Journal of Computational Physics 346 (2017) 318–339

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Approximate solutions of acoustic 3D integral equation and their application to seismic modeling and full-waveform inversion

M. Malovichko [a,*], N. Khokhlov [a], N. Yavich [a], M. Zhdanov [b,c,a]

[a] Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, 141700, Russia
[b] Consortium for Electromagnetic Modeling and Inversion (CEMI), University of Utah, Salt Lake City, UT 84112, USA
[c] TechnoImaging, Salt Lake City, UT 84107, USA

ARTICLE INFO

Article history:
Received 15 August 2016
Received in revised form 9 June 2017
Accepted 10 June 2017
Available online 15 June 2017

Keywords:
Seismology
Acoustics
Integral equation
Full-waveform inversion

ABSTRACT

Over the recent decades, a number of fast approximate solutions of Lippmann–Schwinger equation, which are more accurate than classic Born and Rytov approximations, were proposed in the field of electromagnetic modeling. Those developments could be naturally extended to acoustic and elastic fields; however, until recently, they were almost unknown in seismology. This paper presents several solutions of this kind applied to acoustic modeling for both lossy and lossless media. We evaluated the numerical merits of those methods and provide an estimation of their numerical complexity. In our numerical realization we use the matrix-free implementation of the corresponding integral operator. We study the accuracy of those approximate solutions and demonstrate, that the quasi-analytical approximation is more accurate, than the Born approximation. Further, we apply the quasi-analytical approximation to the solution of the inverse problem. It is demonstrated that, this approach improves the estimation of the data gradient, comparing to the Born approximation. The developed inversion algorithm is based on the conjugate-gradient type optimization. Numerical model study demonstrates that the quasi-analytical solution significantly reduces computation time of the seismic full-waveform inversion. We also show how the quasi-analytical approximation can be extended to the case of elastic wavefield.

1. Introduction

The integral equation (IE) method is a well-known technique for wavefield modeling [18,3,6]. It has a number of attractive properties. It requires discretization for the anomalous volume only, and this drastically reduces the size of the problem in many cases. The IE formulation does nor require boundary conditions. Also, it naturally applies to computation of the Fréchet derivatives and for this reason it is very attractive for the solution of the inverse problem.

The numerical implementation of the IE method is based on discretization of the original integral equation, which in a general case results in a complex-valued dense matrix [17,9,11,27,10,2,4,1,16].

* Corresponding author.
E-mail address: malovichko.mikhail@gmail.com (M. Malovichko).

http://dx.doi.org/10.1016/j.jcp.2017.06.021
0021-9991/

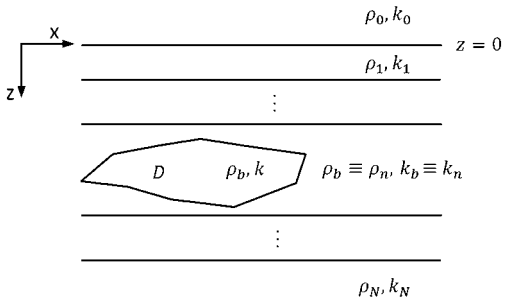

Fig. 1. 3D anomalous volume D embedded in a layered host medium.

Closely connected to the full IE method are the approximate solutions of the integral equation. The most well-known of them are the Born and Rytov approximations. These solutions are quite inaccurate in general. However, for some media, such as low-contrast models, they give decent approximations of the direction of the data gradient relatively fast. That is why they are widely used in the seismic migration and full-waveform inversion [8,26,7,25].

A great progress in this direction has been made within the area of electromagnetic simulation, starting from the pioneered work of [13]. Several approximate solutions, with a wide range of validity, were presented in that study. All of them, including the most notable localized non-linear approximation, are based on the assumption that, the Green's operator of electromagnetic scattering problem is localized. This approach was further advanced and a variety of approximations have been developed, such as the quasi-linear (QL), quasi-analytical (QA) and localized quasi-analytical (LQL) approximations [32,31]. These solutions are proven to be powerful tool for the electromagnetic simulation and, especially, inversion. They rapidly deliver quite an accurate estimation for the data gradient, thus considerably reducing computation time.

An application of a similar approach to acoustic and elastic modeling was proposed in [29], but a few numerical studies were done. To our knowledge, the only application to real problems has been reported in [15], where the QL approximation was used to develop an effective migration algorithm.

In this study we evaluate the localized quasi-linear and quasi-analytical solutions in acoustic modeling and inversion, and compare them to the full IE solution and the Born approximation. We review the underlying theory and effective numerical implementations. We also introduce the method of the Fréchet derivative operator computation based on these approximations, and provide an estimation of the numerical complexity of different solutions. Further, we study the accuracy of approximate solutions in lossy and lossless media. The results of our study show that, the QA approximation is more accurate than the Born approximation, whereas the LQL approximation is not applicable for realistic simulation. We utilize the QA solution for calculation of the Fréchet derivatives within the framework of the conjugate-gradient type minimization. We have tested this approach on a synthetic data set computer simulated for quite complex 3D models, and demonstrate, that the QA estimation can be considered to be an effective tool for modeling and inversion of the full-waveform seismic data.

2. Theory

The earth model is formed by a 3D volume D with anomalous acoustic properties, embedded in a layered host medium (Fig. 1).

Each n-th layer of the host medium is specified by its thickness, h_n, wavenumber, k_n and density, ρ_n. We assume that the anomalous volume is confined to a single layer with $k_n = k_b$ and $\rho_n = \rho_b$. A three-dimensional anomalous volume is characterized by wavenumber $k(\boldsymbol{r})$, where $\boldsymbol{r} = (x, y, z)$ is a radius vector of the point of observation. Let k_a be the anomalous wavenumber such as

$$k_a^2(\boldsymbol{r}) = k^2(\boldsymbol{r}) - k_b^2. \tag{1}$$

For the lossless case we define

$$k(\boldsymbol{r}) = \frac{\omega}{c(\boldsymbol{r})}, \tag{2}$$

$$k_b = \frac{\omega}{c_b}, \tag{3}$$

with $c(\boldsymbol{r}) = c_b(\boldsymbol{r}) + c_a(\boldsymbol{r})$, where $c_a(\boldsymbol{r})$ is the velocity perturbation, ω is the circular frequency.

If the medium is lossy, wavenumber has to be a complex-valued scalar function with $Im(k) \geq 0$ to satisfy the radiation conditions. Let us define a complex-valued wavenumber through attenuation function Q [3]:

$$k(c, Q) = \frac{\omega}{c}\left(1 + \frac{i}{2Q}\right). \tag{4}$$

Fourier transform of the pressure field, p within a layer satisfies the following Helmholtz equation:

$$-\nabla^2 p(\boldsymbol{r}) - k^2 p(\boldsymbol{r}) = f(\boldsymbol{r}), \tag{5}$$

where f is the source term.

The total pressure p can be represented as follows:

$$p(\boldsymbol{r}) = p_a(\boldsymbol{r}) + p_b(\boldsymbol{r}), \tag{6}$$

where p_b is the background part due to host medium and p_a is the anomalous (scattered) part due to the volume with anomalous velocity.

The anomalous field p_a satisfies the following Helmholtz equation [29]:

$$-\nabla^2 p_a(\boldsymbol{r}) - k_b^2 p_a(\boldsymbol{r}) = k_a^2(\boldsymbol{r})\left(p_a(\boldsymbol{r}) + p_b(\boldsymbol{r})\right). \tag{7}$$

The solution of this equation is given by the formula:

$$p_a(\boldsymbol{r}) = \int_D g(\boldsymbol{r}|\boldsymbol{r}') k_a^2(\boldsymbol{r}')\big(p_a(\boldsymbol{r}') + p_b(\boldsymbol{r}')\big) d^3\boldsymbol{r}', \tag{8}$$

where g is the Green's function, primed coordinates denote the source position.

Solving equation (7) in form (8) requires discretization of the anomalous volume only. This reduces the size of the problem drastically in the cases with a relatively compact anomalous domain immersed into a host medium where the background field can be accurately modeled analytically or using a relatively simple numerical methods. In exploration geophysics such an example of a host medium is a layered Earth.

Introducing a contrast function χ,

$$\chi = k_a^2, \tag{9}$$

the solution for the total field can be written as Lippmann–Schwinger equation with respect to p,

$$p(\boldsymbol{r}) = p_b(\boldsymbol{r}) + \mathcal{G}[\chi\, p], \tag{10}$$

where

$$\mathcal{G}[\chi\, p] = \int_D g(\boldsymbol{r}|\boldsymbol{r}')\chi(\boldsymbol{r}') p(\boldsymbol{r}') d^3\boldsymbol{r}'. \tag{11}$$

The background Green's function g satisfies to the following equation:

$$-\nabla^2 g(\boldsymbol{r}|\boldsymbol{r}') - k_b^2 g(\boldsymbol{r}|\boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}'), \tag{12}$$

where $\delta(\boldsymbol{r} - \boldsymbol{r}')$ is the Dirac delta function.

The background field p_b satisfies to the following Helmholtz equation [29]

$$-\nabla^2 p_b(\boldsymbol{r}) - k_b^2 p_b(\boldsymbol{r}) = f(\boldsymbol{r}). \tag{13}$$

In this study we model the field due to the point source $f(\boldsymbol{r}) = \delta(\boldsymbol{r} - \boldsymbol{r}')$, so p_b and g, in fact, satisfy the same equation.

The background Green's function is composed of primary and secondary parts,

$$g = g^p + g^s, \tag{14}$$

which both satisfy the Helmholtz equation. The primary component g^p is Green's function for the free space. When the source and receiver are located in the same layer, g^p is nonzero,

$$g^p = \frac{e^{ikR}}{4\pi R}, \tag{15}$$

with $R = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$.

The solution for the secondary component, g^s, can be obtained for any horizontally stratified earth by using the standard wavenumber integration technique [22]. Here we confine ourselves with a double-halfspace model, which consists of the air half space ($z < 0$, $k_0 = \infty$, $\rho_0 = 0$) and subsurface ($z > 0$, $k_b < \infty$, $\rho_b > 0$). When both receiver and transmitter are located within the subsurface, the expression for g^s is as follows

$$g^s = -\frac{e^{ikR_1}}{4\pi R_1}, \tag{16}$$

where $R_1 = \sqrt{(x-x')^2 + (y-y')^2 + (z+z')^2}$.

In the framework of the integral equation formulation, numerical solutions for p are based on discretization of equation (10). The problem is solved in two main steps. First, equation (10) is treated as an integral equation with respect to field p within domain D ($\boldsymbol{r} \in D$). We call this equation *a domain equation*. Once the field inside the scatterer is found, equation (10) is used again to calculate field p in all required exterior points ($\boldsymbol{r} \notin D$). In this case, equation (10) is called *a field equation*. The first step is by far the most time-consuming and storage-demanding part for realistic problems. Various approximate solutions to the original integral equation are aimed to simplify the domain equation in order to receive computational reduction without a significant loss of accuracy.

2.1. Solutions of the Lippmann–Schwinger equation

We start with the **full integral equation** (IE) method. When $\boldsymbol{r} \in D$, expression (10) is the integral equation with respect to the total field p

$$p(\boldsymbol{r}) - \mathcal{G}[\chi p] = p_b(\boldsymbol{r}). \tag{17}$$

In order to obtain a numerical solution of equation (17) we discretize domain D by $N = N_x N_y N_z$ cubical cells, where N_x,N_y, and N_z are numbers of cells in the corresponding directions. The cells are presumed to be small enough so that p, p_b and χ are constant withing each cell. For simplicity, we assume that all cells have the same dimensions.

After discretization, we obtain the following matrix equation:

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{b}, \tag{18}$$

where $\boldsymbol{u} = (p_1..p_N)$ is the vector of N unknown values of p in the center of each cell, $\boldsymbol{b} = (p_{b,1}..p_{b,N})$ is the vector of known values of the background field in each cell, $\boldsymbol{A}$ is the scattering matrix,

$$\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{G}\boldsymbol{X}, \tag{19}$$

where $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{X} = diag(\chi_1..\chi_N)$ is diagonal matrix formed by contrasts for each cell, $\boldsymbol{G}$ is a dense matrix containing integrals of Green's function over elementary cells D_j:

$$G_{ij} = \int_{D_j} g(\boldsymbol{r}_i|\boldsymbol{r}')d^3\boldsymbol{r}', \quad i, j = 1..N, \quad \boldsymbol{r}_i, \boldsymbol{r}' \in D. \tag{20}$$

We use an iterative solver for problem (18). An effective numerical implementation of the IE method is the topic of a separate paper. However, operation $\boldsymbol{G}\boldsymbol{v}$, where $\boldsymbol{v}$ is an arbitrary complex-valued vector, is important in the context of approximate solutions. It can be implemented via FFT (for example, [1]). The cost of this operation for the half-space host medium, used in this article, is $O(NN_z log(N_x N_y))$. The memory requirements is $O(N)$.

The Born approximation is the most well-known and commonly used one, frequently in the context of migration and full-waveform inversion. It assumes that the scattered field inside the anomalous domain is negligible, that is

$$p \approx p_b. \tag{21}$$

Condition (21) is equivalent to the following condition

$$\int_D g(\boldsymbol{r}|\boldsymbol{r}')\chi(\boldsymbol{r}')p(\boldsymbol{r}')d^3\boldsymbol{r}' \approx 0. \tag{22}$$

It can be seen (for example, [13]) that the range of applicability of the Born approximation is limited to the cases, where the following quantity is small:

$$max\{|\chi|\}H^2, \tag{23}$$

where H is the maximum distance between any two cells within the anomalous volume.

The field inside the scatterer is known immediately after p_b is computed. The anomalous field p_a^B at the receivers is given by the following formula

$$p_a^B(\boldsymbol{r}) = \mathcal{G}[\chi p_b]. \tag{24}$$

The **quasi-linear** (QL) approximation, developed for electromagnetic field by [32] and later extended to acoustics and linear elasticity in [29], is based on an assumption, that the anomalous field within domain D is proportional to the background field via coefficient λ:

$$p_a^{QL}(\boldsymbol{r}) = \lambda(\boldsymbol{r})p_b(\boldsymbol{r}). \tag{25}$$

Thus, for interior points $\boldsymbol{r} \in D$ we have an equation with respect to the unknown scalar field λ:

$$\lambda(\boldsymbol{r})p_b(\boldsymbol{r}) = \mathcal{G}[\chi(1+\lambda)p_b]. \tag{26}$$

In the original paper equation (26) was solved by minimizing the following functional

$$\|\lambda(\boldsymbol{r})p_b(\boldsymbol{r}) - \mathcal{G}[\chi(1+\lambda)p_b]\|_{L_2} \underset{\lambda}{\rightarrow} min. \tag{27}$$

This is convenient if coefficients λ are defined on a grid, which are coarser, than the original discretization. The resulting system is overdetermined and approximate solution can be found numerically. In this paper we assume the system to be consistent (i.e. the discretization is the same) and thus λ is the solution of the equation

$$\lambda p_b(\boldsymbol{r}) - \mathcal{G}[\chi\lambda p_b] = \mathcal{G}[\chi p_b]. \tag{28}$$

The matrix form of equation (28) is as follows:

$$(\boldsymbol{I} - \boldsymbol{GX})\boldsymbol{B}\boldsymbol{\Lambda} = \boldsymbol{GBXe}, \tag{29}$$

where $\boldsymbol{\Lambda} = (\lambda_1..\lambda_N)$, $\boldsymbol{B} = diag(p_{b,1}..p_{b,N})$ and $\boldsymbol{e} = (1..1)$. The right-hand side is, in fact, the Born field p_a^B for interior points and it requires one matrix multiplication only. Also, the number of iterations, M, differs from that for equation (18). Once λ is found, p_a is calculated through (25).

Despite its name, the QL approximation in form (28) provides an exact solution in the sense, that the discrete solution converges to the true solution when grid step size approaches zero. For this reason we did not include the QL approximation in comparative accuracy analysis.

The **quasi-analytical** (QA) approximation is based on the assumption, that the Green's operator $\mathcal{G}$ is localized with respect to λ. That is

$$\mathcal{G}[\lambda\chi p_b] \approx \lambda(\boldsymbol{r})\mathcal{G}[\chi p_b]. \tag{30}$$

With this, equation (28) takes the form

$$\lambda^{QA} p_b(\boldsymbol{r}) - \mathcal{G}[\chi p_b]\lambda^{QA}(\boldsymbol{r}) = \mathcal{G}[\chi p_b], \tag{31}$$

where λ^{QA} is given by

$$\lambda^{QA}(\boldsymbol{r}) = \frac{p_a^B(\boldsymbol{r})}{p_b(\boldsymbol{r}) - p_a^B(\boldsymbol{r})}. \tag{32}$$

From the above equation it follows that

$$p_a^{QA}(\boldsymbol{r}) = \mathcal{G}[\chi p_b(1+\lambda^{QA})]. \tag{33}$$

Thus, the cost of generating the interior field by the QA approximation is roughly equal to the cost of generation p_a^B for interior points, i.e. roughly equals to a single matrix multiplication.

Condition (30) is equivalent to the following condition:

$$\int_D g(\boldsymbol{r}|\boldsymbol{r}')\chi(\boldsymbol{r}')p_b(\boldsymbol{r}')\big[\lambda(\boldsymbol{r}) - \lambda(\boldsymbol{r}')\big]d^3\boldsymbol{r}' \approx 0. \tag{34}$$

Thus, the QA approximation is valid, when Green's function g decays much faster, than variations in λ. This means that, the range of applicability of the QA approximation is wider than that of the Born approximation, since condition (34) is met not only when (23) is satisfied, but also when $\lambda(\boldsymbol{r})$ varies slowly comparing to Green's function.

The extension to the elastic case can be developed in a similar manner (see Appendix A).

The **localized quasi-linear** (LQL) approximation is based on the assumption that Green's operator is localized with respect to the background field p_b as follows:

$$\mathcal{G}[\chi\lambda p_b] \approx \mathcal{G}[\chi\lambda]p_b(\boldsymbol{r}). \tag{35}$$

Under this assumptions, equation (28) can be rewritten in the following form:

$$\lambda^{LQL}(\boldsymbol{r})p_b(\boldsymbol{r}) - \mathcal{G}[\chi\lambda^{LQL}]p_b(\boldsymbol{r}) = \mathcal{G}[\chi]p_b(\boldsymbol{r}) \tag{36}$$

Thus, we have equation with respect to λ^{LQL}

$$\lambda^{LQL}(\boldsymbol{r}) - \mathcal{G}[\chi\lambda^{LQL}] = \mathcal{G}[\chi]. \tag{37}$$

The matrix form is given by

$$(\boldsymbol{I} - \boldsymbol{GX})\boldsymbol{\Lambda} = \boldsymbol{GXe}. \tag{38}$$

Note that, the coefficient matrix for this system is the same as for the original IE system. The right-hand side requires one matrix multiplication. Once λ^{LQL} is found, the solution is given by

$$p_a^{LQL}(\boldsymbol{r}) = \mathcal{G}[\chi\, p_b(1+\lambda^{LQL})]. \tag{39}$$

The LQL approach is independent of the background field and needs to be computed once for multiple seismic sources. This is appealing in view of an inverse problem solution.

Condition (35) is equivalent to the following expression:

$$\int_D g(\boldsymbol{r}|\boldsymbol{r}')\chi(\boldsymbol{r}')\lambda(\boldsymbol{r}')\big[p_b(\boldsymbol{r}) - p_b(\boldsymbol{r}')\big]d^3\boldsymbol{r}' \approx 0. \tag{40}$$

Hence, this approximation is valid not only, when the Born condition (23) is satisfied, but also when $p_b(\boldsymbol{r})$ varies slowly with respect to Green's function.

2.2. Application to inverse scattering

In this section we apply the QA and Born approximations to acoustic inverse problems. The LQL solution is not included here because, as it will be shown below, it is the least accurate method among all, considered in this study.

In order to consider multi-frequency data, let us introduce a model function m that does not depend on frequency,

$$m = \frac{1}{c^2} - \frac{1}{c_b^2}, \tag{41}$$

and vector $\boldsymbol{m}$, that contains discrete values of the models function m at the centers of the model cells, $\boldsymbol{m} = (m_1..m_N)$.

In a lossless medium m and χ are related by the formula $\chi = \omega^2 m$. When the medium is lossy, the factor changes, depending on the particular formulation of the attenuation. For generality, we formulate expressions for the Fréchet derivatives with respect to variable χ.

A general nonlinear acoustic inverse problem can be formulated as a solution of the following operator equation:

$$\mathcal{A}_a(\boldsymbol{m}) = \boldsymbol{d}, \tag{42}$$

where $\mathcal{A}_a$ is acoustic forward modeling operator; and $\boldsymbol{d}$ are acoustic field data. A standard approach to the regularized solution of this problem is based on the minimization of the Tikhonov parametric functional [24,29]:

$$P(\boldsymbol{m}) = \Phi(\boldsymbol{m}) + \alpha S(\boldsymbol{m}) \underset{\alpha,\boldsymbol{m}}{\longrightarrow} min, \tag{43}$$

where

$$\Phi(\boldsymbol{m}) = \|\boldsymbol{W}_d(\mathcal{A}_a(\boldsymbol{m}) - \boldsymbol{d})\|^2 \tag{44}$$

is a misfit functional, and S is a stabilizer; α is a regularization parameter, $\boldsymbol{W}_d$ is a diagonal matrix of data weights. Everywhere throughout this paper we assume a norm to be the L^2-norm, $\|\cdot\| = \|\cdot\|_{L_2}$.

In this paper we use the following stabilizer:

$$S(m) = c_1\|\boldsymbol{W}_L(\boldsymbol{m} - \boldsymbol{m_0})\|^2 + c_2\|(\boldsymbol{m} - \boldsymbol{m_0})\|^2 \tag{45}$$

where $\boldsymbol{W}_L$ represents a discrete analog of operator ∇, c_1 and c_2 are two arbitrary weights, $\boldsymbol{m_0}$ is the initial model.

Most of the inversion methods used in geophysical applications are based on gradient-type optimization schemes (i.e. conjugate-gradient or quasi-Newton methods, such as L-BFGS). In either cases, three operations are required at each inversion iteration: forward modeling, computing products $\boldsymbol{F}^*\boldsymbol{\psi}$ and computing $\boldsymbol{Fq}$, where $\boldsymbol{F}$ is Fréchet derivative of the forward modeling operator, which characterizes the sensitivities of anomalous field p_a at receivers to χ:

$$\boldsymbol{F} = \frac{\partial \boldsymbol{p_a}(\boldsymbol{r})}{\partial \boldsymbol{\chi}}, \quad \boldsymbol{r} \notin D; \tag{46}$$

vector $\boldsymbol{\psi}$ represents an arbitrary scalar field, defined at actual receivers and frequencies; vector $\boldsymbol{q}$ represents an arbitrary scalar field, defined within the anomalous volume; asterisk (*) denotes adjoint operator.

Generally, the task of computing products $\boldsymbol{F}^*\boldsymbol{\psi}$ and $\boldsymbol{Fq}$ is accomplished by using the additional forward modeling calculations, which have been addressed many times in literature [29,30,5,21]. Here we introduce explicit matrix-free expressions for the Fréchet derivatives based on the QA and Born approximations.

2.2.1. Fréchet derivative based on the QA approximation

Expression for the Fréchet derivative based on the QA approximation, $\boldsymbol{F}^{QA}$, for electromagnetic field was given by [28]. Here we develop a similar expression for the acoustic field.

Let us denote

$$\boldsymbol{p}^{\psi} = \boldsymbol{\Gamma}^{*}\boldsymbol{\psi}, \tag{47}$$

where $\boldsymbol{\psi}$ is an arbitrary complex-valued vector in the data space, $\boldsymbol{\Gamma}$ is a rectangular matrix, which contains corresponding volume integrals of Green's function acting from the anomalous volume to receivers,

$$\boldsymbol{\Gamma}_{ij} = \int\limits_{D_j} g(\boldsymbol{r}_i|\boldsymbol{r}')d^3\boldsymbol{r}', \quad i = 1..L, \quad j = 1..N, \quad \boldsymbol{r}_i \notin D, \quad \boldsymbol{r}' \in D, \tag{48}$$

and L is the number of receivers. Vector $\boldsymbol{p}^{\psi}$ can be interpreted as a back propagation of a scalar field $\boldsymbol{\psi}$ under the Born approximation.

With the exception of $\boldsymbol{\Gamma}$, other matrices considered below do not need to be formed explicitly, because they are either diagonal or appear in a matrix–vector product only. In order to save memory, we compute matrices $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}^{*}$ "on-the-fly", when corresponding matrix–vector products are required.

A direct differentiation of both sides of (33) with respect to χ yields:

$$\boldsymbol{F}^{QA} = \boldsymbol{\Gamma}\boldsymbol{\Omega}\boldsymbol{B} + \boldsymbol{\Gamma}\boldsymbol{X}\boldsymbol{\Omega}^{2}\boldsymbol{G}\boldsymbol{B}, \tag{49}$$

where

$$\boldsymbol{\Omega} = diag\left(\frac{1}{1-\gamma_i}\right), \quad i = 1..N, \tag{50}$$

with

$$\gamma(\boldsymbol{r}) = \frac{p_a^B(\boldsymbol{r})}{p_b(\boldsymbol{r})}, \tag{51}$$

and $\boldsymbol{G}$ is the matrix of Green's integral operator $\mathcal{G}$, defined by equation (20).

Hence, the action of the adjoint QA Fréchet derivative operator on a residual vector $\boldsymbol{\psi}$ (the steepest ascent direction for the misfit functional, or the data gradient) can be expressed as follows:

$$\left(\boldsymbol{F}^{QA}\right)^{*}\psi = \overline{\boldsymbol{B}\boldsymbol{\Omega}}\boldsymbol{p}^{\psi} + \overline{\boldsymbol{B}}\,\overline{\boldsymbol{G}}\,\overline{\boldsymbol{X}\boldsymbol{\Omega}}^{2}\boldsymbol{p}^{\psi}, \tag{52}$$

where "bar" denotes complex conjugation. We can rewrite the first term as follows:

$$\overline{\boldsymbol{B}\boldsymbol{\Omega}}\boldsymbol{p}^{\psi} = diag(\overline{\boldsymbol{\Omega}}\boldsymbol{p}^{\psi})\,\overline{\boldsymbol{p}}_b. \tag{53}$$

The second term is equal to

$$\overline{\boldsymbol{B}}\,\overline{\boldsymbol{G}}\,\overline{\boldsymbol{X}\boldsymbol{\Omega}}^{2}\boldsymbol{p}^{\psi} = diag\,(\boldsymbol{u})\,\overline{\boldsymbol{p}}_b, \tag{54}$$

where vector $\boldsymbol{u}$ is given by the following formula:

$$\boldsymbol{u} = \overline{\boldsymbol{G}\boldsymbol{X}\boldsymbol{\Omega}}^{2}\boldsymbol{p}^{\psi}. \tag{55}$$

Vector $\boldsymbol{u}$ can be obtained by the same Green's integral operator:

$$\boldsymbol{u} = \overline{\mathcal{G}[\boldsymbol{X}\boldsymbol{\Omega}^{2}\overline{\boldsymbol{p}}^{\psi}]}. \tag{56}$$

Finally, the QA approximation to the data gradient is given by the following formula:

$$\left(\boldsymbol{F}^{QA}\right)^{*}\boldsymbol{\psi} = \left(diag(\overline{\boldsymbol{\Omega}}\boldsymbol{p}^{\psi}) + diag(\boldsymbol{u})\right)\overline{\boldsymbol{p}}_b. \tag{57}$$

Expression (57) can be interpreted as a product of the conjugate direct field $\boldsymbol{p_b}$ and a residual field $\boldsymbol{\psi}$, back-propagated under the QA approximation. The additional cost for generating the gradient for a single source is roughly equal to one integral transform.

Finally, expression for the product of the Fréchet derivative by an arbitrary vector $\boldsymbol{q}$ is as follows:

$$\boldsymbol{F}^{QA}\boldsymbol{q} = \boldsymbol{\Gamma}\left(\boldsymbol{B}\boldsymbol{\Omega} + \boldsymbol{X}\boldsymbol{\Omega}^{2}\boldsymbol{G}\boldsymbol{B}\right)\boldsymbol{q}. \tag{58}$$

2.2.2. Fréchet derivative based on the Born approximation

The steepest ascend direction for the misfit functional (the data gradient) under the Born approximation is given by the following expression:

$$\left(\boldsymbol{F}^{B}\right)^{*}\boldsymbol{\psi} = diag(\boldsymbol{p}^{\psi})\overline{\boldsymbol{p}}_{b}. \tag{59}$$

It can be interpreted as a product of the conjugate direct field $\boldsymbol{p_b}$ and a residual field $\boldsymbol{\psi}$, back-propagated under the Born approximation.

The product of the Fréchet derivative and an arbitrary vector $\boldsymbol{q}$ can be calculated as follows:

$$\boldsymbol{F}^{B}\boldsymbol{q} = \boldsymbol{\Gamma}\boldsymbol{B}\boldsymbol{q}. \tag{60}$$

2.3. Computational complexity

In this section we estimate the computational complexity of different solutions. We assume that, the total number of the sources is K, and the total number of the receivers for all sources is L.

All methods require computing p_b within the anomalous volume for all K sources. The arithmetic complexity to compute p_b is $O(NK)$, where N is the number of cells within the anomalous volume. In our implementation we do not store elements of matrix $\boldsymbol{\Gamma}$, and calculate them "on-the-fly" instead.

All methods, except the Born approximation, require computation of the elements of matrix $\boldsymbol{G}$. The way the elements of matrix G are computed and stored determines the overall memory requirements. Due to the spatial symmetry of Green's function, the number of unique elements of matrix $\boldsymbol{G}$ is much less than $N^2/2$. For a rectangular uniform grid it is not more than N for a free space host medium, and not more than $N(N_z+1)/2$ for a layered host medium. These amounts can be somewhat reduced when the cells are cubical, but not significantly. If the unique elements of matrix G are precomputed, then the memory requirement becomes $O(NN_z)$ for a layered host medium. This poses a great limitation for models of several tens of millions of cells, unless the memory is distributed across several computational nodes. The elements of $\boldsymbol{G}$ may as well be computed on demand. This strategy reduces the amount of memory for the matrix–vector multiplication operation to the amount needed for a single two-dimensional Fourier transform, $O(N_xN_y)$. The overall memory consumption in this case becomes of $O(N)$. We found the latter approach be almost as fast, as the first one, and use it in this study.

The QL and QA approximations require one additional application of Green's operator for the interior points, which has numerical complexity equal to $O(KNlog(N))$ or $O(KNN_zlog(N_xN_y))$, depending on the type of the host medium (full space of half space, respectively). The LQL solution needs just a single application of the Green's operator for all sources, which reduces the numerical complexity of this step to $O(Nlog(N))$ or $O(NN_zlog(N_xN_y))$, respectively.

The full IE method and QL approximation require solving the system of the corresponding equations for each real-world source. For an iterative solver the arithmetic complexity is equal to $O(KMNlog(N))$ or $O(KMNN_zlog(N_xN_y))$, for a free-space host model or half-space model, respectively, where M is the number of iterations, which is different for IE and QL methods. The LQL approximation requires one system solution for all sources.

Once p inside the scatterer is computed, the field must be mapped to L receivers through equation (8). Here we assume receivers are distant from the anomalous volume. The total cost of this step is $O(NL)$.

The final step is computation of p_b at the receivers. The cost of this step is $O(L)$.

The numerical complexity of computations using different integral formulations is summarized in Table 1. Usually the fourth row determines the overall computational load. However, when this step is not present for a selected method, or in case of large data and small models, the other steps may become of significance.

3. Accuracy of the forward modeling based on the QA, LQL, and Born approximations

In this section we compare accuracy of simulated fields, obtained with the QA, LQL, and Born approximations, with respect to a solution, computed with the full IE method. We will consider simple models, which comprise of a cubic scatterer located in a homogeneous half space. In all cases, shown below, the model setup is fully three-dimensional, with the point sources and receivers. We will study the dependency of the solution accuracy with respect to the simulation frequency, the scatter size and the velocity contrast. We will focus on lossless media, but also study a case of the lossy medium.

The first model, considered in this section is presented in Fig. 2. The model consisted of a cubical anomalous domain $3\times3\times3$ km^3 immersed into a half space. The domain was divided into 2,097,152 cubical cells (128 cells in each direction). The top of the cube was located at a depth of 1 km below the surface. The velocity distribution inside the anomalous volume mimicked salt tectonics (Fig. 2). The background velocity c_b was equal to 4000 m/s. The density was set to 2500 kg/m^3 everywhere below the surface.

In the first numerical experiment we study the accuracy to which the QA and LQL approximations predict the pressure field outside the anomalous domain, and compare these to the IE and Born methods. A point source was placed above the center of anomalous domain, 50 m below the Earth's surface. A receiver line went 50 m below the x-axis spanning x from 0 to 5 km. Fig. 3 shows total pressure field, p, computed along the receiver line at a number of frequencies.

Table 1
Numerical complexity of the different integral formulation methods.

	Step	QL		LQL		QA		Born	
		FSH	LH	FSH	LH	FSH	LH	FSH	LH
1	p_b, $\boldsymbol{r} \in D$	NK	NK	NK	NK	NK	NK	NK	NK
2	El-s of $\boldsymbol{G}$	N	NN_z	N	NN_z	N	NN_z		
3	R.h.s. or λ^{QA}	KQ	KW	Q	W	KQ	KW		
4	Sys. sol.	MKQ	MKW	MQ	MW				
5	p_a, $\boldsymbol{r} \notin D$	NL	NL	NL	NL	NL	NL	NL	NL
6	p_b, $\boldsymbol{r} \notin D$	L	L	L	L	L	L	L	L

Big-O notation is assumed.
FSH – free-space host model,
LH – layered host model,
K – number of the sources,
L – number of the receivers,
$N = N_x N_y N_z$ – number of model cells,
M – number of iterations (depends on the distribution of model parameters and frequency),
$Q = N\log(N)$ – cost of matrix multiplication for free-space,
$W = NN_z \log(N_x N_y)$ – cost of matrix multiplication for layered host medium.
The overall cost of modeling by a selected method for given K, L, Nx, Ny, Nz and the type of the host model is determined by the dominating step in a corresponding column.

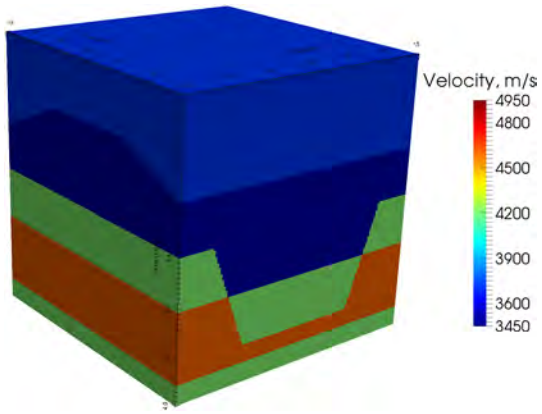

Fig. 2. Model used to test accuracy of different approximate solvers. The velocity distribution within the anomalous volume is shown. The length of the cube is 3 km, its top is placed 1 km below the surface.

We observe that, all approximate solutions work well at 0.1 Hz but start to deteriorate as frequency grows. The best results were provided by the QA approximation though at higher frequencies it is difficult to evaluate the curves visually. Surprisingly, the LQL approximation produced the poorest result out of the three except for the very low frequencies (0.1 and 0.5 Hz), where it outperformed the Born approximation.

In the second numerical experiment, we compared the total pressure response at 3 Hz inside the anomalous domain. As it will be discussed later, this configuration is important for gradient calculations because the gradient is computed inside the model. In this test we used the same source as before. Fig. 4 shows the real part of total pressure at a frequency of 3 Hz in a vertical slice $y = 0$. All approximations where compared to the full IE solution. The best result was provided by the QA approximation. This is due to the fact, that it preserves the phases better than the Born approximation does. The LQL approximation produced the least accurate response.

Numerical tests have revealed, that the visual degree of correspondence between different pressure responses alter considerably, depending on the simulation frequency and location of the observation line with respect to the body. To produce a more clear quantitative indication of the performance of each method, we considered the total field inside the body, shown in Fig. 5. The model consisted of a half space with a background velocity of $c_b = 2000$ m/s, in which a cube with $c_a = 1000$ m/s was embedded. The cube side was 1 km long, and the cube top was located at 1 km below the Earth's surface. The point source was placed above the cube center 50 m below the Earth's surface.

The simulation was carried out within the frequency range of 0.125 Hz to 20 Hz. The cell size was selected in such a way, that the number of cells in one direction was 2^m, where m is the smallest integer for which the number of cells per wavelength was not less than 10. The actual number of cells varied from 2,048 (low frequencies) to 1,048,576 (at 20 Hz). The relative error was computed by the formula:

$$\epsilon = \frac{\|p^{appr} - p^{IE}\|}{\|p^{IE}\|}, \tag{61}$$

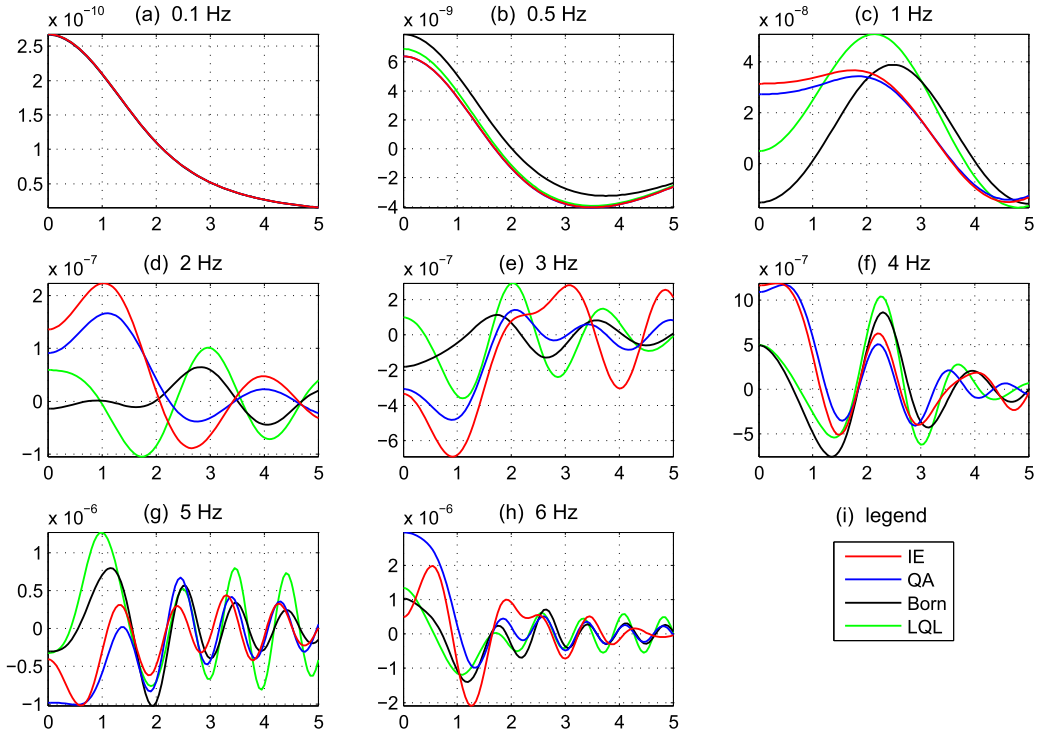

Fig. 3. Anomalous pressure response along a receiver line located 50 m below the x-axis. The real part of the total pressure field is shown. The x-axis is labeled in km. Panels (a)–(h) present plots produced at frequencies 0.1, 0.5, 1, 2, 3, 4, 5, and 6 Hz, respectively. The curves legend is shown in panel (i).

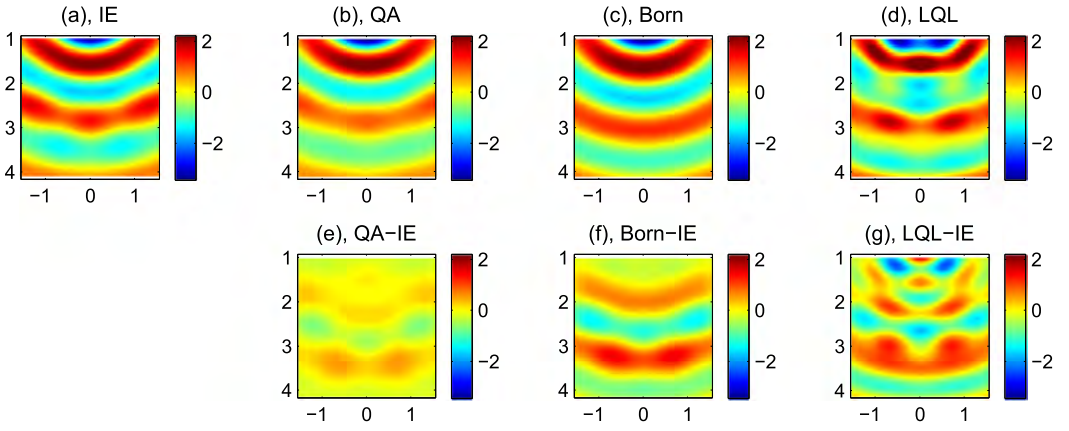

Fig. 4. The real part of anomalous pressure response at 3 Hz in a vertical plain $y = 0$. Panels (a)–(d) show solutions obtained with the IE, QA, Born, and LQL methods, respectively. Panels (e)–(f) show the difference between the QA, Born, and LQL solutions, respectively, and the IE solution. The x- and y-axes are labeled in kilometers. The pressure is given in Pascals multiplied by 10^5.

where $\|\cdot\|$ denotes the L^2-norm, p^{appr} is an approximate solution, p^{IE} is the full solution. The resulting curves are given in Fig. 6.

The field calculated by the QA approximation is closer to the IE curve than two other approximations. This becomes more pronounced at low frequencies, where the QA relative error becomes smaller than the Born one by two orders. The QA error curve comes closer to the Born curve toward the higher end of the frequency range, but always remains lower by 2–3 times. The LQL approximation has the worst accuracy, excluding extremely low frequencies. The zigzag pattern, clearly visible in the LQL curve at the low-frequency end, is linked to the discretization error, which varies according to the actual number of cells per wavelength.

The next numerical experiment was conducted to study dependency of the accuracy of solution on the size of the scatterer. It is well established, that the Born assumption is better satisfied within a scatterer with the limited extent. We might expect the same behavior for the QA approximation. In the following numerical experiment we set the frequency to a constant value of 5 Hz and the cell size to 31.25 m, whereas the scatterer size was varying. The cubical anomalous body

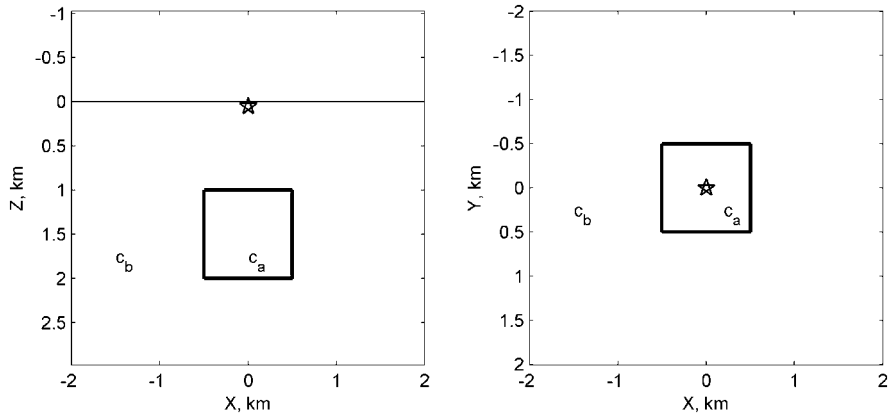

Fig. 5. Model used to test accuracy of different approximate solvers with respect to varying model parameters. The star depicts a position of the source.

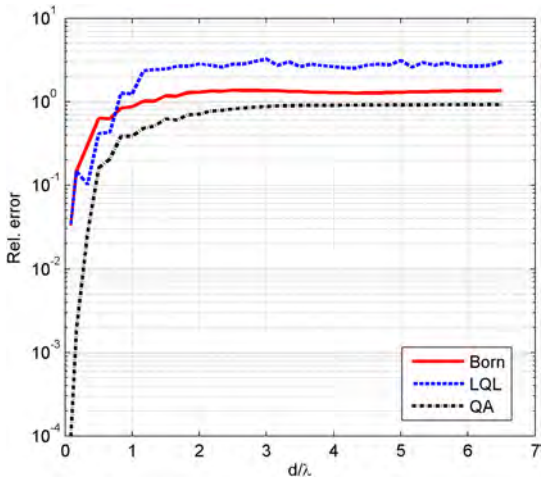

Fig. 6. Relative errors of the pressure response inside anomalous body for the frequency range of 0.125 Hz to 20 Hz, plotted as a function of cube side d divided by wavelength λ; $c_b = 2000$ m/s and $c_a = 1000$ m/s.

was placed at 1 km below the surface with its center being at the z-axis. The geometry of the simulation is described by Fig. 5, except that the cube side was changing from 31.25 m (one cell) to 2000 m (64 cells). The cube top was kept at the same depth of 1 km. The number of cells per wavelength was strictly equal to 10. The velocities were set to $c_b = 2000$ m/s, $c_a = 1000$ m/s. The resulting curves were plotted as a function of the cube side, d, divided by the wavelength within the scatterer, λ (Fig. 7).

These curves behave almost identically to those in Fig. 6, except for the absence of the zigzag pattern (this is because the number of cells per wavelength was not changing). As before, the QA approximation has the smallest errors, which increase with the scatterer size and exceeds 20% when this size is over 0.5 wavelength. Note that the errors of the Born approximation exceeds 20% when the size of the scatterer is just over 0.1 wavelength. For a small scatterer (less than 0.3 wavelength) the LQL approximation performs better than the Born approximation, but when the size exceeds 0.3 wavelength the LQL approximation fails.

The next numerical experiment was intended to study the accuracy of approximate solutions as the function of the velocity contrast. We used a series of models, each of them was identical to that shown in Fig. 5. The background velocity c_b was set to 4000 m/s and the anomalous velocity c_a was varying from -1990 m/s to 2010 m/s with step 200 m/s. Each model was discretized into 2^m cells, where m is the smallest integer for which the number of cells per wavelength is not less than 10. The error curves (Fig. 8) were computed by formula (61).

The most obvious feature in this plot is that the QA approximation has substantially lower error, comparing with the two other solutions, within the entire range of c_a. The difference between the QA curve and the two other approximation grows toward greater positive contrasts, and decreases toward greater negative contrasts. We attribute this behavior to the fact that, slowly oscillating fields make the QA localized condition (34) be satisfied better.

We note that, the area of tolerable precision for all approximations is quite narrow. Even 10 m/s contrast causes the error over 1% for the Born and LQL approximations. The error of the same order of magnitude for the QA solution develops,

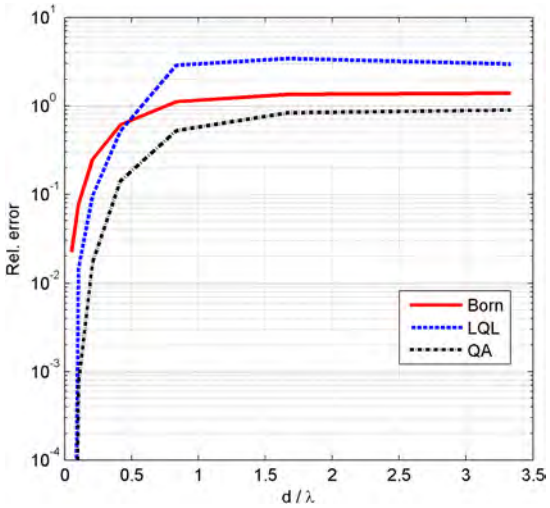

Fig. 7. Relative errors of pressure response at 5 Hz inside the cube of varying size, plotted as a function of cube side d divided by wavelength λ. The cube side d varied from 31.25 m to 2 km; $c_b = 2000$ m/s and $c_a = 1000$ m/s.

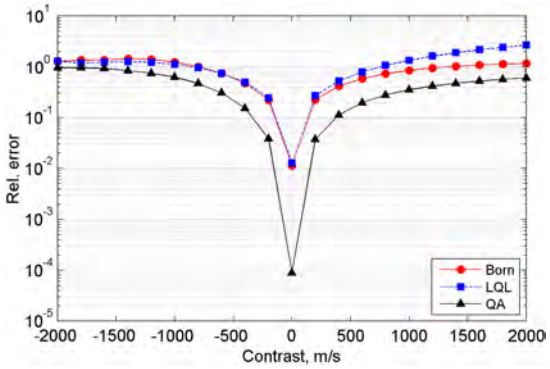

Fig. 8. Relative error of approximate solutions inside scatterer with respect to the IE solution, as a function of anomalous velocity c_a, with $c_b = 4000$ m/s and frequency 5 Hz.

when the absolute value of anomalous velocity exceeds 200 m/s. Thus, accuracy of the QA approximation is several times better, comparing to the Born and LQL approximations, though the error builds very quickly as model contrast grows.

The overall poor precision of the LQL solution, comes from the fact, that Green's function in acoustic medium falls slowly, whereas the background response, induced by the point source, varies quite rapidly. We have observed, that the LQL approximation works better for negative anomalous velocities ($c_a < 0$). We explain this by the fact that, if $c_a > 0$, than p_b oscillates more rapidly inside a model cell. Hence, assumption (40) totally breaks.

From the above results we may expect that, the approximate solutions work better in lossy media. In the following test we have assumed that, the host medium is a half space with a complex-valued wavenumber k_b. In this case, the amplitude of Green's function decays as follows:

$$\frac{e^{-Im(k_b)|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|},$$

provided $Im(k_b) \geq 0$. A faster decaying Green's function makes the integral operator be more localized. The geometry of the model in this test was exactly the same as shown in Fig. 5. We specified the complex-valued wavenumber k as follows:

$$k(c) = \frac{\omega}{c}(1 + 0.5i). \tag{62}$$

Note that, this is extremely strong attenuation compared to what is commonly encountered in seismology. We set c_b to 4000 m/s and ran a series of simulations with c_a varying from -1990 m/s to 2010 m/s with a step of 200 m/s (i.e. the same range, as used above when studying lossless media). The results are shown in Fig. 9. Comparing the resulting curves with Fig. 8 we observe that, the range of velocity contrasts c_a, in which accuracy of the QA approximation is below 1%, becomes at least twice wider than that of the lossless case.

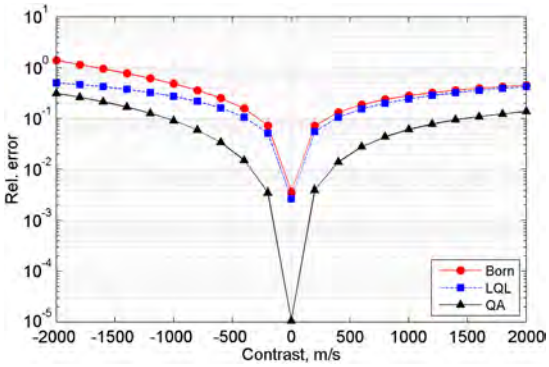

Fig. 9. Relative error of approximate solutions as a function of contract for $c_b = 4000$ m/s and frequency 5 Hz for viscous acoustic medium with $k(c) = (\omega/c)(1 + 0.5i)$.

The accuracy of the LQL and Born approximations have also improved, though they are still highly inaccurate when c_a exceeds several hundreds meter per second. It is interesting to note that, for this highly viscous model the LQL approximation outperforms the Born approximation, especially when $c_a < 0$. For a high negative anomalous velocity, c_a the LQL relative error curve approaches that of the QA approximation, but remains almost one order of magnitude below the relative error of the Born approximation. We attribute this to the fact that, for such a case, p_b oscillates less rapidly, and condition (40) is satisfied better.

The above results show that, the QA, LQL, and Born approximations are accurate within a relatively narrow range of the velocity contrasts.

The LQL solution is the least accurate out of the three approximations, except for the extremely low contrasts or frequencies. This is because the LQL condition (40) is not satisfied with tolerable accuracy for the point acoustic source. The condition holds somewhat better in a lossy medium for negative anomalous velocity $c_a < 0$. We explain this by the fact that, within a low-velocity domain the background Green's function oscillates less rapidly and decays somewhat faster. Hence, the integral operator becomes more localized. This is in good correspondence with the results, obtained in [33] for electromagnetic fields in lossy scatterers with a high-resistive background.

The QA solution is the most accurate out of the three approximations for all models we studied. When medium is viscous, the QA approximation is accurate for greater velocity perturbations. However, even for the QA approximation the error builds up quickly as the model contrast grows, which imposes significant limitations on using these approximations in an accurate seismic forward modeling.

At the same time the QA approximation can be effectively used in place of the Born approximation for computing the data gradient in the framework of the inverse problem solution. The computational time of generating the QA solution is determined by the time required for a single matrix–vector multiplication, whereas the full IE method requires many matrix–vector multiplications. For realistically big models, the cost of generating the QA solution becomes almost equal to the cost of the Born approximation. Thus the QA approximation is very promising for the solution of the acoustic inverse problems.

4. Full-waveform inversion based on the QA approximation

In the previous section we have demonstrated, that the QA approximation requires just a small computational overhead comparing to the Born approximation for large 3D models. Now we will show, that the data gradient, estimated through the QA approximation, is much more accurate, than that based on the Born approximation. Further, we apply it to solving full-waveform acoustic inverse problem using a gradient-type minimization algorithm.

Inversion codes for seismic large-scale 3D full-waveform inversion are frequently based on the conjugate-gradient or quasi-Newton minimization methods (for example, [25]). These minimization strategies require three key steps to be performed at every iteration of the inversion: 1) forward modeling; 2) gradient computation; and 3) estimation of the optimal step for the model update. The data gradient $\boldsymbol{\phi}$ is given by the following expression

$$\boldsymbol{\phi} = \boldsymbol{W}_d \boldsymbol{F}^* \boldsymbol{\psi}, \tag{63}$$

where $\boldsymbol{W}_d$ is, usually, a diagonal matrix of the data weights, $\boldsymbol{\psi}$ is the data residual,

$$\boldsymbol{\psi} = \boldsymbol{d}^{obs} - \boldsymbol{d}^{pr}, \tag{64}$$

$\boldsymbol{d}^{obs}$ is a vector of the observed data, and $\boldsymbol{d}^{pr}$ is a vector of the predicted data calculated for the model, obtained in the previous iteration. The third step is most commonly performed via a linear search, which requires computation of the product $\boldsymbol{Fq}$, where $\boldsymbol{q}$ is a searching direction. In the simplest case, the searching direction, $\boldsymbol{q}$, coincides with the direction, $\boldsymbol{\phi}$, of the data gradient. The cost of computing the products, $\boldsymbol{F}^*\boldsymbol{\psi}$ and $\boldsymbol{Fq}$, is roughly equal to one forward modeling operation.

Thus, the total cost per one inversion iteration is roughly equal to the solution of three forward modeling problems. It is known that, $\boldsymbol{F}^*\boldsymbol{\psi}$ and $\boldsymbol{Fq}$ can be computed with the use of an approximate solution, for example, the Born approximation. Indeed, the Born and Rytov approximations are found to be very effective tool for the gradient computation. Though inaccurate in amplitude, they deliver a decent estimation of the gradient direction and of the optimal step size. Therefore, the computational burden is effectively reduced to just one forward modeling operation.

In this paper, we use the QA approximation to compute the Fréchet derivative and the gradient directions. As we have discussed above, we apply a regularized conjugate gradient (RCG) method to minimize the Tikhonov parametric functional [24]. The RCG algorithm can be summarized as follows [29]:

$$\boldsymbol{r}^n = A_a(\boldsymbol{m}^n) - \boldsymbol{d}, \tag{65a}$$

$$\boldsymbol{u}^n = (\boldsymbol{F}_a^*)^n \boldsymbol{W_d}^2 \boldsymbol{r}^n \xi, \tag{65b}$$

$$\boldsymbol{g}^n = \boldsymbol{u}^n + \alpha^n (c_1^n \boldsymbol{W}_L^2 + c_2^n \boldsymbol{I})(\boldsymbol{m}^n - \boldsymbol{m}^0), \tag{65c}$$

$$\beta^n = \frac{\|\boldsymbol{g}^n\|^2}{\|\boldsymbol{g}^{n-1}\|^2}, \tag{65d}$$

$$\boldsymbol{p}^n = \boldsymbol{g}^n + \beta^n \boldsymbol{p}^{n-1}, \quad \boldsymbol{p}^1 = \boldsymbol{g}^1, \tag{65e}$$

$$k^n = \frac{(\boldsymbol{p}^n, \boldsymbol{g}^n)}{\|\boldsymbol{F}_a^n \boldsymbol{W}_d \boldsymbol{p}^n\|^2 + \alpha^n c_1^n \|\boldsymbol{W}_L \boldsymbol{p}^n\|^2 + \alpha^n c_2^n \|\boldsymbol{p}^n\|^2}, \tag{65f}$$

$$\boldsymbol{m}^n = \boldsymbol{m}^{n-1} - k^n \boldsymbol{p}^n, \tag{65g}$$

where $\boldsymbol{r}^n = \psi$ is a residual at the iteration step n; $\boldsymbol{u}^n$ is the gradient direction; ξ is a frequency-depending factor, which is commented below; $\boldsymbol{g}^n$ is the ascending direction of the parametric functional; $\boldsymbol{p}^n$ is the conjugate direction; $\boldsymbol{F}_a^n$ is the Fréchet derivative matrix of forward modeling operator A_a with respect to vector $\boldsymbol{m}$; $\boldsymbol{W}_d$ is a data weighing matrix with inverse of squared standard deviations on its diagonal, $\boldsymbol{W}_d = diag(1/\sigma_1..1/\sigma_J)$, where J is the number of data points; α^n is a regularization parameter; k^n is a length of the iteration step; and $\boldsymbol{W}_L$ is a discrete analog of ∇ operator. The weights c_1^n and c_2^n are selected at each iteration to balance the contribution from different parts of the stabilizer.

The progress of the inversion is controlled by the value of the normalized misfit, E,

$$E = \frac{\|\boldsymbol{W}_d \boldsymbol{\psi}\|}{\sqrt{J}}, \tag{66}$$

where $\boldsymbol{\psi}$ is the residual vector between observed and predicted data, J is the number of the data points. Value $E = 1$ means that the inversion model fits data to a specified misfit level defined by the data misfit. The algorithm terminates when the target misfit $E = 1$ is reached. If the data misfit E^n during the iteration n not decreases faster than a specified threshold, E_{tol}, i.e. when $|E^n - E^{n-1}|/E^n > E_{tol}$, this is considered as the algorithm stagnation. When it happens, a few attempts with a halved step size k^n are performed. If they have not succeeded, the algorithm terminates prematurely.

In order to prevent high frequencies from dominating in the data gradient in the context of multi-frequency inversion, we use factor $\xi = 1/\omega^4$. The matter of frequency scaling has already been addressed in literature [23,14], so we just briefly restate the main principles here. In our inverse code we assume lossless media, which means that, $\boldsymbol{F}_a = \omega^2 \boldsymbol{F}$. Under certain assumptions, the data residual $\boldsymbol{\psi}$ is proportional to ω^2 factor [23]. Hence, the data gradient, i.e. a product of the form $(\boldsymbol{F}_a)^*\boldsymbol{\psi}$, is proportional to ω^4, which needs to be compensated, when the data at several frequencies are inverted jointly.

The first numerical experiment in this section compares quality of the data gradient at a single frequency. This corresponds to the common practice, when several frequencies are inverted successively from the low to high frequencies. The data gradient, calculated for mono-frequency data, is formed by a narrow band of spatial frequencies and has a high rate of oscillations. Hence, the quality of the data gradient cannot be evaluated by comparing the data gradient at a single iteration with the true structure of the model. Instead, we have performed iterative minimization with two different methods for gradient calculation, but identical in other respects, and compared the final models.

In this test, we used the model shown in (Fig. 2). There was a smoothed version of this model, as depicted in Fig. 10. This model had been obtained by inversion of the data at low frequencies of 0.1 to 0.75 Hz.

The setup is shown in Fig. 11. There were 16 point sources, whose XY coordinates are determined by combinations of four abscissas (−2500 m, −1000 m, 1000 m, 2500 m) and four ordinates (−2500 m, −1000 m, 1000 m, 2500 m). Receivers were distributed on a regular XY grid inside the square $-5750 \le x \le -5750$, $-5750 \le y \le -5750$ with a step of 500 m. All sources and receivers were located at a depth of 50 m below the surface.

The synthetic observed data were simulated at a frequency of 1 Hz using the full IE method. Fig. 11 shows the real part of anomalous pressure response for the source with coordinates $x = y = -1000$ m. We inverted this data set with the minimization scheme described above. The regularization parameter was set to $\alpha = 10^{16}$, and E_{tol} was set to 0.005. The constants c_1^n and c_2^n were selected in such a way to keep the following ratio constant:

$$\frac{c_1^n \|\boldsymbol{W}_L(\boldsymbol{m}^n - \boldsymbol{m}^0)\|^2}{c_2^n \|\boldsymbol{m}^n - \boldsymbol{m}^0\|^2} = 3.$$

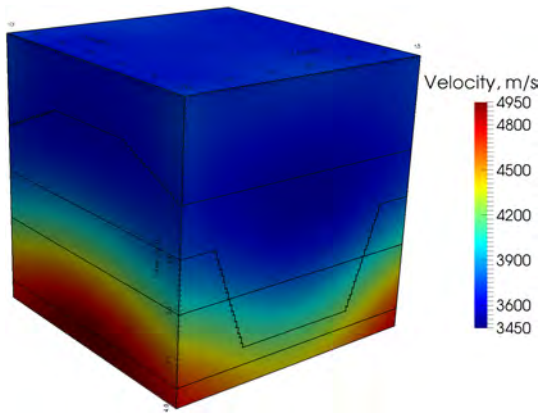

Fig. 10. Inversion test 1. A smoothed version of the velocity shown in Fig. 2, which was used as a starting model for the inversion of 1 Hz data. The model is obtained by the QA inversion of the lower frequencies.

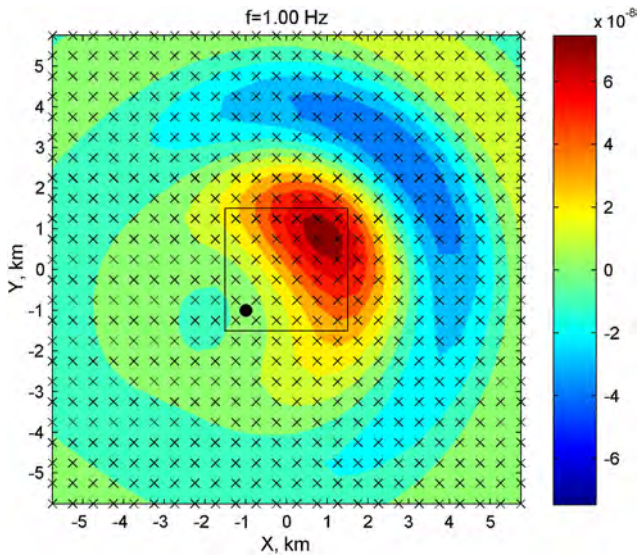

Fig. 11. Inversion test 1. Synthetic observed data at a frequency of 1 Hz in the plane $z = 50$ m for the source $x = y = -1000$ m. The real part of anomalous pressure is shown. Crosses mark the receiver's positions, black circle denotes the source position. A black square outlines the horizontal section of the anomalous volume.

The noise was specified as Gaussian with the standard deviation, σ_n, equal to 5% of the amplitude of the n-th data point. The inversion has started from the smoothed version of the true model (Fig. 10). During the inversion, the forward problem was solved by the full IE method.

Two inversion runs have been made. During the first run, Fréchet derivatives were estimated via the QA approximation. In the second run, Fréchet derivatives were computed through the Born approximation. The two convergence plots (Fig. 12) indicate, that the Born inversion prematurely stopped after 5 iterations with $E = 4.97$, while the QA inversion stopped at iteration 23 with $E = 0.9$, i.e. has reached the target misfit.

Examination of the final models (Fig. 13) confirms, that the QA inversion resulted in a more appropriate models, than the Born inversion did.

In the next numerical experiment we compared the QA and Born gradients by inverting several frequencies together. This is a less common strategy, because it implies larger computational burden and memory consumption, than in the previous test. At the same time, wide-band data recover a range of spatial frequencies of the data gradient, which allows us to make a direct comparison of the gradient with the true model at every iteration of the inversion. This is in contrast with the previous numerical experiment, where the conclusions were derived by inspection of the final models at the end of the iterative process, and might be affected by an interaction with the stabilizing part of the target functional and with other aspects of the minimization strategy.

In this test, we used a more complicated model and a far larger data set than in the previous test. The model consisted of an anomalous rectangular hexahedral domain with a complicated velocity distribution immersed into a homogeneous half space with $c_b = 1500$ m/s. The velocity inside the anomalous domain did not change in the y-direction, except of a

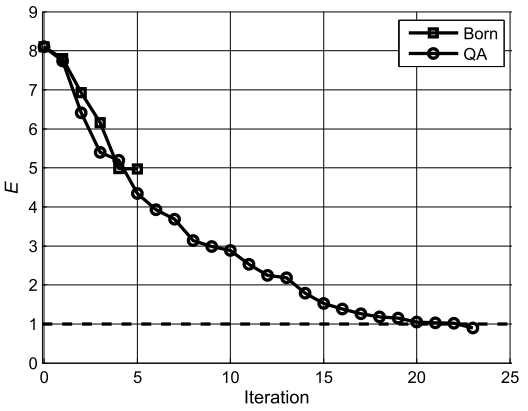

Fig. 12. Inversion test 1. Normalized misfit, E, versus the iteration number during the inversion of the data at a frequency of 1 Hz when the Fréchet derivatives were calculated with the Born (squares) and QA (circles) approximations, respectively.

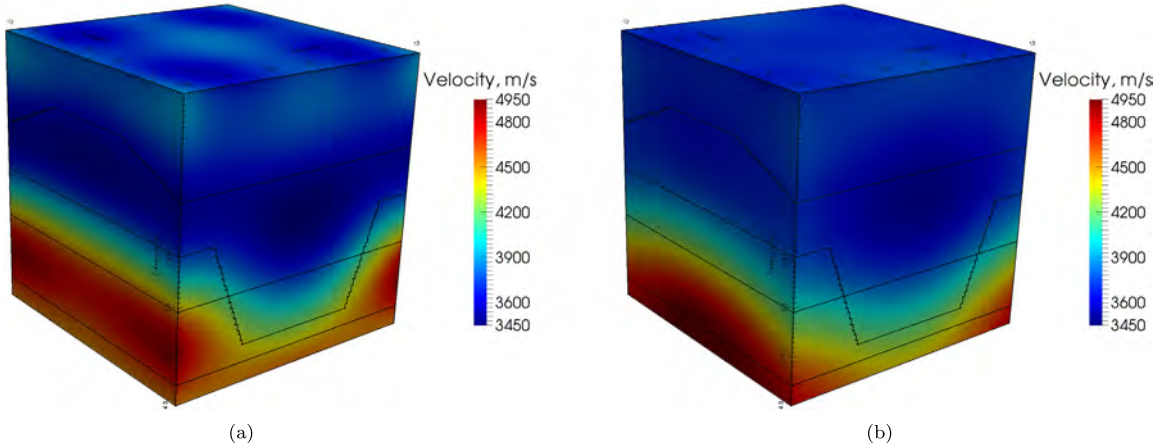

Fig. 13. Inversion test 1. Panel (a) shows the model obtained by the inversion with the QA-based gradient calculations. Panel (b) presents the model obtained by the inversion with the Born-based gradient calculations. The black lines are the boundaries within the true model.

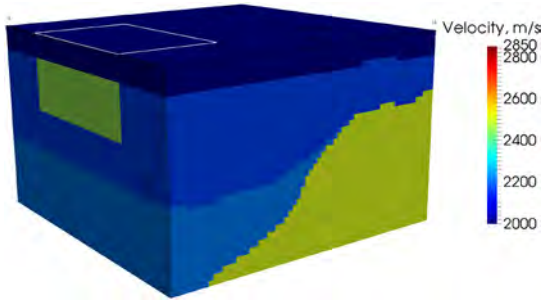

Fig. 14. Inversion test 2. Velocity distribution within the anomalous volume $3 \times 3 \times 1.875$ km^3. The white rectangle marks the projection of the shallow high-velocity prism on the top of the model.

high-velocity rectangular prism on the left side of the model (Fig. 14). The density was set to 2500 kg/m^3 everywhere below the surface. The model was split into 163,840 cubical cells of $46.875 \times 46.875 \times 46.875$ m^3 each. The center of the horizontal projection of the anomalous domain was located at $x = 0$, $y = 0$; its top side was placed 200 m below the surface.

There were 144 sources, uniformly distributed inside the XY square, $-2750 \leq x \leq 2750$, $-2750 \leq y \leq 2750$, with a separation of 500 m, and 841 receivers uniformly distributed in the XY square, $-7000 \leq x \leq 7000$, $-7000 \leq y \leq 7000$, with a separation of 500 m. The synthetic observed data were simulated using the IE method at ten frequencies from 0.1 Hz to 1 Hz with a step of 0.1 Hz. The spectrum of the signal in the source was formed by the Ricker wavelet with the central

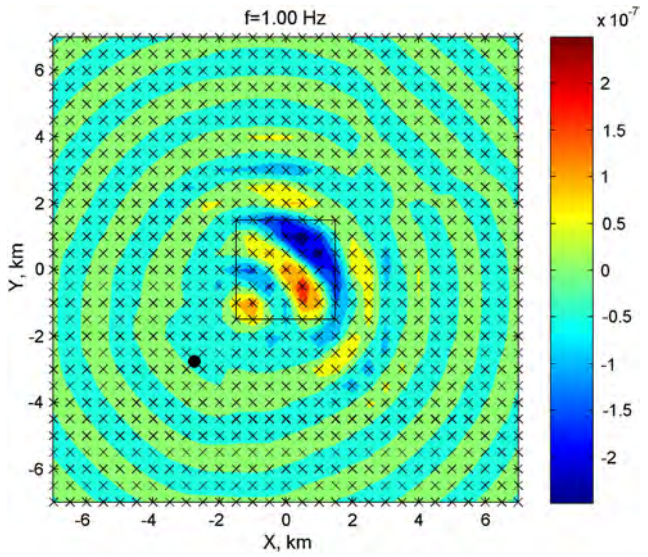

Fig. 15. Inversion test 2. Synthetic data at a frequency of 1 Hz in the plane $z = 50$ m. The real part of the anomalous pressure is shown. Crosses mark the receiver's positions, black circles denote the source's positions. A black square outlines the boundary of the anomalous volume.

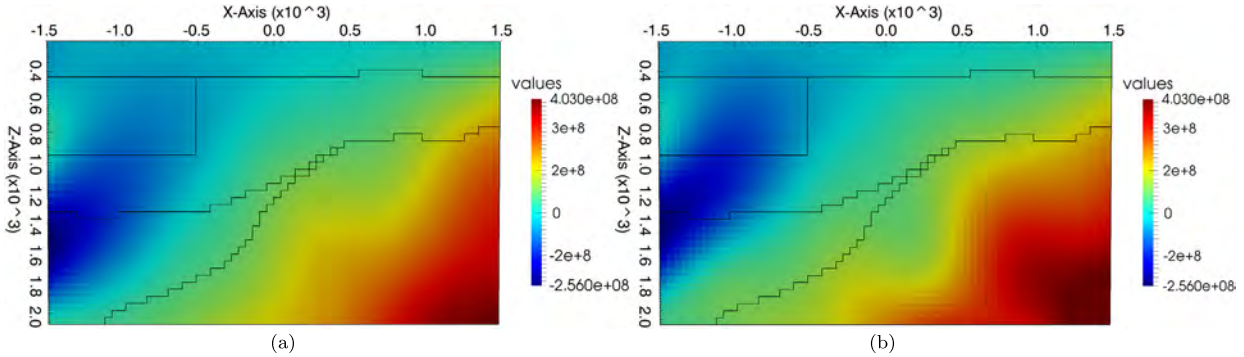

Fig. 16. Inversion test 2. Vertical section of the data gradient inside the anomalous domain in a vertical plain $y = 0$. Panel (a) shows the results obtained with the QA approximation, while panel (b) presents the results of the Born approximation. The black lines mark the boundaries inside the true model.

frequency of 0.5 Hz. As above, the noise standard of noise, σ_n, was set to 5% of the amplitude of the n-th data values. The real part of the anomalous pressure at 1 Hz for the source with coordinates $x = y = -2750$ m is shown in Fig. 15.

We applied the same iterative RCG method of inversion as described above in equations (65). The constants c_1, c_2, and the threshold E_{tol} were selected exactly as it is described above. The regularization parameter α, was selected based on the condition that ratio $\|S\|/\|\Phi\|$ in (43) was equal to 10^{-4} at the beginning of each iteration. The predicted pressure field was calculated by the IE method, while the Fréchet derivatives were computed using the corresponding approximation. The iterative process started with a constant velocity inside the anomalous volume, $c = 2300$ m/s.

Fig. 16 shows the vertical sections of the data gradients, generated at the first iteration by the QA and Born approximations, respectively. The QA data gradient has much more resemblance with the structure of the subsurface, than that of the Born approximation, though both gradients have many artifacts. It might be expected that the QA inversion would converges to a more appropriate model.

Fig. 17 presents the convergence plots of the inversion based on the QA and the Born approximation, respectively. The Born inversion stagnates after iteration 12 with $E = 5.7$, whereas the QA inversion stagnated at iteration 18 with $E = 2.63$. The Born inversion ended up with a low-contrast model (Fig. 18b), which has many artifacts. The QA inversion resulted in a better model (Fig. 18a), in which the main feature (the high-velocity block at the bottom-right side of the model) is clearly visible. Both algorithms were unable to adequately reconstruct the high-velocity prism, which is presented within the original model in the depth range 0.4–0.9 km. Probably, this is due to inadequate resolution of the low-frequency input data.

The seismic full-waveform inversion has the property of strong non-linearity, and the choice of the initial model is critical to avoid local minima. A homogeneous initial model is usually the worst choice. Indeed, the algorithm trapped to local minima in the last example. In real surveys this problem is mitigated by using an appropriate initial model that based on various forms of data processing. The best starting models frequently come from seismic tomographic inversion. In order

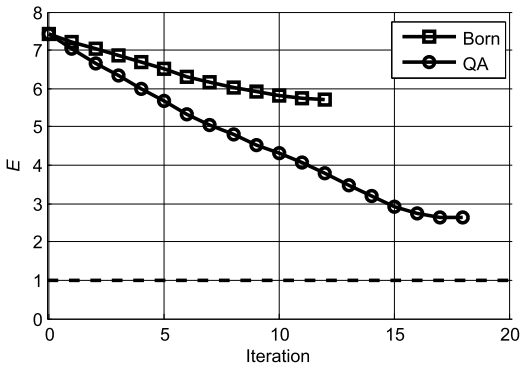

Fig. 17. Inversion test 2. Normalized misfit, E, versus the iteration number during inversion of the multi-frequency data when Fréchet derivatives were calculated with the Born (squares) and QA(circles) approximations, respectively.

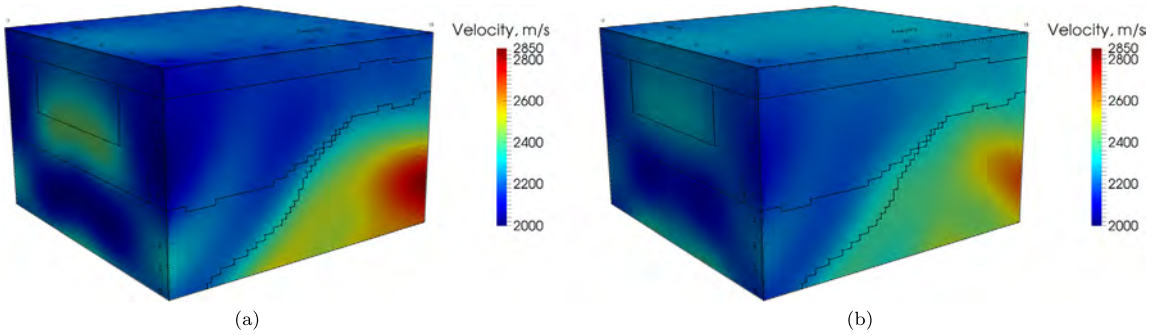

Fig. 18. Inversion test 2. Final velocity models obtained by the inversion of multi-frequency data. Panel (a) shows the QA-based inverse model. Panel (b) shows the Born-based inverse model. The black lines mark the boundaries inside the true model.

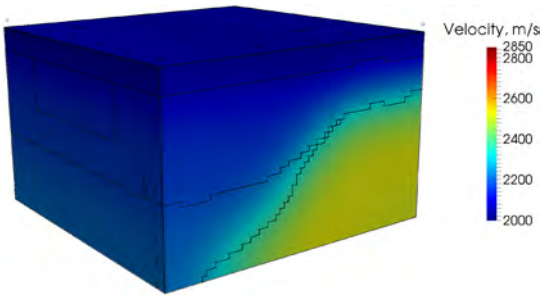

Fig. 19. Inversion test 3. A smooth starting model. This model is obtained from model shown in Fig. 14 by removing the high-velocity prism and subsequent smoothing the velocity distribution. The back lines mark the boundaries inside the true model.

to evaluate this possibility we took the model, shown in Fig. 14, replaced the high-velocity prism (2450 m/s) with the surrounding medium (2150 m/s), and smoothed the model. The rationale is that seismic tomography will likely generate a correct image of the top of the basement, which produces strong reflections, but a compact high-velocity scatterer poses a difficulty for this method. We employed a spline smoothing algorithm described in [12]. The new starting model is presented in Fig. 19.

The inversion parameters and input data were exactly the same as before. The Born inversion stagnated after iteration 2 with $E = 3.2$. The QA inversion almost reached the target misfit stagnating with $E = 1.19$ at iteration 32 (Fig. 20).

Fig. 21 shows the vertical sections of the data gradients, generated at the first iteration by the QA and Born approximations, respectively. As in the previous example, the Born-based gradient is less accurate, though in both cases the high-velocity prism was imaged only roughly.

The QA inversion resulted in a better model (Fig. 22), in which the main features of the true model (the high-velocity block at the bottom-right side and the shallow high-velocity prism) are clearly visible. The quality of the QA model substantially improved with respect to the previous numerical experiment. In the Born-based model the prism footprint is almost invisible. It is interesting to note that, in the previous experiment the Born model had an image of the prism, despite other

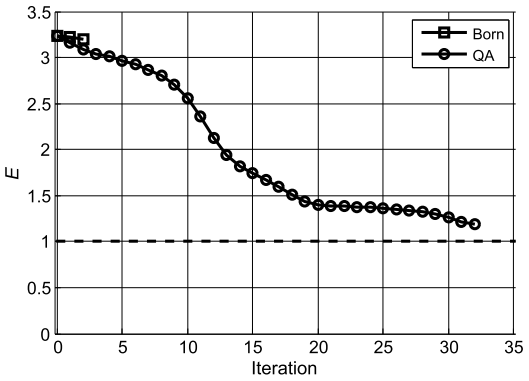

Fig. 20. Inversion test 3. Normalized misfit, E, versus the iteration number during inversion of the multi-frequency data when Fréchet derivatives were calculated with the Born (squares) and QA (circles) approximations, respectively.

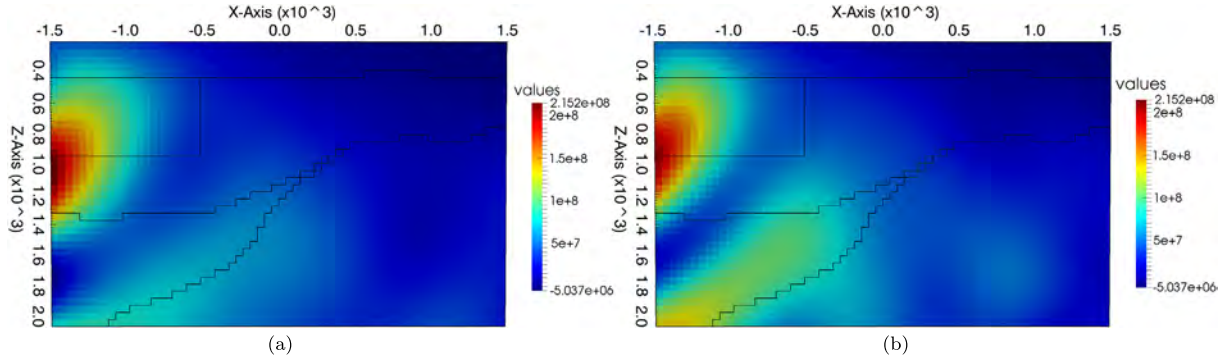

Fig. 21. Inversion test 3. Vertical section of the data gradient inside the anomalous body in a vertical plain $y = 0$. Panel (a) shows the calculations performed with the QA approximation, while panel (b) presents results of the Born approximation. The black lines mark the boundaries inside the true model.

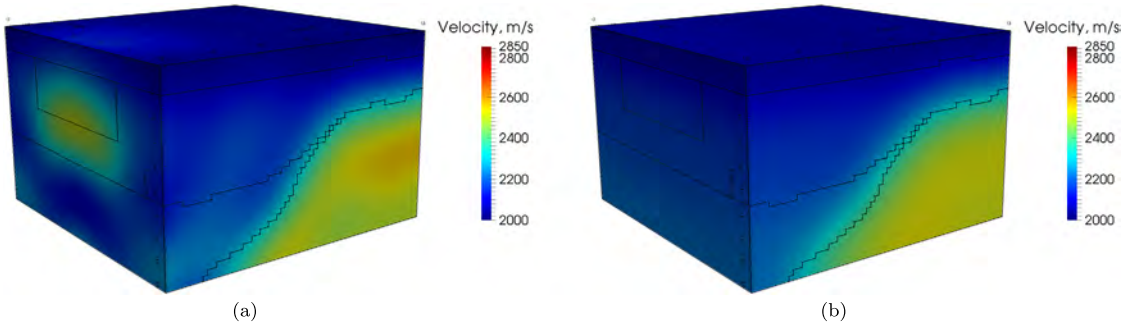

Fig. 22. Inversion test 3. Final velocity model obtained by the inversion of multi-frequency data. Panel (a) shows the QA-based inverse model. Panel (b) shows the Born-based inverse model. The black lines mark the boundaries inside the true model.

parts of the model being greatly distorted, comparing to Fig. 22b. This is because in this example the Born inversion has trapped into a local minimum at the start, and the final model deviated from the initial one only slightly.

5. Conclusions

In this paper, we have examined several approximate solutions of Lippmann–Schwinger equation, which can be treated as a generalization of the classical Born approximation, including the localized quasi-linear (LQL) and quasi-analytical (QA) approximations. We have demonstrated that, the QA approximation is more accurate, than the Born approximation. In contrast to the Born approximation, it accounts for the internal field inside the scatterer. We have also found that the range of applicability of the QA approximation is wider than that of the Born approximation, since it does not necessary require that the size of the scatterer and the velocity contrast will be small, but that the coefficient $\lambda(\boldsymbol{r})$ (relating the anomalous and background fields) will vary relatively slowly comparing to the Green's function. The advantage of the LQL approximation is that it is independent of the background field and needs to be computed once for multiple seismic sources, which extremely

attractive for both forward modeling and inversion. All approximations, Born, QA and LQL, perform better in lossy media than in the lossless case.

In summary, we have found that the QA approximation is the most accurate out of the three approximations for all models we studied; however, even for the QA approximation the accuracy decreases with the increase of the model contrast.

At the same time, we have demonstrated that, the QA approximation can be effectively used in the inversion algorithm for the Fréchet derivative (gradient) calculations, where the "direction" of the gradient is important, but an accurate estimate of the gradient's magnitude is not critical. We have developed an effective algorithm and computer code for full-waveform inversion based on the QA approximation and tested it on several complex models.

We have also demonstrated that the QA approach can be extended to a more general case of elastic media.

Thus, we conclude that, as indicated by the results presented, the practical full-waveform inversion algorithms can be designed with the use of the QA approximation for the data gradient computation. Its main application are the models of moderate to large sizes, up to several million cells, when a rigorous forward modeling, necessary for the use of an adjoint operator method, may become computationally too expansive. The use of the QA approach to compute the Fréchet derivative (gradient) within the inversion procedure reduces computational time greatly, comparing to the rigorous solution.

Acknowledgements

This research was supported by the Russian Science Foundation, project No. 16-11-10188. The authors also acknowledge support from the University of Utah's Consortium for Electromagnetic Modeling and Inversion (CEMI) and TechnoImaging. We are thankful to the Supercomputing Center of Lomonosov Moscow State University for providing computational facility. The authors gratefully acknowledge the computing time granted by Federal Center for Collective Usage at NRC Kurchatov Institute.

Appendix A. The quasi-analytical approximation in the elastic case

In this section we discuss an extension of the QA approximation to the elastic case. We start from the displacement formulation of the propagation of elastic wave in an isotropic inhomogeneous solid, as follows [29,19]:

$$-\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial u_k}{\partial x_k}\right)-\frac{\partial}{\partial x_j}\left(\mu\frac{\partial u_i}{\partial x_j}\right)-\frac{\partial}{\partial x_j}\left(\mu\frac{\partial u_j}{\partial x_i}\right)-\rho\omega^2 u_i=f_i, \tag{A.1}$$

where $\boldsymbol{u}$ is the displacement vector, λ and μ are the Lamé parameters, ρ is density, $\boldsymbol{f}$ is the unit source force. For simplicity we assume constant density but spatially varying λ and μ. Equation (A.1) is supplemented with the radiation boundary conditions.

Let us split the Lamé parameters into a background and anomalous parts, as follows:

$$\lambda=\lambda_b+\lambda_a,\quad \mu=\mu_b+\mu_a. \tag{A.2}$$

We introduce two differential operators, $\mathcal{L}=(\mathcal{L}_1,\mathcal{L}_2,\mathcal{L}_3)$ and $\mathcal{L}_b=(\mathcal{L}_{b1},\mathcal{L}_{b2},\mathcal{L}_{b3})$, as follows:

$$\mathcal{L}_i\boldsymbol{\xi}=-\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial \xi_k}{\partial x_k}\right)-\frac{\partial}{\partial x_j}\left(\mu\frac{\partial \xi_i}{\partial x_j}\right)-\frac{\partial}{\partial x_j}\left(\mu\frac{\partial \xi_j}{\partial x_i}\right), \tag{A.3}$$

$$\mathcal{L}_{bi}\boldsymbol{\xi}=-\frac{\partial}{\partial x_i}\left(\lambda_b\frac{\partial \xi_k}{\partial x_k}\right)-\frac{\partial}{\partial x_j}\left(\mu_b\frac{\partial \xi_i}{\partial x_j}\right)-\frac{\partial}{\partial x_j}\left(\mu_b\frac{\partial \xi_j}{\partial x_i}\right), \tag{A.4}$$

where $\boldsymbol{\xi}$ is an arbitrary vector field. With these definitions, $\boldsymbol{u}$ and $\boldsymbol{u}_b$ satisfy to the following equations:

$$-\mathcal{L}\boldsymbol{u}-\rho\omega^2\boldsymbol{u}=\boldsymbol{f}, \tag{A.5}$$

$$-\mathcal{L}_b\boldsymbol{u}_b-\rho\omega^2\boldsymbol{u}_b=\boldsymbol{f}, \tag{A.6}$$

supplemented with the radiation boundary conditions. Subtracting equation (A.6) from (A.5) and rearranging terms, we arrive to the following equation for he anomalous displacement, $\boldsymbol{u}_a$, as follows:

$$-\mathcal{L}_b\boldsymbol{u}_a-\rho\omega^2\boldsymbol{u}_a=\mathcal{L}_a(\boldsymbol{u}_b+\boldsymbol{u}_a), \tag{A.7}$$

where $\mathcal{L}_a=\mathcal{L}-\mathcal{L}_b$. The solution of (A.7) is given by the following integral:

$$\boldsymbol{u}_a(\mathbf{r})=\int_D \widehat{g}(\mathbf{r}|\mathbf{r}')\cdot\mathcal{L}_a\left(\boldsymbol{u}_b(\mathbf{r}')+\boldsymbol{u}_a(\mathbf{r}')\right)d^3\mathbf{r}', \tag{A.8}$$

where $\widehat{g}$ is the displacement Green's tensor of the second order. Derivation of $\widehat{g}$ can be difficult for arbitrary background medium. In the case of a homogeneous unbounded elastic medium there are closed expressions (for example, see [20]).

By neglecting $\boldsymbol{u}_a$ in (A.8) we obtain the Born approximation:

$$\boldsymbol{u}^B = \int_D \widehat{g}(\boldsymbol{r}|\boldsymbol{r}') \cdot \mathcal{L}_a \boldsymbol{u}_b(\boldsymbol{r}') d^3\boldsymbol{r}'. \tag{A.9}$$

Within framework of the QA approximation, following [29], we introduce a scalar field, λ^{QA}, that links the anomalous and background fields:

$$\boldsymbol{u}_a^{QA} \approx \lambda^{QA} \boldsymbol{u}_b. \tag{A.10}$$

This equality is approximate, because in general $\boldsymbol{u}_a^{QA}$ is related to $\boldsymbol{u}_b$ by a tensor field.

From (A.8) and (A.10) it follows that

$$\boldsymbol{u}_a^{QA} = \boldsymbol{u}^B + \int_D \widehat{g}(\boldsymbol{r}|\boldsymbol{r}') \cdot \mathcal{L}_a(\lambda^{QA}\boldsymbol{u}_b)\, d^3\boldsymbol{r}'. \tag{A.11}$$

Assuming that the main contribution to the integral comes from the neighborhood of the point $\boldsymbol{r} = \boldsymbol{r}'$ and that λ^{QA} is slowly varying, we can write:

$$\boldsymbol{u}_a^{QA} \approx \boldsymbol{u}^B + \lambda^{QA} \int_D \widehat{g}(\boldsymbol{r}|\boldsymbol{r}') \cdot \mathcal{L}_a \boldsymbol{u}_b\, d^3\boldsymbol{r}'. \tag{A.12}$$

Substituting equation (A.10) to (A.12) we obtain the following relation

$$\lambda^{QA}\boldsymbol{u}_b = \boldsymbol{u}^B + \lambda^{QA}\boldsymbol{u}^B. \tag{A.13}$$

By moving $\lambda^{QA}\boldsymbol{u}^B$ to the left side and multiplying both side by $\boldsymbol{u}_b$, we can determine λ^{QA} explicitly as follows:

$$\lambda^{QA} = \frac{\boldsymbol{u}^B \cdot \boldsymbol{u}_b}{(\boldsymbol{u}_b - \boldsymbol{u}^B) \cdot \boldsymbol{u}_b}, \tag{A.14}$$

assuming that the denominator does not vanish. Formulas (A.14) and (A.11) give the solution for the displacement vector in the QA approximation.

References

[1] A. Abubakar, T. Habashy, Three-dimensional visco-acoustic modeling using a renormalized integral equation iterative solver, J. Comput. Phys. 249 (2013) 1–12.
[2] A. Abubakar, P. Van Den Berg, J. Fokkema, Towards non-linear inversion for characterization of time-lapse phenomena through numerical modelling, Geophys. Prospect. 51 (4) (2003) 285–293.
[3] K. Aki, P.G. Richards, Quantitative Seismology, W.R. Freeman and Co., 1980.
[4] L. Amundsen, A.B. Weglein, A. Reitan, On seismic deghosting using integral representation for the wave equation: use of Green's functions with Neumann or Dirichlet boundary conditions, Geophysics 78 (4) (2013) T89–T98.
[5] D.B. Avdeev, Three-dimensional electromagnetic modeling and inversion from theory to application, Surv. Geophys. 26 (2005) 767–799.
[6] N. Bleistein, Mathematical Methods for Wave Phenomena, Academic Press, 1984.
[7] N. Bleistein, J.K. Cohen, J.W. Stockwell, Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion, Interdiscip. Appl. Math., vol. 13, Springer, New York, Berlin, Paris, 2001.
[8] J.K. Cohen, F.G. Hagin, N. Bleistein, Three-dimensional Born inversion with an arbitrary reference, Geophysics 51 (8) (1986) 1552–1558.
[9] H. Freter, An integral equation method for seismic modelling and inversion, in: A. Vogel, A. Sarwar, R. Gorenflo, O. Kounchev (Eds.), Theory and Practice of Geophysical Data Inversion, in: Theory Pract. Appl. Geophys., vol. 5, Vieweg+Teubner Verlag, 1992, pp. 243–264.
[10] L.-Y. Fu, Numerical study of generalized Lippmann–Schwinger integral equation including surface topography, Geophysics 68 (2) (2003) 665–671.
[11] L.-Y. Fu, Y.-G. Mu, H.-J. Yang, Forward problem of nonlinear Fredholm integral equation in reference medium via velocity-weighted wavefield function, Geophysics 62 (2) (1997) 650–656.
[12] D. Garcia, Robust smoothing of gridded data in one and higher dimensions with missing values, Comput. Stat. Data Anal. 54 (2009) 1167–1178.
[13] T.M. Habashy, R.W. Groom, B.R. Spies, Beyond the Born and Rytov approximations: a nonlinear approach to electromagnetic scattering, J. Geophys. Res., Solid Earth 98 (B2) (1993) 1759–1775.
[14] W. Hu, A. Abubakar, T.M. Habashy, Simultaneous multifrequency inversion of full-waveform seismic data, Geophysics 74 (2) (2009) R1–R14.
[15] L.-J. Huang, C.M. Fehler, Quasi-born Fourier migration, Geophys. J. Int. 140 (3) (2000) 521–534.
[16] M. Jakobsen, B. Ursin, Full waveform inversion in the frequency domain using direct iterative t-matrix methods, J. Geophys. Eng. 12 (3) (2015) 400.
[17] S. Johnson, Y. Zhou, M. Berggren, M. Tracy, Acoustic inverse scattering solutions by moment methods and back propagation, in: Theory and Applications, SIAM, 1983, pp. 144–155.
[18] P.M. Morse, H. Feshbach, Methods in Theoretical Physics, McGraw–Hill Book Company, Inc., 1953.
[19] P. Moszo, J. Kristek, M. Galis, The Finite-Difference Modelling of Earthquake Motions, vol. 365, Cambridge University Press, 2014.
[20] Y.H. Pao, V. Varatharajulu, Huygens' principle, radiation conditions, and integral formulas for the scattering of elastic waves, J. Acoust. Soc. Am. 59 (1976) 1361–1371.
[21] R.-E. Plessix, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, Geophys. J. Int. 167 (2) (2006) 495–503.
[22] H. Schmidt, F.B. Jensen, A full wave solution for propagation in multilayered viscoelastic media with application to Gaussian beam reflection at fluid–solid interfaces, J. Acoust. Soc. Am. 77 (3) (1985) 813–825.

[23] L. Sirgue, R.G. Pratt, Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, Geophysics 69 (1) (2004) 231–248.
[24] A.N. Tikhonov, V.Y. Arsenin, Solution of Ill-Posed Problems, H. Winston and Sons, 1977.
[25] J. Virieux, S. Operto, An overview of full-waveform inversion in exploration geophysics, Geophysics 74 (6) (2009).
[26] R.-S. Wu, M.N. Toksoz, Diffraction tomography and multisource holography applied to seismic imaging, Geophysics 52 (1) (1987) 11–25.
[27] R. Zhang, T.J. Ulrych, Seismic forward modeling by integral equation and some practical considerations, in: SEG, 2000, pp. 2329–2332, Chapter 593.
[28] M. Zhdanov, G. Hursan, 3d electromagnetic inversion based on quasi-analytical approximation, Inverse Probl. 16 (2000) 1297.
[29] M.S. Zhdanov, Geophysical Inverse Theory and Regularization Problems, Elsevier, San Diego, CA, 2002.
[30] M.S. Zhdanov, Inverse Theory and Applications in Geophysics, 2nd edition, Elsevier, 2015.
[31] M.S. Zhdanov, V.I. Dmitriev, S. Fang, G. Hursan, Quasi-analytical approximations and series in electromagnetic modeling, Geophysics 65 (6) (2000) 1746–1757.
[32] M.S. Zhdanov, S. Fang, Quasi-linear approximation in 3d electromagnetic modeling, Geophysics 61 (1996) 646–665.
[33] M.S. Zhdanov, E. Tartaras, Three-dimensional inversion of multitransmitter electromagnetic data based on the localized quasi-linear approximation, Geophys. J. Int. 148 (2002) 506–519.