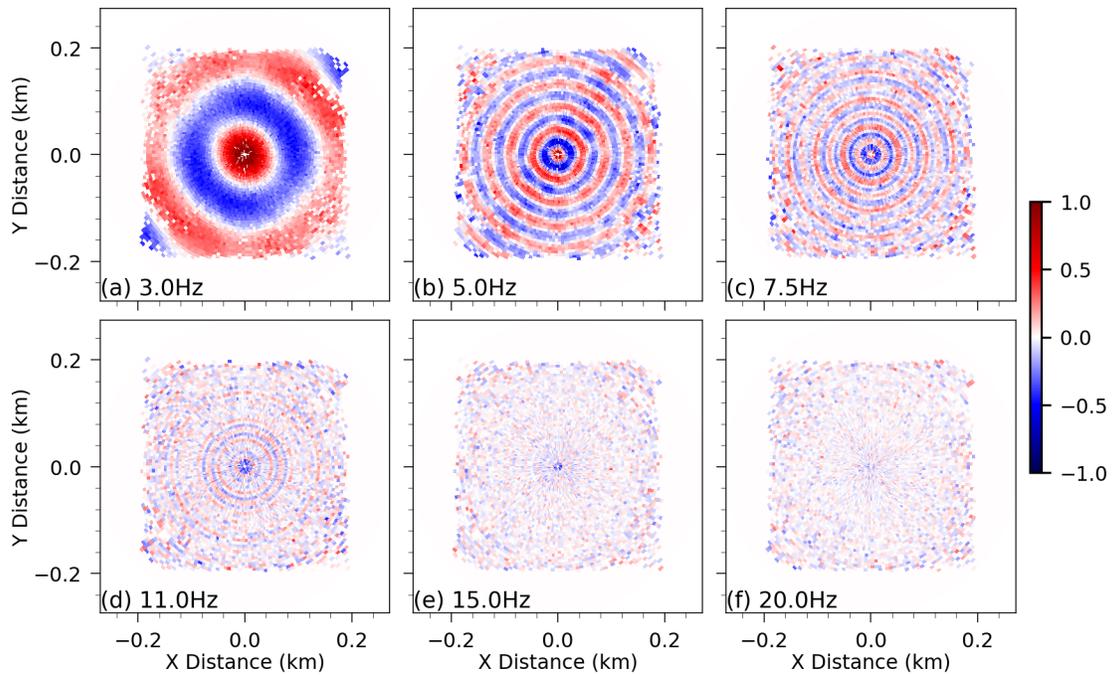


735  
736 **Figure 12.** (a) The source (red dots) distribution and the receiver (blue triangles) array. (b) The random distributed  
737 stations of the receiver array. (c) Examples on the segments of synthetic ambient noise recording. (d) The  
738 distribution of the number of station pairs as a function of the interstation distance and azimuth. (e) The ARF of  
739 the receiver array.

740 To synthesize the data, 10,000 sources randomly distributed over 1.0-1.5 km annular region are  
741 used to excite the vertical component of the Rayleigh wave, as shown by the red dots in Figure 12a.  
742 Each source is assumed to be the vertical point force with random intensity between 0-1. The  
743 source function is a Ricker wavelet with center frequencies distributed randomly between 0-25 Hz.

744 The onset time of the sources is distributed randomly between 0-3,600 s. The waveform at the  
 745 receiver excited by each source is calculated by the dot product of the vertical point force and the  
 746 Green's function shown in Equation 18. The noise data with 1-hour duration is then synthesized by  
 747 summing the waveform for all sources. As shown in Figure 12a and 12b, 200 stations distributed  
 748 randomly within a square array of  $0.2 \text{ km} \times 0.2 \text{ km}$  are designed to record the noise data. Figure  
 749 12c shows some segments of the synthetic ambient noise recording at 40 stations. Figure 12d  
 750 shows the distribution of the number of station pairs as a function of the interstation distance and  
 751 azimuth. Figure 12e shows the ARF of the array.

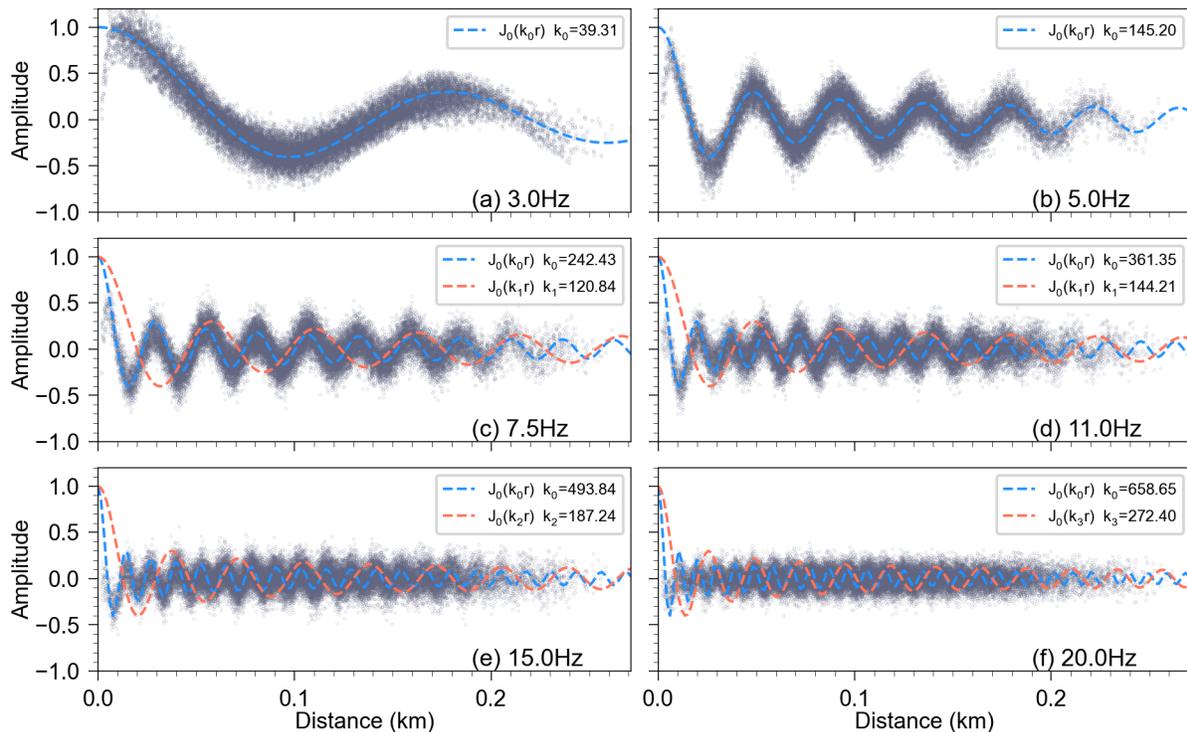
752 **6.1.2. Wavefield and Multimode Dispersion Curves Extracted by WCBF and MCBF**



753 **Figure 13.** The wavefield (real part of the Fourier transformed NCFs) at six selected frequencies across the array  
 754 shown in Figure 12b, plotted as a function of the orientation (azimuth) and interstation distance (radial direction).  
 755 The amplitude in each panel is normalized by the maximum.

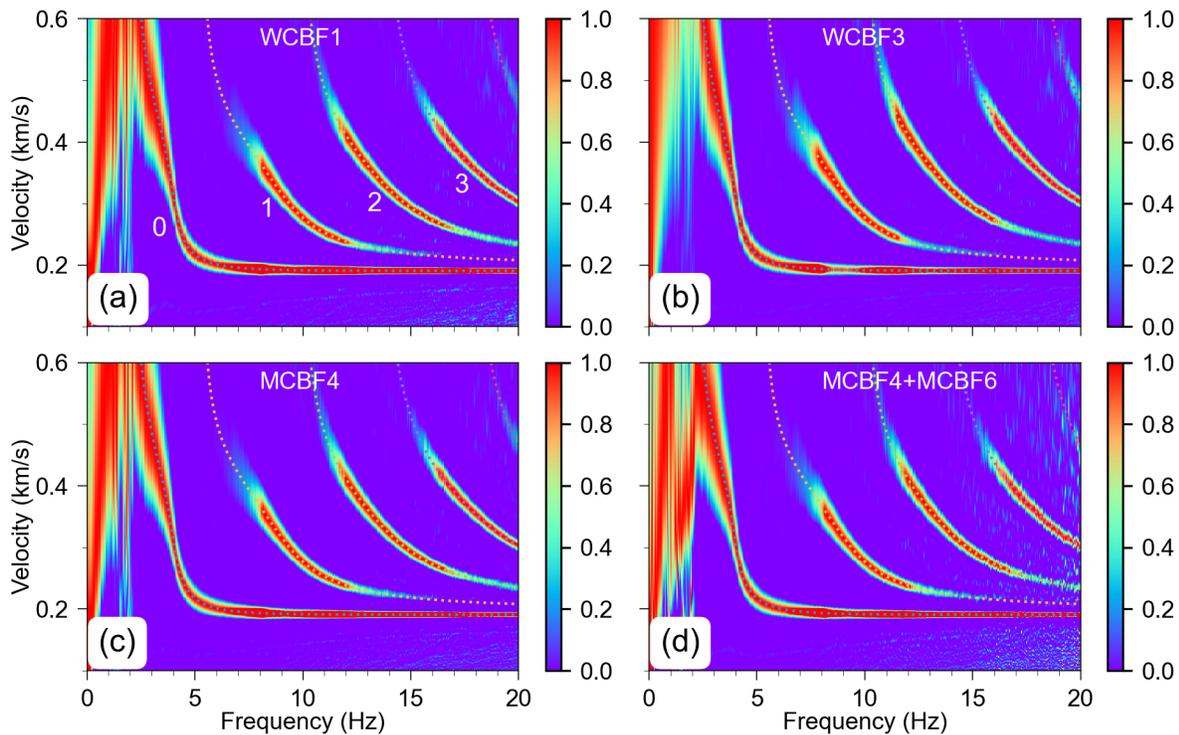
757 Plotted as a function of the orientation and interstation distance, the wavefield at six selected  
 758 frequencies across the array shown in Figure 12b are presented in Figure 13. They are the real part  
 759 of the Fourier transformed NCFs of the interstation inside the array. The amplitude in each panel is  
 760 normalized by the maximum.

761 As we discussed in sections 3.2 and 4.3, under the assumption of lateral isotropy, if the distribution  
 762 of the sources and the interstation orientation is uniform enough, the wavefield would be radially  
 763 symmetric. The wavefield at different azimuth can be projected into the radial direction, as shown  
 764 in Figure 14. The spectrum of NCFs is then just a radial sampling on the cylindrical wavefield across  
 765 the array by arranging the NCFs according to their interstation distance. For F-J method, the integral  
 766 over the distance  $r$  is conducted using the sampling values shown in Figure 14. This operation is  
 767 similar as the azimuth-average in BF.



768  
 769 **Figure 14.** The wavefield variation as a function of radial distance by projecting results at six frequencies in Figure  
 770 13 into one direction. The blue and red dashed lines represent the Bessel functions  $J_0(kr)$  with argument  $kr$ .  
 771 In Figure 14, the blue and red dashed lines represent the Bessel functions  $J_0(kr)$  with argument  
 772  $kr$ . The wavenumber  $k$  associated with the eigenvalues are labeled with corresponding color. It  
 773 can be found the variation of the wavefield can be approximated by  $J_0(kr)$  if only one eigen-  
 774 wavenumber is expected, as illustrated by Figures 14a and 14b where only one mode is observed  
 775 at frequencies 3 and 5 Hz. If more than one eigen-wavenumbers are expected, the wavefield will

776 be the weighted superposition such as  $\sum_n W_n J_0(k_n r)$  as shown in Figure 14c-14f where more  
 777 than one mode is observed at frequencies 7.5, 11, 15 and 20 Hz. It seems that the Bessel function  
 778  $J_0(k_0 r)$  with characteristic wavenumber  $k_0$  of the fundamental mode fit well with the synthetic  
 779 wavefield in Figure 14c-14f. This is because the fundamental mode makes a major contribution to  
 780 the synthetic record (See Figure S3 in the supporting information for the fit of the theoretical and  
 781 synthetic wavefield at frequencies 7.5, 11, 15 and 20 Hz).



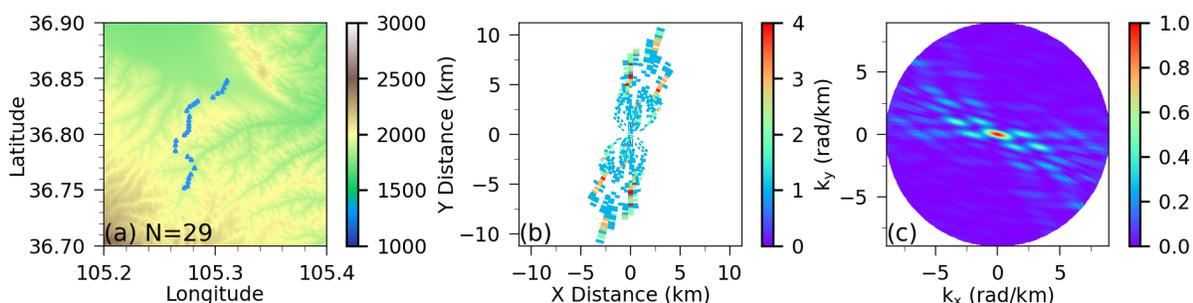
782  
 783 **Figure 15.** The dispersion image obtained by different beamforming schemes. (a) The result of WCBF1. (b) The  
 784 result of WCBF3. (c) The result of MCBF4. (d) The result of MCBF4+MCBF6.

785 Figure 15 shows the dispersion images obtained by different beamforming schemes. The dispersion  
 786 curves of the first four modes (modes 0-3) can be observed clearly from the results of WCBF and  
 787 MCBF. The difference between WCBF and MCBF is slight since the model is lateral isotropic. As we  
 788 observed in Figure 8b for the data from the field data, significant aliasing is not observed in the  
 789 result of MCBF4 in Figure 15c. Even the tiny interferences can be observed at frequencies higher  
 790 than 15 Hz for the result by antialiasing strategy MCBF4+MCBF6, as shown in Figure 15d. As

791 discussed before, we attribute this to the random distribution of stations in the array. Unlike the  
 792 regular sampling, aliasing is not coherent for random sampling. As a result, the aliasing is reduced  
 793 even for the result of MCBF4 for the array with station distributed randomly. As a verification, in  
 794 Figure S4 of the supporting information, we present an example of the synthetic data for the array  
 795 with stations regularly distributed. We find that for the array with stations regularly distributed, the  
 796 manifest aliasing appears for the results of MCBF4, while antialiasing strategy MCBF4+MCBF6 can  
 797 effectively eliminate the aliasing (See Figure S4 in the supporting information).

## 798 6.2. More Examples at Different Scales

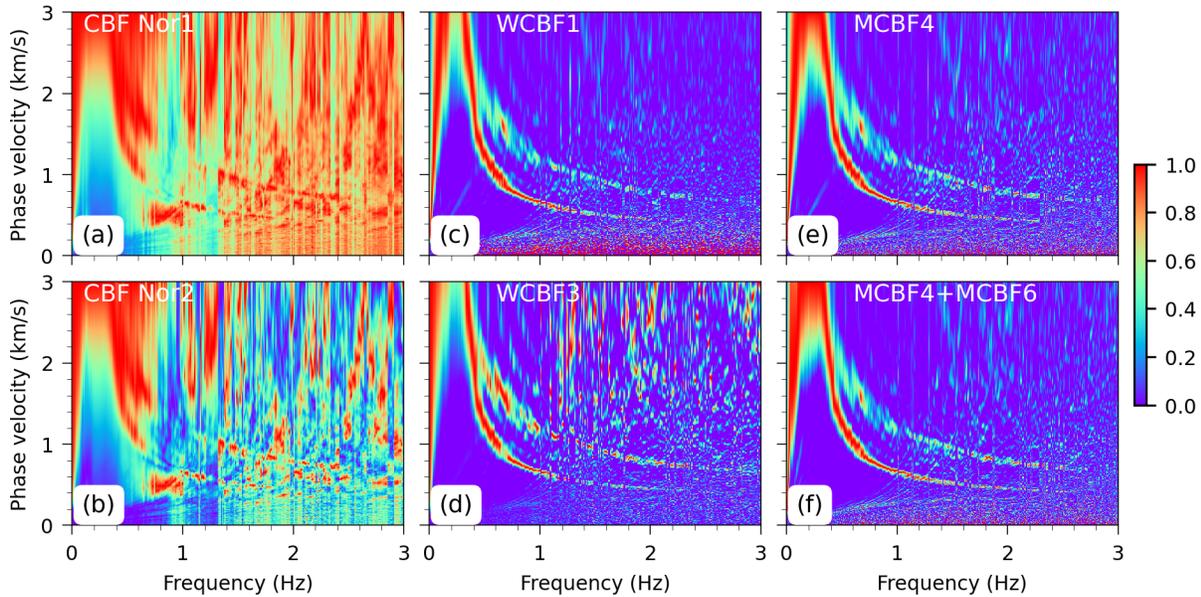
### 799 6.2.1. Example for a Nearly Linear Array in Haiyuan



800  
 801 **Figure 16.** (a) The station distribution of the subset (twenty-nine stations) from Haiyuan array. (b) The distribution  
 802 of the number of station pairs as a function of interstation spacing and azimuth. (c) The ARF of the array in (a).

803 The extraction of multimode Rayleigh waves using WCBF and MCBF are illustrated above based on  
 804 the data from Tongzhou array which is located in the east of Beijing, and mainly on the north China  
 805 plain (Qin et al, 2022). An example using the data from the other array located on the northwest  
 806 China is investigated in this section. The array consists of more than 600 stations with a spacing of  
 807 300m–2km, and observations were conducted around the Haiyuan fault from October 2020 to  
 808 December 2020. The Haiyuan fault is a main active fault in northwest China, and the 1920 M8.0  
 809 earthquake occurred on this fault. The array is designed to investigate the fine structure around  
 810 the fault. The subset stations in the array as shown in Figure 16a were chosen to illustrate the  
 811 extraction of multimode surface waves by WCBF and MCBF. The stations of this subarray are  
 812 roughly linearly distributed. The minimum and maximum interstation distance are  $r_{\min} = 0.27\text{km}$   
 813 and  $r_{\max} = 11.22\text{km}$ , respectively. Figure 16b shows the distribution of the number of station pairs

814 as a function of the interstation spacing and azimuth. Figure 16c shows the ARF of the subarray in  
 815 Figure 16a.



816  
 817 **Figure 17.** The dispersion image obtained by different imaging conditions for the array shown in Figure 16a. (a)  
 818 and (b) are the results of CBF with the same normalization strategies as those used in Figures 1b and 1d,  
 819 respectively. (c) and (d) are the results of WCBF1 and WCBF2, respectively. (e) and (f) are the results of MCBF4  
 820 and MCBF4+MCBF6. Only the positive values are color-coded in (c)-(f).

821 Figure 17 shows the dispersion image obtained by different imaging conditions of beamforming  
 822 for the array consisting of twenty-nine stations shown in Figure 16a. Figure 17a and 17b show the  
 823 results of CBF with the same normalization strategies as those used in Figure 1b and 1d,  
 824 respectively. Figure 17c and 17d are the results of WCBF, where the result of WCBF3 are equivalent  
 825 to that obtained by F-J method. The results of MCBF are given in Figures 17e and 17f. It can be  
 826 found that in order to obtain the dispersion image with the same clarity, larger number of stations  
 827 may be required for conventional CBF. Correction for wavenumber and propagation distance can  
 828 significantly improve the clarity of the dispersion image. Again, consistent with the previous  
 829 discussion, due to the irregular station distribution, the result of MCBF4 is similar to that of the  
 830 designed anti-aliasing scheme MCBF4+MCBF6, and no obvious aliasing is observed.

### 831 6.2.2. Examples for the USArray and ChinArray at Regional Scale

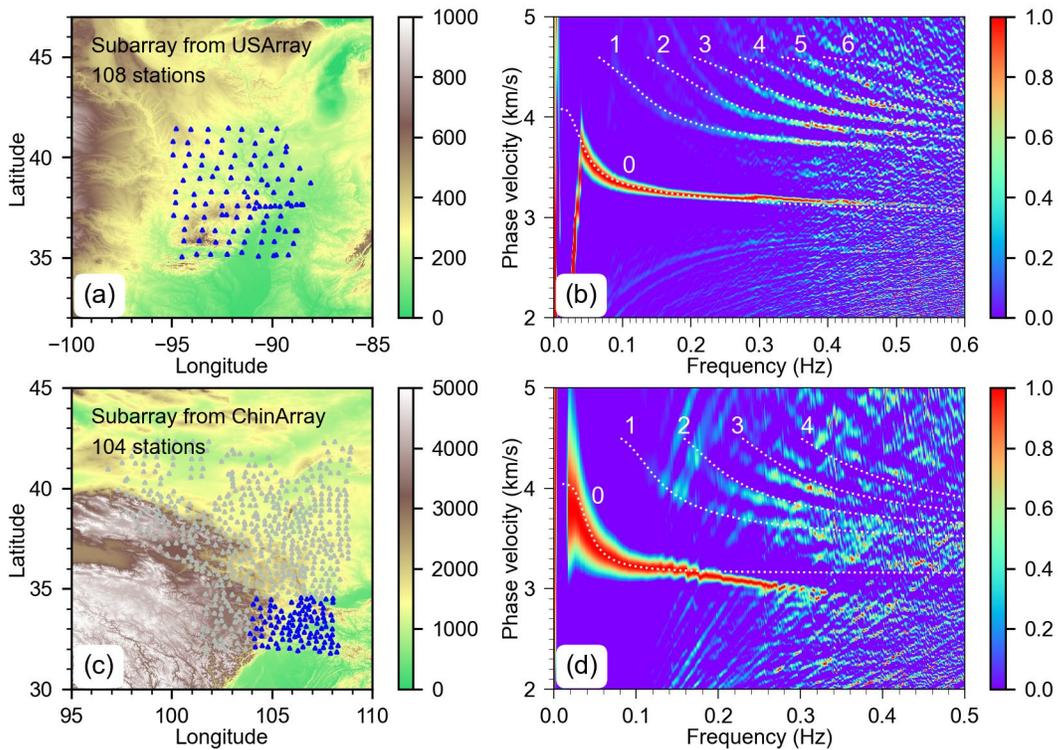
832 The average station spacing of the above two example arrays shown in Figure 1a and Figure 16a is

833 about 1 km. They are subsets sampled from two dense arrays at local scale. The effective frequency  
834 range of the extracted multimode dispersion curves is about 0.3-3 Hz. Two examples for the  
835 subarray at regional scale are presented in this section.

836 One array is the subset sampled from USArray. Figure 18a shows the station map of this array  
837 which contains 108 stations with an average station spacing of 70 Km. A total ninety days  
838 continuous waveform recorded from 1 June 2011 to 31 August 2011 are used. Figure 18b shows  
839 the dispersion image obtained by MCBF. The fundamental (mode 0) and five higher modes (modes  
840 1-5) are clearly observed. The effective frequency range is about 0.03-0.6 Hz, depends on the mode  
841 branches. In such a relatively wide frequency band, different modes correspond to varying  
842 wavelengths and propagation velocities, as a result, the modes usually cannot be separated  
843 apparently in the time domain for the interstation NCFs (See [Figure S5 in the supporting](#)  
844 [information for the interstation NCFs within the array](#)). This means the mode separation in the  
845 time domain is not a requirement for extracting multimode dispersion curves using MCBF. As a  
846 comparison, the theoretical dispersion curves predicted by the reference model are presented by  
847 dashed lines. The reference model is obtained by averaging the model under the array given by  
848 [Shen and Ritzwoller \(2016\)](#).

849 Another array is the subset sampled from ChinArray (Phase II). As shown in Figure 18c, this array  
850 contains 104 stations with an average station interval of 30km. [K. Wang et al. \(2020\)](#) have extracted  
851 the fundamental mode Rayleigh wave between 7 and 35 s using the conventional CBF. We applied  
852 the NCFs, which are the stacking results of three months continuous records (from 1 January 2014  
853 to 31 March 2014), to investigate the extraction of higher modes by MCBF. The resulted dispersion  
854 image is shown in Figure 18d. The white dashed lines denote the dispersion curves predicted by  
855 the averaged model under the array given by [K. Wang et al. \(2020\)](#). For the fundamental mode  
856 Rayleigh wave, the predicted dispersion curve agrees well with the observed one in the frequency  
857 range of 7 -35 s. This is not surprised since the model given by [K. Wang et al. \(2020\)](#) is derived from  
858 the inversion of the fundamental mode Rayleigh dispersion curve at this frequency rang. However,  
859 for the fundamental mode Rayleigh wave at frequencies higher than 0.14 Hz (7 s) and the higher  
860 modes, the predicted dispersion curve deviated from the observed one. This phenomenon can also  
861 be observed for the USArray shown in Figure 18b, where the predicted dispersion curves for the

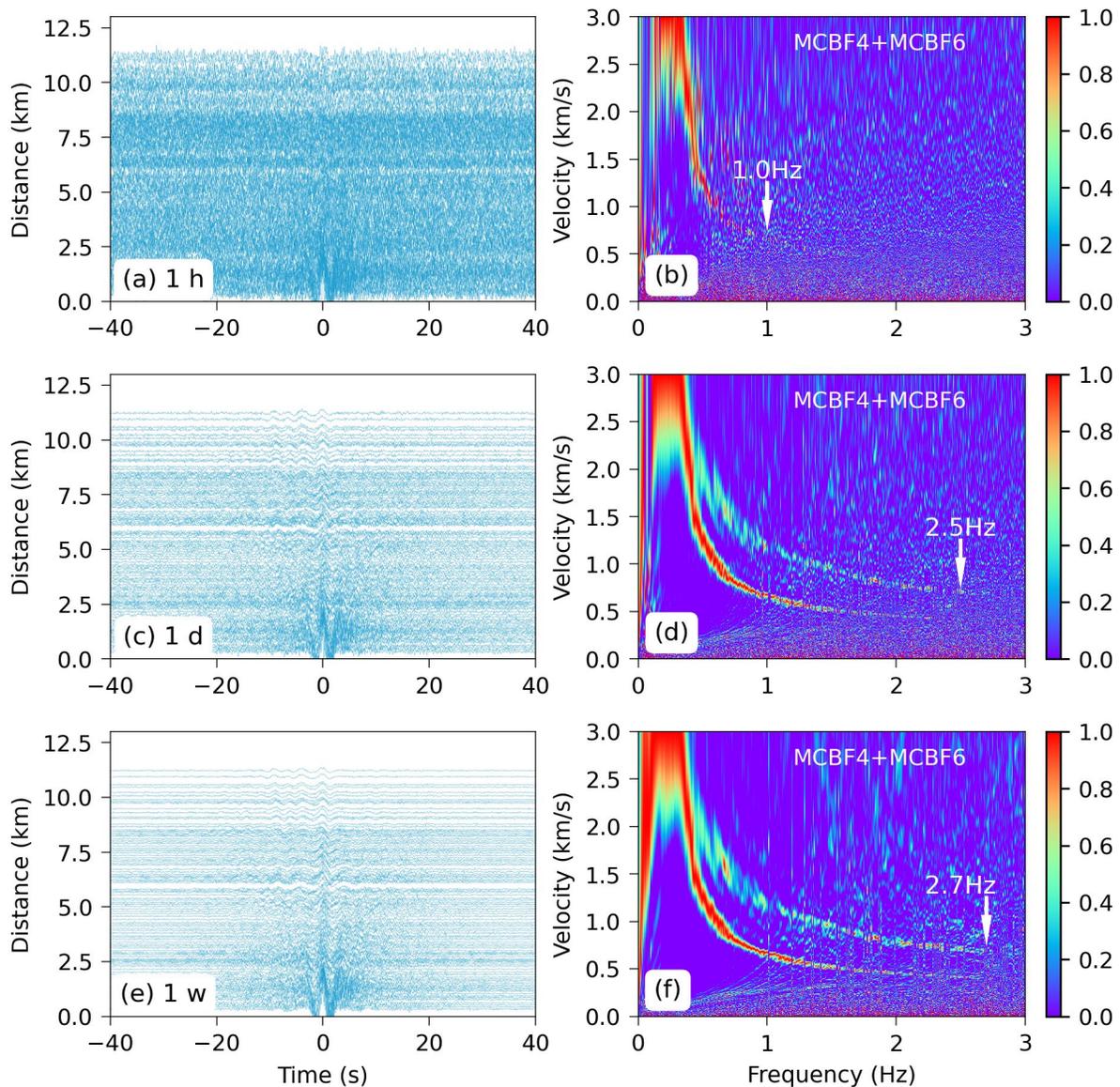
862 higher modes and the fundamental mode with frequencies above 0.3 Hz also deviate from those  
 863 observed. We inverted for a new model using extracted multimode dispersion curves. The  
 864 predicted dispersion curves using the new inverted model agree well with the observed one (See  
 865 [Figure S6 in the supporting information for the inverted model using multimode surface wave and](#)  
 866 [the corresponding predicted dispersion curves](#)). This gives an illustration on the importance of high  
 867 modes in surface wave inversion.  
 868



869  
 870 **Figure 18.** Two examples on the extraction of multimode dispersion curves using MCBF for the array at regional  
 871 scale, which are the subsets of the USArray (a) and ChinArray (c), respectively. The stations inside the array used  
 872 in this paper are presented by blue triangles in left panels. The corresponding dispersion images obtained by  
 873 MCBF are shown in right panels. The white dashed lines in (b) are the predicted dispersion curves for the  
 874 averaged model from [Shen and Ritzwoller \(2016\)](#). In (d), the white dashed lines are the predicted dispersion  
 875 curves for the averaged model from [K.Wang et al. \(2020\)](#).

876 **6.3. The Influence of the Parameters on the Results of MCBF**

877 **6.3.1. The Time Length of the Recordings**



878  
879 **Figure 19.** Bandpass (0.2-5 Hz) filtered NCFs for the interstation within the array shown in Figure 16a, which are  
880 retrieved from recording length of 1 hour (a), 1 day (c) and 1 week (e), respectively. The corresponding dispersion  
881 images obtained by MCBF are presented in right panels.

882 Besides the ability to extract multimode surface wave dispersion curves, another advantage of BF  
883 over traditional NCF-based methods is that phase velocities can be estimated accurately using  
884 quite shorter recordings. For the array data shown in Figure 1a, [Qin et al. \(2022\)](#) found that

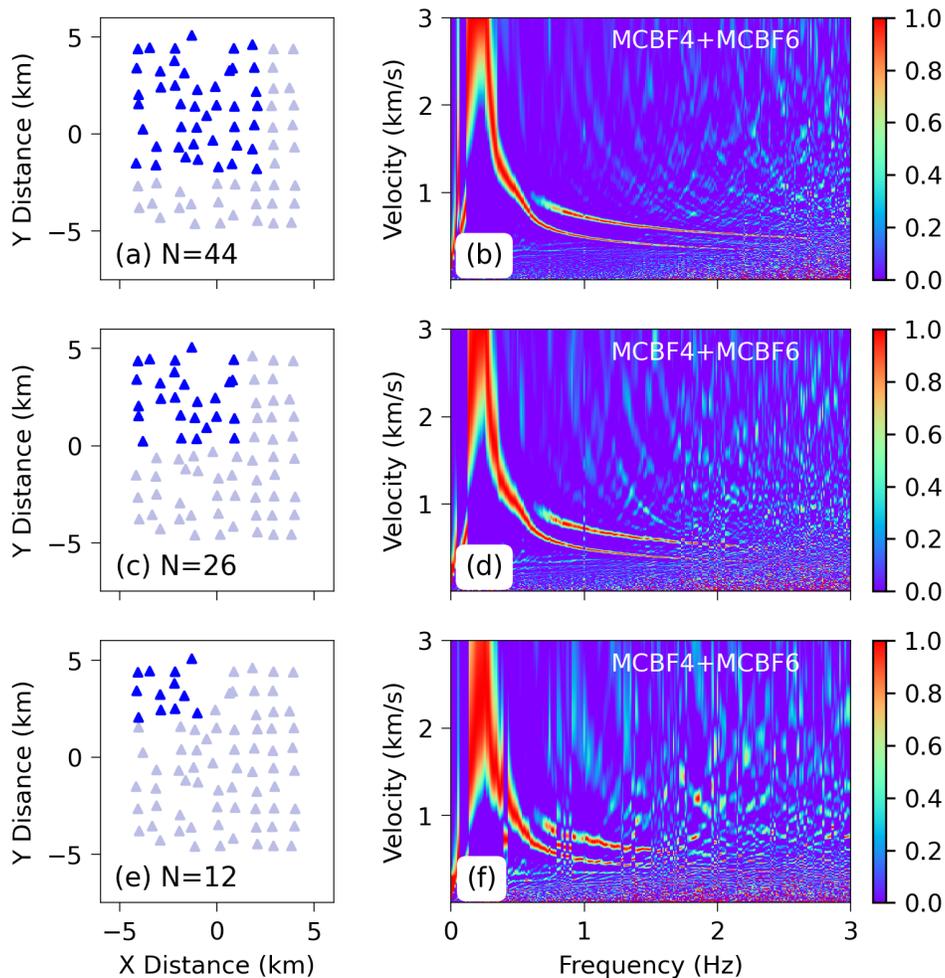
885 recordings as short as a few hours to days are sufficient to extract reliable dispersion curves of the  
886 fundamental and the first higher mode between 0.3 and 1.5 Hz by CBF. This is consistent with the  
887 conclusion of Roux and Ben-Zion (2017), who estimated reliable phase velocities for the  
888 fundamental mode Rayleigh waves using the recording with similar length from California network.  
889 The same conclusion can also be obtained for MCBF for the array shown in Figure 1a (See Figures  
890 S7 and S8 in the supporting information for the corresponding results of MCBF for the array shown  
891 in Figure 1a). Furthermore, for the same recording length, the energy of the dispersion image given  
892 by MCBF is more concentrated on the dispersion branches than that given by CBF, resulting in a  
893 clean dispersion image with few artifacts (e.g., seen by comparing Figure S7 in the supporting  
894 information with Figure 7 in Qin et al. (2021)).

895 For the array shown in Figure 16a, we calculated the NCFs by stacking the different recording  
896 length of 1 hour, 1 day and 1 week, respectively. The bandpass (0.2-5 Hz) filtered NCFs retrieved  
897 from different recording lengths and corresponding dispersion image obtained by MCBF are  
898 presented in Figure 19. As expected, the longer the recording length, the higher the signal-to-noise  
899 ratio (SNR) of the NCFs and, correspondingly, the wider the frequency band of the dispersion  
900 curves resolved by MCBF. Also, the dispersion image looks cleaner for long time recordings due to  
901 fewer artifacts. However, the width of the energy band along two dispersion branches, which is  
902 often an indicator on the precision of the velocity measurement, does not narrow significantly as  
903 the recording length increases. This implies the recording length is not the main factor affecting  
904 the precision of velocity estimation, and it mainly control the resolved frequency band and the  
905 artifacts occurred in the image.

### 906 **6.3.2. The Number of Stations**

907 As discussed in section 2.3, for the time series with the same length, the resolved frequency band  
908 and the resolution of the dispersion measurement by BF depend on the wavefield beneath the  
909 array and ARF. The wavefield mainly depends on the complexity of the velocity structure under the  
910 array, which affects the emergence of interstation NCFs and thus the quality of the dispersion  
911 image. For real arrays with known station distribution, our main concern is the effect of ARF, which  
912 is determined by the array configuration, and thus by the array size, shape, interstation distance

913 and the number of stations inside the array. For the measurements of the azimuth-averaged phase  
 914 velocity, or in other words, assuming the structure beneath the array does not vary with azimuth,  
 915 the number of stations is a major factor affecting the measurement accuracy. This is because the  
 916 algorithm in MCBF can be thought as a discrete summation of an integral with Fourier transformed  
 917 NCFs as the kernel. The accuracy of the integral is determined by the number of NCFs and the  
 918 intervals between them, which are mainly related to the number of stations (rather than array  
 919 shape) for a given array configuration.



920  
 921 **Figure 20.** The dispersion images obtained by MCBF using the subsets of the stations inside the array shown in  
 922 Figure 1a. The size of the subarrays represented by blue triangles in left panels are  $7 \times 7 \text{ km}^2$ ,  $5 \times 5 \text{ km}^2$  and  $3 \times 3$   
 923  $\text{km}^2$ . The number of stations are 44, 26 and 12, respectively. The corresponding dispersion image for these three  
 924 subarrays are presented in right panels. Only the positive value is color-coded in the dispersion images (See  
 925 [Figure S9 in the supporting information where both the positive and negative values are color-coded](#)).

926 The number of stations mainly depends on the size of the array if the station distribution is given.  
927 Therefore, in order to ensure the accuracy of reliable dispersion measurements and sufficient  
928 lateral resolution, a trade-off between the array size and the number of stations within the array  
929 needs to be made when applying BF technique with moving subarrays. For the array data shown  
930 in Figure 1a, [Qin et al. \(2022\)](#) investigated the influence of the number of stations on the estimation  
931 of dispersion curves using CBF through tentative experiments and to determine the appropriate  
932 size of the array. Using the same dataset, we analyze the effect of the station number on the results  
933 of MCBF.

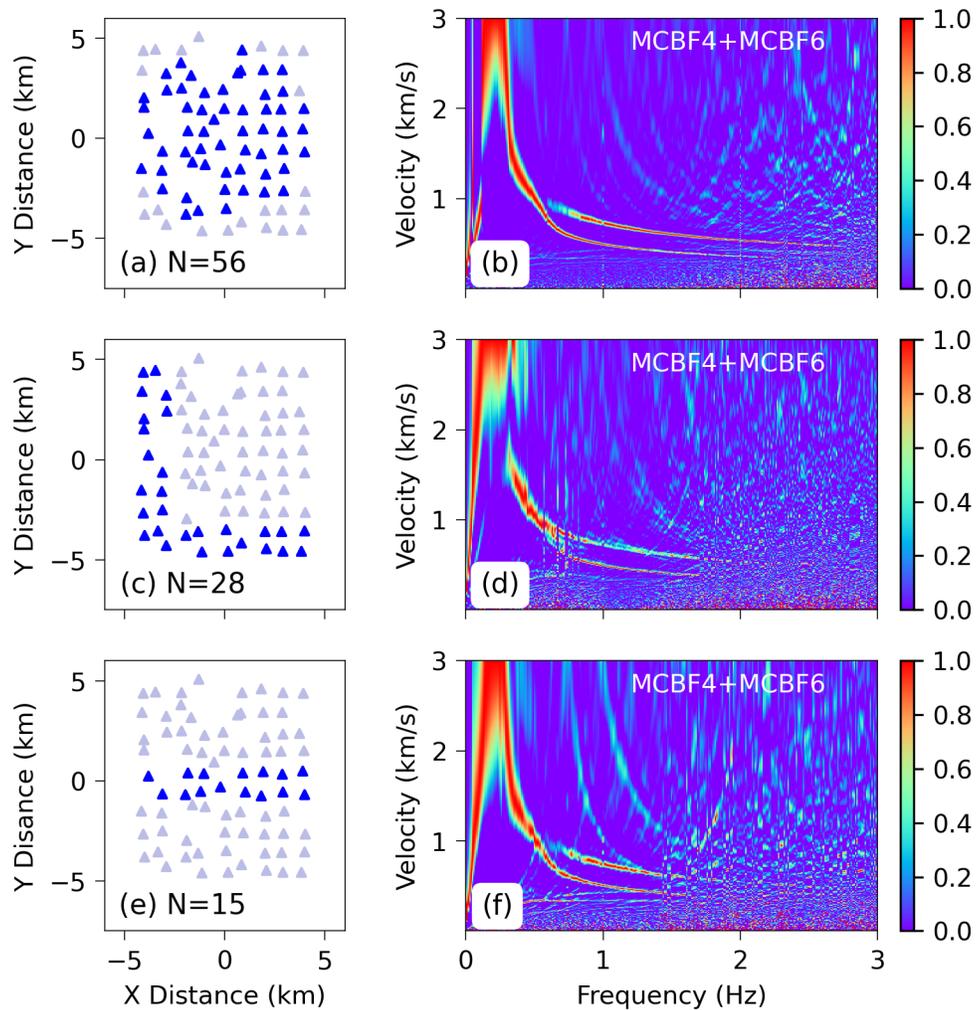
934 The array aperture in Figure 1a is about  $10 \times 10 \text{ km}^2$ . To reduce the array size, we take the subset  
935 of the stations from the upper left corner of the array to get three square subarrays with sizes of  
936  $7 \times 7 \text{ km}^2$ ,  $5 \times 5 \text{ km}^2$  and  $3 \times 3 \text{ km}^2$ . The station distributions of these subarrays are presented by blue  
937 triangles in left panels of Figure 20. The corresponding dispersion images obtained by MCBF are  
938 given in right panels of Figure 20. The number of stations of these three subarrays are 44, 26 and  
939 12, respectively.

940 It can be seen from top to bottom panels in the right column of Figure 20, as the station number  
941 decreases, the reliable frequency band reduces and the energy bands along the two dispersion  
942 branches widen. This implies the reduction in the array size and thus the number of stations  
943 resulted in a narrowing of the resolved frequency range and reduced measurement accuracy,  
944 similar to the observation in [Qin et al. \(2022\)](#) for the results of CBF. However, compared with Figure  
945 6 in [Qin et al. \(2022\)](#), the dispersion image given by MCBF in Figure 20 is much clearer and less  
946 interference from aliasing.

### 947 **6.3.3. The Geometry of the Array**

948 As a function of the azimuth, the beamforming output depends on the shape of the array,  
949 especially the azimuth distribution of the dominant station-pair orientation would seriously affect  
950 the azimuth distribution of the beampower. However, for the azimuth-averaged velocity, the effect  
951 of the array shape on the dispersion image is slight. Figure 21 shows the results of MCBF for three  
952 subsets with different shape sampled from the array shown Figure 1a. The blue triangles in the left  
953 panels of Figure 21 denote the station map of three subsets, which are approximately circular (a),

954 L-shaped (c) and linear (e), respectively. The corresponding result of MCBF for each subset is given  
 955 in the panels of the right column. The energy belts along two dispersion branches can be resolved  
 956 for these three subarrays. Although the reliable frequency range and resolution are slightly  
 957 different. Figures 21b and 21f show that the phase velocities along two mode branches are quite  
 958 similar to those obtained from square arrays in Figures 20b and 20f.



959  
 960 **Figure 21.** Dispersion images obtained by MCBF using the subsets with different shapes of the array in Figure 1a.  
 961 The blue triangles in the left panels denote stations of within three subsets, which are approximately circular (a),  
 962 L-shaped (c) and linear (e), respectively. The station numbers N are labeled in each panel. The corresponding  
 963 result of MCBF for each subset is given in the panels of the right column. Only the positive value is color-coded  
 964 in the dispersion images (See Figure S10 in the supporting information where both the positive and negative  
 965 values are color-coded).

966 Figure 21d shows that the dispersion curve of the fundamental mode intersects that of the first  
967 higher mode at 0.6 Hz for the result of the L-shaped array, slightly different from the results of the  
968 square array shown in Figure 20. This may be due to the fact that the stations involved in the L-  
969 shaped array covers different tectonic units (Qin et al., 2022). The assumption that the model  
970 under the array is laterally isotropic does not hold any more. Generally speaking, for the extraction  
971 of azimuth-averaged velocity, the effect of the array shape on the dispersion curve is much smaller  
972 than that of the station numbers. This conclusion holds for WCBF and CBF, as well as MCBF. As for  
973 the extraction of the azimuth-dependence velocity, the shape of the array would affect the  
974 beamforming output and thus the velocity estimation for WCBF and CBF. However, the array shape  
975 does not significantly affect the results of MCBF, even for the estimation on the azimuth-  
976 dependence velocity.

## 977 **7. Discussion and Conclusions**

978 The deployment of dense array at different scale is becoming a routine operation in observational  
979 seismology. It is therefore becoming possible to directly obtain the lateral variation of the velocity  
980 under the array by applying the array-based surface wave method to the subset of the dense array  
981 through the moving window technology. As two array-based techniques for extracting multimode  
982 surface wave from noise recording, BF and F-J methods have been developed independently,  
983 based on the physical interpretations from different point of view. In this paper, the equivalence  
984 of BF and F-J methods are proved as far as their application in extracting multimode dispersion  
985 curves is concerned. The weighted (WCBF) and modified (MCBF) BF methods are proposed and  
986 their explicit relationship to the F-J method is given. For the finite sampling on spatial wavefield in  
987 practical applications, explicit theoretical representations of the BF technique are given for  
988 different imaging conditions. These representations can be used to investigate analytically the  
989 features of the dispersion image and how the aliasing is eliminated. The proposed methods are  
990 validated both for the synthetic data and the real data from the dense array at different scales. In  
991 summary, the main conclusions of this paper are as follows:

992 (1) For the conventional CBF, the plane wave at a given azimuth is projected into the station pairs  
993 distributed at different orientation. This operation on azimuthal average physically means that

994 the structure beneath the array are laterally uniform. Mathematically, azimuth-average would  
995 express the plane waves represented by the complex exponential function  $e^{ikr \cos \theta}$  as the Bessel  
996 function  $J_0(kr)$ , which results in an additional geometric spread factor of  $1/\sqrt{kr}$  and  $-\pi/4$   
997 phase shift (i.e., the difference of the far field approximation of  $J_0(kr)$  and cosine function  
998  $\cos(kr)$ ).

999 (2) Considering the fact mentioned in (1), WCBF is proposed to correct the geometric spread factor  
1000 (Since the NCF also has  $-\pi/4$  phase shift relative to the plane wave, the phase correction is  
1001 not needed in WCBF). By comparing the operation on the projection of plane waves in WCBF  
1002 and the integration with kernel of  $J_0(kr)$  in F-J, it is found the WCBF is equivalent to F-J,  
1003 differing by a factor  $1/k$ . This means the assumption that the structure under the array is  
1004 laterally homogeneous is also made in F-J method. The F-J transform is theoretically 2D Fourier  
1005 transform of the radially symmetric function. The velocity obtained by WCBF and F-J is the  
1006 azimuth-averaged result. The subtle differences in the azimuth-averaged velocities obtained  
1007 by BF and F-J may originate from the strategy for numerical calculation. If WCBF and F-J are  
1008 used to extract the azimuth-dependence velocity, for instance, by picking the azimuth-  
1009 dependence maximum of the beampower in WCBF or by considering only the NCFs at expected  
1010 azimuth in F-J, the errors caused by the array geometry is inevitable.

1011 (3) Since both WCBF and F-J assume that there is no lateral variation in the velocity structure under  
1012 the array, we proposed a MCBF scheme in which the projection of the plane waves into the  
1013 station pairs is omitted. For a given station pair, only the incident plane wave that is consistent  
1014 with its orientation is considered in MCBF. For the azimuth-averaged velocity, the dispersion  
1015 image with the same even high resolution as that given by WCBF and F-J can be obtained by  
1016 MCBF. Furthermore, in contrast to WCBF and F-J, azimuth-dependence velocity can also be  
1017 measured by MCBF independent of the array geometry theoretically. The explicit relationship  
1018 between the azimuth-averaged MCBF and F-J is given.

1019 (4) For an ideal cylindrical wavefield, the theoretical representations of MCBF for different imaging  
1020 conditions are derived for the finite spatial sampling. The beamforming output of the MCBF for

1021 the imaging conditions with sine and cosine functions as the basis function can be expressed  
1022 as the summation of the products between trigonometric and sinc functions with the  
1023 arguments  $k_n - k$  and  $k_n + k$ . The characteristic wavenumbers  $k_n$  and aliasing wavenumbers  
1024  $-k_n$  are associated with the maximum, or the zero-crossing points between two peaks of the  
1025 beampower, depending on the choice of the basis function. The beamforming output contains  
1026 the items only with the argument  $k_n - k$  or  $k_n + k$  can be achieved by the linear combination  
1027 of different imaging condition. The aliasing wavenumbers can thus be separated and removed.

1028 (5) The validity of the proposed methods is illustrated by synthetic data and the data of the dense  
1029 array at different scales. For the azimuth-averaged velocity, the results are less affected by the  
1030 array geometry, both for WCBF and MCBF. The resolution and reliable frequency range of the  
1031 dispersion image are controlled mainly by the number of stations, which depends on the array  
1032 size and interstation spacing. In addition, the dispersion curves can be estimated with adequate  
1033 accuracy using quite short recordings, for instance over a few hours to days. This offers the  
1034 possibility on the rapid assessment of the medium properties.

### 1035 **Data Availability Statement**

1036 The waveform data of USArray were obtained from the website ([https://ds.iris.edu/ds/nodes/dmc](https://ds.iris.edu/ds/nodes/dmc/data/types/waveform-data)  
1037 [/data/types/waveform-data](https://ds.iris.edu/ds/nodes/dmc/data/types/waveform-data)). The NCFs of the ChinArray (Phase II) for the stations used in the  
1038 paper can be found at <https://doi.org/10.6084/m9.figshare.20304618.v1>. The NCFs of two dense  
1039 arrays at local scale used as the examples in this paper, as well as the Python package for  
1040 generating the dispersion image using different BF schemes based on these NCFs, are also openly  
1041 available from <https://doi.org/10.6084/m9.figshare.20304618.v1>.

### 1042 **Acknowledgments**

1043 This work was supported by the National Natural Science Foundation of China (U1839209) and the  
1044 National Key R & D Program of China (2017YFC1500200).

### 1045 **Appendix A: Convention for the Fourier transform, Hilbert transform and cross-correlation**

1046 The following convention for Fourier transform from the time ( $t$ ) domain to the frequency ( $\omega$ )  
1047 domain is adopted in this paper.

1048

$$f(t) = \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega \quad (A1)$$

$$F(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

1049 Accordingly, for the Fourier transform from the space ( $r$ ) domain to the wavenumber domain ( $k$ )

1050 we use the convention

1051

$$f(r) = \int_{-\infty}^{\infty} F(k) e^{-ikr} dk \quad (A2)$$

$$F(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(r) e^{ikr} dr$$

1052  $f(t)$  and  $F(\omega)$ , as well as  $f(r)$  and  $F(k)$  are two Fourier transform pairs. These conventions

1053 imply the plane wave expression  $e^{-ikx} e^{i\omega t}$  describes the wave propagating in the positive  $x$ , while

1054  $e^{+ikx} e^{i\omega t}$  describes the wave propagating in the negative  $x$ . For the cylindrical wave in the Cartesian

1055 coordinate system with  $z$  downwards, the zero-order Hankel function  $H_0^{(1)}(kr)$  of the first kind

1056 represents the converging wave and the zero-order Hankel function  $H_0^{(2)}(kr)$  of the second kind

1057 represents the wave propagating outwards.

1058 The Hilbert transform of the function  $s(x)$  is defined as

1059

$$s(t) = \mathcal{H} [s(x)] = \frac{1}{\pi} \text{P.V.} \int_{-\infty}^{+\infty} \frac{s(x)}{t-x} d\tau \quad (A3)$$

1060 where P.V. represents the principal value of the Cauchy integral. Under this definition, the

1061 analytical signal  $S(t)$  of the real-value signal  $s(t)$  can be expressed as

1062

$$S(t) = s(t) + i\mathcal{H} [s(t)] \quad (A4)$$

1063 The symbol  $\mathcal{H}$  represents the Hilbert transform.

1064 The cross-correlation of two (complex) signals is defined as

1065

$$C_{fg}(\tau) = \int_{-\infty}^{+\infty} f^*(t) g(t+\tau) dt \quad (A5)$$

1066 where superscript  $*$  represents complex conjugate. If the Fourier transform of  $g(t)$  is  $G(\omega)$ .

1067 According to this definition and the above convention for Fourier transform, we have

1068

$$C_{fg}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F^*(\omega) G(\omega) e^{i\omega\tau} d\omega \quad (A6)$$

1069 i.e.  $C_{fg}(\tau)$  and  $F^*(\omega)G(\omega)$  are Fourier transform pair. Cross-correlation in the time domain  
 1070 corresponds to the product in the frequency domain by taking complex conjugate of one of them.

### 1071 **Appendix B: Theoretical Representation of MCBF**

1072 For an ideal cylindrical wavefield, substituting  $C(r, \omega) = J_0(k_n r)$  into Equation 30. The azimuth-  
 1073 averaged MCBF can be written as

$$\begin{aligned}
 \overline{MCBF}(k, \omega) &= \frac{1}{N^2} \sum_i^N \sum_j^N \sqrt{\pi k r_{ij}} J_0(k_n r_{ij}) e^{i\left(k r_{ij} - \frac{\pi}{4}\right)} \\
 &= \frac{\sqrt{2}}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left( \sum_i^N \sum_j^N \cos\left(k_n r_{ij} - \frac{\pi}{4}\right) e^{i\left(k r_{ij} - \frac{\pi}{4}\right)} \right)
 \end{aligned}
 \tag{B1}$$

1075 In Equation B1, the far field approximation of  $J_0(k_n r)$  shown in Equation 20 is applied. The double  
 1076 summation over the number of stations means the summation over the interstation distance  $r_{ij}$ .  
 1077 Since the geometric spread is corrected by multiplying  $r_{ij}$  in Equation B1, the autocorrelation for  
 1078  $i = j$  is excluded. The cross-correlations for  $i \neq j$  are counted twice for the same interstation  
 1079 distance  $r_{ij} = r_{ji}$ . Equation B1 can then be written as

$$\begin{aligned}
 \overline{MCBF}(k, \omega) &= \frac{2\sqrt{2}}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \sum_j^M \cos(k_n r_j - \frac{\pi}{4}) e^{i(k r_j - \frac{\pi}{4})} \\
 &= \frac{2\sqrt{2}}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \sum_j^M \cos(k_n r_j - \frac{\pi}{4}) \left( \cos(k r_j - \frac{\pi}{4}) + i \sin(k r_j - \frac{\pi}{4}) \right)
 \end{aligned}
 \tag{B2}$$

1081 where  $M = N(N-1)/2$ . Let

$$CC = \sum_j^M \cos(k_n r_j - \frac{\pi}{4}) \cos(k r_j - \frac{\pi}{4}) = \frac{1}{2} \sum_j^M \left[ \cos(k_n r_j - k r_j) + \sin(k_n r_j + k r_j) \right]
 \tag{B3}$$

1083 It is assumed the cylindrical wavefield is sampled spatially in radial distance with equal interval  $\Delta r$ .

1084 We have  $r_j = j\Delta r$ . Equation B3 can be rewritten as

$$CC = \frac{1}{2} \sum_j^M \left[ \cos(k_n - k) j \Delta r + \sin(k_n + k) j \Delta r \right]
 \tag{B4}$$

1086 Applying the Equations 1.342.1 and 1.342.2 in [Gradshteyn and Ryzhik \(2007,P37\)](#)

1087 
$$\sum_{k=0}^n \cos kx = \cos \frac{n+1}{2} x \frac{\sin \frac{nx}{2}}{\sin \frac{x}{2}} + 1$$

1088 
$$\sum_{k=1}^n \sin kx = \sin \frac{n+1}{2} x \frac{\sin \frac{nx}{2}}{\sin \frac{x}{2}}$$

(B5)

1088 Equation B4 can be recast into

1089 
$$CC = \frac{1}{2} \left[ \cos \frac{(M+1)(k_n-k)\Delta r}{2} \frac{\sin \frac{M(k_n-k)\Delta r}{2}}{\sin \frac{(k_n-k)\Delta r}{2}} + \sin \frac{(M+1)(k_n+k)\Delta r}{2} \frac{\sin \frac{M(k_n+k)\Delta r}{2}}{\sin \frac{(k_n+k)\Delta r}{2}} \right]$$

1090 
$$= \frac{M}{2} \left[ \cos \frac{(M+1)(k_n-k)\Delta r}{2} D_M \left( \frac{(k_n-k)\Delta r}{2} \right) + \sin \frac{(M+1)(k_n+k)\Delta r}{2} D_M \left( \frac{(k_n+k)\Delta r}{2} \right) \right]$$

(B6)

1090 where

1091 
$$D_M(x) = \frac{\sin \frac{Mx}{2}}{M \sin \frac{x}{2}}$$

(B7)

1092 is the Dirichlet sinc function. For a small  $\Delta r$  and large  $M$ , we have  $\sin \Delta r = \Delta r$ ,

1093  $R = M\Delta r \approx (M+1)\Delta r$ . Equation B6 can be approximately expressed as

1094 
$$CC \approx \frac{M}{2} \left[ \cos \frac{R(k_n-k)}{2} \text{sinc} \frac{R(k_n-k)}{2} + \sin \frac{R(k_n+k)}{2} \text{sinc} \left( \frac{R(k_n+k)}{2} \right) \right]$$

(B8)

1095 where  $\text{sinc}(x) = \sin x/x$  is the sinc function. Similarly, it can be deduced

1096 
$$CS = \sum_j^M \cos(k_n r_j - \frac{\pi}{4}) \sin(kr_j - \frac{\pi}{4}) = \frac{1}{2} \sum_j^M \left[ \cos(k_n r_j + kr_j) - \sin(k_n r_j - kr_j) \right]$$

1097 
$$= \frac{M}{2} \left[ \cos \frac{(M+1)(k_n+k)\Delta r}{2} D_M \left( \frac{(k_n+k)\Delta r}{2} \right) - \sin \frac{(M+1)(k_n-k)\Delta r}{2} D_M \left( \frac{(k_n-k)\Delta r}{2} \right) \right]$$

(B9)

1098 
$$\approx \frac{M}{2} \left[ \cos \frac{R(k_n+k)}{2} \text{sinc} \frac{R(k_n+k)}{2} - \sin \frac{R(k_n-k)}{2} \text{sinc} \left( \frac{R(k_n-k)}{2} \right) \right]$$

1097 
$$SS = \sum_j^M \sin(k_n r_j - \frac{\pi}{4}) \sin(kr_j - \frac{\pi}{4}) = \frac{1}{2} \sum_j^M \left[ \cos(k_n r_j - kr_j) - \sin(k_n r_j + kr_j) \right]$$

1098 
$$= \frac{M}{2} \left[ \cos \frac{(M+1)(k_n-k)\Delta r}{2} D_M \left( \frac{(k_n-k)\Delta r}{2} \right) - \sin \frac{(M+1)(k_n+k)\Delta r}{2} D_M \left( \frac{(k_n+k)\Delta r}{2} \right) \right]$$

(B10)

1099 
$$\approx \frac{M}{2} \left[ \cos \frac{R(k_n-k)}{2} \text{sinc} \frac{R(k_n-k)}{2} - \sin \frac{R(k_n+k)}{2} \text{sinc} \left( \frac{R(k_n+k)}{2} \right) \right]$$

1098 The modulus of  $CC + iCS$  can be approximated as

1099 
$$|CC + iCS| = \frac{\sqrt{2}}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left( \text{sinc} \frac{R(k_n+k)}{2} + \text{sinc} \frac{R(k_n-k)}{2} \right)$$

(B11)

1100 We therefore have

1101 
$$\overline{MCBF}(k, \omega) = \frac{\sqrt{2}}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} [CC + iCS] \quad (B12)$$

1102 The theoretical representations of the azimuth-averaged MCBF with different imaging conditions  
 1103 can then be written as

1104 
$$\begin{aligned} MCBF1 &= \frac{\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left( \cos \frac{R(k_n - k)}{2} \operatorname{sinc} \frac{R(k_n - k)}{2} + \sin \frac{R(k_n + k)}{2} \operatorname{sinc} \frac{R(k_n + k)}{2} \right) \\ MCBF2 &= \frac{\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left| \cos \frac{R(k_n - k)}{2} \operatorname{sinc} \frac{R(k_n - k)}{2} + \sin \frac{R(k_n + k)}{2} \operatorname{sinc} \frac{R(k_n + k)}{2} \right| \\ MCBF3 &= \frac{\sqrt{2}}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left( \operatorname{sinc} \frac{R(k_n + k)}{2} + \operatorname{sinc} \frac{R(k_n - k)}{2} \right) \\ MCBF4 &= MCBF1 \\ MCBF5 &= \frac{\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left( \cos \frac{R(k_n + k)}{2} \operatorname{sinc} \frac{R(k_n + k)}{2} - \sin \frac{R(k_n - k)}{2} \operatorname{sinc} \frac{R(k_n - k)}{2} \right) \\ MCBF6 &= \frac{\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \left( \cos \frac{R(k_n - k)}{2} \operatorname{sinc} \frac{R(k_n - k)}{2} - \sin \frac{R(k_n + k)}{2} \operatorname{sinc} \frac{R(k_n + k)}{2} \right) \\ MCBF4 + MCBF6 &= \frac{2\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \cos \frac{R(k_n - k)}{2} \operatorname{sinc} \frac{R(k_n - k)}{2} \\ MCBF4 - MCBF6 &= \frac{2\sqrt{2}M}{N^2} \frac{\sqrt{k}}{\sqrt{k_n}} \sin \frac{R(k_n + k)}{2} \operatorname{sinc} \frac{R(k_n + k)}{2} \end{aligned} \quad (B13)$$

1105 Theoretically, MCBF1 is equivalent to MCBF4. However, in practice, the cross spectrum  $C(r, \omega)$ ,  
 1106 i.e., the Fourier transform of NCFs is usually complex rather than the perfect Bessel function  $J_0(kr)$ .  
 1107 We therefore define MCBF1 and MCBF4 with  $C(r, \omega)$  and  $\operatorname{Re}[C(r, \omega)]$ , respectively.

## 1108 References

- 1109 Aki, K. (1957). Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors.  
 1110 *Bulletin of the Earthquake Research Institute*, 35, 415–456.
- 1111 Asten, M. W., & Henstridge, J. D. (1984). Array estimators and the use of microseisms for reconnaissance of  
 1112 sedimentary basins. *Geophysics*, 49(11), 1828–1837. <https://doi.org/10.1190/1.1442333>
- 1113 Baddour, N. (2011). Chapter 1 - Two-Dimensional Fourier Transforms in Polar Coordinates. In P. W. Hawkes (Ed.),  
 1114 *Advances in Imaging and Electron Physics* (Vol. 165, pp. 1–45). Elsevier. [https://doi.org/10.1016/B978-0-12-](https://doi.org/10.1016/B978-0-12-385861-0.00001-4)  
 1115 [385861-0.00001-4](https://doi.org/10.1016/B978-0-12-385861-0.00001-4)
- 1116 Ben-menahem, A., & Singh, S. J. (1968). Multipolar elastic fields in a layered half space. *Bulletin of the*  
 1117 *Seismological Society of America*, 58, 1519–1572.

- 1118 Campillo, M., & Paul, A. (2003). Long-Range Correlations in the Diffuse Seismic Coda. *Science*, 299(5606), 547–  
1119 549. <https://doi.org/10.1126/science.1078551>
- 1120 Capon, J. (1969). High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8),  
1121 1408–1418. <https://doi.org/10/ccpxt7>
- 1122 Chávez-García, F. J., & Luzón, F. (2005). On the correlation of seismic microtremors. *Journal of Geophysical*  
1123 *Research: Solid Earth*, 110(B11). <https://doi.org/10.1029/2005JB003671>
- 1124 Chen, Xiaofei, (1999). Seismogram Synthesis in Multi-layered Half-space Part I . Theoretical Formulations.  
1125 *Earthquake Research in China*, 0(2), 53–78.
- 1126 Forbriger, T. (2003). Inversion of shallow-seismic wavefields: I. Wavefield transformation. *Geophysical Journal*  
1127 *International*, 153(3), 719–734.
- 1128 Gal, M., Reading, A. M., & Ellingsen, S. P. (2019). Short Timescale Analysis of Microseisms and Application to  
1129 Array Calibration. *Journal of Geophysical Research: Solid Earth*, 124(3), 2684–2701.  
1130 <https://doi.org/10.1029/2018JB016959>
- 1131 Gerstoft, P., & Tanimoto, T. (2007). A year of microseisms in southern California. *Geophysical Research Letters*,  
1132 34(20). <https://doi.org/10/d2sts6>
- 1133 Gradshteyn, I. S. & Ryzhik, I. M. (2007). Table of Integrals, Series, and Products, Academic Press, Elsevier Inc.
- 1134 Harkrider, D. G. (1964). Surface waves in multilayered elastic media I. Rayleigh and Love waves from buried  
1135 sources in a multilayered elastic half-space. *Bulletin of the Seismological Society of America*, 54(2), 627–679.
- 1136 Harmon, N., Gerstoft, P., Rychert, C. A., Abers, G. A., Cruz, M. S. de la, & Fischer, K. M. (2008). Phase velocities  
1137 from seismic noise using beamforming and cross correlation in Costa Rica and Nicaragua. *Geophysical Research*  
1138 *Letters*, 35(19). <https://doi.org/10/cjgfrp>
- 1139 Horike, M. (1985). Inversion of Phase Velocity of Long-Period Microtremors to the S-Wave-Velocity Structure  
1140 down to the Basement in Urbanized Areas. *Journal of Physics of the Earth*, 33(2), 59–96.  
1141 <https://doi.org/10/fgpmkc>
- 1142 Hu, S., Luo, S., & Yao, H. (2020). The Frequency-Bessel Spectrograms of Multicomponent Cross-Correlation  
1143 Functions From Seismic Ambient Noise. *Journal of Geophysical Research: Solid Earth*, 125(8).  
1144 <https://doi.org/10.1029/2020JB019630>
- 1145 Lacoss, R. T., Kelly, E. J., & Toksöz, M. N. (1969). Estimation Of Seismic Noise Structure Using Arrays. *Geophysics*,  
1146 34(1), 21–38. <https://doi.org/10.1190/1.1439995>
- 1147 Lobkis, O. & Weaver, R. (2001). On the emergence of the Green's function in the correlations of a diffuse field. *J.*  
1148 *Acoust. Soc. Am.*, 110:3011–3017. <https://doi.org/10.1121/1.1417528>
- 1149 Löer, K., Riahi, N., & Saenger, E. H. (2018). Three-component ambient noise beamforming in the Parkfield area.  
1150 *Geophysical Journal International*, 213(3), 1478–1491. <https://doi.org/10.1093/gji/ggy058>
- 1151 Lu, L. (2021). Revisiting the cross-correlation and SPatial AutoCorrelation (SPAC) of the seismic ambient noise  
1152 based on the plane wave model. *Reviews of Geophysics and Planetary Physics*, 52(2): 123-163

- 1153 Lu, L., Wang, K., & Ding, Z. (2018). Effect of uneven noise source and/or station distribution on estimating the  
1154 azimuth anisotropy of surface waves. *Earthquake Science*, 31(04), 175–186.
- 1155 Morse, P. M. & Feshbach, H. (1953). *Methods of Theoretical Physics*, McGraw-Hill Book Comp., Inc., New York,  
1156 Toronto, London
- 1157 Picozzi, M., Parolai, S., & Bindi, D. (2010). Deblurring of frequency-wavenumber images from small-scale seismic  
1158 arrays. *Geophysical Journal International*, 181(1), 357–368. <https://doi.org/10.1111/j.1365-246X.2009.04471.x>
- 1159 Qin, T., Lu, L., Ding, Z., Feng, X., & Zhang, Y. (2022). High-Resolution 3D Shallow S Wave Velocity Structure of  
1160 Tongzhou, Subcenter of Beijing, Inferred From Multimode Rayleigh Waves by Beamforming Seismic Noise at a  
1161 Dense Array. *Journal of Geophysical Research: Solid Earth*, 127(5), e2021JB023689.  
1162 <https://doi.org/10.1029/2021JB023689>
- 1163 Riahi, N., & Saenger, E. H. (2014). Rayleigh and Love wave anisotropy in Southern California using seismic noise,  
1164 *Geophys. Res. Lett.*, 41, 363–369, doi:10.1002/2013GL058518.
- 1165 Rost, S., & Thomas, C. (2002). ARRAY SEISMOLOGY: METHODS AND APPLICATIONS. <https://doi.org/10/dt7j2t>
- 1166 Roux, P., & Ben-Zion, Y. (2017). Rayleigh phase velocities in Southern California from beamforming short-duration  
1167 ambient noise. *Geophysical Journal International*, 211(1), 450–454. <https://doi.org/10/gb3bkc>
- 1168 Ruigrok, E., Gibbons, S., & Wapenaar, K. (2017). Cross-correlation beamforming. *Journal of Seismology*, 21(3),  
1169 495–508. <https://doi.org/10.1007/s10950-016-9612-6>
- 1170 Seydoux, L., de Rosny, J. & Shapiro, N.M. (2017). Pre-processing ambient noise cross-correlations with equalizing  
1171 the covariance matrix eigenspectrum, *Geophys. J. Int.*, 210(3), 1432–1449
- 1172 Shen, W., and M. H. Ritzwoller (2016), Crustal and uppermost mantle structure beneath the United States, J.  
1173 *Geophys. Res. Solid Earth*, 121, 4306–4342, doi:10.1002/2016JB012887.
- 1174 Smith, M. L. & Dahlen, F. A. (1973). The azimuth dependence of love and rayleigh wave propagation in a slightly  
1175 anisotropic medium. *J. Geophys. Res.*, 78(17):3321-3333.
- 1176 Tokimatsu, K. (1997). Geotechnical site characterization using surface waves. In: Ishihara (ed). Proc. 1<sup>st</sup> Intl. Conf.  
1177 Earthquake eotechnical Engineering, Vol 3, Balkema, pp 1333–1368
- 1178 Tsai, V. C. & Moschetti, M. P. (2010). An explicit relationship between time-domain noise correlation and spatial  
1179 autocorrelation (SPAC) results. *Geophys.J.Int.*, 182(1): 454-460 .
- 1180 Wang, J., Wu, G., & Chen, X. (2019). Frequency-Bessel Transform Method for Effective Imaging of Higher-Mode  
1181 Rayleigh Dispersion Curves From Ambient Seismic Noise Data. *Journal of Geophysical Research: Solid Earth*,  
1182 124(4), 3708–3723. <https://doi.org/10/gmhv6v>
- 1183 Wang, K., Lu, L., Maupin, V., Ding, Z., Zheng, C., & Zhong, S. (2020). Surface Wave Tomography of Northeastern  
1184 Tibetan Plateau Using Beamforming of Seismic Noise at a Dense Array. *Journal of Geophysical Research: Solid*  
1185 *Earth*, 125(4), e2019JB018416. <https://doi.org/10/ghq5kq>
- 1186 Wathelet, M., Jongmans, D., Ohrnberger, M., & Bonnefoy-Claudet, S. (2008). Array performances for ambient  
1187 vibrations on a shallow structure and consequences over Vsinversion. *Journal of Seismology*, 12(1), 1–19.  
1188 <https://doi.org/10/bwd9nr>

- 1189 Woods, J. W., & Lintz, P. R. (1973). Plane waves at small arrays. *GEOPHYSICS*, 38(6), 1023–1041.  
1190 <https://doi.org/10/b58sts>
- 1191 Xi, C., Xia, J., Mi, B., Dai, T., Liu, Y., & Ning, L. (2021). Modified frequency–Bessel transform method for dispersion  
1192 imaging of Rayleigh waves from ambient seismic noise. *Geophysical Journal International*, 225(2), 1271–1280.  
1193 <https://doi.org/10/gh7c26>
- 1194 Xia, J., Miller, R. D., Park, C. B., & Tian, G. (2003). Inversion of high frequency surface waves with fundamental  
1195 and higher modes. *Journal of Applied Geophysics*, 52(1), 45–57. [https://doi.org/10.1016/S0926-9851\(02\)00239-](https://doi.org/10.1016/S0926-9851(02)00239-2)  
1196 2
- 1197 Yamaya, L., Mochizuki, K., Akuhara, T., & Nishida, K. (2021). Sedimentary Structure Derived From Multi-Mode  
1198 Ambient Noise Tomography With Dense OBS Network at the Japan Trench. *Journal of Geophysical Research: Solid*  
1199 *Earth*, 126(6). <https://doi.org/10.1029/2021JB021789>
- 1200 Yokoi, T. and Margaryan, S. (2008). Consistency of the spatial autocorrelation method with seismic  
1201 interferometry and its consequence. *Geophysical Prospecting*, 56: 435–451.
- 1202 Zhou, J., & Chen, X. (2021). Removal of Crossed Artifacts from Multimodal Dispersion Curves with Modified  
1203 Frequency–Bessel Method. *Bulletin of the Seismological Society of America*. <https://doi.org/10/gmr5sx>