Full-waveform adjoint Q tomography in viscoelastic medium with central-frequency measurements

Pan Wenyong¹, Wang Yanfei¹, and Innanen Kristopher Albert Holm²

¹Institute of Geology and Geophysics, Chinese Academy of Sciences ²Department of Geoscience, University of Calgary

November 16, 2022

Abstract

Accurate Q (quality factor) structures can provide important constraints for characterizing subsurface hydrocarbon/water resources in exploration geophysics and interpreting tectonic evolution of the Earth in earthquake seismology. The attenuation effects on seismic amplitudes and phases can be included in forward and inverse modeling by invoking a generalized standard linear solid rheology. Compared to traditional ray-based methods, full-waveform adjoint tomography, which is based on numerical solutions of the visco-elastodynamic wave equation, has the potential to provide more accurate Q models. However, applications of adjoint Q tomography are impeded by the computational complexity of Q sensitivity kernels, and by strong velocity-Q trade-offs. In this study, following the adjoint-state method, we show that the Q (P and S wave quality factors Q_P and Q_S) sensitivity kernels can be constructed efficiently with adjoint memory strain variables. A novel central-frequency difference misfit function is designed to reduce the trade-off artifacts for adjoint Q tomography. Compared to traditional waveform-difference misfit function, this misfit function is less sensitive to velocity variations, and thus is expected to produce fewer trade-off uncertainties. The multiparameter Hessian-vector products are calculated to quantify the resolving abilities of different misfit functions. Comparative synthetic examples are given to verify the advantages of this new misfit function for adjoint Q_P and Q_S tomography. We end with a 3D viscoelastic inversion example designed to simulate a distributed acoustic sensing/vertical seismic profile survey for monitoring of CO₂ sequestration.

Full-waveform adjoint Q tomography in viscoelastic medium with central-frequency measurements

Wenyong Pan¹, Yanfei Wang^{1,2,3}, and Kristopher A. Innanen⁴

4	1 Key Laboratory of Petroleum Resource Research, Institute of Geology and Geophysics, Chinese Academy
5	of Sciences, Beijing 100029, P. R. China
6	² University of Chinese Academy of Sciences, Beijing 100049, P. R. China
7	³ Innovation Academy of Earth Science, Chinese Academy of Sciences, Beijing 100029, P. R. China
8	⁴ Department of Geoscience, University of Calgary, Calgary, AB T2N 1N4 Canada

Key Points:

1

2

3

10	•	An effective formulation based on the adjoint-state method is developed for vis-
11		coelastic full-waveform adjoint tomography
12	•	Multiparameter Hessian-vector products are used to evaluate the resolving abil-
13		ities of different misfit functions for Q perturbations
14	•	The new central-frequency misfit function can invert for the Q_P and Q_S models
15		more accurately by reducing the trade-off uncertainties

Corresponding author: Yanfei Wang, yfwang@mail.iggcas.ac.cn

16 Abstract

Accurate Q (quality factor) structures can provide important constraints for character-17 izing subsurface hydrocarbon/water resources in exploration geophysics and interpret-18 ing tectonic evolution of the Earth in earthquake seismology. The attenuation effects on 19 seismic amplitudes and phases can be included in forward and inverse modeling by in-20 voking a generalized standard linear solid rheology. Compared to traditional ray-based 21 methods, full-waveform adjoint tomography, which is based on numerical solutions of the 22 visco-elastodynamic wave equation, has the potential to provide more accurate Q mod-23 els. However, applications of adjoint Q tomography are impeded by the computational 24 complexity of Q sensitivity kernels, and by strong velocity-Q trade-offs. In this study, 25 following the adjoint-state method, we show that the Q (P and S wave quality factors 26 Q_P and Q_S) sensitivity kernels can be constructed efficiently with adjoint memory strain 27 variables. A novel central-frequency difference misfit function is designed to reduce the 28 trade-off artifacts for adjoint Q tomography. Compared to traditional waveform-difference 29 misfit function, this misfit function is less sensitive to velocity variations, and thus is ex-30 pected to produce fewer trade-off uncertainties. The multiparameter Hessian-vector prod-31 ucts are calculated to quantify the resolving abilities of different misfit functions. Com-32 parative synthetic examples are given to verify the advantages of this new misfit func-33 tion for adjoint Q_P and Q_S tomography. We end with a 3D viscoelastic inversion exam-34 ple designed to simulate a distributed acoustic sensing/vertical seismic profile survey for 35 monitoring of CO_2 sequestration. 36

³⁷ Plain Language Summary

Subsurface Q (quality factor) distributions are important for exploring natural re-38 sources and interpreting Earth's interior. In this study, full-waveform adjoint tomogra-39 phy is developed to improve our ability of mapping subsurface P and S wave quality fac-40 tors $(Q_P \text{ and } Q_S)$ in viscoelastic media. With the adjoint-state method, Q sensitivity 41 kernels can be constructed efficiently with adjoint memory strain variables. A new central-42 frequency difference misfit function is proposed to invert for the Q models. Numerical 43 experiments show that this new misfit function is more sensitive to Q perturbations than 44 standard forms and thus can reconstruct the subsurface Q models more stably by reduc-45 ing the trade-off artifacts caused by the confusions with seismic velocities. 46

47 **1** Introduction

As seismic waves propagate through the Earth's interior, partial energy is trans-48 formed into heat due to intrinsic attenuation, leading to amplitude dissipation and ve-49 locity dispersion of the waveforms (H. Liu et al., 1976). Attenuation effects are commonly 50 quantified with quality factor Q. Subsurface Q variations are associated with crack den-51 sity, fluid-saturation, partial melt, chemical composition (etc.) of the Earth materials 52 (Mavko & Nur, 1979; Berryman, 1988; Hauksson & Shearer, 2006), and thus maps of Q53 represent potentially useful interpretive tools. In exploration/monitoring seismology, high-54 resolution Q profiles can provide complement constraints for characterizing fluid-filled/gas-55 chimney/gas-hydrate reservoirs (Innanen, 2011; Operto & Miniussi, 2018; Y. Wang et 56 al., 2020); furthermore, knowledge of Q allows compensation procedures to produce high-57 fidelity seismic images by correcting for amplitude loss and phase distortions (T. Zhu et 58 al., 2014; Shen et al., 2018). In earthquake seismology, joint interpretation of elastic ve-59 locities and Q anomalies have been shown to improve our understandings of crust/mantle 60 structures and tectonic evolution (Romanowicz, 1995; Cao & Romanowicz, 2004; H. Zhu 61 et al., 2013; Bao et al., 2016; Debayle et al., 2020). 62

Most current subsurface Q models are obtained by ray-based tomography methods with spectral-ratio (Bath, 1974), central-frequency (Quan & Harris, 1997) or t^* (Eberhart-Phillips & Chadwick, 2002) measurements. However, these ray-based approaches may

give inaccurate Q values in complex geologic areas due to the inaccurate forward mod-66 eling operators. Anelasticity of a geological volume can be approximated using a phe-67 nomenological model represented mechanically by a combination of springs and dash spots. 68 The system constructed by several parallel standard linear elements is referred to as the generalized standard linear solid (GSLS) rheology (Carcione et al., 1988a). With the GSLS 70 model, dissipation and dispersion of propagating waves can be simulated by solving a 71 set of differential equations with the superposition of parallel relaxation mechanisms (Robertsson 72 et al., 1994; Blanch et al., 1995). High-resolution Q models with arbitrary spatial vari-73 ability can in principle be reconstructed through full-waveform adjoint tomography, which 74 has emerged as a powerful tool to obtain subsurface properties in exploration (Tarantola, 75 1984; Virieux & Operto, 2009; Brossier et al., 2009; Yuan et al., 2015; Operto & Miniussi, 76 2018; Yao et al., 2018; Pan et al., 2018) and earthquake (Tromp et al., 2005; P. Chen et 77 al., 2007; Tape et al., 2009; Fichtner & Trampert, 2011a; H. Zhu & Tromp, 2013; Tong 78 et al., 2014; M. Chen et al., 2017; Krischer et al., 2018; Dong & Yang, 2020) seismology. 79

However, the progress of full-waveform adjoint Q tomography is significantly lag-80 ging behind, in comparison to that of velocity tomography. The high computational com-81 plexity of constructing Q sensitivity kernels, and the strong multiparameter couplings 82 between Q and other elastic properties, are both important reasons for this. Effective 83 inversion strategies and algorithms for reliable Q estimation are still underdeveloped. In 84 this study, we have developed an effective formulation for constructing the Q sensitiv-85 ity kernels with adjoint memory strain variables, and applied them in the minimization 86 of a new central-frequency misfit function, designed to reduce trade-off artifacts. Devel-87 opments of these methods are introduced in the following. 88

In full-waveform adjoint tomography, the sensitivity kernels of elastic properties 89 can be calculated efficiently by cross-correlating the forward and adjoint wavefields based 90 on the adjoint-state method (Q. Liu & Tromp, 2006; Plessix, 2006). However, Q is not 91 explicitly included as a parameter in the rheological bodies of time domain viscoelastic 92 wave equation. Instead, the damping effects are modeled by determining the relaxation 93 parameters through a least-squares inversion to approximate a constant Q (Bohlen, 2002; Fichtner & van Driel, 2014). This complicates the derivation and calculation of Q sen-95 sitivity kernels. In previous studies, discrepancies exist between different frameworks for 96 constructing the sensitivity kernels in viscoelastic full-waveform adjoint tomography. Charara 97 et al. (2000) defined the attenuation parameters as the differences between "relaxed" and 98 "unrelaxed" moduli and derived the corresponding sensitivity kernels based on Born ap-99 proximation (Wu, 1989; Bostock et al., 2001). Assuming constant Q, Tromp et al. (2005) 100 derived the Q sensitivity kernels based on the Kolsky-Futterman model (Kolsky, 1952; 101 Futterman, 1962). In this study, we introduce a different, fully viscoelastic formulation; 102 we apply an adjoint-state method (Q. Liu & Tromp, 2008; Fichtner & van Driel, 2014) 103 to derive viscoelastically-based Q_P and Q_S sensitivity kernels, and show that they can 104 be efficiently calculated via the adjoint memory strain variables. 105

The problem of interparameter trade-off significantly complicates the multiparam-106 eter full-waveform adjoint tomography (Tarantola, 1986; Fichtner & Trampert, 2011b; 107 Operto et al., 2013; Innanen, 2014; Alkhalifah & Plessix, 2014; Pan et al., 2018; Kazei 108 & Alkhalifah, 2019). In viscoelastic media, velocity errors tend to produce strong arti-109 facts in the inverted Q models, which impedes the progress of adjoint Q tomography, in-110 creases the inversion uncertainties and may produce misleading interpretations for geo-111 physicists (Brossier, 2011; Keating & Innanen, 2020). Off-diagonal blocks within the in-112 verse multiparameter Hessian accounts for suppressing the trade-off artifacts (Pan et al., 113 2018). However, approximating the inverse Hessian with a small number of linear con-114 jugate gradient iterations does not seem to reduce the trade-off artifacts obviously. This 115 is a strong motivation to engage in designing specific misfit function that are inherently 116 less-exposed to trade-off errors (Karaoğlu & Romanowicz, 2018; Pan & Wang, 2020). Pan 117 and Wang (2020) showed that the amplitude-based misfit functions can resolve the Q 118

anomalies more effectively than traditional waveform-difference (WD) misfit function. 119 However, it is difficult to control the absolute seismic amplitudes, which are influenced 120 by many other factors including transmission/reflection loss, source and receiver coupling, 121 etc. Seismic attenuation reduces high-frequencies more rapidly than low-frequencies lead-122 ing to central-frequency downshifts of the seismic data (Hauge, 1981; Quan & Harris, 123 1997; T. Zhu et al., 2017), which is more stable and broadly-applicable across the full 124 bandwidth. Noting this, we introduce a new central-frequency difference (CD) misfit func-125 tion for adjoint Q_P and Q_S tomography in this study. The corresponding new adjoint 126 source is derived for calculating the Q sensitivity kernels. Because central-frequency vari-127 ation is mainly caused by intrinsic attenuation, the new CD misfit function is relatively 128 insensitive to velocity perturbations and amplitude fluctuations. In order to evaluate the 129 trade-off errors between different physical parameters, traditionally scattering radiation 130 patterns (Tarantola, 1986; Operto et al., 2013; Alkhalifah & Plessix, 2014; Pan et al., 131 2016) are qualitatively interpreted. In this paper, to evaluate the advantage of our method, 132 we quantify the velocity-Q trade-offs within WD and CD misfit functions using multi-133 parameter Hessian-vector products (Pan et al., 2018; Krischer et al., 2018). With the 134 quantitative and waveform-based measures, the new CD misfit function is expected to 135 be more sensitive to Q variations than the standard WD misfit and produce more ac-136 curate Q models. 137

We justify all conclusions with numerical experiments. First, we analyze the sen-138 sitivity of central-frequency corresponding to velocity and Q variations and evaluate the 139 resolving abilities of WD and CD misfit functions for Q perturbations with multiparam-140 eter Hessian-vector products. Then, a synthetic inversion example involving isolated velocity-141 Q anomalies is carried out, followed by a synthetic global tomography example with more 142 realistic structures based on the Rayleigh wave phase speed map (Trampert & Wood-143 house, 1995). We have found that the Q sensitivity kernels and inverted Q models ob-144 tained by WD misfit function are significantly damaged by the mappings from velocity 145 errors. However, when the new CD misfit function is applied instead, the Q anomalies 146 in the sensitivity kernels and inverted models can be identified more clearly suffering from 147 fewer trade-off artifacts. Finally, we carry out a 3D synthetic data inversion experiment, 148 designed to reflect a realistic CO_2 injection and storage monitoring effort (Yu et al., 2019). 149 Distributed acoustic sensing (DAS) has emerged as a popular technology for seismic ac-150 quisition with the advantages of low-cost, dense sampling, higher sensitivity to lower fre-151 quencies, etc., which make it particularly appropriate for permanent deployment in bore-152 hole environments (Cox et al., 2012; Daley et al., 2016; Ajo-Franklin et al., 2019). DAS 153 data can be incorporated relatively straightforwardly into seismic tomography algorithms, 154 either in isolation or combination with standard geophone data (Eaid et al., 2020). In 155 this study, we apply the proposed algorithms to 3D viscoelastic full-waveform adjoint 156 tomography with DAS data acquired in a vertical seismic profile survey (DAS-VSP) con-157 figuration. These numerical experiments act as a feasibility study and are suggestive of 158 the effectiveness and potential advantages of the proposed algorithms for adjoint Q to-159 mography in both earthquake and exploration seismology. 160

This paper is organized as follows. The basic principle of viscoelastic full-waveform adjoint tomography is first reviewed. Forward modeling problem in viscoelastic medium based on the GSLS model is introduced in the Appendix. Then, viscoelastic sensitivity kernels based on the adjoint-state method are derived. Multiparameter Hessian-vector products and the new CD misfit function are then introduced. In the numerical modeling section, synthetic examples are given to verify the advantages of proposed algorithms.

¹⁶⁷ 2 Theory and Methods

168

2.1 Full-waveform adjoint tomography in viscoelastic medium

In a standard formulation of viscoelastic full-waveform adjoint tomography, parameters representing geological properties of the medium are iteratively updated by minimizing the direct difference between simulated data \mathbf{u} and observed data \mathbf{d} , constrained by the visco-elastodynamic wave equation. This is embodied in the classical WD misfit function (Q. Liu & Tromp, 2006; Plessix, 2006):

$$\boldsymbol{\chi}\left(\mathbf{m},\mathbf{u},\boldsymbol{\lambda}\right) = \frac{1}{2}\sum_{\mathbf{x}_{r}}\int_{0}^{t'}\left[\mathbf{u}\left(\mathbf{m}\right) - \mathbf{d}\right]^{2}dt - \int_{0}^{t'}\int_{\Omega}\boldsymbol{\lambda}\cdot\left(\rho\partial_{t}^{2}\mathbf{u} - \nabla\cdot\mathbf{T} - \boldsymbol{f}^{s}\right)d\mathbf{x}dt, \quad (1)$$

where **m** is the model vector including different physical properties, \mathbf{x}_r indicate the receiver locations, t' is the maximum recording time, Ω indicates the full subsurface volume, $\boldsymbol{\lambda}$ is the Lagrangian multiplier, "·" means divergence, f^s is the source term and **T** is the stress tensor, determined by the entire history of strain fields via Boltzmann principle (Dahlen & Tromp, 1998; Aki & Richards, 2002):

$$T_{ij} = \kappa \delta_{ij} * \partial_t \varepsilon_{kk} + \mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) * \partial_t \varepsilon_{kl}, \tag{2}$$

where "*" means convolution, δ_{ij} is the Kronecker delta, the subscripts i, j, k and l take on the values of x, y, z, κ and μ are the relaxation functions of bulk and shear moduli:

$$\kappa = \kappa^R \left(1 + \frac{1}{P} \sum_{p=1}^P \tau_\kappa e^{-t/\tau^{\sigma_p}} \right) H(t),$$

$$\mu = \mu^R \left(1 + \frac{1}{P} \sum_{p=1}^P \tau_\mu e^{-t/\tau^{\sigma_p}} \right) H(t),$$
(3)

where κ^R and μ^R are the "relaxed" bulk and shear moduli at time $t = +\infty$ correspond-169 ing to low-frequency limit, P is the maximum number of relaxation mechanisms, H is 170 the Heaviside step function, $\tau^{\sigma p}$ is the stress relaxation time of the *p*th Maxwell body, 171 τ_{κ} and τ_{μ} measure the attenuation strengths of bulk and shear moduli, respectively. Equa-172 tion (2) is not well-suited for solving initial value problem as the whole history of strain 173 fields are required (Pan & Wang, 2020). The convolutional constitutive relationship can 174 be eliminated by introducing memory variables with the superposition of several par-175 allel relaxation mechanisms based on the generalized standard linear solid (GSLS) model. 176 Details about how this is included in the forward modeling solutions are provided in Ap-177 pendix A. 178

Newton optimization framework is the common approach for solving the full-waveform inverse problem. The search direction $\Delta \mathbf{m}$ at qth iteration can be obtained by solving the following Newton linear system:

$$\mathbf{H}_q \Delta \mathbf{m}_q = -K_q,\tag{4}$$

where K_q and \mathbf{H}_q are the sensitivity kernel (or gradient), based on first derivative of the misfit function with respect to the model parameters, and the Hessian, based on second derivative, respectively. In large-scale inverse problems, explicit calculation of the Hessian \mathbf{H} and its inverse are generally too computationally expensive to be possible. In response to this, optimization methods have been developed in which inverse Hessian approximations are used to precondition the gradient updates (Y. Wang & Yuan, 2005; Métivier et al., 2013; He & Wang, 2020). We employ a quasi-Newton *l*-BFGS method in our study to construct the search direction, following a "two-loop recursion" scheme (Nocedal & Wright, 2006). In the inversion process, the model parameters are updated by

$$\mathbf{m}_{q+1} = \mathbf{m}_q + \alpha_q \Delta \mathbf{m}_q,\tag{5}$$

where α_q is the step length for scaling the model updates and obtained using the backtracking line search method (Nocedal & Wright, 2006).

¹⁸¹ 2.2 Viscoelastic sensitivity kernels

Within the time domain GSLS viscoelastic wave equation, Q parameters are not 182 explicitly assigned to the rheological bodies. Instead, the relaxation parameters are se-183 lected to approximate a constant Q via an optimization procedure. This makes it cum-184 bersome to derive the Q sensitivity kernels, and introduces discrepancies of different meth-185 ods for attenuation estimation. Charara et al. (2000) defined the attenuation parame-186 ters as the differences between "relaxed" and "unrelaxed" moduli. The sensitivity ker-187 nels of "unrelaxed" moduli and attenuation parameters are derived based on Born ap-188 proximation; Tromp et al. (2005) derived the Q sensitivity kernels based on the Kolsky-189 Futterman model (Kolsky, 1952; Futterman, 1962) and frequency domain Born scatter-190 ing integral. The expressions of Q sensitivity kernels are the same with the moduli sen-191 sitivity kernels, but lead to different adjoint source in frequency domain. Pan and Wang 192 (2020) used Tromp's approach to calculate the Q sensitivity kernels in viscoelastic ad-193 joint tomography. However, the Kolsky-Futterman model used to derive the adjoint source 194 is not consistent with the GSLS model used in forward modeling, and this discrepancy 195 may produce additional uncertainty in the inversion results. 196

Thus, in this study, we make use of a fully self-consistent derivation and obtain the viscoelastic sensitivity kernels directly within the adjoint-state method (Q. Liu & Tromp, 2008; Fichtner & van Driel, 2014). The bulk and shear moduli quality factors (Q_{κ} and Q_{μ}) are first incorporated into the convolutional constitutive relation explicitly with the enforcement of $\tau_{\kappa} = Q_{\kappa}^{-1}$ and $\tau_{\mu} = Q_{\mu}^{-1}$ (Fichtner & van Driel, 2014), which gives

$$T_{ij} = \kappa^R \delta_{ij} \left(1 + \frac{1}{P} \sum_{p=1}^P \frac{1}{Q_\kappa} e^{-t/\tau^{\sigma_P}} \right) H(t) * \partial_t \varepsilon_{kk} + \mu^R \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) \left(1 + \frac{1}{P} \sum_{p=1}^P \frac{1}{Q_\mu} e^{-t/\tau^{\sigma_P}} \right) H(t) * \partial_t \varepsilon_{kl}.$$
(6)

Then, inserting equation (6) into equation (1) gives the augmented Lagrangian misfit function as:

$$\boldsymbol{\chi}\left(\mathbf{m},\mathbf{u},\boldsymbol{\lambda}\right) = \frac{1}{2} \sum_{\mathbf{x}_{r}} \int_{0}^{t'} \left[u_{i}-d_{i}\right]^{2} dt - \int_{0}^{t'} \int_{\Omega} \lambda_{i} \left(\rho \partial_{t}^{2} u_{i}-f_{i}^{s}\right) d\mathbf{x} dt + \int_{0}^{t'} \int_{\Omega} \lambda_{i} \partial_{j} \left[\kappa^{R} \delta_{ij} \left(1+\frac{1}{P} \sum_{p=1}^{P} \frac{1}{Q_{\kappa}} e^{-t/\tau^{\sigma_{P}}}\right) H\left(t\right) * \partial_{t} \varepsilon_{kk}\right] d\mathbf{x} dt - \int_{0}^{t'} \int_{\Omega} \lambda_{i} \partial_{j} \left[\frac{2}{3} \mu^{R} \delta_{ij} \left(1+\frac{1}{P} \sum_{p=1}^{P} \frac{1}{Q_{\mu}} e^{-t/\tau^{\sigma_{P}}}\right) H\left(t\right) * \partial_{t} \varepsilon_{kk}\right] d\mathbf{x} dt + \int_{0}^{t'} \int_{\Omega} \lambda_{i} \partial_{j} \left[2 \mu^{R} \left(1+\frac{1}{P} \sum_{p=1}^{P} \frac{1}{Q_{\mu}} e^{-t/\tau^{\sigma_{P}}}\right) H\left(t\right) * \partial_{t} \varepsilon_{ij}\right] d\mathbf{x} dt.$$

$$(7)$$

Thus, following the adjoint-state method, we can directly derive the viscoelastic sensitivity kernels. Variation of the Lagrangian misfit function (equation (7)) due to the perturbations of model properties $(\Delta u_i, \Delta \rho, \Delta \kappa, \Delta \mu, \Delta Q_\kappa \text{ and } \Delta Q_\mu)$ is given by

$$\begin{split} \Delta \boldsymbol{\chi} \left(\mathbf{m}, \mathbf{u}, \boldsymbol{\lambda} \right) &= \int_{0}^{t'} \int_{\Omega} \sum_{\mathbf{x}_{r}} \left[u_{i} - d_{i} \right] \Delta u_{i} d\mathbf{x} dt \\ &- \int_{0}^{t'} \int_{\Omega} \left[\rho \partial_{t}^{2} \lambda_{i} - \partial_{j} \left(\kappa \delta_{ij} \delta_{kl} * \partial_{t} \tilde{\varepsilon}_{kl} \right) \right] \Delta u_{i} d\mathbf{x} dt \\ &+ \int_{0}^{t'} \int_{\Omega} \partial_{j} \left[\mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) * \partial_{t} \tilde{\varepsilon}_{kl} \right] \Delta u_{i} d\mathbf{x} dt \\ &- \int_{0}^{t'} \int_{\Omega} \Delta \rho \lambda_{i} \partial_{t}^{2} u_{i} d\mathbf{x} dt \\ &- \int_{0}^{t'} \int_{\Omega} \left(\Delta \kappa * \partial_{t} \tilde{\varepsilon}_{kk} \right) \varepsilon_{ii} d\mathbf{x} dt \\ &- \int_{0}^{t'} \int_{\Omega} \left[\Delta \mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) * \partial_{t} \tilde{\varepsilon}_{kl} \right] \varepsilon_{ij} d\mathbf{x} dt \\ &+ \int_{0}^{t'} \int_{\Omega} \left[\Delta Q_{\kappa} \frac{\kappa^{R}}{Q_{\kappa}^{2}} \sum_{p=1}^{P} \tau^{\sigma p} \tilde{\epsilon}_{kk}^{p} \right) \varepsilon_{ii} d\mathbf{x} dt \\ &+ \int_{0}^{t'} \int_{\Omega} \left[\Delta Q_{\mu} \frac{\mu^{R}}{Q_{\mu}^{2}} \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) \sum_{p=1}^{P} \tau^{\sigma p} \tilde{\epsilon}_{kl}^{p} \right] \varepsilon_{ij} d\mathbf{x} dt \\ \end{split}$$

where $\tilde{\varepsilon}_{ij} = \partial_j \lambda_i$ is the Lagrangian strain field, and $\tilde{\epsilon}_{kl}^p$ are the Lagrangian memory strain variables:

$$\tilde{\epsilon}_{kl}^{p} = \frac{1}{P\tau^{\sigma p}} e^{-t/\tau^{\sigma p}} H\left(t\right) * \partial_{t} \tilde{\varepsilon}_{kl}.$$
(9)

The Lagrangian is stationary with respect to wavefield perturbation Δu_i in the absence of the perturbations $\Delta \rho$, $\Delta \kappa$, $\Delta \mu$, ΔQ_{κ} and ΔQ_{μ} (Q. Liu & Tromp, 2006). Setting the coefficient of Δu_i as zero gives the adjoint-state equation:

$$\rho \partial_t^2 \lambda_i - \partial_j \left[\kappa \delta_{ij} \delta_{kl} * \partial_t \tilde{\varepsilon}_{kl} + \mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) * \partial_t \tilde{\varepsilon}_{kl} \right] = f_{i,\text{WD}}^{\dagger}, \quad (10)$$

where $f_{i,\text{WD}}^{\dagger} = -(d_i - u_i)$ is the adjoint source of WD misfit function. Variation of the Lagrangian misfit function can be reduced to

$$\Delta \chi = -\left(\frac{\Delta\rho}{\rho}K_{\rho} + \frac{\Delta\kappa}{\kappa}K_{\kappa} + \frac{\Delta\mu}{\mu}K_{\mu} + \frac{\Delta Q_{\kappa}}{Q_{\kappa}}K_{Q_{\kappa}} + \frac{\Delta Q_{\mu}}{Q_{\mu}}K_{Q_{\mu}}\right),\tag{11}$$

where K_{ρ} , K_{κ} , K_{μ} , $K_{Q_{\kappa}}$ and $K_{Q_{\mu}}$ are the corresponding sensitivity kernels (or gradients) for ρ , κ , μ , Q_{κ} and Q_{μ} models, respectively. Their explicit expressions are listed in the following:

$$K_{\rho} = -\int_{0}^{t'} \rho u_i^{\dagger} \partial_t^2 u_i dt, \qquad (12a)$$

$$K_{\kappa} = -\int_{0}^{t'} \left(\kappa * \partial_t \varepsilon_{kk}^{\dagger}\right) \varepsilon_{ii} dt, \qquad (12b)$$

$$K_{\mu} = -\int_{0}^{t'} \left[\mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) * \partial_{t} \varepsilon_{kl}^{\dagger} \right] \varepsilon_{ij} dt, \qquad (12c)$$

$$K_{Q_{\kappa}} = \int_{0}^{t'} \left(\frac{\kappa^{R}}{Q_{\kappa}} \sum_{p=1}^{P} \tau^{\sigma p} \epsilon_{kk}^{p,\dagger} \right) \varepsilon_{ii} dt, \qquad (12d)$$

$$K_{Q_{\mu}} = \int_{0}^{t'} \left[\frac{\mu^{R}}{Q_{\mu}} \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) \sum_{p=1}^{P} \tau^{\sigma p} \epsilon_{kl}^{p,\dagger} \right] \varepsilon_{ij} dt,$$
(12e)

where u_i^{\dagger} is the adjoint displacement field defined as the time-reversed Lagrangian field, $\varepsilon_{kl}^{\dagger}$ and $\varepsilon_{kl}^{p,\dagger}$ are the adjoint strain fields and adjoint memory strain variables. The viscoelastic media is commonly described by the velocity-Q parameterization associated with V_P , V_S , Q_P and Q_S . According to the relations between these different physical parameters (see equations (15) and (16) in Pan and Wang (2020)), the sensitivity kernels for velocity (V_P and V_S) and the corresponding quality factors (Q_P and Q_S) can be derived following the chain rule, as listed in the following:

$$K_{V_P} = -\int_0^{t'} 2\left(\rho V_P^2 \delta_{ij} * \partial_t \varepsilon_{kk}^{\dagger} - W \kappa^R \delta_{ij} \sum_{p=1}^P \tau^{\sigma p} \epsilon_{kk}^{p,\dagger}\right) \varepsilon_{ij} dt,$$
(13a)

$$K_{V_S} = -\int_0^{t'} 4\left(\rho V_S^2 * \partial_t \varepsilon_{ij}^{\dagger} - \rho V_S^2 \delta_{ij} * \partial_t \varepsilon_{kk}^{\dagger} + \frac{W}{2} \kappa^R \delta_{ij} \sum_{p=1}^P \tau^{\sigma p} \epsilon_{kk}^{p,\dagger} \right) \varepsilon_{ij} dt, \qquad (13b)$$

$$K_{Q_P} = \frac{V_P^2}{Q_P \left(V_P^2 - V_S^2\right)} \int_0^{t'} \left(\kappa^R \delta_{ij} \sum_{p=1}^P \tau^{\sigma_P} \epsilon_{kk}^{p,\dagger}\right) \varepsilon_{ij} dt,$$
(13c)

$$K_{Q_S} = \frac{2}{Q_S} \int_0^{t'} \left(\mu^R \sum_{p=1}^P \tau^{\sigma p} \epsilon_{ij}^{p,\dagger} - \frac{1}{3} \mu^R \delta_{ij} \sum_{p=1}^P \tau^{\sigma p} \epsilon_{kk}^{p,\dagger} \right) \varepsilon_{ij} dt - \frac{V_S^2}{Q_S \left(V_P^2 - V_S^2\right)} \int_0^{t'} \left(\kappa^R \delta_{ij} \sum_{p=1}^P \tau^{\sigma p} \epsilon_{kk}^{p,\dagger} \right) \varepsilon_{ij} dt,$$
(13d)

where $\varepsilon_{ij}^{\dagger}$ and $\epsilon_{ij}^{p,\dagger}$ represent the adjoint strain fields and adjoint memory strain variables, and the coefficient W is

$$W = \frac{(Q_S - Q_P) V_P^2 V_S^2}{Q_P Q_S (V_P^2 - V_S^2)^2}.$$
(14)

Here, we ignore the influence of density in the inversion process. The velocity sensitiv-197 ity kernels K_{V_P} and K_{V_S} consist of the terms associated with adjoint stress fields and 198 adjoint memory strain variables, which are different from those in purely elastic media 199 (Tromp et al., 2005; Pan et al., 2018). The Q sensitivity kernels K_{Q_P} and K_{Q_S} are con-200 structed with forward strain fields and adjoint memory strain variables, which charac-201 terize the damping effects of seismic waves. Compared to the approach given in Tromp 202 et al. (2005), this formulation is theoretically more complete because the same GSLS model 203 is used in the forward and adjoint simulations. 204

205

2.3 Interparameter mapping and Hessian-vector product

The problem of interparameter trade-off is arguably the main practical challenge in the application of multiparameter adjoint tomography, and has been intensively investigated by geophysicists in recent years (Operto et al., 2013; Innanen, 2014; Alkhalifah & Plessix, 2014; Pan et al., 2016). In attenuating media, the inverted Q models can be easily damaged by the trade-off artifacts (Kamei & Pratt, 2013). For instance, considering simultaneous estimation of V_P and Q_P , the Newton linear system (equation (4)) can be written as:

$$\begin{bmatrix} \mathbf{H}_{V_P V_P} & \mathbf{H}_{V_P Q_P} \\ \mathbf{H}_{Q_P V_P} & \mathbf{H}_{Q_P Q_P} \end{bmatrix} \begin{bmatrix} \Delta V_P \\ \Delta Q_P \end{bmatrix} = -\begin{bmatrix} K_{V_P} \\ K_{Q_P} \end{bmatrix},$$
(15)

where the off-diagonal blocks $\mathbf{H}_{V_PQ_P}$ and $\mathbf{H}_{Q_PV_P}$ in multiparameter Hessian consists of mixed Fréchet derivatives of V_P and Q_P and characterize their interparameter couplings. The first-order sensitivity kernels K_{V_P} and K_{Q_P} can be re-formulated as the summation of Hessian-vector products (Pan et al., 2018):

$$K_{V_P} = -\mathbf{H}_{V_P V_P} \Delta V_P - \mathbf{H}_{V_P Q_P} \Delta Q_P, \tag{16a}$$

$$K_{Q_P} = -\mathbf{H}_{Q_P V_P} \Delta V_P - \mathbf{H}_{Q_P Q_P} \Delta Q_P, \tag{16b}$$

where the products of diagonal Hessian blocks $\mathbf{H}_{V_P V_P}$ and $\mathbf{H}_{Q_P Q_P}$ with the model perturbations of ΔV_P and ΔQ_P form the diagonal Hessian sensitivity kernels. The products of off-diagonal Hessian blocks $\mathbf{H}_{Q_P V_P}$ and $\mathbf{H}_{V_P Q_P}$ with the model perturbations of ΔV_P and ΔQ_P form the "second-order" trade-off (or contamination) sensitivity kernels, which serve as the mappings between V_P and Q_P (Pan et al., 2018). Similarly, interparameter trade-offs exist between V_S and Q_S , V_P and V_S , ρ and Q_S , etc., but in this study, we will focus on the trade-offs between velocity and Q.

The trade-offs of different physical parameters are commonly analyzed qualitatively with radiation patterns (Tarantola, 1986; Operto et al., 2013; Alkhalifah & Plessix, 2014; Pan et al., 2016), which measure the scattering amplitudes of seismic waves due to local model perturbations. However, radiation patterns are limited in their ability to incorporate dispersion effects due to Q perturbations (Keating & Innanen, 2020). Multiplying multiparameter Hessian with local model perturbations approximates the matrix columns, which are referred to as point-spread functions and provides one effective tool to quantify the trade-offs (Fichtner & Leeuwen, 2015; Pan et al., 2018; Sager et al., 2018). In this study, we calculate the products of multiparameter Gauss-Newton Hessian with model perturbation vectors to analyze the velocity-Q trade-offs. The Hessian-vector products can be calculated efficiently using the adjoint-state method (Métivier et al., 2013; Pan et al., 2018) (see Appendix B). For instance, product of off-diagonal block $\mathbf{H}_{Q_PV_P}$ with model perturbation ΔV_P can be calculated by

$$\mathbf{H}_{Q_P V_P} \Delta V_P = \left\langle \frac{2V_P^2}{Q_P \left(V_P^2 - V_S^2\right)} \left(\kappa^R \delta_{ij} \sum_{p=1}^P \tau^{\sigma p} \epsilon_{kk}^{p,\dagger} \right) \varepsilon_{ij} \\
\times \left(\rho V_P^2 \delta_{i'j'} * \partial_t \varepsilon_{k'k'}^{\dagger} - W \kappa^R \delta_{i'j'} \sum_{p=1}^P \tau^{\sigma p} \epsilon_{k'k'}^{p,\dagger} \right) \Delta V_P \varepsilon_{i'j'} \right\rangle,$$
(17)

where $\langle \cdot \rangle$ means summations of time and space samples for sake of compactness.

2.4 Central-frequency difference misfit function

214

Previous studies have revealed that when using WD misfit function, velocity errors produce strong trade-off artifacts in the inverted Q models (Mulder & Hak, 2009; Brossier, 2011; Pan & Wang, 2020), with both velocity and Q variations contributing to the waveform residuals. In principle, the inverse Hessian corrects many of the trade-off errors, but, in applying the truncated Newton approach, and using a small number of conjugate gradient iterations, we have observed limited reduction of trade-off artifacts.

Intrinsic attenuation leads to obvious amplitude reduction and central-frequency downshift of the seismic data. Designing specific misfit functions measuring amplitude and central-frequency variations represents one alternative approach to reduce the tradeoff artifacts without increasing computational cost. Because seismic amplitudes are also influenced by many other factors including instrument responses, radiation patterns, etc, the amplitude-based misfit functions are easily affected by unexpected amplitude fluctuations. Therefore, in this study, we introduce a new central-frequency difference (CD) misfit function for adjoint Q tomography:

$$\boldsymbol{\chi}_{\rm CD}\left(Q\right) = \frac{1}{2} \sum_{\mathbf{x}_r} \left[f_i^c\left(Q\right) - f_{i,\rm obs}^c \right]^2,\tag{18}$$

where f_i^c and $f_{i,obs}^c$ indicate the central-frequencies of synthetic and observed data, respectively. We determine central-frequency of the signal by (Berkhout, 1984; Barnes, 1993)

$$f_{i}^{c} = \frac{1}{N} \int_{0}^{+\infty} f \mathcal{A}_{i}^{2}(f) \, df, \tag{19}$$

where f indicates frequency, \mathcal{A}_i represents amplitude spectrum of the seismic data and the coefficient N in the denominator is

$$N = \int_0^{+\infty} \mathcal{A}_i^2(f) \, df. \tag{20}$$

Variation of the misfit function with respect to Q perturbation is given by

$$\Delta \boldsymbol{\chi}_{\rm CD} \left(Q \right) = \sum_{\mathbf{x}_r} \left[f_i^c \left(Q \right) - f_{i,\rm obs}^c \right] \Delta f_i^c, \tag{21}$$

where Δf_i^c indicates the variation of central-frequency due to Q perturbation:

$$\Delta f_i^c = \left[\int_0^{+\infty} f \mathcal{A}_i(f) df \right] \frac{2\mathcal{R} \left[\tilde{u}_i(f) \Delta \tilde{u}_i(f) \right]}{N \mathcal{A}_i(f)} - \left[\int_0^{+\infty} f \mathcal{A}_i^2(f) df \int_0^{+\infty} \mathcal{A}_i(f) df \right] \frac{2\mathcal{R} \left[\tilde{u}_i(f) \Delta \tilde{u}_i(f) \right]}{N^2 \mathcal{A}_i(f)},$$
(22)

where \tilde{u}_i is the synthetic data in frequency domain, the symbol \mathcal{R} means real part and $\Delta \tilde{u}_i$ indicates the data variation caused by Q perturbation. Inserting equation (22) into equation (21) gives:

$$\Delta \boldsymbol{\chi}_{\rm CD} = \sum_{\mathbf{x}_r} \left[f_i^c - f_{i,\text{obs}}^c \right] \frac{2\mathcal{R} \left[\tilde{u}_i \left(f \right) \Delta \tilde{u}_i \left(f \right) \right]}{N \mathcal{A}_i \left(f \right)} \\ \times \left[\int_0^{+\infty} f \mathcal{A}_i \left(f \right) df - \frac{1}{N} \int_0^{+\infty} f \mathcal{A}_i^2 \left(f \right) df \int_0^{+\infty} \mathcal{A}_i \left(f \right) df \right].$$
(23)

Thus, the adjoint source for calculating Q sensitivity kernels using the central-frequency misfit function χ_{CD} in frequency domain can be obtained as:

$$\tilde{f}_{i,\text{CD}}^{s,\dagger}(f) = \sum_{\mathbf{x}_{r}} \frac{2\mathcal{R}\left[\tilde{u}_{i}\left(f\right)\left(f_{i}^{c}-f_{i,\text{obs}}^{c}\right)\right]}{N\mathcal{A}_{i}\left(f\right)} \times \left[\int_{0}^{+\infty} f\mathcal{A}_{i}\left(f\right)df - \frac{1}{N}\int_{0}^{+\infty} f\mathcal{A}_{i}^{2}\left(f\right)df \int_{0}^{+\infty} \mathcal{A}_{i}\left(f\right)df\right].$$
(24)

Because central-frequency of the seismic data is mainly controlled by Q variations, the new CD misfit function is expected to invert for the Q models independent of velocity variations. This efficacy can be quantified by applying multiparameter Hessian-vector products to the problem of resolution analysis for Q inversion emerging from this new misfit function.

3 Numerical Experiments

In this section, we first analyze the seismic waveforms with velocity and Q variations and evaluate the sensitivities of different misfit functions for Q perturbations using multiparameter Hessian-vector products. Then, inversion experiments are carried out for full-waveform adjoint Q tomography in 2D/3D viscoelastic medium. The numerical experiments are performed using open-source package SeisElastic2D_1.1 (Pan et al., 2020) based on the spectral-element forward modeling method (Komatitsch & Tromp, 2005).

233 234

3.1 Sensitivity analysis of seismic waveforms and multiparameter Hessianvector products

235 Seismic forward modeling experiments in viscoelastic media are carried out by per-236 turbing Q and velocity models. The 2D initial viscoelastic model is homogeneous with 237 $V_P=2V_S=5.50 \text{ km/s}, Q_P=Q_S=150 \text{ and density } \rho=3000 \text{ kg/m}^3$, as shown in Figure 1a.



Figure 1. (a) shows the initial homogeneous model for seismic waveform analysis with single source-receiver pair; The source and receiver are arranged at (X=31.5 km, Y=31.5 km) and (X=68.5 km, Y=68.5 km), respectively. (b) shows the acquisition geometry for calculating the multiparameter Hessian-vector products; The red stars and white circles indicate the locations of sources and receivers. Locations of the 4 sources are (X=31.5 km, X=31.5 km), (X=31.5 km, Y=68.5 km), (X=68.5 km, Y=31.5 km) and (X=68.5 km, Y=68.5 km), respectively. (c) and (d) show the perturbation vectors of velocity ($\Delta V_P = 2\Delta V_S = 0.275$ km/s) and Q ($\Delta Q_P = \Delta Q_S = -125$).

The maximum sizes of the model in X and Y directions are both 100 km. A Ricker wavelet 238 with dominant frequency of 0.50 Hz is used as the source time function. The initial ho-239 mogeneous model is used as benchmark for the forward modeling experiments and re-240 ferred to as Model I. We first change the background Q_P and Q_S values from 150 to 25, 241 which is referred to as Model II. Based on Model II, we apply -5% and -10% pertur-242 bations to V_P and V_S models, which are referred to as Model III and IV, respectively. 243 The density model is kept unchanged. Properties of the 4 viscoelastic models are illus-244 trated in Table 3.1. 245

Then, forward modeling experiments are performed with single source-receiver pair. 246 The recorded vertical (z) component data are plotted for comparison, as shown in Fig-247 ure 2a. Figures 2b and 2c present the corresponding amplitude spectra of P and S waves, 248 respectively. The central-frequencies calculated using equation (19) are also labeled. As 249 can be seen that when decreasing Q_P and Q_S from 150 to 25, the waveform amplitudes 250 reduce significantly. Central-frequencies of P and S waves are also reduced from 0.46 and 251 0.45 to 0.42 and 0.38, respectively. For Model III and IV, the waveforms lag obviously 252 due to velocity perturbations. However, the central-frequencies almost do not change, 253 meaning that velocity perturbations produce limited influences on central-frequencies 254 of the seismic data. Similar phenomena can also be observed in horizontal (x) compo-255 nent of the recordings (Figure S1). 256



Figure 2. (a) shows the comparison of seismic waveforms (z component) calculated using Model I (solid-blue line), II (solid-black line), III (dashed-orange line) and IV (dash-dotted-purple line). The red and blue shading areas indicate the time windows for extracting P and S waves. (b) shows the corresponding amplitude spectra of P waves; (c) shows the corresponding amplitude spectra of S waves. f^c (Hz) indicate central-frequencies of P and S waves. Colors of the amplitude spectra and f^c are consistent with those of the models.

Models	Velocity (km/s)	Quality factor	Density (kg/m^3)
I	$V_P = 5.50$ $V_S = 2.75$	$Q_P = 150$ $Q_S = 150$	3000
II	$V_P = 5.50$ $V_S = 2.75$	$Q_P = 25$ $Q_S = 25$	3000
III	$V_P = 5.23$ $V_S = 2.61$	$Q_P = 25$ $Q_S = 25$	3000
IV	$V_P = 4.95$ $V_S = 2.48$	$Q_P = 25$ $Q_S = 25$	3000

 Table 1. Properties of the Viscoelastic Models for Forward Modeling



Figure 3. (a) shows the Hessian-vector products $(\mathbf{H}_{V_P V_P}^{\mathrm{wd}} \Delta V_P, \mathbf{H}_{V_P Q_P}^{\mathrm{wd}} \Delta Q_P, \mathbf{H}_{Q_P V_P}^{\mathrm{wd}} \Delta V_P)$ and $\mathbf{H}_{Q_P Q_P}^{\mathrm{wd}} \Delta Q_P$) associated with ΔV_P and ΔQ_P calculated using WD misfit function; (b) shows the Hessian-vector products $(\mathbf{H}_{V_S V_S}^{\mathrm{wd}} \Delta V_S, \mathbf{H}_{V_S Q_S}^{\mathrm{wd}} \Delta Q_S, \mathbf{H}_{Q_S V_S}^{\mathrm{wd}} \Delta V_S \text{ and } \mathbf{H}_{Q_S Q_S}^{\mathrm{wd}} \Delta Q_S)$ associated with ΔV_S and ΔQ_S calculated using WD misfit function; (c) shows the Hessian-vector products $(\mathbf{H}_{V_P V_P}^{cd} \Delta V_P, \mathbf{H}_{V_P Q_P}^{cd} \Delta Q_P, \mathbf{H}_{Q_P V_P}^{cd} \Delta V_P \text{ and } \mathbf{H}_{Q_P Q_P}^{cd} \Delta Q_P)$ associated with ΔV_P and ΔQ_P calculated using CD misfit function; (d) shows the Hessian-vector products ($\mathbf{H}_{V_S V_S}^{cd} \Delta V_S$, $\mathbf{H}_{V_SQ_S}^{cd}\Delta Q_S, \mathbf{H}_{Q_SV_S}^{cd}\Delta V_S \text{ and } \mathbf{H}_{Q_SQ_S}^{cd}\Delta Q_S)$ associated with ΔV_S and ΔQ_S calculated using CD misfit function.

Next, we proceed forward further by calculating the multiparameter Hessian-vector products for inversion sensitivity analysis with different misfit functions. The distribu-258 tions of 4 sources and 289 receivers are presented in Figure 1b. We apply velocity (-5%)259 and $Q (\Delta Q_P = \Delta Q_S = -125)$ perturbations at center parts of the models, as presented 260 in Figures 1c and 1d. The recorded P waves are used to calculate the products of mul-261 tiparameter Hessian with perturbation vectors ΔV_P and ΔQ_P . The recorded S waves 262 are used to calculate the multiparameter Hessian-vector products associated with ΔV_S 263 and ΔQ_S . The Hessian-vector products calculated using WD and CD misfit functions 264 are presented in Figure 3. When using WD misfit function, the Hessian-vector products 265 $\mathbf{H}_{V_P V_P}^{\mathrm{wd}} \Delta V_P$ and $\mathbf{H}_{V_S V_S}^{\mathrm{wd}} \Delta V_S$ dominate the panels, as shown in Figures 3a and 3b. This 266 means that even for -5% velocity perturbations, the WD misfit function is much more 267 sensitive to velocity perturbations. However, when using CD misfit function, the pan-268 els are dominated by Hessian-vector products $\mathbf{H}_{Q_PQ_P}^{cd}\Delta Q_P$ and $\mathbf{H}_{Q_SQ_S}^{cd}\Delta Q_S$, suggest-269 ing that the CD misfit function is more sensitive to Q perturbations. These observations 270 in the numerical experiments suggest that the new CD misfit function will be more ef-271 fective to invert for Q models suffering from limited influences of velocity errors. 272



Figure 4. (a) The target V_S model structure; (b) The target model perturbation ΔV_S ; (c) The target Q_S model structure; (c) The target model perturbation ΔQ_S .

3.2 Isolated velocity-Q anomaly example

273

A synthetic example with isolated velocity-Q anomalies is designed to examine the 274 performances of WD and CD misfit functions for Q inversion. The initial velocity and 275 Q models are the same with those in previous example. The target V_S model is created 276 by embedding two positive perturbations in the homogeneous background, as shown in 277 Figure 4a. Figure 4b shows the target model perturbation ΔV_S . The initial and target 278 V_P models are created with $V_P/V_S=2$. The target Q_S model is created by embedding 279 two strong anomalies (Q=25) in the homogeneous background, as shown in Figure 4c. 280 Figure 4d shows the target model perturbation ΔQ_S . The initial and target Q_P mod-281 els are created with $Q_P = Q_S$. The Q and velocity anomalies are isolated, which helps 282 evaluate and visualize the interparameter mappings. Furthermore, to examine the in-283 fluences of velocity errors, the velocity perturbations vary from +5% to +10% of the back-284 ground velocities. 285

Figure 5 shows the Q_P and Q_S sensitivity kernels calculated using WD and CD 286 misfit functions for comparison. With +5% velocity perturbations, the Q sensitivity ker-287 nels calculated using WD misfit function are blurred by the trade-off artifacts, as indi-288 cated by the arrows in Figures 5a and 5b. When increasing velocity perturbations to +10%, 289 these artifacts in the Q sensitivity kernels become stronger, as shown in 5b and 5d. How-290 ever, when using the new CD misfit function, for +5% and +10% velocity perturbations, 291 the Q anomalies in the sensitivity kernels can be identified clearly suffering from fewer 292 trade-off artifacts. Then, we carry out inversion experiments for Q estimation in the pres-293 ence of +5% velocity perturbations. The P and S waves in both z and x component data 294 are extracted to invert for Q_P and Q_S models, respectively. The recovered perturbations 295 $\Delta Q_P^{\rm wd}$ and $\Delta Q_S^{\rm wd}$ by WD misfit function are contaminated by trade-off artifacts seriously, 296 as indicated by the arrows in Figures 5i and 5j. However, with CD misfit function, the 297 velocity errors produce very limited influences on the recovered perturbations $\Delta Q_P^{\rm ed}$ and 298



Figure 5. (a) and (b) are the Q_P and Q_S sensitivity kernels $K_{Q_P}^{wd}$ and $K_{Q_S}^{wd}$ calculated using WD misfit function with +5% velocity perturbations; (c) and (d) are the Q_P and Q_S sensitivity kernels $K_{Q_P}^{wd}$ and $K_{Q_S}^{wd}$ calculated using WD misfit function with +10% velocity perturbations; (e) and (f) are the sensitivity kernels $K_{Q_P}^{cd}$ and $K_{Q_S}^{cd}$ calculated using CD misfit function with +5% velocity perturbations; (g) and (h) are the sensitivity kernels $K_{Q_P}^{cd}$ and $K_{Q_S}^{cd}$ calculated using CD misfit function with +5% velocity perturbations; (g) and (h) are the sensitivity kernels $K_{Q_P}^{cd}$ and $K_{Q_S}^{cd}$ calculated using CD misfit function with +10% velocity perturbations; (i) and (j) are the inverted ΔQ_P^{wd} and ΔQ_S^{wd} model perturbations using WD misfit function with +5% velocity perturbations; (k) and (l) are the corresponding inverted ΔQ_P^{cd} and ΔQ_S^{cd} model perturbations using CD misfit function. The dotted-black lines indicate boundaries of the Q anomalies.

²⁹⁹ $\Delta Q_S^{\rm cd}$. The Q anomalies are clearly resolved, as shown in Figures 5k and 5l. These ob-³⁰⁰ servations verify the advantages of CD misfit function for reliable Q inversion.

301

3.3 Global tomography example

In this subsection, following Yuan et al. (2016) and Tromp and Bachmann (2019), 302 we design a synthetic global tomography example with realistic structures based on the 303 Rayleigh wave phase speed map provided by Trampert and Woodhouse (1995). Figure 304 6a shows the target V_S model, which is discretized with 50 and 100 elements in Latitude 305 and Longitude directions corresponding to one meter per degree (see section 5.2 in Yuan 306 et al. (2016)). The initial V_S model is homogeneous with a constant value of 4.5 km/s. 307 Figure 6b shows the target model perturbation ΔV_S . The target and initial V_P models 308 are created with $V_P/V_S=2$. The initial Q_S model is homogeneous with a constant value 309 of 150. The target Q_S model is created by embedding one strong anomaly ($Q_S=25$) in 310 the homogeneous background, as shown in Figure 6c. Figure 6d shows the target model 311 perturbation ΔQ_S . The target and initial Q_P models are created with $Q_P = Q_S$. A num-312 ber of 32 sources and 293 receivers are selected as the acquisition network, as indicated 313 by the red stars and white circles in Figure 6. Ray coverage of the network is plotted in 314



Figure 6. (a) and (b) are the target V_S velocity model and perturbation ΔV_S ; The red stars and white circles indicate the locations of sources and stations. (c) and (d) are the target Q_S model and perturbation ΔQ_S ; (e) shows the ray coverage of the seismic acquisition; (f) shows the locations of one source and two receivers (A and B) for waveform comparison.

Figure 6e. A 0.60 kHz dominant frequency Ricker wavelet is used as the source to generate the observed data.

For the inversion experiments, in stage I, we first invert for V_P and V_S models us-317 ing cross-correlation (CC) traveltime misfit function (Luo & Schuster, 1991; Yuan et al., 318 2016) by expanding the frequency band from [0 kHz, 0.3 kHz] to [0 Hz, 0.6 kHz], which 319 helps reduce non-linearity of the inverse problem. P and S waves are extracted from z320 and x component data to invert for V_P and V_S models separately. Then, the velocity struc-321 tures are refined using WD misfit function. The final reconstructed velocity perturba-322 tions are presented in Figure 7. Reductions of the CC and WD data misfits are provided 323 in Figure S2. The velocity structures around the Q anomalies appear to be under-estimated 324 because of wrong Q models, as indicated by the black arrows. Because the ray coverage 325 is limited, the whole velocity structures are not perfectly recovered. 326

In stage II of the inversion experiments, we use WD and CD misfit functions to in-327 vert for Q models. The Q_P and Q_S sensitivity kernels are calculated for comparison, as 328 shown in Figure 8. In the Q sensitivity kernels calculated using WD misfit function, the 329 Q anomalies can be identified. However, the structures are distorted to some extent. The 330 trade-off artifacts caused by residual velocity errors also appear, as indicated by the red 331 arrows in Figures 8a and 8b. Figures 8e and 8f show the corresponding inverted Q model 332 perturbations by WD misfit function, wherein the Q structures are distorted and strong 333 trade-off artifacts are clearly visible, as indicated by the red arrows. These false Q struc-334 tures may result in misleading interpretations of subsurface models. The Q sensitivity 335 kernels and model perturbations obtained using CD misfit function are presented in Fig-336



Figure 7. (a) and (b) are the inverted velocity model perturbations $\Delta V_P^{\text{cc-wd}}$ and $\Delta V_S^{\text{cc-wd}}$ using CC and WD misfit functions in stage I.

ures 8c-8d and Figures 8g-8h. Even though boundaries of the Q anomalies suffer from some smearing effects, the whole structures match with the target Q models more closely without containing strong artifacts.

In Figure 9, the recorded z component P and S waves calculated using the target, 340 initial and inverted models are plotted. The locations of one source and two receivers 341 (A and B) are presented in Figure 6f. Because the seismic waves propagating from the 342 source to receiver B pass trough the strongly attenuative zones, the amplitudes and central-343 frequencies reduce more obviously that those of the waves recorded at A. Compared to 344 the observed data, the synthetic data calculated using initial models experience ampli-345 tude and central-frequency reductions. Furthermore, compared to WD misfit function, 346 the waveforms and central-frequencies of the data calculated using the inverted Q mod-347 els by CD misfit function match with those of the observed data more closely. Compar-348 isons of the x component waveforms are provided in Figure S3. The CD data misfits also 349 reduce much more faster than the WD data misfits for Q inversion, as shown in Figure 350 S2. From these numerical experiments and observations, we have proved that the new 351 CD misfit function shows superiority to overcome the parameter trade-off difficulty and 352 recover subsurface Q models more reliably. 353

354

3.4 3D viscoelastic inversion example

Finally, the proposed algorithms are applied to 3D viscoelastic full-waveform tomography with data emulating a distributed acoustic sensing (DAS) experiment acquired in a walk-away vertical seismic profile configuration (DAS-VSP). This synthetic example is designed to emulate a carbon capture and storage monitoring with DAS-VSP data in practical conditions (Hall et al., 2019).

Figures 10a and 10b show the target and initial 3D V_S models with complex to-360 pographic variations. The maximum sizes of the model in X, Y and Z directions are 1000 361 m, 1000 m and 500 m, respectively. The target and initial 3D V_P models are created with 362 $V_P/V_S=3$. Figure 10c shows the target 3D Q_S model. The initial 3D Q_S model is ho-363 mogeneous with $Q_S=200$. The target 3D Q_S model is created by embedding one atten-364 uating layer $(Q_s=20)$ overlapping with the low velocity formation at near-surface. The 365 target and initial 3D Q_P models are created with $Q_P = Q_S$. The density model is homo-366 geneous with a constant value of 3000 kg/m^3 and is kept unchanged in the inversion. Fig-367 ure 11a shows the 3D mesh grids for discretizing the model and describing the complex 368 topography. Figure 11b shows the acquisition settings. Twenty-five sources are arranged 369 regularly on top surface of the model, as indicated by the red spheres in Figure 11b. The 370 fiber-optic is deployed in the borehole with a recording interval of 1 m from 20 m to 500 371



Figure 8. (a) and (b) are the Q_P and Q_S sensitivity kernels $K_{Q_P}^{wd}$ and $K_{Q_S}^{wd}$ calculated using WD misfit function; (c) and (d) are the Q_P and Q_S sensitivity kernels $K_{Q_P}^{cd}$ and $K_{Q_S}^{cd}$ calculated using CD misfit function; (e) and (f) are the inverted Q model perturbations ΔQ_P^{wd} and ΔQ_S^{wd} using WD misfit function; (g) and (h) are the inverted Q model perturbations ΔQ_P^{cd} and ΔQ_S^{cd} using CD misfit function. The dotted-black lines depict the boundaries of Q anomalies. The red stars and white circles indicate the locations of sources and stations.



Figure 9. (a) shows the P waves recorded at A calculated using target (black PA-T), initial models (blue PA-I) and the inverted models by WD (orange PA-WD) and CD (purple PA-CD) misfit functions; (b) shows the S waves recorded at A calculated using target (black S-A-T), initial models (blue S-A-I) and the inverted models by WD (orange S-A-WD) and CD (purple S-A-CD) misfit functions; (c) shows the P waves recorded at B calculated using target (black PB-T), initial models (blue PB-I) and the inverted models by WD (orange PB-WD) and CD (purple PB-CD) misfit functions; (d) shows the S waves recorded at B calculated using target (black S-B-T), initial models (blue S-B-I) and the inverted models by WD (orange S-B-WD) and CD (purple PB-CD) misfit functions; (e), (f), (g) and (h) are the corresponding amplitude spectra of the P and S waves in (a), (b), (c) and (d). f^c (kHz) indicate central-frequencies of P and S waves. Colors of the amplitude spectra and f^c are consistent with those of the waves.



Figure 10. (a) The target 3D V_S velocity model with complex topographic variations; (b) The initial 3D V_S velocity model; (c) The target 3D Q_S model.



Figure 11. (a) The 3D unstructured mesh; (b) The distributions of sources (red spheres) and optical-fiber (white spheres).

m, as indicated by the white spheres in Figure 11b. Figure 12 shows the 3D slices extracted from the target model perturbations ΔV_S and ΔQ_S .

The source used for forward modeling is a 30 Hz dominant frequency Ricker wavelet. We first transform the data of vertical strain-rate into z component displacement data using the frequency-wavenumber method (Daley et al., 2016), which is then used as the observed data for inversion. The forward and adjoint simulations in 3D viscoelastic medium are performed on a local workstation with 16 cores (3.10 GHz) and 2 NVIDIA GeoRTX2080Ti Graphic Processing Unit (GPU) cards.

In the inversion experiments, we first invert for V_P and V_S models using CC and 380 WD misfit functions within frequency band [1 Hz, 30 Hz] in stage I. The 3D slices ex-381 tracted from the reconstructed velocity model perturbations ΔV_P and ΔV_S are presented 382 in Figure 13. Reductions of the normalized CC and WD data misfits are presented in 383 Figure S4. Because the source-receiver illumination is limited, only the velocity struc-384 tures around the borehole are well recovered. Some formations are distorted due to the 385 influences of wrong Q models. In stage II, we apply the CD misfit function to invert for 386 the near-surface Q_P and Q_P anomalies. The reconstructed model perturbations ΔQ_P 387 and ΔQ_S are presented in Figure 14. Reductions of the normalized CD data misfits are 388 presented in Figure S4. Even though some artifacts appear in the deeper parts of the 389 inverted ΔQ_P and ΔQ_S model perturbations, the strong Q_P and Q_S anomalies at near-390 surface are resolved clearly. Figure 15 shows the comparisons of shot gathers and traces 391 (at the depths of -250 m and -300 m) calculated using target, initial and inverted mod-392 els. Waveforms and central-frequencies calculated using the inverted velocity and Q mod-393

els match those of the observed data closely. These results indicate that the proposed algorithms are applicable to DAS-VSP data for Q_P and Q_S inversion.

396 4 Discussions

In this paper, the engine for viscoelastic forward and adjoint simulations is based on the GSLS rheology in time domain. However, the attenuation effects on propagating seismic waves can be modeled using many other physical mechanisms in time or frequency domain. Thus, the velocity and Q sensitivity kernels in equation (13) are only applicable to viscoelastic inversion experiments based on the GSLS model. The formulas of calculating the sensitivity kernels should depend on the physical mechanism in forward simulation. Furthermore, it is necessary to evaluate the uncertainties and modeling errors in viscoelastic adjoint tomography with different physical mechanisms.

Theoretically, central-frequency of the seismic data is mainly controlled by Q vari-405 ations. In the numerical experiments, we find that velocity perturbations result in fluc-406 tuations of the amplitude spectra. In practical seismic data, the amplitude spectra are always dithered due to noise. These effects can result in errors for calculating central-408 frequencies using equation (19), which will lead to unexpected uncertainties for Q inver-409 sion. Smoothing the amplitude spectra can remove these fluctuations and dithering ef-410 fects. Variation of the central-frequency can also be measured by cross-correlating the 411 amplitude spectra of synthetic and observed data, which is similar to the CC traveltime 412 misfit function but leads to different adjoint source for calculating Q sensitivity kernels. 413 Stability and accuracy of these two different approaches for determining central-frequency 414 difference will be examined in our future studies. 415

416 5 Conclusions

In this study, we develop theory and methods for full-waveform adjoint Q tomog-417 raphy in 2D/3D viscoelastic medium. The Q sensitivity kernels are constructed efficiently 418 with adjoint memory strain variables based on the adjoint-state method. Multiparam-419 eter Hessian-vector products provide an effective tool for quantifying the trade-offs be-420 tween velocity and Q. Compared to traditional waveform-difference misfit function, the 421 proposed new central-frequency misfit function is more sensitive to Q variations and can 422 invert for the Q models more reliably by reducing the trade-off artifacts caused by ve-423 locity errors. Synthetic inversion examples in 2D/3D viscoelastic medium verify that the 424 proposed algorithms are effective and advantageous for adjoint Q tomography, which is 425 applicable in both exploration and earthquake seismology. 426

427 Acknowledgments

This research is supported by National Natural Science Foundation of China (Grant 428 No. E1115401), IGGCAS Research Start-up Funds (Grant No. E0515402), IGGCAS grant 429 2019031 and CAS innovation program ZDBS-LY-DQC003. K. A. Innanen is supported 430 by CREWES project and NSERC grants (CRDPJ 461179-13 and CRDPJ 543578-19). 431 The forward modeling codes in 2D and 3D viscoelastic medium are available at https:// 432 github.com/PanIGGCAS/SeisElastic2D_1.1/specfem2d and https://github.com/geodynamics/ 433 specfem3d. The full-waveform adjoint tomography codes can be found at https://github .com/PanIGGCAS/SeisElastic2D_1.1/SeisElastic2D. The data, models and plotting 435 scripts used in this research can be downloaded from http://doi.org/10.6084/m9.figshare 436 .13655186. 437



Figure 12. (a) shows the Y-Z slice of the target model perturbation ΔV_S at X=400 m; (b) shows the X-Z slice of the target model perturbation ΔV_S at Y=400 m; (c) shows the X-Y slice of the target model perturbation ΔV_S at Z=-320 m; (d) shows the 3D view of these slices in ΔV_S ; (e), (f) and (g) are the Y-Z, X-Z and X-Y slices of target model perturbation ΔQ_S at X=400 m, Y=400 m and Z=-100 m, respectively; (h) is the 3D view of these slices in ΔQ_S .



Figure 13. (a) shows the Y-Z slice of the inverted model perturbation ΔV_P at X=400 m; (b) shows the X-Z slice of the inverted model perturbation ΔV_P at Y=400 m; (c) shows the X-Y slice of the inverted model perturbation ΔV_P at Z=-320 m; (d) shows the 3D view of these slices in ΔV_P ; (e), (f) and (g) are the Y-Z, X-Z and X-Y slices of inverted model perturbation ΔV_S at X=400 m, Y=400 m and Z=-320 m, respectively; (h) is the 3D view of these slices in ΔV_S .



Figure 14. (a) shows the Y-Z slice of the inverted model perturbation ΔQ_P at X=400 m; (b) shows the X-Z slice of the inverted model perturbation ΔQ_P at Y=400 m; (c) shows the X-Y slice of the inverted model perturbation ΔQ_P at Z=-100 m; (d) shows the 3D view of these slices in ΔQ_P ; (e), (f) and (g) are the Y-Z, X-Z and X-Y slices of inverted model perturbation ΔQ_S at X=400 m, Y=400 m and Z=-100 m, respectively; (h) is the 3D view of these slices in ΔQ_S .



Figure 15. (a), (b) and (c) show the shot gathers of P waves calculated using target, initial and inverted models with source at (X=100 m, Y=200 m, Z=0 m); (d), (e) and (f) show the corresponding shot gathers of S waves; (g) shows the P waves recorded at A (Z=-250 m) calculated using target (black PA-T), initial (blue PA-I) and inverted models (purple PA-CD); (h) shows the S waves recorded at A calculated using target (black S-A-T), initial (blue S-A-I) and inverted models (purple S-A-CD); (i) shows the P waves recorded at B (Z=-300 m) calculated using target (black PB-T), initial (blue PB-I) and inverted models (purple PB-CD); (j) shows the S waves recorded at B calculated using target (black S-B-T), initial (blue S-B-I) and inverted (purple S-B-CD); (k), (l), (m) and (m) are the corresponding amplitude spectra of the P and S waves in (g), (h), (i) and (j). f_c (Hz) indicate central-frequencies. Colors of the amplitude spectra and f^c are consistent with those of the waves.

Appendix A Review of forward modeling in viscoelastic medium based on GSLS model

Different physical mechanisms and approaches (Kolsky, 1952; Biot, 1956; Futterman, 1962; Emmerich & Korn, 1987; Carcione et al., 1988a, 1988b; Robertsson et al., 1994; T. Zhu & Carcione, 2014; N. Wang et al., 2019) have been suggested to model the anelastic behaviors of propagating seismic waves in real Earth's medium. The system constructed by the parallel connection of several standard linear elements is referred to as the generalized standard linear solid rheology (H. Liu et al., 1976; Aki & Richards, 2002). In this study, the GSLS model is adopted as the forward modeling engine for adjoint Q tomography in time domain. The convolution integral in the constitutive relationship (equation (2)) can be eliminated by taking partial derivative of time on its both sides, which yields (Blanch et al., 1995)

$$\partial_t T_{ij} = \kappa^R \delta_{ij} \delta_{kl} \left[(\tau_\kappa + 1) \, \partial_t \varepsilon_{kl} - \sum_{p=1}^P \tau_\kappa \epsilon_{kl}^p \right] + \mu^R \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) \left[(\tau_\mu + 1) \, \partial_t \varepsilon_{kl} - \sum_{p=1}^P \tau_\mu \epsilon_{kl}^p \right],$$
(A1)

where ϵ_{kl}^p are the introduced memory strain variables describing anelastic characteristics of the seismic waves (Pan & Wang, 2020):

$$\epsilon_{kl}^{p} = \frac{1}{P\tau^{\sigma p}} e^{-t/\tau^{\sigma p}} H\left(t\right) * \partial_{t} \varepsilon_{kl},\tag{A2}$$

and time derivative of the memory variables satisfies the following equation (Pan & Wang, 2020):

$$\partial_t \epsilon_{kl}^p = \frac{1}{P \tau^{\sigma p}} \partial_t \varepsilon_{kl} - \frac{1}{\tau^{\sigma p}} \epsilon_{kl}^p. \tag{A3}$$

Thus, wave propagation in viscoelastic medium can be modeled by numerically solving the equation of motion and equations (A1) and (A3) with the superposition of several parallel relaxation mechanisms. The quality factor Q is defined in frequency domain as (Blanch et al., 1995; Bohlen, 2002)

$$\tilde{Q}_{M}^{-1}(\omega,\tau_{M},\tau^{\sigma p}) = \frac{\mathcal{I}[M(\omega)]}{\mathcal{R}[M(\omega)]} = \left[\tau_{M}\sum_{p=1}^{P}\frac{\omega\tau^{\sigma p}}{1+(\omega\tau^{\sigma p})^{2}}\right] \left[1+\tau_{M}\sum_{p=1}^{P}\frac{(\omega\tau^{\sigma p})^{2}}{1+(\omega\tau^{\sigma p})^{2}}\right]^{-1},$$
(A4)

where M indicates Fourier transform of the time derivative of the modulus relaxation function, ω denotes angular frequency, \mathcal{I} means imaginary part.

440

441

To approximate a nearly constant value of Q_{ref} within the seismic frequency band by several Maxwell bodies, the relaxation parameters τ_M and $\tau^{\sigma p}$ can be obtained by minimizing the distance between Q_{ref}^{-1} and \tilde{Q}_M^{-1} (equation (A4)). Emmerich and Korn (1987) solved this problem using the classical linear optimization approach. Fichtner and van Driel (2014) obtained the relaxation parameters with Monte Carlo approaches and found that the optimal relaxation parameters can be determined with the enforcement of $\tau_M = Q_M^{-1}$ for the whole set of Q values. In this study, following Blanc et al. (2016), we introduce the objective function:

$$\mathcal{J}\left(\left\{\tau_{M},\tau^{\sigma p}\right\};P\right) = \int_{\omega_{0}}^{\omega'} \left\{\sum_{p=1}^{P} \frac{\tau_{M}}{P} \frac{\omega \tilde{Q}_{\mathrm{ref}}\left(\omega\right) \left[\left(\tau^{\sigma p}\right)^{-1} - \omega \tilde{Q}_{\mathrm{ref}}^{-1}\left(\omega\right)\right]}{\left(\tau^{\sigma p}\right)^{-2} + \omega^{2}} - 1\right\}^{2} d\omega, \qquad (A5)$$

where ω_0 and ω' indicate the minimum and maximum angular frequencies. The relaxation parameters are solved using the *SolvOpt* algorithm (Kappel & Kuntsevich, 2000) based on the iterative Sher's method (Sher, 1985) with positivity constraints of σ

based on the iterative Shor's method (Shor, 1985) with positivity constraints of $\tau_M >$

⁴⁴⁵ 0 and $\tau^{\sigma p} > 0$ (see section 2.3 in Blanc et al. (2016)). Similarly, with the enforcement ⁴⁴⁶ of $\tau_M = Q_M^{-1}$, Q can be explicitly introduced in the constitutive relation and equation ⁴⁴⁷ of motion, which enables and facilitates the derivation of Q sensitivity kernels in adjoint ⁴⁴⁸ tomography. In this study, a number of 3 relaxation mechanisms is used to approximate ⁴⁴⁹ constant Q in the forward modeling experiments.

Appendix B Hessian-vector product calculation in viscoelastic medium with the adjoint-state method

Products of multiparameter Hessian with an arbitrary vector can be calculated efficiently with the first-order adjoint-state method (Métivier et al., 2013; Pan et al., 2018). We can first consider minimizing the following augmented Lagrangian misfit function:

$$\tilde{\boldsymbol{\chi}}(\mathbf{m},\mathbf{u},\boldsymbol{\Lambda}) = \sum_{\mathbf{x}_r} \int_0^{t'} \int_{\Omega} \left[\mathbf{u}^{\ddagger} v - \boldsymbol{\Lambda} \cdot \left(\rho \partial_t^2 \mathbf{u} - \nabla \cdot \mathbf{T} - \boldsymbol{f}^s \right) \right] d\mathbf{x} dt, \tag{B1}$$

where "‡" means complex conjugate, Λ is a new Lagrangian multiplier and v is an arbitrary function. Variation of misfit function with respect to the perturbations of model parameters ($\Delta \kappa$, $\Delta \mu$, ΔQ_{κ} and ΔQ_{μ}) and wavefield (Δu_n) can be obtained as

$$\begin{split} \Delta \tilde{\boldsymbol{\chi}} \left(\mathbf{m}, \mathbf{u}, \mathbf{\Lambda} \right) &= \int_{0}^{t} \int_{\Omega} \sum_{\mathbf{x}_{r}} \left[v - \rho \partial_{t}^{2} \Lambda_{i} + \partial_{j} \left(\kappa \delta_{ij} \delta_{kl} * \partial_{t} \tilde{\Gamma}_{kl} \right) \right] \Delta u_{n} d\mathbf{x} dt \\ &+ \int_{0}^{t'} \int_{\Omega} \partial_{j} \left[\mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) * \partial_{t} \tilde{\Gamma}_{kl} \right] \Delta u_{n} d\mathbf{x} dt \\ &- \int_{0}^{t'} \int_{\Omega} \left[\left(\Delta \kappa * \partial_{t} \tilde{\Gamma}_{kk} \right) \Gamma_{ii} - \left(\Delta Q_{\kappa} \frac{\kappa^{R}}{Q_{\kappa}^{2}} \sum_{p=1}^{P} \tau^{\sigma p} \tilde{\Pi}_{kk}^{p} \right) \varepsilon_{ii} \right] d\mathbf{x} dt \end{split}$$
(B2)
$$&- \int_{0}^{t'} \int_{\Omega} \left[\Delta \mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) * \partial_{t} \tilde{\Gamma}_{kl} \right] \varepsilon_{ij} d\mathbf{x} dt \\ &+ \int_{0}^{t'} \int_{\Omega} \left[\Delta Q_{\mu} \frac{\mu^{R}}{Q_{\mu}^{2}} \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) \sum_{p=1}^{P} \tau^{\sigma p} \tilde{\Pi}_{kl}^{p} \right] \varepsilon_{ij} d\mathbf{x} dt, \end{split}$$

where $\tilde{\Gamma}_{ij} = \partial_j \Lambda_i$ is the Lagrangian strain field and $\tilde{\Pi}_{kl}^p$ are the corresponding Lagrangian memory strain variables. In equation (B2), we ignore density ρ for compactness. The adjoint-state equation can be obtained by setting the coefficient of Δu_n as zero:

$$\rho \partial_t^2 \Lambda_i - \partial_j \left[\kappa \delta_{ij} \delta_{kl} * \partial_t \tilde{\Gamma}_{kl} + \mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} - \frac{2}{3} \delta_{ij} \delta_{kl} \right) * \partial_t \tilde{\Gamma}_{kl} \right] = v, \qquad (B3)$$

where v serves as the adjoint source and solution of equation (B3) can be obtained by convolving Green's function G_{in} with v:

$$\Lambda_i(t) = \int_0^{t'} G_{in}(t - t'') v(t) dt''.$$
 (B4)

To calculate the Hessian-vector product, we can first derive the gradients of misfit function (equation (B1)) with respect to the model parameters. For example, gradient for κ is:

$$\nabla_{\kappa} \tilde{\boldsymbol{\chi}} = -\sum_{\mathbf{x}_{r}} \int_{\Omega} \int_{0}^{t'} \int_{0}^{t'} \Gamma_{ii}\left(t\right) \partial_{t} \partial_{k} G_{kn}\left(t'-t-t''\right) v\left(t''\right) dt'' dt d\mathbf{x}.$$
 (B5)

Product of Gauss-Newton Hessian with model perturbation vector can be calculated by replacing v in equation (B5) with Jacobian-vector product. For example, product of the Jacobian matrix with model perturbation vector $\Delta \kappa$ is

$$\mathbf{J}_{\kappa}\left(t^{\prime\prime}\right) = -\int_{0}^{t^{\prime}} \partial_{t} \partial_{k^{\prime}} G_{k^{\prime}n^{\prime}}\left(t^{\prime\prime} - t^{\prime\prime\prime}\right) \Delta \kappa \Gamma_{i^{\prime}i^{\prime}}\left(t^{\prime\prime\prime}\right) dt^{\prime\prime\prime}.$$
 (B6)

Then, replacing v in equation (B5) with equation (B6) gives the Hessian-vector product $\mathbf{H}_{\kappa\kappa}\Delta\kappa$:

$$\mathbf{H}_{\kappa\kappa}\Delta\kappa = \sum_{\mathbf{x}_{r}} \int_{\Omega} \int_{0}^{t'} \int_{0}^{t'} \Gamma_{ii}\left(t\right) \partial_{t}\partial_{k}G_{kn}\left(t'-t-t''\right) \\ \times \int_{0}^{t'} \partial_{t}\partial_{k'}G_{k'n'}\left(t''-t'''\right)\Delta\kappa\Gamma_{i'i'}\left(t'''\right)dt'''dt''dtd\mathbf{x}.$$
(B7)

452 More detailed introduction for calculating the Hessian-vector product can be found in

Appendix B of Pan et al. (2018). The multiparameter Hessian-vector products associated with other physical parameters or misfit functions can be derived following the same procedure.

456 **References**

463

464

465

469

- Ajo-Franklin, J. B., Dou, S., Lindsey, N. J., Monga, I., Tracy, C., Robertson, M., ...
 Li, X. (2019). Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection. *Scientific Reports*, 9, 1–14.
- Aki, K., & Richards, P. G. (2002). *Quantitative seismology* (2nd ed.). University
 Science Books.
 - Alkhalifah, T., & Plessix, R. E. (2014). A recipe for practical full-waveform inversion in anisotropic media: An analytic parameter resolution study. *Geophysics*, 79(3), R91–R101.
- Bao, X., A, D. C., Jin, G., James, B. J., & Shen, Y. (2016). Imaging Rayleigh
 wave attenuation with USArray. *Geophysical Journal International*, 206(1),
 241–259.
 - Barnes, A. E. (1993). Instantaneous spectral bandwidth and dominant frequency with applications to seismic reflection data. *Geophysics*, 58(3), 419–428.
- Bath, M. (1974). Spectral analysis in geophysics: Developments in solid earth geo physics. Elsevier Science Publishing Co.
- Berkhout, A. J. (1984). Seismic resolution: Resolving power of acoustical echo tech niques. Geophysical Press.
- Berryman, J. G. (1988). Seismic wave attenuation in fluid-saturated porous media.
 Pure Appl. Geophys., 128, 423–432.
- Biot, M. A. (1956). Theory of propagation of elastic waves in a fluid-saturated porous solid, I, Low-frequency range. J. Acoust. Soc. Am., 28(0), 168–178.
- ⁴⁷⁹ Blanc, E., Komatitsch, D., Chaljub, E., Lombard, B., & Xie, Z. (2016). Highly
 ⁴⁸⁰ accurate stability-preserving optimization of the Zener viscoelastic model,
 ⁴⁸¹ with application to wave propagation in the presence of strong attenuation.
 ⁴⁸² *Geophysical Journal International*, 205(1), 427–439.
- Blanch, J. O., Robertsson, J. O. A., & Symes, W. W. (1995). Modeling of a constant Q: Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique. *Geophysics*, 60(1), 176–184.
- Bohlen, T. (2002). Parallel 3-D viscoelastic finite difference seismic modelling. Com *puter & Geosciences*, 28, 845–870.
- Bostock, M. G., Rondenay, S., & Shragge, J. (2001). Multiparameter twodimensional inversion of scattered teleseismic body waves 1. Theory for oblique incidence. Journal of Geophysical Research, 106(12), 30771–30782.
- Brossier, R. (2011). Two-dimensional frequency-domain visco-elastic full wave form inversion: Parallel algorithms, optimization and performance. Computer
 & Geosciences, 37, 444–445.
- Brossier, R., Operto, S., & Virieux, J. (2009). Robust frequency-domain full waveform inversion using the l1 norm. Geophysical Research Letters, 36,
 L20310.

- Cao, A., & Romanowicz, B. (2004). Hemispherical transition of seismic attenuation at the top of the Earth's inner core. *Earth Planet. Sci. Lett.*, 228, 243–253.
- Carcione, J. M., Kosloff, D., & Kosloff, R. (1988a). Viscoacoustic wave propagation
 simulation in the earth. *Geophysics*, 53, 769–777.
- Carcione, J. M., Kosloff, D., & Kosloff, R. (1988b). Wave propagation simulation
 in a linear viscoelastic medium. *Geophysical Journal International*, 95(3), 597–611.
 - Charara, M., Barnes, C., & Tarantola, A. (2000). Full waveform inversion of seismic data for a viscoelastic medium, in Methods and Applications of Inversion. Springer.

505

- Chen, M., Niu, F., Tromp, J., Lenardic, A., Lee, C. A., Cao, W., & Ribeiro, J.
 (2017). Lithospheric foundering and underthrusting imaged beneath Tibet.
 Nature Communications, 8(0), 15659.
- Chen, P., Zhao, L., & Chapman, C. H. (2007). Full 3D tomography for the crustal
 structure of the Los Angeles region. Bulletin of the Seismological Society of
 America, 97(4), 1094–1120.
- ⁵¹³ Cox, B., Wills, P., Kiyashchenko, D., Jeff, M., Jorge, L., Bourne, S., ... Roy, J.
 (2012). Distributed Acoustic Sensing for Geophysical Measurement, Monitoring and Verification. CSEG Recorder, 37, 7–13.
- Dahlen, F. A., & Tromp, J. (1998). Theoretical global seismology. Princeton University Press.
- Daley, T. M., Miller, D. E., Dodds, K., Cook, P., & Freifeld, B. M. (2016). Field
 testing of modular borehole monitoring with simultaneous distributed acoustic
 sensing and geophone vertical seismic profiles at Citronelle, Alabama. Geo physical Prospecting, 64, 1318–1334.
- Debayle, E., Bodin, T., & Durand, S. (2020). Seismic evidence for partial melt below
 tectonic plates. *Nature*, 586(7830), 555–559.
- ⁵²⁴ Dong, X., & Yang, D. (2020). Reanimating lithosphere of North China: new insights ⁵²⁵ from full-waveform seismic tomography. *Science Bulletin*, 65(10), 775–777.
- Eaid, M. V., Keating, S. D., & Innanen, K. A. (2020). Multiparameter seismic
 elastic full-waveform inversion with combined geophone and shaped fiber-optic
 cable data. *Geophysics*, 85(6), R537–R552.
- Eberhart-Phillips, D., & Chadwick, M. (2002). Three-dimensional attenuation model
 of the shallow Hikurangi subduction zone in the Raukumara Peninsula, New
 Zealand. J. Geophys. Res., 107(B2), 2033.
- Emmerich, H., & Korn, M. (1987). Incorporation of attenuation into time-domain
 computations of seismic wave fields. *Geophysics*, 52, 1252–1264.
- Fichtner, A., & Leeuwen, T. (2015). Resolution analysis by random probing. J. Geophys. Res. Solid Earth, 120, 5549–5573.
- Fichtner, A., & Trampert, J. (2011a). Hessian kernels of seismic data functionals based upon adjoint techniques. *Geophysical Journal International*, 185(2), 775–798.
- Fichtner, A., & Trampert, J. (2011b). Resolution analysis in full waveform inversion.
 Geophysical Journal International, 187(3), 1604–1624.
- Fichtner, A., & van Driel, M. (2014). Models and Fréchet kernels for frequency (in)dependent Q. Geophysical Journal International, 198, 1878–1889.
- Futterman, W. I. (1962). Dispersive body waves. Journal of Geophysical Research,
 67, 5279–5291.
- Hall, K. W., Bertram, K. L., Bertram, M. B., Innanen, K. A., & Lawton, D. C.
- (2019). Simultaneous accelerometer and optical fibre multi-azimuth walk-away
 VSP experiment, Newell County, Alberta, Canada. SEG Technical Program
 Expanded Abstracts, 5340–5344.
- Hauge, P. (1981). Measurements of attenuation from vertical seismic profiles. Geophysics, 46, 1548–1558.
- Hauksson, E., & Shearer, P. S. (2006). Attenuation models (Q_P and Q_S) in three

552	dimensions of the southern California crust: Inferred fluid saturation at seis-
553	mogenic depths. J. Geophys. Res., 111, B05302.
554	He, Q., & Wang, Y. (2020). Inexact Newton-type methods based on Lanczos or-
555	thonormal method and application for full waveform inversion. Inverse Prob-
556	lems, 36, 115007.
557	Innanen, K. A. (2011). Inversion of the seismic AVF/AVA signatures of highly atten-
558	uative targets. $Geophysics$, $76(1)$, R1–R14.
559	Innanen, K. A. (2014). Seismic AVO and the inverse Hessian in precritical reflection
560	full waveform inversion. Geophysical Journal International, 199, 717–734.
561	Kamei, R., & Pratt, G. R. (2013). Inversion strategies for visco-acoustic waveform
562	inversion. Geophysical Journal International, $194(2)$, $859-884$.
563	Kappel, F., & Kuntsevich, A. (2000). An implementation of Shor's r-algorithm.
564	Comput. Optim. Appl., 15(2), 193–205.
565	Karaoğlu, H., & Romanowicz, B. (2018). Global seismic attenuation imaging using
566	full-waveform inversion: a comparative assessment of different choices of misfit
567	functionals. Geophysical Journal International, 212(2), 807–826.
568	Kazei, V., & Alkhalifah, T. (2019). Scattering radiation pattern atlas: What
569 570	anisotropic elastic properties can body waves resolve? J. Geophys. Res. Solid Earth, 124(3), 2781–2811.
571	Keating, S., & Innanen, K. A. (2020). Parameter crosstalk and leakage between spa-
572	tially separated unknowns in viscoelastic full-waveform inversion. Geophysics,
573	85(4), R397 - R408.
574	Kolsky, H. (1952). The propagation of stress pulses in viscoelastic solids. Philosophi-
575	cal Magazine, 1, 693–710.
576	Komatitsch, D., & Tromp, J. (2005). Introduction to the spectral-element method
577	for 3-D seismic wave propagation. Geophysical Journal International, 139,
578	806–822.
579	Krischer, L., Fichtner, A., Boehm, C., & Igel, H. (2018). Automated Large-Scale
580	Full Seismic Waveform Inversion for North America and the North Atlantic. J .
581	Geophys. Res. Solid Earth, 123 , $5902-5928$.
582	Liu, H., Anderson, D. L., & Kanamori, H. (1976). Velocity dispersion due to aneias- ticity implications for science and months composition. Combus. I. P. As
583	tron Soc $\sqrt{7(1)}$ 41–58
504	Liu O & Tromp I (2006) Finite-Frequency Kernels Based on Adjoint Methods
586	Bulletin of Seismological Society of America, 96(6), 2383–2397.
587	Liu, Q., & Tromp, J. (2008). Finite-frequency sensitivity kernels for global seismic
588	wave propagation based upon adjoint methods. Geophysical Journal Interna-
589	tional, 174, 265-286.
590	Luo, Y., & Schuster, G. 1. (1991). Wave-equation traveitime inversion. <i>Geophysics</i> , 56, 645, 652
591	50, 045-055.
592	Ceonbusics //(0) 161-178
595	Métivier I. Brossier B. Virieux I. & Operto S. (2013) Full waveform inver-
505	sion and the truncated Newton method SIAM Journal On Scientific Comput-
596	ing. 35, B401–B437.
597	Mulder, W. A., & Hak, B. (2009). An ambiguity in attenuation scattering imaging.
598	Geophysical Journal International, 178, 1614–1624.
599	Nocedal, J., & Wright, S. J. (2006). Numerical optimization. Springer.
600	Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Métivier, L., &
601	Virieux, J. (2013). A guided tour of multiparameter full waveform inversion
602	with multicomponent data: from theory to practice. The Leading Edge, 32,
603	1040 - 1054.
604	Operto, S., & Miniussi, A. (2018). On the role of density and attenuation in 3D
605	multi-parameter visco-acoustic VTI frequency-domain FWI: an OBC case

606	study from the North Sea.	Geophysical Journal	$\ International,$	213, 2037-
607	2059.			

- Pan, W., Geng, Y., & Innanen, K. A. (2018). Interparameter trade-off quantifi cation and reduction in isotropic-elastic full-waveform inversion: synthetic
 experiments and Hussar land data set application. *Geophysical Journal Inter- national*, 213, 1305–1333.
- Pan, W., Innanen, K. A., Margrave, G. F., Fehler, M. C., Fang, X., & Li, J. (2016).
 Estimation of elastic constants for HTI media using Gauss-Newton and full Newton multiparameter full-waveform inversion. *Geophysics*, 81(5), R275–
 R291.
 - Pan, W., Innanen, K. A., & Wang, Y. (2020). SeisElastic2D: An open-source package for multiparameter full-waveform inversion in isotropic-, anisotropic- and visco-elastic media. *Computers & Geosciences*, 145, 104586.

617

618

636

637

655

- Pan, W., & Wang, Y. (2020). On the influence of different misfit functions for atten uation estimation in viscoelastic full-waveform inversion: synthetic study. *Geo- physical Journal International*, 221(2), 1292–1319.
- Plessix, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167, 495–503.
- Quan, Y., & Harris, J. M. (1997). Seismic attenuation tomography using the frequency shift method. *Geophysics*, 62, 895–905.
- Robertsson, J. O., Blanch, J. O. A., & Symes, W. W. (1994). Viscoelastic finitedifference modeling. *Geophysics*, 59, 1444–1456.
- Romanowicz, B. (1995). A global tomographic model of shear attenuation in the upper mantle. J. Geophys. Res. Solid Earth, 100(B7), 12375–12394.
- Sager, K., Ermert, L., Boehm, C., & Fichtner, A. (2018). Towards full waveform am bient noise inversion. *Geophysical Journal International*, 212(1), 566–590.
- Shen, Y., Biondi, B., & Clapp, R. (2018). Q-model building using one-way wave equation migration Q analysis Part 1: Theory and synthetic test. Geo physics, 83(2), S93–S109.
 - Shor, N. (1985). *Minimization Methods for Non-Differentiable Functions*. Springer-Verlag.
- Tape, C., Liu, Q., Maggi, A., & Tromp, J. (2009). Adjoint tomography of the south ern California crust. *Science*, 325, 988–992.
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approxima tion. Geophysics, 49, 1259–1266.
- Tarantola, A. (1986). A strategy for nonlinear elastic inversion of seismic reflection
 data. *Geophysics*, 51, 1893-1903.
- Tong, P., Zhao, D., Yang, D., Yang, X., Chen, J., & Liu, Q. (2014). Wave-equationbased travel-time seismic tomography-Part 1: Method. Solid Earth, 5(2),
 1151–1168.
- Trampert, J., & Woodhouse, J. H. (1995). Global phase-velocity maps of Love
 and Rayleigh-waves between 40 and 150 seconds. *Geophysical Journal Interna*tional, 122(2), 675–690.
- Tromp, J., & Bachmann, E. (2019). Source encoding for adjoint tomography. Geophysical Journal International, 218(3), 2019–2044.
- Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time
 reversal, and banana-doughnut kernels. *Geophysical Journal International*,
 160, 195-216.
 - Virieux, A., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6), WCC1–WCC26.
- Wang, N., Li, J., Borisov, D., Gharti, H. N., Shen, Y., Zhang, W., & Savage, B.
 (2019). Modeling Three-Dimensional Wave Propagation in Anelastic Models
 With Surface Topography by the Optimal Strong Stability Preserving Runge-Kutta Method. J. Geophys. Res. Solid Earth, 124(1), 890–907.

- Wang, Y., Ning, Y., & Wang, Y. (2020). Fractional Time Derivative Seismic Wave
 Equation Modeling for Natural Gas Hydrate. *Energies*, 13(22), 5901.
- Wang, Y., & Yuan, Y. (2005). Convergence and regularity of trust region methods for nonlinear ill-posed inverse problems. *Inverse Problems*, 21, 821–838.
- Wu, R. (1989). The perturbation method in elastic wave scattering. *Pure and Applied Geophysics*, 131(4), 605–637.
- Yao, G., Silva, N. V., Warner, M., & Kalinicheva, T. (2018). Separation of migration
 and tomography modes of full-waveform inversion in the plane wave domain.
 J. Geophys. Res. Solid Earth, 123, 1486–1501.
- Yu, X., Ahmadinia, M., Shariatipour, S. M., Lawton, D., Osadetz, K., & Saeedfar,
 A. (2019). Impact of Reservoir Permeability, Permeability Anisotropy and
 Designed Injection Rate on CO₂ Gas Behavior in the Shallow Saline Aquifer
 at the CaMI Field Research Station, Brooks, Alberta. Natural Resources
 Research, 29(4), 2735–2752.
- Yuan, Y. O., Simons, F. J., & Bozdağ, E. (2015). Multiscale adjoint waveform to mography for surface and body waves. *Geophysics*, 80(5), R281–R302.
 - Yuan, Y. O., Simons, F. J., & Tromp, J. (2016). Double-difference adjoint seismic tomography. *Geophysical Journal International*, 206(3), 1599–1618.

- Zhu, H., Bozdağ, E., Duffy, T. S., & Tromp, J. (2013). Seismic attenuation beneath
 Europe and the North Atlantic: Implications for water in the mantle. Earth
 and Planetary Science Letters, 381, 1–11.
- ⁶⁸² Zhu, H., & Tromp, J. (2013). Mapping tectonic deformation in the crust and upper ⁶⁸³ mantle beneath Europe and the North Atlantic Ocean. *Science*, 341, 871–875.
- Zhu, T., Ajo-Franklin, J. B., & Daley, T. M. (2017). Spatiotemporal changes of seismic attenuation caused by injected CO2 at the Frio-II pilot site, Dayton, TX, USA. J. Geophys. Res. Solid Earth, 122, 7156–7171.
- Zhu, T., & Carcione, J. M. (2014). Theory and modelling of constant-Q P- and S waves using fractional spatial derivatives. *Geophysical Journal International*, 196, 1787–1795.
- Zhu, T., Harris, J. M., & Biondi, B. (2014). Q-compensated reverse-time migration.
 Geophysics, 79(3), R471–R484.