Testing performances of the optimal sequence estimation and autoregressive method in the frequency domain for estimating eigenfrequencies and zonal structure coefficients of low-frequency normal modes

Josipa Majstorović, Séverine Rosat, Sophie Lambotte, Yves Rogister

Geophysical Journal International

Geophys. J. Int. (2019) **216,** 1157–1176
Advance Access publication 2018 November 15
GJI Seismology

doi: 10.1093/gji/ggy483

Testing performances of the optimal sequence estimation and autoregressive method in the frequency domain for estimating eigenfrequencies and zonal structure coefficients of low-frequency normal modes

Josipa Majstorović [1], Séverine Rosat,[1] Sophie Lambotte[1] and Yves Rogister[2]

[1] *Institut de Physique du Globe de Strasbourg UMR* 7516, *Université de Strasbourg/EOST, CNRS,* 5 *rue Descartes,* 67084 *Strasbourg, France. Email: jmajstorovic@unistra.fr*

[2] *Ecole et Observatoire des Sciences de la Terre, Université de Strasbourg,* 5 *rue Descartes,* 67084 *Strasbourg, France*

Accepted 2018 November 14. Received 2018 November 12; in original form 2018 September 12

SUMMARY

Improvement of global 3-D Earth density and velocity models is based in part on measurements of Earth's normal mode eigenfrequencies and splitting function coefficients. Despite many methods developed inconsistency in measurements still exists and it is difficult to understand which results are more precise, that is, which methods introduce less systematic biases in the measurements. Therefore, the main goal of this study is to test the performances of typically used techniques in low-frequency normal mode studies: the optimal sequence estimation stacking method and the autoregressive method in the frequency domain, where validation of the estimates is performed with the phasor walkout method. Motivations for their utilization are their easy and fast implementation and their accurate performances when it comes to eigenfrequency estimates. For this purpose, we first perform the analysis with synthetic seismograms in order to evaluate how the station distributions and noise levels impact the estimates of eigenfrequencies and structure coefficients. Synthetic seismograms are calculated for a 3-D realistic earth model, which includes Earth's rotation as well as ellipticity and other lateral heterogeneities. They were computed by means of normal mode summation and a perturbation theory for modes up to 1 mHz. The three methods above are also applied to long-period seismometer and superconducting gravimeter data recorded after six earthquakes of magnitude greater than 8.3. Finally, our study shows that the optimal sequence estimation is sensitive to the station distribution under the noise influence, while the autoregressive method for frequency estimation gives us reasonable estimates within the estimated error bars. Moreover, we present new estimates of eigenfrequencies and Q-factors for ${}_0S_2$, ${}_0S_3$, ${}_2S_1$ and ${}_3S_1$ multiplets. A new value for the c_{20} structure coefficient of ${}_0S_2$ multiplet -0.7233 ± 0.0623 μHz is obtained.

Key words: Structure of the Earth; Fourier analysis; Surface waves and free oscillations.

1 INTRODUCTION

Until the present day, the interior of the Earth represents a great challenge in geophysics. The normal mode studies have long provided some of the essential discoveries. For example, they contributed to building spherically symmetric earth models (Dziewonski & Anderson 1981) as well as 3-D models of lateral heterogeneities (Ritsema *et al.* 2011; Moulik & Ekström 2014; Koelemeijer *et al.* 2016). However, even though normal mode studies contributed in the estimation of the lateral density variations (Ishii & Tromp 1999, 2001; Trampert *et al.* 2004), the resolution of density is still controversial (Resovsky & Ritzwoller 1999; Romanowicz 2001; Kuo & Romanowicz 2002; Al-Attar *et al.* 2012; Akbarashrafi *et al.* 2017). During the years many methods were developed, from the well-established techniques to retrieve eigenmode frequencies and quality factors, such as stripping (Gilbert 1971; Ritzwoller *et al.* 1986) and stacking methods (Courtier *et al.* 2000). Further, techniques like the autoregressive and nonlinear fit of a resonance function have also been widely used (e.g. Chao & Gilbert 1980; Rosat *et al.* 2005; Ding & Shen 2013a), while splitting function coefficients are usually determined through iterative nonlinear spectral fitting (Woodhouse & Giardini 1985; Ritzwoller *et al.* 1986, 1988; Giardini *et al.* 1987, 1988; Resovsky & Ritzwoller 1998; Deuss *et al.* 2011, 2013). For more complete insight into existing techniques the

reader is referred to the papers by Masters & Gilbert (1983), Ding & Shen (2013a) and Ding & Chao (2015a). Despite these various studies and methods, substantial uncertainties and inconsistency in the singlets eigenfrequencies and splitting function coefficient measurements still exist (Pachhai *et al.* 2016; Akbarashrafi *et al.* 2017).

A reason for this is the inherent problem of spectral leakage and mode–mode interference, which introduces a systematic bias in spectral peak measurements in frequency domain and consequently also in the split singlet frequency measurements (Guoming *et al.* 1983). This issue has been more or less solved by introducing data tapering, however, this solution is more efficient for well-isolated low-frequency modes (Dahlen 1982a). The next essential problem is the presence of noise in recorded displacements. Substantial studies have been done on the estimation of station noise levels (e.g. Rosat *et al.* 2003; Rosat & Hinderer 2011) and noise sources (Widmer-Schnidrig 2003, and references therein). Nevertheless, the questions of how the noise deteriorates estimates, how it introduces bias in methods and what is the level of sensitivity of the methods are still important to understand.

The calculation of realistic normal mode displacements is another problem in normal mode studies. The problem is twofold, since one needs to truncate the normal mode infinite set to a finite one and also accurately implement the mode coupling theory (Dahlen 1968, 1969; Woodhouse & Dahlen 1978; Woodhouse 1980, 1983; Woodhouse & Giardini 1985). It is known that self- and group-coupling approximations introduce biases over the full-coupling approximation (SC, GC, FC, respectively; Deuss & Woodhouse 2001, 2004; Irving *et al.* 2008, 2009; Al-Attar *et al.* 2012; Yang & Tromp 2015; Akbarashrafi *et al.* 2017). These theoretical errors cannot be ignored for the frequencies higher than 1 mHz, where FC calculations are necessary to obtain sufficiently accurate spectra (Yang & Tromp 2015). They also have for sure affected earlier studies where the measurements substantially depend on the comparison of the synthetically calculated normal mode displacements with the observations. However, since in this study we do not implement those measurements, we just acknowledge these theoretical errors as being significant.

There are still important issues in normal mode studies, which have been acknowledged, but not properly scrutinized. In this study the focus is set on the inevitable presence of noise in the records and the number of stations used during the measurement process. For example, for the signal-to-noise ratio (SNR) $\leq$50 the aforementioned effects have higher errors than theoretical errors introduced by the SC, GC and FC approximations. Therefore, the objective of this paper is to measure sensitivity and to test performance of the commonly used methods under the influence of noise and number of stations. The tested methods are the stacking method, called the optimal sequence estimation (OSE) introduced by Ding & Shen (2013a), and the autoregressive method for the estimation of normal mode's parameters introduced by Chao & Gilbert (1980; ARFD80). The OSE was first applied in the search for the Slichter modes (Slichter 1961), which are the three translational modes of the inner core (Ding & Shen 2013a; Ding & Chao 2015c), but its application extended to the retrieval of other normal modes (Ding & Chao 2015a; Zeng & Shen 2017, 2018) as well as to the pole tide signals (Ding & Chao 2016). It has been proven that OSE has better performance than other stacking methods, such as the spherical harmonic stacking (SHS; Buland *et al.* 1979) and the multistation experiment (MSE; Courtier *et al.* 2000), since it was developed on the basis of these two methods using the principle of the noise-term elimination (Zeng & Shen 2017). The OSE has been successfully extended to transverse components (Ding & Chao 2015a) and applied in the GC approximation, where nearby modes were grouped as an isolated cluster (Zeng & Shen 2017). In terms of frequency range it has been successfully applied in the SC approximation from 0.309 mHz ($_0S_2$) to 9.865 mHz ($_{27}S_2$) (Ding & Shen 2013a; Ding & Chao 2015b; Zeng & Shen 2017) and in the spheroidal–spheroidal GC approximation from 1.413 mHz ($_4S_1$) to 2.822 mHz ($_6S_3$) (Zeng & Shen 2018). Most studies claim that OSE in the SC approximation is performing accurately and as a consequence it can be applied to the modes with $f \leq 1.5$ mHz (Zeng & Shen 2017, 2018). Furthermore, even though the OSE method proved to be the foremost among the stacking methods, some limitations have been identified. For example, the dependence on records' SNRs and the number of stations used in stacking (Ding & Shen 2013a; Zeng & Shen 2017). For the estimation of the harmonic function parameters we chose the ARFD80 method because it has been proven to be very successful too. This method is fast, highly accurate, multimode estimation suitable and easy to implement (Chao & Gilbert 1980; Masters & Gilbert 1983; Ding & Shen 2013b; Ding & Chao 2015a,b,c; Zeng & Shen 2017, 2018) and it has been applied to the same frequency range as OSE, since most studies use these two methods together.

Motivated by the previous findings, we decide first to test these methods on synthetic seismograms. Experiments contain two main parts, one is the gradually addition of noise in our records and the other part is the usage of different networks in OSE. The purpose of both parts is the quantification of noise and network effect on the estimates of eigenfrequencies, quality factors and amplitudes. The importance of noise effect is implicit, on the other hand the network effect is linked with the OSE feature. Theoretically, the measured frequency of a normal mode should be the same anywhere on the Earth and the usage of different station network should have the same results. Consequently, if we want to measure the split frequencies of specific modes one only needs to consider those stations which are not located on the nodal lines of the eigendisplacements (Masters *et al.* 2000; Häfner & Widmer-Schnidrig 2013). Thus, the second part of this work aims at testing whether any systematic bias is introduced by the network effect. We also validate the accuracy of our complex eigenfrequency estimate graphically using the phasor walkout method (Zürn & Rydelek 1994). Furthermore, once measured, the split eigenfrequencies are used to retrieve the splitting function coefficients using a perturbation theory of the first order (Dahlen 1974; Ritzwoller *et al.* 1988; Widmer *et al.* 1992; Häfner & Widmer-Schnidrig 2013).

Since we are interested in the method's performances, our tests include only one source mechanism, also only one multiplet chain at the time and we focus on the low-frequency modes, where we can use the SC approximation. The interest in the highly precise measurement of the low-frequency split eigenfrequencies is valuable since it is known that any existing density model should fit split frequencies perfectly (Widmer-Schnidrig 2003). For suitable candidates, we chose to work with the lowest frequency multiplet chain $_0S_2$–$_0T_2$–$_2S_1$–$_0S_3$, spanning from 0.309 to 0.468 mHz, and the lowest frequency multiplet chain where there is a significant interference between adjacent multiplets, $_0T_5$–$_2S_2$–$_1S_3$–$_3S_1$, spanning from 0.928 to 0.943 mHz.

The paper is organized as follows. In the first section we review the theoretical basis of the methods used and we describe the applied procedure. In the second section we show the results of our synthetic experiments for two different chains of multiplets for different station distributions and gradually increased noise levels. In the third section, we summarize our analysis on real observations,

based on long-period seismometer and superconducting gravimeter (SG) data after six earthquakes of magnitude larger than 8.3, for multiplets ${}_0S_2$, ${}_2S_1$, ${}_0S_3$ and ${}_3S_1$. Finally, we present new estimates of eigenfrequencies, quality factors and associated splitting function coefficients.

2 THEORY AND METHODS

For the purpose of analysing complex eigenfrequencies and splitting function coefficients we established a protocol: first, we stack our records using OSE stacking method; second, we calculate complex eigenfrequencies, quality factors (further on Q-factors), and amplitudes of target singlets using ARFD80 method; third, we check the validity of our eigenfrequency estimates using the phasor walkout method; and finally from the eigenfrequency we estimate splitting function coefficients. Further, we give a short introduction to the first-order splitting theory and the methods used in the analysis.

2.1 First-order perturbation theory for isolated multiplet

Degenerate frequencies of spherically symmetric non-rotating earth models are split by rotation, ellipticity and lateral heterogeneities. In most studies, rotation and ellipticity are known and lateral heterogeneities, which are represented by velocity and density perturbations, are the quantities to estimate. The lateral heterogeneities are described by the splitting function coefficients, c_{st}, which are linearly related to the model parameters. Estimation of the splitting function coefficients is a highly nonlinear inverse problem which is commonly performed iteratively from a starting model in spectral domain (Ritzwoller *et al.* 1986, 1988; Giardini *et al.* 1987, 1988; Li *et al.* 1991). This procedure has been frequently used in the past several years specially for building a catalogue of c_{st} coefficients (Resovsky & Ritzwoller 1998; Deuss *et al.* 2011, 2013; Koelemeijer *et al.* 2013). Besides, Ritzwoller *et al.* (1986) and Widmer *et al.* (1992) proposed retrieving information about splitting function coefficients directly from the estimated normal mode eigenfrequencies. This approach is based on a first-order perturbation theory for isolated multiplets and the assumption that the lateral heterogeneity is predominantly zonal which is also valid for some isolated multiplets (Woodhouse & Dahlen 1978; Woodhouse 1980; Dahlen & Tromp 1998). Moreover, for multiplets below 1 mHz where the splitting is dominated by the effect of rotation it is known that the dominant heterogeneity sensed is axisymmetric (Widmer *et al.* 1992). All listed assumptions are valid only for the low-frequency modes and thus they are the main focus in this study.

For an isolated multiplet and aspherical earth model the displacement recorded at the position $\boldsymbol{r}$ and triggered by the source at $\boldsymbol{r}_0$ can be represented as (Woodhouse & Girnius 1982)

$$\boldsymbol{u}(\boldsymbol{r}, t) = \mathrm{Re}\left[\boldsymbol{r}_k(\boldsymbol{r})^T \mathrm{e}^{\mathrm{i}H_k t} \boldsymbol{s}_k(\boldsymbol{r}_0) \mathrm{e}^{\mathrm{i}\bar{\omega}_k t}\right], \tag{1}$$

where subscript k stands for the multiplet index and is a shorthand for indices (n, l, q) where $n \geq 0$ is overtone number, $l \geq 0$ is the harmonic degree and q stands for either S or T, thereby designating spheroidal or toroidal modes. Each multiplet is made up of $(2l + 1)$ singlets, which are degenerate when the earth model is spherical and non-rotating, labelled with azimuthal order m. Therefore, $\boldsymbol{r}_k$ is the $(2l + 1)$ receiver vector defined with the displacement eigenfunction (the superscript T denotes the transpose). $\boldsymbol{s}_k(\boldsymbol{r}_0)$ is the $(2l + 1)$ source vector describing the excitation of each singlet and it depends on the moment rate tensor $\mathbf{M}$ and strain tensor $\mathbf{E}_k$. $\bar{\omega}_k$ is the degenerate eigenfrequency of the multiplet and H_k is the $(2l + 1) \times (2l + 1)$ splitting matrix. The splitting matrix contains perturbations due to rotation, ellipticity and lateral heterogeneities and is defined as

$$H_{mm'} = \bar{\omega}_k(a + mb + m^2 c)\delta_{mm'} + \sum_{\substack{s=0 \\ s \text{ even}}}^{2l} \sum_{t=-s}^{s} \gamma_{st}^{mm'} c_{st} \tag{2}$$

where aforementioned rotation and elliptical effects are defined with parameter b due to the first-order effect of the Coriolis force, whereas a and c are due to ellipticity and the second-order rotational effects (Dahlen & Sailor 1979). $\delta_{mm'}$ is the Kronecker delta symbol and c_{st} are, aforementioned, splitting function coefficients which are linearly related to the perturbations to the spherical structure. Additionally, $\gamma_{st}^{mm'}$ are real coefficients which can be easily evaluated (Woodhouse 1980; Ritzwoller *et al.* 1986; Dahlen & Tromp 1998). These coefficients are non-zero only if s is even, $0 \leq s \leq 2l$ and $t = m - m'$, consequently these conditions define the kind of structure a specific isolated multiplet is sensitive to. Obtaining the splitting matrix, calculating its eigenvalues and adding them to the multiplet degenerate eigenfrequency, one can calculate the singlet split eigenfrequencies, ω_m, within the multiplet. In the special case where eq. (2) is diagonal, resulting in a singlet being sensitive only to rotation, ellipticity and even degree axisymmetric aspherical structure ($t = 0$), eq. (2) simplifies to

$$H_{mm} = \bar{\omega}_k(a + mb + m^2 c) + \sum_{\substack{s=0 \\ s \text{ even}}}^{2l} \gamma_{s0}^{mm} c_{s0}, \tag{3}$$

which leads us to the expression of the split eigenfrequency

$$\omega_m = \bar{\omega}_k(1 + a + mb + m^2 c) + \sum_{\substack{s=0 \\ s \text{ even}}}^{2l} \gamma_{s0}^{mm} c_{s0}. \tag{4}$$

Relation (4) shows that if one measures the split eigenfrequencies ω_m and further subtracts the first term on the right-hand side, calculated for a given earth model, from the left-hand side one can estimate axisymmetric splitting function coefficients sensitive to even degree structures. In the next sections we will describe how the singlet eigenfrequency can be estimated from the displacement recorded at the Earth's surface.

2.2 The complex eigenfrequency estimation

The recorded ground displacement after an earthquake at any point at the surface of the Earth is a discrete time-series. It represents the supersposition of discrete modes of oscillation that can be represented in the complex domain with exponential functions. The measurement of exponential parameters, for example, amplitudes and exponential arguments, will be done using the ARFD80 method, which is based on the Prony technique for extracting exponential signals from time-series. Only the basic outline will be presented, and for further discussion the reader is referred to the papers by Chao & Gilbert (1980) and Chao (1990).

A displacement represented as discrete time-series of superimposed decaying, complex exponential functions can be written as

$$x(t) = \sum_{j=1}^{M} [A_j \mathrm{e}^{\mathrm{i}\sigma_j t} + A_j^* \mathrm{e}^{-\mathrm{i}\sigma_j^* t}], \quad t = \Delta t, 2\Delta t, \ldots, N\Delta t, \tag{5}$$

where M is the number of recorded modes, Δt is sampling rate, N is the number of data samples, A_j are the complex amplitudes, σ_j are the complex frequencies that can be written in terms of eigenfrequencies and decay rates as $\sigma_j = \omega_j + i\alpha_j$, and $*$ denotes the

complex conjugate. Complex amplitudes A_j and complex frequencies σ_j are unknowns to be determined. For this purpose, eq. (5) can be represented by a recursive system of linear difference equations of order $2M$:

$$x(t) = \sum_{i=1}^{2M} S_i x(t - i\Delta t), \quad t = 2M\Delta t + \Delta t, \ldots, N\Delta t, \tag{6}$$

where S_i are real constant coefficients. If for the demonstration we put $M = 1$, then relations (5) and (6) become

$$x(t) = A_1 e^{i\sigma_1 t} + A_1^* e^{-i\sigma_1^* t}, \quad t = \Delta t, 2\Delta t, \ldots, N\Delta t, \tag{7}$$

$$\begin{bmatrix} x(3\Delta t) \\ \vdots \\ x(N\Delta t) \end{bmatrix} = \begin{bmatrix} x(2\Delta t) & x(\Delta t) \\ \vdots & \vdots \\ x(N\Delta t - \Delta t) & x(N\Delta - 2\Delta t) \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}. \tag{8}$$

Substituting eq. (7) into the first equation of eq. (8) (Masters, G. lectures' notes https://igppweb.ucsd.edu/guy/sio227b/) with a little bit algebra it is easy to see that

$$\begin{aligned} S_2 &= -e^{-2\alpha_1 \Delta t}, \\ S_1 &= 2\cos(\omega_1 \Delta t) e^{-\alpha_1 \Delta t}, \end{aligned} \tag{9}$$

thus, by solving for coefficients S_1 and S_2 we can found frequency ω_1 and decay rate α_1 and Q-factor $Q_1 = \omega_1/2\alpha_1$.

The calculation of the real coefficients S_i using eq. (6) is heavy in the time domain, while the number of excited modes after the earthquake is large and unknown. The problem is resolved by considering that the basis functions $e^{i\sigma_j}$ are separated into individual peaks in frequency domain. Basically, the main idea behind the ARFD80 method is to Fourier transform eq. (8) in order to solve a system of linear equations for narrow frequency band containing the small number of modes. The advantage of this procedure is to focus on the narrow frequency range where the mode is expected to occur. This enables resizing the linear system (8). As shown earlier, if $M = 1$ the system to solve consists of two unknowns and $K > 2M$ linear equations, where usually $3 \leq K \leq 5$ and it represents 3 or 5 frequencies that describe the position of the spectral peak in the frequency domain we study. The described analysis plainly depends on the Fourier transform and hence bears all technical advantages and disadvantages of the Fourier transform. To reduce the side-band levels of the spectral peaks due to spectral leakage we perform tapering with a Hann window and to refine the waveform frequency resolution we perform zero padding to each column in eq. (8).

Even though ARFD80 is capable of analysing small group of modes, in this paper the focus is set on analysing one mode at once, hence it was important to observe a spectral peak in frequency domain that belongs to the target singlet. When the singlet is visible we choose at least five frequencies that describe the position of the spectral peak in frequency domain, one which defines the main peak and two on each side of the main peak. After solving linear system (8) in frequency domain and obtaining coefficients S_i, calculating frequencies, decay rates and Q-factors we also estimate complex amplitudes (Chao & Gilbert 1980) used in bootstrap experiments for the calculation of parameter's standard deviations following Häfner & Widmer-Schnidrig (2013). Simultaneously, for every estimate we also measure the SNR as the ratio of peak amplitude of the singlet over the root-mean-square amplitude of two narrow frequency bands targeting the singlet. Once the signal frequency is obtained we validate our estimates using the phasor walkout method described in Section 2.4.

2.3 Data stacking method

Successful estimation of harmonic function parameters such as complex frequency, Q-factor and amplitude, substantially depends on the representation of the spectral peak in frequency domain. One can improve spectral peak representation by increasing the time-series length, however bearing in mind the trade-off between frequency resolution and noise level. An alternative option is to use stacking methods to enhance the SNR of the target signal. Here we will give a short introduction to the OSE (Ding & Shen 2013a).

If we consider only the radial component of the displacement at the surface of the anelastic earth model and introduce the singlet complex eigenfrequency σ_m, relation (1) gives

$$u_R(\Omega, t) = \sum_m {}_nU_l Y_l^m(\Omega) s_m(\boldsymbol{r_0}) e^{i\sigma_m t} \tag{10}$$

where ${}_nU_l$ is the value taken by the radial eigenfunction at the surface, $Y_l^m(\Omega)$ is the spherical harmonic function of degree l and order m and $\Omega = (\theta, \phi)$ are the colatitude and longitude of the receiver. For brevity we will introduce $\epsilon_m = {}_nU_l s_m(\boldsymbol{r_0})$ and $A_m = \epsilon_m e^{i\sigma_m t}$, therefore eq. (10) becomes

$$u_R(\Omega, t) = \sum_m A_m Y_l^m(\Omega), \tag{11}$$

and for $j = 1, \ldots, N$ stations from relation (11) one can form a multistack

$$\mathbf{U} = \mathbf{YA}, \tag{12}$$

where

$$\mathbf{U} = \begin{bmatrix} u_R(\Omega_1, t) \\ u_R(\Omega_2, t) \\ \vdots \\ u_R(\Omega_N, t) \end{bmatrix}, \tag{13}$$

$$\mathbf{Y} = \begin{bmatrix} Y_L^{-L}(\Omega_1) & Y_L^{-L+1}(\Omega_1) & \ldots & Y_L^{L}(\Omega_1) \\ Y_L^{-L}(\Omega_2) & Y_L^{-L+1}(\Omega_2) & \ldots & Y_L^{L}(\Omega_2) \\ \vdots & & & \vdots \\ Y_L^{-L}(\Omega_N) & Y_L^{-L+1}(\Omega_N) & \ldots & Y_L^{L}(\Omega_N) \end{bmatrix}, \tag{14}$$

$$\mathbf{A} = \begin{bmatrix} \varepsilon_{-L} \exp^{i\sigma_{-L} t} \\ \varepsilon_{-L+1} \exp^{i\sigma_{-L+1} t} \\ \vdots \\ \varepsilon_{L} \exp^{i\sigma_{L} t} \end{bmatrix}, \tag{15}$$

for $t = 1, \ldots, n$ time samples. In eq. (12) $\mathbf{U}$ is an $[N \times n]$ matrix of radial-component observations, $\mathbf{Y}$ is an $[N \times (2L + 1)]$ matrix of spherical harmonics and $\mathbf{A}$ is a $[(2L + 1) \times n]$ matrix to be inverted for which each row consists of only one singlet for the target (N, L) multiplet. If $N > (2L + 1)$, system (12) can be solved by the general least-squares procedure

$$\mathbf{A} = (\mathbf{Y}^T \mathbf{pY})^{-1} \mathbf{Y}^T \mathbf{pU} \tag{16}$$

where $\mathbf{p}$ is a weight whose values can be chosen to be inversely proportional to the SNR of the target mode. Eq. (16) defines the OSE method.

Separating each singlet resonance function within a multiplet in matrix $\mathbf{A}$ by yielding only one spectral peak in spectral domain corresponding to the target singlet gives OSE advantage over, for example, the least-squares spectral analysis (Vaníček 1969), the product spectral analysis (Smylie 1992) and the cross-spectrum

analysis (Hinderer *et al.* 1995), where all singlets within the multiplet would appear together in only one spectrum. Furthermore, the OSE method does not require the information about the source mechanism, which gives it advantage over, for example, the singlet stripping method (Gilbert 1971; Ritzwoller *et al.* 1986). For more comparison we recommend the paper by Ding & Shen (2013a).

2.4 The phasor walkout method and test of goodness

As previously stated the performance of the ARFD80 method plainly depends on the representation of a spectral peak in frequency domain, hence the difficulty still remains whether the spectral peak observed in the frequency domain corresponds to a harmonic signal or noise. One way to address this problem is to use the phasor walkout method revisited by Zürn & Rydelek (1994). Essentially, it is a graphical representation of Fourier transform for a test frequency where one estimates the complex contribution of the Fourier transfom for each sample and sums the corresponding vectors in the complex plane. The shape and behaviour of the resulting vector pattern, the walkout, give us information about the signal properties.

The derivation process is quite straightforward. If we consider a time-series x_j, for $j = 1, \ldots, N$ equidistantly sampled with sampling rate Δt, which may consist of a signal with the frequency f_0, the phasor walkout contributions obtained for the test frequency f_s are

$$p_j(f_s) = x_j \mathrm{e}^{-\mathrm{i}2\pi f_s (j-1)\Delta t}, \quad j = 1, \ldots, N. \tag{17}$$

Notably, these contributions are complex and they are recognized as vectors in a 2-D space. Deriving the full phasor walkout pattern is achieved by adding these vectors graphically and successively in the complex plane. For a better understanding it is useful to look at function (17) where x_j defines the scaling of walkout and the exponential factor defines the walkout rotation by an angle $2\pi f_s \Delta t$. First, the walkout has a loop-like circular shape if the scaling factor is constant. This loop shape is completely defined by the rotation angle, that is, the full loop of 2π is closed in K steps, where $K = 2\frac{f_n}{f_s}$, with f_n being Nyquist frequency ($f_n = \frac{1}{2\Delta t}$). Consequently, the loop-like circular shapes are K polygons. If K is an integer the polygons are aligned and if K is not an integer the polygons are rotated with respect to each other. If we now consider that x_j is noise, one would no longer have polygons with the constant phase change in successive samples, but randomly changing phases which would result in a random walk.

One important case to consider is a simple harmonic function $x_j = A_0 \cos(2\pi f_0 (j-1)\Delta t)$ in which case the phasor walkout for tested frequency f_s is

$$p_j(f_s) = \frac{A_0}{2}\left(\mathrm{e}^{\mathrm{i}2\pi(f_0 - f_s)(j-1)\Delta t} + \mathrm{e}^{-\mathrm{i}2\pi(f_0 + f_s)(j-1)\Delta t}\right). \tag{18}$$

Assuming that $f_0 \neq f_s$ the shape of eq. (18) is still a polygon even though a more complicated one because each axis, the real and complex ones, now contains the sum of two trigonometric functions instead of one trigonometric function. The rotation angle is now changed and it is $2\pi(f_0 - f_s)\Delta t$, hence the full circle is now reached with K steps, where $K = \frac{2f_n}{|f_0 - f_s|}$. An interesting feature is obtained when $f_0 = f_s$, so eq. (18) becomes

$$p_j(f_s) = \frac{A_0}{2}\left(1 + \mathrm{e}^{-\mathrm{i}4\pi f_0 (j-1)\Delta t}\right). \tag{19}$$

Examining eq. (19) one can note that the complex part is bounded by a sine function, while the real part is progressively advancing because of the constant term, number one, in its definition. Strikingly, this means that the phasor walkout of a harmonic signal, when $f_0 = f_s$, is going to gradually progress without curving, hence showing linearity. This is the most useful feature of the phasor walkout: by examining the phasor walkout pattern, one can conclude that the tested signal is really a harmonic signal with frequency f_0.

Furthermore, it is important to consider the case of signals consisting of two harmonic functions with very close frequencies because this is often the case in normal mode studies. Work on this topic has also been done by Zürn & Rydelek (1994). Due to the very similar harmonic function frequencies in frequency domain the phasor walkout of each harmonic function now depends on the relative amplitude between the two functions. The phasor walkout at the frequency of the signal with larger amplitude has rapid phase changes, while the phasor walkout of the signal with smaller amplitude has slower phase changes and additional loops. However, the most important is that both walkouts show linear progress, which is not the case if the tested frequencies are different from the true harmonic frequencies. Different cases are studied in the Appendix of this paper.

The situation is becoming even more complicated when the signals are contaminated by noise. The regularity of the phasor walkout in the presence of noise is more or less deteriorated. In such circumstances it is useful to use a range of tested frequencies to calculate the phasor walkouts. Theoretically, if the estimated frequency f_e is the true frequency of the harmonic function only the phasor walkout with $f_s = f_e$ should be a straight line and for all other tested frequencies the phasor walkouts should be softly curved. As stated before differences between slightly different tested frequencies are sometimes difficult to observe due to the presence of noise. To overcome this problem, we propose to compute the coefficient of determination R^2 (Draper & Smith 1998) on our phasor walkouts. Basically, this includes fitting a linear model on our phasor walkout and then estimating R^2 coefficient, which is a statistical feature determining how well the fitted model describes the data. This would be an additional statistical test to show how much our phasors are actually straight or curved. Hence, having a basic linear model $Y = aX + b$ one should estimate coefficients a and b, Y being the complex part of phasor walkout and X the real part of the walkout. The next step is to build a linear model Y_m using estimated coefficients a and b and X data: $Y_m = aX + b$. Ending up with observed data Y and modelled data Y_m one can calculate the coefficient of determination R^2 defined by

$$R^2 = 1 - \frac{\sum_i (Y_i - Y_{m,i})^2}{\sum_i (Y_i - \bar{Y})^2}, \tag{20}$$

where $\bar{Y}$ is the mean of the Y_i. Values of R^2 fall between 0 and 1, with a value 1 indicating that all variance is accounted for by the model, that is, all of the data points fall perfectly on the regression line. Therefore, the coefficient of the determination for our five tested frequencies should be the largest for $f_s = f_e$ if f_e is the true frequency of the harmonic function. Finally, the R^2-test does not provide us with the frequency estimates, but with the statistical values for our estimated frequencies.

3 SYNTHETIC EXPERIMENTS

To test how the methods are influenced by different noise levels and station distributions, experiments are conducted by simulating records using real earthquakes parameters. Synthetic seismograms

are calculated for a 3-D earth model by means of normal mode summation and perturbation theory. Reference basis functions, eigenfrequencies and the associated eigenfunctions for a spherically symmetric, non-rotating earth model are obtained for PREM model (Dziewonski & Anderson 1981) using MINEOS software package (Woodhouse 1988) with a cut-off frequency of 80 mHz. Perturbations due to rotation, ellipticity and lateral heterogeneities are introduced by computing the splitting matrix H_k for the GC approximation (Dahlen & Tromp 1998, p. 643) using S40RTS 3-D earth model (Ritsema *et al.* 2011). Therefore, shear wave velocity perturbation $\delta \ln V_s$ is calculated from S40RTS model, where compressional wave velocity perturbation are scaled by $\delta \ln V_p = 0.5 \delta \ln V_s$ and density perturbation by $\delta \ln \rho = 0.3 \delta \ln V_s$. Source vector is calculated using the Global CMT Catalog solutions (Dziewonski *et al.* 1981; Ekström *et al.* 2012).

The analyses are done for the spheroidal modes of two multiplet chains, namely ${}_0S_2 - {}_0T_2 - {}_2S_1 - {}_0S_3$ and ${}_0T_5 - {}_2S_2 - {}_1S_3 - {}_3S_1$, that have frequencies <1.5 mHz. Target modes are considered in the GC approximation, that is, the multiplets within the chain were coupled and isolated from all the other modes (Deuss & Woodhouse 2001).

For the ${}_0S_2 - {}_0T_2 - {}_2S_1 - {}_0S_3$ chain we generate seismograms for the magnitude 9.0 Tohoku earthquake that occurred on 2011 March 11 at 05:46 UTC with the epicentre approximately 70 km east of the Oshika Peninsula of Tohoku. Station database is built from the Global Seismograph Network (GSN–IRIS/USGS; http://www.fdsn.org/networks/detail/IU/) consisting of 81 stations. From this general network we define four groups of stations, namely 'full' network which contains all stations (81 stations), 'northern' network with stations in northern hemisphere (53 stations), 'southern' network with stations in southern hemisphere (28 stations) and 'selected' network which contains only 10 stations chosen for specific reasons, explained further in the text. Noise is generated using a random function that draws scalars from the standard normal distribution with zero mean and variance one and is added to the synthetic time-series. If the station displacement is denoted by x_s and if the noise function is denoted by x_n the simple relation between the two is established by

$$x_{s+n} = x_s + \sigma_n x_n, \tag{21}$$

σ_n being the noise standard deviation with the acceleration unit nm s^{-2}, which we gradually increase in our analysis. For every experiment that consists of different networks and noise levels the outputs are estimated frequency, Q-factor and amplitude of the target singlet. Further, our results are represented in terms of frequencies only, which we need to estimate the splitting function coefficients.

From the three spheroidal modes in the ${}_0S_2 - {}_0T_2 - {}_2S_1 - {}_0S_3$ chain, we show results for ${}_0S_2$ and ${}_2S_1$, while the results for ${}_0S_3$ are more or less similar to the ones for ${}_0S_2$. To successfully analyse ${}_0S_2$ one needs to use at least 4 hr of data to separate this multiplet from the next multiplet in frequency domain, 2.5 d to separate singlets within the multiplet, while the 1.1 Q cycle is 19 d for having the optimal SNR (Dahlen 1982a). We choose 15-d-long records which are zero-padded till 90 d to obtain a frequency resolution of order 10^{-8} Hz with a sampling rate of 10 s. After the 2011 event, amplitudes of $m = \pm 1$ singlets are better excited than amplitudes of $m = \pm 2$ singlets and specially compared to $m = 0$ (Ding & Shen 2013b). For this reason, the selected network contains only 10 stations for which the $m = 0$ singlet is visibly excited. The results for the best excited singlet $m = 1$ of ${}_0S_2$, noted ${}_0S_2^1$, are in Fig. 1. The results, quite straightforward, confirm that when noise is gradually added the SNR decreases and standard deviations increase. Furthermore, the results for the full network are overall foremost while the estimates have the smallest standard deviations and relative errors. Besides, estimates become more scattered around the true synthetic value (dashed line) with noise increment. However, not all frequency estimates contain the true synthetic values within their standard deviations. Conclusions drawn for the frequency estimates are also applicable to the estimates of Q-factors. In the case of Q-factors, relative error for full network is not more than 6.2 per cent ($\sigma_n =$ 9 nm s^{-2}) while it increases to 50 per cent ($\sigma_n = 9$ nm s^{-2}) for selected network.

Further, to better understand our results we examine the phasor walkout for five test frequencies $f_s = \{f_e - 2\sigma_f, f_e - \sigma_f, f_e, f_e + \sigma_f, f_e + 2\sigma_f\}$ and three cases: (1) full network with noise level $\sigma_n = 1$ nm s^{-2} and relative error 1.4 ppm, (2) full network with $\sigma_n = 9$ nm s^{-2} and 66.6 ppm (the largest for this network, frequency with its standard deviation does not include the synthetic value), (3) selected network with $\sigma_n = 8$ nm s^{-2} and 251.1 ppm (the largest for the whole data set). The results shown in Fig. 2 demonstrate that the phasor walkout method successfully distinguishes between five different test frequencies and additionally indicates that the centre one is the true frequency of the stacked signal when the noise level is only $\sigma_n = 1$ nm s^{-2}. This encourages us to further use this method, when the noise levels are higher. Moreover, mutual comparison of these three cases demonstrates how noise deteriorates our estimates: with more noise our phase is consistently abrupt causing twisting of the walkout pattern, which nevertheless has the propensity to straightness. However, we can expect that in some cases, where this twisting is more prominent, the R^2-test will be difficult to accomplish. For the second case we can ask ourselves: why do the estimated frequencies not include the true synthetic value within their standard deviations? Our goal is to understand where the bias is coming from, OSE or ARFD80 method. From the theoretical background, if the ARFD80 method failed in estimating frequency, which means that the bias is not coming from the stacking, by plotting the phasor walkout one should obtain a curved line for $f_s = f_e$. However, if the estimated frequency is the true frequency of this harmonic signal one should have a straight line. In the latter case, it means that the OSE method introduced a bias by producing a slight peak shift of the singlet due to noise input. The results in Fig. 2 confirm that the bias is introduced with OSE method because the R^2-test is the highest for the central frequencies indicating that ARFD80 method estimated true frequency of stacked signal. Additionally, we can confirm that the bias introduced with OSE method is generated by the noise input, since the estimates of eigenfrequencies without added noise are not biased. Furthermore, observing the phasor walkout graphs and R^2-tests of the three cases it is evident that the first case has the highest R^2-test, while the third case, the lowest R^2-test, meaning that the first case holds the highest linearity and the third one the lowest.

Studying the relationship between relative errors and SNRs within all experiments conducted for the ${}_0S_2$ multiplet, thus considering all singlets, one can note that for each SNR one can associate a range of possible relative errors. Meaning that it does not necessarily means that solution with larger SNR has smaller relative error and thus better estimate. Considering this fact, we decide to compare relationship between standard deviations estimated with the bootstrap method and the one estimated from the maximum relative error from all our experiments (i.e. 200 experiments for 5 singlets, 10 noise levels and 4 networks) for a specific SNR. In Fig. 3 we plot the aforementioned standard deviations versus SNR in log scale. Standard deviations estimated from maximum relative

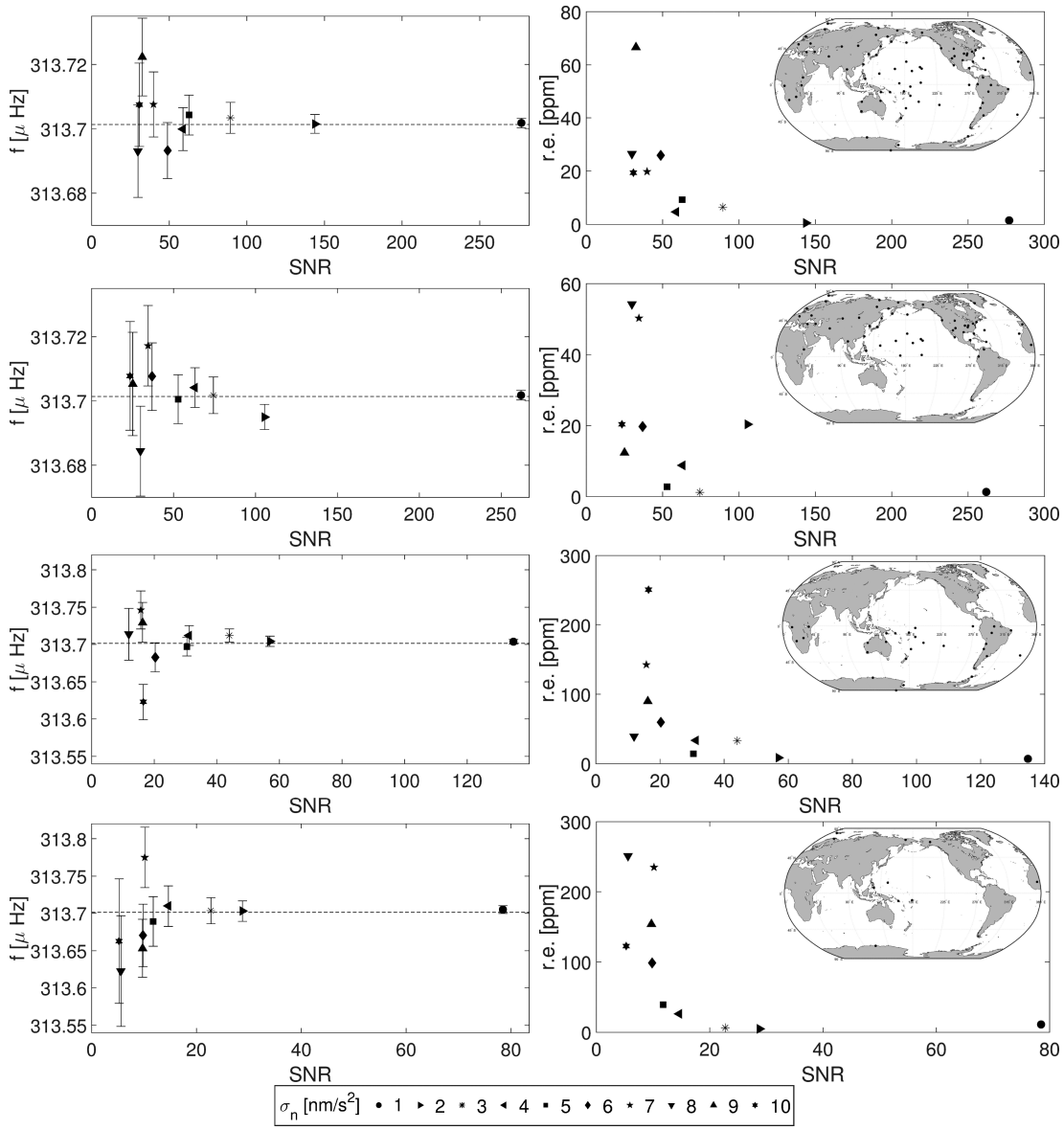

Figure 1. Synthetic experiments for ${}_0S_2^1$ singlet conducted for four station distributions: (1) full (the first row), (2) northern (the second row), (3) southern (the third row), (4) selected (the fourth row) networks. Figures on the left-hand side represent the estimated frequencies versus SNR for 10 different noise levels. Figures on the right-hand side represent the relative errors of frequencies with regard to synthetic value (dashed line) versus SNR for 10 noise levels. Be careful, the vertical scales of the left-hand side figures are optimized, thus the full and the northern networks have the same scale and the southern and the selected networks too.

errors are obtained using the formula

$$\sigma_{\text{r.e.}}(\text{SNR}) = \max \text{r.e.}(\text{SNR})\, f_{\text{syn}}, \tag{22}$$

where max r.e.(SNR) is, as stated before, a maximum relative error for the specific SNR. Surprisingly, both estimates of standard deviations are exponentially dependent on SNR. However, standard deviations, recalculated from the relative errors, are for most cases higher than the bootstrap standard deviations. This implies that for most SNRs, especially for the lower values, we can have bootstrap standard deviations which are smaller than the actual differences between synthetic values and estimated values. We can conclude that even though we have good precision on our estimates we have poor accuracy due to the biases introduced by all used methods.

From the built catalogue consisting of frequencies and their associated standard deviations for each singlet within the ${}_0S_2$ multiplet using eq. (4) we can calculate the splitting function coefficients associated for each network and noise level. For that we need the values of a, b and c parameters and of the degenerate frequency $\bar{\omega}_k$. The parameters a, b, and c are estimated by fitting a parabola to the synthetic singlets frequencies for a rotating hydrostatic ellipsoidal earth model derived from PREM. Please note that these parameters could also be computed directly using explicit formulas form Dahlen & Sailor (1979). $\bar{\omega}_k$ is a degenerate eigenfrequency of the spherical non-rotating PREM model. In the next step we subtract the first term in eq. (3) from singlet frequencies to arrive at

$$\omega_m^{\text{residual}} = \gamma_{20}^{mm} c_{20} + \gamma_{40}^{mm} c_{40}, \tag{23}$$

because for ${}_0S_2$ azimuthal order is 2, thus $s = 2, 4$. Since it was found from some studies of aspherical structures that the degree 2 structure is much larger in amplitude than the degree 4 (Ritzwoller

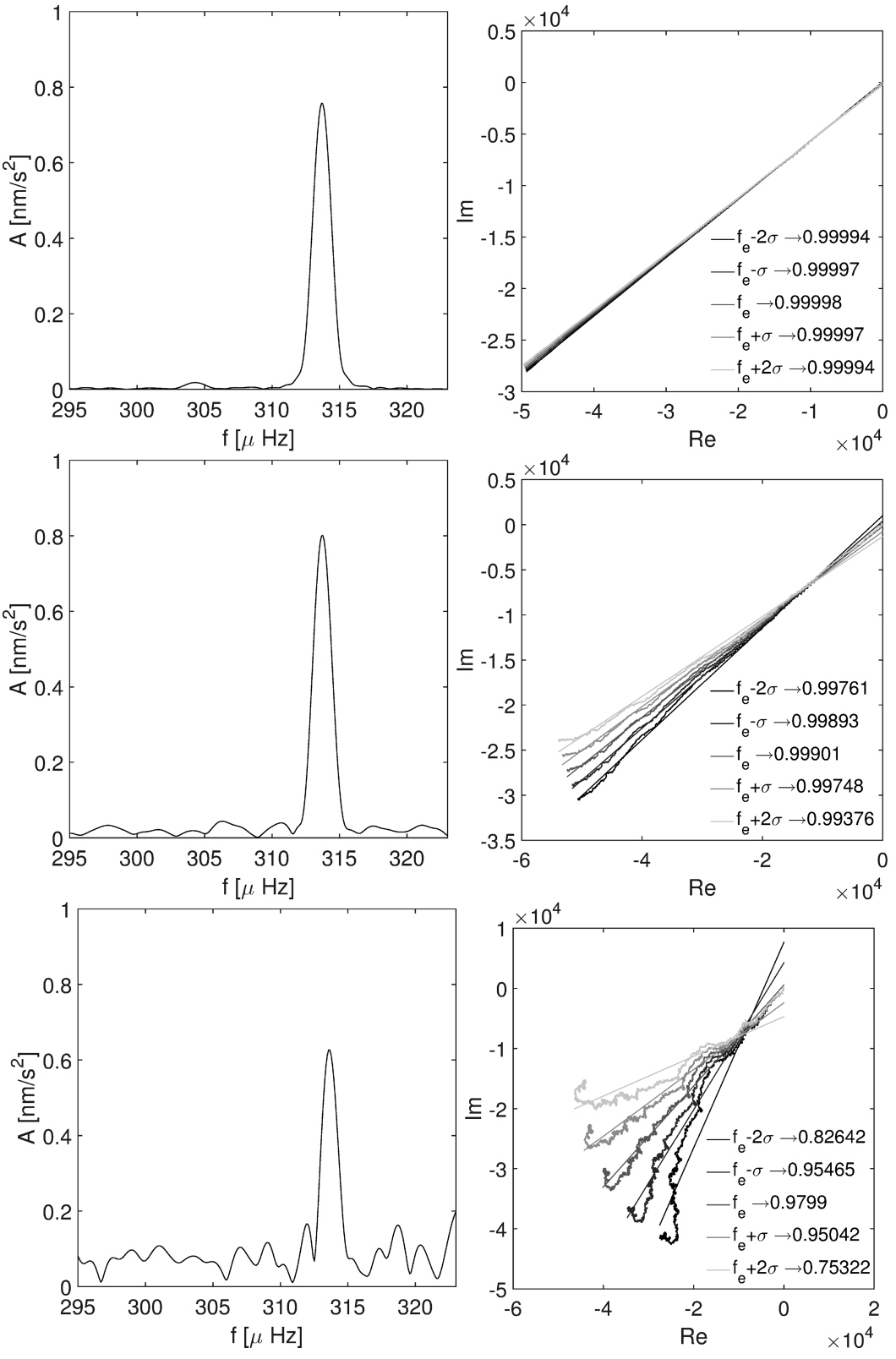

Figure 2. The Fourier transform of OSE stacked signals (left) with associated phasor walkouts and applied R^2-test values (right) for five test frequencies with f_e being the estimated frequency of stacked signal on left and σ the standard deviation. Three cases are shown: full network with noise level $\sigma_n = 1$ nm s^{-2}(top), full network with $\sigma_n = 9$ nm s^{-2} (middle) and selected network with $\sigma_n = 8$ nm s^{-2} (bottom). Results are for ${}_0S_2^1$ singlet.

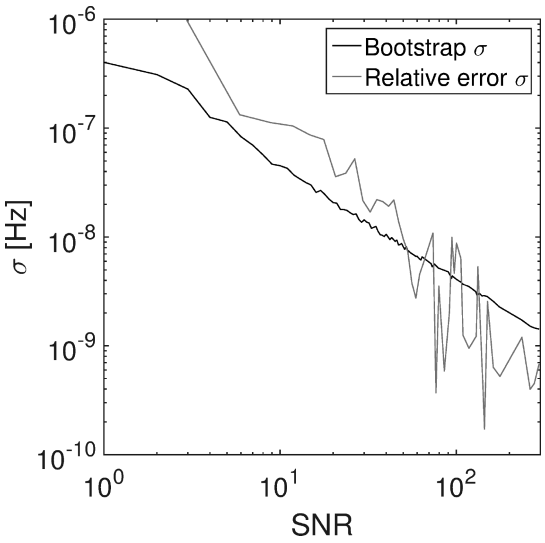

Figure 3. Estimated frequency bootstrap standard deviations (black line) and standard deviations estimated from frequency relative errors (grey line) versus SNRs for all experiments of ${}_0S_2$ multiplet.

et al. 1986; Widmer *et al.* 1992) we decide to fit only for degree 2 structure in eq. (23) (e.g. for the S40RTS model c_{20} is 6 times larger than c_{40}). The real coefficient γ_{20}^{mm} is calculated using the SC approximation theory (Dahlen & Tromp 1998). The results for c_{20} splitting function coefficients are shown in Fig. 4. They are calculated using non-weighted and weighted ordinary least squares method where weights are obtained from frequency standard deviations. To be consistent with observed values, we estimate the referent value (dashed line in Fig. 4) in the same manner. That is, from the synthetic singlet frequencies, calculated for a rotating earth model with lateral heterogeneities (S40RTS model), we first remove the effects of rotation and ellipticity using previously estimated a, b, and c parameters and next we fit just for c_{20} using relation (23). We are aware that this procedure may introduce a small bias in our referent value, however this is necessary for the value to be comparable with the measured ones [due to the diagonalization of the splitting matrix (Dahlen & Tromp 1998, p. 643)].

It is usually expected that the weighted solutions have overall better results than the non-weighted solutions, however Fig. 4 shows that the non-weighted solutions have similar results as weighted solutions specially when SNR is large and there are enough stations. As analysis progresses the non-weighted solutions become expectedly more deteriorated and off balanced. It is important to emphasize that each splitting function coefficients is obtained considering five biased frequencies and their associated standard deviations, however it seems that this effect is the most relevant for all non-weighted solutions and for weighted solutions in case of 10 stations, impacting both the structure coefficients and their associated standard deviations. The latter one is evident due to the fact that the standard deviations are in some occasions uncorrelated with the number of stations used or noise level amplitude (Fig. 4). Besides, it is encouraging that weighted c_{20} coefficients are close to and contain synthetic values within the estimated standard deviations for most of the experiments.

For comparison, let us consider now the case of the $m = -1$ singlet of ${}_2S_1$ multiplet. In this case the selected network contains stations where all three singlets of this multiplet are excited. For the analysis we use around 12-d-long time-series, which are zero padded to 72 d. The conclusions made for the ${}_0S_2$ multiplet are valid for this multiplet too. However, since this mode is poorly excited, the number of stations used in the stacking method becomes more critical than for ${}_0S_2$. During the experiments, not for all combinations of noise levels and station distributions singlets emerge during the stacking. The relative errors become quite large compared to ${}_0S_2$, ranging from 56 ppm (full network) to 323 ppm (selected network) for $\sigma_n = 1$ nm s^{-2} and phasor walkouts become much more complex with more loops indicating the higher noise influence (Fig. 5). During the analysis the number of used stations and associated records' SNR become more relevant. If we have two stacked signals with the same SNR, one of them could be built for the singlet with less excited amplitude, more stations and lower noise level, and the other stacked signal could be built for the singlet with better excited amplitude, less stations and higher noise level. Even though we end up with the same SNR, the ARFD80 method frequently fails in estimating the true frequency of stacked signal for the first case according to the R^2-test. Instead, the estimated frequencies are within the estimated standard deviations. Hence, it seems that ARFD80 is sensitive to the singlet starting amplitude to be stacked, that is the method is becoming unreliable when starting signals of singlets are buried in noise.

The performances of the used methods are tested on the higher frequency multiplet chain ${}_0T_5-{}_2S_2-{}_1S_3-{}_3S_1$, characterized by more prominent mode-mode interferences. It is known from the PREM model prediction that three modes ${}_2S_2$, ${}_1S_3$ and ${}_3S_1$ have very different Q-factors, but similar frequencies. Their Q-factors are approximately 96, 283 and 827 and frequencies 937.85, 939.83 and 943.95 μHz, respectively. The frequency difference between ${}_2S_2$ and ${}_3S_1$ is 6.1 μHz, which means that one needs 2.8 d to separate them in frequency domain, but the ${}_2S_2$ has the Q-cycle duration of 1.18 d, thus it is impossible to observe this mode without interference of ${}_3S_1$ and ${}_1S_3$. The quickly decaying ${}_2S_2$ mode is always predominated by slower decaying ${}_3S_1$ and ${}_1S_3$ modes. The observation of the ${}_1S_3$ mode has the similar issues due to the presence of ${}_3S_1$ mode. The measurement of ${}_3S_1$ is easier than the other two, however, the existing overlapping with ${}_1S_3$ mode introduces the contamination in the measurement of the ${}_3S_1$ splitting frequencies (Rogister 2003; Roult *et al.* 2010; Shen & Wu 2012; Ding & Shen 2013a,b; Chao & Ding 2014). Although there are several studies that measured the ${}_3S_1$ frequencies (Roult *et al.* 2010; Shen & Wu 2012; Ding & Shen 2013a,b; Chao & Ding 2014) only in the work by Chao & Ding (2014) the frequencies of all three modes ${}_2S_2$, ${}_1S_3$, ${}_3S_1$ are recovered using spherical harmonics stacking (Buland *et al.* 1979) in the SC approximation.

In this study, the aforementioned characteristics of this chain prevent us from conducting the station distribution analysis, while it is already difficult to find combination of stations where some singlets are excited above noise level. Once the best combination of stations has been found we perform noise analysis. The OSE method turns out to be less successful in generating isolated singlets than in the case of the ${}_0S_2$ multiplet chain and the phasor walkouts consistently generate beating patterns indicating the existence of neighboring singlets. The example for the relatively successful OSE resonance functions for ${}_1S_3$ multiplet is showed in Fig. 6. The figure is showing the comparison between the synthetic resonance functions obtained for the cases of the GC approximation considering ${}_0T_5-{}_2S_2-{}_1S_3-{}_3S_1$ and the SC approximation considering only ${}_1S_3$.

In our knowledge there is only one study that claims that ${}_2S_2$, ${}_1S_3$ and ${}_3S_1$ multiplets can be successfully recovered in the SC approximation. According to the recent work, OSE is successfully performing for the higher frequency modes in the SC approximation

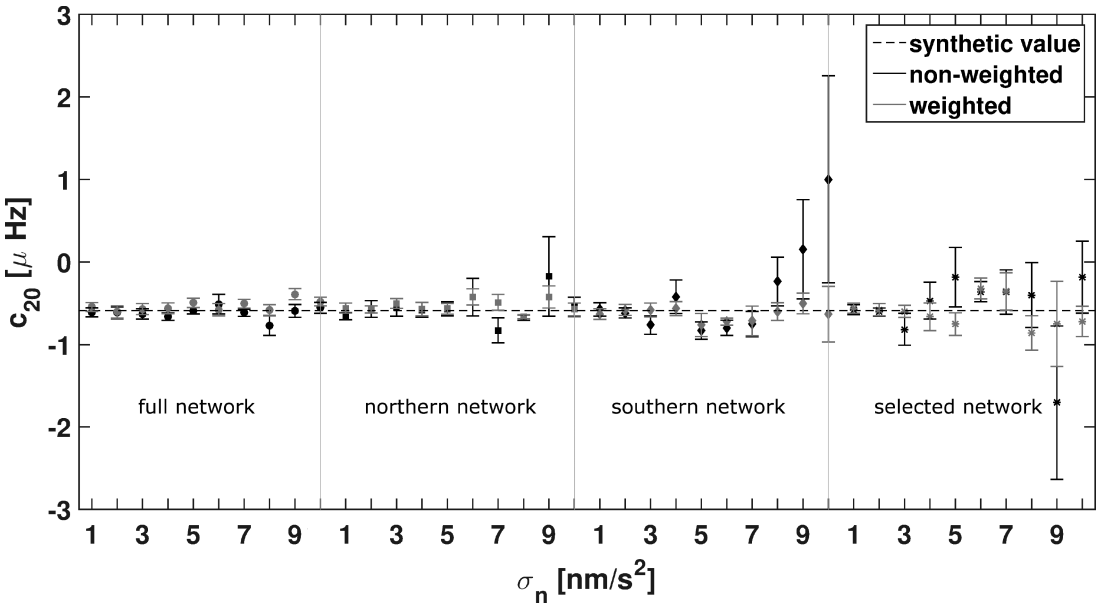

Figure 4. Splitting function coefficient c_{20} inverted from eigenfrequencies estimated in synthetic experiments for ${}_0S_2$ multiplet considering four different station distributions and ten noise levels.

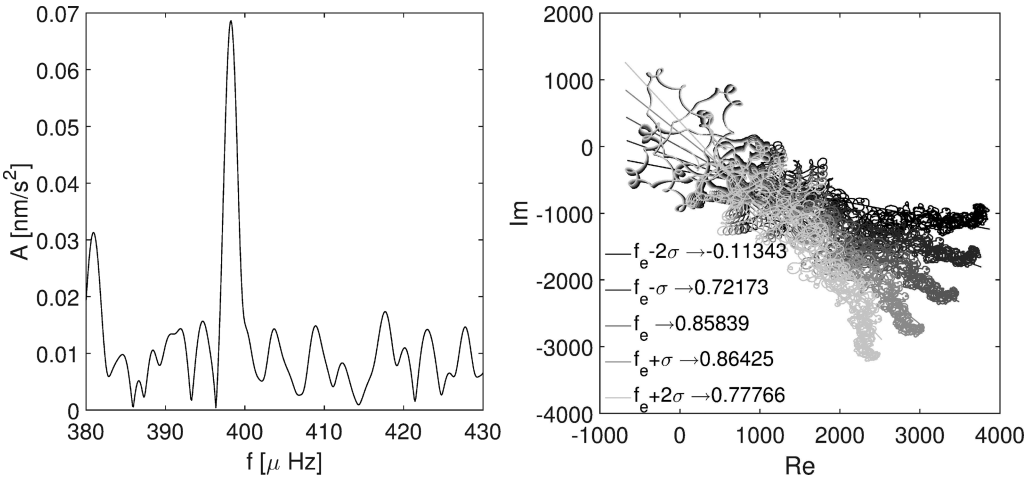

Figure 5. Fourier transform of OSE signal (left) with associated phasor walkout and R^2-test values (right) for five test frequencies with f_e being the estimates frequency of stacked signal on left and σ the standard deviation. Results correspond to selected network with $\sigma_n = 1$ nm s^{-2} for ${}_2S_1^{-1}$ singlet with SNR of 5.8.

(Zeng & Shen 2017) and in the GC approximation (Zeng & Shen 2018), however Fig. 6 seems to suggest us that OSE is not working sufficiently good for the coupled multiplets in the SC approximation. For the ${}_3S_1$ multiplet there are more studies which claim that this multiplet could be recovered in the SC approximation (Roult *et al.* 2010; Shen & Wu 2012; Ding & Shen 2013a,b).

4 FREQUENCIES AND STRUCTURE COEFFICIENTS FROM OBSERVATIONS

The methods and procedure conducted on synthetic data are now applied to the long-period seismometer and superconducting gravimeter data recorded after the six latest earthquakes of magnitude greater than 8.3. Seismogram database is built considering earthquakes with epicentres off the west coast of Sumatra island, Indonesia in 2004, off the coast of central Chile in 2010, off the Pacific coast of Tohoku, Japan in 2011, near the Indonesian province of Aceh in 2012, in the Sea of Okhotsk in western Pacific Ocean in 2013 and offshore from Illapel, Chile in 2015. The long-period STS-1 and STS-2 seismometer data are requested from IRIS service for LHZ channel with 125 stations in total. The RDseed software (http://ds.iris.edu/ds/nodes/dmc/software/downloads/rdseed/) is used to read SEED volumes and to retrieve SAC files. Next the Python package Obspy (Beyreuther *et al.* 2010; Megies *et al.* 2011; Krischer *et al.* 2015) is used to perform instrument deconvolution and finally the TSoft software (Van Camp & Vauterin 2005) is used to clear time-series of glitches, small gaps, etc. The accepted records are decimated to 60 s after low-pass filtering and cut 5 hr after the earthquake. 12-d-long time records (rarely 10 d due to deteriorated time-series) are then used. Seismometer data are not corrected for the atmospheric pressure effect because most of the barometric data are missing. Thus to be consistent during the analysis we skip this part. Gravimeter database is built considering the same earthquakes and downloaded from the IGETS website (http://igets.u-strasbg.fr/). The downloaded files are first merged,

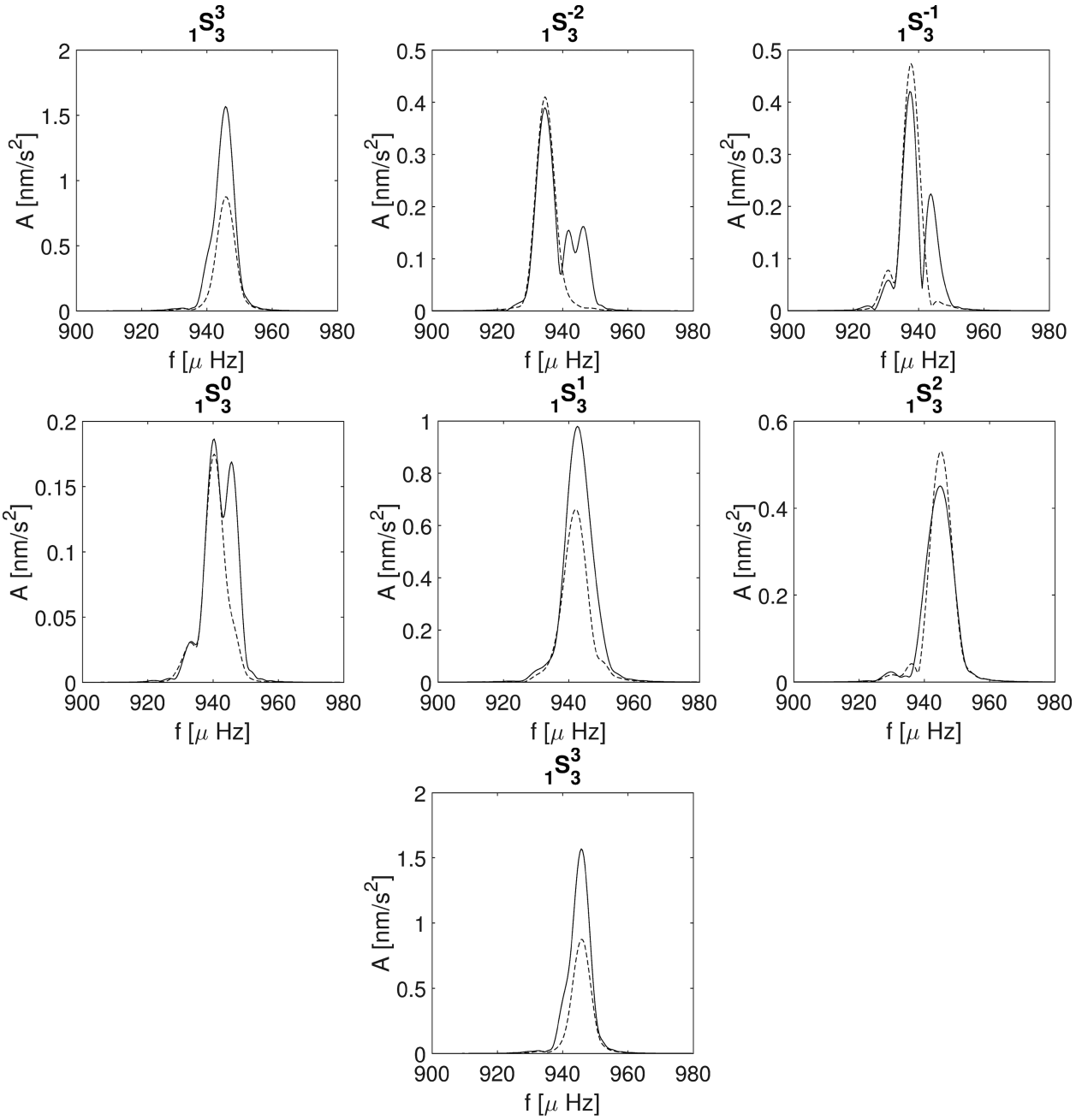

Figure 6. Comparison between the synthetic resonance functions obtained for group coupling (solid line) and self-coupling (dashed line) approximations for the singlets of $_1S_3$ multiplet.

then the instrumental and pressure scale factors are applied. Further, the gaps and spikes are fixed and finally we apply a high-pass least-squares filter with a cut-off frequency of 0.1 mHz to remove tides and other long-period effects. Atmospheric pressure effects are finally corrected using a nominal admittance of -3 nm s^{-2} hPa^{-1} (Zürn & Widmer-Schnidrig 1995).

In view of the synthetic experimental results we base our analysis of observed data on the SNRs. That is, after choosing the target singlet we estimate SNR from each seismogram and gravimetric record and organize them by decreasing SNR and eventually choose the first half of the records, with the highest SNRs. Chosen records are stacked and the SNR of this newly stacked signal is calculated, which we call $\mathrm{SNR_{base}}$. In the next steps, we either add or remove one signal from the chosen first half, perform again the stacking and calculation of the SNR of newly stacked signal, which we call $\mathrm{SNR_{base+1}}$ if we add signal and $\mathrm{SNR_{base-1}}$ if we remove signal. If either $\mathrm{SNR_{base+1}} > \mathrm{SNR_{base}}$ or $\mathrm{SNR_{base-1}} > \mathrm{SNR_{base}}$ we continue adding or removing n signals until either $\mathrm{SNR_{base+n}} < \mathrm{SNR_{base+(n-1)}}$ or $\mathrm{SNR_{base-n}} < \mathrm{SNR_{base-(n-1)}}$. Once we find the combination of records which produces the highest SNR for target singlet, we perform the calculation of complex frequencies with ARFD80 method and additionally validate our eigenfrequencies with the phasor walkout representation.

For the purpose of comparing observed results with synthetic results, we decide to analyse the same multiplets as in synthetic experiments, namely $_0S_2$, $_0S_3$, $_2S_1$ and $_3S_1$. There have been numerous

Table 1. Estimated eigenfrequencies (in μHz) and Q-factors with associated standard deviations, SNRs and number of stations N used in the stacking for ${}_0S_2$ and ${}_0S_3$ multiplets compared with published values. Methods applied: $*$ Lorentzian fitting, $\diamond$ nonlinear iterative least squares inversion, $\circ$ multitapers, $\triangledown$ ARFD80, $\triangleleft$ ensemble empirical mode decomposition, $\triangleright$ AR-spectrum and ARFD80.

${}_0S_2$	m	−3	−2	−1	0	1	2	3
This work	f	–	299.939	304.619	309.226	313.830	318.431	–
		–	±0.013	±0.009	±0.018	±0.012	±0.013	–
	Q	–	412.8	477.3	481.7	484.8	469.6	–
		–	±14.6	±13.7	±26.2	±17.8	±17.2	–
	SNR	–	39	53	27	39	40	–
	N	–	33	35	60	68	50	–
Roult *et al.* (2010)$*$	f	–	299.98	304.47	309.22	313.74	318.35	–
Deuss *et al.* (2011)$^\diamond$	f	–	299.93	304.63	309.28	313.86	318.40	–
Rosat *et al.* (2012)$*$	f	–	299.96	304.58	309.25	313.83	318.44	–
		–	±0.022	±0.051	±0.033	±0.046	±0.021	–
Häfner & Widmer-Schnidrig (2013)$^\circ$	f	–	299.948	304.612	309.269	313.840	318.429	–
		–	±0.009	±0.006	±0.016	±0.005	±0.009	–
	Q	–	486	473	512	488	498	–
		–	±16	±9	±29	±8	±15	–
Ding & Shen (2013b)$^\triangledown$	f	–	299.958	304.588	309.263	313.835	318.422	–
		–	±0.0081	±0.0046	±0.011	±0.0014	±0.0074	–
	Q	–	509.4	484.7	394.4	520.2	532.7	–
		–	±12.1	±9.3	±14.3	±8.1	±10.1	–
Shen & Ding (2014)$^\triangleleft$	f	–	299.994	304.618	309.278	313.865	318.424	–
		–	±0.011	±0.0078	±0.0091	±0.0077	±0.0096	–
Ding & Chao (2015b)$^\triangleright$	f	–	299.967	304.587	309.372	313.850	318.396	–
		–	±0.014	±0.0078	±0.050	±0.0069	±0.013	–
${}_0S_3$								
This work	f	461.705	464.132	466.459	468.724	470.781	472.838	474.724
		±0.023	±0.012	±0.020	±0.040	±0.020	±0.012	±0.015
	Q	418.2	410.1	439.9	298.5	433.9	438.0	408.6
		18.3	9.8	17.0	15.1	16.8	9.5	10.5
	SNR	25.14	46.10	27.95	17	27.37	49.38	40.66
	N	50	55	51	39	51	55	50
Roult *et al.* (2010)$*$	f	461.60	464.17	466.40	468.60	470.76	472.75	474.70
Rosat *et al.* (2012)$*$	f	461.67	464.24	466.39	–	470.84	472.66	474.74
		±0.054	±0.078	±0.036	–	±0.030	±0.078	±0.068
Ding & Shen (2013b)$^\triangledown$	f	461.623	464.219	466.535	468.549	470.657	472.843	474.831
		±0.0049	±0.0018	±0.0026	±0.0052	±0.0024	±0.0017	±0.0035
	Q	351.5	418.5	348.4	424.4	356.9	397.7	417.2
		±19.1	±9.2	±15.8	±22.5	±14.5	±10.0	±16.3
Shen & Ding (2014)$^\triangleleft$	f	461.618	464.161	466.397	468.650	470.734	472.816	474.727
		±0.018	±0.011	±0.026	±0.030	±0.022	±0.011	±0.019

studies dedicated to the estimates of the frequencies of the aforementioned multiplets, however the first study of all singlets of ${}_0S_2$ and ${}_0S_3$ multiplets was done by Buland *et al.* (1979), for the triplet ${}_2S_1$ by Rosat *et al.* (2003) and for ${}_3S_1$ by Chao & Gilbert (1980). Our process of searching for the stacked signal with the highest SNR resulted in obtaining different station distributions, different numbers of stations and different percentages of used earthquakes for each singlet. The estimated split frequencies, Q-factors and associated standard deviations along with the SNR of stacked signals and number of stations used are shown in Tables 1 and 2 together with published values. During the analysis the OSE method successfully isolates singlets and the phasor walkout graphs with the associated R^2-tests indicate that the true eigenfrequencies of stacked signals are within the estimated standard deviations. The example of stacked signals for ${}_2S_1$ and ${}_3S_1$ is shown in Fig. 7. It is important to note that the phasor walkout patterns for observed data and synthetic experiments are quite different. The walkout pattern for real data is characterized with repeated loops and twists, where the first characteristics indicate the presence of a harmonic function of close frequency with the dominant amplitude compared with the amplitude of tested signal and the second characteristics indicate the noise presence and the phase abruptness, which is expected due to the imperfections of the observed data, such as remaining glitches.

The new set of estimated eigenfrequencies with associated standard deviations is used to calculate a new axisymmetric splitting function coefficient of second order. As for the synthetic case, to be able to use eq. (23), we subtract the a, b, c parameters from the split eigenfrequencies. In this case a, b, c parameters are estimated by fitting the frequencies for the PREM model, calculated for a rotating ellipsoidal oceanless earth model (Rogister 2003). The results for non-weighted and weighted (where weights are frequency standard deviations) solution are shown in Table 3 and compared with published values. The splitting function coefficients from Häfner & Widmer-Schnidrig (2013) are estimated by subtracting a, b, c parameters as in this work.

At this stage it is difficult to thoroughly compare results of this study with previously published values, considering that all other studies used different methods and data to obtain the estimates. Thus, it is difficult to conclude which study is less biased and more

Table 2. Estimated eigenfrequencies (in μHz) and Q-factors with associated standard deviations, SNRs and number of stations N used in the stacking for $_2S_1$ and $_3S_1$ triplets compared with published values. Methods applied: $*$ Lorentzian fitting, $^\diamond$ nonlinear iterative least squares inversion, $^\times$ OSE and ARFD80, $^\triangledown$ ARFD80, $^\triangleleft$ ensemble empirical mode decomposition, $^\triangleright$ AR-spectrum and ARFD80, $^\otimes$ multistation experiment technique, † spherical harmonic stacking and ARFD80.

$_2S_1$	m	−1	0	1
This work	f	398.854	405.290	410.880
		±0.347	±0.316	±0.108
	Q	250.5	391.0	404.8
		98.9	237.6	91.0
	SNR	3.3	2.6	8.2
	N	34	23	34
Roult *et al.* (2010)$*$	f	397.70	403.94	410.63
Deuss *et al.* (2011)$^\diamond$	f	397.92	405.18	410.45
Rosat *et al.* (2012)$*$	f	398.10	–	410.82
		±0.98	–	±0.18
Ding & Shen (2013a)$^\times$	f	397.982	–	411.051
		±0.12	–	±0.055
Ding & Shen (2013b)$^\triangledown$	f	398.662	405.014	410.768
		±0.0085	±0.0027	±0.0012
	Q	365.9	448.3	385.6
		±20.3	±15.7	±11.8
Shen & Ding (2014)$^\triangleleft$	f	398.363	404.757	410.810
		±0.043	±0.039	±0.019
Ding & Chao (2015b)$^\triangleright$	f	398.174	404.955	410.806
		±0.2	±0.079	±0.064
$_3S_1$				
This work	f	942.565	944.570	945.661
		±0.019	±0.032	±0.005
	Q	873.2	649.5	903.9
		±31.3	±28.1	±9.5
	SNR	30.42	20.91	103
	N	82	12	82
Roult *et al.* (2010)$*$	f	942.56	944.19	945.79
Shen & Wu (2012)$^\otimes$	f	942.598	944.113	945.864
		±0.42	±0.27	±0.21
Ding & Shen (2013a)$^\times$	f	942.267	944.765	945.763
		±0.022	±0.051	±0.018
Ding & Shen (2013b)$^\triangledown$	f	942.426	944.713	945.612
		±0.0025	±0.0017	±0.0046
	Q	943.8	773.6	629.5
		±12.5	±10.1	±18.4
Chao & Ding (2014)†	f	942.57	944.20	945.76
		±0.028	±0.092	±0.034
	Q	801	625	650
		±20	±31	±22

relevant. Summarizing the information from the ten published studies used in Tables 1 and 2, we note that most studies used no more than three earthquakes in their analysis, where all of them used Sumatra earthquake from 2004. Next, the majority used around 11 stations, with the exception of Deuss *et al.* (2011) where they even used 300 spectra for some modes, Chao & Ding (2014) 96 records, Ding & Chao (2015b) 46 records and Roult *et al.* (2010) 247 records. However, it is not always clear whether all the records are used in the analysis of all target modes in their studies. Most of the studies used SG records, except (Deuss *et al.* 2011), Chao & Ding (2014) and Roult *et al.* (2010) who used only seismograms. Further, the methods for obtaining the modal parameters are different (see Tables 1 and 2). The estimation of standard errors is performed either by using the bootstrap method along with the weighted mean (Häfner & Widmer-Schnidrig 2013; Ding & Shen 2013b; Shen & Ding 2014), or just the bootstrap method (Ding & Chao 2015b), or using the error analysis from Dahlen (1982b) (Rosat *et al.* 2012) or the error analysis from Chao & Gilbert (1980) (Ding & Shen 2013a; Chao & Ding 2014) and the least-square fitting Shen & Wu (2012). Methodologically, only Ding & Shen (2013a) used the same process analysis as in this paper, however they used only 8 and 11 SG records for obtaining the estimates for $_2S_1$ and $_3S_1$, respectively, and thus their results may be more biased. Further, directly comparing the measured eigenfrequencies and their standard deviations we can conclude that not all measurements agree within their standard errors. Whether it is a problem in the small data sets within some studies or biases introduced with the methods, it is difficult to comment. We would have the complete information by comparing the SNR values and used number of stations for each estimate, unfortunately, this information is usually missing. Considering the error analysis, the error estimated by the bootstrap method is the statistical error, while the error analysis from Dahlen (1982b) and Chao & Gilbert (1980) give the formal errors. In the view of our synthetic tests, we show that for the most relevant SNRs the estimated error

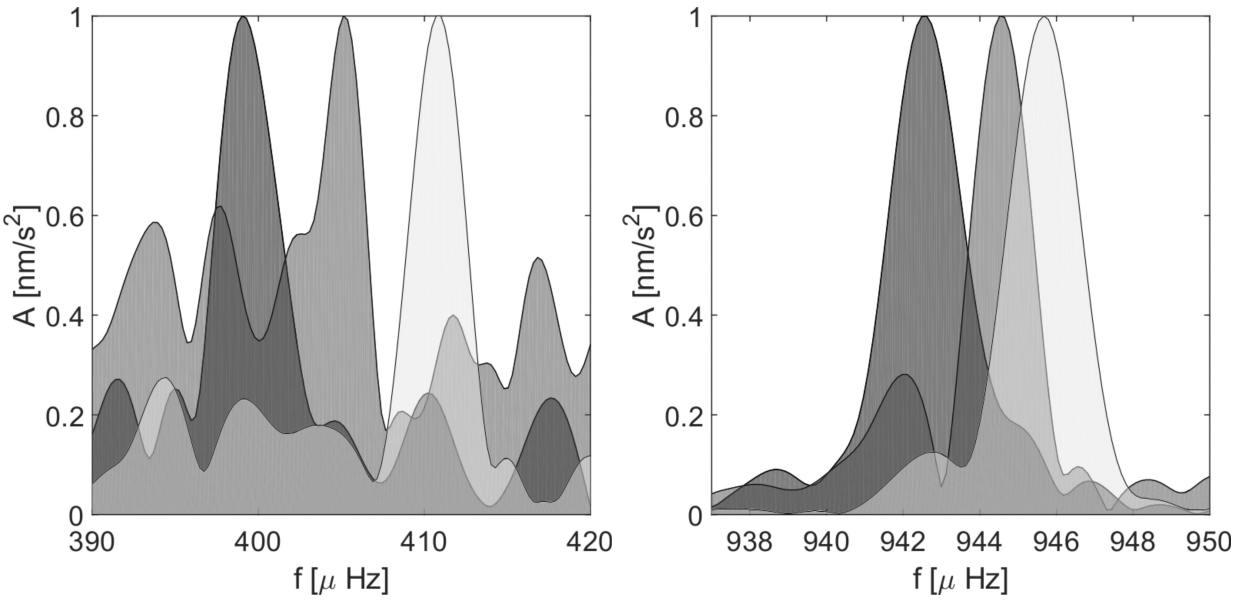

Figure 7. $_2S_1$ and $_3S_1$ singlets obtained using the OSE stacking method applied on 7 and 12-d-long time-series, respectively. The amplitude values are 0.08, 0.05, 0.20 nm s^{-2} for $_2S_1$ and 0.51, 0.25, 0.81 nm s^{-2} for $_3S_1$.

Table 3. Second-order axisymmetric structure coefficients for $_0S_2$ compared with published values. The structure coefficients are computed from the singlet frequencies (∗) or the nonlinear iterative least-squares inversion (◦).

References	c_{20} (μHz)
This work (weighted)∗	-0.7233 ± 0.0623
This work (non-weighted)∗	-0.7428 ± 0.0633
Ritzwoller *et al.* (1986)∗	-0.30 ± 0.65
Deuss *et al.* (2011)◦	-0.66 ± 0.32
Häfner & Widmer-Schnidrig (2013) (weighted)∗	-0.7404 ± 0.0466
Häfner & Widmer-Schnidrig (2013) (non-weighted)∗	-0.6902 ± 0.0398

by the bootstrap method is understimated. However, one needs to remember that this error reflects the bias from the used OSE and ARFD80 methods. For our tests it is shown that using around and more than 50 stations in the analysis is generating satisfying estimates. Compared to the frequency estimates, the Q-factor estimates are more scattered, but this has been already known since it is more difficult to measure amplitude than frequency of a mode. The structure coefficient estimates are satisfying since they agree well with published values, bearing in mind that they are estimated using frequencies and their standard deviations obtained for different station distributions. It is true that we can always argue that our estimates are close to published values, however the question of accuracy remains. Only a truthfully comparison of all relevant methods in the same condition can tell which study is the most precise.

5 CONCLUSIONS

We have quantified the effects of different station distributions and noise levels on the estimation of eigenfrequencies, Q-factors and splitting function coefficients of the gravest seismic modes. A specific protocol was used: the OSE method for stacking signals, the ARFD80 method for obtaining the estimates of harmonic parameters and finally the phasor walkout together with R^2-test for validating the estimated eigenfrequencies. The methods were tested on synthetically generated data and on observations. Synthetic experiments have shown that when noise is gradually added, the SNR of the stacked signals decreases and standard deviations of estimated frequencies and Q-factors increase, being overall more scattered around the true synthetic value. Results are foremost when more records are included in the stacking. More importantly it has been shown that OSE method is sensitive to different station distributions under the noise influence. Furthermore, the performances of ARFD80 method become deteriorated when the stacking signal is obtained with the less excited signals, that is, when the input signals have SNR ≤ 2. However, even for those cases the estimates are within the standard deviations. Moreover, it turns out that the standard deviations calculated with the bootstrap method are not sufficient to include all biases introduced with the methods, that is to say our standard deviations are for most cases underestimated. Thus, even though we have a good precision on our estimates, the accuracy can be poor. This analysis showed that we do not need *a priori* model to estimate structure coefficients, but results may be biased. Finally, we have proposed new eigenfrequency and Q-factor estimates for $_0S_2$, $_0S_3$, $_2S_1$ and $_3S_1$ and also new estimates for the axisymmetric degree-2 structure coefficient from $_0S_2$ eigenfrequency measurements. The results are in good agreement with previously published values, even though the methods are different and thus presumably introduce different biases into the estimates. With the synthetic test performed, we are confident in our new estimates,

specially for the ones obtained with more than 50 stations. Finally, future tomographic models that use the splitting function measurements should take into account the existing biases mentioned in this work.

ACKNOWLEDGEMENTS

The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms, related metadata, and/or derived products used in this study. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscienceand EarthScope (SAGE) Proposal of the National Science Foundation under Cooperative Agreement EAR-1261681. We also thank the managers of superconducting gravimeters for sharing their data within the IGETS (former GGP) (data downloaded from http://isdc.gfz-potsdam.de) (Voigt *et al.* 2016). We are grateful to David Al-Attar for his detailed and careful remarks on this manuscript and to an anonymous reviewer for helpful comments.

REFERENCES

Akbarashrafi, F., Al-Attar, D., Deuss, A., Trampert, J. & Valentine, A.P., 2017. Exact free oscillation spectra, splitting functions and the resolvability of Earth's density structure, *Geophys. J. Int.,* **213**(1), 58–76.

Al-Attar, D., Woodhouse, J.H. & Deuss, A., 2012. Calculation of normal mode spectra in laterally heterogeneous earth models using an iterative direct solution method, *Geophys. J. Int.,* **189**(2), 1038–1046.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J., 2010. ObsPy: a python toolbox for seismology, *Seismol. Res. Lett.,* **81**(3), 530–533.

Buland, R., Berger, J. & Gilbert, F., 1979. Observations from the IDA network of attenuation and splitting during a recent earthquake, *Nature,* **277,** 358–362.

Chao, B.F., 1990. On the use of maximum entropy/autoregressive spectrum in harmonic analysis of time series, *Pure appl. Geophys.,* **134**(2), 303–311.

Chao, B.F. & Ding, H., 2014. Spherical harmonic stacking for the singlets of Earth's normal modes of free oscillation, *Geophys. Res. Lett.,* **41**(15), 5428–5435.

Chao, B.F. & Gilbert, F., 1980. Autoregressive estimation of complex eigenfrequencies in low frequency seismic spectra, *Geophys. J. Int.,* **63**(3), 641–657.

Courtier, N.,et &al., 2000. Global superconducting gravimeter observations and the search for the translational modes of the inner core, *Phys. Earth planet. Inter.,* **117,** 3–20.

Dahlen, F.A., 1968. The normal modes of a rotating, elliptical earth, *Geophys. J. R. astr. Soc.,* **16**(4), 329–367.

Dahlen, F.A., 1969. The normal modes of a rotating, elliptical Earth-II near-resonance multiplet coupling, *Geophys. J. R. astr. Soc.,* **18**(4), 397–436.

Dahlen, F.A., 1974. Inference of the Lateral Heterogeneity of the Earth from the eigenfrequency spectrum: a linear inverse problem, *Geophys. J. Int.,* **38**(1), 143–167.

Dahlen, F.A., 1982a. Attenuation in the earth at low frequencies, *Geophys. J. Int.,* **69**(2), 537–549.

Dahlen, F.A., 1982b. The effect of data windows on the estimation of free oscillations parameters, *Geophys. J. R. astr. Soc.,* **69,** 537–549.

Dahlen, F.A. & Sailor, R.V., 1979. Rotational and elliptical splitting of the free oscillations of the Earth, *Geophys. J. Int.,* **58**(3), 609–623.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology,* Princeton University Press.

Deuss, A., Ritsema, J. & Van Heijst, H., 2011. Splitting function measurements for Earth's longest period normal modes using recent large earthquakes, *Geophys. Res. Lett.,* **38**(4), 1–5.

Deuss, A., Ritsema, J. & van Heijst, H., 2013. A new catalogue of normal-mode splitting function measurements up to 10 mHz, *Geophys. J. Int.,* **193**(2), 920–937.

Deuss, A. & Woodhouse, J.H., 2001. Theoretical free-oscillation spectra: the importance of wide band coupling, *Geophys. J. Int.,* **146**(3), 833–842.

Deuss, A. & Woodhouse, J.H., 2004. Iteration method to determine the eigenvalues and eigenvectors of a target multiplet including full mode coupling, *Geophys. J. Int.,* **159**(1), 326–332.

Ding, H. & Chao, B.F., 2015a. Data stacking methods for isolation of normal-mode singlets of Earth's free oscillation: Extensions, comparisons, and applications, *J. geophys. Res.,* **120**(7), 5034–5050.

Ding, H. & Chao, B.F., 2015b. Detecting harmonic signals in a noisy time-series: The z-domain autoregressive (AR-z) spectrum, *Geophys. J. Int.,* **201**(3), 1287–1296.

Ding, H. & Chao, B.F., 2015c. The Slichter mode of the Earth: Revisit with optimal stacking and autoregressive methods on full superconducting gravimeter data set, *J. geophys. Res.,* **120**(10), 7261–7272.

Ding, H. & Chao, B.F., 2016. Solid pole tide in global GPS and superconducting gravimeter observations: Signal retrieval and inference for mantle anelasticity, *Earth planet. Sci. Lett.,* **459,** 244–251.

Ding, H. & Shen, W.-B., 2013a. Search for the Slichter modes based on a new method: Optimal sequence estimation, *J. geophys. Res.,* **118**(9), 5018–5029.

Ding, H. & Shen, W.-B., 2013b. Determination of the complex frequencies for the normal modes below 1mHz after the 2010 Maule and 2011 Tohoku earthquakes, *Ann. Geophys.,* **56**(5), 1–15.

Draper, N.R. & Smith, H., 1998. *Applied Regression Analysis,* 3rd edn, Wiley-Interscience..

Dziewonski, A.M. & Anderson, D.L., 1981. Preliminary reference earth model, *Phys. Earth planet. Inter.,* **25**(4), 297–356.

Dziewonski, A.M., Chou, T. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.,* **86,** 2825–2852.

Ekström, G., Nettles, M. & Dziewonski, A.M., 2012. The global CMT project 2004-2010: centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.,* **200–201,** 1–9.

Giardini, D., Li, X.-D. & Woodhouse, J.H., 1987. Three-dimensional structure of the Earth from splitting in free-oscillation spectra, *Nature,* **325,** 405–411.

Giardini, D., Li, X.-D. & Woodhouse, J.H., 1988. Splitting Functions of Long-Period Normal Modes of the Earth, *J. geophys. Res.,* **93**(B11), 13716.

Gilbert, F., 1971. The diagonal sum rule and averaged eigenfrequencies, *Geophys. J. Int.,* **23**(1), 119–123.

Guoming, X., Knopoff, L. & Zürn, W., 1983. Variations of period and Q of free oscillations due to mode overlap, *Geophys. J. Int.,* **72**(3), 709–719.

Hinderer, J., Crossley, D. & Jensen, O., 1995. A search for the Slichter triplet in superconducting gravimeter data, *Phys. Earth planet. Inter.,* **90**(3–4), 183–195.

Häfner, R. & Widmer-Schnidrig, R., 2013. Signature of 3-D density structure in spectra of the spheroidal free oscillation 0S2, *Geophys. J. Int.,* **192,** 285–294.

Irving, J. C. E., Deuss, A. & Andrews, J., 2008. Wide-band coupling of Earth's normal modes due to anisotropic inner core structure, *Geophys. J. Int.,* **174**(3), 919–929.

Irving, J. C. E., Deuss, A. & Andrews, J., 2009. Normal mode coupling due to hemispherical anisotropic structure in Earth's inner core, *Geophys. J. Int.,* **178**(2), 962–975.

Ishii, M. & Tromp, J., 1999. Normal-mode and free-air gravity constraints on lateral variations in velocity and density of the Earth's mantle, *Science,* **285,** 1231–1236.

Ishii, M. & Tromp, J., 2001. Even-degree lateral variations in the Earth's mantle constrained by free oscillations and the free-air gravity anomaly, *Geophys. J. Int.,* **145**(1), 77–96.

Koelemeijer, P., Deuss, A. & Ritsema, J., 2013. Observations of core-mantle boundary Stoneley modes, *Geophys. Res. Lett.,* **40**(11), 2557–2561.

Koelemeijer, P., Ritsema, J., Deuss, A. & Van Heijst, H.-J., 2016. SP12RTS: a degree-12 model of shear- and compressional-wave velocity for Earth's mantle, *Geophys. J. Int.,* **204**(2), 1024–1039.

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C. & Wassermann, J., 2015. ObsPy: a bridge for seismology into the scientific Python ecosystem, *Comput. Sci. Discovery,* **8**(1), 1–17.

Kuo, C. & Romanowicz, B., 2002. On the resolution of density anomalies in the Earth's mantle using spectral fitting of normal-mode data, *Geophys. J. Int.,* **150**(1), 162–179.

Li, X.-D., Giardini, D. & Woodhouse, J.H., 1991. Large-scale three-dimensional even-degree structure of the Earth from splitting of long-period normal modes, *J. geophys. Res.,* **96**(B1), 551–557.

Masters, G. & Gilbert, F., 1983. Attenuation in the earth at low frequencies, *Phil. Trans. R. Soc. Lond., A,* **308**(1504), 479–522.

Masters, G., Laske, G. & Gilbert, F., 2000. Autoregressive estimation of the splitting matrix of free-oscillation multiplets, *Geophys. J. Int.,* **141**(1), 25–42.

Megies, T., Beyreuther, M., Barsch, R., Krischer, L. & Wassermann, J., 2011. ObsPy - what can it do for data centers and observatories?, *Ann. Geophys.,* **54**(1), 47–58.

Moulik, P. & Ekström, G., 2014. An anisotropic shear velocity model of the Earth's mantle using normal modes, body waves, surface waves and long-period waveforms, *J. geophys. Int.,* **199**(3), 1713–1738.

Pachhai, S., Tkalčić, H. & Masters, G., 2016. Estimation of splitting functions from Earth's normal mode spectra using the neighbourhood algorithm, *Geophys. J. Int.,* **204**(1), 111–126.

Resovsky, J.S. & Ritzwoller, M.H., 1998. New and refined constraints on three-dimensional Earth structure from normal modes below 3 mHz, *J. geophys. Res.,* **103**(B1), 783–810.

Resovsky, J.S. & Ritzwoller, M.H., 1999. Regularization uncertainty in density models estimated from normal mode data, *Geophys. Res. Lett.,* **26**(15), 2319–2322.

Ritsema, J., Deuss, A., Van Heijst, H.J. & Woodhouse, J.H., 2011. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.,* **184**(3), 1223–1236.

Ritzwoller, M., Masters, G. & Gilbert, F., 1986. Observations of anomalous splitting and their interpretation in terms of aspherical structure, *J. geophys. Res.,* **91**(B10), 10 203–10 228.

Ritzwoller, M., Masters, G. & Gilbert, F., 1988. Constraining aspherical structure with low-degree interaction coefficients: application to uncoupled multiplets, *J. geophys. Res.,* **93**(B6), 6369–6396.

Rogister, Y., 2003. Splitting of seismic-free oscillations and of the Slichter triplet using the normal mode theory of a rotating, ellipsoidal earth, *Phys. Earth planet. Inter.,* **140**(1–3), 169–182.

Romanowicz, B., 2001. Can we resolve 3D density heterogeneity in the lower mantle?, *Geophys. Res. Lett.,* **28**(6), 1107–1110.

Rosat, S. & Hinderer, J., 2011. Noise levels of superconducting gravimeters: updated comparison and time stability, *Bull. seism. Soc. Am.,* **101**(3), 1233–1241.

Rosat, S., Hinderer, J., Crossley, D. & Rivera, L., 2003. The search for the Slichter mode: comparison of noise levels of superconducting gravimeters and investigation of a stacking method, *Phys. Earth planet. Inter.,* **140,** 183–202.

Rosat, S., Sato, T., Imanishi, Y., Hinderer, J., Tamura, Y., McQueen, H. & Ohashi, M., 2005. High-resolution analysis of the gravest seismic normal modes after the 2004 Mw = 9 Sumatra earthquake using superconducting gravimeter data, *Geophys. Res. Lett.,* **32**(13), L13304.

Rosat, S., Sato, T., Imanishi, Y., Hinderer, J., Tamura, Y., McQueen, H. & Ohashi, M., 2012. Correction to "High-resolution analysis of the gravest seismic normal modes after the 2004 $M_w = 9$ Sumatra earthquake using superconducting gravimeter data", *Geophys. Res. Lett.,* **39**(22), L22601.

Roult, G., Roch, J. & Clévédé, E., 2010. Observation of split modes from the 26th December 2004 Sumatra–Andaman mega-event, *Phys. Earth planet. Inter.,* **179,** 45–59.

Shen, W.-B. & Ding, H., 2014. Observation of spheroidal normal mode multiplets below 1 mHz using ensemble empirical mode decomposition, *Geophys. J. Int.,* **196**(3), 1631–1642.

Shen, W.-B. & Wu, B., 2012. A case study of detecting the triplet of 3S1 using superconducting gravimeter records with an alternative data preprocessing technique, *Ann. Geophys.,* **55**(2), 293–300.

Slichter, L.B., 1961. The fundamental free mode of the Earth's inner core, *Proc. Natl. Acad. Sci. USA,* **47**(2), 186–190.

Smylie, D.E., 1992. The inner core translational triplet and the density near Earth's center, *Science,* **255**(1), 1678–1682.

Trampert, J., Deschamps, F., Resovsky, J. & Yuen, D., 2004. Probabilistic tomography maps chemical heterogeneities throughout the lower mantle, *Science,* **306,** 853–856.

Van Camp, M. & Vauterin, P., 2005. Tsoft: graphical and interactive software for the analysis of time series and Earth tides, *Comput. Geosci.,* **31**(5), 631–640.

Vanícek, P., 1969. Approximate spectral analysis by least-squares fit, *Astrophys. Space Sci.,* **4**(4), 387–391.

Voigt, C. *et al.*, 2016. Report on the Data Base of the International Geodynamics and Earth Tide Service (IGETS), *(Scientific Technical Reports STR - Data; 16/08), Potsdam : GFZ German Research Centre for Geosciences,* 24p.

Widmer-Schnidrig, R., 2003. What can Superconducting Gravimeters contribute to normal mode seismology, *Bull. seism. Soc. Am.,* **93**(3), 1370–1380.

Widmer, R., Masters, G. & Gilbert, F., 1992. Observably split multiplets data analysis and interpretation in terms of large scale aspherical structure, *Geophys. J. Int.,* **111**(3), 559–576.

Woodhouse, J.H., 1980. The coupling and attenuation of nearly resonant multiplets in the Earth's free oscillation spectrum, *Geophys. J. R. astr. Soc.,* **61**(2), 261–283.

Woodhouse, J.H., 1983. The joint inversion of seismic waveforms for lateral variations in earth structure and earthquake source parameters, *Geophys. J. R. astr. Soc.,* **85,** 366–397.

Woodhouse, J.H., 1988. The calculation of eigenfrequencies and eigenfunctions of the tree oscillations of the Earth and the Sun, in *Seismological Algorithms, Computational Methods and Computer Programs,* pp. 321–370, ed. Doornbos, D.J., Academic Press.

Woodhouse, J.H. & Dahlen, F.A., 1978. The effect of a general aspherical perturbation on the free oscillations of the Earth, *Geophys. J. R. astr. Soc.,* **53**(2), 335–354.

Woodhouse, J.H. & Giardini, D., 1985. Inversion for the splitting function of isolated low order normal mode multiplets, *EOS, Trans. Am. geophys. Un.,* **66,** 300.

Woodhouse, J.H. & Girnius, T.P., 1982. Surface waves and free oscillations in a regionalized earth model, *Geophys. J. R. astr. Soc.,* **68**(3), 653–673.

Yang, H.-Y. & Tromp, J., 2015. Synthetic free-oscillation spectra: an appraisal of various mode-coupling methods, *Geophys. J. Int.,* **203**(2), 1179–1192.

Zeng, S.-Y. & Shen, W.-B., 2017. Stripping the narrowband coupled multiplets based on seismic records using the OSE, *Phys. Earth planet. Inter.,* **272,** 58–67.

Zeng, S.-Y. & Shen, W.-B., 2018. Observations of the singlets of higher-degree modes based on the OSE, *J. Earth Sci.,* 1–11.

Zürn, W. & Rydelek, P.A., 1994. Revisiting the Phasor-Walkout method for detailed investigation of harmonic signals in time series, *Surveys in Geophysics,* **15**(4), 409–431.

Zürn, W. & Widmer-Schnidrig, R., 1995. On noise reduction in vertical seismic records below 2 mHz using local barometric pressure, *Geophys. Res. Lett.,* **22**(24), 3537–3540.

APPENDIX : THE PHASOR WALKOUT

The important reference for the phasor walkout method is paper by Zürn & Rydelek (1994), where they revisited the method and performed several synthetic tests. This appendix is a complementary work to their paper.

For the signal defined as

$$x_j = A_k \cos(2\pi f_k (j-1)\Delta t + \theta_k) \mathrm{e}^{-\alpha_k (j-1)\Delta t}, \quad j = 1, \ldots, N, \tag{A1}$$

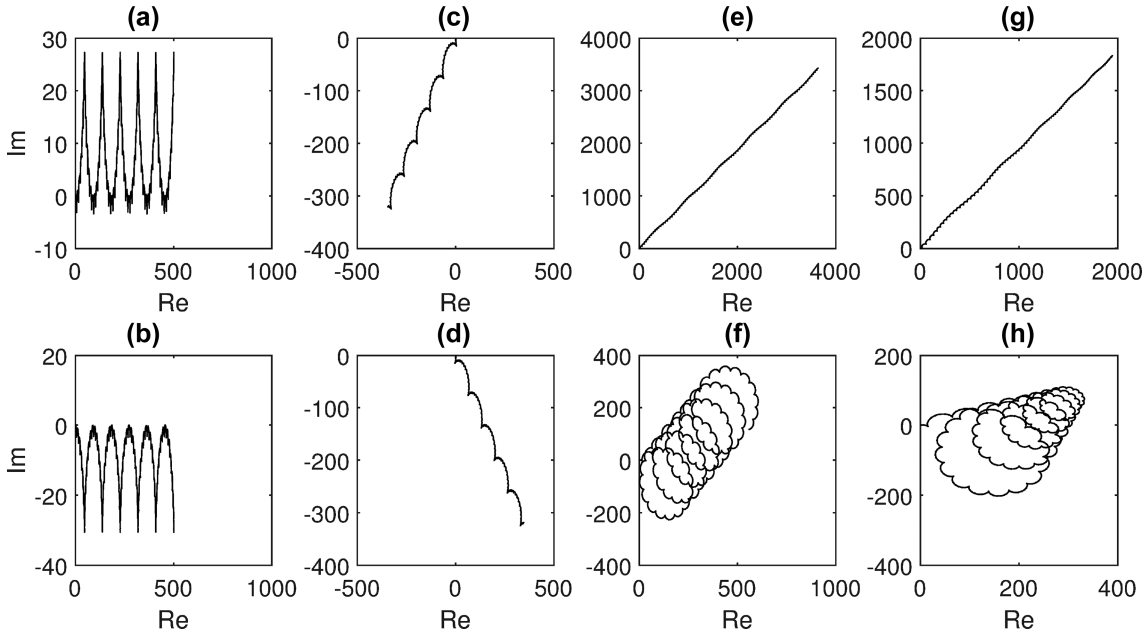

Figure A1. The phasor walkout for a signal with two harmonic functions of frequencies $f_1 = 9$ mHz and $f_2 = 10.1$ mHz for several different cases. From (a) to (g) the tested frequency is $f_s = f_1$, while for (b) to (h) $f_s = f_2$. For cases (a) and (b), the starting signal is defined by (A3) with $A_1 = A_2 = 1$ [nm s^{-2}], $\theta_1 = \theta_2 = 0$. In (c) and (d), phases $\theta_1 = 0.75$ and $\theta_2 = -0.75$ are added. In (e) and (f), phases are the same $\theta_1 = \theta_2 = 0.75$ and amplitudes are added with $A_1 = 10$ [nm s^{-2}] and $A_2 = 1$ [nm s^{-2}]. In (g) and (h), phases and amplitudes stay the same but a decay rate defined with $\exp\left(-\frac{\pi f_1}{100}\mathrm{i}\Delta t\right)$ is added.

the phasor walkout is defined as

$$p_j(f_s = f_k) = \frac{1}{2}A_k \mathrm{e}^{-\alpha_k(j-1)\Delta t}\{[\cos(\theta_k) + \cos(4\pi f_k(j-1)\Delta t + \theta_k)] + \mathrm{i}[\sin(\theta_k) - \sin(4\pi f_k(j-1)\Delta t + \theta_k)]\} \quad \text{(A2)}$$

and for constants $A_k > 0$, $\theta_k > 0$ and condition $|\sin(\theta_k)| > |\sin(4\pi f_k(j-1)\Delta t + \theta_k)|$, the phasor walkout is placed in the first quadrant of the Cartesian 2-D system, where the x-axis is a real axis and the y-axis is an imaginary axis. In other case, $|\sin(\theta_k)| < |\sin(4\pi f_k(j-1)\Delta t + \theta_k)|$, the phasor would be placed in the fourth quadrant of the Cartesian system. This is also true if $\theta_k < 0$. Both axes have constant value in terms of phase angle θ_k, which implies that the phasor walkout is progressing in the straight inclined line. If for some reason the amplitude is negative, which can be the case due to stacking, the combined effect of amplitude and phase angle defines the quadrant of the Cartesian system in which the phasor walkout is placed. That is, besides the first and fourth quadrant the phasor walkout could also be placed in the second and the third quadrants. The mentioned effects are demonstrated using the signal containing two harmonic functions of close frequencies, called 'beats'. In this case the signal is defined as

$$x_j = A_1\cos(2\pi f_1 j\Delta t + \theta_1) + A_2\cos(2\pi f_2 j\Delta t + \theta_2), \quad \text{(A3)}$$

where $f_1 \approx f_2$. The study is carried out by calculating the phasor walkout for tested frequencies $f_s = f_1$ and $f_s = f_2$ for several different cases:

(i) setting $\theta_1 > 0$, $\theta_2 < 0$, $|\theta_1| = |\theta_2|$, $|A_1| = |A_2| > 0$;
(ii) setting $|\theta_1| = |\theta_2| > 0$, $|A_1| > |A_2| > 0$;
(iii) setting $|\theta_1| = |\theta_2| > 0$, $|A_1| > |A_2| > 0$ and decay rate $\mathrm{e}^{\left(-\frac{\pi f_1}{Q}\mathrm{i}\Delta t\right)}$ to both harmonic functions.

The results are shown in Fig. A1. The phasor walkout for the basic signal defined in eq. (A3) with $A_1 = A_2 = 1$ [nm s^{-2}], $\theta_1 = \theta_2 = 0$ and tested for the frequencies of both harmonic functions is shown in (a) for $f_s = f_1$ and in (b) for $f_s = f_2$. It is important to note the regularity of this walkout, which is progressive advantage due to the fact that the tested frequencies are the same as the frequencies of one of the harmonic functions within the signal. In the next example, in panels (c) and (d), study case (i) is shown. Added phases cause the inclination of the walkout to have either positive imaginary axis for $\theta > 0$ (case c) or negative imaginary axis for $\theta < 0$ (case d). In panels (e) and (f), study case (ii) is shown. Now, even though both cases show linearity, the phasor walkout of the harmonic function with smaller amplitude (f) has slower phase change and additional loops, while the harmonic function with larger amplitude has rapid phase changes (e). The final study case (iii) is shown in panels (g) and (h). Adding decay rate is affecting the phasor walkout's amplitude causing the final walkout to have cone shape. In conclusion, the phasor walkout of the two harmonic functions of similar frequencies heavily depends on the amplitude ratio of the harmonic functions within the signal. Having more similar amplitudes will cause walkouts to have straight lines and thus enabling easier conclusion about the existence of particular harmonic function in the signal.

Next, we simulate signal containing five ${}_0S_2$ singlets, where each singlet is defined as (A1) with $\theta_k = 0$. Therefore, in case of ${}_0S_2$ mode k is 5. The signal is 15-d-long with a sampling rate of 60 s. We simulated ten cases which parameters are summarized in Table A1. Results are shown in Figs. A2 and A3. The frequency used for calculating the walkout in each case is marked with the asterisk, for example, in case j, we tested for the ${}_0S_2^{-1}$ singlet. Compared to the previous case the differences between singlets frequencies are about 100 times smaller introducing a considerable effect of the coupling between the singlets—the walkouts become more complex. In the experiments from a to e, we fix the same amplitude for all singlets and gradually increase the number of singlets in the signal from case a having one singlet, ${}_0S_2^{-2}$, to case e having all five singlets. Adding more singlets is affecting the phasor walkout of the first singlet obviously. While in the first case the walkout is simple cone-like, in

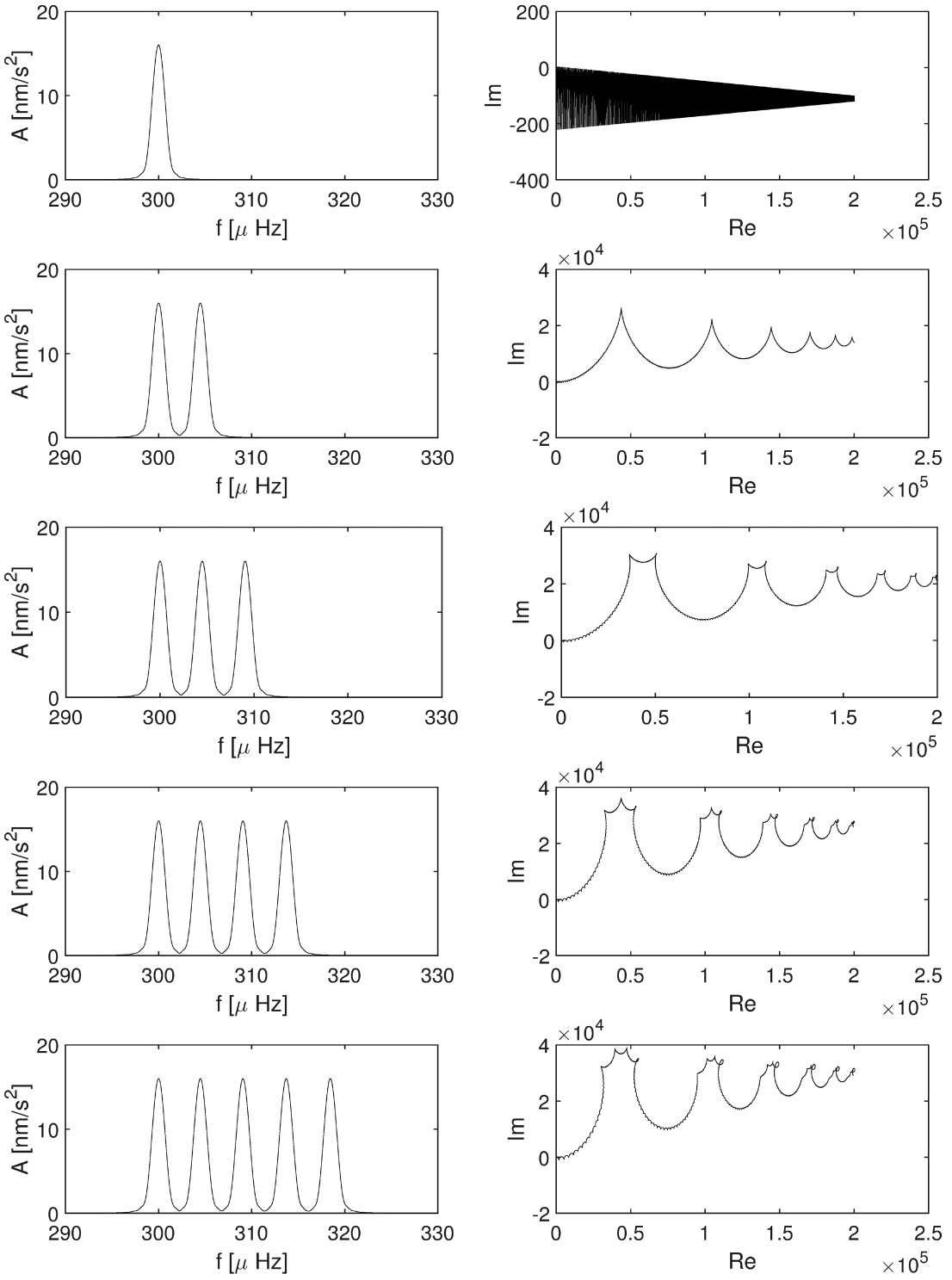

Figure A2. Synthetic signals (left) and appropriate phasor walkouts (right) for the experiment setups *a*, *b*, *c*, *d*, and *e* from Table A1. In all setups the phasor walkout is calculated for the ${}_0S_2^{-2}$ singlet.

other cases the number and complexity of loops are more prominent. In the experiments from *f* to *j* we change the amplitudes of singlets and test for different singlets. In case *f*, we calculate the walkout for the ${}_0S_2^0$ singlet and even though the signal is the same as in case *e* the walkout output is much different. At the same time it is also more difficult to argue about the straightness even though a regularity exists. Similar conclusions are valid for the cases *h* and *i*. We can claim that regularity exists, however, the straightness of the walkout is completely deteriorated. Therefore, it is reasonable to conclude that small singlets in close vicinity of singlets with higher amplitude are completely underdominated, resulting in phasor walkouts which are not straight but adopt loop-like shapes.

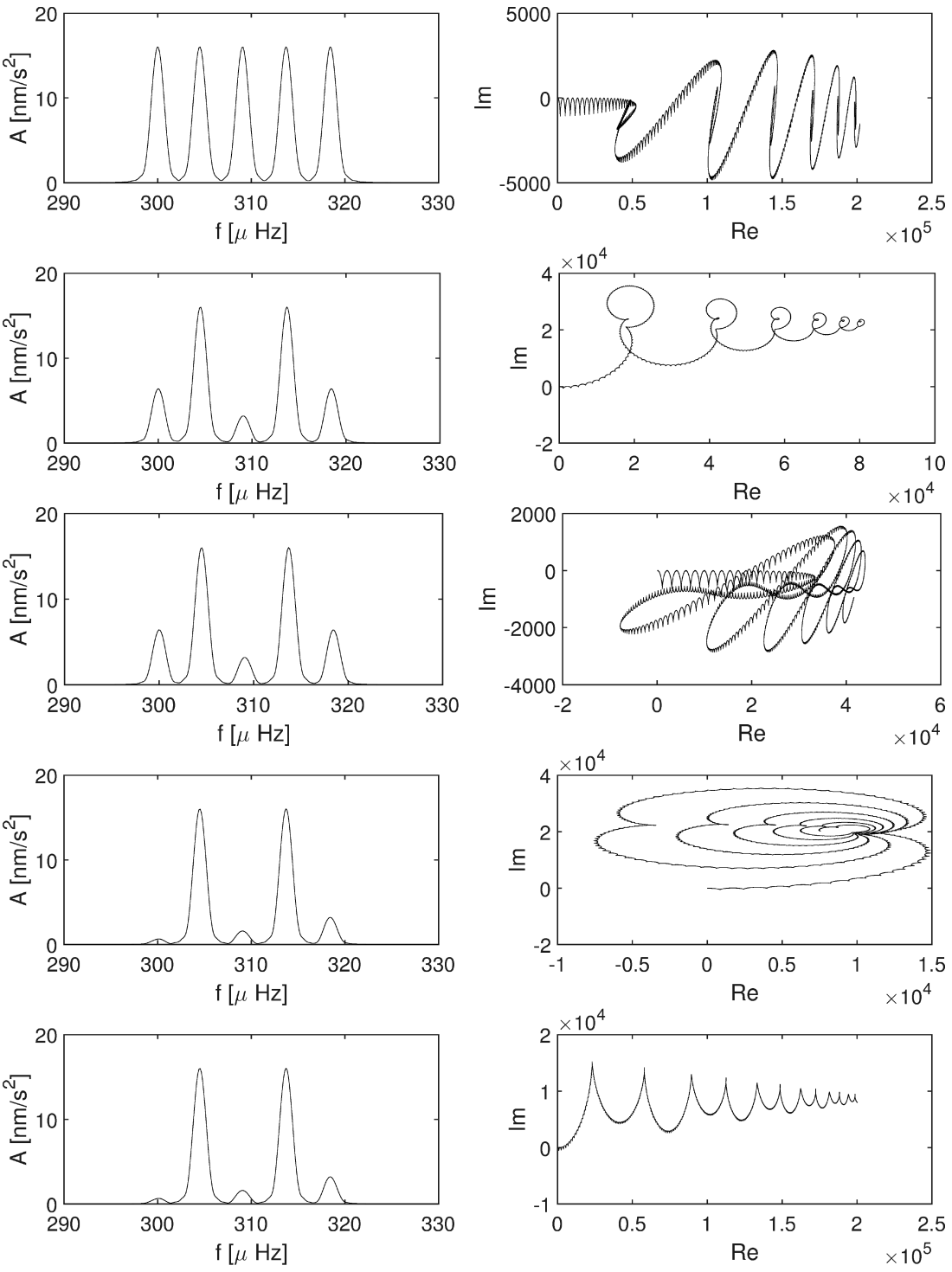

Figure A3. Synthetic signals (left) and appropriate phasor walkouts (right) for the experiment setups *f*, *g*, *h*, *i*, and *j* from Table A1. In setups *f* and *h* the phasor walkout is calculated for the ${}_0S_2^0$ singlet, in *g* and *i* for ${}_0S_2^{-2}$ singlet and in *j* for ${}_0S_2^{-1}$ singlet.

Table A1. Singlet parameters used to calculate signals, where f [μHz] is the frequency of the signal, Q the quality-factor and A [nm s^{-2}] amplitude. Amplitude marked with the asterisk points to the singlet which frequency is used as test frequency for obtaining the phasor walkout patters in Figs. A2 and A3.

	Mode	${}_0S_2^{-2}$	${}_0S_2^{-1}$	${}_0S_2^{0}$	${}_0S_2^{1}$	${}_0S_2^{2}$
	f	300.001	304.493	309.064	313.716	318.452
	Q	494.6	501.8	509.3	517.0	525.0
a	A	50∗	0	0	0	0
b	A	50∗	50	0	0	0
c	A	50∗	50	50	0	0
d	A	50∗	50	50	50	0
e	A	50∗	50	50	50	50
f	A	50	50	50∗	50	50
g	A	20∗	50	10	50	20
h	A	20	50	10∗	50	20
i	A	2∗	50	5	50	10
j	A	2	2∗	5	50	10