Geophysical Journal International

Geophys. J. Int. (2014) **199,** 717–734 doi: 10.1093/gji/ggu291
GJI Seismology

Seismic AVO and the inverse Hessian in precritical reflection full waveform inversion

Kristopher A. Innanen

Department of Geoscience, University of Calgary, Calgary, Alberta, Canada, T2N 1N4. E-mail: k.innanen@ucalgary.ca

Accepted 2014 July 23. Received 2014 July 22; in original form 2014 February 1

SUMMARY

We analyse the role played by amplitude-variation-with-offset (AVO) information in the construction of full waveform inversion (FWI) updates from pre-critical seismic reflection data. A mixed continuous-discrete formulation of frequency-domain FWI is found to conveniently expose issues such as parameter cross-talk to quantitative interpretation. Two types of approximate inverse Hessian operator, differing in which off-diagonal elements are retained and which are neglected, are derivable within this formulation. The first, which emphasizes correlations between different parameters collocated in space, is referred to as an approximation of *parameter-type*. The second, which emphasizes correlations between spatially separated elements, is referred to as an approximation of *space-type*. The two approximations are such that, if they are both made simultaneously, the result is an update involving only the diagonal elements of the Hessian. The parameter-type inverse Hessian approximation appears to be capable of performing the kind of data manipulations through which linearized AVO-inversion and inverse scattering methods avoid cross-talk. This is confirmed with an analytic calculation of the first iteration in the reconstruction of a single acoustic boundary. The parameter-type FWI update, in this special case, reduces to linear admixtures of reflection coefficient versus angle which are consistent with those produced by direct AVO inversion. The parameter-type components in the full inverse Hessian represent, in this sense, the generalization of standard linearized AVO inversion within FWI.

Key words: Inverse theory; Theoretical seismology; Wave propagation.

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

1 INTRODUCTION

In seismic full waveform inversion, or FWI (Lailly 1983; Tarantola 1984), it is understood that the inverse Hessian operator corrects many of the amplitude-related errors for which gradient-based methods are criticized. The elements of the inverse Hessian as retained in Gauss-Newton methods (Pratt *et al.* 1998), have, for instance, been tied to the task of illumination compensation (Shin *et al.* 2001; Virieux & Operto 2009) . In problems involving updates in multiple parameters, the importance of the Hessian is even greater. For instance, a multiparameter gradient-based update is exposed to cross-talk, or the attribution of data amplitudes to variations in the wrong medium properties. It is evident that the off-diagonal elements of the inverse Hessian operator are involved in the resolution of this type of amplitude ambiguity (Operto *et al.* 2013). Consider, for instance, that there are N discrete points in the Earth to which we assign M independent parameters, such that the Hessian is of size NM×NM, organizable into an M×M block matrix of N×N submatrices. Only in the M(M−1) off-diagonal submatrices are terms found which couple different parameters. Therefore, only these parts of the inverse Hessian have the mathematical wherewithal to suppress cross-talk.

The increased rate of convergence to be expected from cross-talk suppression is in principle very significant, when measured in numbers of iterations. Unfortunately, with the inclusion of the inverse Hessian, the associated increase in computation per iteration is generally so large as to reverse any savings. Trade-offs between convergence rate and per-iteration computational cost can be made by approximating the Hessian, but here caution is needed. Gauss-Newton iteration, in neglecting many important off-diagonal Hessian elements, will for instance not account for cross-talk. Approximate inverse Hessians must be carefully constructed if they are not to lose this or any other subtle feature.

FWI has advanced significantly in recent years, on exploration and monitoring scales and beyond (Bleibinhaus *et al.* 2007; Fichtner 2010). Progress has been reported in multiparameter inversion (Brossier *et al.* 2010; Operto *et al.* 2013; Plessix *et al.* 2013), and in inversion of reflection data (Brossier *et al.* 2013; Wang *et al.* 2013), but a fully realized methodology for pre-critical reflection data has not as yet been achieved, in part because of longstanding issues such as missing wavelengths (Kelly *et al.* 2013). There is evidence, however, that the

© The Authors 2014. Published by Oxford University Press on behalf of the Royal Astronomical Society

incorporation of well-control could be a practical source for missing wavelengths (Margrave *et al.* 2012b), especially when supported by seismic surveys in which we generate and measure increasingly low temporal frequencies (Margrave *et al.* 2012a; Plessix *et al.* 2012).

Assuming that inclusion of pre-critical reflection modes in FWI is a meaningful possibility, there remain pressing theoretical questions. Several decades of experience in the geophysics community with amplitude-variation-with-offset, or AVO (Castagna & Backus 1993; Foster *et al.* 2010), have clarified that reflection data amplitudes require multiple parameters to be properly accounted for. Managing issues like parameter cross-talk, therefore, will be central to seamless incorporation of pre-critical reflection data amplitudes into FWI. AVO inversion methods (Hampson & Russell 1990) and their inverse scattering generalizations (Raz 1981; Clayton & Stolt 1981), are, to first order, free of cross-talk. Weighted differences of reflection strengths across angle, offset or wavenumber are employed in these algorithms to separate out the influence of each parameter individually. Does a properly assembled full-Newton FWI update do the same thing automatically? Where in the inverse Hessian operators are to be found the mechanisms which suppress pre-critical cross-talk? How, in other words, is AVO information to be incorporated into FWI?

Our purpose here is to quantitatively address these questions, by formulating a version of frequency-domain FWI which clearly exposes the role of the Hessian and gradient in managing pre-critical reflection amplitudes. We adopt a mixed continuous-discrete approach, in which the distributions of Earth properties in space are treated continuously but individual parameters are incorporated with discrete mathematical quantities. Close analysis of the Hessian has, in other FWI studies (Fichtner & Trampert 2011; Operto *et al.* 2013; Métivier *et al.* 2013, 2014), allowed physical (i.e. scattering) interpretations of approximate forms to be carried out; something similar becomes possible in our scheme. We obtain two simple approximations of the Hessian operator, each leading to a different approximate update formula. We refer to these as parameter-type and space-type approximations.

The central purpose in this paper being to extract insight into FWI, we next examine the parameter-type update formula as applied to an idealized problem. We consider the reconstruction of a 1-D earth model, in which acoustic parameters κ (bulk modulus) and ρ (density) vary across a single horizontal interface, by minimizing an objective function defined in terms of a discrete sum over lateral wavenumbers, or equivalently angles, and a continuous sum over temporal frequencies. There are three reasons for this choice. First, in this configuration there is no phase mismatch between measured and modelled fields, and so it effectively isolates the AVO aspect of the problem of interest. Secondly, every ingredient in this problem, including the data (Weglein *et al.* 1986), can be supplied analytically, and in terms of ideal results, which are in this case the FWI step-lengths taking us directly to the correct answer. The output of any provisional FWI scheme can thus be judged by comparing it against this ideal. And third, the acoustic two-parameter problem, being significantly simpler than the three-parameter elastic problem, but containing the same essential ideas, acts as a manageable proxy for realistic seismic amplitude analysis.

We identify several important properties of the inverse Hessian components associated with the parameter-type approximation:

(i) *Computation and storage.* It is not strictly necessary to construct, store and invert the Hessian operator in practice (Saad 2003; Operto *et al.* 2006; Brossier *et al.* 2009; Métivier *et al.* 2013); nevertheless it is worth pointing out that this operator in its raw form is of size NM×NM (M and N defined as in the first paragraph). In contrast, each iteration of an approximate scheme involving parameter-type inverse Hessian elements requires an M×M matrix for each of the N output model elements, which is a significant decrease.

(ii) *The internal ability to diagnose ill-posedness.* Reflection data involving, for example, a single fixed incidence angle, cannot constrain variations in more than one parameter. Neither gradient-based methods nor methods involving solely diagonal Hessian elements have internal checks to confirm that an input data set is sufficient in this sense. Updates involving parameter-type inverse Hessian components are, in contrast, halted automatically in the presence of insufficient data.

(iii) *Balanced updates.* Convergence of the iterative FWI procedure could be hindered if by using, for instance, different groups of angle values, different step-lengths are determined. The parameter-type update contains components of the inverse Hessian which act to balance the update across angles.

(iv) *Cross-talk suppression.* A primary concern in linearized inverse scattering research (Raz 1981; Clayton & Stolt 1981; Stolt & Weglein 1985; Weglein *et al.* 1986; Innanen & Weglein 2007) is to discover if and how multiple parameters can be individually determined from reflection amplitudes, without cross-talk (the term cross-talk is not itself used in those references). The inverse scattering framework has been understood from such investigations to generalize AVO inversion (Stolt & Weglein 1985; Weglein *et al.* 2003). Parameter-type inverse Hessian components endow the FWI update with a level of protection from cross-talk comparable to that of standard linear AVO inversion/inverse scattering.

In Section 2 we set out the mixed continuous-discrete FWI formulation, beginning with simple univariate, bivariate and trivariate minimization templates and from these expressing minimization schemes for one parameter (e.g. scalar wave), two parameter (e.g. acoustic wave) and three parameter (e.g. isotropic-elastic wave) problems. In Section 3, we show by retaining and/or rejecting certain off-diagonal portions of the inverse Hessian how the parameter-type and space-type updates arise in this formulation. We also show that when both operate simultaneously the update lapses to one involving only the diagonal elements of the Hessian. In Section 4, we analyse the parameter-type approximation by considering the first update to a homogeneous background medium in the reconstruction of a single acoustic boundary. Using analytic data we fill in items (ii) to (iv) on the above list. In Section 5, we show on a 'map' of the inverse Hessian where issues such as cross-talk are managed, and comment on the general consistency of AVO inversion methods and a FWI update of parameter type.

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

2 MINIMIZATION FORMULAS

In this section, we set out mathematical expressions for full Newton FWI updates in forms suitable for the forthcoming approximations and analysis.

2.1 Scalar minimization

In scalar minimization, a scalar functional of one function of space is used:

$$\Phi_1\left(s(\mathbf{r})\right), \tag{1}$$

with s (ultimately) representing the scalar wave velocity as a function of the position vector $\mathbf{r}$. A Taylor's series expansion of the derivative of Φ_1 with respect to the model is truncated at

$$\frac{\partial\Phi_1\left(s+\delta s\right)}{\partial s(\mathbf{r})} \approx g_1(\mathbf{r}) + \int \mathrm{d}\mathbf{r}' H_1(\mathbf{r},\mathbf{r}')\delta s(\mathbf{r}'), \tag{2}$$

where

$$g_1(\mathbf{r}) = \frac{\partial\phi_1\left(s\right)}{\partial s(\mathbf{r})}, \quad H_1(\mathbf{r},\mathbf{r}') = \frac{\partial^2\Phi_1\left(s\right)}{\partial s(\mathbf{r})\partial s(\mathbf{r}')}. \tag{3}$$

The Newton step is the δs for which the left-hand side of eq. (2) is zero. Omitting the $\approx$ sign, we have

$$g_1(\mathbf{r}) = -\int \mathrm{d}\mathbf{r}' H_1(\mathbf{r},\mathbf{r}')\delta s(\mathbf{r}'), \tag{4}$$

which is inverted to obtain the basic functional form for the update:

$$\delta s(\mathbf{r}) = -\int \mathrm{d}\mathbf{r}' H_1^{-1}(\mathbf{r},\mathbf{r}')g_1(\mathbf{r}'), \tag{5}$$

where H_1^{-1} is the inverse of H_1 in the sense that

$$\int \mathrm{d}\mathbf{r}'' H_1^{-1}(\mathbf{r},\mathbf{r}'')H_1(\mathbf{r}'',\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'), \tag{6}$$

and where $\delta(\mathbf{r})$ is the Dirac delta function. At the nth iteration, given a current model $s^{(n)}(\mathbf{r})$ and the nth gradient and Hessian functions, eq. (5) allows us to compute the update

$$s^{(n+1)}(\mathbf{r}) = s^{(n)}(\mathbf{r}) + \delta s^{(n)}(\mathbf{r}). \tag{7}$$

2.2 Acoustic minimization

We next consider a scalar functional of two functions, $s_\kappa(\mathbf{r})$ and $s_\rho(\mathbf{r})$, which will ultimately represent the distributions in space of bulk modulus κ and density ρ:

$$\Phi_2\left(s_\kappa(\mathbf{r}), s_\rho(\mathbf{r})\right). \tag{8}$$

The extension of the scalar relationship in eq. (4) to the acoustic problem involves a 2×2 matrix whose elements are integral operators:

$$\begin{bmatrix} g_\kappa(\mathbf{r}) \\ g_\rho(\mathbf{r}) \end{bmatrix} = -\int \mathrm{d}\mathbf{r}' \begin{bmatrix} H_{\kappa\kappa}(\mathbf{r},\mathbf{r}') & H_{\kappa\rho}(\mathbf{r},\mathbf{r}') \\ H_{\rho\kappa}(\mathbf{r},\mathbf{r}') & H_{\rho\rho}(\mathbf{r},\mathbf{r}') \end{bmatrix} \begin{bmatrix} \delta s_\kappa(\mathbf{r}') \\ \delta s_\rho(\mathbf{r}') \end{bmatrix}, \tag{9}$$

where

$$H_{\kappa\kappa}(\mathbf{r},\mathbf{r}') = \frac{\partial^2\Phi_2\left(s_\kappa,s_\rho\right)}{\partial s_\kappa(\mathbf{r})\partial s_\kappa(\mathbf{r}')}, \quad H_{\kappa\rho}(\mathbf{r},\mathbf{r}') = \frac{\partial^2\Phi_2\left(s_\kappa,s_\rho\right)}{\partial s_\kappa(\mathbf{r})\partial s_\rho(\mathbf{r}')},$$

and

$$H_{\rho\kappa}(\mathbf{r},\mathbf{r}') = \frac{\partial^2\Phi_2\left(s_\kappa,s_\rho\right)}{\partial s_\rho(\mathbf{r})\partial s_\kappa(\mathbf{r}')}, \quad H_{\rho\rho}(\mathbf{r},\mathbf{r}') = \frac{\partial^2\Phi_2\left(s_\kappa,s_\rho\right)}{\partial s_\rho(\mathbf{r})\partial s_\rho(\mathbf{r}')},$$

are the four Hessian functions, extensions of the single scalar function $H_1(\mathbf{r},\mathbf{r}')$. The overall Hessian operator is symmetric; an additional useful symmetry will be found in Section 4. We next form an inverse counterpart to this relationship which, though not general, has merit in that it is easy to analyse. Defining a generalized determinant of the form

$$\mathcal{H}_2(\mathbf{r},\mathbf{r}') = \int \mathrm{d}\mathbf{r}'' \left[H_{\kappa\kappa}(\mathbf{r},\mathbf{r}'')H_{\rho\rho}(\mathbf{r}'',\mathbf{r}') - H_{\kappa\rho}(\mathbf{r},\mathbf{r}'')H_{\rho\kappa}(\mathbf{r}'',\mathbf{r}')\right], \tag{10}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

the two parameter acoustic Newton step becomes

$$\begin{bmatrix} \delta s_\kappa(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} = \int \mathrm{d}\mathbf{r}' \mathcal{H}_2^{-1}(\mathbf{r},\mathbf{r}') \int d\mathbf{r}'' \begin{bmatrix} -H_{\rho\rho}(\mathbf{r}',\mathbf{r}'') & H_{\kappa\rho}(\mathbf{r}',\mathbf{r}'') \\ H_{\rho\kappa}(\mathbf{r}',\mathbf{r}'') & -H_{\kappa\kappa}(\mathbf{r}',\mathbf{r}'') \end{bmatrix} \begin{bmatrix} g_\kappa(\mathbf{r}'') \\ g_\rho(\mathbf{r}'') \end{bmatrix}, \tag{11}$$

where the elements of the gradient vector

$$\mathbf{g}_2 = \begin{bmatrix} g_\kappa(\mathbf{r}) \\ g_\rho(\mathbf{r}) \end{bmatrix} \tag{12}$$

are, explicitly,

$$g_\kappa(\mathbf{r}) = \frac{\partial \Phi_2(s_\kappa, s_\rho)}{\partial s_\kappa(\mathbf{r})} \quad \text{and} \quad g_\rho(\mathbf{r}) = \frac{\partial \Phi_2(s_\kappa, s_\rho)}{\partial s_\rho(\mathbf{r})}, \tag{13}$$

and where $\mathcal{H}_2^{-1}$ is the inverse of $\mathcal{H}_2$ in the sense that

$$\int \mathrm{d}\mathbf{r}'' \mathcal{H}_2^{-1}(\mathbf{r},\mathbf{r}'')\mathcal{H}_2(\mathbf{r}'',\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'). \tag{14}$$

Eq. (11), though not a general formula for the inversion of eq. (9), is applicable in cases where the integral operators in each element of the 2×2 system in eq. (9) commute with each other. The cases we consider in Sections 3 and 4 all either meet this condition or can be temporarily assumed to do so without introducing inconsistencies into our analysis. After the formula in eq. (11) is evaluated using the nth gradient and Hessian functions, it may be added to the nth model iteration to complete the update:

$$\begin{bmatrix} s_\kappa^{(n+1)}(\mathbf{r}) \\ s_\rho^{(n+1)}(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} s_\kappa^{(n)}(\mathbf{r}) \\ s_\rho^{(n)}(\mathbf{r}) \end{bmatrix} + \begin{bmatrix} \delta s_\kappa^{(n)}(\mathbf{r}) \\ \delta s_\rho^{(n)}(\mathbf{r}) \end{bmatrix}. \tag{15}$$

2.3 Isotropic-elastic minimization

We finally consider a scalar functional of three functions $s_P(\mathbf{r})$, $s_S(\mathbf{r})$ and $s_\rho(\mathbf{r})$, corresponding for instance to P-wave velocity V_P, S-wave velocity V_S and density ρ:

$$\Phi_3(s_P(\mathbf{r}), s_S(\mathbf{r}), s_\rho(\mathbf{r})). \tag{16}$$

The associated update has the form

$$\begin{bmatrix} \delta s_P(\mathbf{r}) \\ \delta s_S(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} = \int d\mathbf{r}' \mathcal{H}_3^{-1}(\mathbf{r},\mathbf{r}') \int \mathrm{d}\mathbf{r}'' \begin{bmatrix} \mathcal{H}_{SS\rho\rho}^{\rho SS\rho} & \mathcal{H}_{\rho SP\rho}^{PS\rho\rho} & \mathcal{H}_{PSS\rho}^{SSP\rho} \\ \mathcal{H}_{\rho PS\rho}^{SP\rho\rho} & \mathcal{H}_{PP\rho\rho}^{\rho PP\rho} & \mathcal{H}_{SPP\rho}^{PPS\rho} \\ \mathcal{H}_{SP\rho S}^{\rho PSS} & \mathcal{H}_{\rho PPS}^{PP\rho S} & \mathcal{H}_{PPSS}^{PSSP} \end{bmatrix} \begin{bmatrix} g_P(\mathbf{r}'') \\ g_S(\mathbf{r}'') \\ g_\rho(\mathbf{r}'') \end{bmatrix}, \tag{17}$$

where

$$\mathcal{H}_{ijkl}^{pqrs}(\mathbf{r}',\mathbf{r}'') = -\int d\mathbf{r}''' \left[H_{ij}(\mathbf{r}',\mathbf{r}''')H_{kl}(\mathbf{r}''',\mathbf{r}'') - H_{pq}(\mathbf{r}',\mathbf{r}''')H_{rs}(\mathbf{r}''',\mathbf{r}'')\right], \tag{18}$$

and where the H_{ij} (the i, j taking on values of P, S and ρ as needed) are

$$H_{ij}(\mathbf{r},\mathbf{r}') = \frac{\partial^2 \Phi_3(s_P, s_S, s_\rho)}{\partial s_i(\mathbf{r})\partial s_j(\mathbf{r}')}. \tag{19}$$

The generalized determinant is

$$\mathcal{H}_3(\mathbf{r},\mathbf{r}') = \int \mathrm{d}\mathbf{r}'' \left[H_{PP}(\mathbf{r},\mathbf{r}'')\mathcal{H}_{SS\rho\rho}^{\rho SS\rho}(\mathbf{r}'',\mathbf{r}') + H_{PS}(\mathbf{r},\mathbf{r}'')\mathcal{H}_{S\rho\rho P}^{SP\rho\rho}(\mathbf{r}'',\mathbf{r}') + H_{P\rho}(\mathbf{r},\mathbf{r}'')\mathcal{H}_{SP\rho S}^{SS\rho P}(\mathbf{r}'',\mathbf{r}')\right], \tag{20}$$

and the inverse of $\mathcal{H}_3$, namely $\mathcal{H}_3^{-1}$, is defined such that

$$\int d\mathbf{r}'' \mathcal{H}_3^{-1}(\mathbf{r},\mathbf{r}'')\mathcal{H}_3(\mathbf{r}'',\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'). \tag{21}$$

With these quantities in hand at the nth step we may use eq. (17) to form the $n+1$th model iterate via

$$\begin{bmatrix} s_P^{(n+1)}(\mathbf{r}) \\ s_S^{(n+1)}(\mathbf{r}) \\ s_\rho^{(n+1)}(\mathbf{r}) \end{bmatrix} = \begin{bmatrix} s_P^{(n)}(\mathbf{r}) \\ s_S^{(n)}(\mathbf{r}) \\ s_\rho^{(n)}(\mathbf{r}) \end{bmatrix} + \begin{bmatrix} \delta s_P^{(n)}(\mathbf{r}) \\ \delta s_S^{(n)}(\mathbf{r}) \\ \delta s_\rho^{(n)}(\mathbf{r}) \end{bmatrix}. \tag{22}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

3 APPROXIMATE NEWTON UPDATES

Two different kinds of Hessian approximation natural to this mixed continuous-discrete formulation can be made. Both involve neglecting some off-diagonal parts of the Hessian, and retaining others. We will refer to them as the *parameter-type* and *space-type* approximations, to distinguish which parts of the Hessian are retained in each case.

3.1 Parameter-type Hessian approximation

The parameter-type approximation involves retaining as non-negligible the on- and off-diagonal block elements of the 2×2 and 3×3 systems in eqs (11) and (17), but neglecting the off-diagonal components internal to each of these operator elements.

3.1.1 Parameter-type approximation, scalar case

A scalar Hessian function $H_1(\mathbf{r}, \mathbf{r}')$ of the kind used in Section 2.1 contributes off-diagonally to the update when $\mathbf{r} \neq \mathbf{r}'$, and on-diagonally when $\mathbf{r} = \mathbf{r}'$. An approximate Hessian in which we have neglected off-diagonal contributions can therefore be written

$$H_1(\mathbf{r}, \mathbf{r}') \approx h(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}'). \tag{23}$$

When eq. (23) is substituted into eq. (5), we obtain the update formula

$$\delta s(\mathbf{r}) \approx -\frac{1}{h(\mathbf{r})} g(\mathbf{r}), \tag{24}$$

which can be used, given the nth gradient $g^{(n)}(\mathbf{r})$ and the Hessian function $h^{(n)}(\mathbf{r})$, to determine the update $\delta s^{(n)}(\mathbf{r})$. Because $-1/h(\mathbf{r})$ is a diagonal operator, in the scalar case the parameter-type approximation is equivalent to preconditioning the gradient with the diagonal elements of the inverse Hessian.

3.1.2 Parameter-type approximation, acoustic case

In eq. (23) the parts of the function $H(\mathbf{r}, \mathbf{r}')$ corresponding to the case $\mathbf{r} = \mathbf{r}'$ were retained, and all other parts were neglected. In the acoustic case the same approximation involves assuming the forms

$$H_{\kappa\kappa}(\mathbf{r}, \mathbf{r}') \approx h_{\kappa\kappa}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}'), \quad H_{\kappa\rho}(\mathbf{r}, \mathbf{r}') \approx h_{\kappa\rho}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}'), \quad H_{\rho\kappa}(\mathbf{r}, \mathbf{r}') \approx h_{\rho\kappa}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}'), \quad H_{\rho\rho}(\mathbf{r}, \mathbf{r}') \approx h_{\rho\rho}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}'). \tag{25}$$

Substituting eqs (25) into eq. (10), we firstly find that

$$\mathcal{H}_2(\mathbf{r}, \mathbf{r}') \approx \left[h_{\kappa\kappa}(\mathbf{r})h_{\rho\rho}(\mathbf{r}) - h_{\kappa\rho}(\mathbf{r})h_{\rho\kappa}(\mathbf{r})\right] \delta(\mathbf{r} - \mathbf{r}'), \tag{26}$$

which has the straightforward inverse

$$\mathcal{H}_2^{-1}(\mathbf{r}, \mathbf{r}') \approx \frac{\delta(\mathbf{r} - \mathbf{r}')}{h_{\kappa\kappa}(\mathbf{r})h_{\rho\rho}(\mathbf{r}) - h_{\kappa\rho}(\mathbf{r})h_{\rho\kappa}(\mathbf{r})}. \tag{27}$$

If eq. (25) holds, each of the four integral operator elements of the full Hessian in eq. (9) is diagonal. Because these operators commute, we may legitimately substitute eqs (25) and (27) into eq. (11), arriving at the parameter-type update formula:

$$\begin{bmatrix} \delta s_\kappa(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} \approx -\frac{1}{h_{\kappa\kappa}(\mathbf{r})h_{\rho\rho}(\mathbf{r}) - h_{\rho\kappa}(\mathbf{r})h_{\kappa\rho}(\mathbf{r})} \begin{bmatrix} h_{\rho\rho}(\mathbf{r}) & -h_{\kappa\rho}(\mathbf{r}) \\ -h_{\rho\kappa}(\mathbf{r}) & h_{\kappa\kappa}(\mathbf{r}) \end{bmatrix} \begin{bmatrix} g_\kappa(\mathbf{r}) \\ g_\rho(\mathbf{r}) \end{bmatrix}, \tag{28}$$

which can be used given the nth gradients and Hessian functions to determine the update $[\delta s_\kappa^{(n)}(\mathbf{r}), \delta s_\rho^{(n)}(\mathbf{r})]^T$. More clearly here than in the scalar problem, we observe that in the parameter-type approximation, inverse Hessian elements corresponding to coupling between the parameters, namely $h_{\kappa\rho}(\mathbf{r})$ and $h_{\rho\kappa}(\mathbf{r})$, are retained, whereas off-diagonal elements coupling model parameters at different points in space are neglected. The detailed forms taken by $h_{\kappa\kappa}(\mathbf{r})$, $h_{\kappa\rho}(\mathbf{r})$, $h_{\rho\kappa}(\mathbf{r})$ and $h_{\rho\rho}(\mathbf{r})$ are determined when a particular FWI problem is formulated (i.e. a data set, objective function, etc.). In Section 4, we will show this when we work through a particular example.

3.1.3 Parameter-type approximation, isotropic-elastic case

In the isotropic-elastic parameter-type approximate Hessian, starting from the formulas in Section 2.3, we make the approximations

$$H_{ij}(\mathbf{r}, \mathbf{r}') \approx h_{ij}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}'). \tag{29}$$

This simplifies the generalized determinant in eq. (20) to $\mathcal{H}_3(\mathbf{r}, \mathbf{r}') \approx h_3(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')$, where

$$h_3(\mathbf{r}) = h_{PP}(\mathbf{r})\left[h_{SS}(\mathbf{r})h_{\rho\rho}(\mathbf{r}) - h_{S\rho}(\mathbf{r})h_{\rho S}(\mathbf{r})\right] + h_{PS}(\mathbf{r})\left[h_{SP}(\mathbf{r})h_{\rho\rho}(\mathbf{r}) - h_{S\rho}(\mathbf{r})h_{\rho P}(\mathbf{r})\right] + h_{P\rho}(\mathbf{r})\left[h_{SP}(\mathbf{r})h_{\rho S}(\mathbf{r}) - h_{SS}(\mathbf{r})h_{\rho P}(\mathbf{r})\right]. \tag{30}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

We then have $\mathcal{H}_3^{-1}(\mathbf{r}, \mathbf{r}') \approx h_3^{-1}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')$, and when this is substituted into eq. (17) we arrive at the parameter-type approximate update formula:

$$\begin{bmatrix} \delta s_P(\mathbf{r}) \\ \delta s_S(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} \approx \frac{-1}{h_3(\mathbf{r})} \begin{bmatrix} h^{\rho SS\rho}_{SS\rho\rho}(\mathbf{r}) & h^{PS\rho\rho}_{\rho SP\rho}(\mathbf{r}) & h^{SSP\rho}_{PSS\rho}(\mathbf{r}) \\ h^{SP\rho\rho}_{\rho PS\rho}(\mathbf{r}) & h^{\rho PP\rho}_{PP\rho\rho}(\mathbf{r}) & h^{PPS\rho}_{SPP\rho}(\mathbf{r}) \\ h^{\rho PSS}_{SP\rho S}(\mathbf{r}) & h^{PP\rho S}_{\rho PPS}(\mathbf{r}) & h^{PSSP}_{PPSS}(\mathbf{r}) \end{bmatrix} \begin{bmatrix} g_P(\mathbf{r}) \\ g_S(\mathbf{r}) \\ g_\rho(\mathbf{r}) \end{bmatrix}, \tag{31}$$

where the elements of the generalized cofactor matrix are

$$h^{pqrs}_{ijkl}(\mathbf{r}) = h_{ij}(\mathbf{r})h_{kl}(\mathbf{r}) - h_{pq}(\mathbf{r})h_{rs}(\mathbf{r}),$$

with the indices referring as in Section 2.3 to combinations of the three elastic parameters.

3.2 Space-type Hessian approximation

The space-type Hessian approximation emphasizes correlations between model elements separated in space, while neglecting correlations between different parameters.

3.2.1 Space-type approximation, acoustic case

The acoustic space-type approximation is found by suppressing the off-diagonal elements of the 2×2 system in eq. (11), that is by assuming that the Hessian functions coupling κ and ρ are negligible. With $H_{\kappa\rho}$ and $H_{\rho\kappa}$ set to nil, the generalized determinant inverted in eq. (11) simplifies to

$$\mathcal{H}_2(\mathbf{r}, \mathbf{r}') \approx \int d\mathbf{r}'' H_{\kappa\kappa}(\mathbf{r}, \mathbf{r}'')H_{\rho\rho}(\mathbf{r}'', \mathbf{r}'). \tag{32}$$

If $H_{\rho\rho}^{-1}$ and $H_{\kappa\kappa}^{-1}$ are the inverses of $H_{\rho\rho}$ and $H_{\kappa\kappa}$ in the sense of eq. (6), it follows that $\mathcal{H}_2$ in eq. (32) has the inverse

$$\mathcal{H}_2^{-1}(\mathbf{r}, \mathbf{r}') \approx \int d\mathbf{r}'' H_{\rho\rho}^{-1}(\mathbf{r}, \mathbf{r}'')H_{\kappa\kappa}^{-1}(\mathbf{r}'', \mathbf{r}'). \tag{33}$$

In order for eq. (11) to be correctly used, the nonzero operator elements of the Hessian in eq. (9) must commute, which now appears as the requirement that

$$\int d\mathbf{r}'' H_{\rho\rho}^{-1}(\mathbf{r}, \mathbf{r}'')H_{\kappa\kappa}^{-1}(\mathbf{r}'', \mathbf{r}') = \int d\mathbf{r}'' H_{\kappa\kappa}^{-1}(\mathbf{r}, \mathbf{r}'')H_{\rho\rho}^{-1}(\mathbf{r}'', \mathbf{r}'). \tag{34}$$

After substitution of eq. (33) into eq. (11), along with the neglect of $H_{\kappa\rho}$ and $H_{\rho\kappa}$ and the use of relation (34), we obtain the space-type formula

$$\begin{bmatrix} \delta s_\kappa(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} \approx -\int d\mathbf{r}' \begin{bmatrix} H_{\kappa\kappa}^{-1}(\mathbf{r}, \mathbf{r}')g_\kappa(\mathbf{r}') \\ H_{\rho\rho}^{-1}(\mathbf{r}, \mathbf{r}')g_\rho(\mathbf{r}') \end{bmatrix}. \tag{35}$$

Here, the inverse Hessian retains the ability to influence the update at some particular point in space using values of the gradient at other points in space. However, it has lost the ability to affect the bulk modulus update with density gradient values, and vice versa. Eq. (35) is derivable directly from (9), without the assumption of commuting integral operators. So, eq. (34) can be viewed as a convenient restriction, which allows us to make consistent use of eq. (11), but which can be relaxed if needed.

3.2.2 Space-type approximation, isotropic elastic case

The elastic space-type approximate Hessian and the associated updates are, as in the acoustic case, brought about by neglecting all Hessian functions H_{ij} in the 3×3 matrix in the Newton step in eq. (17) which couple different parameters together. The cofactor matrix becomes

$$\begin{bmatrix} H_{SS}H_{\rho\rho} & 0 & 0 \\ 0 & H_{PP}H_{\rho\rho} & 0 \\ 0 & 0 & H_{PP}H_{SS} \end{bmatrix}, \tag{36}$$

and the generalized determinant reduces to

$$\mathcal{H}_3(\mathbf{r}, \mathbf{r}') \approx \int d\mathbf{r}'' \int d\mathbf{r}''' H_{PP}(\mathbf{r}, \mathbf{r}'')H_{SS}(\mathbf{r}'', \mathbf{r}''')H_{\rho\rho}(\mathbf{r}''', \mathbf{r}'). \tag{37}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

This leads through arguments identical to those in Section 3.2.1 to the three-parameter space-type update

$$\begin{bmatrix} \delta s_P(\mathbf{r}) \\ \delta s_S(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} \approx -\int d\mathbf{r}' \begin{bmatrix} H_{PP}^{-1}(\mathbf{r},\mathbf{r}')g_P(\mathbf{r}') \\ H_{SS}^{-1}(\mathbf{r},\mathbf{r}')g_S(\mathbf{r}') \\ H_{\rho\rho}^{-1}(\mathbf{r},\mathbf{r}')g_\rho(\mathbf{r}') \end{bmatrix}. \tag{38}$$

3.3 Preconditioning with diagonal elements of the inverse Hessian

If both types of off-diagonal inverse Hessian element are neglected simultaneously, which can be accomplished for instance by setting

$$H_{\kappa\kappa}^{-1}(\mathbf{r},\mathbf{r}') \approx h_{\kappa\kappa}^{-1}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')$$
$$H_{\rho\rho}^{-1}(\mathbf{r},\mathbf{r}') \approx h_{\rho\rho}^{-1}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') \tag{39}$$

in eq. (35), we obtain for the acoustic update

$$\begin{bmatrix} \delta s_\kappa(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} \approx -\begin{bmatrix} h_{\kappa\kappa}^{-1}(\mathbf{r})g_\kappa(\mathbf{r}) \\ h_{\rho\rho}^{-1}(\mathbf{r})g_\rho(\mathbf{r}) \end{bmatrix}. \tag{40}$$

A similar application of both approximations in the isotropic-elastic update in eq. (17) leads to an update of the form

$$\begin{bmatrix} \delta s_P(\mathbf{r}) \\ \delta s_S(\mathbf{r}) \\ \delta s_\rho(\mathbf{r}) \end{bmatrix} \approx -\begin{bmatrix} h_{PP}^{-1}(\mathbf{r})g_P(\mathbf{r}) \\ h_{SS}^{-1}(\mathbf{r})g_S(\mathbf{r}) \\ h_{\rho\rho}^{-1}(\mathbf{r})g_\rho(\mathbf{r}) \end{bmatrix}. \tag{41}$$

The inverse Hessians are evidently fully diagonal under these substitutions, indicating that a simultaneous application of the parameter-type and space-type approximations is equivalent to a preconditioning of the gradient with the diagonal elements of the inverse Hessian.

4 FULL WAVEFORM INVERSION AND AVO

The formulas in the previous sections are convenient for the quantitative interpretation of the machinery of a seismic FWI update. A characterization of how this machinery copes with multiparameter reflection issues, primarily cross-talk, and how it is related to AVO analysis and inversion, can be accomplished by working a specific example. We will consider the problem of reconstruction of a single acoustic boundary using the parameter-type formula in eq. (28).

4.1 Linearized acoustic AVO equations

We invoke the Aki-Richards approximation (Aki & Richards 2002) in the acoustic limit, expressed in terms of FWI update quantities. Let a plane wave impinge at angle θ on a planar interface separating a halfspace with density and bulk modulus values ρ_0 and κ_0 from a halfspace with values ρ_1 and κ_1 (Fig. 1). The reflection coefficient for this configuration is given by

$$R(\theta) = \frac{1-\Omega(\theta)}{1+\Omega(\theta)}, \tag{42}$$

where

$$\Omega(\theta) = \left(\frac{\rho_0}{\rho_1}\right)\sqrt{\frac{\kappa_0}{\kappa_1}\frac{\rho_1}{\rho_0}}\left(\frac{1}{\cos\theta}\right)\sqrt{1-\frac{\kappa_1}{\kappa_0}\frac{\rho_0}{\rho_1}\sin^2\theta}. \tag{43}$$

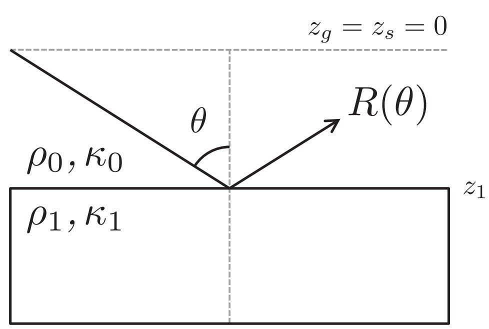

Figure 1. Geometry of a reflection from a single horizontal planar acoustic interface.

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

Defining the reciprocal modulus and density parameters in a manner consistent with Sections 2 and 3,

$$s_\kappa = \frac{1}{\kappa_1}, \quad s_{\kappa 0} = \frac{1}{\kappa_0}, \quad s_\rho = \frac{1}{\rho_1}, \quad s_{\rho 0} = \frac{1}{\rho_0}, \tag{44}$$

the reflection coefficient, to first order in the changes

$$\delta s_\kappa = s_\kappa - s_{\kappa 0}, \quad \delta s_\rho = s_\rho - s_{\rho 0}, \tag{45}$$

and $\sin^2\theta$, is

$$R(\theta) \approx -\frac{1}{4}\frac{1}{\cos^2\theta}\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right) - \frac{1}{4}\cos^2\theta\left(\frac{\delta s_\rho}{s_{\rho 0}}\right). \tag{46}$$

Hereafter the words *to first order* will be used in the sense above, that is meaning *true for small parameter contrasts and incidence angles*. This is standard in AVO analysis (Castagna & Backus 1993). The significance of these equations to our understanding of reflection FWI is as follows. If the initial FWI medium model is homogeneous, and the data are due to a reflection from a single acoustic interface at depth z_1, then the ideal update, that is that which takes us in a single step to the correct model, is

$$\delta s_\kappa(z) = \delta s_\kappa S(z - z_1), \qquad \delta s_\rho(z) = \delta s_\rho S(z - z_1), \tag{47}$$

where

$$S(z) = \begin{cases} 0, & z < 0 \\ 1, & z > 0 \end{cases} \tag{48}$$

is the Heaviside function, and where δs_κ and δs_ρ on the right-hand sides of eq. (47) satisfy eq. (46) to first order.

4.2 The (k_g, x_s, ω) domain

We next restrict the unknown medium to be variable in depth z only, and suppose it to be overlain by line sources and receivers, arranged along, and perpendicular to, a lateral axis x. It is convenient to begin by considering the data and modelled wavefields in the (k_g, x_s, ω) domain, where ω is the angular frequency, k_g is the Fourier conjugate to the lateral receiver coordinate x_g, and x_s is the lateral source coordinate. The wavefield P observed at depth z, due to a source at (x_s, z_s), in these coordinates will therefore be expressed as

$$P(k_g, z, x_s, z_s, \omega). \tag{49}$$

During the nth FWI iteration we assume access to the modelled field

$$G_n(k_g, z, x_s, z_s, \omega) \tag{50}$$

everywhere in the nth medium iterate.

4.3 Ingredients for a FWI update

4.3.1 *Acoustic wave equations*

We will assume that the data are full bandwidth, source-wavelet deconvolved measurements of an acoustic field P. Such a field satisfies

$$\left[\nabla \cdot \left(\frac{1}{\rho(\mathbf{r}_g)}\right)\nabla + \frac{\omega^2}{\kappa(\mathbf{r}_g)}\right] P(\mathbf{r}_g, \mathbf{r}_s, \omega) = \delta(\mathbf{r}_g - \mathbf{r}_s), \tag{51}$$

where $\mathbf{r}_g$ is the observation point, and $\mathbf{r}_s$ is the location of an idealized impulsive source. The derivatives ∇ are taken with respect to $\mathbf{r}_g$. In terms of the reciprocals of κ and ρ the equation is:

$$\left[\nabla \cdot s_\rho(\mathbf{r}_g)\nabla + \omega^2 s_\kappa(\mathbf{r}_g)\right] P(\mathbf{r}_g, \mathbf{r}_s, \omega) = \delta(\mathbf{r}_g - \mathbf{r}_s). \tag{52}$$

The FWI problem is to update the nth background model $\left(s_{\kappa 0}^{(n)}, s_{\rho 0}^{(n)}\right)$, that is to form $\left(\delta s_\kappa^{(n)}, \delta s_\rho^{(n)}\right)$ such that at the point $\mathbf{r}$ we may calculate

$$s_\rho^{(n+1)}(\mathbf{r}) = s_{\rho 0}^{(n)}(\mathbf{r}) + \delta s_\rho^{(n)}(\mathbf{r}),$$
$$s_\kappa^{(n+1)}(\mathbf{r}) = s_{\kappa 0}^{(n)}(\mathbf{r}) + \delta s_\kappa^{(n)}(\mathbf{r}). \tag{53}$$

The parameters $s_{\kappa 0}^{(n)}$ and $s_{\rho 0}^{(n)}$ are known, as are the solutions for the fields in the medium characterized by these parameters, namely G_n, where

$$\left[\nabla \cdot s_{\rho 0}^{(n)}(\mathbf{r}_g)\nabla + \omega^2 s_{\kappa 0}^{(n)}(\mathbf{r}_g)\right] G_n(\mathbf{r}_g, \mathbf{r}_s, \omega) = \delta(\mathbf{r}_g - \mathbf{r}_s). \tag{54}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

These solutions are the modelled fields as discussed in Section 4.2. Combining eqs (52)–(54), to first order in δs_κ and δs_ρ the difference $\delta G_n = P - G_n$ can be written

$$\delta G_n(\mathbf{r}_g, \mathbf{r}_s, \omega) \approx -\int \mathrm{d}\mathbf{r}' G_n(\mathbf{r}_g, \mathbf{r}', \omega)\left[\nabla \cdot \delta s_\rho^{(n)}(\mathbf{r}')\nabla + \omega^2 \delta s_\kappa^{(n)}(\mathbf{r}')\right] G_n(\mathbf{r}', \mathbf{r}_s, \omega), \tag{55}$$

where the derivatives ∇ are taken with respect to the integration variable. Eq. (55) allows us to derive acoustic sensitivities in any convenient domain. Expressing the position vectors in a 2D Cartesian system $\mathbf{r} = (x, z)$, and taking the Fourier transform with respect to the lateral observation variable x_g, we generate two special cases of eq. (55). First, we set $\delta s_\kappa^{(n)}(x', z') = \delta s_\kappa^{(n)}(z)\delta(z - z')$ and $\delta s_\rho^{(n)}(x', z') = 0$, and compute

$$\frac{\partial G_n(k_g, z_g, x_s, z_s, \omega)}{\partial s_\kappa^{(n)}(z)} = \lim_{\delta s_\kappa^{(n)} \to 0} \frac{\delta G_n(k_g, z_g, x_s, z_s, \omega)}{\delta s_\kappa^{(n)}(z)}, \tag{56}$$

from which we find the sensitivity for the depth-dependent bulk modulus parameter:

$$\frac{\partial G_n(k_g, z_g, x_s, z_s, \omega)}{\partial s_\kappa^{(n)}(z)} = -\omega^2 \int dx' G_n(k_g, z_g, x', z, \omega) G_n(x', z, x_s, z_s, \omega). \tag{57}$$

Secondly, we set $\delta s_\rho^{(n)}(x', z') = \delta s_\rho^{(n)}(z)\delta(z - z')$ and $\delta s_\kappa^{(n)}(x', z') = 0$, and upon forming the appropriate limit compute the sensitivity for a depth-dependent density parameter:

$$\frac{\partial G_n(k_g, z_g, x_s, z_s, \omega)}{\partial s_\rho^{(n)}(z)} = \int \mathrm{d}x' \left[\frac{\partial G_n(k_g, z_g, x', z, \omega)}{\partial z}\frac{\partial G_n(x', z, x_s, z_s, \omega)}{\partial z} - G_n(k_g, z_g, x', z, \omega)\frac{\partial^2 G_n(x', z, x_s, z_s, \omega)}{\partial x'^2}\right]. \tag{58}$$

4.3.2 *Green's functions*

We now consider the first ($n = 0$) iteration of the FWI problem, assuming a homogeneous acoustic medium as the background. Solving eq. (54) with spatially constant κ_0 and ρ_0 (or equivalently s_{κ_0} and s_{ρ_0}), we obtain two useful forms (Clayton & Stolt 1981)

$$G_0(k_g, z, x_s, z_s, \omega) = \rho_0 e^{-ik_g x_s}\frac{e^{iq_g|z - z_s|}}{i2q_g}, \tag{59}$$

and

$$G_0(x_g, z, x_s, z_s, \omega) = \frac{\rho_0}{2\pi}\int \mathrm{d}k' e^{ik'(x_g - x_s)}\frac{e^{iq'|z - z_s|}}{i2q'}, \tag{60}$$

where

$$q_g = \frac{\omega}{c_0}\sqrt{1 - \frac{k_g^2 c_0^2}{\omega^2}}, \quad q' = \frac{\omega}{c_0}\sqrt{1 - \frac{k'^2 c_0^2}{\omega^2}}. \tag{61}$$

Here, $c_0 = \sqrt{\kappa_0/\rho_0} = \sqrt{s_{\rho_0}/s_{\kappa_0}}$ is the constant background acoustic wave velocity.

4.3.3 *Analytic data and residuals*

The data set, $D = P(k_g, z_g, x_s, z_s, \omega)$, is the wavefield evaluated on the surface $z = z_g$. Let us now assume these D correspond to a reflection from a single horizontal interface separating two infinite half-spaces, due to a single source at the origin, $z_s = x_s = 0$. Further, let that origin lie on the measurement surface such that $z_g = 0$. The data measured above the interface at z_1 can be expressed analytically as

$$D(k_g, \omega) = P(k_g, 0, 0, 0, \omega) = \rho_0 \frac{1}{i2q_g} + \rho_0 R(\theta)\frac{e^{i2q_g z_1}}{i2q_g}, \tag{62}$$

where the first term is the component of the wavefield propagating directly between the (coincident) source and receiver depths, and the second term is the reflection coming from the interface at depth z_1. The coefficient $R(\theta)$ was given exactly in eq. (42) and approximately in eq. (43). In Fig. 2 the geometric configuration through which θ is related to the Fourier quantities k_g, q_g and ω is illustrated. This geometry provides the following useful relations:

$$\tan\theta = \frac{k_g}{q_g}, \quad \cos\theta = \frac{q_g}{\omega\sqrt{\rho_0/\kappa_0}} = \frac{q_g c_0}{\omega}, \tag{63}$$

and for a differential element of frequency

$$\mathrm{d}\omega = \mathrm{d}(2q_g)\left(\frac{c_0}{2\cos\theta}\right). \tag{64}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

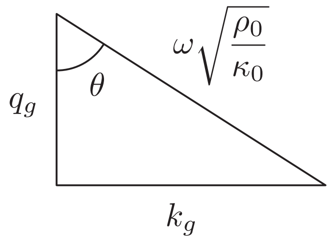

Figure 2. Geometry of harmonic plane wavenumbers, frequencies and angles.

The complex conjugates of the residuals $\delta P_n = P - G_n$ are required to calculate an FWI update. In the first iteration, the modelled field is the same as the first term in eq. (62), that is the full solution in a homogeneous medium:

$$G_0(k_g, 0, 0, 0, \omega) = \rho_0 \frac{1}{i2q_g}. \tag{65}$$

The complex conjugate of the difference between eqs (65) and (62) is, then,

$$\delta P_0^*(k_g, \omega) = -\rho_0 R(\theta) \frac{e^{-i2q_g z_1}}{i2q_g}. \tag{66}$$

4.3.4 Objective function

In geophysical inversion problems, objective functions are designed to penalize data misfit and steer away from undesirable models (Oldenburg 1984; Parker 1994). Here to stay consistent with standard FWI theory we will employ a simple objective function based on data misfit, but we note parenthetically that more sophisticated examples for seismic reflection data do exist (Symes & Carazzone 1991; Guitton & Symes 2003). An objective function appropriate for the depth-varying, two-parameter acoustic FWI problem is Φ_2 where

$$\Phi_2(s_\kappa, s_\rho) = \frac{1}{2} \sum_{k_g} \int d\omega |\delta P|^2, \tag{67}$$

which measures the square of the residuals, summed over the experimental variables k_g and ω, with the sum being continuous over ω and discrete over k_g.

4.4 Acoustic gradients for depth-varying models

The unknown Earth properties s_κ and s_ρ, and therefore the gradient and Hessian functions also, have been restricted to be functions of z only. Gradients as defined in, for instance, eq. (13), evaluated using the objective function in eq. (67), are of the form

$$g_\kappa^{(n)}(z) = -\sum_{k_g} \int d\omega \frac{\partial G_n(k_g, z_g, x_s, z_s, \omega)}{\partial s_\kappa^{(n)}(z)} \delta P_n^*, \tag{68}$$

and

$$g_\rho^{(n)}(z) = -\sum_{k_g} \int d\omega \frac{\partial G_n(k_g, z_g, x_s, z_s, \omega)}{\partial s_\rho^{(n)}(z)} \delta P_n^*. \tag{69}$$

Using the sensitivities in eqs (57) and (58), the gradient for the bulk modulus parameter in eq. (68) becomes

$$g_\kappa^{(n)}(z) = \sum_{k_g} \int d\omega \omega^2 \int dx' G_n(k_g, z_g, x', z, \omega) G_n(x', z, x_s, z_s, \omega) \delta P_n^*, \tag{70}$$

and the gradient for the density parameter in eq. (69) becomes

$$g_\rho^{(n)}(z) = -\sum_{k_g} \int d\omega \int dx' \left[\frac{\partial G_n(k_g, z_g, x', z, \omega)}{\partial z} \frac{\partial G_n(x', z, x_s, z_s, \omega)}{\partial z} - G_n(k_g, z_g, x', z, \omega) \frac{\partial^2 G_n(x', z, x_s, z_s, \omega)}{\partial x'^2} \right] \delta P_n^*. \tag{71}$$

Substituting eqs (59) and (60) into these formulas, and setting $z_g = z_s = x_s = 0$, we obtain for the first FWI update (i.e. for the $n = 0$ case) gradients of the form

$$g_\kappa^{(0)}(z) = -\frac{\rho_0^2}{4} \sum_{k_g} \int d\omega \; e^{i2q_g z} \left(\frac{\omega^2}{q_g^2} \right) \delta P_0^*, \tag{72}$$

and

$$g_\rho^{(0)}(z) = -\frac{\rho_0^2}{4} \sum_{k_g} \int d\omega \; e^{i2q_g z} \left(1 - \frac{k_g^2}{q_g^2} \right) \delta P_0^*. \tag{73}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

4.4.1 Transformation from k_g to angle

It is now convenient to transform the sum over k_g to a sum over incidence angle, using the relations in eq. (63). Eqs (72)–(73) become

$$g_\kappa^{(0)}(z) = -\frac{\rho_0^2 c_0^2}{4} \sum_\theta \int \mathrm{d}\omega\ e^{i2q_g z} \left(\frac{1}{\cos^2\theta}\right) \delta P_0^*, \tag{74}$$

and

$$g_\rho^{(0)}(z) = -\frac{\rho_0^2}{4} \sum_\theta \int \mathrm{d}\omega\ e^{i2q_g z} \left(1 - \tan^2\theta\right) \delta P_0^*. \tag{75}$$

4.5 Criticism of gradient-based updates

With eqs (74) and (75) in hand, the first ($n = 0$) gradient-based FWI update in the reconstruction of the single acoustic boundary illustrated in Fig. 1 can be analysed. Let us first suppose we are given only one angle of data, θ_1. By substituting the analytic residuals in eq. (66) into the formula in eq. (74), evaluated using θ_1, we obtain

$$\begin{aligned} g_\kappa^{(0)}(z|\theta_1) &= \frac{\rho_0^2 c_0^2}{4} \int \mathrm{d}\omega\ e^{i2q_g z} \left(\rho_0 \frac{R(\theta_1)}{\cos^2\theta_1} \frac{e^{-i2q_g z_1}}{i2q_g}\right) \\ &= \frac{R(\theta_1)\rho_0^3 c_0^2}{4\cos^2\theta_1} \int \mathrm{d}\omega \left(\frac{e^{i2q_g(z-z_1)}}{i2q_g}\right) \\ &= \frac{R(\theta_1)\rho_0^3 c_0^3}{8\cos^3\theta_1} \int \mathrm{d}(2q_g) \left(\frac{e^{i2q_g(z-z_1)}}{i2q_g}\right) \\ &= R(\theta_1)\frac{\pi\rho_0^3 c_0^3}{4}\left(\frac{1}{\cos^3\theta_1}\right) S(z - z_1), \end{aligned} \tag{76}$$

where in the second last step we made use of eq. (64), and in the last step we recognized the integral as the inverse Fourier transform of the spectrum of the step function defined in eq. (48). Similarly the density gradient is found to be

$$\begin{aligned} g_\rho^{(0)}(z|\theta_1) &= \frac{R(\theta_1)\pi c_0 \rho_0^3 (1 - \tan^2\theta_1)}{4\cos\theta_1} S(z - z_1) \\ &\approx R(\theta_1)\frac{\pi c_0 \rho_0^3}{4}(\cos\theta_1) S(z - z_1), \end{aligned} \tag{77}$$

the last approximation being appropriate for small (i.e. precritical) angles in which $\tan^2\theta \approx \sin^2\theta$, a condition already in place in our AVO approximations in Section 4.1.

It is to be expected that a two-parameter inverse problem cannot be solved with a single angle of data. In the progression through eqs (76) and (77) we have not as yet encountered any problem in attempting to do so (a fact we will later identify as a serious flaw in gradient-based multiparameter updates), but to accommodate the possibility that N angles of data are required, we further evaluate the formulas in eqs (74) and (75) for the set Θ where

$$\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}, \tag{78}$$

as follows (dropping the $\approx$ sign):

$$\begin{aligned} g_\kappa^{(0)}(z|\Theta) &= \frac{c_0^3 \rho_0^3 \pi}{4} \left[\sum_j^N \frac{R(\theta_j)}{\cos^3\theta_j}\right] S(z - z_1) \\ g_\rho^{(0)}(z|\Theta) &= \frac{c_0 \rho_0^3 \pi}{4} \left[\sum_j^N R(\theta_j)\cos\theta_j\right] S(z - z_1). \end{aligned} \tag{79}$$

We will now use the linearized acoustic AVO relation (eq. 46) as a means to study the efficacy of gradient-based FWI updates. Setting $N = 2$ in eq. (79), and substituting eq. (46) for each instance of the reflection coefficient, $R(\theta_1)$ and $R(\theta_2)$, we obtain to first order

$$\begin{aligned} g_\kappa^{(0)}(z|\theta_1, \theta_2) \approx &-\frac{c_0^3 \rho_0^3 \pi}{16}\left[\frac{1}{\cos^5\theta_1} + \frac{1}{\cos^5\theta_2}\right]\left(\frac{\delta s_\kappa}{s_{\kappa_0}}\right) S(z - z_1) \\ &-\frac{c_0^3 \rho_0^3 \pi}{16}\left[\frac{1}{\cos\theta_1} + \frac{1}{\cos\theta_2}\right]\left(\frac{\delta s_\rho}{s_{\rho_0}}\right) S(z - z_1), \end{aligned} \tag{80}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

and

$$g_\rho^{(0)}(z|\theta_1,\theta_2) \approx -\frac{c_0\rho_0^3\pi}{16}\left[\cos^3\theta_1+\cos^3\theta_2\right]\left(\frac{\delta s_\rho}{s_{\rho 0}}\right)S(z-z_1)$$
$$-\frac{c_0\rho_0^3\pi}{16}\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left(\frac{\delta s_\kappa}{s_{\kappa 0}}\right)S(z-z_1). \quad (81)$$

In a gradient-based approach, a line search (Pratt 1999) would normally at this point be carried out to determine a factor μ and thereby the updates $\mu g_\kappa^{(0)}(z|\theta_1,\theta_2)$ and $\mu g_\rho^{(0)}(z|\theta_1,\theta_2)$. The scalar μ can, in principle, correct the discrepancy between one (though not both) of the gradients in eqs (80) and (81) and the corresponding ideal update taking the form of a multiplicative factor. Therefore, in criticizing these gradients in their capacity as updates, we must allow that any such discrepancy has been corrected.

We judge efficacy by comparing eqs (80) and (81) to the benchmark updates in eq. (47). The comparison leads to two main criticisms of the gradient-based update. For the first (and less serious) of the two, let us suppose that the second line in eq. (80), and the first line in eq. (81), were both discovered to be negligible relative to their counterparts. There would then be two positive points to make about the gradients as updates. First, the space dependences of both updates would be found to be correct, being zero above the depth z_1 and finite below on account of the function S. This fact—while positive—is not all that remarkable, as we chose to study an example in which the initial medium and the actual medium agree for depths $z < z_1$. And second, below z_1 the bulk modulus update would be found to be proportional to the ideal update magnitude, δs_κ, and likewise the density update would be found to be proportional to the ideal magnitude δs_ρ. However, there remain prefactors which directly depend on the two angles used, θ_1 and θ_2. The gradients, that is, differ depending on which set of data angles are used.

For the second, and more serious, criticism, we drop the temporary assumption that any of the terms in eqs (80) and (81) are negligible, and recognize that the gradients are each proportional to both ideal updates, δs_κ and δs_ρ. The gradient-based update in any one parameter is, in general, a linear mixture of the ideal updates for that parameter and all others. No multiplicative factor exists which sets the second term in eq. (80) to zero but not the first, so a line search cannot suppress this source of error. Consequently, we conclude that a gradient-based update will see density leak uncontrollably into the bulk modulus estimate, and vice versa. This is how cross-talk appears in our mode of analysis.

Notice, however, that having come so far in our example, we can contrive a solution for this instance of cross-talk. Because (80) and (81) are simultaneous equations for δs_κ and δs_ρ in terms of known and/or determined quantities, we may solve the system for the ideal updates. Were we to do this, the result would be expressions for δs_κ and δs_ρ comprising linear mixtures of the two gradients. This solution (though it is of limited practical interest because of how specific a problem it solves) indicates that gradients can potentially be properly mixed to produce well-formed multiparameter updates. Rather than take this route, we will instead show how and where in the inverse Hessian a similar procedure is automatically carried out.

4.6 Acoustic Hessian functions

We will next quantify the extent to which the above criticisms of gradient-based updates are addressed by the components of the inverse Hessian retained in the parameter-type approximation. We start by computing the four two-parameter Hessian functions which appear in the acoustic Newton update formula (eq. 11). We begin with the function $H_{\kappa\kappa}^{(n)}(z,z')$ which at iteration n is

$$H_{\kappa\kappa}^{(n)}(z,z') = \frac{\partial^2\Phi_2\left(s_\kappa^{(n)},s_\rho^{(n)}\right)}{\partial s_\kappa^{(n)}(z)\partial s_\kappa^{(n)}(z')} = \frac{\partial}{\partial s_\kappa^{(n)}(z)}g_\kappa^{(n)}(z'). \quad (82)$$

Using the form for the gradient determined in the previous section, and the Green's functions in eqs (59)–(61), we find after some manipulation that, for $n = 0$,

$$H_{\kappa\kappa}^{(0)}(z,z') = \sum_{k_g}\int \mathrm{d}\omega\frac{\omega^4\rho_0^4}{(i2q_g)^4}e^{i2q_g(z'-z)} + \text{terms in } \delta P_0^*, \quad (83)$$

where we have distinguished parts of $H_{\kappa\kappa}^{(n)}(z,z')$ which depend on the residuals from parts which do not. We will presently find that the former do not contribute to the parameter-type update. Similarly,

$$H_{\kappa\rho}^{(n)}(z,z') = \frac{\partial^2\Phi_2\left(s_\kappa^{(n)},s_\rho^{(n)}\right)}{\partial s_\kappa^{(n)}(z)\partial s_\rho^{(n)}(z')} = \frac{\partial}{\partial s_\kappa^{(n)}(z)}g_\rho^{(n)}(z'), \quad (84)$$

which for the $n = 0$ case leads to

$$H_{\kappa\rho}^{(0)}(z,z') = \sum_{k_g}\int \mathrm{d}\omega\frac{\omega^2\rho_0^4}{(i2q_g)^4}\left(q_g^2-k_g^2\right)e^{i2q_g(z'-z)} + \text{terms in } \delta P_0^*, \quad (85)$$

and likewise, from

$$H_{\rho\kappa}^{(n)}(z,z') = \frac{\partial^2\Phi_2\left(s_\kappa^{(n)},s_\rho^{(n)}\right)}{\partial s_\rho^{(n)}(z)\partial s_\kappa^{(n)}(z')} = \frac{\partial}{\partial s_\rho^{(n)}(z)}g_\kappa^{(n)}(z'), \quad (86)$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

at $n = 0$ we obtain

$$H^{(n)}_{\rho\kappa}(z, z') = \sum_{k_g} \int d\omega \frac{\omega^2 \rho_0^4}{(i2q_g)^4} \left(q_g^2 - k_g^2\right) e^{i2q_g(z'-z)} + \text{terms in}\, \delta P_0^*, \tag{87}$$

which confirms the expectation that the symmetry $H^{(0)}_{\kappa\rho} = H^{(0)}_{\rho\kappa}$ exists in the 2×2 system in eq. (28), at least in this specific form. Finally, from

$$H^{(n)}_{\rho\rho}(z, z') = \frac{\partial^2 \Phi_2\left(s^{(n)}_\kappa, s^{(n)}_\rho\right)}{\partial s^{(n)}_\rho(z) \partial s^{(n)}_\rho(z')} = \frac{\partial}{\partial s^{(n)}_\rho(z)} g^{(n)}_\rho(z'), \tag{88}$$

we obtain at $n = 0$

$$H^{(0)}_{\rho\rho}(z, z') = \sum_{k_g} \int d\omega \frac{\rho_0^4}{(i2q_g)^4} (q_g^2 - k_g^2)^2 e^{i2q_g(z'-z)} + \text{terms in}\, \delta P_0^*. \tag{89}$$

4.6.1 *Transformation from k_g to angle*

We again use the relations in eqs (63) and (64), to re-write the Hessian functions in eqs (82)–(89) as

$$H^{(0)}_{\kappa\kappa}(z, z') = \frac{1}{16} \sum_\theta \frac{\rho_0^4 c_0^4}{\cos^4\theta} \int d\omega e^{i2q_g(z'-z)} + \text{terms in}\, \delta P_0^*, \tag{90}$$

as well as

$$H^{(0)}_{\kappa\rho}(z, z') \approx \frac{1}{16} \sum_\theta \rho_0^4 c_0^2 \int d\omega e^{i2q_g(z'-z)} + \text{terms in}\, \delta P_0^*, \tag{91}$$

and

$$H^{(0)}_{\rho\rho}(z, z') \approx \frac{1}{16} \sum_\theta \rho_0^4 \cos^4\theta \int d\omega e^{i2q_g(z'-z)} + \text{terms in}\, \delta P_0^*, \tag{92}$$

which, along with the symmetry $H^{(0)}_{\kappa\rho}(z, z') = H^{(0)}_{\rho\kappa}(z, z')$, provides all four components. The $\approx$ symbol has been again invoked wherever the approximation $\tan^2\theta \approx \sin^2\theta$ is used. (We have already confined ourselves to regimes of small angle in Sections 4.1 and 4.5.)

4.7 Calculation of the parameter-type update

As established in Section 3.1, we seek portions of the Hessian functions which can be written

$$\begin{aligned}
H^{(0)}_{\kappa\kappa}(z, z') &\approx h^{(0)}_{\kappa\kappa}(z)\delta(z - z'),\\
H^{(0)}_{\kappa\rho}(z, z') &\approx h^{(0)}_{\kappa\rho}(z)\delta(z - z'),\\
H^{(0)}_{\rho\kappa}(z, z') &\approx h^{(0)}_{\rho\kappa}(z)\delta(z - z'),\\
H^{(0)}_{\rho\rho}(z, z') &\approx h^{(0)}_{\rho\rho}(z)\delta(z - z'),
\end{aligned} \tag{93}$$

that is, which are diagonal in their depth-dependence. If we can find parts of the Hessian functions in eqs (90)–(92) whose space dependences are delta functions, the coefficients $h^{(0)}_{\kappa\kappa}$, $h^{(0)}_{\kappa\rho}$, $h^{(0)}_{\rho\kappa}$ and $h^{(0)}_{\rho\rho}$ can then be directly read from the results. It is not difficult to show that the residual-independent parts of the Hessian functions fit this bill. For instance, in eq. (90), we find by neglecting the term proportional to δP^*,

$$\begin{aligned}
H^{(0)}_{\kappa\kappa}(z, z') &\approx \sum_\theta \frac{c_0^4 \rho_0^4}{16\cos^4\theta} \int d\omega e^{i2q_g(z'-z)}\\
&\approx \sum_\theta \frac{c_0^5 \rho_0^4}{32\cos^5\theta} \int d(2q_g) e^{i2q_g(z'-z)}\\
&\approx \sum_\theta \frac{c_0^5 \rho_0^4 \pi}{16\cos^5\theta} \delta(z - z').
\end{aligned} \tag{94}$$

The residual-dependent parts have no comparable terms. The coefficient $h^{(0)}_{\kappa\kappa}(z)$ is, therefore, in the single interface case, the constant $h^{(0)}_{\kappa\kappa}$ where

$$h^{(0)}_{\kappa\kappa} = \sum_\theta \frac{c_0^5 \rho_0^4 \pi}{16\cos^5\theta}. \tag{95}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

Likewise we may find

$$h_{\kappa\rho}^{(0)} = h_{\rho\kappa}^{(0)} = \sum_{\theta} \frac{c_0^3 \rho_0^4 \pi}{16 \cos\theta}, \tag{96}$$

and

$$h_{\rho\rho}^{(0)} = \sum_{\theta} \frac{c_0 \rho_0^4 \pi}{16} \cos^3\theta. \tag{97}$$

With eqs (80)–(81) and (95)–(97), we now have all the necessary ingredients to calculate the first parameter-type update, as it would arise for data from a single interface:

$$\begin{bmatrix} \delta s_{\kappa}^{(0)}(z) \\ \delta s_{\rho}^{(0)}(z) \end{bmatrix} \approx -\frac{1}{h_{\kappa\kappa}^{(0)} h_{\rho\rho}^{(0)} - h_{\rho\kappa}^{(0)} h_{\kappa\rho}^{(0)}} \begin{bmatrix} h_{\rho\rho}^{(0)} & -h_{\kappa\rho}^{(0)} \\ -h_{\rho\kappa}^{(0)} & h_{\kappa\kappa}^{(0)} \end{bmatrix} \begin{bmatrix} g_{\kappa}^{(0)}(z) \\ g_{\rho}^{(0)}(z) \end{bmatrix}. \tag{98}$$

4.8 Analysis of the parameter-type update

We next fill in the four-point list of the traits of the parameter-type update given the introduction, with the exception of item one, concerning the computational cost of the update, which is established by its general form.

4.8.1 A check for ill-posedness

Suppose we were attempting to solve for updates in both modulus and density using a single angle of data, a problem which in standard AVO practice is well known to be underdetermined. We have already seen that the gradients are as easily constructed using one angle as they are from two or more. A gradient-based step is, evidently, insensitive to the ill-posedness of the one-angle/two-parameter problem. The same is not true of a parameter-type update. As an example, let us attempt to solve for the density update with the parameter-type formula in eq. (98) using one angle of data. The coefficients $h_{\kappa\kappa}^{(0)}$, $h_{\kappa\rho}^{(0)}$, $h_{\rho\kappa}^{(0)}$ and $h_{\rho\rho}^{(0)}$, calculated using eqs (95)–(97) with their sums limited to one angle, θ_1, form the determinant

$$h_{\kappa\kappa}^{(0)} h_{\rho\rho}^{(0)} - h_{\kappa\rho}^{(0)} h_{\rho\kappa}^{(0)} = \left(\frac{c_0^5 \rho_0^4 \pi}{16 \cos^5\theta_1} \right) \left(\frac{c_0 \rho_0^4 \pi \cos^3\theta_1}{16} \right) - \left(\frac{c_0^3 \rho_0^4 \pi}{16 \cos\theta_1} \right)^2 = 0. \tag{99}$$

The update

$$\delta s_{\rho}^{(0)}(z) \approx -\frac{h_{\kappa\kappa}^{(0)} g_{\rho}^{(0)}(z) - h_{\rho\kappa}^{(0)} g_{\kappa}^{(0)}(z)}{h_{\kappa\kappa}^{(0)} h_{\rho\rho}^{(0)} - h_{\kappa\rho}^{(0)} h_{\rho\kappa}^{(0)}}, \tag{100}$$

is, then, immediately halted by an undefined step length. This stopping condition is produced by the interaction of on- and off-diagonal elements of the inverse Hessian (e.g. $h_{\kappa\kappa}^{(0)}$ and $h_{\kappa\rho}^{(0)}$, respectively), and so is not found in either gradient-based or Gauss-Newton updates. Thus, whereas gradient-based and Gauss-Newton methods are in danger of proceeding with or without adequate data, in a parameter-type update, internal rank calculations on small (e.g. 2×2 or 3×3) matrices derived from the Hessian quickly diagnose ill-posedness. This is item two on the list.

4.8.2 Cross-talk and angle dependence

We will next establish that, to leading order, the portions of the inverse Hessian retained in a parameter-type approximation suppress cross-talk while simultaneously producing an angle-balanced update. Consider the bulk modulus update at the first iteration, $\delta s_{\kappa}^{(0)}$, which is formed through the calculation of the first element of eq. (98):

$$\delta s_{\kappa}^{(0)}(z) = -\frac{h_{\rho\rho}^{(0)} g_{\kappa}^{(0)}(z) - h_{\kappa\rho}^{(0)} g_{\rho}^{(0)}(z)}{h_{\kappa\kappa}^{(0)} h_{\rho\rho}^{(0)} - h_{\kappa\rho}^{(0)} h_{\rho\kappa}^{(0)}}. \tag{101}$$

The simplest well-posed inversion involves two angles of data. Evaluating eqs (95)–(97) with sums over angles θ_1 and θ_2, and using the gradient expressions in eqs (80)–(81), which have been expressed in terms of the ideal updates δs_κ and δs_ρ, we assemble the components of eq. (101) one at a time. The two numerator quantities are

$$\begin{aligned} h_{\rho\rho}^{(0)} g_{\kappa}^{(0)}(z|\theta_1, \theta_2) = &-\frac{c_0^4 \rho_0^7 \pi^2}{256} \left[\cos^3\theta_1 + \cos^3\theta_2 \right] \left[\frac{1}{\cos^5\theta_1} + \frac{1}{\cos^5\theta_2} \right] \left(\frac{\delta s_\kappa}{s_{\kappa 0}} \right) S(z - z_1) \\ &-\frac{c_0^4 \rho_0^7 \pi^2}{256} \left[\cos^3\theta_1 + \cos^3\theta_2 \right] \left[\frac{1}{\cos\theta_1} + \frac{1}{\cos\theta_2} \right] \left(\frac{\delta s_\rho}{s_{\rho 0}} \right) S(z - z_1), \end{aligned} \tag{102}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

and

$$h^{(0)}_{\kappa\rho}g^{(0)}_{\rho}(z|\theta_1,\theta_2) = -\frac{c_0^4\rho_0^7\pi^2}{256}\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left[\cos^3\theta_1+\cos^3\theta_2\right]\left(\frac{\delta s_\rho}{s_{\rho_0}}\right)S(z-z_1)$$
$$-\frac{c_0^4\rho_0^7\pi^2}{256}\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left[\frac{1}{\cos\theta_1}+\frac{1}{\cos\theta_2}\right]\left(\frac{\delta s_\kappa}{s_{\kappa_0}}\right)S(z-z_1). \quad (103)$$

In the numerator these two quantities are subtracted:

$$h^{(0)}_{\rho\rho}g^{(0)}_{\kappa}(z|\theta_1,\theta_2) - h^{(0)}_{\kappa\rho}g^{(0)}_{\rho}(z|\theta_1,\theta_2) = -\left(\frac{c_0^4\rho_0^7\pi^2}{256}\right)\left[\frac{\cos^3\theta_1}{\cos^5\theta_2}-\frac{2}{\cos\theta_1\cos\theta_2}+\frac{\cos^3\theta_2}{\cos^5\theta_1}\right]\left(\frac{\delta s_\kappa}{s_{\kappa_0}}\right)S(z-z_1). \quad (104)$$

Meanwhile the determinant in the reciprocal evaluates to

$$\frac{1}{h^{(0)}_{\kappa\kappa}h^{(0)}_{\rho\rho}-h^{(0)}_{\kappa\rho}h^{(0)}_{\rho\kappa}} = \left(\frac{256}{c_0^6\rho_0^8\pi^2}\right)\left[\frac{\cos^3\theta_1}{\cos^5\theta_2}-\frac{2}{\cos\theta_1\cos\theta_2}+\frac{\cos^3\theta_2}{\cos^5\theta_1}\right]^{-1}. \quad (105)$$

Recalling that $s_{\kappa_0} = (\rho_0 c_0^2)^{-1}$, the negative of the product of eqs (104) and (105) is

$$\delta s^{(0)}_\kappa(z) = -\frac{h^{(0)}_{\rho\rho}g^{(0)}_\kappa(z) - h^{(0)}_{\kappa\rho}g^{(0)}_\rho(z)}{h^{(0)}_{\kappa\kappa}h^{(0)}_{\rho\rho}-h^{(0)}_{\kappa\rho}h^{(0)}_{\rho\kappa}} = s_{\kappa_0}\left(\frac{\delta s_\kappa}{s_{\kappa_0}}\right)S(z-z_1) = \delta s_\kappa S(z-z_1), \quad (106)$$

which is the parameter-type result, accurate to leading order in δs_κ, δs_ρ and $\sin^2\theta$. An identical calculation of the parameter-type update for the density leads to

$$\delta s^{(0)}_\rho(z) = \frac{1}{\rho_0}\left(\frac{\delta s_\rho}{s_{\rho_0}}\right)S(z-z_1) = \delta s_\rho S(z-z_1), \quad (107)$$

recollecting that $s_{\rho_0} = 1/\rho_0$.

There are two aspects of this calculation to draw attention to. First, the lack of cross-talk. In our mode of analysis, if a proposed update in any one parameter, for example $\delta s^{(0)}_\kappa(z)$, can be shown to be proportional to the benchmark updates in any *other* parameter (e.g. $\delta s_\rho(z)$ in eq. 47), it is said to suffer from cross-talk. The κ gradient, $g_\kappa(z|\theta_1, \theta_2)$, being evaluated for its ability to act, in isolation, as a bulk modulus update, was seen in eq. (80) to be strongly affected in this sense. In contrast, the gradients in the parameter-type update do not act in isolation, but are weighted and subtracted from each other. The weighted subtraction, seen in eq. (104), produces a cancellation between terms proportional to δs_ρ, similar to our contrived solution to gradient-based cross-talk in Section 4.5. The combination of the off- and on-diagonal elements of the small discrete cofactor matrix, then, that is what remains of the inverse Hessian after the parameter-type approximation is made, is the mechanism for suppressing cross-talk.

Secondly, the balance of the update across angles. The numerator, eq. (104), in spite of its positive traits, is still a function of the specific angles of data included in the inversion (in this case, θ_1 and θ_2). However, all remaining angle-dependent coefficients match those of the determinant, and so they are suppressed by the denominator. These two points correspond with items three and four on the list of traits of the parameter-type update given in the introduction. To first order in δs_κ, δs_ρ, and $\sin^2\theta$, the parameter-type inverse Hessian components produce an update which is indistinguishable from the ideal benchmark.

5 DISCUSSION

Seismic FWI is multivariate in two ways: there are multiple parameters to be solved for, and each parameter must be solved for at each point in space. In a full Newton update, the off-diagonal elements of the inverse Hessian affect the update through coupling of these variables with each other. Some couple different parameters at the same point in space; others couple the same parameters at different points in space; many do a mixture of the two. The computational intractability of the full Newton update in FWI is a strong motivation to form simplified approximations of the inverse Hessian, which may be designed to retain one or other type of such coupling. Another is that through analysis of various approximations, a sort of 'map' of the Hessian can be sketched out, highlighting where and how important aspects of the multiparameter seismic inverse problem might be managed in a FWI updating scheme. Such a map could influence FWI calculations directly, by suggesting forms for approximate inverse Hessian operators to be constructed and applied to gradients; however, it would also provide a framework in which ad hoc preconditioning using existing technology (e.g. AVO algorithms) could be properly characterized.

A diagram representing the general two-parameter (acoustic) Hessian as realized in eq. (9), organized similarly to Fig. 3 in the paper by Operto *et al.* (2013), is presented in Fig. 3(a). Each of the four block elements represents an integral operator coupling parameters to each other, with coupling between different parameters occurring in the off-diagonal blocks. For instance, the top right element represents the operator $\int d\mathbf{r}' H_{\kappa\rho}(\mathbf{r},\mathbf{r}')[\cdot]$, which measures coupling between κ and ρ. Each of these four block elements also internally contains both on-diagonal and off-diagonal elements, with off-diagonal elements coupling different spatial locations to one another. An approximation in which we emphasize coupling between parameters is illustrated in Fig. 3(b). Here, the four block elements are each assumed to be internally diagonal, meaning we neglect correlations between different points in space, but all four such block elements are retained, meaning we do not neglect correlations between parameters. This leads to what we have referred to as the parameter-type inverse Hessian approximation. An

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

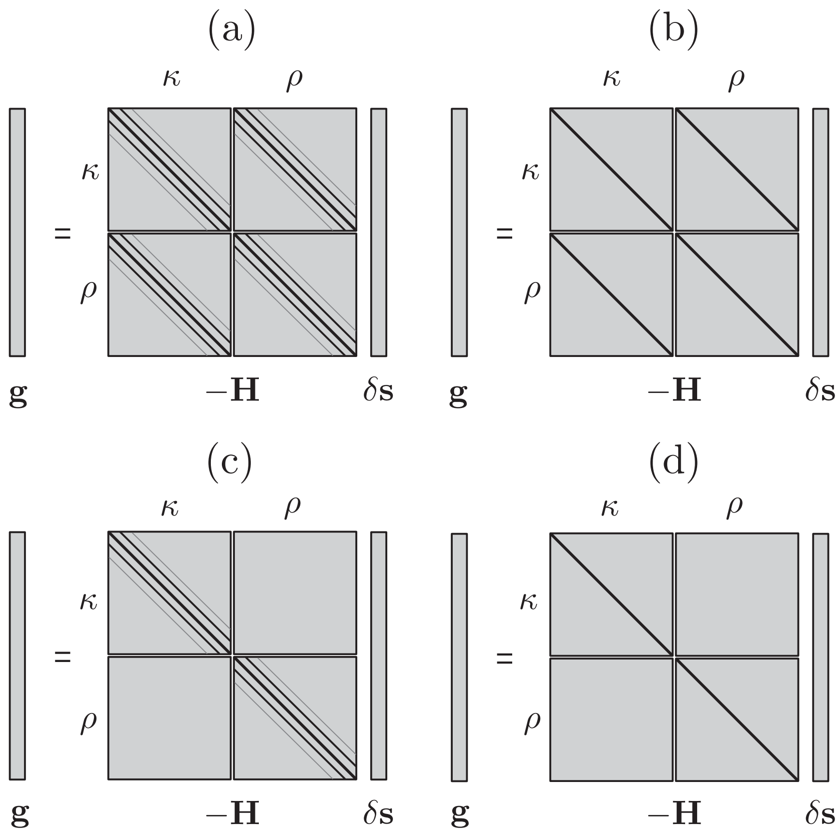

Figure 3. Schematic diagrams of the two-parameter acoustic Hessian as expressed mathematically in eq. (9), organized into a 2×2 block operator representing the four modes of coupling between bulk modulus κ and density ρ. (a) The full Hessian; (b) Hessian components retained in the parameter-type inverse formula, with the diagonals of each block element retained; (c) Hessian components retained in the space-type inverse formula, with the full block elements on the diagonal of the 2×2 system retained; (c) only diagonal Hessian components are retained.

approximate Hessian which, conversely, neglects off-diagonal block elements, but which retains the off-diagonal elements internal to each remaining block element, is illustrated in Fig. 3(c). The update resulting from inversion of this Hessian we have referred to as the space-type approximation. Neglecting both types of off-diagonal element simultaneously leads to a retention of diagonal Hessian elements only (Fig. 3d).

The parameter-type approximation, when incorporated in a FWI update, has been interpreted (eq. 106) as acting to generate balanced and, to first order, cross-talk free updates. In the special case of layered media, in which horizontal planar interfaces are to be reconstructed, it carries out steps comparable to those in standard linear AVO inversion. This correspondence raises additional questions. Through it, we see that the parameter-type approximation does not incorporate all of the interparameter coupling produced by the physics of a seismic reflection. The full AVO problem is intrinsically nonlinear, with the nonlinearity of the reflection amplitudes leading to an additional, and wholly different kind of interparameter coupling (Innanen 2013). Only the residual-dependent portions of the inverse Hessian, neglected in the parameter-type approximation, have the wherewithal to incorporate even approximate forms of this type of nonlinearity. Establishing the extent to which a full Newton step can incorporate true AVO nonlinearity is a matter of ongoing research.

The essential consistency of the first iteration of a parameter-type FWI update and direct AVO inversion is best seen by considering the two in terms of formulas for estimation of $(\delta s_\kappa / s_{\kappa_0})$ and/or $(\delta s_\rho / s_{\rho_0})$. For instance, the bulk modulus update in eq. (101) evaluated for the single-interface case with two angles θ_1 and θ_2 can be written

$$\left(\frac{\delta s_\kappa^{(0)}}{s_{\kappa_0}}\right) = F_\kappa^{\mathrm{fwi}}(\theta_1, \theta_2) R(\theta_1) + F_\kappa^{\mathrm{fwi}}(\theta_2, \theta_1) R(\theta_2), \tag{108}$$

where

$$F_\kappa^{\mathrm{fwi}}(A, B) = \left(\frac{\cos^3 A}{\cos^3 B} - \frac{\cos A}{\cos B}\right)\left(\frac{\cos^3 A}{\cos^5 B} - \frac{2}{\cos A \cos B} + \frac{\cos^3 B}{\cos^5 A}\right)^{-1}. \tag{109}$$

If the linear AVO relations in eq. (46) are directly inverted with the same pair of angles we obtain instead

$$\left(\frac{\delta s_\kappa^{(0)}}{s_{\kappa_0}}\right) = F_\kappa^{\mathrm{avo}}(\theta_1, \theta_2) R(\theta_1) + F_\kappa^{\mathrm{avo}}(\theta_2, \theta_1) R(\theta_2), \tag{110}$$

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

where

$$F_\kappa^{\mathrm{avo}}(A, B) = 4\cos^2 B \left(\frac{\cos^2 B}{\cos^2 A} - \frac{\cos^2 A}{\cos^2 B}\right)^{-1}, \tag{111}$$

a result which has been shown (in a slightly different form) to be a special case of linearized acoustic inverse scattering (Zhang & Weglein 2009). A parameter-type FWI update, posed as in Sections 2– 4, and direct AVO inversion, are equivalent to the extent that these two formulas agree. There are differences between the way the functions F_κ^{fwi} and F_κ^{avo} work. Given two or more linear equations in two or more unknowns, any individual unknown must be determined through differencing. Eq. (108) effects differencing because the cofactor matrix part of $F_\kappa^{\mathrm{fwi}}(A, B)$ changes sign under the exchange of A and B. In contrast, eq. (110) effects differencing because the determinant part of $F_\kappa^{\mathrm{avo}}(A, B)$ changes sign under a similar exchange. The mathematical provenance of cross-talk suppression being slightly different between AVO to FWI, it is not therefore appropriate to call the two perfectly equivalent, at least in their internal workings. However, because both to first-order produce identical determinations of the updates δs_κ etc., we may view them as being formally reconciled.

6 CONCLUSIONS

FWI is currently undergoing large-scale development and deployment in academic and industrial domains. Its ultimate relevance and role in the seismic processing toolbox within these domains is unknown. Part of the heavy lifting to be done in progressing FWI research—and establishing its role and relevance—involves not only making algorithms, but making those algorithms intelligible, and easily exposable to analysis. We are here concerned with one aspect of this large problem, namely understanding what the roles of the quantities familiar in FWI (e.g. gradients and Hessian operators) are in incorporating the pre-critical reflection amplitude information we normally subject to AVO analysis. We know almost instinctively, from years of collective progress in AVO inversion in the geophysics community, how angle variations in the reflection coefficient can be combined in sums and differences to separate out the influence of various parameters—to suppress cross-talk, in the language of FWI. The purpose of this paper has been to determine whether, where and how in the mathematics of FWI something similar automatically occurs.

A challenge is that full Hessians are complex and difficult to analyse (as well as store and invert). We can, however, devise simple formulas for the inverse Hessian if we assume its block elements to have certain commutation properties. The parameter-type inverse Hessian, which is one such approximation, contains machinery which is qualitatively reminiscent of AVO inversion. Assembling the ingredients for a FWI update based on this approximation, and applying them to the problem of reconstructing a single acoustic boundary, we find formulas which are indeed consistent with AVO inversion (see, e.g. eqs 108 and 110). To finalize the comparison, we re-write the linearized form of the acoustic reflection coefficient in terms of the idealized two parameter FWI updates we would find in a perfect world, and substitute these into formulas (108) and (110). The parameter-type updates, in contrast to, for example, gradient-based updates, are seen to reduce to the idealized formulas. We conclude that precritical reflection FWI and AVO inversion can, in this sense, be formally reconciled.

For simplicity in this paper we have analysed the acoustic update formula (Section 2.3), rather than the elastic formula (Section 2.4). Linear acoustic AVO is a weak approximation to linear *PP* elastic AVO (*PP* meaning involving scattering from an incident *P* wave to an outgoing *P* wave), and has no capacity to model shear (*S*) wave reflections or conversions. However, the basic mathematical form of the acoustic AVO equations mimics that of each individual (*PP*, *PS*, *SP*, *SS*) component of the linearized elastic AVO equations. Therefore, the results and demonstrations in Section 4, which focus on mathematical forms and 'machinery', in favour of the prediction of actual numbers, can be expected to carry over in essence to the elastic-isotropic case also. An interesting line of inquiry is to extend the present analysis to accommodate nonlinear behaviour also, as it affects AVO through the full Zoeppritz equations, and FWI through the residual-dependent portions of the inverse Hessian and nonlinear sensitivities. Such an extension will, in contrast to the current linear analysis, engender many differences between acoustic and elastic results. That is a matter of ongoing research.

In spite of the fact that we have underlined several of its admirable traits, the primary value of the parameter-type inverse Hessian is unlikely to lie in its isolated use. It seems more constructive to view the analysis in this paper as having highlighted a quasi-independent system within the full inverse Hessian which has an identifiable role to play. If precritical reflection amplitudes are expected to contribute to multiparameter updates in a particular FWI problem, the system isolated in the parameter-type approximation should be expected to have a tangible positive effect on the outcome.

ACKNOWLEDGEMENTS

This work was funded by the CREWES project and an NSERC Discovery grant. CREWES sponsors and research personnel are thanked for their support. Technical discussions with Gary Margrave, Matt Yedlin and Alison Malcolm are gratefully acknowledged.

REFERENCES

Aki, K. & Richards, P.G., 2002. *Quantitative Seismology,* 2nd edn, University Science Books.

Bleibinhaus, F., Hole, J.A., Ryberg, T. & Fuis, G.S., 2007. Structure of the California Coast Ranges and San Andreas Fault at SAFOD from seismic waveform inversion and reflection imaging, *J. geophys. Res.,* **112,** doi:10.1029/2006JB004611.

Brossier, R., Operto, S. & Virieux, J., 2009. 2D elastic frequency-domain full-waveform inversion for imaging complex onshore structures, *EAGE Expanded Abstracts*.

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022

Brossier, R., Operto, S. & Virieux, J., 2010. Which data residual norm for robust elastic frequency-domain full waveform inversion?, *Geophysics,* **75,** WCC105–WCC118.

Brossier, R., Operto, S. & Virieux, J., 2013. Toward data-domain waveform inversion of reflected waves, *SEG Expanded Abstracts,* pp. 892–897.

Castagna, J.P. & Backus, M., 1993. *Offset-Dependent Reflectivity: Theory and Practice of AVO Analysis,* SEG.

Clayton, R.W. & Stolt, R.H., 1981. A Born-WKBJ inversion method for acoustic reflection data, *Geophysics,* **46**(11), 1559–1567.

Fichtner, A., 2010. *Full Seismic Waveform Modelling and Inversion,* Springer.

Fichtner, A. & Trampert, J., 2011. Hessian kernels of seismic data functionals based on adjoint techniques, *Geophys. J. Int.,* **185**.

Foster, D.J., Keys, R.G. & Lane, F.D., 2010. Interpretation of AVO anomalies, *Geophysics,* **75,** 75A3–75A13.

Guitton, A. & Symes, W.W., 2003. Robust inversion of seismic data using the Huber norm, *Geophysics,* **68**(4), 1310–1319.

Hampson, D. & Russell, B.H., 1990. AVO inversion: theory and practice, *SEG Expanded Abstracts,* 23–27 September, San Francisco, CA, USA.

Innanen, K.A., 2013. Coupling in amplitude variation with offset and the Wiggins approximation, *Geophysics,* **78**(4), N21–N33.

Innanen, K.A. & Weglein, A.B., 2007. On the construction of an absorptive-dispersive medium model via direct linear inversion of reflected seismic primaries, *Inverse Prob.,* **23,** 2289–2310.

Kelly, S., Ramos-Martinez, J., Zou, K. & Tsimelzon, B., 2013. Inversion of refractions and reflections by full-waveform inversion for marine streamer data: classification of problem types and solution methods, *Leading Edge,* **32,** 1130–1138.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations, *Conference on Inverse Scattering, Theory and Application, Society of Industrial and Applied Mathematics, Expanded Abstracts,* pp. 206–220.

Margrave, G.F. *et al.*, 2012. A low-frequency seismic field experiment, *SEG Expanded Abstracts,* pp. 1–5.

Margrave, G.F., Ferguson, R.J. & Hogan, C.M., 2012. Full waveform inversion by iterative depth migration and impedance estimation using well control, *EAGE Expanded Abstracts*.

Métivier, L., Brossier, R., Virieux, J. & Operto, S., 2013. Full waveform inversion and the truncated Newton method, *SIAM J. Sci. Comput.,* **35**(2), B401–B437.

Métivier, L., Bretaudeau, F., Brossier, R., Operto, S. & Virieux, J., 2014. Full waveform inversion and the truncated Newton method: quantitative imaging of complex subsurface structures, *Geophys. Prospect.,* doi: 10.1111/1365-2478.12136.

Oldenburg, D.W., 1984. An introduction to linear inverse theory, *IEEE Trans. Geosci. Remote Sens.,* **GE-22,** 665–674.

Operto, S., Virieux, J., Dessa, J.X. & Pascal, G., 2006. Crustal seismic imaging from multifold ocean bottom seismometer data by frequency-domain full-waveform tomography: application to the eastern Nankai trough, *J. geophys. Res.,* **111,** B09306, doi:10.1029/2005JB003835.

Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Metivier, L. & Virieux, J., 2013. A guided tour of multiparameter full-waveform inversion with multicomponent data: from theory to practice, *Leading Edge,* **32,** 1040–1054.

Parker, R.L., 1994. *Geophysical Inverse Theory,* 1st edn, Princeton Univ. Press.

Plessix, R.-E., Baeten, G., de Maag, J.W., ten Kroode, F. & Rujie, Z., 2012. Full waveform inversion and distance separated simultaneous sweeping: a study with a land seismic data set, *Geophys. Prospect.,* **60**(4), 733–747.

Plessix, R.-E., Milcik, P., Rynia, H., Stopin, A., Matson, K. & Abri, S., 2013. Multiparameter full-waveform inversion: marine and land examples, *Leading Edge,* **32,** 1030–1038.

Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain, part 1: theory and verification in a physical scale model, *Geophysics,* **64**(3), 888–901.

Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss-newton and full newton mthods in frequency-space seismic waveform inversion, *Geophys. J. Int.,* **133,** 341–362.

Raz, S., 1981. Direct reconstruction of velocity and density profiles from scattered field data, *Geophysics,* **46,** 832–836.

Saad, Y., 2003. *Iterative Methods for Sparse Linear Systems,* 1st edn, SIAM.

Shin, C., Yoon, K., Marfurt, K.J., Park, K., Yang, D., Lim, H.Y., Chung, S. & Shin, S., 2001. Efficient calculation of a partial-derivative wavefield using reciprocity for seismic imaging and inversion, *Geophysics,* **66**(6), 1856–1863.

Stolt, R.H. & Weglein, A.B., 1985. Migration and inversion of seismic data, *Geophysics,* **50**(12), 2458–2472.

Symes, W.W. & Carazzone, J.J., 1991. Velocity inversion by differential semblance optimization, *Geophysics,* **56,** 654–663

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics,* **49,** 1259–1266.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74**(6), WCC1–WCC26.

Wang, S., Chen, F., Zhang, H. & Shen, Y., 2013. Reflection-based full waveform inversion (RFWI) in the frequency domain, *SEG Expanded Abstracts,* pp. 877–881.

Weglein, A.B., Violette, P.B. & Keho, T.H., 1986. Using multiparameter born theory to obtain certain exact multiparameter inversion goals, *Geophysics,* **51,** 1069–1074.

Weglein, A.B. *et al.*, 2003. Inverse scattering series and seismic exploration, *Inverse Problems,* **19,** R27–R83.

Zhang, H. & Weglein, A.B., 2009. Direct nonlinear inversion of multiparameter 1D acoustic media using inverse scattering subseries, *Geophysics,* **74**(6), WCD29–WCD39.

Downloaded from https://academic.oup.com/gji/article/199/2/717/619452 by guest on 16 August 2022