



DANIEL LEAL MACEDO

**SCATTERING-BASED DECOMPOSITION OF  
SENSITIVITY KERNELS OF ACOUSTIC FULL  
WAVEFORM INVERSION**

**DECOMPOSIÇÃO BASEADA EM TEORIA DE  
ESPALHAMENTO DOS NÚCLEOS DE  
SENSIBILIDADES DA INVERSÃO DE ONDA  
COMPLETA ACÚSTICA**

CAMPINAS, SP  
2014







UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA MECÂNICA  
E INSTITUTO DE GEOCIÊNCIAS

DANIEL LEAL MACEDO

**SCATTERING-BASED DECOMPOSITION OF SENSITIVITY OF  
ACOUSTIC FULL WAVEFORM INVERSION**

**DECOMPOSIÇÃO BASEADA EM TEORIA DE ESPALHAMENTO DOS  
NÚCLEOS DE SENSIBILIDADES DA INVERSÃO DE ONDA  
COMPLETA ACÚSTICA**

Thesis presented to the Mechanical Engineering Faculty and Geosciences Institute of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Petroleum Sciences and Engineering in the area of Reservoirs and Management.

Tese apresentada à Faculdade de Engenharia Mecânica e Instituto de Geociências da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Ciências e Engenharia de Petróleo na área de Reservatórios e Gestão.

**Orientador: Prof. Dr. Joerg Dietrich Wilhelm Schleicher**

Este exemplar corresponde à versão final da tese defendida pelo aluno Daniel Leal Macedo, e orientada pelo Prof. Dr. Joerg Dietrich Wilhelm Schleicher.

CAMPINAS, SP  
2014



Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Área de Engenharia e Arquitetura  
Luciana Pietrosanto Milla - CRB 8/8129

M151s      Macedo, Daniel Leal, 1975-  
Scattering-based decomposition of sensitivity kernels of acoustic full waveform inversion / Daniel Leal Macedo. – Campinas, SP : [s.n.], 2014.

Orientador: Joerg Dietrich Wilhelm Schleicher.  
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica e Instituto de Geociências.

1. Propagação de ondas. 2. Ondas sísmicas. 3. Inversão (Geofísica). I. Schleicher, Joerg Dietrich Wilhelm, 1964-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Mecânica. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Decomposição baseada em teoria de espalhamento dos núcleos de sensibilidade da inversão de onda completa acústica

**Palavras-chave em inglês:**

Wave propagation

Seismic waves

Inversion (Geophysics)

**Área de concentração:** Reservatórios e Gestão

**Titulação:** Doutor em Ciências e Engenharia de Petróleo

**Banca examinadora:**

Joerg Dietrich Wilhelm Schleicher [Orientador]

Jesse Carvalho Costa

Ricardo Caetano Azevedo Biloti

Claudio Guerra Cardoso

Djalma Manoel Soares Filho

**Data de defesa:** 29-08-2014

**Programa de Pós-Graduação:** Ciências e Engenharia de Petróleo





UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA MECÂNICA  
E INSTITUTO DE GEOCIÊNCIAS

TESE DE DOUTORADO

SCATTERING-BASED DECOMPOSITION OF THE  
SENSITIVITY KERNELS FOR ACOUSTIC FULL  
WAVEFORM INVERSION

Autor: Daniel Leal Macedo

Orientador: Prof. Dr. Joerg Dietrich Wilhelm Schleicher

A banca examinadora composta pelos membros abaixo aprovou esta tese:

Prof. Dr. Joerg Dietrich Wilhelm Schleicher, Presidente  
DMA/IMECC/UNICAMP

Prof. Dr. Ricardo Caetano Azevedo Biloti  
DMA/IMECC/UNICAMP

Prof. Dr. Jessé Carvalho Costa  
IG/UFGA

Dr. Claudio Guerra Cardoso  
PETROBRAS/RJ

Dr. Djalma Manoel Soares Filho  
PETROBRAS/RJ

Campinas, 29 de agosto de 2014.



---



# Acknowledgements

It is said that the path is far more important than the final destination. This summarises perfectly my feelings along these years dedicated to this thesis. Now, it is the time to acknowledge and thank some important people and institutions that helped me all along this journey.

My beloved parents, Carlos and Hélia, I thank you with all my heart and soul. Without your help, support and love I could not have done it. I love you so much.

My sister Karla, and brother Eduardo, I thank you for the unshakeable belief in my potential, which helped me to face those few moments when dark clouds prevented me from seeing the clear sky. I love you.

Luciana e my daughter Clara, I thank you for all the happy moments I shared with you. You being part of my life is one of the most valuable gifts life gave me during my PhD. You are going to be part of me forever.

I would like to thank my advisor, Joerg Schleicher, for the unwearying disposition to help, discuss and collaborate with me since before the PhD itself.

I would like to thank Ivan Vasconcelos for opening many doors on my professional life; for being, together with Joerg, the corner stone of my academic development, with enlightening discussions; and for being a good friend in foreign lands.

I would like to thank all the friends I made at the Computational Geophysics Laboratory (LGC), the ones that are already gone and the newcomers. Among the students and the technical staff I could find valuable discussions, helpful technical support and, most important of all, good friends.

I would like to thanks the Petroleum Science and Engineering Program staff, in special, Sonia and Michelle. They were always helpful, thoughtful and competent.

Last, but not the least, I would like to thank also the Brazilian National Research Council (CNPq) and the Science and Technology National Institute for Petroleum Geophysics (INCT-GP), as well as Petrobras, Schlumberger, and the sponsors of the Wave Inversion Technology (WIT) Consortium for The financial support.







# Abstract

While in principle nonlinear gradient-based full-waveform inversion (FWI) is capable of handling all aspects of wave propagation contained in the data, including full nonlinearity, in practice, it is limited due to its notorious sensitivity to the choice of the starting model. To help addressing model-convergence issues in FWI, we study a decomposition based on the scattering theory that allows to break the acoustic-wavefield sensitivity kernels with respect to model parameters into background and singular parts. The estimates for both background perturbation and/or singular-part perturbation obtained with the subkernels' adjoints are components of the estimate obtained with the total kernel's adjoint. Our numerical experiments shows the feasibility of our main claim: the decomposition into subkernels allows to backproject the scattered-wavefield residuals only so as to obtain reasonable background-model perturbation estimates. In an experiment with restricted acquisition geometry (reflection data, narrow offset), the multiple-scattering subkernels take advantage of medium self-illumination provided by the scattered wavefields. This self-illumination provides better estimates, with longer wavelength content.

**Key words:** Inverse theory; Full waveform inversion; Wave scattering and diffraction; Computational seismology; Theoretical seismology; Wave propagation







# Resumo

A inversão de onda completa (FWI, do inglês “*full waveform inversion*”) não-linear baseada em gradientes (métodos de descida) é, a princípio, capaz de levar em conta todos os aspectos da propagação de onda contida nos dados sísmicos. Porém, FWI baseada em gradientes é limitada pela sua bem conhecida sensibilidade no que diz respeito à escolha do modelo inicial. Com o intuito de melhor entender algumas questões relacionadas à convergência do modelo na FWI, nós estudamos uma decomposição baseada na teoria de espalhamento que permite dividir os núcleos de sensibilidade dos campos de onda acústica em função dos parâmetros do modelo em duas partes: uma relativa ao componente de fundo, outra relativa à componente singular do modelo. Estimativas para a perturbação de fundo, bem como para a perturbação da parte singular do modelo obtidas com os adjuntos destes subnúcleos são componentes da estimativa obtida com o adjunto do núcleo total de sensibilidade. Os experimentos numéricos suportam a tese de que a decomposição em subnúcleos permite que se retroprojete somente os resíduos do campo de onda espalhado de modo a obter estimativas razoáveis da perturbação de fundo do modelo. Em um experimento com geometria de aquisição restrita (dados de reflexão com afastamento curto), os subnúcleos baseados em espalhamento múltiplo se aproveitam da autoiluminação do meio devido às ondas múltiplamente espalhadas. A autoiluminação fornece estimativas melhores com conteúdo espectral mais rico nas baixas frequências.

**Palavras-chave:** Problemas inversos; Inversão de onda completa; espalhamento de onda e difração; Sísmica computacional; Teoria sísmica; Propagação de onda





# Contents

<b>List of Figures</b>	<b>xvii</b>
<b>List of Tables</b>	<b>xxvii</b>
<b>List of Symbols, Variables and their Definitions</b>	<b>xxix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 FWI overview . . . . .	2
1.1.1 Forward Modelling . . . . .	3
1.1.2 Multiscale approach . . . . .	5
1.1.3 Optimisation methods . . . . .	7
1.1.4 Adjoint-state methods . . . . .	12
1.1.5 Challenges . . . . .	14
1.2 Kernel decomposition . . . . .	16
<b>2 Kernel Decomposition</b>	<b>21</b>
2.1 Full-waveform inversion and sensitivity kernels . . . . .	21
2.1.1 Conventional formulation . . . . .	22
2.1.2 Decomposed perturbation . . . . .	26
2.2 Physical interpretation of the forward decomposition . . . . .	33
2.2.1 Old information, brand new point of view . . . . .	33
2.2.2 Multiple scattering/internal reflection uncovered . . . . .	33
2.2.3 Background and singular-part contributions . . . . .	36
2.2.4 Physical interpretation of the kernels: naming the terms . . . . .	38
2.3 Inversion: backprojecting the residuals into model space . . . . .	41

2.3.1	Conventional formulation . . . . .	41
2.3.2	Decomposed formulation . . . . .	42
2.3.3	Residual information leakage . . . . .	43
<b>3</b>	<b>Numerical Experiments and Discussion</b>	<b>45</b>
3.1	Experimental Procedure . . . . .	45
3.1.1	Forward Decomposition . . . . .	45
3.1.2	Backprojecting residuals . . . . .	46
3.2	Model gallery . . . . .	47
3.3	Experiment gallery . . . . .	51
3.3.1	Experiment 1: Constant background perturbation of Model I, forward decomposition . . . . .	51
3.3.2	Experiment 2: Lens-shaped background perturbation of Model I, backprojecting residuals . . . . .	58
3.3.3	Experiment 3: Constant background perturbation of Model II, forward decomposition . . . . .	64
3.3.4	Experiment 4: Perturbing the scatterers' position of Model II, forward decomposition . . . . .	68
3.3.5	Experiment 5: Perturbing the scatterers' position of Model II, backprojecting residuals . . . . .	71
3.3.6	Experiment 6: Lens-shaped background perturbation of Model III, forward decomposition . . . . .	76
3.3.7	Experiment 7: Lens-shaped background perturbation of Model III, backprojecting residuals . . . . .	80
3.3.8	Experiment 8: Lens-shaped background perturbation of Model IV, backprojecting residuals . . . . .	87
<b>4</b>	<b>Application: Scattering-based Sensitivity Kernels for Time-lapse Differential Wave- form Inversion</b>	<b>93</b>
4.1	Time-lapse survey as a FWI scattering problem . . . . .	94
4.2	Numerical experiments . . . . .	95
4.2.1	Large offset surveys . . . . .	96
4.2.2	Narrow offset surveys . . . . .	104



## CONTENTS

---

<b>5</b>	<b>Final Remarks</b>	<b>111</b>
5.1	Thesis' follow-ups . . . . .	115
	<b>References</b>	<b>117</b>
<b>A</b>	<b>Mathematical Definitions</b>	<b>125</b>
<b>B</b>	<b>Generalised decomposition and its relation to the proposed decomposition</b>	<b>127</b>
B.1	Generalised decomposition . . . . .	127
B.2	Four-fold decomposition of a perturbed model and wavefield . . . . .	128
B.3	Linearisation of the fourfold decomposition . . . . .	131
<b>C</b>	<b>Estimates for the Model Perturbations</b>	<b>135</b>
C.1	Estimate of the background perturbation $\delta K_{B,BS}^{\text{est}}$ . . . . .	135
C.2	Estimate of the background perturbation $\delta \rho_{B,BS}^{\text{est}}$ . . . . .	136
C.3	Estimate of the background perturbation $\delta K_{B,b\nu B}^{\text{est}}$ . . . . .	138
C.4	Estimate of the background density perturbation $\delta \rho_{B,b\nu B}^{\text{est}}$ . . . . .	140
<b>D</b>	<b>Superposition of Subkernels of the Singular-wavefield Residual</b>	<b>143</b>





# List of Figures

1.1	[Extracted from Bunks et al. (1995)] Heuristic illustration of the multigrid method applied to a nonlinear problem. Frame (a) illustrates a 1-D objective function and frames (b)-(e) illustrate the same objective function at ever increasing scale lengths. A gradient method applied to the shortest scale succeeds in finding the local minimum closest to the starting point, whereas the gradient method applied to the longest scale finds the global minimum (at that scale) regardless of the starting point. A fining up procedure helps find the estimate of the position of the global minimum for the fine scale problem. . . . .	6
1.2	Line search with three trial step lengths and a second-order polynomial interpolation. The parabola vertex defines the efficient step length. . . . .	8
1.3	Convergence in the steepest descent method for a quadratic objective function. . . .	11
1.4	Conventionally, the exact model (top) is thought of as being decomposed into a smooth background (bottom left), supposedly known, and a singular scattering part (bottom right), to be determined under the assumption that it is small against the background. . . . .	17
1.5	Here, the exact model (top) is thought of as being decomposed into a slightly wrong estimate of the smooth background (centre left), and a slightly wrong singular scattering part (bottom left), both supposed to be known, plus perturbations to the background (centre right) <i>and</i> the singular part (bottom right), to be determined under the assumption that they are small against their known estimates. . . . .	19
2.1	A sketch of the proposed decomposition. The components of the reference model are thought to represent a (smooth) background model and an estimate of the singular part of the model ( $\mathbf{m}_0 = \mathbf{m}_B + \mathbf{m}_S$ ). The two components of the perturbation part are the perturbations of these two model components ( $\delta\mathbf{m} = \delta\mathbf{m}_B + \delta\mathbf{m}_S$ ). . . .	28

2.2	Different terms showing different levels of interaction of single- and multiple- scattered/reflected information. The colour code is the one adopted in eqs. (2.35), (2.36), (2.37) and (2.38). From top left, in clockwise direction: (a) and (b) show contribution from single-scattering events, difference between them resting on which operator is used to generate the secondary source to be extrapolated; (c) brings contribution from multiple-scattering events from receiver-side; (d) shows the contribution of the strong multiple-scattering events. . . . .	34
2.3	Subkernels obtained from the decomposition of the sensitivity kernel $\bar{U}$ for bulk-modulus perturbation. Kernel $\bar{V}$ for density perturbation is decomposed analogously.	37
2.4	Physical meaning of the subkernels. Each one of the cartoons shows three elements: <i>source-side wavefield</i> ; the <i>operator</i> generating the secondary source; and the <i>receiver-side wavefield extrapolator</i> . Subcaptions indicate the new name and correspondence to the previous, numbered nomenclature. . . . .	40
3.1	Unperturbed velocity Model I (velocity in $m/s$ ). The first layer has a velocity of $1500 m/s$ , and the second layer a velocity of $2500 m/s$ . Density is kept constant ( $2200 kg/m^3$ ). The source and receiver positions are indicated by black star and triangle respectively. . . . .	48
3.2	Unperturbed velocity Model II (velocity in $m/s$ ) containing 40 randomly distributed scatterers, with constant background velocity ( $2205 m/s$ ) and density ( $2200 kg/m^3$ ). One source (black star) and two control receivers (black triangles) – named as cr154 ( <i>top</i> ) and cr283 ( <i>bottom</i> ) – were used in the forward decompositions experiments. The other sources and receivers (white stars and triangles respectively) were used in the backpropagation experiment. . . . .	48
3.3	Unperturbed velocity Model III (velocity in $m/s$ ) containing 60 randomly distributed scatterers, with constant background velocity ( $2205 m/s$ ) and density ( $2200 kg/m^3$ ). For the numerical experiments, we used the same acquisition geometry as in Model II. . . . .	50
3.4	Unperturbed velocity Model IV (velocity in $m/s$ ) containing 150 randomly distributed scatterers, with constant background velocity ( $2205 m/s$ ) and density ( $2200 kg/m^3$ ). . . . .	50



## LIST OF FIGURES

---

- 3.5 EXPERIMENT 1. True total wavefield residual  $\delta p^{\text{true}} = p - p_0$  at cr550. Observe two distinct signals: the direct wavefield residual (peak at 0.27 s) and the reflected/scattered wavefield residual (peak at 0.36 s). . . . . 52
- 3.6 EXPERIMENT 1. Linearised background-wavefield residuals  $\delta p_B$  (blue line) evaluated at cr550 as compared to the true one (red line),  $\delta p_B^{\text{true}}$ . . . . . 53
- 3.7 EXPERIMENT 1. Terms related to single-scattering effects at the medium perturbations (blue lines):  $\delta p_{S,-BCLB}$  (top),  $\delta p_{S,BBB}$  (centre), and their sum (bottom). The red line true singular residual  $\delta p_S^{\text{true}}$ . Separately,  $\delta p_{S,-BCLB}$  and  $\delta p_{S,BBB}$  carry the largest energy. Their sum demonstrates that they are approximately complementary. 53
- 3.8 EXPERIMENT 1. Multiple-scattering contributions (blue lines), from top to bottom:  $\delta p_{S,BBS}$ ,  $\delta p_{S,SBB}$  and  $\delta p_{S,SBS}$ . In all figures,  $\delta p_S^{\text{true}}$  is depicted in red lines for comparison. The first two terms carry most of the energy of the singular residual  $\delta p_S$ . . . . . 54
- 3.9 EXPERIMENT 1. Non-symmetrical contributions (blue lines), first four traces from top to bottom:  $\delta p_{S,-BLS}$ ,  $\delta p_{S,bVB}$ ,  $\delta p_{S,bVS}$  and  $\delta p_{S,bVB} + \delta p_{S,bVS}$  respectively. In these traces,  $\delta p_S^{\text{true}}$  is depicted in red lines for comparison. The last trace show both  $-\delta p_{S,-BLS}$  (black line) and  $\delta p_{S,bVB} + \delta p_{S,bVS}$  (dark green line) contributions normalised to their respective maximum absolute value. . . . . 56
- 3.10 EXPERIMENT 1. Comparing  $\delta p^{\text{true}}$  (red line) with evaluated counter-parts. In black,  $\delta p$  (evaluated with eq. (2.14)) and, in blue,  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all but the 1<sup>st</sup>, 2<sup>nd</sup> and 7<sup>th</sup> terms of eq. (2.33) respectively). See the perfect match between blue and black lines. . . . . 57
- 3.11 EXPERIMENT 2. Perturbation in the bulk-modulus background model,  $\delta K_B$ . The scalebar shows, in percentage, the perturbation magnitude. . . . . 58
- 3.12 EXPERIMENT 2. *Top*: Bulk modulus estimate  $\delta K^{\text{est}}$ . The others are all background bulk-modulus-estimate contributions. *From left to right, top to bottom*:  $\delta K_{B,B}^{\text{est}}$ ;  $\delta K_{B,-BCLB}^{\text{est}}$ ;  $\delta K_{B,-BS}^{\text{est}}$ ;  $\delta K_{B,-SB}^{\text{est}}$ ;  $\delta K_{B,-BB}^{\text{est}}$  and  $\delta K_{B,-SS}^{\text{est}}$ . The first of them in the upper-left corner,  $\delta K_{B,B}^{\text{est}}$ , is result of background residual backprojection; the rest of them are result of singular residual backprojection. The subtraction of  $\delta K_{B,B}^{\text{est}}$  (upper-left corner) with  $\delta K_{B,-BCLB}^{\text{est}}$  (upper-right corner) yields the conventional estimate  $\delta K_{\text{conv}}^{\text{est}}$  (Figure 3.13, upper-left corner). . . . . 60

- 3.13 EXPERIMENT 2. *Top*: Bulk modulus estimate  $\delta K^{\text{est}}$ . *Upper-left corner*: Bulk modulus estimate  $\delta K_{\text{conv}}^{\text{est}}$  obtained with the conventional approach. The others are all background bulk-modulus-estimate contributions. *Upper-right corner*: the background estimate  $\delta K_B^{\text{est}} = \delta K_{B,B}^{\text{est}} + \delta K_{B,S}^{\text{est}}$ . *Lower-left corner*:  $\delta K_{B,B}^{\text{est}}$ , result of background residual backprojection. *Lower-right corner*:  $\delta K_{B,S}^{\text{est}}$ , result of singular residual backprojection (sum of all five contributions). . . . . 61
- 3.14 EXPERIMENT 2. First four estimates are obtained by backprojecting the background residual with the adjoint sub-kernels  $\bar{U}_{S,j}^\dagger$ , i.e.,  $\bar{\delta K}_{j(B)}^{\text{est}}$  (cross-terms). *From left to right, top to bottom*:  $\bar{\delta K}_{BS(B)}^{\text{est}} = \bar{U}_{S,BS} \hat{p}_B$ ;  $\bar{\delta K}_{BB(B)}^{\text{est}} = \bar{U}_{S,BB} \hat{p}_B$ ;  $\bar{\delta K}_{SS(B)}^{\text{est}} = \bar{U}_{S,SS} \hat{p}_B$ ;  $\bar{\delta K}_{SB(B)}^{\text{est}} = \bar{U}_{S,SB} \hat{p}_B$ . The next two figures are each one of the sums seen in the right-hand side of eq. (3.11). *From left to right*:  $\left(\sum_j \bar{U}_{S,j}^\dagger\right) \hat{p}_B$  and  $\left(\sum_j \bar{U}_{S,j}^\dagger\right) \hat{p}_S$ . At the lower-left corner is displayed the summation of the previous sums, i.e.,  $\bar{\delta K}^{\text{est}}$  of eq. (3.11). At the lower-right corner, the bulk-modulus estimate  $\delta K^{\text{est}}$  for comparison. . . . . 62
- 3.15 EXPERIMENT 3. Linearised background-wavefield residuals  $\delta p_B$  (blue line) evaluated at cr154 as compared to the true ones (red line),  $\delta p_B^{\text{true}}$ . . . . . 65
- 3.16 EXPERIMENT 3. Terms related to single-scattering effects at the medium perturbations evaluated at cr154 (blue lines):  $\delta p_{S,-BLB}$  (top),  $\delta p_{S,BBB}$  (centre), and their sum (bottom). The red line true singular residual  $\delta p_S^{\text{true}}$ . Separately,  $\delta p_{S,-BLB}$  and  $\delta p_{S,BBB}$  carry the largest energy. Their sum,  $\delta p_{S,-BWB}$ , demonstrates that they are approximately complementary. . . . . 66
- 3.17 EXPERIMENT 3.: First two traces from top to bottom:  $\delta p_{S,-BLS}$  and  $\delta p_{S,bVB} + \delta p_{S,bVS}$ , respectively. In these traces,  $\delta p_S^{\text{true}}$  is depicted in red lines for comparison. The last trace show both  $\delta p_{S,BLS}$  (black line) and  $\delta p_{S,bVB} + \delta p_{S,bVS}$  (dark green line) contributions normalised to their respective absolute maximum value. . . . . 67
- 3.18 EXPERIMENT 3. Comparing  $\delta p^{\text{true}}$  (red line) with evaluated counter-parts at cr154. In black,  $\delta p$  (evaluated with eq. (2.14)); in dark green,  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all terms of eq. (2.33) respectively); and in blue,  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all but the 1<sup>st</sup>, 2<sup>nd</sup> and 7<sup>th</sup> terms of eq. (2.33) respectively). The blue and black lines are indistinguishable. . . . . 68

## LIST OF FIGURES

---

- 3.19 EXPERIMENT 3. Multiple-scattering contributions (blue lines), from top to bottom:  $\delta p_{S,BBS}$ ,  $\delta p_{S,SB B}$  and  $\delta p_{S,SB S}$ . These are the terms that carry most of the energy of the singular residual  $\delta p_S$ . In the last trace we can see the summation of the first two traces depicted above (dark green line) and the  $\delta p_S$  evaluated with all five contributions ( $BBB$ ,  $SB B$ ,  $SB S$ ,  $BBS$  and  $-B\mathcal{L}B$ ). In all figures,  $\delta p_S^{\text{true}}$  is depicted in red lines for comparison. . . . . 69
- 3.20 EXPERIMENTS 4 AND 5. Only the singular part of the model is perturbed, i.e.,  $\delta K_S \neq 0$ . This is achieved by modifying the vertical and horizontal coordinates of the scatterers randomly within a given range. The cool (blue) spots show the unperturbed positions, while the hot (red) ones tell us the new positions. The intensity of the spots is related to the magnitude of the perturbation. The green background corresponds to no perturbation. . . . . 70
- 3.21 EXPERIMENT 4. Traces evaluated at cr154. Total residual,  $\delta p$ , given by equation (2.14), in black, and singular-wavefield residual,  $\delta p_S$ , equation (2.32), in blue, match almost perfectly. Good match between linearised and true residual,  $\delta p_S$  (in blue) and  $\delta p_S^{\text{true}}$  (in red) indicates the good quality of the Born approximation in this case, which is a consequence of the small perturbations in the model. . . . . 71
- 3.22 EXPERIMENT 4. All four terms that contribute to the total singular residual at cr154. *From top to bottom:*  $\delta p_{S,B S B}$ ,  $\delta p_{S,S S B}$ ,  $\delta p_{S,B S S}$ ,  $\delta p_{S,S S S}$ . The single-scattering term,  $\delta p_{S,B S B}$ , accounts for almost all of the complete singular residual. . . . . 72
- 3.23 EXPERIMENT 5. Backprojecting the total (singular) wavefield residual,  $\delta p^{\text{true}} = \delta p_S^{\text{true}}$ . Cartoons indicate the (sub)kernels used. *Top:* Bulk modulus perturbation estimate,  $\delta K^{\text{est}}$ , obtained with the first equation of (2.49). *Bottom:* Singular part of bulk modulus perturbation estimate,  $\delta K_S^{\text{est}}$ , obtained with the second equation of (3.14). Colour scale clipped at 0.6 for better comparison. . . . . 73
- 3.24 EXPERIMENT 5. Each one of the contributions to the singular perturbation estimate. From top to bottom:  $\delta K_{S,B S}^{\text{est}}$ ,  $\delta K_{S,S B B}^{\text{est}}$ ,  $\delta K_{S,S S S}^{\text{est}}$  and  $\delta K_{S,S B}^{\text{est}}$ . . . . . 74

3.25	EXPERIMENT 5. <i>Top</i> : Singular part of Bulk modulus estimate, $\delta K_{S,BB}^{\text{est}}$ , obtained with the single-scattering related kernel. <i>Center</i> : Conventional Bulk modulus estimate, $\delta K_{\text{conv}}^{\text{est}}$ , obtained with the background-wavefield related kernel. Note that these estimates are very similar to the true perturbation observed in Figure 3.20, the conventional one being a little better. <i>Bottom</i> : Background bulk-modulus estimate, $\delta K_B^{\text{est}}$ , obtained with first equation of (3.14) (residual leakage). Colour scale clipped at 0.6 for better comparison. . . . .	75
3.26	EXPERIMENTS 6 AND 7. Perturbation in the bulk-modulus background model, $\delta K_B$ . The scalebar shows, in percentage, the perturbation magnitude. . . . .	77
3.27	EXPERIMENT 6. Comparison between the total wavefield residual, $\delta p$ (black lines) and sum of background and singular residuals, $\delta p_B + \delta p_S$ (blue lines). <i>Top</i> : Evaluated at control receiver cr154. <i>Bottom</i> : Evaluated at control receiver cr283. . . . .	77
3.28	EXPERIMENT 6. Single-scattering terms evaluated at control receiver cr154 (blue lines): $\delta p_{S,BBB}$ (top), $\delta p_{S,-B\mathcal{L}B}$ (centre), and their sum $\delta p_{S,BWB}$ , compared to the true singular-wavefield residual $\delta p_S^{\text{true}}$ (red lines). As expected, these terms are very small because the Fresnel zone of the direct wavefield does not intersect with the lens-shaped perturbed area. Compare with Figure 3.29. . . . .	79
3.29	EXPERIMENT 6. Single-scattering terms evaluated at control receiver cr283 (blue lines): $\delta p_{S,BBB}$ (top), $\delta p_{S,-B\mathcal{L}B}$ (centre), and their sum $\delta p_{S,BWB}$ , compared to the true singular-wavefield residual $\delta p_S^{\text{true}}$ (red lines). Each of these terms carries much energy, but their sum contributes very little to the singular-wavefield residual. Compare with Figure 3.28. . . . .	80
3.30	EXPERIMENT 6. Terms with the most important contributions to the singular residual evaluated at cr154. <i>From top to bottom</i> : $\delta p_{S,BBS}$ , $\delta p_{S,SB\mathcal{B}}$ and $\delta p_{S,SBS}$ . . .	81
3.31	EXPERIMENT 6. Terms with the most important contributions to the singular residual evaluated at cr283. <i>From top to bottom</i> : $\delta p_{S,BBS}$ , $\delta p_{S,SB\mathcal{B}}$ and $\delta p_{S,SBS}$ . . .	82
3.32	EXPERIMENT 7. Backprojecting residuals. <i>From top to bottom</i> : Bulk modulus perturbation estimate, $\delta K^{\text{est}}$ , from the total wavefield residual, $\delta p^{\text{true}}$ (equation (2.49)); Background bulk-modulus perturbation estimate, $\delta K_{B,B}^{\text{est}}$ , from background residual, $\delta p_B^{\text{true}}$ (first contribution in first equation of (3.7)); Background bulk-modulus perturbation estimate, $\delta K_{B,S}^{\text{est}}$ , from singular residual, $\delta p_S^{\text{true}}$ (second contribution in first equation of (3.7)); and the sum $\delta K_{B,B}^{\text{est}} + \delta K_{B,S}^{\text{est}}$ . . . . .	83



## LIST OF FIGURES

---

- 3.33 EXPERIMENT 7. Backprojecting residuals. Estimates related to eq. (3.11). *From top to bottom:*  $\sum_j \bar{\delta K}_{j(B)}^{\text{est}} = \left( \sum_j \bar{U}_{S,j}^\dagger \right) \hat{\delta p}_B$ ,  $\sum_j \bar{\delta K}_{B,j}^{\text{est}} = \left( \sum_j \bar{U}_{S,j}^\dagger \right) \hat{\delta p}_S$  and their sum. The last is the total perturbation estimate  $\delta K^{\text{est}}$ , put here for comparison. 84
- 3.34 EXPERIMENT 7. Backprojecting singular-wavefield residual  $\delta p_S^{\text{true}}$ . *Top:* Subkernel  $\bar{U}_{B,BS}$  used for backprojection. *Bottom:* Subkernel  $\bar{U}_{B,SB}$  used for backprojection. Note the very strong similarity between both estimates apart from the region delimited by the receiver and source lines. It suggests a redundancy that might be explored. . . . . 85
- 3.35 EXPERIMENT 7. Backprojecting singular-wavefield residual  $\delta p_S^{\text{true}}$  recorded at cr283 from the source at 1000 m (black star). *Top:* Subkernel  $\bar{U}_{B,BS}$  used for backprojection. The singularities act as receivers. *Bottom:* Subkernel  $\bar{U}_{B,SB}$  used for backprojection. Here, the singularities act as sources. This gives us further insight into the congruence of the parts of Figure 3.34. . . . . 86
- 3.36 EXPERIMENT 8. Backprojecting residuals from “surface” receivers. *Top:* Bulk-modulus perturbation estimate  $\delta K^{\text{est}}$  from total wavefield residual  $\delta p^{\text{true}}$ , *Center:* Bulk-modulus perturbation estimate  $\delta K_{B,B}^{\text{est}}$  from background residual  $\delta p_B^{\text{true}}$ , *Bottom:* Bulk-modulus perturbation estimate  $\delta K_{B,S}^{\text{est}}$  from singular residual  $\delta p_S^{\text{true}}$ . . . . 88
- 3.37 EXPERIMENT 8. Cross-term estimate resulting from the sum  $\left( \sum_j \bar{U}_{S,j}^\dagger \right) \hat{\delta p}_B$ . . . 89
- 3.38 EXPERIMENT 8. Single-scattering contribution  $\delta K_{B,-B\text{LB}}^{\text{est}}$ . . . . . 89
- 3.39 EXPERIMENT 8. Bulk-modulus perturbation estimates. *From top to Bottom:*  $\delta K_{B,BS}^{\text{est}}$ ,  $\delta K_{B,SB}^{\text{est}}$ ,  $\delta K_{B,BB}^{\text{est}}$  and conventional estimate  $\delta K_{\text{conv}}^{\text{est}}$  (backpropagation of complete residual  $\delta p$  with only background model). . . . . 90
- 3.40 EXPERIMENT 8. *From top to bottom:* Single pair source-receiver contributions to conventional estimate  $\delta K_{\text{conv}}^{\text{est}}$ ,  $\delta K_{B,BS}^{\text{est}}$ ,  $\delta K_{B,SB}^{\text{est}}$ , and  $\delta K_{B,BB}^{\text{est}}$ . . . . . 92
- 4.1 Center parts of the baseline (a and b) and monitor (c and d) models for Model Sets 1 (a and c) and 2 (b and d). . . . . 95
- 4.2 SET 1 - LARGE OFFSET. Estimates for bulk-modulus perturbation in the absence of a strong reflector. (a) Singular part estimate according to equation (4.1); (b) Conventional perturbation estimate using wave propagation in the background model. The last is substantially close to contribution  $\delta K_{S,BB}^{\text{est}}$  (see Figure 4.4c, right). Both figures are normalised to their common maximum for comparison . . . . . 97

4.3	SET 1 - LARGE OFFSET. The same estimates seen in Figure 4.2 ( $\delta K_S^{\text{est}}$ and $\delta K_{\text{conv}}^{\text{est}}$ ) but now each one is normalised to their own maximum. . . . .	98
4.4	SET 1 - LARGE OFFSET. Individual contributions to $\delta K_S^{\text{est}}$ are shown in the left column. In the right column we see the results of succesively summing up the contributions. Thus, the last one is the result of summing up all four contributions. . . . .	99
4.5	SET 1 - LARGE OFFSET. <i>Top</i> : the norm of the residual (as a measure of the energy) at the receivers (horizontal axis) as a function of shots (vertical axis). <i>Bottom</i> : Result of stacking (integrating) the above function over sources. The energy of the residuals concentrate between 500 to 1000 $m$ and 2000 to 2500 $m$ . . . . .	101
4.6	SET 2 - LARGE OFFSET. Estimates for bulk-modulus perturbation in the presence of a strong reflector. (a) Singular part estimate according to equation (4.1); (b) Conventional perturbation estimate using wave propagation in the background model. The last is substantially close to contribution $\delta K_{S,BB}^{\text{est}}$ (see Figure 4.7c, right); (c) Same as before, but with dashed ray paths indicating the back-illumination from the strong reflector. . . . .	102
4.7	SET 1 - LARGE OFFSET. Individual contributions to $\delta K_S^{\text{est}}$ are shown in the left column. In the right column the result of summing up the current contribution with all the previous ones. which makes, for instance, the first estimate in the left to be no more than $\delta K_{S,BS}^{\text{est}}$ ; on its turn, the last one is the result of summing up all four contributions. . . . .	103
4.8	SET 2 - LARGE OFFSET. <i>Top</i> : the norm of the residual (as a measure of the energy) at the receivers (horizontal axis) as a function of shots (vertical axis). <i>Bottom</i> : Result of stacking (integrating) the above function over sources. . . . .	104
4.9	SET 1 - NARROW OFFSET. Narrow-offset estimates of the perturbation in the absence of a strong reflector. (a) Singular estimate with all contributions; (b) conventional estimate. . . . .	105
4.10	SET 1 - NARROW OFFSET. <i>From top to bottom, left column</i> : contributions to singular perturbation estimate $\delta K_{S,BS}^{\text{est}}$ , $\delta K_{S,BB}^{\text{est}}$ , $\delta K_{S,SS}^{\text{est}}$ and $\delta K_{S,SB}^{\text{est}}$ . <i>Right column</i> : Cumulative summation of individual contributions. . . . .	106
4.11	SET 1 - NARROW OFFSET. Summation of the multiple-scattering contributions $\delta K_{S,SS}^{\text{est}}$ and $\delta K_{S,SB}^{\text{est}}$ . . . . .	107

## LIST OF FIGURES

---

4.12	SET 2 - NARROW OFFSET. Narrow-offset estimates of the perturbation in the presence of a strong reflector. (a) Singular estimate with all contributions; (b) conventional estimate. . . . .	108
4.13	SET 2 - NARROW OFFSET. Summation of the multiple-scattering contributions $\delta K_{S,SS}^{\text{est}}$ and $\delta K_{S,SB}^{\text{est}}$ . . . . .	109
4.14	SET 2 - NARROW OFFSET. <i>From top to bottom, right column:</i> contributions to singular perturbation estimate $\delta K_{S,BS}^{\text{est}}$ , $\delta K_{S,BB}^{\text{est}}$ , $\delta K_{S,SS}^{\text{est}}$ and $\delta K_{S,SB}^{\text{est}}$ . <i>Left column:</i> Cumulative summation of individual contributions. . . . .	110

## LIST OF FIGURES

---



# List of Tables

3.1	Unperturbed model gallery. $n\#$ stands for the number of samples in the $\#$ -axis. $d\#$ is the sampling interval in the $\#$ -axis; $x$ - and $z$ -axis are given in meters; $t$ -axis is given in milliseconds. the origin for all axis was always $o\# = 0$ . . . . .	49
3.2	Experiment gallery. . . . .	51
3.3	Experiment 3: $L^2$ -norm as a measure of energy of all observed terms at cr154. The terms are sorted in descending order of energy. The terms related to single scattering effects are the strongest, followed by the ones related to multiple-scattering. The braces with attached values show the energy of the corresponding summed traces, indicating destructive interference of the individual contributions (see also Figures 3.16). . . . .	67
3.4	Experiment 6: $L^2$ -norm as a measure of energy at cr154 and cr283. The braces indicate summation of the related traces, the value attached to them being the energy of the resulting trace. . . . .	78

## **LIST OF TABLES**

---

# List of Symbols, Variables and their definitions

Here, all the variable dependencies are shown. Throughout the text some dependencies of the variables may be (partially) omitted. The terms in parenthesis are usually omitted for sake of simplicity.

## Basic Variables

$\mathbf{x}$	Spatial position
$\mathbf{x}_s$	Point-source position
$\mathbf{x}_g$	Point-receiver position
$t$	Time
$\omega$	Angular frequency
$S(t)$	Source wavelet

## Wavefields

$p(\mathbf{x}, t; x')$	Total pressure wavefield
$p_0(\mathbf{x}, t; x')$	Reference wavefield
$p_s(\mathbf{x}, t; x')$	Scattered wavefield
$p_B(\mathbf{x}, t; x')$	Background (part of the reference) wavefield
$p_S(\mathbf{x}, t; x')$	Singular (part of the reference) wavefield
$\delta p(\mathbf{x}, t; x')$	Total-wavefield perturbation/residual
$\delta p_B(\mathbf{x}, t; x')$	Background-wavefield perturbation/residual
$\delta p_S(\mathbf{x}, t; x')$	Singular-wavefield perturbation/residual
$\delta p_{S,i}(\mathbf{x}, t; x')$	Singular-wavefield residual term/component/contribution
$G_0(\mathbf{x}, t; x')$	Reference (medium) Green's function

## LIST OF SYMBOLS AND VARIABLES

---

$G_B(\mathbf{x}, t; \mathbf{x}')$	Background (medium) Green's function
$G_S(\mathbf{x}, t; \mathbf{x}')$	Singular (part of the reference medium) Green's function

### *Wavefield superscripts*

Used in any wavefield residual.

$K$	Residual is caused by bulk modulus perturbation
$\rho$	Residual is caused by density perturbation

### **Model parameters**

$\mathbf{m}(\mathbf{x})$	Model parameter vector
$\mathbf{m}_0(\mathbf{x})$	Reference model parameter vector
$\delta\mathbf{m}(\mathbf{x})$	Reference model perturbation vector
$K(\mathbf{x})$	Bulk modulus
$K_0(\mathbf{x})$	Reference bulk modulus
$K_B(\mathbf{x})$	Background (part of the reference) bulk modulus
$K_S(\mathbf{x})$	Singular (part of the reference) bulk modulus
$K_i(\mathbf{x})$	i-th part of bulk modulus
$\rho(\mathbf{x})$	Density
$\rho_0(\mathbf{x})$	Reference density
$\rho_B(\mathbf{x})$	Background (part of the reference) density
$\rho_S(\mathbf{x})$	Singular (part of the reference) density
$\rho_i(\mathbf{x})$	i-th part of density
$\delta K(\mathbf{x})$	Reference bulk modulus perturbation
$\delta K_B(\mathbf{x})$	Background bulk modulus perturbation
$\delta K_S(\mathbf{x})$	Singular bulk modulus perturbation
$\delta K_{B,B}^{\text{est}}(\mathbf{x})$	Background bulk modulus perturbation estimate from background wavefield residual backprojection
$\delta K_{B,S}^{\text{est}}(\mathbf{x})$	Background bulk modulus perturbation estimate from singular wavefield residual backprojection
$\delta K_{S,S}^{\text{est}}(\mathbf{x})$	Singular bulk modulus perturbation estimate from singular wavefield residual backprojection



## LIST OF SYMBOLS AND VARIABLES

---

$\delta K_{B,i}^{\text{est}}(\mathbf{x})$	i-th background bulk modulus perturbation estimate contribution from singular wavefield residual backprojection with i-th component of the first derivatives of singular wavefield w.r.t. background/singular bulk modulus
$\delta K_{B,i(B)}^{\text{est}}(\mathbf{x})$	i-th background bulk modulus perturbation estimate contribution from background wavefield residual backprojection with i-th component of the first derivatives of singular wavefield w.r.t. background/singular bulk modulus
$\delta K_{S,i}^{\text{est}}(\mathbf{x})$	i-th singular bulk modulus perturbation estimate contribution from singular wavefield residual backprojection with i-th component of the first derivatives of singular wavefield w.r.t. background/singular bulk modulus
$\delta \rho(\mathbf{x})$	Reference density perturbation
$\delta \rho_B(\mathbf{x})$	Background bulk modulus perturbation
$\delta \rho_S(\mathbf{x})$	Singular bulk modulus perturbation

### *Model parameter superscripts*

Used in any model parameter perturbation.

est	The perturbation is an estimate from wavefield backprojection
-----	---

## Operators

They can be applied onto any wavefield. Their space-dependency is inherited from the model parameters they are function of.

$\mathcal{L}(K, \rho)[\cdot]$	(Acoustic) wave operator
$\mathcal{L}_0(K_0, \rho_0)[\cdot]$	Reference (acoustic) wave operator
$\mathcal{L}_B(K_B, \rho_B)[\cdot]$	Background (acoustic) wave operator
$\mathcal{L}_j(K_{i=0,\dots,j}, \rho_{i=0,\dots,j})[\cdot]$	j-th level (acoustic) wave operator
$\mathcal{V}(K, K_0, \rho, \rho_0)[\cdot]$	(Acoustic) scattering potential
$\mathcal{V}_0(K_0, K_B, \rho_0, \rho_B)[\cdot]$	Reference (acoustic) scattering potential
$\mathcal{V}_j(K_{i=0,\dots,j}, \rho_{i=0,\dots,j})[\cdot]$	j-th level (acoustic) scattering potential
$\delta \mathcal{L}(\delta K, K, \delta \rho, \rho)[\cdot]$	(Acoustic) secondary potential
$\delta \mathcal{L}_B(\delta K_B, K_B, \delta \rho_B, \rho_B)[\cdot]$	Background (acoustic) secondary potential
$\delta \mathcal{B}(\delta K_B, K_0, \delta \rho_B, \rho_0)[\cdot]$	Background part of (acoustic) secondary potential
$\delta \mathcal{S}(\delta K_S, K_0, \delta \rho_S, \rho_0)[\cdot]$	Singular part of (acoustic) secondary potential

**Fréchet derivatives and its components**

$\Phi(\mathbf{x}, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of wavefield w.r.t. full model parameters
$\overline{U}(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of wavefield w.r.t. full bulk modulus
$\overline{V}(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of wavefield w.r.t. full density
$\overline{U}_B(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of background wavefield w.r.t. background bulk modulus
$\overline{V}_B(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of background wavefield w.r.t. background density
$\overline{U}_S^B(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of singular wavefield w.r.t. background bulk modulus
$\overline{U}_S^S(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of singular wavefield w.r.t. singular bulk modulus
$\overline{V}_S^B(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of singular wavefield w.r.t. background density
$\overline{V}_S^S(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	First derivatives of singular wavefield w.r.t. singular density
$\overline{U}_{S,i}(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	i-th component of the first derivatives of singular wavefield w.r.t. background/singular bulk modulus
$\overline{V}_{S,i}(\mathbf{x}_i, t; \mathbf{x}_s, \mathbf{x}_g)$	i-th component of the first derivatives of singular wavefield w.r.t. background/singular density

# 1. Introduction

For many years, the most common imaging techniques were based on ray theory, such as Kirchhoff migration. Lately, as the industry has been facing geologically more complex areas, migration methods based on wavefield extrapolation have been usually applied. This is because in such areas the ray theory is not capable of describing the complexity of wavefield propagation. The development started with one-way wave-equation migration and culminated more recently with two-way wave-equation techniques. All this became possible because of new acquisition techniques which give better illumination of the subsurface, and more powerful computational resources.

Those new imaging methods require more and more refined earth models. Even though migration has advanced quickly with the raise of available computer power, constructing these models is still mostly ray- and travelttime-based. However, in recent years, one model-building tool, based on the two-way wave- equation, has been extensively studied and developed for Earth modelling: *full waveform inversion (FWI)*.

The basic idea behind FWI is optimisation: minimisation of an *objective function*<sup>1</sup> that measures, somehow, the difference between the observed seismic data and synthetic data, numerically modelled using an estimated earth model. In the last decades, many studies on FWI were carried out. Something that was impractical for industrial applications some years ago is now becoming a established element of the seismic processing flow in the exploration and monitoring stages of the oil and gas production. For an overview and the state of the art, the reader is referred to the works of Virieux and Operto (2009), Plessix (2012) and Asnaashari et al. (2013).

A series of papers published in the eighties (Lailly, 1983; Tarantola, 1984a, 1986) introduced the gradient-based FWI method in applied geophysics. In essence, these methods rely on the fact that a model can be updated iteratively with the help of so-called *sensitivity kernels* (Tromp et al., 2005). The sensitivity kernels are the volumetric density of misfit function derivatives with

---

<sup>1</sup>Also known as *misfit* or *objective* functional.

respect to the model parameters. These sensitivity kernels are related to the *Fréchet derivatives* of the seismogram with respect to the model parameters. Under the framework of the adjoint state-method (Tromp et al., 2005; Plessix, 2006; Fichtner and Trampert, 2011), the construction of these sensitivity kernels involves the interaction between the forward wavefield from the survey sources with the backward-modelled adjoint residuals.

Before outlining the decomposition proposed here and the reasons behind it we make a brief overview on the main topics on FWI related to the subjects discussed in the coming chapters.

## 1.1 FWI overview

Two general approaches can be used when dealing with nonlinear inversion methods: global optimisation or local iterative descent methods.

Global optimisation techniques, such as Monte Carlo, although able to avoid local-minima, have limited application in solving seismic inversion problems due to the huge amount of parameters to be determined (Gauthier et al., 1986). As a matter of fact, even with the great improvement on the computational capabilities of the Oil and Gas Industry in the last thirty years, we are far from seeing this methods applied successfully.

Lailly (1983) and Tarantola (1984a) presented a nonlinear gradient- based local optimisation which is considered the cornerstone of the least- squares based inversion in the geophysics community. Besides the techniques itself, these works are important because they showed the relation of gradient- based methods with the well known two-way wave equation reverse-time migration (RTM). The main goal of nonlinear inversion is to deal with not only primary reflections but with the full information content in the seismogram such as diving waves, supercritical reflections, refractions and multiple-scattered waves.

While in principle capable of handling all aspects of wave propagation that influence the data, including full nonlinearity, in practice nonlinear gradient-based FWI is limited by its notorious sensitivity to the choice of the starting model. This is so because for short-offset acquisition of reflection data, the seismic wavefield is rather insensitive to high/intermediate wavelengths (Gauthier et al., 1986; Mora, 1987) which prevents the gradient-based local optimisation from circumventing the existence of innumerable local minima in the usual highly non-convex objective function. Thus, the initial model, usually assumed to be smooth, must, at least, kinematically explain the observed data.



## 1.1 FWI overview

---

### 1.1.1 Forward Modelling

Modelling plays a key role in FWI. The main steps of gradient-based inversion rely critically in the ability of the modelling scheme to predict all non-linear events taking place in a geologically complex area. Two opposing forces come into play when choice must be made: accuracy and computational cost.

Three main questions must be answered before choosing the modeller: which physics theory behind wave equation will be used to approximate the real wave propagation phenomena in the subsurface, which numerical discretisation method better fits the problem, and in which domain we will implement the method. All three aspects are interconnected.

As FWI intent to use the full seismic information recorded, the physical theory describing the modelled wavefield is very important. If, for instance, one chooses acoustic wave equation to describe the wavefield propagation, information about the subsurface carried by shear and converted waves would not be put to use, not helping or even leading the inversion process to a wrong solution. On the other hand, the higher the complexity of the physical theory describing the wavefield propagation (no doubt the visco-elastic model would do a better job than the acoustic approximation), the higher the number of variables. In Section 1.1.5, we comment on the challenges that multiparameter inversion brings up. Nowadays, 2D elastic and 3D acoustic inversion are the most common approach.

Among the main discretisation methods, finite-difference is, no doubt, the most popular one although other more sophisticated methods such as finite-element or finite-volume may be used. Marfurt (1984) quantitatively compares finite-difference and finite-element solutions of the scalar and elastic hyperbolic wave equations for the most popular implicit and explicit finite-difference and finite-element methods. Infrastructural criteria (such as computational cost, accuracy, available machine and memory, flexibility) and model-related criteria (Poisson's ratio, medium and/or mesh inhomogeneity, structural complex boundary conditions) must be taken into account when choosing the discretisation to be used.

#### **Time domain *versus* frequency domain**

Along with the choice of discretisation method, we also must define its domain. Although time-domain and frequency-domain methods are analytically equivalent – to some extent the non-specialist (or end user) need not be concerned in which domain the results are obtained (Pratt, 1999) – this choice depends on the size of the problem and on the available computational power.

Both early theoretical developments and practical applications were done in time domain (Taratola, 1984a; Gauthier et al., 1986; Mora, 1987). Pratt and Worthington (1990) first brought the development of inversion in the frequency domain, and Pratt et al. (1996) showed one of the first real data application of this approach.

The frequency-domain implementation has shown important advantages over time-domain approach:

- It introduces a natural multiscale approach by progressing from the lowest frequencies to the higher ones which helps to mitigate the notorious non-linearity of the inverse problem (Pratt et al., 1996; Bunks et al., 1995; Virieux and Operto, 2009).
- It allows to decimate the data by performing the inversion with a limited number of frequency components of wide-aperture/wide-azimuth surveys which helps to manage and process smaller volumes of data with large number of sources (Pratt et al., 1996; Sirgue and Pratt, 2004)

Beside these advantages which are independent of the discretisation method used, we can identify other advantages when finite-difference methods are applied in the frequency domain. According to Plessix (2007), in 2D problems the linear system resulting from the frequency-domain scheme can be solved with a direct solver, and the frequency-domain multishot migration algorithm is about one order of magnitude faster than its time-domain counterpart, because the LU decomposition of the matrix of the system has to be performed only once per frequency.

On the other hand, with realistic 3D problems, the linear system from a frequency-domain implementation may be difficult to solve with a direct solver, because of the high memory requirements. Plessix (2007) discuss many possible alternative approaches to solve large 3D problems in the frequency domain and evaluate the result given by Helmholtz preconditioned iterative solver.

As a matter of fact, the 3D problems brought the time-domain implementations back into use. The time-domain method is expensive: the forward and time-reversed wavefield propagation may have high computational cost depending on the time interval needed to avoid dispersion and aliasing. However, as pointed out by Vigh et al. (2009), the fact that cluster computers are getting more and more efficient for most processing steps, including time-domain RTM, indicates that it is feasible to perform time-domain waveform inversion of large 3D data sets on current hardware.

Preconditioning is another issue that influences the choice of the inversion domain. According to the domain, specific methodologies can be applied to precondition either the data residuals or the

## 1.1 FWI overview

---

gradient, so that they improve the convergence and/or the linearity of the inverse problem (Shipp and Singh, 2002).

The main advantages frequency-domain implementations naturally introduce (multiscale approach and frequency decimation) can be achieved by time-domain approaches, if one, for instance, extracts the monochromatic solution by discrete Fourier transform over the time steps (Sirgue et al., 2008). This permits the data to be decimated by choosing an arbitrary number of frequencies, allowing the multiscale approach to be applied.

### 1.1.2 Multiscale approach

It is well known that the main theoretical difficulty for nonlinear seismic inversion is the presence of numerous local minima in the objective function (Gauthier et al., 1986; Mora, 1987; Bunks et al., 1995; Pratt et al., 1996; Clement et al., 2001; Virieux and Operto, 2009; Asnaashari et al., 2013). These local minima prevent iterative techniques from finding the global minimum unless the initial model for the inverted parameter is already near enough from the global solution.

Among the various approaches developed to circumvent this problem, the proposal by Bunks et al. (1995) of decomposing the seismic inversion problem by scale proved itself very effective. It is based on the observation that the complexity of the objective function is directly related to the dominant length scale of the subsurface model. In other words, at smooth models where long wavelengths are dominant there are fewer local minima and those that remain are further apart from each other. Rough models with dominant short wavelengths generate rough objective functions with numerous local minima (see Figure 1.1). Using iterative inversion methods, one must start inverting with the long-period data used to constrain the long wavelength component of the model. The final model of this first step is used as initial model for the next round of iterations performed with higher frequency data, so that shorter wavelength structures are included in the model updates. This process may go on until the data with the highest frequency have been used in the inversion process.

Note that performing waveform inversion in a multiscale manner, lower frequencies can provide several scalable options, such as changing the shots in conjunction with incrementing the bin sizes in all directions (Vigh et al., 2009). Beside this, the forward modelling can benefit with the change of scale which also helps in decreasing the computational cost in the first stages of the inversion process.

There is one important thing to note: the success of the multi-scale approach is empirical,

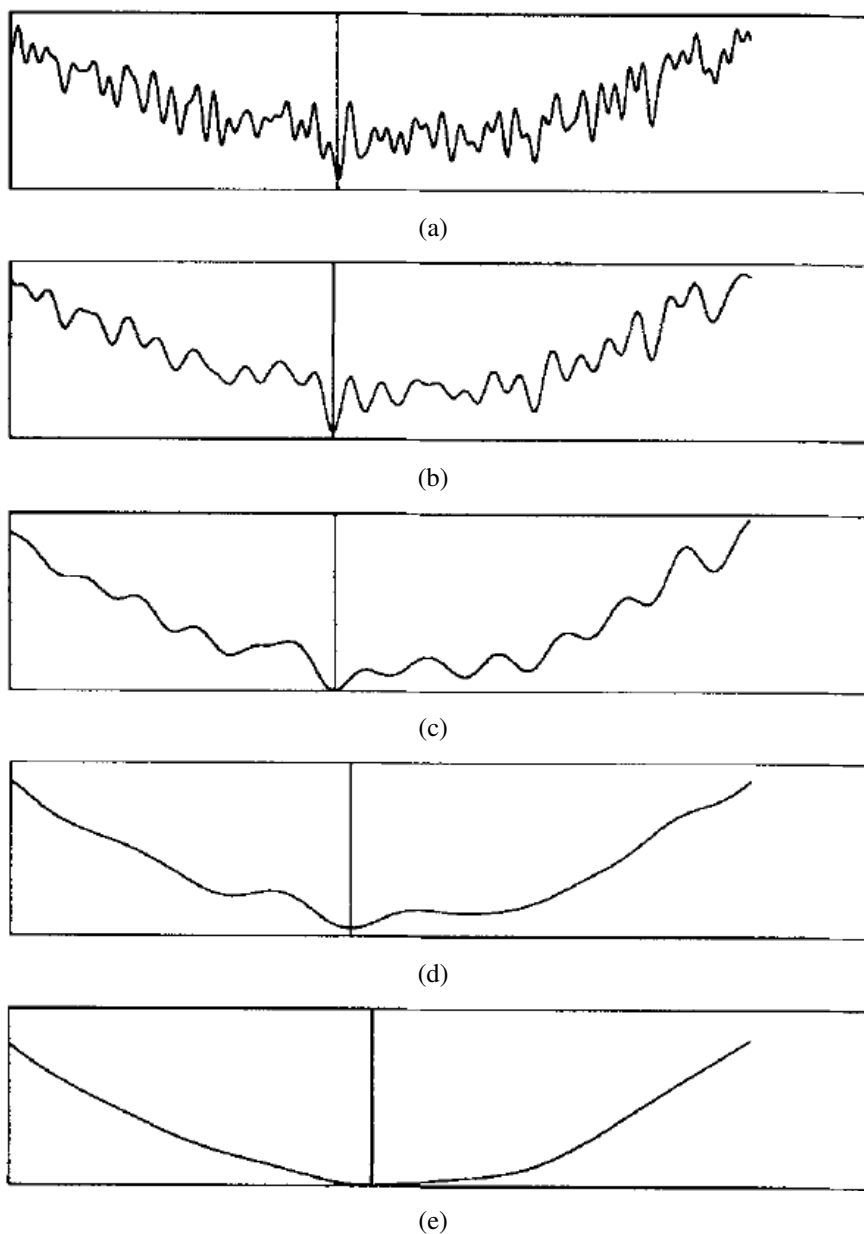


Figure 1.1: [Extracted fromBunks et al. (1995)] Heuristic illustration of the multigrid method applied to a nonlinear problem. Frame (a) illustrates a 1-D objective function and frames (b)-(e) illustrate the same objective function at ever increasing scale lengths. A gradient method applied to the shortest scale succeeds in finding the local minimum closest to the starting point, whereas the gradient method applied to the longest scale finds the global minimum (at that scale) regardless of the starting point. A fining up procedure helps find the estimate of the position of the global minimum for the fine scale problem.

## 1.1 FWI overview

---

meaning that there is no guarantee that it works under different circumstances. In other words, there is no theoretical proof that the transitions from one scale length to the other have the property that the global solution at one scale is in the convex neighbourhood of the global minimum at the next scale (Bunks et al., 1995; Fichtner, 2010).

### 1.1.3 Optimisation methods

As pointed out before, due to the huge size of the seismic inversion problems, the nonlinear minimisation of the objective function must rely on gradient- based iterative methods also known as *descent methods*. We do not intent to make a detailed review of all the known descent methods but to give a brief overview of some of them. For further, deeper references the reader is referred to Fichtner (2010) and Luemberger and Ye (2010).

#### General descent method

For this brief overview we are going to use the notation found in Fichtner (2010). Let us consider a model  $\mathbf{m}$  and as objective function  $\chi(\mathbf{m})$  that must be minimised. The idea behind iterative optimisation method is, starting from the best known initial model  $\mathbf{m}_0$ , to successively update the current model  $\mathbf{m}_i$  to a new model  $\mathbf{m}_{i+1}$  such that

$$\mathbf{m}_{i+1} = \mathbf{m}_i + \gamma_i \mathbf{h}_i \text{ with } \chi(\mathbf{m}_{i+1}) < \chi(\mathbf{m}_i), \quad (1.1)$$

where the particular choice of the *descent direction*,  $\mathbf{h}_i$ , and the *step length*,  $\gamma_i > 0$  depends on the minimisation scheme and are not arbitrary. The condition  $\chi(\mathbf{m}_{i+1}) < \chi(\mathbf{m}_i)$  is necessary but not sufficient to guarantee convergence. In this matter the choice of the step length is crucial.

In the general case, the descent direction  $\mathbf{h}_i$  is given by

$$\mathbf{h}_i = -\mathbf{A} \cdot \nabla_{\mathbf{m}} \chi(\mathbf{m}_i), \quad (1.2)$$

where  $\mathbf{A}$  is a positive definite matrix whose choice depends on the descent method. Therefore, the key point of any descent method is the evaluation of the gradient  $\nabla_{\mathbf{m}} \chi(\mathbf{m})$ . The most efficient way of doing it is using the *Adjoint-state method* (see Section 1.1.4).

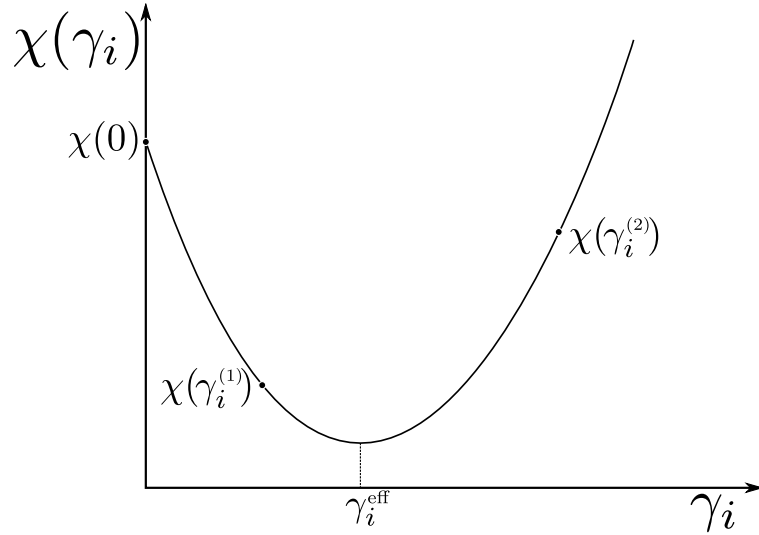


Figure 1.2: Line search with three trial step lengths and a second-order polynomial interpolation. The parabola vertex defines the efficient step length.

### Step length and line search

Besides the gradient evaluation, finding an optimal/efficient step length is also very important. It is this choice that ensures the convergence process and the correctness regarding unit dimensionality.

The optimal value for  $\gamma_i$  in a convex minimisation problem is given by

$$\gamma_i = \frac{\mathbf{h}_i \cdot \nabla_m \chi(\mathbf{m}_i)}{\mathbf{h}_i \cdot \mathbf{H}_\chi(\mathbf{m}_i) \cdot \mathbf{h}_i}, \quad (1.3)$$

where  $\mathbf{H}_\chi(\mathbf{m}_i)$  is the Hessian of the objective function  $\chi$ .

An alternative way to evaluate the step length is usually used to avoid the expensive task of evaluating the Hessian: *the line search*. The idea behind this procedure is simple (see Figure 1.2): once given the current model  $\mathbf{m}_i$  and the descent direction  $\mathbf{h}_i$  at the  $i$ -th iteration, we must evaluate the objective function  $\chi(\mathbf{m}_i + \gamma_i^{(k)} \mathbf{h}_i)$  for some trial step lengths,  $\gamma_i^{(k)}$  ( $k = 1, \dots, n$ ), and use this points to interpolate a polynomial of degree  $n - 1$  as an approximation of  $\chi(\mathbf{m}_i + \gamma_i \mathbf{h}_i)$ . The minimum of this polynomial defines the efficient step length  $\gamma_i^{\text{eff}}$ .



## 1.1 FWI overview

---

### Some well-established descent methods

As said before, the main difference between descent methods is how the descent direction  $\mathbf{h}_i$  is defined/evaluated.

**Steepest descent.** In the steepest descent method we define  $\mathbf{h}_i = -\nabla_m \chi(\mathbf{m}_i)$ , i.e, the steepest descent direction is in the opposite direction of the gradient of the objective function. Or, put in another way, one must define the matrix  $\mathbf{A}$  as the identity. The step length is defined through a line search.

**Newton's method.** In this method, both first- and second-derivative information are used to define the descent direction what significantly improving the convergence speed given that the initial model is near enough of the (supposedly) global minimum. The descent direction is determined by the solution of the linear system

$$\mathbf{H}_\chi(\mathbf{m}_i) \cdot \mathbf{h}_i = -\nabla_m \chi(\mathbf{m}_i), \quad (1.4)$$

where the Hessian  $\mathbf{H}_\chi(\mathbf{m}_i)$  brings the objective function second- derivative information. Note that this is equivalent to choosing the matrix  $\mathbf{A}$  from eq. (1.2) as

$$\mathbf{A} = \mathbf{H}_\chi^{-1}(\mathbf{m}_i). \quad (1.5)$$

The model update is, then, given by  $\mathbf{m}_{i+1} = \mathbf{m}_i + \mathbf{h}_i$  ( $\gamma_i = 1$ ).

**Regularised Newton's method.** Usually, the Hessian is nearly singular what requires some kind of regularisation for its inversion. Hence, the model update in regularised Newton's method is done by:

$$\mathbf{m}_{i+1} = \mathbf{m}_i - [\mathbf{H}_\chi(\mathbf{m}_i) + \beta \mathbf{I}]^{-1} \cdot \nabla_m \chi(\mathbf{m}_i), \quad (1.6)$$

where the parameter  $\beta$  can be found through a line search. Note that for big values of  $\beta$  this methods tends to behave as the steepest descent, while, for small values of  $\beta$  the Hessian tends to be dominant which makes the method closer to Newton's.

**Gauss-Newton method.** Variations on the Newton's method are commonly used to avoid the expensive cost of evaluating the full Hessian when dealing with non-linear least-square problems.

One of these is the Gauss–Newton method which uses the *approximate Hessian*. For instance, as seen in Fichtner and Trampert (2011), for the least-square objective function

$$\chi(\mathbf{d}(\mathbf{m})) = \frac{1}{2}(\mathbf{d}(\mathbf{m}) - \mathbf{d}_0)^T(\mathbf{d}(\mathbf{m}) - \mathbf{d}_0), \quad (1.7)$$

with synthetic data  $\mathbf{d}$  and observed data  $\mathbf{d}_0$ , the Hessian matrix is given by

$$\mathbf{H}_\chi = \mathbf{G}^T \cdot \mathbf{G} + (\mathbf{d}(\mathbf{m}) - \mathbf{d}_0)^T \cdot \nabla_m \mathbf{G}, \quad (1.8)$$

here  $\mathbf{G} = \nabla_m \mathbf{d}$  is the jacobian matrix of  $\mathbf{d}$ . In the Gauss-Newton method we use the inverse of the approximate Hessian  $\tilde{\mathbf{H}}_\chi = \mathbf{G}^T \cdot \mathbf{G}$  to update the model in this method:

$$\mathbf{m}_{i+1} = \mathbf{m}_i - \tilde{\mathbf{H}}_\chi^{-1} \cdot \nabla_m \chi(\mathbf{m}_i), \quad (1.9)$$

Pratt et al. (1998) presents a very instructive comparison between three gradient- based optimisations techniques presented here: steepest descent, Gauss-Newton and Newton’s methods. He developed the formulation of these methods in the discretised frequency-space domain using a finite-difference approach.

**Gauss–Newton method variations: Levenberg and Levenberg–Marquardt methods.** We can use regularised variations of the Gauss–Newton method. The *Levenberg method* uses the identity matrix for regularisation:

$$\mathbf{m}_{i+1} = \mathbf{m}_i - \left[ \tilde{\mathbf{H}}_\chi(\mathbf{m}_i) + \beta \mathbf{I} \right]^{-1} \cdot \nabla_m \chi(\mathbf{m}_i). \quad (1.10)$$

In the Levenberg–Marquardt method the identity matrix is replaced by the diagonal of the approximate Hessian:

$$\mathbf{m}_{i+1} = \mathbf{m}_i - \left[ \tilde{\mathbf{H}}_\chi(\mathbf{m}_i) + \beta \text{diag} \tilde{\mathbf{H}}_\chi(\mathbf{m}_i) \right]^{-1} \cdot \nabla_m \chi(\mathbf{m}_i). \quad (1.11)$$

**Conjugate-gradient methods.** The steepest descent method may suffer from slow convergence. The successive descent directions may not be linear independent (or even parallel indeed) causing the minimisation path to “zig-zag”, increasing the number of iterations in the minimisation process to reach a point close enough to the minimum (see Figure 1.3). On the other hand, Newton’s

## 1.1 FWI overview

---

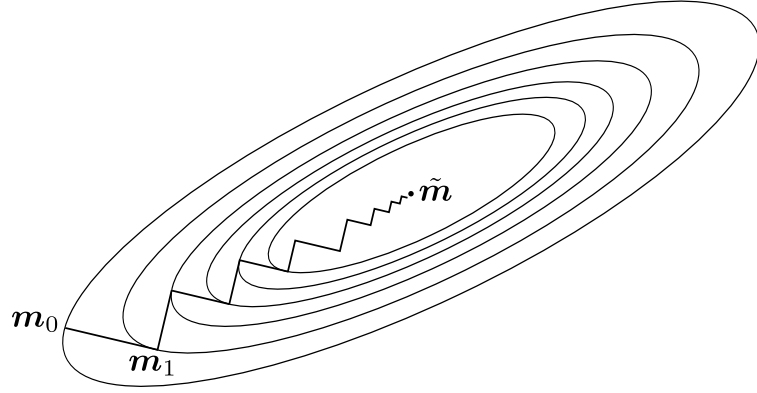


Figure 1.3: Convergence in the steepest descent method for a quadratic objective function.

method can accelerate the convergence if the initial model is close enough to the minimal solution but at the high expense of evaluating and storing the Hessian.

The *conjugate-gradient (C-G) methods* can be considered as being somewhat intermediate between the steepest descent method and the Newton's method. The C-G method was originally designed and analysed for the quadratic problem where the Hessian is a constant positive definite matrix, but extensions and approximations have been made to deal with more general nonlinear problems (Hestenes and Stiefel, 1952; Fletcher and Reeves, 1964; Polak and Ribiere, 1969).

The idea behind this method is to build descent directions that satisfy the orthogonality conditions below:

$$\mathbf{h}_i \cdot \nabla_m \chi(\mathbf{m}_j) = 0, \quad (1.12)$$

$$\mathbf{h}_i \cdot \mathbf{H}_\chi \cdot \mathbf{h}_j = 0, \quad (1.13)$$

for any  $0 \leq i < j \leq n$ , where  $n$  is the dimension of the model space. Eq. (1.12) states that the gradient of the objective function at the current model has to be orthogonal to all previous descent directions. Eq. (1.13) states that the current descent direction must be  $\mathbf{H}_\chi$ -orthogonal to all the previous descent directions. In other words, the directions  $\mathbf{h}_i$  and  $\mathbf{h}_j$  are *conjugate with respect to*  $\mathbf{H}_\chi$ . These relations guarantee that, in a pure quadratic problem, the solution would be found in, at most,  $n$  iterations.

The C-G method is based on the update of the model according to eq. (1.1)

$$\mathbf{m}_{i+1} = \mathbf{m}_i + \gamma_i \mathbf{h}_i, \quad (1.14)$$

where the first descent direction is given by the steepest one,  $\mathbf{h}_0 = -\nabla_m \chi(\mathbf{m}_0)$ . The step length at each iteration is given by eq. (1.3) in the case of quadratic problem where the Hessian is known.

The key element, though, is how the next descent direction is evaluated. The next direction  $\mathbf{h}_{i+1}$  depends on the current direction  $\mathbf{h}_i$  and is given by

$$\mathbf{h}_{i+1} = -\nabla_m \chi(\mathbf{m}_{i+1}) + \beta_i \mathbf{h}_i, \quad (1.15)$$

where the parameter  $\beta_i$  is given by

$$\beta_i = \frac{\|\nabla_m \chi(\mathbf{m}_{i+1})\|_2^2}{\|\nabla_m \chi(\mathbf{m}_i)\|_2^2}. \quad (1.16)$$

In the more general case of a nonlinear objective function, the orthogonality relations stated in eqs. (1.12) and (1.13) do not hold and convergence may be achieved in more than  $n$  iterations. The process above can be applied in the general case with a line search to evaluate the step length  $\gamma_i$  since the Hessian is not easily available. In this case, it is commonly referred as the *Fletcher-Reeves method* (Fletcher and Reeves, 1964). One well-known variation of the method is the *Polak-Ribière method* (Polak and Ribiere, 1969) where  $\beta_i$  is defined as

$$\beta_i = \frac{\nabla_m \chi(\mathbf{m}_{i+1}) \cdot [\nabla_m \chi(\mathbf{m}_{i+1}) - \nabla_m \chi(\mathbf{m}_i)]}{\nabla_m \chi(\mathbf{m}_i) \cdot \nabla_m \chi(\mathbf{m}_i)}. \quad (1.17)$$

Another one is the *Hestenes-Stiefel method* (Hestenes and Stiefel, 1952) where  $\beta_i$  is given by

$$\beta_i = \frac{\nabla_m \chi(\mathbf{m}_{i+1}) \cdot [\nabla_m \chi(\mathbf{m}_{i+1}) - \nabla_m \chi(\mathbf{m}_i)]}{\mathbf{h}_i \cdot [\nabla_m \chi(\mathbf{m}_{i+1}) - \nabla_m \chi(\mathbf{m}_i)]}. \quad (1.18)$$

### 1.1.4 Adjoint-state methods

As seen above, all the descent optimisations techniques require the computation of the gradient of the objective function with respect to the model parameters in order to define the descent direction at each iteration.

Usually, the objective function depends on the model parameters via the state variables (in case of wave propagation, the wavefield  $\mathbf{u}$ ), i.e.,  $\chi(\mathbf{m}) = \chi(\mathbf{u}(\mathbf{m}))$ . It means that, due to the chain rule, the gradient of the objective function depends on the gradient of the wavefield with respect to the model parameter,  $\nabla_m \mathbf{u}$ . The latter gradient is known as the *Fréchet derivatives*, *Jacobian* or

## 1.1 FWI overview

---

*Sensitivity matrix.* The computation of this gradient may be very expensive.

The *adjoint-state method* is a mathematical tool that allows us to compute the gradient of an objective functional with respect to the model parameters very efficiently without the need of explicitly building the sensitivity matrix (Virieux and Operto, 2009; Fichtner, 2010). Its formulation can be developed with the help of the perturbation theory (Fichtner and Trampert, 2011) or with the Lagrange multiplier method for constrained optimisation (Bunks et al., 1995). Plessix (2006) presents a very good review of the adjoint state method and the development of its formulations.

It is not the scope of this introduction to develop the formalism of the adjoint state method. But to clarify the relation of this method with the our proposed decomposition we summarise one example of the application of time domain adjoint-state method for a constant-density acoustic problem found in Plessix (2006). In this case the wave operator is  $\mathcal{L} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2$ . With initial boundary conditions, the pressure wavefield  $p$  due to the source  $f$  satisfies:

$$p(\mathbf{x}, 0) = 0, \quad (1.19)$$

$$\frac{\partial p(\mathbf{x}, 0)}{\partial t} = 0, \quad (1.20)$$

$$\mathcal{L}[p(\mathbf{x}, t)] = f(\mathbf{x}, t). \quad (1.21)$$

The objective function is defined

$$\chi(m) = \frac{1}{2} \sum_{s,r} \int_0^T (S_{s,r} p(\mathbf{x}, t) - p_{s,r}^{\text{obs}}(\mathbf{x}, t))^2 dt, \quad (1.22)$$

where  $T$  is the observed time, and  $S_{s,r}$  is the restriction operator onto the receiver position. The model parameter is the squared slowness,  $m = \frac{1}{c^2}$ .

The adjoint-state system reads

$$q(\mathbf{x}, 0) = 0, \quad (1.23)$$

$$\frac{\partial q(\mathbf{x}, 0)}{\partial t} = 0, \quad (1.24)$$

$$\mathcal{L}[q(\mathbf{x}, t)] = \sum_r S_{s,r}^T (S_{s,r} p(\mathbf{x}, T - t) - p_{s,r}^{\text{obs}}(\mathbf{x}, T - t)). \quad (1.25)$$

$q$  satisfies the same wave equation that  $p$ , but with a different source term. In this case  $\mathcal{L}$  is self-

adjoint.  $S_{s,r}p(\mathbf{x}, T - t) - p_{s,r}^{\text{obs}}(\mathbf{x}, T - t)$  is the residual in reverse time. The gradient of  $\chi$  reads:

$$\frac{\partial \chi}{\partial m}(\mathbf{x}) = - \sum_s \int_0^T q(\mathbf{x}, T - t) \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} dt. \quad (1.26)$$

Generally speaking, the evaluation of the objective function gradient for waveform inversion with the aid of the adjoint state method may be summarised in three steps (Tromp et al., 2005):

- Modelling the direct (regular) wavefield from the source with the current model.
- Time-reversed modelling (backpropagation) of the residual between observed and modelled wavefield with the current model. The resulting wavefield is known as *adjoint wavefield*.
- Cross-correlate the appropriate temporal/spatial derivatives of the regular and adjoint wavefields.

Note that this procedure is very similar to that of RTM. The main difference is that the recorded wavefield is backpropagated in RTM instead of the wavefield residuals. Lailly (1983) was the first one to draw this connection between wave-equation migration and waveform inversion.

### 1.1.5 Challenges

Great theoretical and practical improvements have been achieved, since the first papers on FWI in the 80's. But many challenges still remain. If we could summarise all the challenges in just one sentence it would be: "Perform a 3D elastic FWI in realistic data set from a not so accurate initial model with reasonable time and computational cost."

### 3D

The main limitation regarding the 3D FWI is the the computational cost in processing and/or memory requirements of the forward modelling.

Some successful studies with 3D cases were made both with time-domain finite- difference methods (Sirgue et al., 2008) or frequency-domain finite-difference direct solvers (Ben-Hadj-Ali et al., 2008) for synthetic data. Differently from the 2D case where frequency-domain solvers proved to be the best option due to the smaller processing requirements, the extensive memory needs of the frequency-domain finite-difference modellers push towards the time-domain approach

## 1.1 FWI overview

---

for 3D problems (Sirgue et al., 2008). Efforts to circumvent the high memory costs are being made with the use of frequency-domain iterative solvers (Plessix, 2007, 2009).

Nevertheless, nowadays most of the 3D FWI implementations in industry are done in the time-domain, at least when it comes to modelling, seizing all the expertise and infrastructure (software and hardware) already available for RTM.

### **Monoparameter *versus* multiparameter FWI**

Due to the current computational limitations, only monoparameter acoustic FWI is feasible when dealing with real-sized 3D problems (Plessix, 2012).

Such underlying physical assumption definitely restrain all the potential high resolution that FWI can offer. According to Barnes and Charara (2009), acoustic full-waveform inversion is applicable only when the S-wave velocity and the density fields are smooth enough to reduce the amplitude-versus-offset (AVO) effect, or when the near-offset seismograms are inverted with a good starting model. This excludes the most realistic cases, specially marine reflection surveys where wide-aperture seismic data is a common place and significant AVO and azimuthal anisotropic effects should be observed in the data (Virieux and Operto, 2009). As mentioned before, the seismic modelling used in the FWI algorithm should, in principle, honour all the physics of wave propagation because.

When dealing with more than one parameter class, the ill-posedness of the inversion problem increases either because more degrees of freedom come into play, or because the sensitivity of the inversion with respect to each parameter may vary, or still because these parameters can be more or less coupled (Virieux and Operto, 2009). Studying the radiation pattern of each parameter class as a function of the aperture angle (Wu and Aki, 1985; Tarantola, 1986) can help understanding the parameters coupling in order to devise the best set of parameter class to be used in the inversion. This will also help to design hierarchical strategies to reduce the ill-posedness of FWI.

Nevertheless, multiparameter FWI with extensions to elasticity, attenuation and anisotropy remains challenging (Plessix, 2012; Virieux et al., 2009). Most of the successful examples of elastic FWI found in the literature are 2D cases (see, e.g., Crase et al., 1990; Shipp and Singh, 2002; Sears et al., 2008; Brossier and Operto, 2009). Either in acoustic or elastic assumption, density proved itself a parameter difficult to reconstruct. Some works involving density inversion are found in literature (Choi et al., 2008; Bae et al., 2012; Jeong et al., 2012). Reconstruction of anisotropic parameters by FWI is probably one of the most undeveloped and challenging fields of investigation



(Virieux and Operto, 2009). Some studies on this subjects were published very recently (Warner et al., 2013; Kamath and Tsvankin, 2013).

### **New objective functions**

Classically, the objective function in FWI is based on the least-squared norm. This norm assumes that the noise distribution is Gaussian (Tarantola, 1987) which may not always be true. Crase et al. (1990) was the first one to drive a study on using more robust minimisation criteria in the inversion process to handle many types of noise. After that many others tried to use alternative minimisation criteria such as the Huber norm (Guitton and Symes, 2003; Ha et al., 2009) and Logarithmic objective functions (Shin and Min, 2006; Shin et al., 2007; Bednar et al., 2007; Pyun et al., 2007).

### **Initial model construction**

The ultimate goal of FWI would be to obtain a high definition final model from a poorly detailed initial model. Unfortunately, even with all the development in acquisition technology which provide denser data, especially at low frequencies, we are far from that. In this way, an approach that simultaneously inverts for the transmission, reflection and refraction data in complex terrains without a good initial model is still under research (Plessix, 2012).

It implies that before performing the full waveform inversion itself, we must build a initial model as accurate as possible. Nowadays, the most usual approaches to build a starting model are reflection tomography (Woodward et al., 2008), migration-based velocity analysis (Sava and Biondi, 2004; Symes, 2008), first-arrival traveltimes tomography (Hole, 1992), stereotomography (Lambaré, 2008), and Laplace- or Laplace-Fourier-domain inversion (Shin and Cha, 2008; Shin and Ha, 2008; Shin and Ho Cha, 2009).

## **1.2 Kernel decomposition**

To help addressing model convergence issues in FWI, we analyse the role of nonlinearity in the sensitivity kernels. We demonstrate that the above-mentioned insensitivity to high/intermediate wavelengths (Gauthier et al., 1986; Mora, 1987) is partly caused by the conventional choice of smooth initial models, considering all singular parts of the model as unknown (see Figure 1.4).

## 1.2 Kernel decomposition

---

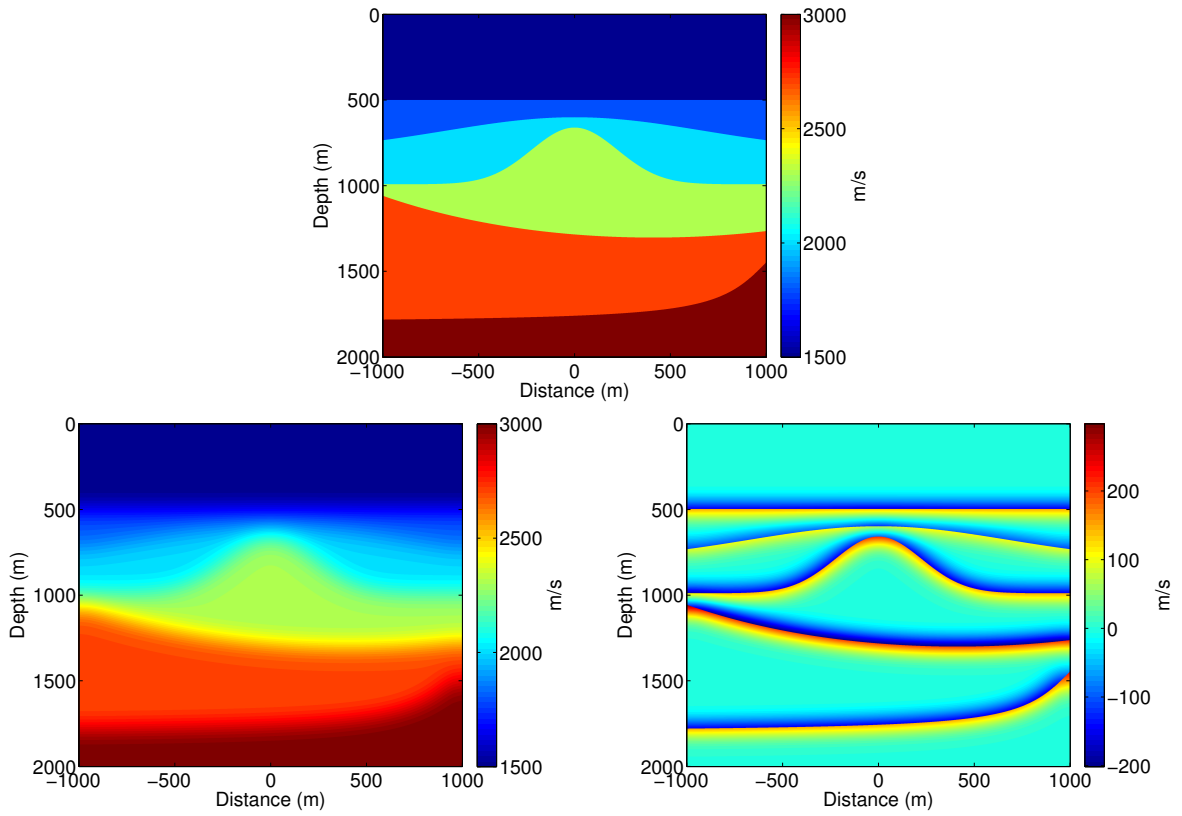


Figure 1.4: Conventionally, the exact model (top) is thought of as being decomposed into a smooth background (bottom left), supposedly known, and a singular scattering part (bottom right), to be determined under the assumption that it is small against the background.

To do so, we use a scattering-based approach (Vasconcelos et al., 2009). Assuming an acoustic medium, we reparameterise the subsurface model in terms of background and singular components for both bulk modulus and density. This leads to a decomposition of the data into a background wavefield that is sensitive only to the background model, and a singular wavefield sensitive to both background and singular model components. Both the background and singular parts of the model are then subject to perturbations that cause wavefield residuals (see Figure 1.5). This decomposition is motivated by the fact that conventional initial models, for example obtained from tomographic or migration-velocity-analysis methods, are only imperfect estimates of the true smooth background, but that initial estimates for the singular part are available, for example from migrated images or a priori geologic information. Focusing on the model backprojections from the singular data only, we provide expressions for the Fréchet-derivative sensitivity kernels for all four model parameters (perturbations of both smooth and singular parts of compressibility and density).

We demonstrate that the *forward* decomposition is successful in bringing out subkernels that unravel different levels of non-linearity with respect to data and model. This, in turn, can be translated into different levels of interaction between non-, single-, and multiple-scattered information that otherwise would be hidden in the undecomposed sensitivity kernels. Moreover, we predict that part of the answer to the problem of lacking low-frequency information on the model should lie in utilising scattered wavefields, because these travel through the medium long enough to carry this information (Snieder et al., 2002). Moreover, we show numerical results of using those subkernels to backproject the *scattered residual* only into model space and obtain model perturbation estimates.

We envision broad potential applications for the scattering-based FWI kernel decomposition we propose in this work. These include devising multiscale FWI algorithms, and improving velocity model building in the image domain using extended image (EI) gathers (Rickett and Sava, 2002; Sava and Fomel, 2003; Symes, 2008; Sava and Vasconcelos, 2009, 2011). As shown by Vasconcelos et al. (2009, 2010), there is a connection between extended image conditions and the interferometry formalism: the Extended Images behave like locally scattered wavefields in the image domain.

Important to say that we do not present nor perform an inversion in this thesis. In a fully iterative implementation, the gradient of the objective function is evaluated, a descent direction is defined, a step length is computed that minimises the objective function when the parameters are varied along the descent direction and, in this way a new model estimate is generated and used in the next iteration. As done in Pratt and Worthington (1990), we restrict the calculations to a single

## 1.2 Kernel decomposition

---

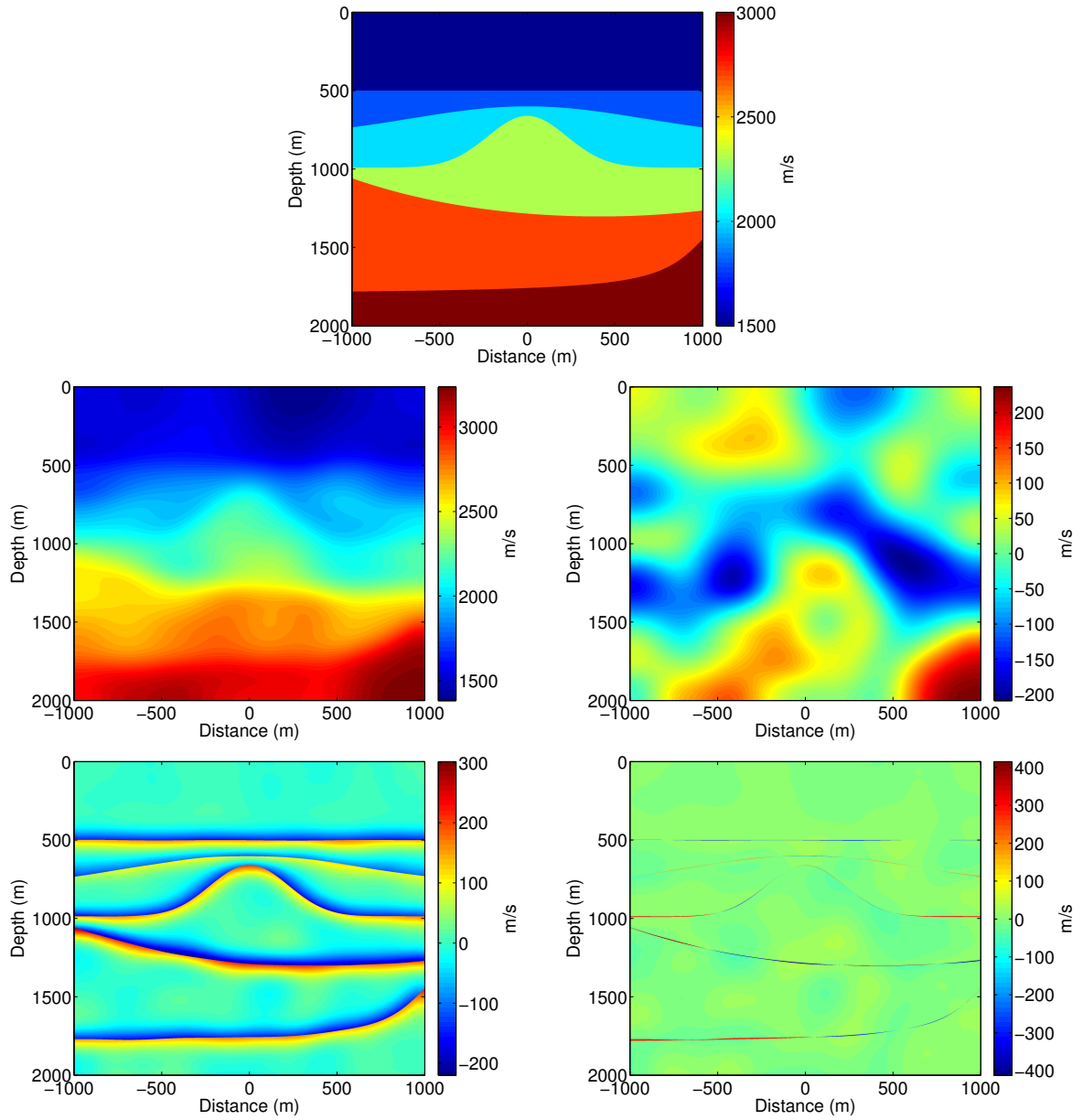


Figure 1.5: Here, the exact model (top) is thought of as being decomposed into a slightly wrong estimate of the smooth background (centre left), and a slightly wrong singular scattering part (bottom left), both supposed to be known, plus perturbations to the background (centre right) *and* the singular part (bottom right), to be determined under the assumption that they are small against their known estimates.

iteration only, evaluating and decomposing the sensitivity kernels which, as mentioned before, has connections with the objective function gradient.

The thesis is organised as follows. In Chapter 2 we present the mathematical development of the forward decomposition and discuss its physical meaning. We also present its adjoint formulation in order to obtain expressions of the decomposed estimates of model perturbations. When discussing the role of the scattering theory in our decomposition we developed a  $N$ -fold generalised decomposition of the complete wave operator in  $N - 1$  scattering potentials.

In Chapter 3, we perform numerical experiments to validate both the forward and the backprojection decompositions. Using different models and acquisition geometries, we gain further insight in the physical interpretation of the previous chapter which helped gaining insight in the possible application of the proposed decomposition.

In Chapter 4 we present an application of the decomposed subkernels in a time-lapse problem. In this examples we show how the enhanced illumination due to the backscattered waves can help delineating the time-lapse change. Chapter 5 summarises our conclusions and presents our final remarks.

Appendix A brings some mathematical definitions used in this thesis. Appendix ?? presents the derivation of the secondary sources for the singular-wavefield residuals, a key concept to the decompositions of the Fréchet derivatives kernels. In Appendix C, we developed the mathematical expressions for components of the bulk modulus and density perturbation estimates. Appendix D shows the result of the superposition of two related single-scattering subkernels and discuss its physical meaning. Finally, in Appendix B, we present the relation of the  $N$ -fold generalised decomposition in scattering theory with the one proposed here.

## 2. Kernel Decomposition

In this chapter, we present the mathematical derivation and physical interpretation of the proposed decomposition. Rather than a complex mathematical reinterpretation, our proposition requires from the reader an exercise of change of perspective when using the scattering theory in the inversion process. In an effort of making things clearer we avoided to use known terms with meanings different from the established one. Therefore, unusual but meaningful terms are used when referring to concepts and phenomena similar to well established ones (see Page xxix).

This chapter is organised as follows. In Section 2.1, we present the forward decomposition. To do so we first briefly review the conventional formulation of Tarantola (1984a) and relate this to the classical scattering theory (Section 2.1.1). Next, we generalise the conventional model and wavefield decomposition to an  $N$ -fold decomposition (Section ??), before entering the topic of our proposed sensitivity-kernel decomposition with two perturbed parts (Section 2.1.2).

Section 2.2 is devoted to the physical interpretation and discussion of the result from the previous section. We discuss the new insights our approach allows (Section 2.2.1), in particular regarding the role of multiple internal reflections/scattering in the decomposition (Section 2.2.2), and we identify the individual contributions of background and singular perturbations to the total singular-wavefield residual (Section 2.2.3).

In Section 2.3, we use the adjoint of the subkernels derived in the previous section to obtain expressions for the model-perturbation estimates.

### 2.1 Full-waveform inversion and sensitivity kernels

Sensitivity kernels are the core of the iterative gradient-based FWI schemes. They are used to backproject the difference between the observed and modeled data (residuals) into the model space in order to obtain parameter estimates used to iteratively update from a starting model.

### 2.1.1 Conventional formulation

In the conventional formulation of FWI, the residual is considered as a wavefield scattered by unknown perturbations in the known reference model. If these perturbations are considered small, linearisation is possible. In that case, one can find an expression for the Fréchet derivatives of the wavefield with respect to the model parameters which leads to the sensitivity kernel.

#### Wave equation

The acoustic wave equation describes the propagation of a pressure wavefield throughout a medium characterised by two model parameters, the bulk modulus, and density. For a point source and given the initial conditions, the acoustic wavefield  $p$  is uniquely defined by

$$\mathcal{L}[p(\mathbf{x}, t; \mathbf{x}_s)] = \delta(\mathbf{x} - \mathbf{x}_s)S(t), \quad (2.1)$$

where the right-hand side is the source term, with  $\mathbf{x}_s$  denoting the source position and  $S(t)$  describing the source wavelet. On the left-hand side of eq. (2.1), the acoustic wave-equation differential operator  $\mathcal{L}$ , also simply referred to as the wave operator, is given by

$$\mathcal{L} = \frac{1}{K(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\rho(\mathbf{x})} \nabla \right). \quad (2.2)$$

Here the bulk modulus  $K$  and the density  $\rho$  are the model parameters that constitute the model parameter vector  $\mathbf{m}$ . In most parts of the theoretical description we will use the closed form of eq. (2.1). In this way, the general considerations apply in the same way to any other form of the wave equation. For specific examples and the numerical tests, we will use the explicit form of the acoustic wave equation.

The solution of eq. (2.1) constitutes a non-linear relation with respect to the model parameters that can be represented as

$$p = f(\mathbf{m}), \quad (2.3)$$

where  $f$  denotes the functional relating the wavefield  $p$  and the model vector  $\mathbf{m}$ . Ideally,  $f$  depends on the model parameters only, but in practice, this functional also depends on the method used to solve the forward problem.



## 2.1 Full-waveform inversion and sensitivity kernels

---

### Scattering theory: medium and wavefield decomposition

In the classical scattering theory (Rodberg and Thaler, 1967; Newton, 1982; Bleistein et al., 2001; Stolt and Weglein, 2012), the total wavefield  $p$  propagating in a medium with bulk modulus  $K$  and density  $\rho$  is interpreted as being composed of two components: a reference wavefield,  $p_0$ , and a scattered wavefield,  $p_s$ , such that  $p = p_0 + p_s$ . The reference wavefield  $p_0$  is generated by the original source  $\delta(\mathbf{x} - \mathbf{x}_s)S(t)$ , but propagates in a known (usually assumed smooth) reference medium, described by bulk modulus  $K_0$  and density  $\rho_0$ . The scattered wavefield  $p_s$  is generated when the reference wavefield runs into the unknown (scattering) part of the medium, described by  $K_s$  and  $\rho_s$ , satisfying  $K = K_0 + K_s$  and  $\rho = \rho_0 + \rho_s$ . In most practical applications, it is assumed that the known part of the medium contains only low-frequency information, i.e., no sharp contrasts, reflectors and/or diffractors. On the other hand, the unknown part contains the complete high-frequency information.

This medium and wavefield decomposition transforms equation (2.1) into the following set of differential equations,

$$\mathcal{L}_0 [p_0(\mathbf{x}, t; \mathbf{x}_s)] = \delta(\mathbf{x} - \mathbf{x}_s)S(t), \quad (2.4)$$

$$\mathcal{L} [p_s(\mathbf{x}, t; \mathbf{x}_s)] = -\mathcal{V} [p_0(\mathbf{x}, t; \mathbf{x}_s)], \quad (2.5)$$

where

$$\mathcal{L}_0 = \frac{1}{K_0(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\rho_0(\mathbf{x})} \nabla \right) \quad (2.6)$$

is the wave operator in the reference medium, and where the difference between the wave operators in the total and reference models,

$$\mathcal{V} = \mathcal{L} - \mathcal{L}_0, \quad (2.7)$$

is known as the scattering potential. For brevity, we will refer to  $\mathcal{L}$  and  $\mathcal{L}_0$  as the total and reference wave operators, respectively. Note that the source term in eq. (2.4) is the original one of eq. (2.1), while the one in eq. (2.5) describes secondary sources excited by the reference wavefield  $p_0$ . These secondary sources are non-zero only at points where  $K_s \neq 0$  and/or  $\rho_s \neq 0$ .

Alternatively, eq. (2.5) can be expressed as

$$\mathcal{L}_0 [p_s(\mathbf{x}, t; \mathbf{x}_s)] = -\mathcal{V} [p_0(\mathbf{x}, t; \mathbf{x}_s) + p_s(\mathbf{x}, t; \mathbf{x}_s)], \quad (2.8)$$

with the physical interpretation that the illuminated contrasts that act as the scattered-wavefield

sources are illuminated by the total wavefield to generate a scattered wavefield propagating in the reference medium.

Eq. (2.8) is nothing more than a wave-equation describing the propagation of the scattered wavefield. The secondary sources can be forward propagated with the help of the Green's function of eq. (2.4),  $G_0(\mathbf{x}, t; \mathbf{x}_s)$ . In other words, with the knowledge of this Green's function in the reference medium, the scattered wavefield is described by

$$p_s(\mathbf{x}, t; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' G_0(\mathbf{x}, t; \mathbf{x}') * \mathcal{V} [p_0(\mathbf{x}', t; \mathbf{x}_s) + p_s(\mathbf{x}', t; \mathbf{x}_s)] . \quad (2.9)$$

Here, the symbol  $*$  denotes time-convolution, and  $\mathbb{V}$  is a volume that contains all scatterers. Eq. (2.9) is an exact integral equation for the scattered wavefield  $p_s$ . The assumption  $p_s \ll p_0$  at all points  $\mathbf{x}'$  within volume  $\mathbb{V}$  leads to the Born approximation, which replaces the total wavefield in the integrand by the reference wavefield to approximate  $p_s(\mathbf{x}, t; \mathbf{x}_s)$  by its single-scattering contributions.

As mentioned above, in most practical applications the reference model is taken to be a smooth background model. In that case, all high-frequency contributions to the model make part of the scattering potential. Note, however, that other interpretations of  $p_0$  and  $p_s$  are also possible without changing the theory above. By no means we are constrained to considering the reference and scattering potential as the low and high spatial frequency components of the subsurface model. Actually, this decomposition can be completely arbitrary: once the reference model is defined, the scattering potential is set by its definition – eq. (2.7) – and the wavefields are determined by eqs. (2.4) and (2.5). In the same way, there are no restrictions regarding the order of magnitude of the scattering potential with respect to the reference model and, therefore, of the scattering wavefield with respect to the reference wavefield. Obviously, this is not true when approximations are done (see Section 2.1.2).

### **Linearisation and sensitivity kernel representation**

The most common way to make practical use of the classical medium and wavefield decomposition described in Section 2.1.1 is to assume that the scattering parts of the medium can be considered small perturbations to the known reference model. Then,  $p_s \ll p_0$  and the problem can be linearised (Tarantola, 1984a; Bleistein et al., 2001). To describe this in terms of eq. (2.3), let us denote these small perturbations in the model parameters by  $\delta \mathbf{m}$ , i.e.,  $\mathbf{m} = \mathbf{m}_0 + \delta \mathbf{m}$ . Then, we

## 2.1 Full-waveform inversion and sensitivity kernels

---

can write according to eq. (2.3)

$$p = p_0 + p_s = f(\mathbf{m}_0 + \delta\mathbf{m}) = f(\mathbf{m}_0) + \Phi \delta\mathbf{m} + O(\|\delta\mathbf{m}^2\|), \quad (2.10)$$

where  $\Phi$  are the first derivatives of the wavefield with respect to the model parameters, also known as Fréchet derivatives. If only first-order terms in the model perturbations are taken into account, the scattered wavefield is approximately equal to the linearised wavefield perturbation, i.e.,

$$p_s \approx \delta p = \Phi \delta\mathbf{m}. \quad (2.11)$$

The expression for  $\Phi$  is obtained with the help of *secondary* or *adjoint sources*. These can be understood as pseudo-sources that give rise to the wavefield residual, seen as a wavefield itself. The secondary sources are, of course, functions of the perturbations in the model parameters. Their expressions are obtained when small model parameter perturbations  $K_s \approx \delta K$  and  $\rho_s \approx \delta\rho$  are introduced in the wave operator  $\mathcal{L}$  of eq. (2.2). As shown by Tarantola (1984a), introducing these perturbations and retaining only first-order contributions yields

$$\mathcal{L}_0 [\delta p(\mathbf{x}, t; \mathbf{x}_s)] = -\delta\mathcal{L} [p_0(\mathbf{x}, t; \mathbf{x}_s)], \quad (2.12)$$

where the secondary potential,  $\delta\mathcal{L}$ , is defined for the acoustic wave equation as

$$\delta\mathcal{L} = -\left\{ \frac{\delta K(\mathbf{x})}{K_0^2(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{\delta\rho(\mathbf{x})}{\rho_0^2(\mathbf{x})} \nabla \right) \right\}. \quad (2.13)$$

Eq. (2.12) is nothing else but the linearised version of eqs. (2.5) or (2.8), i.e.,  $\delta\mathcal{L}$  of eq. (2.13) is the linearised version of operator  $\mathcal{V}$  in eq. (2.7). Thus, in full analogy to eq. (2.9), the linearised wavefield perturbation can be represented using the Green's function of eq. (2.4) as

$$\delta p(\mathbf{x}_g, t; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3\mathbf{x}' G_0(\mathbf{x}_g, t; \mathbf{x}') * \delta\mathcal{L} [p_0(\mathbf{x}', t; \mathbf{x}_s)], \quad (2.14)$$

In other words, this equation represents the Born approximation mentioned above.

The discretised version of eq. (2.14) can be written in the frequency domain as

$$\hat{\delta p} = \hat{\Phi} \delta\mathbf{m} = \begin{bmatrix} \bar{U} & \bar{V} \end{bmatrix} \begin{bmatrix} \frac{\delta\bar{K}}{\delta\rho} \end{bmatrix} = \bar{U} \delta\bar{K} + \bar{V} \delta\bar{\rho} = \hat{\delta p}^K + \hat{\delta p}^\rho. \quad (2.15)$$

This equation defines the *Fréchet derivatives* of the total wavefield with respect to bulk modulus and density,  $\bar{U}$  and  $\bar{V}$ , respectively. Superscripts  $K$  and  $\rho$  denote the bulk-modulus and density contributions. Moreover, a hat indicates that the variable is in the frequency domain and the overline characterises a discretised scalar quantity (see Appendix A for further details).

The components of eq. (2.14) express the wavefield perturbation at a receiver at  $\mathbf{x}_g$  due to bulk-modulus and density perturbations. For example, in the time domain the bulk-modulus contribution reads explicitly

$$\delta p^K(\mathbf{x}_g, t; \mathbf{x}_s) = \int_{\mathbb{V}} d^3\mathbf{x}' \left[ \frac{1}{K_0^2(\mathbf{x}')} G_0(\mathbf{x}', t; \mathbf{x}_g) * \frac{\partial^2 p_0}{\partial t^2}(\mathbf{x}', t; \mathbf{x}_s) \right] \delta K(\mathbf{x}'), \quad (2.16)$$

where we have used the reciprocity property of the Green's function. In eq. (2.16), we can identify the expression in brackets as the integral-operator kernel of the bulk-modulus Fréchet derivative, i.e., the bulk-modulus sensitivity kernel for the source-receiver pair  $(\mathbf{x}_s, \mathbf{x}_g)$ , which is defined as volumetric density of the corresponding Fréchet derivative (Fichtner and Trampert, 2011)<sup>1</sup>. Its Fourier transform is  $\bar{U}$  in eq. (2.15). A similar procedure provides the corresponding expression for the Fréchet derivative and sensitivity kernel with respect to density.

As mentioned before, the separation of the model and wavefields is completely arbitrary. Therefore, the linearised treatment remains valid as long as the reference model is sufficiently close to the true model. This means, however, that validity problems may arise in practice because of the choice of the reference model. In complex geological settings, the conventional choice of a smooth version of the model as the reference model and all high-frequency contributions as the scattering part may fail, because these model parts, and consequently also the resulting wavefields, may be of the same order. In such a situation, the above linearisation is no longer valid. To proceed with a linearised inversion, we must decompose the model in other ways.

## 2.1.2 Decomposed perturbation

Note that the approach described above – considering the the unknown part of the model as the scattering potential and linearizing it to obtain an expression for the scattered wavefield which,

---

<sup>1</sup> Actually Fichtner and Trampert (2011) define the sensitivity kernel as the volumetric density of the Fréchet derivative of the objective function w.r.t. the model parameters. But if the  $L^2$ -norm of the residual is used as objective function its Fréchet derivative has direct relation with the Fréchet derivatives of the wavefield w.r.t the model parameters (Pratt et al., 1998). Thus, we extended the concept of sensitivity kernel to the volumetric density of the wavefield Fréchet derivative w.r.t the model parameters.

## 2.1 Full-waveform inversion and sensitivity kernels

---

thus, become the residual – is not the only way to use the scattering theory in FWI. An alternative way is to further decompose the reference and perturbation parts themselves. For the linearised inversion this is the simpler procedure because it allows to start from the linearised equations. Therefore, we will develop this type of decomposition in more detail below.

Actually, such a redecomposition could be done in as many parts as convenient. In Appendix B, Section B.1, we show that the decomposition described by eqs. (2.4) and (2.5) can be immediately generalised, because eq. (2.4) is of the same type as eq. (2.1) and can be decomposed in the same way into two equations and so on. Of these, we consider the most basic one, i.e., a decomposition into two components each (see Figure 2.1). The components of the reference model are, in principle, completely arbitrary for we fully known the reference model. For convenience, from an application point of view, we think of them as representing a (smooth) background model and an estimate of the singular part of the model. The two components of the perturbation part are the perturbations of these two model components. Correspondingly, we will have four wavefield components, being the background and singular-part wavefields and their respective perturbations. For the purpose of linearised inversions, we will consider the perturbations small in comparison to the reference components.

In principle, it is also possible to decompose the model into several parts according to eq. (B.8) and consider perturbations to *all* model and corresponding wavefield contributions. Such a generalisation to a perturbation of all terms to an  $N$ -fold decomposition is conceptually straightforward, but the resulting mathematical treatment is too extensive and will be left as a possible thesis follow up.

Moreover, the case of two medium and wavefield components is the most practical one. We often have access to initial approximations for both a background and a singular part of the medium, e.g., a smooth migration velocity model and the reflector positions in the corresponding migrated image. Then, the appropriate procedure is to invert for perturbations of these two model parts using the corresponding wavefield perturbations.

Therefore, we assume from now on that the reference model can be decomposed into background and singular parts as

$$K_0 = K_B + K_S \quad \text{and} \quad \rho_0 = \rho_B + \rho_S, \quad (2.17)$$

which are of the same order. Correspondingly, the reference wavefield will be separated into background and singular parts, i.e.,  $p_0 = p_B + p_S$ , also assumed to be of the same order. For

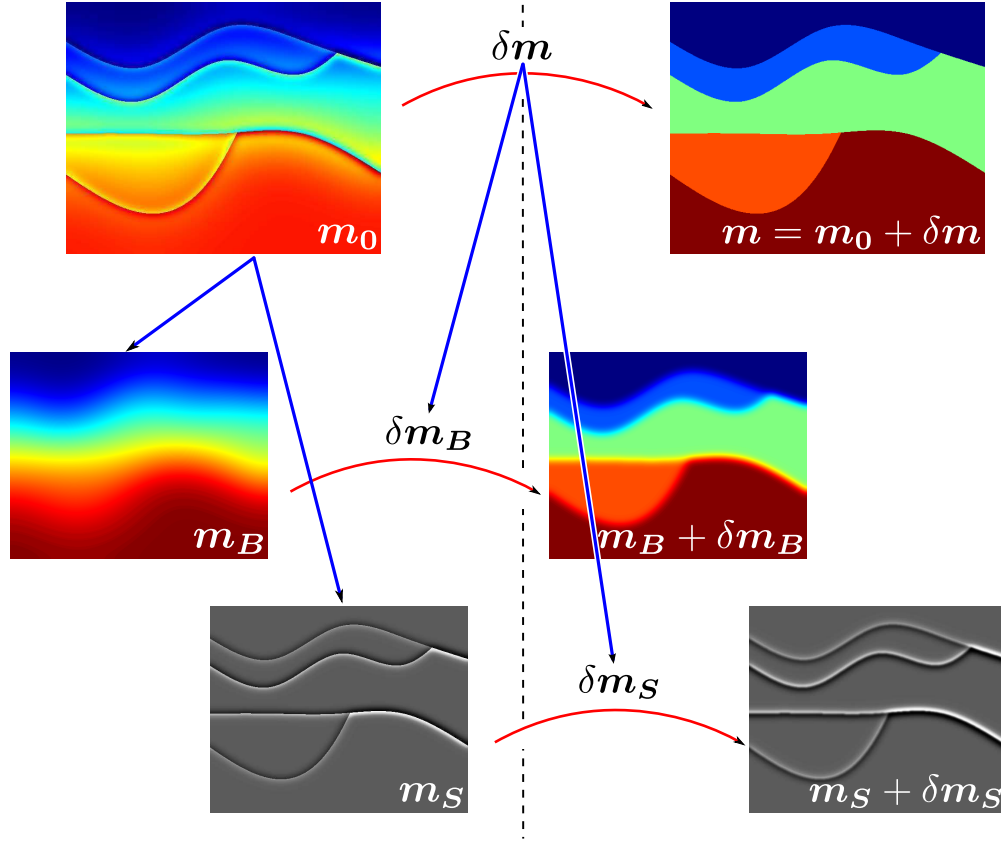


Figure 2.1: A sketch of the proposed decomposition. The components of the reference model are thought to represent a (smooth) background model and an estimate of the singular part of the model ( $m_0 = m_B + m_S$ ). The two components of the perturbation part are the perturbations of these two model components ( $\delta m = \delta m_B + \delta m_S$ ).

brevity, we will refer to the background and singular parts of the reference model and wavefield as the background and singular models and wavefields, respectively.

With decomposition (2.17), we are ready to investigate how perturbations of both parts of the reference model affect the set of equations in Section 2.1.1. In other words, we are interested in answering the following questions: Once perturbations are introduced in the background and/or in the singular model, what will be the resulting residuals  $\delta p_B$  and  $\delta p_S$  of the background and singular wavefields,  $p_B$  and  $p_S$ ? What are the sensitivity kernels for these wavefields? And what is the expression for the individual secondary sources for each of these wavefields?

After a similar analysis as outlined above for the perturbation of the undecomposed wavefield, we find that the background and singular wavefield residuals as a function of the new background

## 2.1 Full-waveform inversion and sensitivity kernels

---

and singular model perturbations can be written in discretised form in the frequency domain as

$$\begin{bmatrix} \hat{\delta p}_B \\ \hat{\delta p}_S \end{bmatrix} = \begin{bmatrix} \bar{U}_B & \bar{V}_B & 0 & 0 \\ \bar{U}_S^B & \bar{V}_S^B & \bar{U}_S^S & \bar{V}_S^S \end{bmatrix} \begin{bmatrix} \bar{\delta K}_B \\ \bar{\delta \rho}_B \\ \bar{\delta K}_S \\ \bar{\delta \rho}_S \end{bmatrix}. \quad (2.18)$$

Here,  $\bar{\delta K}_B$ ,  $\bar{\delta \rho}_B$ ,  $\bar{\delta K}_S$  and  $\bar{\delta \rho}_S$  are the discretised versions of the perturbations of the background model parameters  $\delta K_B$  and  $\delta \rho_B$ , and the perturbations of the singular model parameters  $\delta K_S$  and  $\delta \rho_S$ , which are defined in such a way that

$$\delta K = \delta K_B + \delta K_S \quad \text{and} \quad \delta \rho = \delta \rho_B + \delta \rho_S. \quad (2.19)$$

Moreover,  $\bar{U}_B$  and  $\bar{V}_B$  are the discretised sensitivity kernels of the background wavefield with respect to the background bulk modulus and density,  $\bar{U}_S^B$  and  $\bar{V}_S^B$  are the discretised sensitivity kernels of the singular wavefield with respect to the background bulk modulus and density, and  $\bar{U}_S^S$  and  $\bar{V}_S^S$  are the discretised sensitivity kernels of the singular wavefield with respect to the singular bulk modulus and density, respectively.

### Sensitivity subkernels

Let us now demonstrate how we arrive at eq. (2.18). In full analogy to subsection 2.1.1, the decomposition of eq. (2.4) with  $p_0 = p_B + p_S$  leads to the system

$$\mathcal{L}_B [p_B(\mathbf{x}, t; \mathbf{x}_s)] = \delta(\mathbf{x} - \mathbf{x}_s) S(t), \quad (2.20)$$

$$\mathcal{L}_0 [p_S(\mathbf{x}, t; \mathbf{x}_s)] = -\mathcal{V}_0 [p_B(\mathbf{x}, t; \mathbf{x}_s)], \quad (2.21)$$

where  $\mathcal{V}_0 = \mathcal{L}_0 - \mathcal{L}_B$  is the reference scattering potential. Note that the background wavefield  $p_B$  excites the secondary sources for the singular wavefield  $p_S$ . With the help of the Green's function  $G_0$  in the reference medium, the singular wavefield is

$$p_S(\mathbf{x}, t; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' G_0(\mathbf{x}, t; \mathbf{x}') * \mathcal{V}_0 [p_B(\mathbf{x}', t; \mathbf{x}_s)]. \quad (2.22)$$

Each line of eq. (2.18) is a consequence of introducing medium parameter perturbations in eqs. (2.20) and (2.21). Perturbing the former equation will lead us to the secondary sources of the



background-wavefield residual which, in the end, will give us the first line of eq. (2.18). Perturbing the latter will lead us to the secondary sources of the singular-wavefield residual which will give us the second line of the same equation.

### Secondary sources for the background-wavefield residual

We note that eq. (2.20), which describes the most basic component of the wavefield in any decomposition, has the same functional form of that for the total wavefield, eq. (2.1). Therefore, small perturbation of  $K_B$  and  $\rho_B$  leads to the situation discussed in Subsection 2.1.1 with subscripts  $_0$  replaced by  $_B$ ;  $\delta p$  by  $\delta p_B$ ;  $\delta K$  by  $\delta K_B$ ; and  $\delta \mathcal{L}$  by  $\delta \mathcal{L}_B$  in eqs. (2.12) and (2.13). As a consequence, the equation for the corresponding linearised residual  $\delta p_B$  of the background wavefield has the form of eq. (2.12), i.e.,

$$\mathcal{L}_B [\delta p_B(\mathbf{x}, t; \mathbf{x}_s)] = -\delta \mathcal{L}_B [p_B(\mathbf{x}, t; \mathbf{x}_s)], \quad (2.23)$$

where  $\mathcal{L}_B$  denotes the acoustic wave operator in the background medium. Moreover, the background secondary potential,  $\delta \mathcal{L}_B$ , is defined as

$$\delta \mathcal{L}_B = - \left\{ \frac{\delta K_B(\mathbf{x})}{K_B^2(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{\delta \rho_B(\mathbf{x})}{\rho_B^2(\mathbf{x})} \nabla \right) \right\}. \quad (2.24)$$

The solution to equation (2.23) can again be found by convolution with a Green's function. With the help of the Green's function  $G_B$  in the background medium, we can evaluate the background-wavefield residual as

$$\delta p_B(\mathbf{x}, t; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' G_B(\mathbf{x}, t; \mathbf{x}') * \delta \mathcal{L}_B [p_B(\mathbf{x}', t; \mathbf{x}_s)], \quad (2.25)$$

of which the first line of eq. (2.18) is the discrete representation in the frequency domain.

Using the Green's function's reciprocity, the explicit expression for the bulk-modulus part of the sensitivity kernel in eq. (2.25) becomes

$$\delta p_B^K(\mathbf{x}_g, t; \mathbf{x}_s) = \int_{\mathbb{V}} d^3 \mathbf{x}' \left[ \frac{1}{K_B^2(\mathbf{x}')} G_B(\mathbf{x}', t; \mathbf{x}_g) * \frac{\partial^2 p_B}{\partial t^2}(\mathbf{x}', t; \mathbf{x}_s) \right] \delta K_B(\mathbf{x}'), \quad (2.26)$$

where the expression in brackets is the sensitivity kernel of the background wavefield with respect to the background bulk modulus for the source-receiver pair  $(\mathbf{x}_s, \mathbf{x}_g)$ . In other words, its Fourier

## 2.1 Full-waveform inversion and sensitivity kernels

---

transform is represented by  $\bar{U}_B$  in eq. (2.18). Correspondingly, an expression for  $\bar{V}_B$  in eq. (2.18) can be obtained from the density part of  $\delta\mathcal{L}_B$  in eq. (2.25).

### Secondary sources for the singular-wavefield residual

The derivation of the second line in eq. (2.18) starts from the linearised eq. (2.12). To derive the secondary sources of the singular-wavefield residual, we decompose the perturbed wavefield as  $\delta p = \delta p_B + \delta p_S$ , where  $\delta p_B$  is the background residual which satisfies equation (2.23). From the linearised equation (2.12), taking into account the wavefield decomposition, i.e.,  $p = p_B + p_S$ , we find

$$\mathcal{L}_0 [\delta p_B + \delta p_S] = -\delta\mathcal{L} [p_B + p_S], \quad (2.27)$$

$$\mathcal{L}_0 [\delta p_B] + \mathcal{L}_0 [\delta p_S] = -\delta\mathcal{L} [p_B] - \delta\mathcal{L} [p_S], \quad (2.28)$$

$$\mathcal{L}_0 [\delta p_S] = -\delta\mathcal{L} [p_B] - \delta\mathcal{L} [p_S] - \mathcal{L}_0 [\delta p_B]. \quad (2.29)$$

Using the identity  $\mathcal{L}_0 = \mathcal{V}_0 + \mathcal{L}_B$  yields

$$\mathcal{L}_0 [\delta p_S] = -\delta\mathcal{L} [p_B] - \delta\mathcal{L} [p_S] - \mathcal{V}_0 [\delta p_B] - \mathcal{L}_B [\delta p_B]. \quad (2.30)$$

Upon the use of equation (2.23), this can be recast into the form

$$\mathcal{L}_0 [\delta p_S] = -\mathcal{V}_0 [\delta p_B] - \delta\mathcal{L} [p_B] - \delta\mathcal{L} [p_S] + \delta\mathcal{L}_B [p_B]. \quad (2.31)$$

Here,  $\delta\mathcal{L}$  is the linearised secondary potential given by eq. (2.13), with the model parameters and perturbations decomposed according to eqs. (2.17) and (2.19), respectively. Note that in linear approximation, equations (2.30) and (??) are equivalent. However, since  $\mathcal{L}_B [\delta p_B]$  and  $\delta\mathcal{L}_B [p_B]$  are different operators applied to different wavefields, practical results may slightly differ when calculating one or the other. The size of the difference allows to assess the importance of nonlinear effects.

An alternative, more tedious way to derive eq. (2.31) is to introduce the model perturbations  $\delta K_B$ ,  $\delta K_S$ ,  $\delta\rho_B$ , and  $\delta\rho_S$  in eq. (2.21). Other possibility is shown in Appendix ??.

Eq. (2.31) allows us to evaluate the singular-wavefield residual,  $\delta p_S$ , resulting in

$$\delta p_S(\mathbf{x}, t; \mathbf{x}_s) = \int_{\mathbb{V}} d^3\mathbf{x}' G_0(\mathbf{x}, t; \mathbf{x}') * \Delta s(\mathbf{x}', t; \mathbf{x}_s), \quad (2.32)$$

where  $G_0(\mathbf{x}, t; \mathbf{x}')$  is the Green's function of the reference wave operator  $\mathcal{L}_0$  as before. Discretisation of this expression under consideration of the contributions to  $\Delta s$  leads to the second line of eq. (2.18).

### Decomposition of the singular-wavefield residual

Let us now make the contributions to these sensitivity kernels more explicit. For this purpose we consider the Green's function  $G_0$ , too, as composed by background and singular-part wavefields, i.e.  $G_0 = G_B + G_S$ , where  $G_B$  satisfies eq. (2.20) with an impulsive point source, and where  $G_S$  satisfies eq. (2.21) for secondary sources excited by  $G_B$ . In other words,  $G_B$  is the Green's function in the background medium, and  $G_S$  is its difference to the Green's function in the reference medium. We will refer to  $G_B$  and  $G_S$  as background and singular Green's functions, respectively.

With this notation, expression (2.32) can be decomposed into 8 terms as

$$\delta p_S(\mathbf{x}_g, t; \mathbf{x}_s) = \sum_{i=1}^{n=8} \delta p_{S,i}(\mathbf{x}_g, t; \mathbf{x}_s) = \quad (2.33)$$

$$- \int_{\mathbb{V}} d^3 \mathbf{x}' G_S(\mathbf{x}_g, t; \mathbf{x}') * \mathcal{V}_0 [\delta p_B(\mathbf{x}', t; \mathbf{x}_s)] \quad (2.33.1)$$

$$- \int_{\mathbb{V}} d^3 \mathbf{x}' G_B(\mathbf{x}_g, t; \mathbf{x}') * \mathcal{V}_0 [\delta p_B(\mathbf{x}', t; \mathbf{x}_s)] \quad (2.33.2)$$

$$- \int_{\mathbb{V}} d^3 \mathbf{x}' G_S(\mathbf{x}_g, t; \mathbf{x}') * \delta \mathcal{L} [p_B(\mathbf{x}', t; \mathbf{x}_s)] \quad (2.33.3)$$

$$- \int_{\mathbb{V}} d^3 \mathbf{x}' G_B(\mathbf{x}_g, t; \mathbf{x}') * \delta \mathcal{L} [p_B(\mathbf{x}', t; \mathbf{x}_s)] \quad (2.33.4)$$

$$- \int_{\mathbb{V}} d^3 \mathbf{x}' G_S(\mathbf{x}_g, t; \mathbf{x}') * \delta \mathcal{L} [p_S(\mathbf{x}', t; \mathbf{x}_s)] \quad (2.33.5)$$

$$- \int_{\mathbb{V}} d^3 \mathbf{x}' G_B(\mathbf{x}_g, t; \mathbf{x}') * \delta \mathcal{L} [p_S(\mathbf{x}', t; \mathbf{x}_s)] \quad (2.33.6)$$

$$+ \int_{\mathbb{V}} d^3 \mathbf{x}' G_S(\mathbf{x}_g, t; \mathbf{x}') * \delta \mathcal{L}_B [p_B(\mathbf{x}', t; \mathbf{x}_s)] \quad (2.33.7)$$

$$+ \int_{\mathbb{V}} d^3 \mathbf{x}' G_B(\mathbf{x}_g, t; \mathbf{x}') * \delta \mathcal{L}_B [p_B(\mathbf{x}', t; \mathbf{x}_s)], \quad (2.33.8)$$

where Green's function reciprocity was used once again. The terms are provisionally numbered from top to bottom, i.e., term  $\delta p_{S,1}$  is given in line (2.33.1) and so forth. In the next section, we

## 2.2 Physical interpretation of the forward decomposition

---

will analyse and name these terms more appropriately based on their physical meaning.

## 2.2 Physical interpretation of the forward decomposition

### 2.2.1 Old information, brand new point of view

It is important to recognise that splitting the wavefield residual,  $\delta p$ , into separate background and singular terms,  $\delta p_B$  and  $\delta p_{S,i}$ , does not add new contributions as compared to the classical form of the undecomposed wavefield residual given by expression (2.14), if the unperturbed medium  $m_0$  is assumed to contain both a background and a singular part. The significance of the splitting lies on the possibility to analyse the contributions individually, and to recognise the importance of multiple-scattering contributions originating from the non-smooth part in the reference model.

First of all, we immediately recognise that the eighth term (2.33.8) of eq. (2.33) is the negative of the background-wavefield residual, eq. (2.25). This leads to

$$\begin{aligned}\delta p_S(\mathbf{x}_g, t; \mathbf{x}_s) - \delta p_{S,8}(\mathbf{x}_g, t; \mathbf{x}_s) &= \delta p_S(\mathbf{x}_g, t; \mathbf{x}_s) + \delta p_B(\mathbf{x}_g, t; \mathbf{x}_s) \\ &= \delta p(\mathbf{x}_g, t; \mathbf{x}_s) = \sum_{i=1}^{n=7} \delta p_{S,i}(\mathbf{x}_g, t; \mathbf{x}_s),\end{aligned}\quad (2.34)$$

where we have used that  $\delta p = \delta p_B + \delta p_S$ . In other words, the sum of the first seven terms above yields the same total wavefield perturbation as the conventional formulation.

### 2.2.2 Multiple scattering/internal reflection uncovered

All of the contributions of eq. (2.33) have the very same structure that reveals the physics and the importance of each one of these terms. They consist of a convolution of an extrapolator term with a secondary-source term. The extrapolator describes the wave process that takes the energy from the secondary sources to the receiver. The secondary sources are constituted by the potential operator of the involved scattering mechanism and the source wavefield that excited this mechanism in the first place.

The meaning and importance of the contribution depends on the specific model decomposi-

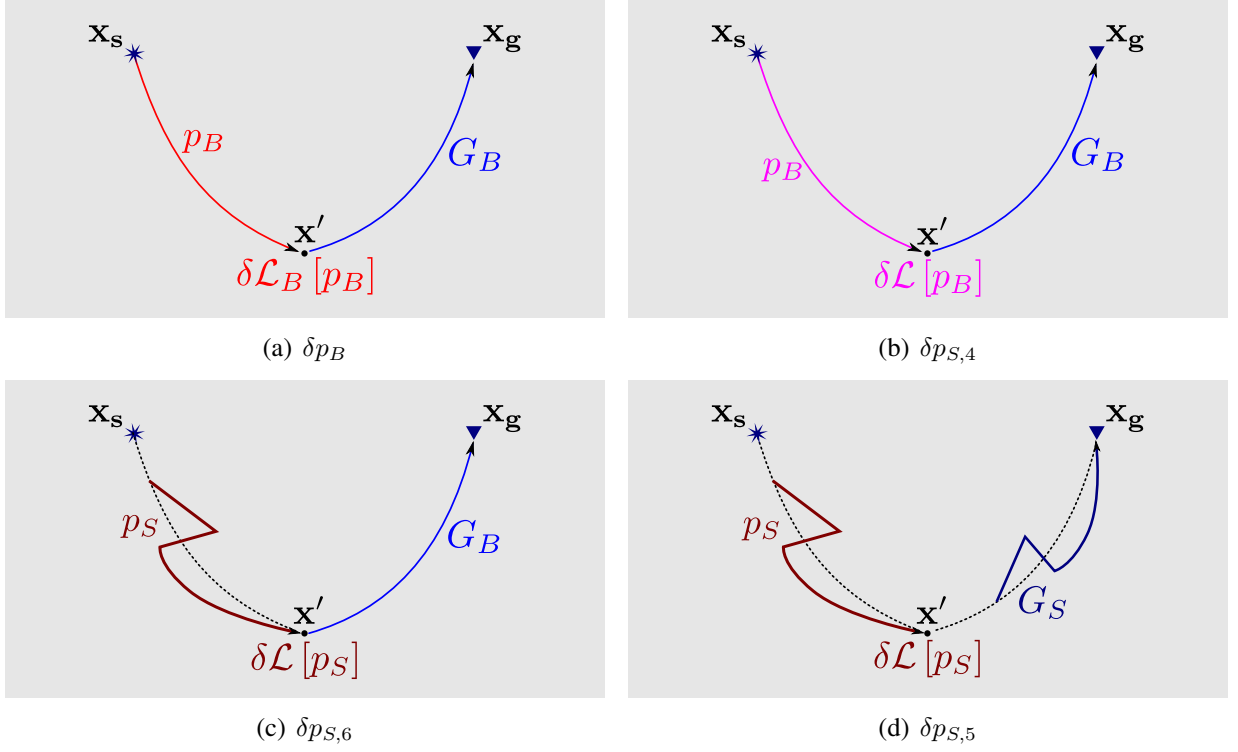


Figure 2.2: Different terms showing different levels of interaction of single- and multiple- scattered/reflected information. The colour code is the one adopted in eqs. (2.35), (2.36), (2.37) and (2.38). From top left, in clockwise direction: (a) and (b) show contribution from single-scattering events, difference between them resting on which operator is used to generate the secondary source to be extrapolated; (c) brings contribution from multiple-scattering events from receiver-side; (d) shows the contribution of the strong multiple-scattering events.

tion. Here, we assume the decomposition to follow the classical approach, i.e., in a background model carrying the low spatial-frequency components and a singular model carrying the high ones. Then, the decomposition will make explicit the different levels of interaction between single- and multiple-scattered/reflected information within the data. Below, we will analyse the respective physics behind some of the individual terms (see also Figure 2.2).

It is worth noting that decomposing the linearised expressions means that there is still a single-scattering assumption involved. While the decomposed approach considers full multiple scattering in the (singular part of the) reference model, it considers only single scattering at the (unknown) model perturbations. In other words, both in the conventional approach and in the one proposed here, the wavefield residuals are evaluated with the help of the Born approximation, i.e., the extrapolators  $G_0(\mathbf{x}_g, t; \mathbf{x}')$ ,  $G_B(\mathbf{x}_g, t; \mathbf{x}')$ , and  $G_S(\mathbf{x}_g, t; \mathbf{x}') = G_0(\mathbf{x}_g, t; \mathbf{x}') - G_B(\mathbf{x}_g, t; \mathbf{x}')$  in eqs. (2.14)

## 2.2 Physical interpretation of the forward decomposition

and (2.33) are wavefields propagating in the unperturbed models. No higher-order interactions among the model perturbations are taken into account.

We are now ready to analyse the individual terms of the wavefield residuals. Note that there are terms based solely on single scattering at model perturbations without scattering at singular parts of the reference model. Two examples are the background residual,  $\delta p_B$  of eq. (2.25), and the singular-residual contribution,  $\delta p_{S,4}$  of eq. (2.33.4). These are given by

$$\delta p_B(\mathbf{x}_g, t; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' \quad \overbrace{G_B(\mathbf{x}_g, t; \mathbf{x}')}^{\text{background-wavefield extrapolator}} * \overbrace{\delta \mathcal{L}_B[p_B(\mathbf{x}', t; \mathbf{x}_s)]}^{\text{background-wavefield excited background secondary source}} \quad (2.35)$$

and

$$\delta p_{S,4}(\mathbf{x}_g, t; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' \quad \overbrace{G_B(\mathbf{x}_g, t; \mathbf{x}')}^{\text{background-wavefield extrapolator}} * \overbrace{\delta \mathcal{L}[p_B(\mathbf{x}', t; \mathbf{x}_s)]}^{\text{background-wavefield excited total secondary source}}, \quad (2.36)$$

respectively. Figures 2.2(a) and (b) display cartoons that graphically explain these single-scattering phenomena, showing that for these terms both the source-side wavefield exciting the secondary sources and the receiver-side extrapolator are wavefields that propagate in the unperturbed background model. These figures also highlight with a colour code the difference between them: the operator involved in the scattering mechanism. For  $\delta p_B$ , only the relative background perturbations  $\frac{\delta K_B}{K_B}$  and  $\frac{\delta \rho_B}{\rho_B}$  generate secondary sources, while for  $\delta p_{S,4}$ , the relative total medium perturbations  $\frac{\delta K_B + \delta K_S}{K_0}$  and  $\frac{\delta \rho_B + \delta \rho_S}{\rho_0}$  do so.

Parts (c) and (d) of Figure 2.2 represent examples of terms that are single scattered at the model perturbations and multiple scattered at the singularities of the unperturbed model. The contribution  $\delta p_{S,6}$  of term (2.33.6) has the form

$$\delta p_{S,6}(\mathbf{x}_g, t; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' \quad \overbrace{G_B(\mathbf{x}_g, t; \mathbf{x}')}^{\text{background-wavefield extrapolator}} * \overbrace{\delta \mathcal{L}[p_S(\mathbf{x}', t; \mathbf{x}_s)]}^{\text{singular-wavefield excited total secondary source}}. \quad (2.37)$$

Here, multiple scattering is observed in the source-side wavefield as indicated in its graphical representation in Figure 2.2c. The difference to the single-scattering term  $\delta p_{S,4}$  (Figure 2.2b) is

that the wavefield that excites the secondary source at  $\mathbf{x}'$  now originates at the singularities in the reference medium described by the operator difference  $\mathcal{V}_0 = \mathcal{L}_0 - \mathcal{L}_B$ .

An example of strongly multiple-scattered events is the term (2.33.5), i.e.,

$$\delta p_{S,5}(\mathbf{x}_g, t; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' \quad \overbrace{G_S(\mathbf{x}_g, t; \mathbf{x}')}^{\text{singular-wavefield extrapolator}} * \overbrace{\delta \mathcal{L}[p_S(\mathbf{x}', t; \mathbf{x}_s)]}^{\text{singular-wavefield excited total secondary source}} . \quad (2.38)$$

Both the secondary source and the extrapolator are built up by unperturbed singular wavefields (see Figure 2.2d) generated by the differences between the reference medium and its background part.

Similar analyses can be done for the remaining terms to understand the physical meaning of each contribution. However, before rewriting eq. (2.33) accordingly, we need to recognise that terms  $\delta p_{S,i}$  ( $i = 3, 4, 5, 6$ ) can be further decomposed. We discuss this in the next section. In Section 2.2.4, we will then be ready to rewrite eq. (2.33) in a form that allows to see all different levels of interaction among the single and multiple scattered/reflected wavefield existing in each one of the terms.

### 2.2.3 Background and singular-part contributions

Actually, the secondary potential  $\delta \mathcal{L}$ , defined in eq. (2.13), consists of two components. Explicitly, it reads

$$\delta \mathcal{L} = - \left\{ \left( \frac{\delta K_B + \delta K_S}{(K_B + K_S)^2} \right) \frac{\partial^2}{\partial t^2} - \nabla \cdot \left[ \left( \frac{\delta \rho_B + \delta \rho_S}{(\rho_B + \rho_S)^2} \right) \nabla \right] \right\} . \quad (2.39)$$

In other words, the secondary potential can be written as  $\delta \mathcal{L} = \delta \mathcal{B} + \delta \mathcal{S}$ , with

$$\delta \mathcal{B} = - \left\{ \left( \frac{\delta K_B}{(K_B + K_S)^2} \right) \frac{\partial^2}{\partial t^2} - \nabla \cdot \left[ \left( \frac{\delta \rho_B}{(\rho_B + \rho_S)^2} \right) \nabla \right] \right\} \quad (2.40)$$

and

$$\delta \mathcal{S} = - \left\{ \left( \frac{\delta K_S}{(K_B + K_S)^2} \right) \frac{\partial^2}{\partial t^2} - \nabla \cdot \left[ \left( \frac{\delta \rho_S}{(\rho_B + \rho_S)^2} \right) \nabla \right] \right\} , \quad (2.41)$$

where both  $\delta \mathcal{B}$  and  $\delta \mathcal{S}$  depend on both unperturbed parts of the model, but only on the perturbation of one of them. Particularly, notice that  $\delta \mathcal{B} \neq \delta \mathcal{L}_B$  (compare to eq. (2.24)). The operator  $\delta \mathcal{B}$

## 2.2 Physical interpretation of the forward decomposition

---

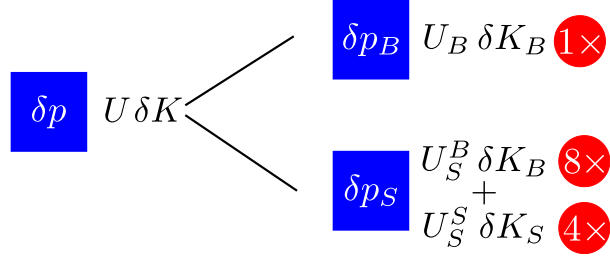


Figure 2.3: Subkernels obtained from the decomposition of the sensitivity kernel  $\bar{U}$  for bulk-modulus perturbation. Kernel  $\bar{V}$  for density perturbation is decomposed analogously.

is called *background part of secondary potential*, and  $\delta S$  is called *singular part of secondary potential*.

Therefore, each of the terms  $\delta p_{S,3}$  to  $\delta p_{S,6}$  possesses two contributions, corresponding to the parts  $\delta \mathcal{B}$  and  $\delta \mathcal{S}$  of  $\delta \mathcal{L}$ . As an example, let us take the third contribution  $\delta p_{S,3}$ . It can be written as a sum of a background and a singular part, viz.,

$$\delta p_{S,3}(\mathbf{x}_g, t; \mathbf{x}_s) = \delta p_{S,3B}(\mathbf{x}_g, t; \mathbf{x}_s) + \delta p_{S,3S}(\mathbf{x}_g, t; \mathbf{x}_s), \quad (2.42)$$

where each of the two terms on the right side still consists of two contributions, corresponding to the bulk-modulus and density perturbations. The background bulk-modulus contribution has the form

$$\delta p_{S,3B}^K(\mathbf{x}_g, t; \mathbf{x}_s) = \int_{\mathbb{V}} d^3 \mathbf{x}' \left[ \frac{1}{(K_B(\mathbf{x}') + K_S(\mathbf{x}'))^2} G_S(\mathbf{x}', t; \mathbf{x}_g) * \frac{\partial^2 p_B}{\partial t^2}(\mathbf{x}', t; \mathbf{x}_s) \right] \delta K_B(\mathbf{x}'). \quad (2.43)$$

The term in brackets describes the  $\delta p_{S,3}$  part of the sensitivity kernel of the singular-wavefield residual with respect to the background of the bulk modulus, the discretised Fourier transform of which is  $\bar{U}_S^B$  of eq. (2.18). Thus, we denote this contribution by  $\bar{U}_{S,3}^B$ . Correspondingly, there are parts of the sensitivity kernel of the singular-wavefield residual with respect to the singular part of the bulk modulus,  $\bar{U}_S$ , and therefore called  $\bar{U}_{S,3}^S$ , and to the background and singular parts of the density,  $\bar{V}_{S,3}^B$  and  $\bar{V}_{S,3}^S$ . The terms  $\delta p_{S,4}$ ,  $\delta p_{S,5}$ , and  $\delta p_{S,6}$  can be decomposed analogously, giving rise to corresponding contributions to both sensitivity kernels. Note that in all four cases, the contribution to the background sensitivity kernel is equal to that to the singular-part sensitivity kernel, i.e.,  $\bar{U}_{S,i}^B = \bar{U}_{S,i}^S = \bar{U}_{S,i}$  and  $\bar{V}_{S,i}^B = \bar{V}_{S,i}^S = \bar{V}_{S,i}$  ( $i = 3, 4, 5, 6$ ).

In this way, we arrive at thirteen subkernels (Figure 2.3) for each parameter of the model, i.e.,



bulk modulus and density. These are one subkernel to evaluate the background-wavefield residual and twelve subkernels to evaluate the singular-wavefield residual. Of the latter, eight depend on perturbations of the smooth part of the model and four depend on perturbations of the singular part. Therefore, eq. (2.18) can be rewritten as

$$\begin{aligned} \begin{bmatrix} \hat{\delta p}_B \\ \hat{\delta p}_S \end{bmatrix} &= \begin{bmatrix} \bar{U}_B & \bar{V}_B & 0 & 0 \\ \bar{U}_S^B & \bar{V}_S^B & \bar{U}_S^S & \bar{V}_S^S \end{bmatrix} \begin{bmatrix} \overline{\delta K}_B \\ \overline{\delta \rho}_B \\ \overline{\delta K}_S \\ \overline{\delta \rho}_S \end{bmatrix} \\ &= \begin{bmatrix} \bar{U}_B & \bar{V}_B & 0 & 0 \\ \sum_i \bar{U}_{S,i} & \sum_i \bar{V}_{S,i} & \sum_j \bar{U}_{S,j} & \sum_j \bar{V}_{S,j} \end{bmatrix} \begin{bmatrix} \overline{\delta K}_B \\ \overline{\delta \rho}_B \\ \overline{\delta K}_S \\ \overline{\delta \rho}_S \end{bmatrix}, \end{aligned} \quad (2.44)$$

with  $i = 1, 2, 3, 4, 5, 6, 7, 8$  and  $j = 3, 4, 5, 6$ , the numbering referred to the one displayed in eq. (2.33).

## 2.2.4 Physical interpretation of the kernels: naming the terms

The provisional numbered notation was useful to count the contributions to the wavefield residuals, but it is not very helpful to identify them. For completeness, we even needed to introduce another subscript to identify the individual contributions to  $\delta p_{S,i}$  ( $i = 3, 4, 5, 6$ ) as indicated in eq. (2.42). For this reason, we rename all terms based on their physical meaning as  $\delta p_{S,\alpha\beta\gamma}$ , where each of the three subscripts  $\alpha$ ,  $\beta$ , and  $\gamma$  stands for a physical action involved in the generation of the contribution. The first subscript,  $\alpha$ , represents the wavefield responsible for the propagation of the contribution on the source side, the second subscript,  $\beta$ , stands for the potential operator describing the secondary sources, and the third subscript,  $\gamma$ , designates the wavefield responsible for the propagation of the contribution on the receiver side. The receiver-wavefield subscript  $\gamma$  is either  $B$  for the background wavefield or  $S$  for the singular wavefield. The source-wavefield subscript  $\alpha$  can, in addition to  $B$  and  $S$ , also be  $b$ , representing the wavefield residual due to the background perturbation,  $\delta p_B$ . The potential index  $\beta$  can take the values  $\mathcal{L}$ ,  $\mathcal{B}$ ,  $\mathcal{S}$  and  $\mathcal{V}$ , representing the background secondary potential  $\delta \mathcal{L}_B$ , background part  $\delta \mathcal{B}$  and singular part  $\delta \mathcal{S}$  of the secondary potential, and the scattering potential  $\mathcal{V}_0$ , respectively. For instance, the background contribution  $\delta p_{S,3B}$  to  $\delta p_{S,3}$ , the bulk-modulus contribution of which is explicated in eq. (2.43),

## 2.2 Physical interpretation of the forward decomposition

---

becomes  $\delta p_{S,BBS}$ , indicating that the background wavefield ( $B$ ) excites secondary sources at the background perturbations of the reference medium ( $B$ ), which are propagated to the receiver by the singular Green's function ( $S$ ). Correspondingly, its singular-part contribution  $\delta p_{S,3S}$  becomes  $\delta p_{S,BSS}$ .

Figure 2.4 shows cartoons representing the physical interpretation of all thirteen bulk-modulus subkernels of eq. (2.33). The subcaptions state the new name of each term and its old one in the previous, numbered nomenclature of eq. (2.33). We still need an additional index (for example, a superscript  $K$  or  $\rho$ ) to indicate whether the contribution originates from a perturbation of the bulk modulus or density. However, in our numerical examples below, where we restrict ourselves to bulk-modulus perturbations, we will omit this additional index for the sake of simplicity.

The subscripts of the Fréchet derivatives contributions  $\bar{U}_{S,i}$  and  $\bar{V}_{S,i}$  to the sensitivity kernels can be replaced accordingly. However, as seen in the previous section, what would be the contributions  $\bar{U}_{S,BBS}$  and  $\bar{U}_{S,BSS}$  are the same, making the potential subscript  $\beta$  needless for this pair. The same happens to the pairs:  $\bar{U}_{S,BBB}$  and  $\bar{U}_{S,BSB}$ ;  $\bar{U}_{S,SBS}$  and  $\bar{U}_{S,SSS}$ ; and  $\bar{U}_{S,SSB}$  and  $\bar{U}_{S,SSS}$ . So, for these 4 pairs we simplify the notation dropping the potential subscript when referring to the corresponding discretised Fréchet derivative. Therefore, we restate the following definitions from (2.44):

$$\bar{U}_S^B = \sum_i \bar{U}_{S,i}, \quad (2.45)$$

$$\bar{U}_S^S = \sum_j \bar{U}_{S,j}, \quad (2.46)$$

where  $i = b\mathcal{V}S, b\mathcal{V}B, BS, BB, SS, SB, -B\mathcal{L}S, -B\mathcal{L}B$ , and  $j = BS, BB, SS, SB$ . By analogy, the same can be done to the Fréchet derivatives  $\bar{V}_S^B$  and  $\bar{V}_S^S$ .

Two examples of the use of the new nomenclature are

$$\hat{\delta p}_{S,b\mathcal{V}S}^K = \bar{U}_{S,b\mathcal{V}S} \bar{\delta K}_B, \quad (2.47)$$

$$\hat{\delta p}_{S,BSB}^\rho = \bar{V}_{S,BB} \bar{\delta \rho}_S. \quad (2.48)$$

Let us analyse the sub- and superscripts of the first example (2.47). There, we have one of the contributions to the singular-wavefield residual ( $_S$ ) due to perturbation in the bulk-modulus model ( $^K$ ). The perturbation refers to the background model ( $_b$  indicates dependence on the background-wavefield residual which is present only if the background model is perturbed). the contribution is

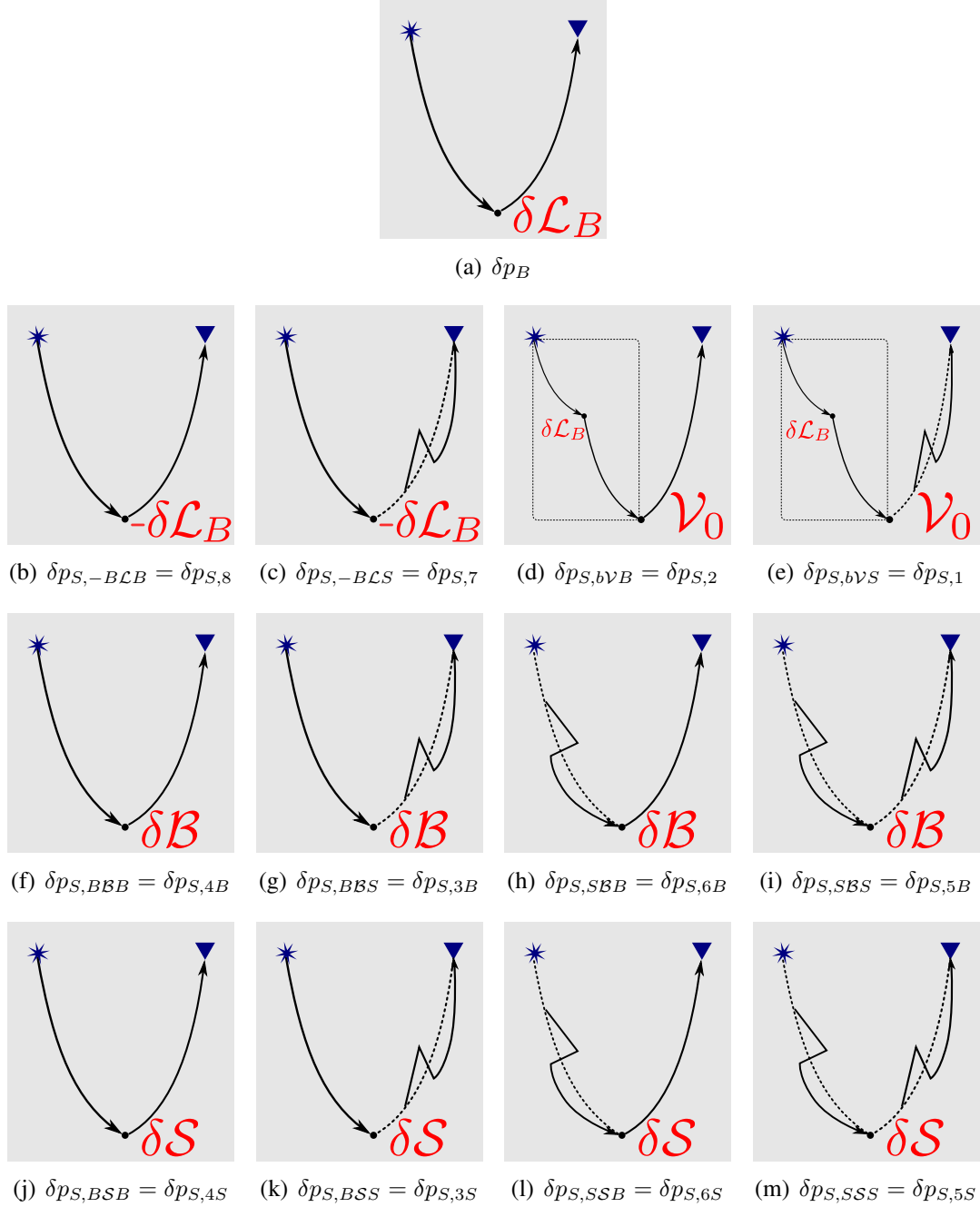


Figure 2.4: Physical meaning of the subkernels. Each one of the cartoons shows three elements: *source-side wavefield*; the *operator* generating the secondary source; and the *receiver-side wavefield extrapolator*. Subcaptions indicate the new name and correspondence to the previous, numbered nomenclature.

## 2.3 Inversion: backprojecting the residuals into model space

---

individualised by the subscript  $_{b\mathcal{V}\mathcal{S}}$  which details the elements involved in its generation. Firstly, we identify the source-side wavefield index ( $_b$ ) which indicates the wavefield that excites the secondary source. In this case, it is the background-wavefield residual. Secondly, the scattering mechanism index ( $_{\mathcal{V}}$ ) which represent the potential operator applied onto the source-side wavefield. Here, it is the reference scattering potential. And finally, the extrapolator index ( $_{\mathcal{S}}$ ) showing the wavefield that propagates the secondary source to the receivers. In this example, we have the singular Green's function.

On the second example (2.48) we have again one of the contributions to the singular-wavefield residual ( $_{\mathcal{S}}$ ), but now due to perturbation in the density model ( $^{\rho}$ ). The perturbation is done in the singular model (the index  $_{\mathcal{S}}$  stands for operator  $\delta\mathcal{S}$  which depends only on the singular- model perturbation). The contribution is label with the subscript  $_{BSB}$ . This indicates that the secondary source is excited by the background wavefield ( $_{B}$ ), the singular part of the secondary potential operator  $\delta\mathcal{S}$  as scattering mechanism operator, and the background wavefield as extrapolator ( $_{B}$ ).

## 2.3 Inversion: backprojecting the residuals into model space

The purpose of deriving expressions for the individual contributions to the wavefield residuals is their use in linearised inversion in order to update the model. This is achieved by backprojecting the residuals into the model space with the help of the adjoint Fréchet derivatives (sensitivity kernels) and using them to obtain estimates of the perturbation with the technique known as *adjoint-state method* (Tromp et al., 2005; Fichtner, 2010; Fichtner and Trampert, 2011).

### 2.3.1 Conventional formulation

From the total wavefield residual  $\delta p$ , estimates for the bulk-modulus and density perturbations,  $\delta K^{\text{est}}$  and  $\delta \rho^{\text{est}}$ , respectively, are obtained by the adjoint operation to eq. (2.15). This is achieved by backprojecting the wavefield residual according to

$$\begin{bmatrix} \delta K^{\text{est}} \\ \delta \rho^{\text{est}} \end{bmatrix} = \begin{bmatrix} \overline{U}^{\dagger} \\ \overline{V}^{\dagger} \end{bmatrix} \hat{\delta p}, \quad (2.49)$$

where superscript  $^{\dagger}$  denotes adjoint operator.

### 2.3.2 Decomposed formulation

Under the separation into background and singular components proposed here, estimates of background and singular model perturbations can be evaluated individually by backprojecting the residuals. The perturbation estimates are given by the adjoint to eq. (2.44) as follows

$$\begin{bmatrix} \delta K_B^{\text{est}} \\ \delta \rho_B^{\text{est}} \\ \delta K_S^{\text{est}} \\ \delta \rho_S^{\text{est}} \end{bmatrix} = \begin{bmatrix} \delta K_{B,B}^{\text{est}} + \sum_i \delta K_{B,i}^{\text{est}} \\ \delta \rho_{B,B}^{\text{est}} + \sum_i \delta \rho_{B,i}^{\text{est}} \\ \sum_j \delta K_{S,j}^{\text{est}} \\ \sum_j \delta \rho_{S,j}^{\text{est}} \end{bmatrix} = \begin{bmatrix} \bar{U}_B^\dagger & \sum_i \bar{U}_{S,i}^\dagger \\ \bar{V}_B^\dagger & \sum_i \bar{V}_{S,i}^\dagger \\ 0 & \sum_j \bar{U}_{S,j}^\dagger \\ 0 & \sum_j \bar{V}_{S,j}^\dagger \end{bmatrix} \begin{bmatrix} \hat{\delta p}_B \\ \hat{\delta p}_S \end{bmatrix}, \quad (2.50)$$

where the summation indices again take the values  $i = b\mathcal{V}S, b\mathcal{V}B, BS, BB, SS, SB, -B\mathcal{L}S, -B\mathcal{L}B$  and  $j = BB, BS, SB, SS$ .

The explicit meaning of these expressions is analogous to the representation of the estimate in Tarantola (1984a). At each point  $\mathbf{x}$  of the model, each contribution to the perturbation estimate is the cross-correlation between the direct wavefield from source with the backpropagated residual from the receivers. For example, estimates  $\delta K_{B,BS}^{\text{est}}$  and  $\delta K_{B,b\mathcal{V}B}^{\text{est}}$  reads explicitly (see derivation in Appendix C)

$$\delta K_{B,BS}^{\text{est}}(\mathbf{x}) = \sum_s \sum_g \int d\omega \frac{-\omega^2}{K_0^2(\mathbf{x})} \overbrace{\hat{p}_B^*(\mathbf{x}, \omega; \mathbf{x}_s)}^{\text{direct wavefield}} \underbrace{\hat{G}_S^*(\mathbf{x}, \omega; \mathbf{x}_g) \hat{\delta p}_S(\mathbf{x}_g, \omega; \mathbf{x}_s)}_{\text{back-propagation of } \hat{\delta p}_S} \quad (2.51)$$

and

$$\begin{aligned} \delta K_{B,b\mathcal{V}B}^{\text{est}}(\mathbf{x}) = & - \sum_s \sum_g \int d\omega \frac{-\omega^2}{K_B^2(\mathbf{x})} \overbrace{\hat{p}_B^*(\mathbf{x}, \omega; \mathbf{x}_s)}^{\text{direct wavefield}} \times \\ & \underbrace{\int_{\mathbb{V}} d^3 \mathbf{x}' \underbrace{\hat{G}_B^*(\mathbf{x}, \omega; \mathbf{x}')}_{\text{2nd back-prop.}} \overbrace{\left( \frac{1}{K_0(\mathbf{x}')} - \frac{1}{K_B(\mathbf{x}')} \right)}^{\text{scattering potential at } \mathbf{x}'} \underbrace{\hat{G}_B^*(\mathbf{x}', \omega; \mathbf{x}_g) \hat{\delta p}_S(\mathbf{x}_g, \omega; \mathbf{x}_s)}_{\text{1st back-propagation of } \hat{\delta p}_S \text{ to } \mathbf{x}'}}_{\text{}}. \end{aligned} \quad (2.52)$$

Note that there are singular-wavefield residuals that must be backpropagated once, e.g., from  $\mathbf{x}_g$  to  $\mathbf{x}$  as for estimate  $\delta K_{B,BS}^{\text{est}}$ , or twice, e.g., from  $\mathbf{x}_g$  to  $\mathbf{x}'$  and from there to  $\mathbf{x}$  as seen for estimate  $\delta K_{B,b\mathcal{V}B}^{\text{est}}$ . To get the explicit expressions of the other sensitivity kernels, one just needs to do the proper substitutions of the wavefields and potentials.

It is to be noted in this context that in analogy to the wavefields  $p_0$  and  $G_0$  in the conventional

## 2.3 Inversion: backprojecting the residuals into model space

---

approach, the wavefields  $p_B$ ,  $G_B$ ,  $p_S$ , and  $G_S$  in the decomposed approach are known, because all of them are calculated in the known reference model. We stress once more that the decomposition of reference model into background and singular parts is completely arbitrary, not relying on any kind of approximation.

### 2.3.3 Residual information leakage

Ideally, one part of the model is perfectly known, so that only perturbations of the other part need to be considered. However, backprojecting the wavefield residuals may lead to non-zero perturbation estimates for unperturbed parameters. We call this effect *residual information leakage*. Generally speaking, when backprojected into the model space, data residuals caused by a perturbation of the background model can leak to the estimates of the singular part and vice-versa. From expression (2.50), it is clear that there will be non-zero estimates for both  $\delta K_B$  and  $\delta K_S$  independently of whether only the singular-wavefield residual or both, background and singular residuals are non-zero (there is no way to have only a non-zero background-wavefield residual). The only way to avoid the leakage would be to use the inverse that would correctly distribute the power among the estimates. Unfortunately this is not the usual scenario in practice when we are dealing with non-invertible operators.

This matter is particularly important under the present decomposition. Since the sensitivity subkernels related to the terms  $\delta p_{S,BBB}$  and  $\delta p_{S,BSB}$  are identical ( $\overline{U}_{S,BB}$ ), so will be the estimates for both singular and background perturbations  $\delta K_{B,BB}^{\text{est}}$  and  $\delta K_{S,BB}^{\text{est}}$ , irrespective of whether the backprojecting residuals are due to a perturbation of the background model alone or of both parts. Additional criteria will be needed to decide which of these parameter perturbations really caused the wavefield perturbations.



## 3. Numerical Experiments and Discussion

We performed numerical experiments using the open-source software package *Madagascar* to study and validate the proposed decomposition, and to better understand the individual contributions, particularly when used to backproject residuals to build perturbation estimates. For this purpose, we constructed a set of different models with different perturbations and used different acquisition geometries (see Tables 3.1 and 3.2).

### 3.1 Experimental Procedure

For each of the models used in the numerical experiments we evaluated either the forward decomposition (direct problem – Section 2.1.2) only, the backprojection of the residuals (inverse problem – Section 2.3) only, or both.

All numerical experiments rely on bulk-modulus perturbations only and use a constant density of  $2200 \text{ kg/m}^3$ . The wavefield modelling was carried out with staggered-grid vector-acoustic time-domain finite difference second-order in time, fourth-order in space. The input medium parameters were bulk modulus and density. Thus, the displayed velocity values in all figures hereafter were calculated from the bulk-modulus and density values used in the computations, using the relationship  $v_p = \sqrt{K/\rho}$ . We used monopole point-sources of volume density injection rate with a Ricker signal with a peak frequency of 30 Hz as source signature.

#### 3.1.1 Forward Decomposition

All forward-modelling decomposition experiments followed these basic steps:

- i. We defined an unperturbed reference model and its background part, and modelled the corresponding reference and background wavefields  $p_0$  and  $p_B$  (the singular wavefield was evaluated by its definition,  $p_S = p_0 - p_B$ ).



- ii. After model perturbation, we modelled the corresponding perturbed wavefields and determined the true residuals  $\delta p^{\text{true}}$ ,  $\delta p_B^{\text{true}}$ , and  $\delta p_S^{\text{true}}$  from the differences.
- iii. Using the model perturbation, we calculated the linearised residuals  $\delta p$ ,  $\delta p_B$ , and  $\delta p_S$  of the total, background and singular wavefields according to eqs. (2.14), (2.25), and (2.33) respectively, and compared them to the true ones.

In all forward experiments we performed single-shot simulations. The point-source position is indicated in the corresponding model figures by a black star. The wavefields were recorded at control receivers represented in the model figures by black triangles.

All the linearised wavefield residuals are result of Born approximation and have the general format below:

$$\int_{\mathbb{V}} d^3 \mathbf{x}' G(\mathbf{x}_g, t; \mathbf{x}') * s(\mathbf{x}', t; \mathbf{x}_s), \quad (3.1)$$

where  $G(\mathbf{x}_g, t; \mathbf{x}')$  is the Green's function of the acoustic wave operator for a given set of known model parameters, and  $s(\mathbf{x}', t; \mathbf{x}_s)$  is the secondary source, result of a differential operator applied onto a known wavefield. As an analytic solution of such integral is impossible, we evaluate this expression numerically by modeller injection. In other words, whenever needed, we applied the appropriate differential operator on the known wavefield to build the secondary source  $s$  and injected it as source input at all grid points in the modeler with the appropriate set of model parameters related to the Green's function  $G(\mathbf{x}_g, t; \mathbf{x}')$ .

In the figures related to the forward-decomposition experiments (Experiments 1, 3, 4 and 6), red lines represent the true residuals and blue lines the corresponding linearly approximated ones. To the left of the wavefield comparisons, we place the associated cartoons of Figure 2.4 showing the corresponding Green's function and secondary source to help recall the physical meaning of the studied residual. The modelling of the linearised wavefield residual introduces an amplitude-error factor. For better comparison, we scaled the linearised residuals to match the amplitude of the corresponding true wavefield residuals. We then used the same scale factor for the display of all individual contributions.

### 3.1.2 Backprojecting residuals

Correspondingly, all the residual-backprojection experiments followed a procedure consisting of the following steps:

## 3.2 Model gallery

---

- i. True-residual backprojection: By means of the proper extrapolator in the time domain, we backpropagated (once or twice) the true total, background, and singular residuals (see Section 3.1.1, item ii) for a given source from a given receiver into the target area.
- ii. Cross-correlation with direct wavefield: We cross-correlated in the frequency domain the backpropagated wavefield with the proper direct wavefield (reference, background, or singular) from the corresponding source.
- iii. Stack: We stacked the resulting wavefield over frequency, sources, and receivers.

In the backprojecting experiments we performed non-simultaneous multiple-shot experiments. Source and receiver positions are marked in all models by white stars and triangles, respectively.

In the figures that depict the results of backprojecting experiments (Experiments 2, 5, 7 and 8) we employ variations of the cartoons of Figure 2.4 to indicate which (sub)kernel is used to backproject the residuals and obtain the bulk-modulus estimates. Each cartoon exhibits the kind of source-side wavefield (reference, background or singular), the kind of receiver-side Green's function used to backpropagate the residual (some of them are built with two backpropagations), and the differential operator applied before the cross-correlation between wavefields. For better comparison, we normalised the total Bulk-modulus perturbation estimates to their maximum absolute value. Individual contributions to the estimates are depicted using the same scale factor.

## 3.2 Model gallery

The numerical experiments presented in this chapter were performed departing from four different unperturbed models summarised in Table 3.1. This table also contains the description of the discretisation in each model and the acquisition geometry used in each experiment involving such model.

Figure 3.1 shows the unperturbed velocity Model I. It consists of a single planar reflector at 350  $m$  depth, separating two homogeneous halfspaces with velocities 1500  $m/s$  and 2500  $m/s$ . We chose a simple model for the first numerical experiment because, in this way, it is easy to recognise and understand each of the decomposition components.

For our decomposition, we chose the background part of Model I to be a constant velocity model with the first layer's velocity. The unperturbed singular part is then a Heaviside step function in the vertical direction with magnitude of 1000  $m/s$  and step at 350  $m$ . The vertical and horizontal

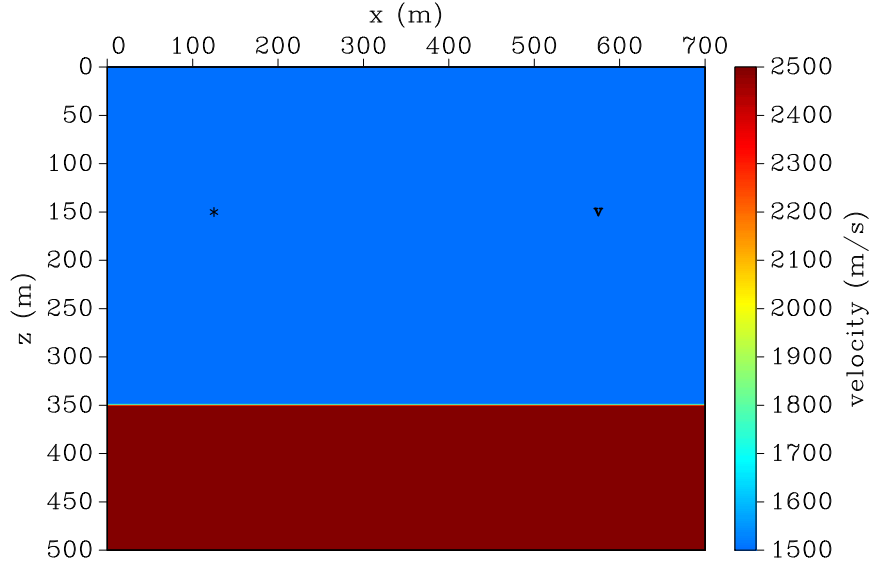


Figure 3.1: Unperturbed velocity Model I (velocity in  $m/s$ ). The first layer has a velocity of  $1500 m/s$ , and the second layer a velocity of  $2500 m/s$ . Density is kept constant ( $2200 kg/m^3$ ). The source and receiver positions are indicated by black star and triangle respectively.

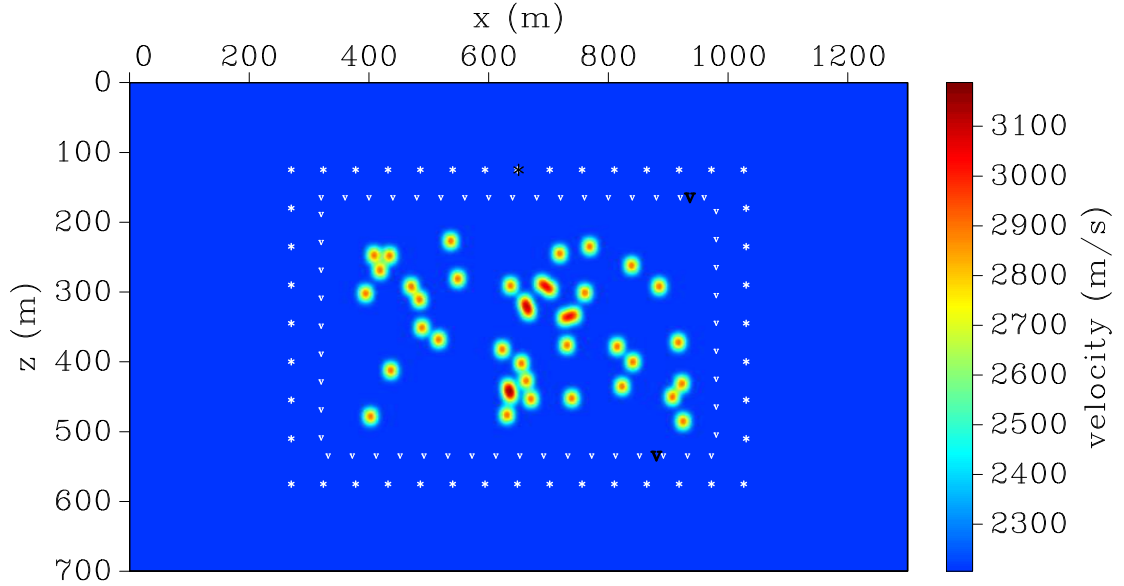


Figure 3.2: Unperturbed velocity Model II (velocity in  $m/s$ ) containing 40 randomly distributed scatterers, with constant background velocity ( $2205 m/s$ ) and density ( $2200 kg/m^3$ ). One source (black star) and two control receivers (black triangles) – named as cr154 (*top*) and cr283 (*bottom*) – were used in the forward decompositions experiments. The other sources and receivers (white stars and triangles respectively) were used in the backpropagation experiment.

### 3.2 Model gallery

Model	Name	References	Acquisition geometry	Sample number	Sampling interval
I	Single reflector	Fig. 3.1	Single receiver	$nx = 701$ $nz = 501$ $nt = 2151$	$dx = 1.0$ $dz = 1.0$ $dt = 0.2$
II	40 scatterers	Fig. 3.2	Surrounding	$nx = 651$ $nz = 701$ $nt = 4001$	$dx = 2.0$ $dz = 1.0$ $dt = 0.2$
III	60 scatterers	Fig. 3.3	Surrounding	$nx = 651$ $nz = 701$ $nt = 4001$	$dx = 2.0$ $dz = 1.0$ $dt = 0.2$
IV	150 scatterers	Fig. 3.4	Surface	$nx = 651$ $nz = 701$ $nt = 4001$	$dx = 2.0$ $dz = 1.0$ $dt = 0.2$

Table 3.1: Unperturbed model gallery.  $n\#$  stands for the number of samples in the  $\#$ -axis.  $d\#$  is the sampling interval in the  $\#$ -axis;  $x$ - and  $z$ -axis are given in meters;  $t$ -axis is given in milliseconds. the origin for all axis was always  $o\# = 0$ .

grid spacing was 1  $m$ , and the time-marching step 0.2  $ms$ . The recorded wavefields were then resampled to a 2  $ms$  interval.

The next three models were set to see the role of multiple-scattering effect in the proposed decomposition.

Figure 3.2 shows the unperturbed multiple-scatterers velocity Model II. It consists of a background part with a constant velocity of 2205  $m/s$ , and a singular part composed of 40 randomly distributed scatterers. Each of these scatterers consists of a 16  $m \times 16 m$  perturbation that was smoothed horizontal and vertically to create a bell-shaped high-velocity zone with a maximum velocity difference of 798  $m/s$  (some of the scatterers overlap partially, possibly resulting in larger velocity difference). The full acquisition geometry was used to study the effects of multiple scattering independently of the ones due to a lack of illumination. The vertical grid spacing was 1  $m$ , the horizontal one 2  $m$ , and the time-marching step was 0.2  $ms$ . We resampled the recorded wavefields to a 6  $ms$  interval. Sources were spaced by 55  $m$  and the receivers by 4  $m$ .

Figure 3.3 shows the unperturbed velocity Model III. It resembles the previous unperturbed model, the only difference being the larger number of scatterers: 60. This is done to enhance the multiple- scattering phenomena, increasing the amount of information on the background pertur-

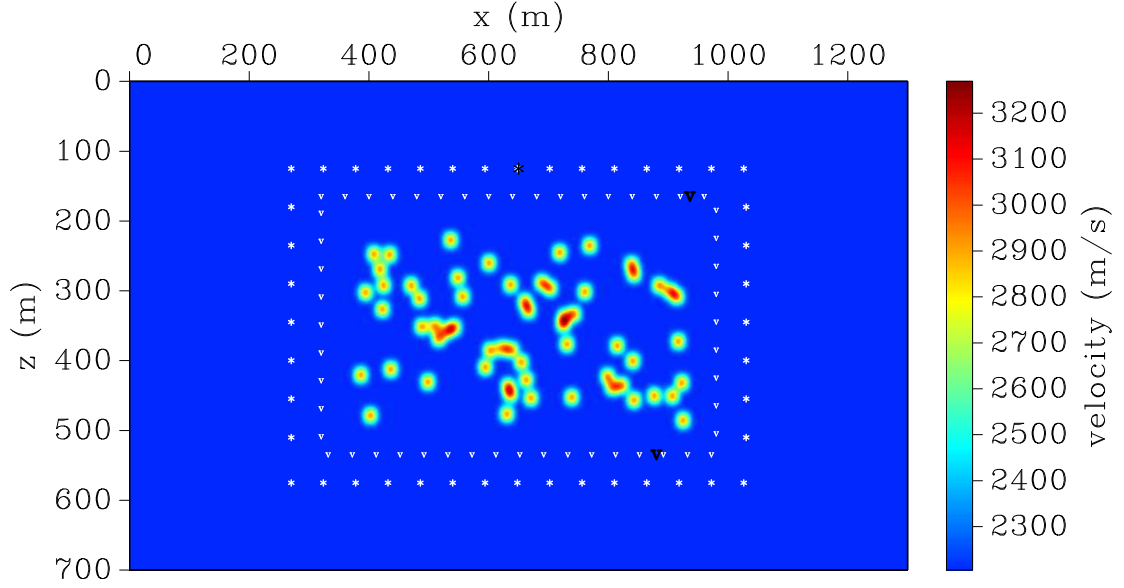


Figure 3.3: Unperturbed velocity Model III (velocity in  $m/s$ ) containing 60 randomly distributed scatterers, with constant background velocity ( $2205 m/s$ ) and density ( $2200 kg/m^3$ ). For the numerical experiments, we used the same acquisition geometry as in Model II.

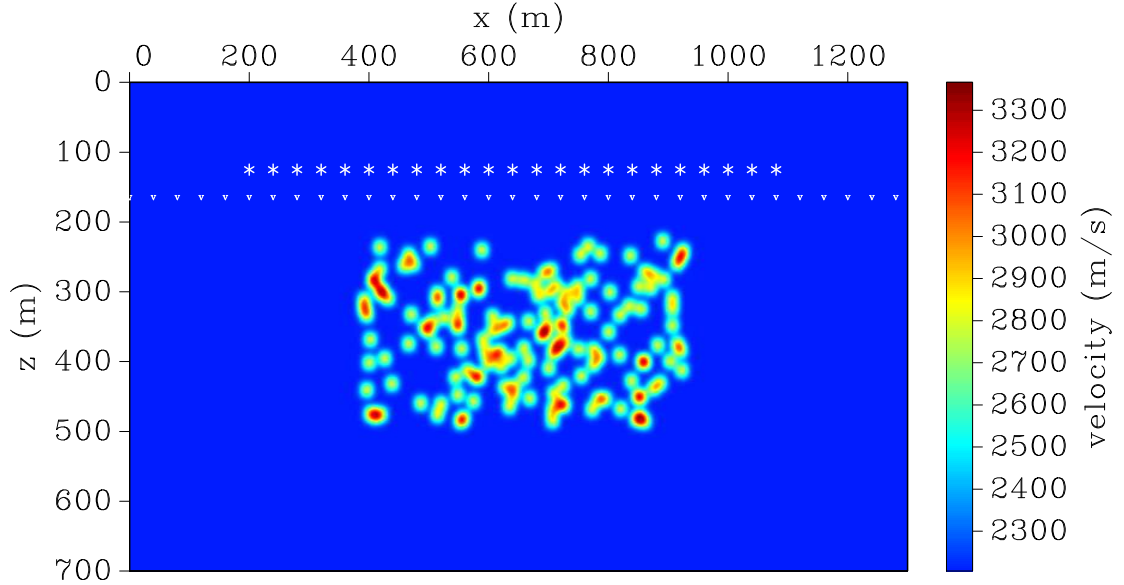


Figure 3.4: Unperturbed velocity Model IV (velocity in  $m/s$ ) containing 150 randomly distributed scatterers, with constant background velocity ( $2205 m/s$ ) and density ( $2200 kg/m^3$ ).

### 3.3 Experiment gallery

Experiment	Unperturbed model	Perturbation	Modality
1	Model I	Constant background	Forward Decomposition
2		Lens-shaped background	Backprojecting Residuals
3	Model II	Constant background	Forward Decomposition
4		Scatterers' positions	Forward Decomposition
5			Backprojecting Residuals
6	Model III	Lens-shaped background	Forward Decomposition
7			Backprojecting Residuals
8	Model IV	Lens-shaped background	Backprojecting Residuals

Table 3.2: Experiment gallery.

bation carried by the singular wavefield. We used the same grid spacing, time-marching step and resampling interval.

Finally, in Figure 3.4 we can see the unperturbed Model IV. It consists of 150 randomly distributed scatterers with the same grid spacing, time-marching step, resampling interval as the two previous models. The main difference between this and the previous ones lies in the acquisition geometry: source and receivers are restricted to the top portion of the model, simulating a surface acquisition. This was done to study the influence of the uneven illumination on our results.

## 3.3 Experiment gallery

Table 3.2 summarises all the experiments in the four models presented in this chapter. For each experiment is indicated which model was used, the kind of perturbation imposed (details will be seen in the coming sections) and which experiment modality was performed (forward decomposition or residual backprojection). Each next subsection is devoted to one of the eight experiments.

### 3.3.1 Experiment 1: Constant background perturbation of Model I, forward decomposition

In Experiment 1, we kept the singular part of the model unperturbed and introduced a 10% perturbation to the complete background part of Model I, i.e.,  $\delta K_B = 0.1K_B = 150 \text{ m/s}$  and  $\delta K_S = 0 \text{ m/s}$  everywhere in the model. According to equation (2.44), this yields a perturbation in

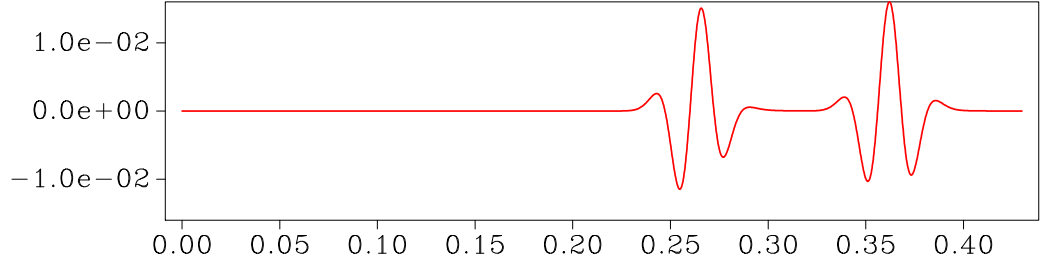


Figure 3.5: EXPERIMENT 1. True total wavefield residual  $\delta p^{\text{true}} = p - p_0$  at cr550. Observe two distinct signals: the direct wavefield residual (peak at 0.27 s) and the reflected/scattered wavefield residual (peak at 0.36 s).

both the background and singular wavefields,

$$\begin{bmatrix} \hat{\delta p}_B \\ \hat{\delta p}_S \end{bmatrix} = \begin{bmatrix} \bar{U}_B \\ \sum_i \bar{U}_{S,i} \end{bmatrix} \delta K_B, \quad (3.2)$$

where the subscript  $i$  takes the values  $BB, BS, SB, SS, -B\mathcal{L}B, -B\mathcal{L}S, b\mathcal{V}B$  and  $b\mathcal{V}S$ . These contributions correspond to the terms represented graphically by Figures 2.4a to 2.4i.

We recorded the true and calculated wavefields at a receiver (hereafter called a *control receiver*) named cr550 (black triangle in Figure 3.1). The true total wavefield residual can be seen in Figure 3.5. Two events can be clearly distinguished: the direct wavefield residual (peak at 0.27 s) and the reflected/scattered wavefield residual (peak at 0.36 s).

We start by comparing the true and estimated values of the wavefield perturbations at the control receiver and their individual contributions according to the decomposed expression (3.2).

The first contribution to be studied is the background-wavefield residual  $\delta p_B$  (Figure 3.6). We observe a good match between both waveforms at the receiver.

Next, we address the singular-wavefield residual  $\delta p_S$  at the control receiver. As shown in eq. (2.33) and illustrated in Figure 2.4,  $\delta p_S$  consists of twelve contributions, four of which are due to perturbations of the singular part of the model (see Figures 2.4j to 2.4m). Since we have not perturbed the singular part in Experiment 1, we have only eight non-zero contributions.

Based on the conventional practice of representing the reference model with the background model alone, it might be expected that the *single-scattering contributions*  $\delta p_{S,-B\mathcal{L}B}$  and  $\delta p_{S,B\mathcal{B}B}$  to the singular residual should be the most important ones. But when we inspect Figure 3.7 we can see that they are closely related to the background residual rather than to the singular one.

### 3.3 Experiment gallery

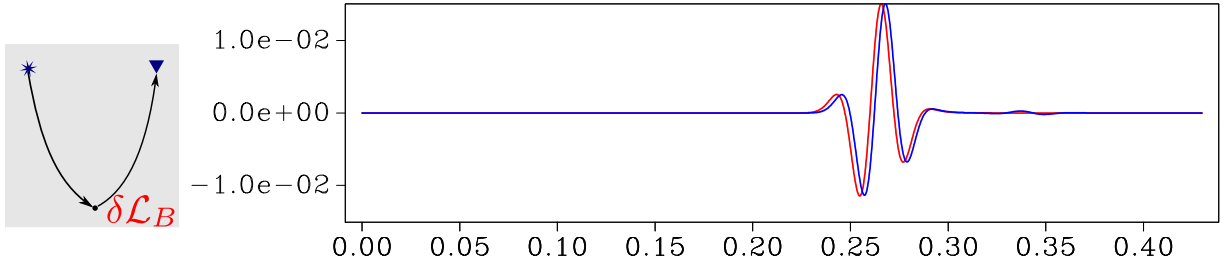


Figure 3.6: EXPERIMENT 1. Linearised background-wavefield residuals  $\delta p_B$  (blue line) evaluated at cr550 as compared to the true one (red line),  $\delta p_B^{\text{true}}$ .

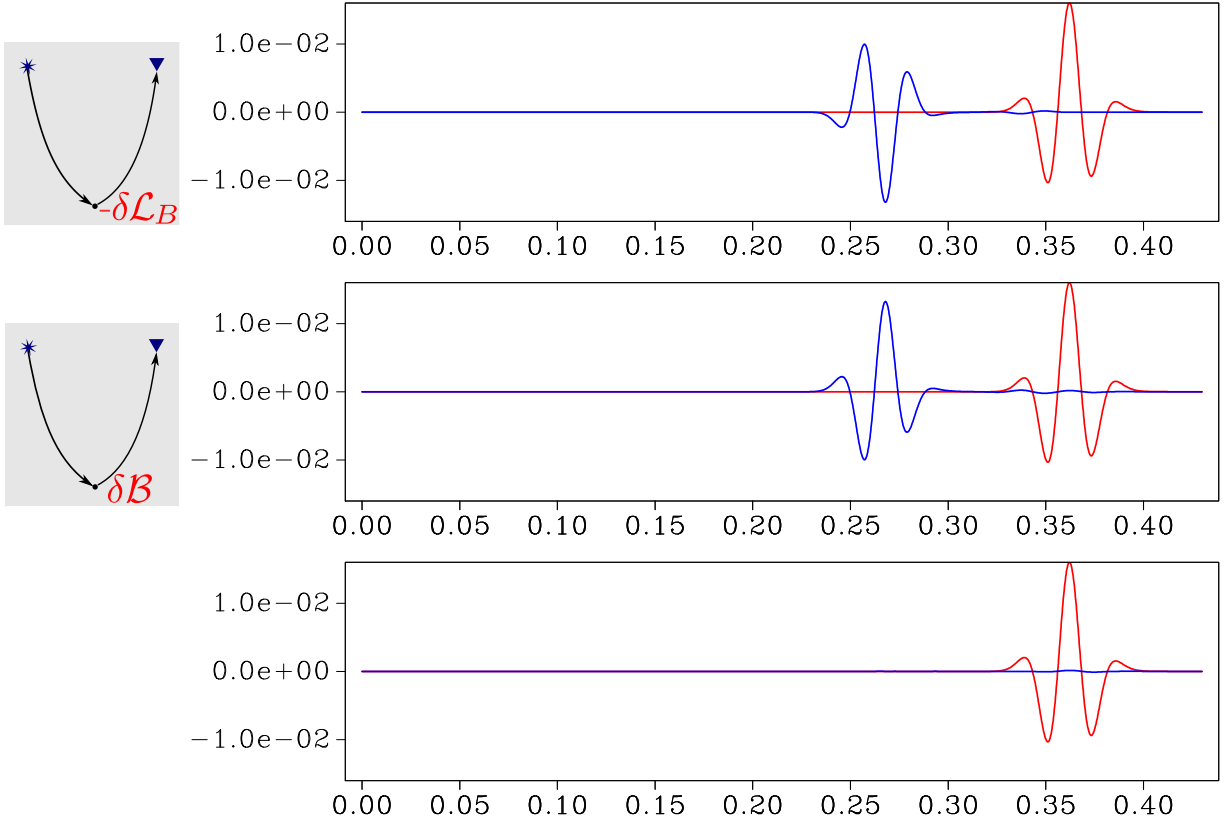


Figure 3.7: EXPERIMENT 1. Terms related to single-scattering effects at the medium perturbations (blue lines):  $\delta p_{S,-B\mathcal{L}B}$  (top),  $\delta p_{S,B\mathcal{B}B}$  (centre), and their sum (bottom). The red line true singular residual  $\delta p_S^{\text{true}}$ . Separately,  $\delta p_{S,-B\mathcal{L}B}$  and  $\delta p_{S,B\mathcal{B}B}$  carry the largest energy. Their sum demonstrates that they are approximately complementary.

Actually, as stated before, the term  $\delta p_{S,-B\mathcal{L}B}$  is the opposite of the background residual,  $\delta p_B$ .



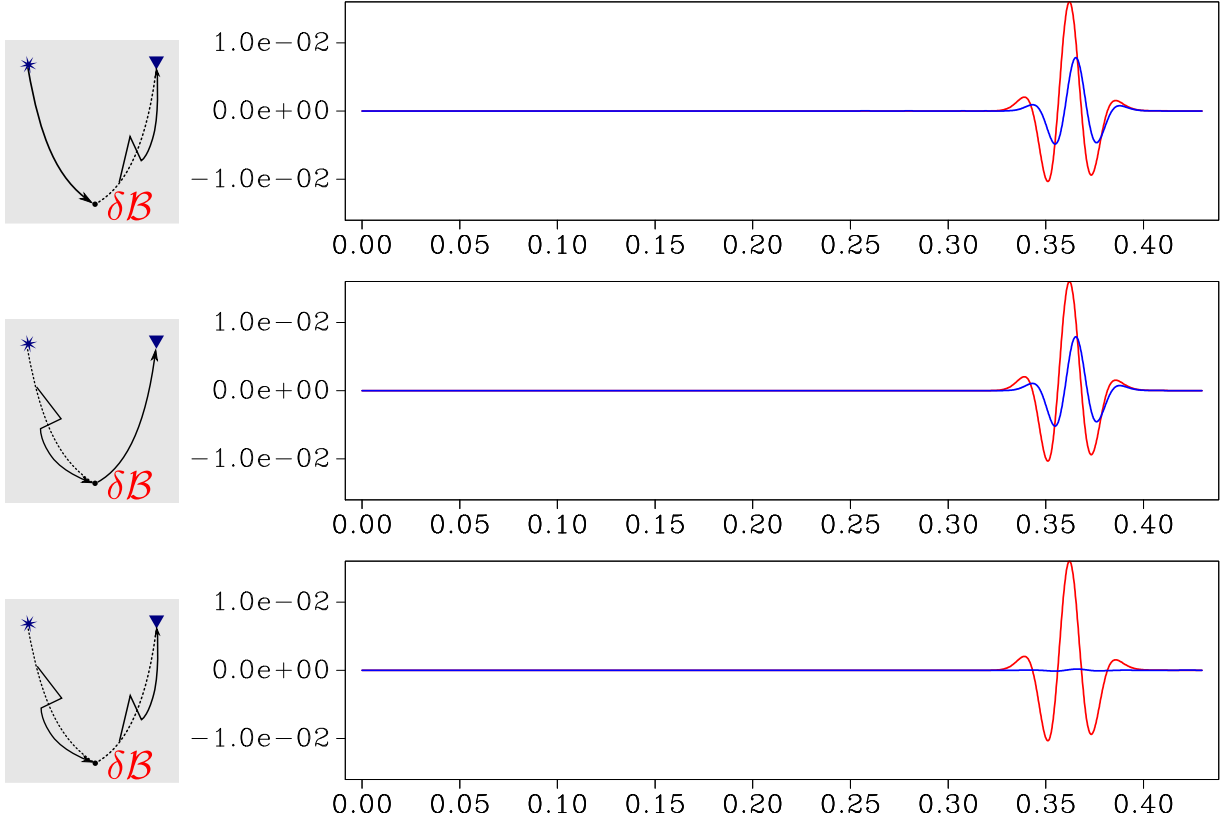


Figure 3.8: EXPERIMENT 1. Multiple-scattering contributions (blue lines), from top to bottom:  $\delta p_{S,BBS}$ ,  $\delta p_{S,SBB}$  and  $\delta p_{S,SBS}$ . In all figures,  $\delta p_S^{\text{true}}$  is depicted in red lines for comparison. The first two terms carry most of the energy of the singular residual  $\delta p_S$ .

From the mathematical expressions of  $\delta p_{S,-B\mathcal{L}B}$  and  $\delta p_{S,B\mathcal{B}B}$ , we know that both terms slightly differ in the secondary-source potential operator applied to the source-side background wavefield. Because of the negative sign of  $\delta p_{S,-B\mathcal{L}B}$ , they are almost complementary to one another and, in this experiment, end up canceling almost completely their main contribution. In Appendix D we show the resulting expression of the superposition of these terms,  $\delta p_{S,-B\mathcal{W}B}$ . The physical reason for this behaviour is that the influence of the background perturbation is already accounted for in the perturbation of the background wavefield  $\delta p_B$ . As we going to see later in this chapter,  $\delta p_{S,-B\mathcal{W}B}$  is non zero only when the background perturbation overlaps the singular component of the model within the Fresnel zone of the total wavefield, which is not the case in this experiment.

The next terms we analyse are the *multiple-scattering contributions*  $\delta p_{S,BBS}$ ,  $\delta p_{S,SBB}$  and  $\delta p_{S,SBS}$ . In Figure 3.8 we see that the first two terms are the ones that answer for most of the

### 3.3 Experiment gallery

---

energy carried by the singular residual,  $\delta p_S$ . The small contribution coming from the strong multiple-scattering term  $\delta p_{S,BBS}$  is explained by the lack of other scatterers/reflectors the single reflector could interact with, leaving the very weak self-interaction of the reflector the only source of energy to this term.

Finally, we analyse the remaining three *non-symmetrical contributions*  $\delta p_{S,-BLS}$ ,  $\delta p_{S,bVB}$  and  $\delta p_{S,bVS}$  for which no corresponding contributions are present in eq. (2.33) and Figure 2.4. These contributions are depicted in the Figure 3.9. At first view, they seem to bear no resemblance to each other. The contribution  $\delta p_{S,-BLS}$  carries about as much energy as the related contribution  $\delta p_{S,BBS}$ , though with a different signal. Actually they have the same relation as terms  $\delta p_{S,-BLB}$  and  $\delta p_{S,BBB}$  have with each other: both terms are almost complementary to each other, differing only slightly in the secondary-source potential operator applied to the source-side singular wavefield. As in the single-scattering situation above, the superposition term  $\delta p_{S,-BWS}$  carries energy only if the background perturbations overlap with the medium singularities (Appendix D). On the other hand, the terms  $\delta p_{S,bVB}$  and  $\delta p_{S,bVS}$  (second and third plots in Figure 3.9) apparently carry no energy.

That being said, let's take a closer look in these two terms:  $\delta p_{S,bVB}$  and  $\delta p_{S,bVS}$ . Their summation in the frequency domain yields (angular frequency  $\omega$  omitted):

$$\widehat{\delta p_{S,bVB}}(\mathbf{x}_g; \mathbf{x}_s) + \widehat{\delta p_{S,bVS}}(\mathbf{x}_g; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' \left( \widehat{G}_B(\mathbf{x}_g; \mathbf{x}') + \widehat{G}_S(\mathbf{x}_g; \mathbf{x}') \right) \mathcal{V}_0 \left[ \widehat{\delta p_B}(\mathbf{x}'; \mathbf{x}_s) \right]. \quad (3.3)$$

Recognising  $G_B + G_S = G_0$  and inserting  $\widehat{\delta p_B}$  from the Fourier transform of eq. (2.25) yields

$$\widehat{\delta p_{S,bVB}}(\mathbf{x}_g; \mathbf{x}_s) + \widehat{\delta p_{S,bVS}}(\mathbf{x}_g; \mathbf{x}_s) = \int_{\mathbb{V}} d^3 \mathbf{x}' \widehat{G}_0(\mathbf{x}_g; \mathbf{x}') \mathcal{V}_0 \left[ \int_{\mathbb{V}} d^3 \mathbf{x} \widehat{G}_B(\mathbf{x}'; \mathbf{x}) \delta \mathcal{L}_B [\widehat{p_B}(\mathbf{x}; \mathbf{x}_s)] \right]. \quad (3.4)$$

Note that  $\mathcal{V}_0$  acts on  $\mathbf{x}'$ , so it can be brought into the integral on  $\mathbf{x}$  and be applied only onto  $G_B$ . After changing the order of integration, we find

$$\widehat{\delta p_{S,bVB}}(\mathbf{x}_g; \mathbf{x}_s) + \widehat{\delta p_{S,bVS}}(\mathbf{x}_g; \mathbf{x}_s) = \int_{\mathbb{V}} d^3 \mathbf{x} \underbrace{\int_{\mathbb{V}} d^3 \mathbf{x}' \widehat{G}_0(\mathbf{x}_g; \mathbf{x}') \mathcal{V}_0 \left[ \widehat{G}_B(\mathbf{x}'; \mathbf{x}) \right]}_{-\widehat{G}_s(\mathbf{x}_g; \mathbf{x})} \delta \mathcal{L}_B [\widehat{p_B}(\mathbf{x}; \mathbf{x}_s)], \quad (3.5)$$

where we recognise the inner integral highlighted by the curly braces to be the Fourier transform

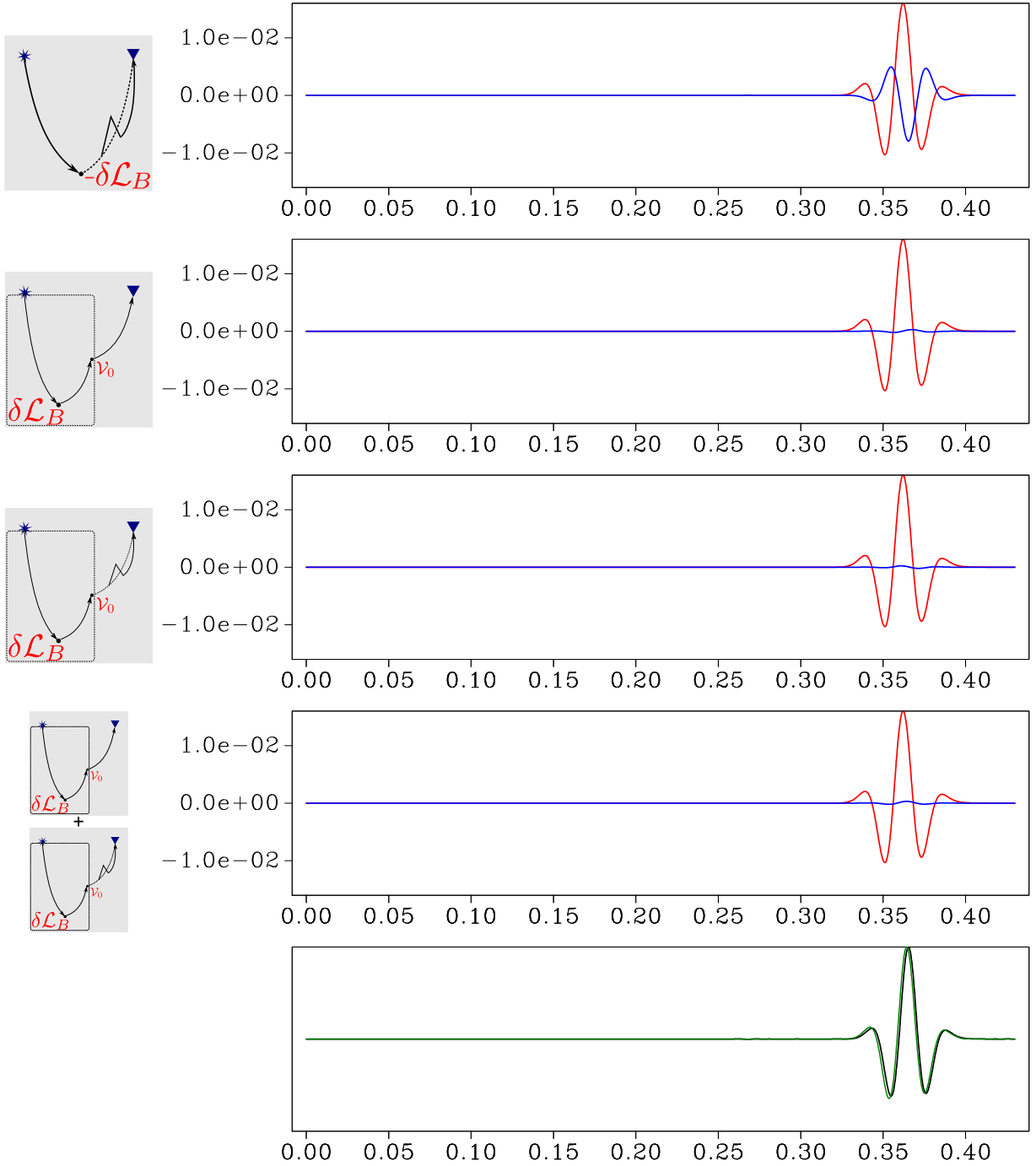


Figure 3.9: EXPERIMENT 1. Non-symmetrical contributions (blue lines), first four traces from top to bottom:  $\delta p_{S,-BLS}$ ,  $\delta p_{S,bVB}$ ,  $\delta p_{S,bVS}$  and  $\delta p_{S,bVB} + \delta p_{S,bVS}$  respectively. In these traces,  $\delta p_S^{\text{true}}$  is depicted in red lines for comparison. The last trace show both  $-\delta p_{S,-BLS}$  (black line) and  $\delta p_{S,bVB} + \delta p_{S,bVS}$  (dark green line) contributions normalised to their respective maximum absolute value.

### 3.3 Experiment gallery

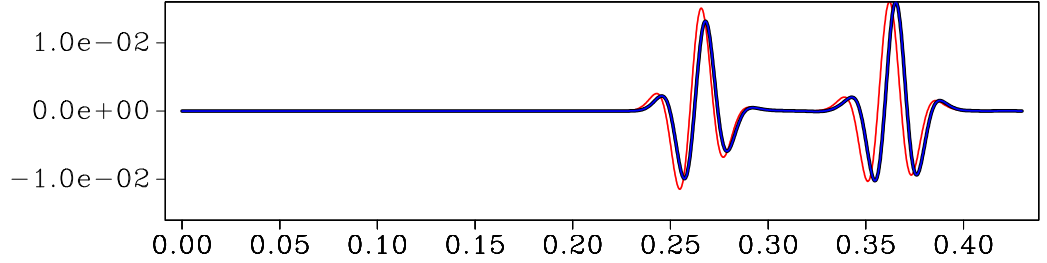


Figure 3.10: EXPERIMENT 1. Comparing  $\delta p^{\text{true}}$  (red line) with evaluated counter-parts. In black,  $\delta p$  (evaluated with eq. (2.14)) and, in blue,  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all but the 1<sup>st</sup>, 2<sup>nd</sup> and 7<sup>th</sup> terms of eq. (2.33) respectively). See the perfect match between blue and black lines.

of  $-G_S(\mathbf{x}_g, t; \mathbf{x})$  in full analogy to eq (2.22). Therefore, according to eq. (2.33.7) and Figure 2.4c:

$$\begin{aligned} \widehat{\delta p}_{S,bVB}(\mathbf{x}_g; \mathbf{x}_s) + \widehat{\delta p}_{S,bVS}(\mathbf{x}_g; \mathbf{x}_s) &= - \int_{\mathbf{V}} d^3 \mathbf{x} \widehat{G}_s(\mathbf{x}_g; \mathbf{x}) \delta \mathcal{L}_B [\widehat{p}_B(\mathbf{x}; \mathbf{x}_s)] \\ &= -\widehat{\delta p}_{S,BLS}(\mathbf{x}_g; \mathbf{x}_s). \end{aligned} \quad (3.6)$$

How to explain the disagreement between the theoretical prediction above and the numerical results depicted in Figure 3.9 (first and fourth traces)? The evaluation of the Born approximation via modeller injection introduces an amplitude error. The contribution  $\delta p_{S,BLS}$  is evaluated with *one* injection of the secondary source from the source at the background perturbation position  $\mathbf{x}$ . On their turn, both contributions  $\delta p_{S,bVB}$  and  $\delta p_{S,bVS}$  are evaluated with the injection of *two* secondary sources: the first is done at the background perturbation position  $\mathbf{x}$  with the secondary source generated by the wavefield coming from the source; the second at the scattering potential position  $\mathbf{x}'$  with the secondary source generated by the wavefield coming from  $\mathbf{x}$ . Thus, these contributions suffer twice from the amplitude error of the modeller. To compare the contributions without the influence of the amplitude error, the last part of Figure 3.9 shows both  $-\delta p_{S,-BLS}$  (black line) and  $\delta p_{S,bVB} + \delta p_{S,bVB}$  (dark gree line) normalised to their maximum value. We can observe that the normalised contributions agree very well.

Another way to verify the result seen in eq. (3.6) is by comparing  $\delta p^{\text{true}}$ ,  $\delta p$  (evaluated with eq. (2.14)),  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all but the 1<sup>st</sup>, 2<sup>nd</sup> and 7<sup>th</sup> terms of eq. (2.33) respectively). This is shown in Figure 3.10. When  $\delta p_{S,BLS}$ ,  $\delta p_{S,bVB}$  and  $\delta p_{S,bVB}$  are *not* taken into account the match between  $\delta p$  and  $\delta p_B + \delta p_S$  is perfect.

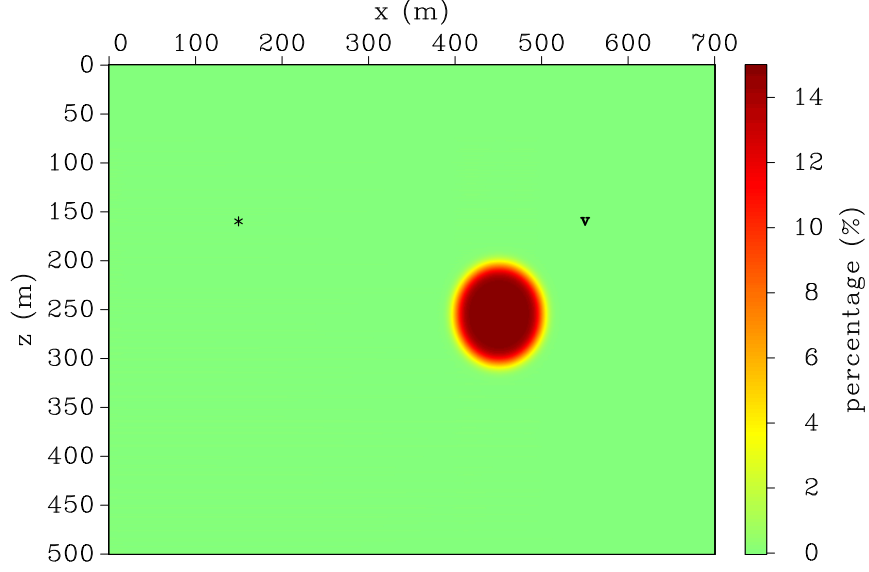


Figure 3.11: EXPERIMENT 2. Perturbation in the bulk-modulus background model,  $\delta K_B$ . The scalebar shows, in percentage, the perturbation magnitude.

### 3.3.2 Experiment 2: Lens-shaped background perturbation of Model I, backprojecting residuals

In Experiment 2 we departed from the same unperturbed, single-reflector Model I as the previous one (see Figure 3.1) but introduced a spatially limited (lens-shaped) perturbation in its background as shown in Figure 3.11. As in the previous experiment, we kept the singular part unperturbed. Again, we have perturbations in both the background and singular wavefields according to equation (3.2).

In this experiment, we are interested in the behaviour of the contributions in the backprojections. Therefore, we backproject both the true background and singular residuals,  $\delta p_B^{\text{true}}$  and  $\delta p_S^{\text{true}}$  with the proper adjoint kernel as seen in eq. (2.50). Based on the result presented in eq. (3.6) we can restate that equation as:

$$\begin{bmatrix} \delta K_B^{\text{est}} \\ \delta K_S^{\text{est}} \end{bmatrix} = \begin{bmatrix} \delta K_{B,B}^{\text{est}} + \sum_i \delta K_{B,i}^{\text{est}} \\ \sum_j \delta K_{S,j}^{\text{est}} \end{bmatrix} = \begin{bmatrix} \bar{U}_B^\dagger & \sum_i \bar{U}_{S,i}^\dagger \\ 0 & \sum_j \bar{U}_{S,j}^\dagger \end{bmatrix} \begin{bmatrix} \hat{\delta p}_B \\ \hat{\delta p}_S \end{bmatrix}, \quad (3.7)$$

where the summation indices take only the values  $i = BS, BB, SS, SB, -B\mathcal{L}B$  and  $j = BB, BS, SB, SS$ . Notice, as previously determined, that the estimates  $\delta K_{B,b\nu S}^{\text{est}}$ ,  $\delta K_{B,b\nu B}^{\text{est}}$  and  $\delta K_{B,-B\mathcal{L}S}^{\text{est}}$

### 3.3 Experiment gallery

---

cancel each other out.

Despite its simplicity, this is a most elucidative experiment which clearly exemplifies the decomposition discussed here. The first result is displayed in Figure 3.12. There we can see the bulk-modulus estimate  $\delta K^{\text{est}}$  (*top*). Here and in the following residual-backprojection experiments,  $\delta K^{\text{est}}$  is always obtained with the backpropagation of the total wavefield residual  $\delta p$  according to the first equation of (2.49), i.e., it is evaluated using the reference Green's function in the adjoint to equation (2.16). In this experiment,  $\delta K^{\text{est}}$  is the estimate to be decomposed. After that, from left to right, from top to bottom in Figure 3.12, we observe the six components of the background bulk modulus estimate  $\delta K_B^{\text{est}}$  expressed in the first equation of (3.7). In the upper-left corner we have the bulk modulus estimate  $\delta K_{B,B}^{\text{est}}$  result of backprojecting the background residual  $\delta p_B$  followed by the contributions resulting from the backprojection of the singular residual  $\delta p_S$  with the five adjoint sub-kernels:  $\delta K_{B,-B\mathcal{L}B}^{\text{est}}$ ;  $\delta K_{B,-BS}^{\text{est}}$ ;  $\delta K_{B,-SB}^{\text{est}}$ ;  $\delta K_{B,-BB}^{\text{est}}$  and  $\delta K_{B,-SS}^{\text{est}}$  respectively. Comparing each contribution with the estimate  $\delta K^{\text{est}}$  in the top figure one can clearly identify all the different “pieces” that compose the complete estimate.

Analysing each one of the contributions shown in Figure 3.12, we can draw some conclusions about their physical meaning. Term  $\delta K_{B,B}^{\text{est}}$  is related to the direct-wave Fresnel zone. Both  $\delta K_{B,-B\mathcal{L}B}^{\text{est}}$  and  $\delta K_{B,-BB}^{\text{est}}$  are related to the RTM migration operator, their difference residing in the presence or not of the singular part of the reference model, i.e., which of the factors,  $\frac{1}{K_B}$  or  $\frac{1}{K_0}$ , is used in the scattering mechanism operator. The contribution  $\delta K_{B,-SB}^{\text{est}}$  is related to the receiver-side reflecting Fresnel zone;  $\delta K_{B,-BS}^{\text{est}}$  to the source-side reflecting Fresnel zone; and  $\delta K_{B,-SS}^{\text{est}}$  to the reflector self-interaction Fresnel zone.

This simple example can be used to give more insight into the differences of three important estimates:

- the conventional perturbation estimate  $\delta K_{\text{conv}}^{\text{est}}$  obtained with the conventional approach, i.e., backprojecting the total residual  $\delta p$  with the adjoint kernel based on the background wavefields;
- the total perturbations estimate  $\delta K^{\text{est}}$  obtained with aid of first line of eq. (2.49); and
- the background perturbation estimate  $\delta K_B^{\text{est}} = \delta K_{B,B}^{\text{est}} + \delta K_{B,S}^{\text{est}}$  obtained as seen in the first line of eq. (2.50).

Figure 3.13 shows these estimates for comparison. In the top, we can see the total estimate  $\delta K^{\text{est}}$ . In the upper-left corner we can see the conventional estimate  $\delta K_{\text{conv}}^{\text{est}}$  obtained with the conventional

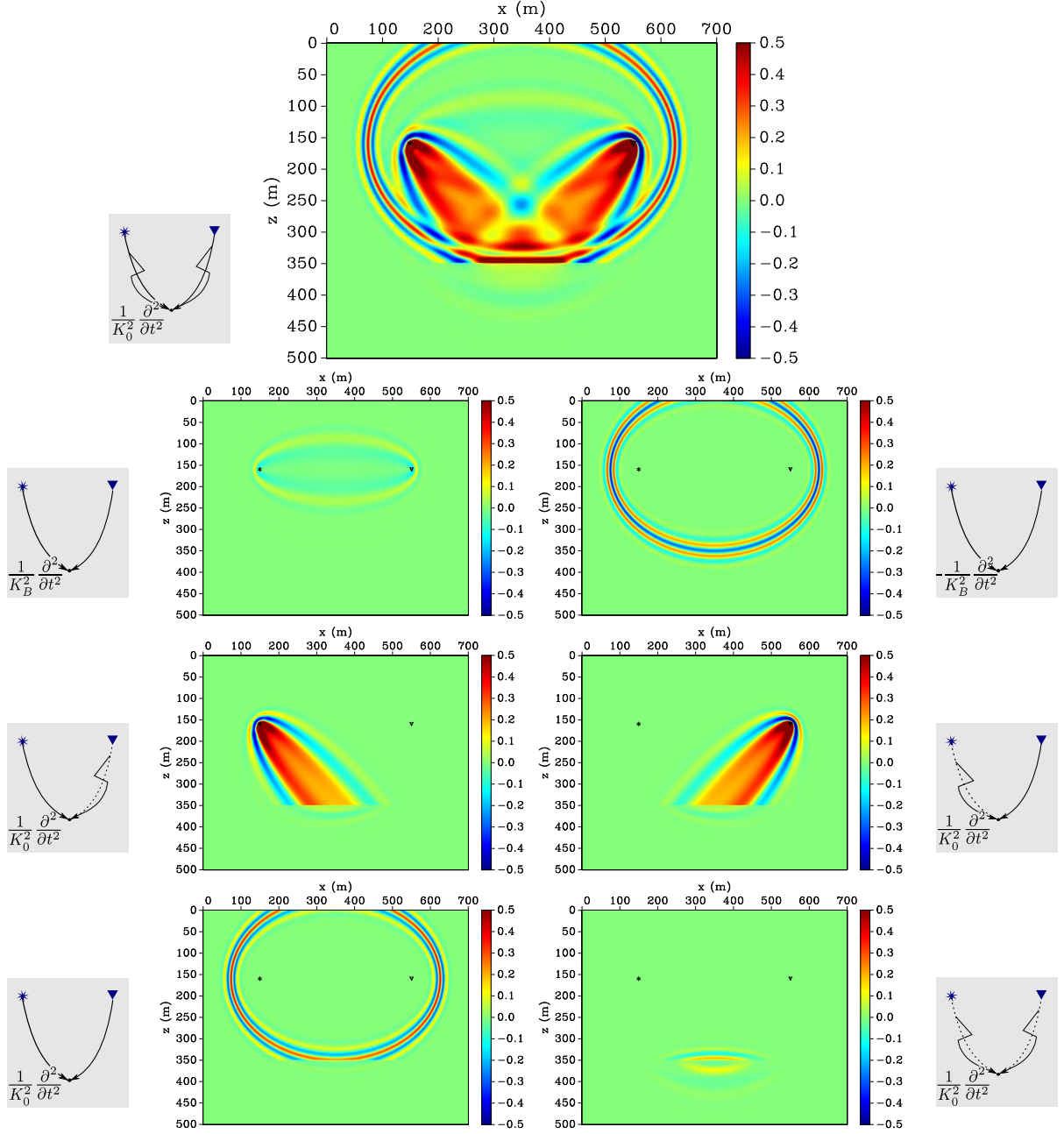


Figure 3.12: EXPERIMENT 2. *Top*: Bulk modulus estimate  $\delta K^{\text{est}}$ . The others are all background bulk-modulus-estimate contributions. *From left to right, top to bottom*:  $\delta K_{B,B}^{\text{est}}$ ;  $\delta K_{B,-B\mathcal{L}B}^{\text{est}}$ ;  $\delta K_{B,-BS}^{\text{est}}$ ;  $\delta K_{B,-SB}^{\text{est}}$ ;  $\delta K_{B,-BB}^{\text{est}}$  and  $\delta K_{B,-SS}^{\text{est}}$ . The first of them in the upper-left corner,  $\delta K_{B,B}^{\text{est}}$ , is result of background residual backprojection; the rest of them are result of singular residual backprojection. The subtraction of  $\delta K_{B,B}^{\text{est}}$  (upper-left corner) with  $\delta K_{B,-B\mathcal{L}B}^{\text{est}}$  (upper-right corner) yields the conventional estimate  $\delta K_{\text{conv}}^{\text{est}}$  (Figure 3.13, upper-left corner).

### 3.3 Experiment gallery

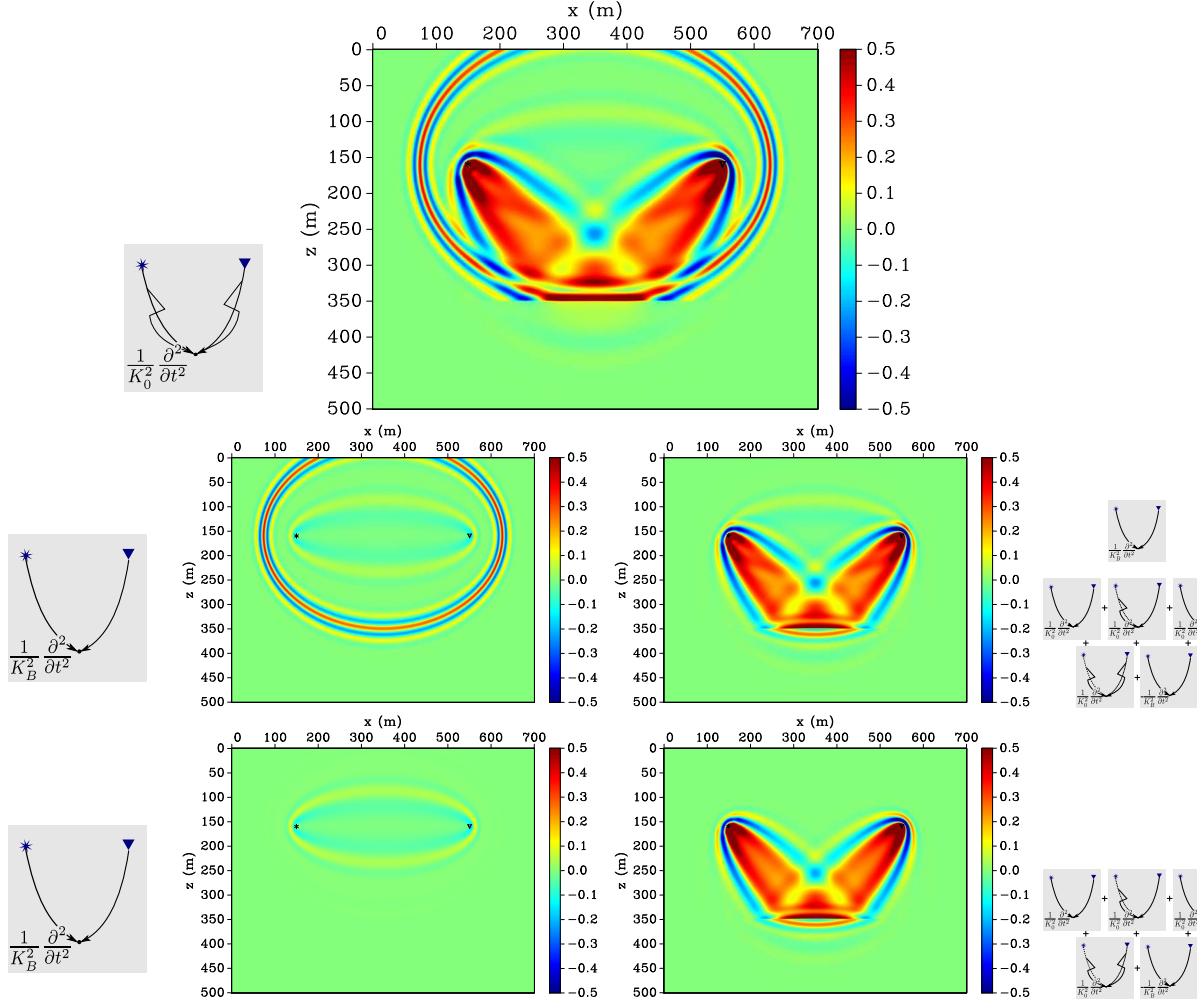


Figure 3.13: EXPERIMENT 2. *Top*: Bulk modulus estimate  $\delta K^{\text{est}}$ . *Upper-left corner*: Bulk modulus estimate  $\delta K_{\text{conv}}^{\text{est}}$  obtained with the conventional approach. The others are all background bulk-modulus-estimate contributions. *Upper-right corner*: the background estimate  $\delta K_B^{\text{est}} = \delta K_{B,B}^{\text{est}} + \delta K_{B,S}^{\text{est}}$ . *Lower-left corner*:  $\delta K_{B,B}^{\text{est}}$ , result of background residual backprojection. *Lower-right corner*:  $\delta K_{B,S}^{\text{est}}$ , result of singular residual backprojection (sum of all five contributions).

approach. It is constructed by backprojecting the total residual  $\delta p$  with the adjoint kernel based on the background wavefields, which, according to our definition, would be  $U_B^\dagger$ . Mathematically:

$$\begin{aligned} \overline{\delta K}_{\text{conv}}^{\text{est}} &= \overline{U}_B^\dagger \widehat{\delta p} = \overline{U}_B^\dagger (\widehat{\delta p}_B + \widehat{\delta p}_S) = \overline{U}_B^\dagger \widehat{\delta p}_B - \overline{U}_{S,-B\mathcal{L}B}^\dagger \widehat{\delta p}_S \\ &= \overline{\delta K}_{B,B}^{\text{est}} - \overline{\delta K}_{B,-B\mathcal{L}B}^{\text{est}} \end{aligned} \quad (3.8)$$



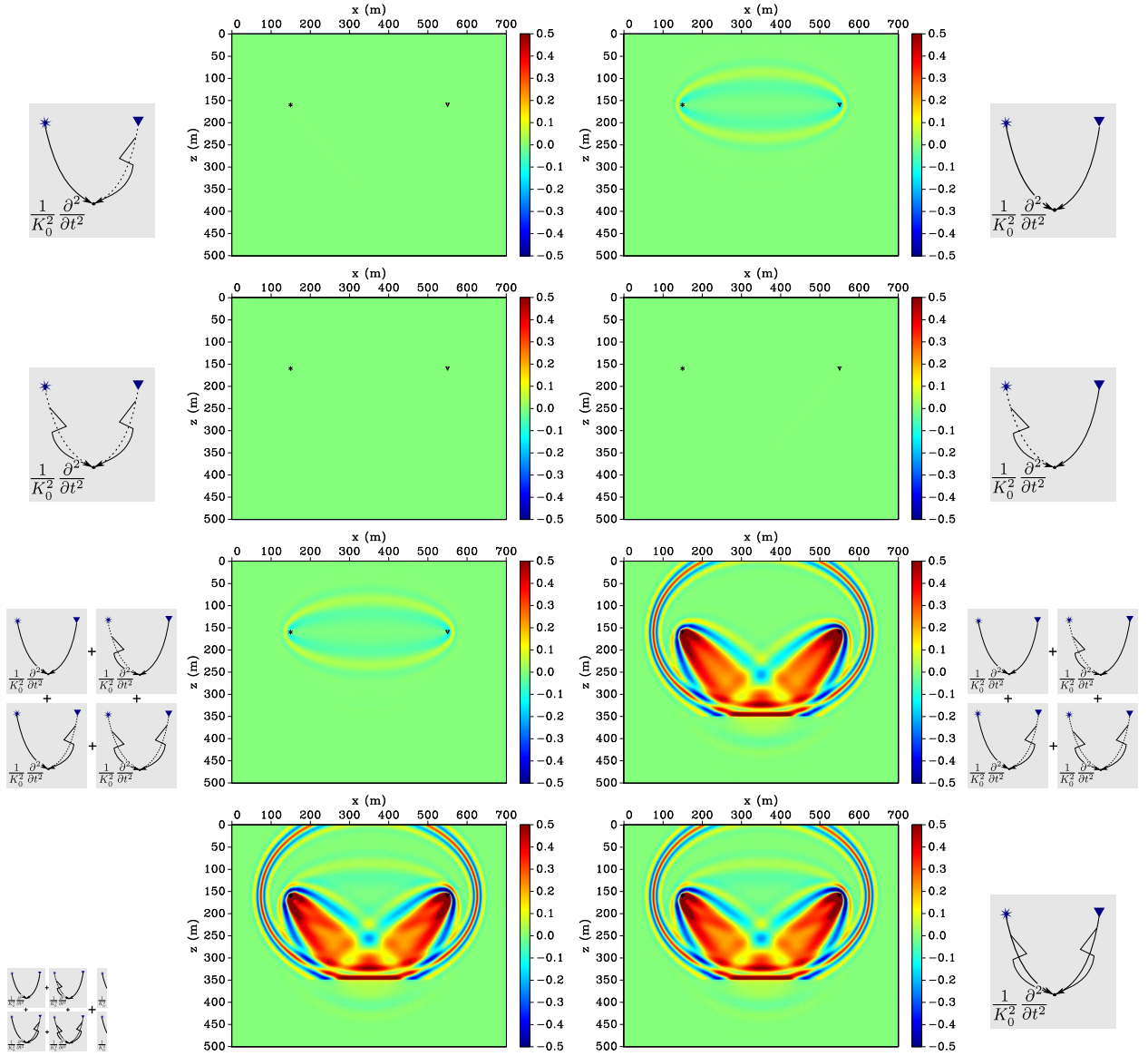


Figure 3.14: EXPERIMENT 2. First four estimates are obtained by backprojecting the background residual with the adjoint sub-kernels  $\bar{U}_{S,j}^\dagger$ , i.e.,  $\bar{\delta K}_{j(B)}^{\text{est}}$  (cross-terms). *From left to right, top to bottom:*  $\bar{\delta K}_{BS(B)}^{\text{est}} = \bar{U}_{S,BS} \hat{p}_B$ ;  $\bar{\delta K}_{BB(B)}^{\text{est}} = \bar{U}_{S,BB}^\dagger \hat{p}_B$ ;  $\bar{\delta K}_{SS(B)}^{\text{est}} = \bar{U}_{S,SS}^\dagger \hat{p}_B$ ;  $\bar{\delta K}_{SB(B)}^{\text{est}} = \bar{U}_{S,SB}^\dagger \hat{p}_B$ . The next two figures are each one of the sums seen in the right-hand side of eq. (3.11). *From left to right:*  $\left(\sum_j \bar{U}_{S,j}^\dagger\right) \hat{p}_B$  and  $\left(\sum_j \bar{U}_{S,j}^\dagger\right) \hat{p}_S$ . At the lower-left corner is displayed the summation of the previous sums, i.e.,  $\bar{\delta K}^{\text{est}}$  of eq. (3.11). At the lower-right corner, the bulk-modulus estimate  $\delta K^{\text{est}}$  for comparison.

### 3.3 Experiment gallery

---

for we know that  $\bar{U}_B = -\bar{U}_{S,-B\mathcal{L}B}$ . It show that the conventional estimate is obtained with the subtractions of the two contributions depicted in the second row of Figure 3.12. Definitely, there is a lot of information missing in  $\delta K_{\text{conv}}^{\text{est}}$  when compared to  $\delta K^{\text{est}}$ . From this observation we can already conclude that it is important to include an estimate of  $K_S$  in the starting model.

Next, in the upper-right corner of Figure 3.13, we can see the background estimate  $\delta K_B^{\text{est}}$  which is the summation of the estimate  $\delta K_{B,B}^{\text{est}}$  (lower-left corner) with the estimate  $\delta K_{B,S}^{\text{est}}$  (lower-right corner). The former is the result of background residual backprojection (seen also if Figure 3.12), the last is the sum of all five contributions  $\delta K_{B,i}^{\text{est}}$  with  $i = BS, BB, SS, SB, -B\mathcal{L}B$ , result of the singular residual backprojection. As the contributions  $\delta K_{B,BB}^{\text{est}}$  and  $\delta K_{B,-B\mathcal{L}B}^{\text{est}}$  are almost complementary, almost all information of  $\delta K_{B,S}^{\text{est}}$  comes from the estimates  $\delta K_{B,BS}^{\text{est}}, \delta K_{B,SS}^{\text{est}}, \delta K_{B,SB}^{\text{est}}$ .

The most important difference between the conventional estimate  $\delta K_{\text{conv}}^{\text{est}}$  and the background estimate  $\delta K_B^{\text{est}}$  (upper-left and -right corners of Figure 3.13) resides in the illuminated region by each estimate. The conventional estimates practically does not see the subsurface. The only information from deep region comes from the short-wavelength migration-related contribution  $-\delta K_{B,-B\mathcal{L}B}^{\text{est}}$ . On the other hand, the estimate  $\delta K_B^{\text{est}}$  benefits from the back-illumination provided by the single reflector, which allows the estimate to see a broader, deeper area with a larger wavelength.

As in this experiment only perturbation in the background was introduced, i.e.,  $\delta K = \delta K_B$ , one might expect that  $\delta K^{\text{est}}$  and  $\delta K_B^{\text{est}} = \delta K_{B,B}^{\text{est}} + \delta K_{B,S}^{\text{est}}$  would give the same estimate which is not true when we compare the first (*top*) and the last (*lower-right*) estimates of Figure 3.13. To understand why these two estimates do not agree, we must go back to the forward problem formulation. As  $\hat{\delta p} = \hat{\delta p}_B + \hat{\delta p}_S$ , from equations (2.15) and (3.2) we have that

$$\hat{\delta p} = \hat{\delta p}_B + \hat{\delta p}_S = \bar{U} \bar{\delta K}_B = \bar{U}_B \bar{\delta K}_B + \bar{U}_{S,-B\mathcal{L}B} \bar{\delta K}_B + \left( \sum_j \bar{U}_{S,j} \right) \bar{\delta K}_B, \quad (3.9)$$

where  $j = BS, BB, SS, SB$ . Because  $\bar{U}_B = -\bar{U}_{S,-B\mathcal{L}B}$ , we have

$$\bar{U} = \sum_j \bar{U}_{S,j}. \quad (3.10)$$

Thus we obtain from  $\bar{\delta K}^{\text{est}} = \bar{U}^\dagger (\hat{\delta p}_B + \hat{\delta p}_S)$  and  $\bar{\delta K}_B^{\text{est}} = \bar{U}_B^\dagger \hat{\delta p}_B + \left( \sum_i \bar{U}_{S,i}^\dagger \right) \hat{\delta p}_S$ :

$$\bar{\delta K}^{\text{est}} = \left( \sum_j \bar{U}_{S,j}^\dagger \right) \hat{\delta p}_B + \left( \sum_j \bar{U}_{S,j}^\dagger \right) \hat{\delta p}_S \quad (3.11)$$

$$\overline{\delta K}_B^{\text{est}} = \overline{U}_B^\dagger \widehat{\delta p}_B + \overline{U}_{S,-B\mathcal{L}B}^\dagger \widehat{\delta p}_S + \left( \sum_j \overline{U}_{S,j}^\dagger \right) \widehat{\delta p}_S \quad (3.12)$$

where  $i = BS, BB, SS, SB, -B\mathcal{L}B$  and  $j = BS, BB, SS, SB$ . We see that, besides the terms already known, some new *cross-terms* come into play in the expression of the estimate  $\overline{\delta K}^{\text{est}}$ :  $\overline{U}_{S,j}^\dagger \widehat{\delta p}_B$ . We are going to refer to these cross-terms as  $\overline{\delta K}_{j(B)}^{\text{est}}$  indicating that these are components of the total bulk-modulus estimate ( $\overline{\delta K}^{\text{est}}$ ) obtained by backprojecting the background residual (thus, the index  $(B)$ ) with the subkernel  $\overline{U}_{S,j}$ .

In Figure 3.14, we analyse the components of  $\overline{\delta K}^{\text{est}}$  in eq. (3.11) ( $\overline{\delta K}_B^{\text{est}}$ , in eq. (3.12), was already displayed in Figure 3.13, lower-left corner). The first conclusion, comparing the last two estimates, is that the decomposition in equation (3.11) correctly accounts for all components of  $\overline{\delta K}^{\text{est}}$ : the first four estimates of Figure 3.14 correspond to the first sum in the right-hand side; and the last four estimates of Figure 3.12 correspond to the second sum. Nevertheless, the one relevant contribution among those resulting of background residual backpropagation – first four estimate of the Figure – is  $\overline{\delta K}_{BB(B)}^{\text{est}} = \overline{U}_{S,BB}^\dagger \widehat{\delta p}_B$  which, in this case, is exactly the same as  $\overline{\delta K}_{B,B}^{\text{est}} = \overline{U}_B^\dagger \widehat{\delta p}_B$ . Therefore the actual difference between  $\delta K_B^{\text{est}}$  and  $\delta K^{\text{est}}$  is the presence (or not) of the contribution  $\overline{\delta K}_{B,-B\mathcal{L}B}^{\text{est}}$  which, if present, cancels the migration- related contribution  $\overline{\delta K}_{B,BB}^{\text{est}}$ .

### 3.3.3 Experiment 3: Constant background perturbation of Model II, forward decomposition

In the first experiment with this model, we kept the scatterers in place and introduced a 10% perturbation to the complete background part of Model II, i.e.,  $\delta K_B \neq 0$  and  $\delta K_S = 0$ . As in the previous experiment, this yields a perturbation in both the background and singular wavefields.

We fired a single source and recorded the calculated wavefields at two control receivers (black star and triangles in Figure 3.2, respectively). The control receivers are named cr154 (*above the scatterers*) and cr283 (*below the scatterers*).

First, we compare the true and estimated values of the wavefield perturbations and its components at the control receiver. In this experiment, we are focusing our analysis on the response at cr154.

The first contribution to be studied is the background-wavefield residual  $\delta p_B$ . We observe a very good match between waveforms at receiver cr154 (Figure 3.15) located relatively close to the source. There are, though, some small fluctuation in the linearised residual  $\delta p_B$  (blue line) at 0.15

### 3.3 Experiment gallery

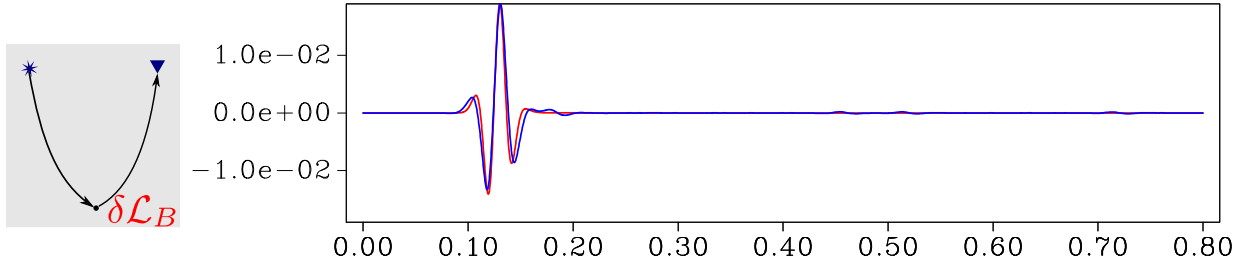


Figure 3.15: EXPERIMENT 3. Linearised background-wavefield residuals  $\delta p_B$  (blue line) evaluated at cr154 as compared to the true ones (red line),  $\delta p_B^{\text{true}}$ .

(right after the main signal), 0.50 and 0.70 s. These are none but the residual reflections at the top border, bottom and right borders (overlapping), and left border, respectively, despite the absorbing boundaries conditions used in the modeller.

Next, we address the singular-wavefield residual  $\delta p_S$  and its eight components due to perturbations of the background model.

We start by analysing the terms  $\delta p_{S,-B\mathcal{L}B}$  and  $\delta p_{S,B\mathcal{B}B}$ . An indication of the importance of each contributing term is the energy it carries. Table 3.3 shows the energy carried by each of the eight terms contributing to the singular residual observed at cr154 as measured by means of its  $L^2$ -norm. The table shows that the most energetic contributions correspond to single-scattering terms represented by these two terms. But this is by no means a rule. Here, as in Experiment 1, we can see that these terms are related to the background residual rather than to the singular one. If, by any reason, the background residual carries little energy, so will these terms or, at least, the estimate  $\delta p_{S,-B\mathcal{L}B}$ . However, differently from the previous forward experiment, their summation,  $\delta p_{S,-B\mathcal{W}B}$ , has small but non-negligible contribution (Figure 3.16, last trace). This happens because in this example the background perturbation that overlaps the singular part of the model (scatterers) lies within the Fresnel zone of the complete non-perturbed wavefield. Nevertheless, its contribution is secondary.

As seen in Experiment 1, we expected the terms  $\delta p_{S,-B\mathcal{L}S}$  and  $\delta p_{S,b\nu S} + \delta p_{S,b\nu B}$  (first two traces of Figure 3.17) to cancel each other, although, due to numerical issues already discussed, this does not happen. As before, we normalise both terms to their respective maximum value and compare them (last trace of the same figure). The match, although good, is not as perfect as seen in the last trace of Figure 3.9 indicating that numerically they do not cancel each other out completely. However, in order to support that this fact is due to numerical issues rather than a theoretical flaw, we can compare  $\delta p^{\text{true}}$ ,  $\delta p$  (evaluated with eq. (2.14)),  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all

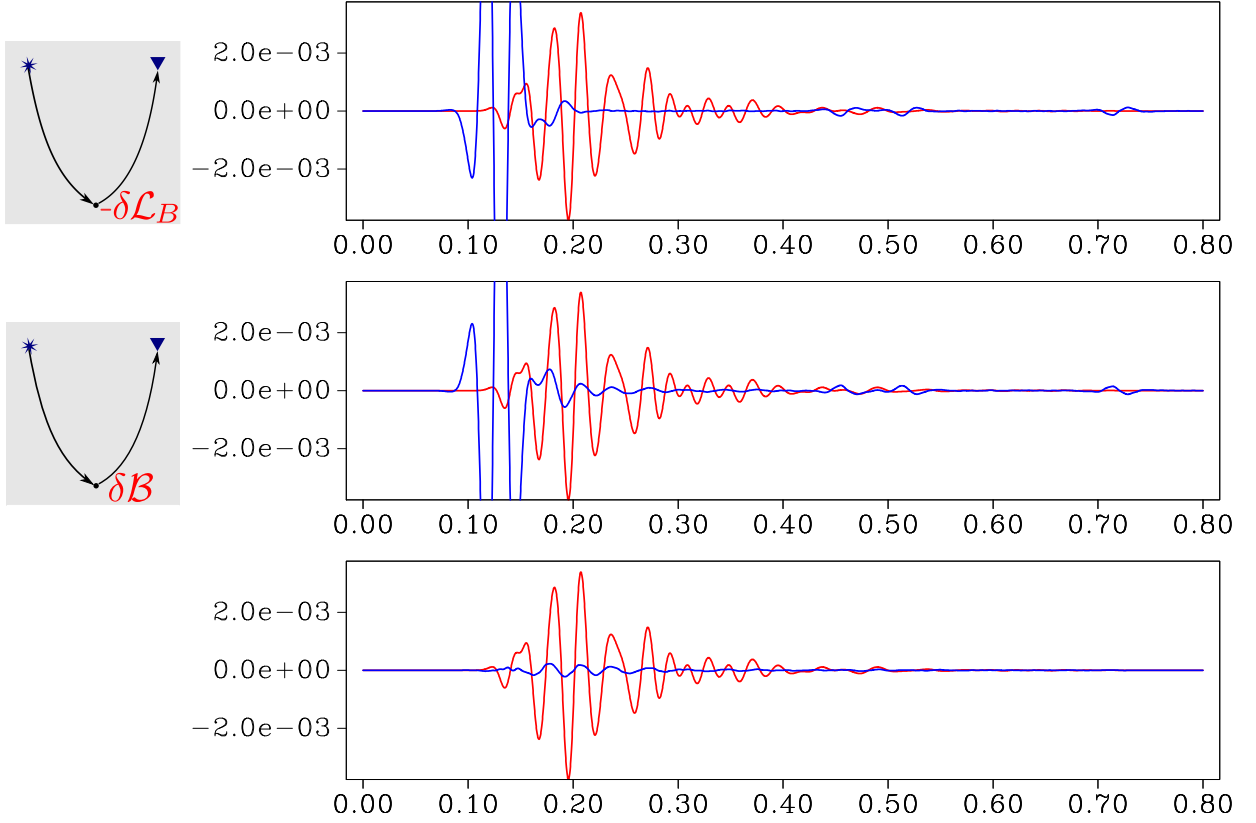


Figure 3.16: EXPERIMENT 3. Terms related to single-scattering effects at the medium perturbations evaluated at cr154 (blue lines):  $\delta p_{S,-B\mathcal{L}B}$  (top),  $\delta p_{S,B\mathcal{B}B}$  (centre), and their sum (bottom). The red line true singular residual  $\delta p_S^{\text{true}}$ . Separately,  $\delta p_{S,-B\mathcal{L}B}$  and  $\delta p_{S,B\mathcal{B}B}$  carry the largest energy. Their sum,  $\delta p_{S,-B\mathcal{W}B}$ , demonstrates that they are approximately complementary.

terms of eq. (2.33) respectively) and  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all but the 1<sup>st</sup>, 2<sup>nd</sup> and 7<sup>th</sup> terms of eq. (2.33) respectively). This is shown in Figure 3.18. When  $\delta p_{S,B\mathcal{L}S}$ ,  $\delta p_{S,b\nu B}$  and  $\delta p_{S,b\nu B}$  are *not* taken into account the match between  $\delta p$  and  $\delta p_B + \delta p_S$  is close to perfect.

Therefore, hereafter, we refer only to the terms  $\delta p_{S,i}$  with  $i = B\mathcal{B}S, B\mathcal{B}B, S\mathcal{B}S, S\mathcal{B}B, -B\mathcal{L}B$  as the background-dependent contributions to the singular residual.

That leaves the last three terms,  $\delta p_{S,B\mathcal{B}S}$ ,  $\delta p_{S,S\mathcal{B}S}$  and  $\delta p_{S,S\mathcal{B}B}$  as the most important contributions to the total singular-wavefield residual  $\delta p_S$ . The upper three parts of Figure 3.19 compare these source-side multiple-scattering contributions (blue lines) individually to  $\delta p_S^{\text{true}}$  (red lines). We see that the contributions  $\delta p_{S,B\mathcal{B}S}$  and  $\delta p_{S,S\mathcal{B}B}$  do similar jobs trying to match the true singular residual. The strong multiple-scattering contribution  $\delta p_{S,S\mathcal{B}S}$  apparently plays a secondary role.

### 3.3 Experiment gallery

Term	$L^2$ -norm ( $\times 10^{-5}$ )	Term	$L^2$ -norm ( $\times 10^{-5}$ )
$\delta p_{S,BBB}$	387.59	$\delta p_{S,BBS}$	65.40
$\delta p_{S,-B\mathcal{L}B}$	386.67	$\delta p_{S,SBS}$	17.27
$\delta p_{S,SB\mathcal{B}}$	80.32	$\delta p_{S,b\nu B}$	2.38
$\delta p_{S,-B\mathcal{L}S}$	70.59	$\delta p_{S,b\nu S}$	1.64

Table 3.3: Experiment 3:  $L^2$ -norm as a measure of energy of all observed terms at cr154. The terms are sorted in descending order of energy. The terms related to single scattering effects are the strongest, followed by the ones related to multiple-scattering. The braces with attached values show the energy of the corresponding summed traces, indicating destructive interference of the individual contributions (see also Figures 3.16).

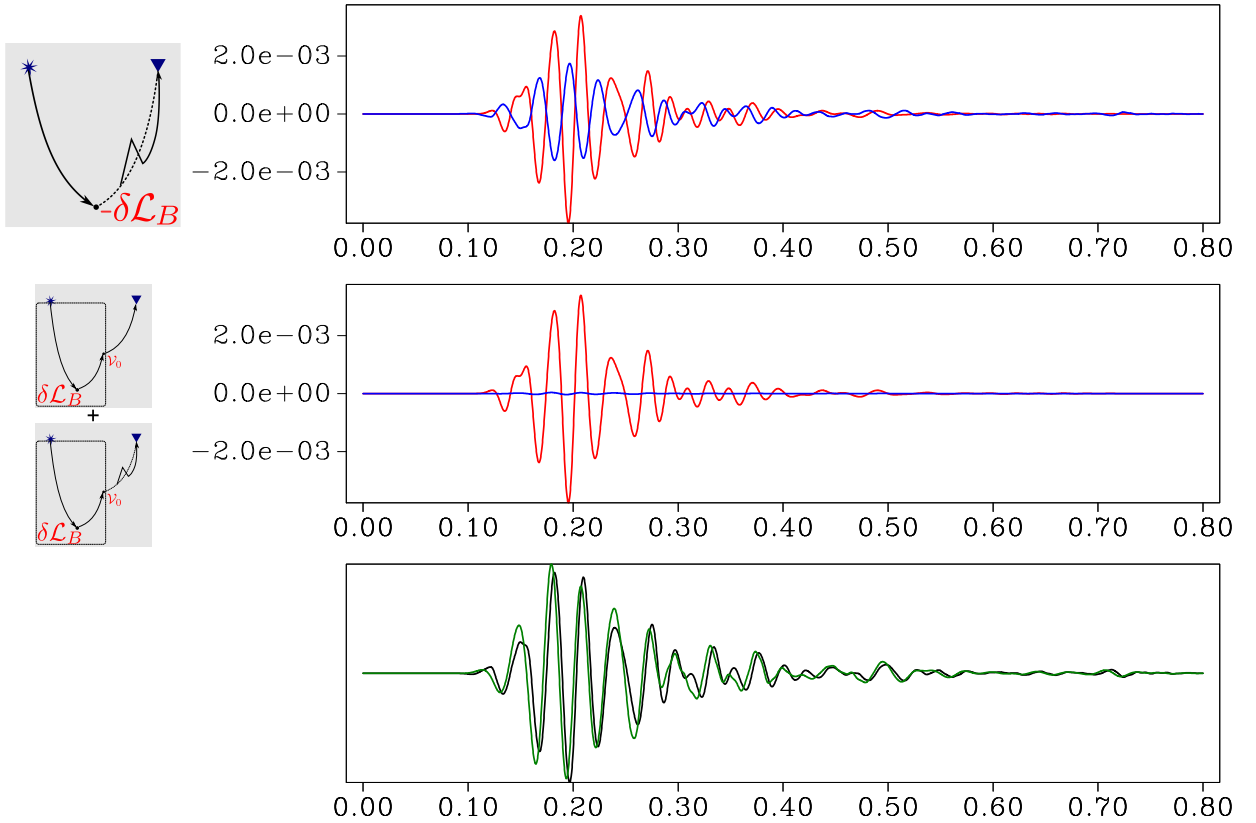


Figure 3.17: EXPERIMENT 3.: First two traces from top to bottom:  $\delta p_{S,-B\mathcal{L}S}$  and  $\delta p_{S,b\nu B} + \delta p_{S,b\nu S}$ , respectively. In these traces,  $\delta p_S^{\text{true}}$  is depicted in red lines for comparison. The last trace show both  $\delta p_{S,-B\mathcal{L}S}$  (black line) and  $\delta p_{S,b\nu B} + \delta p_{S,b\nu S}$  (dark green line) contributions normalised to their respective absolute maximum value.

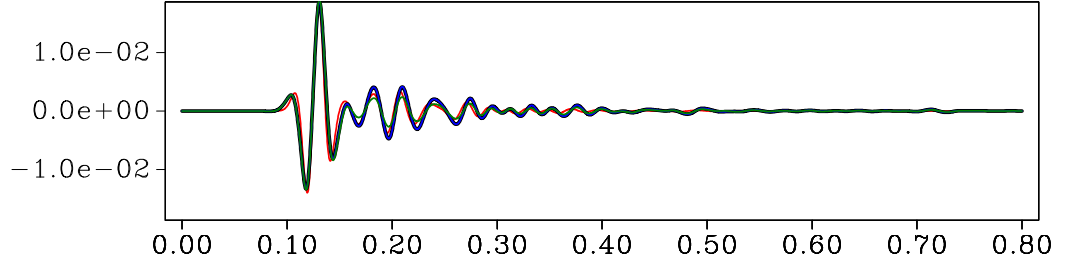


Figure 3.18: EXPERIMENT 3. Comparing  $\delta p^{\text{true}}$  (red line) with evaluated counter-parts at cr154. In black,  $\delta p$  (evaluated with eq. (2.14)); in dark green,  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all terms of eq. (2.33) respectively); and in blue,  $\delta p_B + \delta p_S$  (evaluated with eq. (2.25) and all but the 1<sup>st</sup>, 2<sup>nd</sup> and 7<sup>th</sup> terms of eq. (2.33) respectively). The blue and black lines are indistinguishable.

The last trace of Figure 3.19 demonstrates that the sum of the individual contributions  $\delta p_{S,BBS}$  and  $\delta p_{S,SB B}$  carry practically all the information contained in the sum of all five contributions.

The main conclusion from this experiment is that the multiple-scattering contributions to the perturbed wavefield can carry important information on the perturbation of the background model. Therefore, it might be very helpful to include the singular-part information into the reference model, even for an inversion of the background model only.

### 3.3.4 Experiment 4: Perturbing the scatterers' position of Model II, forward decomposition

Here, in opposition to the previous experiment, we kept the background of Model II unchanged and perturbed the singular part by randomly dislocating the scatterers up to 6  $m$  in both the vertical and horizontal directions. Figure 3.20 shows the normalised values of the medium perturbations, where a value of -1 indicates the strongest reduction in velocity and +1 the strongest increase. In other words, in this experiment we have  $\delta K_B = 0$  and  $\delta K_S \neq 0$ .

According to equation (2.44), this yields a perturbation only of the singular wavefield as

$$\begin{bmatrix} \hat{\delta p}_B \\ \hat{\delta p}_S \end{bmatrix} = \begin{bmatrix} 0 \\ \left( \sum_j \bar{U}_{S,j} \right) \delta K_S \end{bmatrix}, \quad (3.13)$$

with  $j = BSB, BSS, SSB$ , and  $SSS$ . These individual contributions correspond to the cartoons in Figures 2.4j to 2.4m. Source and control receivers for the single-shot modelling are in the same position as in the constant background perturbation experiment (black star and triangles in

### 3.3 Experiment gallery

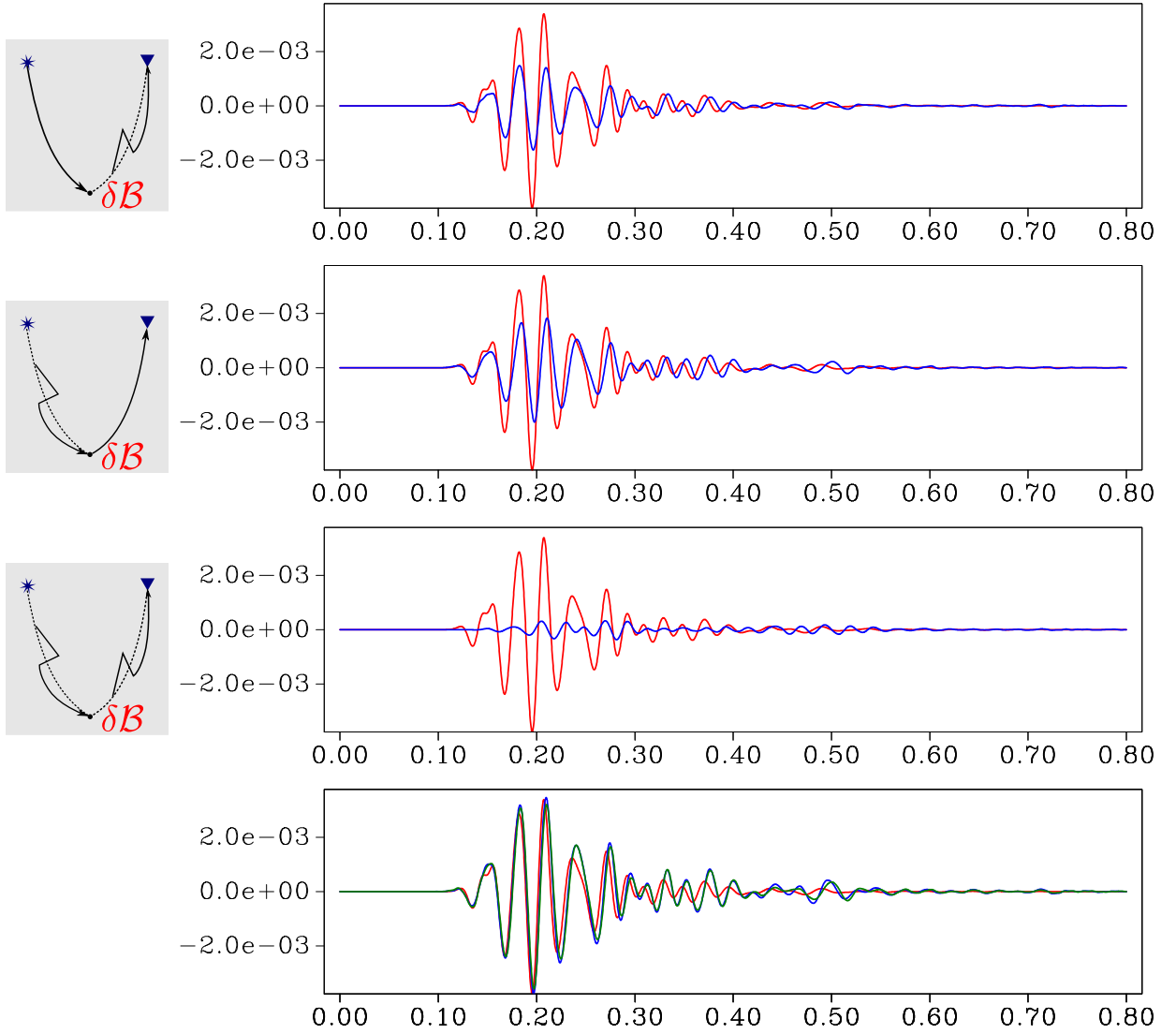


Figure 3.19: EXPERIMENT 3. Multiple-scattering contributions (blue lines), from top to bottom:  $\delta p_{S,BBS}$ ,  $\delta p_{S,SBB}$  and  $\delta p_{S,SBS}$ . These are the terms that carry most of the energy of the singular residual  $\delta p_S$ . In the last trace we can see the summation of the first two traces depicted above (dark green line) and the  $\delta p_S$  evaluated with all five contributions ( $BBS$ ,  $SBB$ ,  $SBS$ ,  $BBS$  and  $-BCLB$ ). In all figures,  $\delta p_S^{\text{true}}$  is depicted in red lines for comparison.

Figure 3.2). In this experiment we analyse the wavefield residuals at cr154.

Figure 3.21 compares the decomposed singular-wavefield residual  $\delta p_S$  (blue line) to its undecomposed counterpart  $\delta p = \delta p_S$  (black line) and to the true residual  $\delta p_S^{\text{true}}$  (red line) calculated by the difference of the wavefields in the perturbed and unperturbed models. Since in Experiment 4



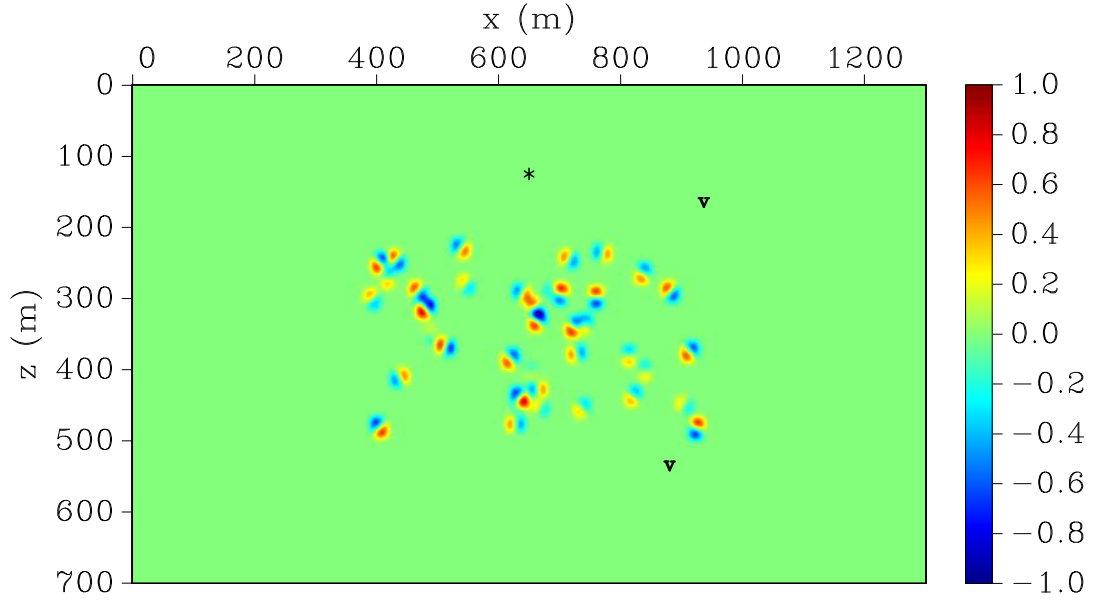


Figure 3.20: EXPERIMENTS 4 AND 5. Only the singular part of the model is perturbed, i.e.,  $\delta K_S \neq 0$ . This is achieved by modifying the vertical and horizontal coordinates of the scatterers randomly within a given range. The cool (blue) spots show the unperturbed positions, while the hot (red) ones tell us the new positions. The intensity of the spots is related to the magnitude of the perturbation. The green background corresponds to no perturbation.

the total model perturbation  $\delta K$  is equal to the singular-part perturbation  $\delta K_S$ , we have  $\delta p_B = 0$  and, therefore, the true total residual  $\delta p^{\text{true}}$  equals the true singular residual  $\delta p_S^{\text{true}}$ . The good match between  $\delta p_S$  and  $\delta p_S^{\text{true}}$  indicates that the Born approximation is valid in this situation, as expected.

In Figure 3.22, we see all four terms that contribute to  $\delta p_S$  at cr154. As we can see, in terms of energy, all of them are of the same order,  $\delta p_{S,SSS}$  being a little weaker than the other three. None of the contribution cancels another one, and their sum yields a good match with the true singular residual (Figure 3.21, blue line). The individual best match is doubtlessly achieved by the single-scattering term  $\delta p_{S,BSB}$  (Figure 3.22, top), but the importance of the other terms to improve the match with the true residual is visible.

### 3.3 Experiment gallery

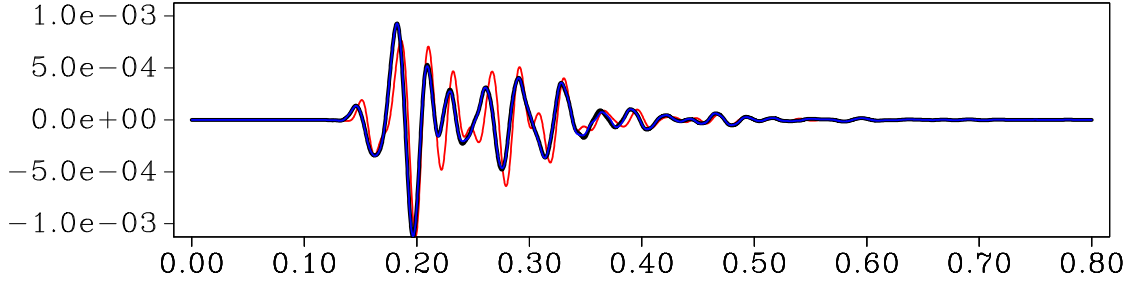


Figure 3.21: EXPERIMENT 4. Traces evaluated at cr154. Total residual,  $\delta p$ , given by equation (2.14), in black, and singular-wavefield residual,  $\delta p_S$ , equation (2.32), in blue, match almost perfectly. Good match between linearised and true residual,  $\delta p_S$  (in blue) and  $\delta p_S^{\text{true}}$  (in red) indicates the good quality of the Born approximation in this case, which is a consequence of the small perturbations in the model.

#### 3.3.5 Experiment 5: Perturbing the scatterers' position of Model II, backprojecting residuals

In this experiment, using the same unperturbed and perturbed models of the previous experiment, we estimate the total model perturbation  $\delta K$  which, as mentioned before, is equal to the singular-part perturbation  $\delta K_S$ . Therefore, we have  $\delta p_B = 0$  and the true total residual  $\delta p^{\text{true}}$  equals the true singular residual  $\delta p_S^{\text{true}}$ . Thus, this is the residual we need to backproject in this experiment. We simulated 44 shots and recorded the wavefields at a receiver array surrounding the scatterers (white stars and triangles in Figure 3.2 respectively). According to eq. (3.7), the bulk-modulus estimates are then given by

$$\begin{bmatrix} \delta K_B^{\text{est}} \\ \delta K_S^{\text{est}} \end{bmatrix} = \begin{bmatrix} \sum_i \bar{U}_{B,i}^\dagger \hat{\delta p}_S^{\text{true}} \\ \sum_j \bar{U}_{S,j}^\dagger \hat{\delta p}_S^{\text{true}} \end{bmatrix}, \quad (3.14)$$

where the summation indices  $i$  and  $j$  take the values explained in connection with equation (3.7).

Figure 3.23 compares  $\delta K_S^{\text{est}}$  to the bulk-modulus perturbation estimate  $\delta K^{\text{est}}$  obtained with the undecomposed approach using the first equation of (2.49). As expected, the summation of the four terms yielding  $\delta K_S^{\text{est}}$  (Figure 3.23, bottom) almost exactly leads to the same result as the total bulk-modulus perturbation estimate,  $\delta K^{\text{est}}$  (Figure 3.23, top). This indicates that the decomposition was successful in the sense that the summation of all components recovers the complete estimate. This

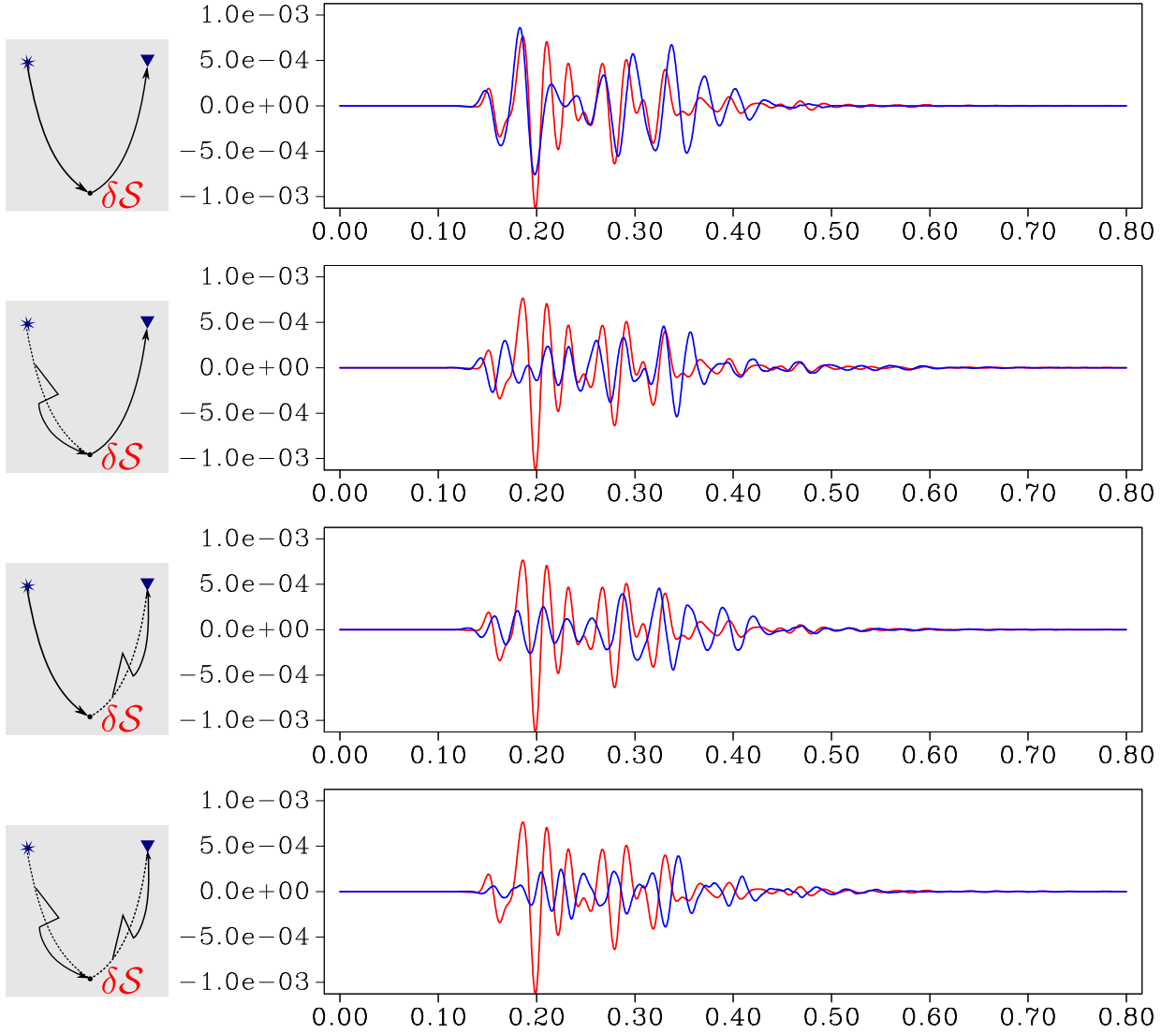


Figure 3.22: EXPERIMENT 4. All four terms that contribute to the total singular residual at cr154. From top to bottom:  $\delta p_{S,BSB}$ ,  $\delta p_{S,SSB}$ ,  $\delta p_{S,BSS}$ ,  $\delta p_{S,SSS}$ . The single-scattering term,  $\delta p_{S,BSB}$ , accounts for almost all of the complete singular residual.

was expected since  $\delta p_B = 0$  which, from eq. (3.11), yields

$$\overline{\delta K}^{\text{est}} = \left( \sum_j \overline{U}_{S,j} \right) \widehat{\delta p}_S = \overline{\delta K}_S^{\text{est}}. \quad (3.15)$$

But when we compare  $\delta K_S^{\text{est}}$  (Figure 3.23, bottom) with the true perturbation  $\delta K_S$  (Figure 3.20)

### 3.3 Experiment gallery

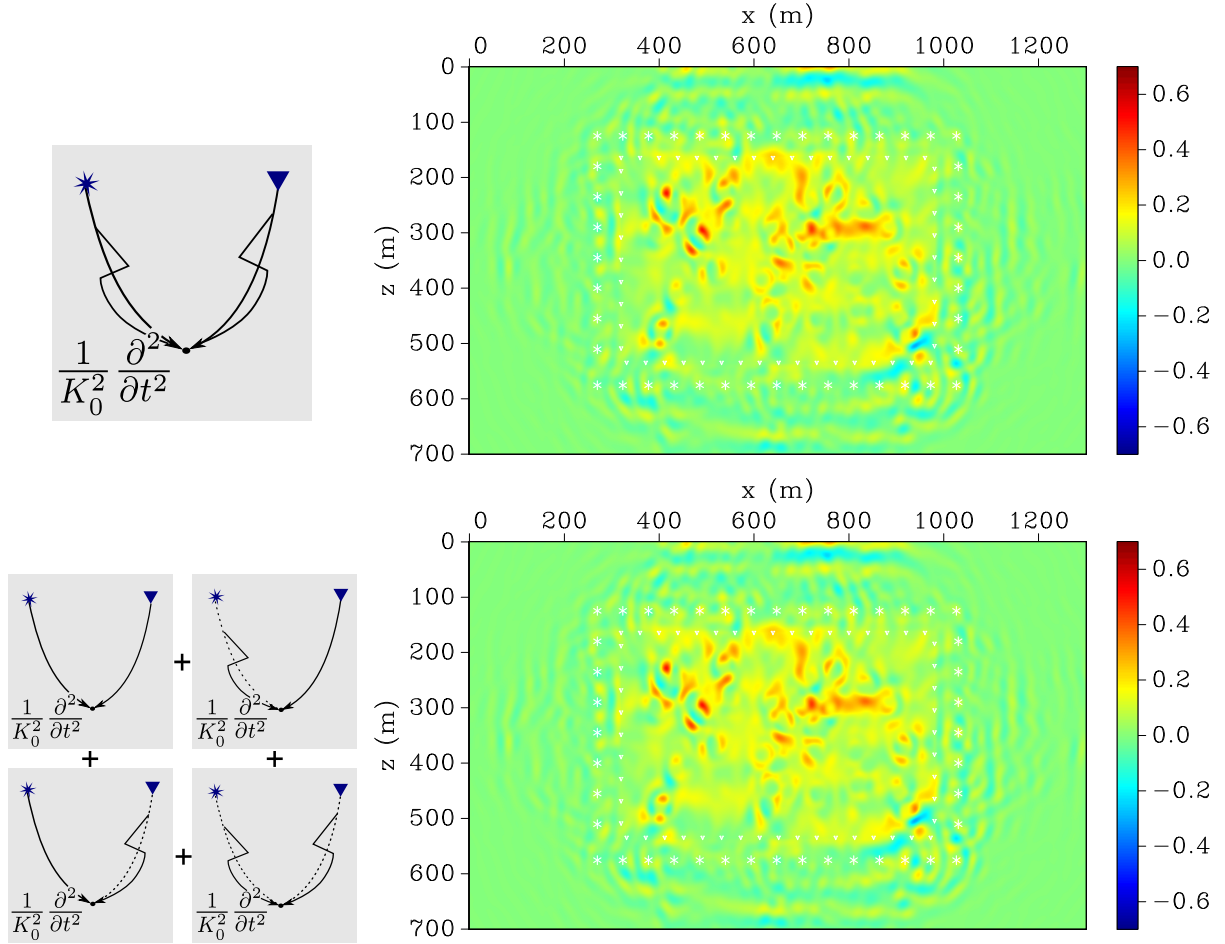


Figure 3.23: EXPERIMENT 5. Backprojecting the total (singular) wavefield residual,  $\delta p_S^{\text{true}} = \delta p_S^{\text{true}}$ . Cartoons indicate the (sub)kernels used. *Top*: Bulk modulus perturbation estimate,  $\delta K_S^{\text{est}}$ , obtained with the first equation of (2.49). *Bottom*: Singular part of bulk modulus perturbation estimate,  $\delta K_S^{\text{est}}$ , obtained with the second equation of (3.14). Colour scale clipped at 0.6 for better comparison.

we see that the estimate is somewhat far from the true perturbation.

To better understand this issue, we must inspect each contribution separately. Figure 3.24 displays all four of them. Comparing  $\delta K_{S,BB}^{\text{est}}$  with  $\delta K_S$  (Figure 3.20), we see that the single-scattering contribution alone gives a better perturbation estimate. In Figure 3.25, top, we can see the estimate  $\delta K_{\text{conv}}^{\text{est}}$  which is traditionally used when performing conjugate gradient FWI. Comparing this estimate with the single-scattering contribution  $\delta K_{S,BB}^{\text{est}}$  (displayed also in Figure 3.25, top, for better comparison) we can observe that the former gives an even better image when compared

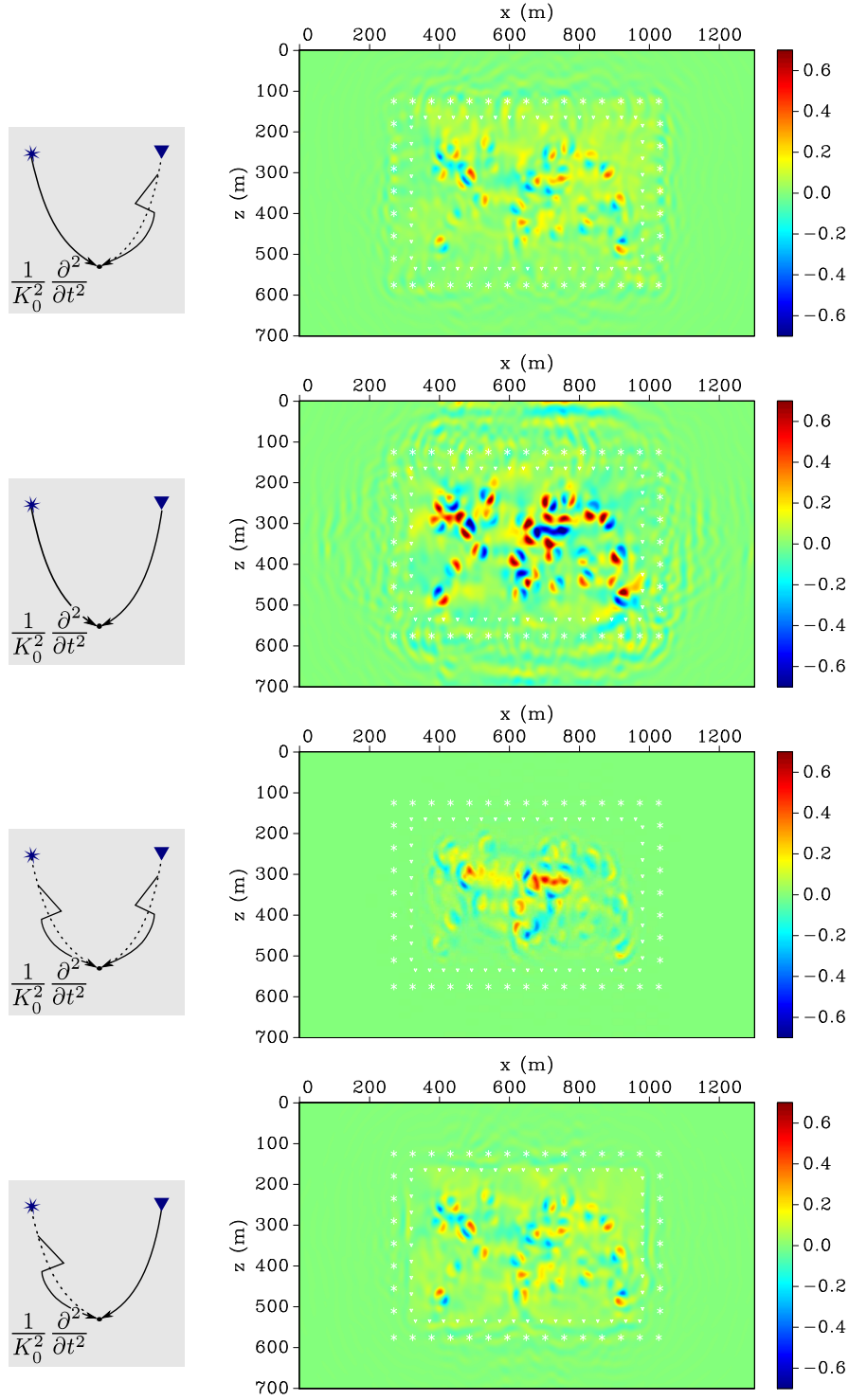


Figure 3.24: EXPERIMENT 5. Each one of the contributions to the singular perturbation estimate. From top to bottom:  $\delta K_{S,BS}^{\text{est}}$ ,  $\delta K_{S,BB}^{\text{est}}$ ,  $\delta K_{S,SS}^{\text{est}}$  and  $\delta K_{S,SB}^{\text{est}}$ .

### 3.3 Experiment gallery

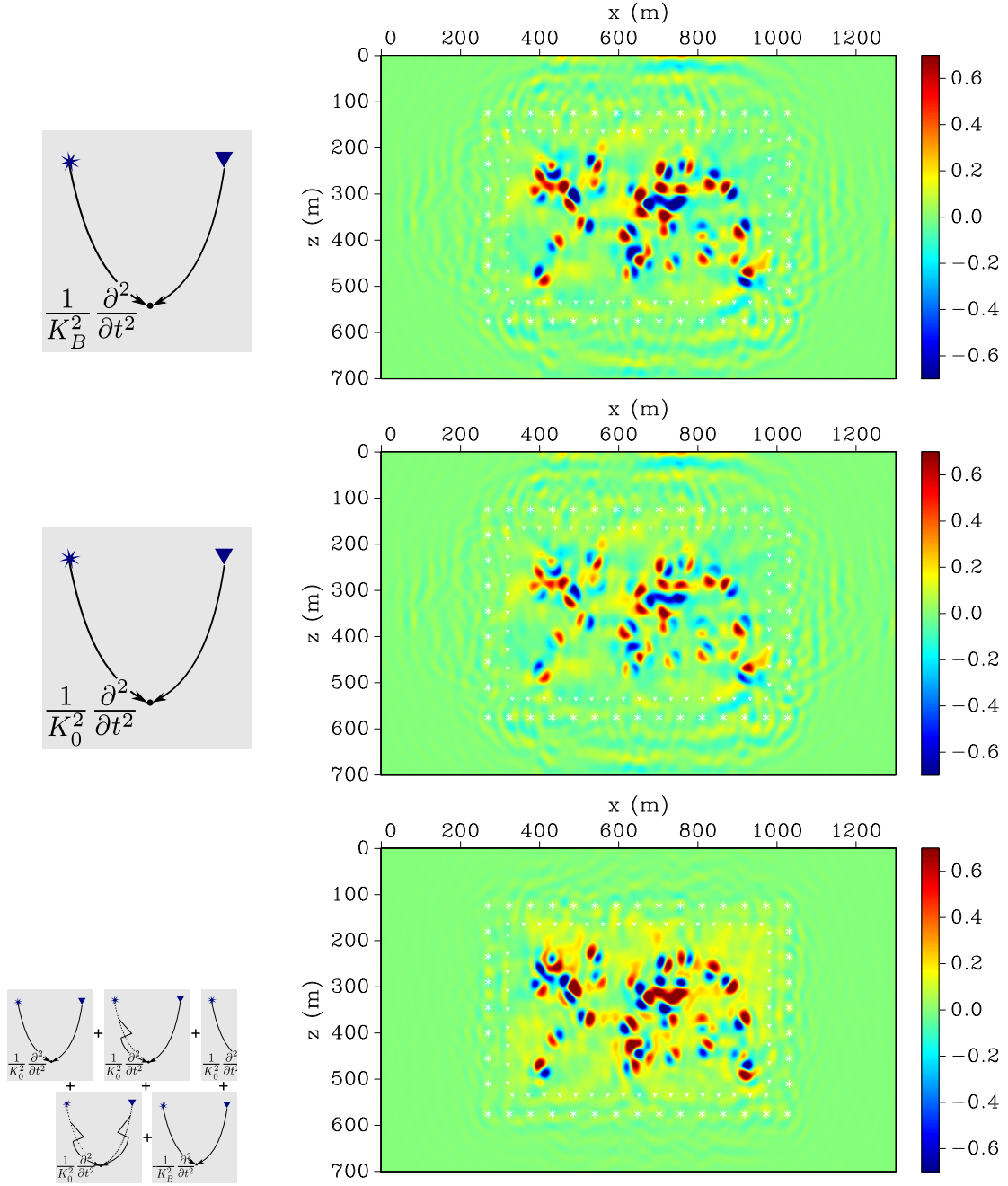


Figure 3.25: EXPERIMENT 5. *Top*: Singular part of Bulk modulus estimate,  $\delta K_{S,BB}^{est}$ , obtained with the single-scattering related kernel. *Center*: Conventional Bulk modulus estimate,  $\delta K_{conv}^{est}$ , obtained with the background-wavefield related kernel. Note that these estimates are very similar to the true perturbation observed in Figure 3.20, the convetional one being a little better. *Bottom*: Background bulk-modulus estimate,  $\delta K_B^{est}$ , obtained with first equation of (3.14) (residual leakage). Colour scale clipped at 0.6 for better comparison.

with the true perturbation  $\delta K_S$  (Figure 3.20). This is not by chance. Many works (Mora, 1987, 1988; Jannane et al., 1989; Pratt, 1999) show that reflection data are able to properly resolve only the high frequency information of the model with the single- scattering term. Results like this one motivate the conventional choice of letting the background medium alone represent the reference medium. Its high quality is expected for this example, because the assumption that the background medium is fully known and all perturbations only affect the singular part, is exactly satisfied.

Finally note that in this experiment, only  $\delta K_S^{\text{est}}$  should be non-zero. However, we see from eq. (3.14) that  $\delta K_B^{\text{est}}$  look quite similar to  $\delta K_S^{\text{est}}$ . Numerical evaluation confirms that  $\delta K_B^{\text{est}}$  is indeed non-zero, too (see Figure 3.25b). This is a consequence of residual information leakage since  $\delta p_S$  is generated by a singular-part perturbation only. Surprisingly, this estimate is way more focussed and clean from artefacts than all the other estimates, although its polarity is reversed.

Residual leakage occurs because the inversion uses the adjoint operation instead of the inverse. In the estimate of  $\delta K_B$ , the dominant contributions are  $\delta K_{B,BB}^{\text{est}}$  (Figure 3.25, top) and  $\delta K_{B,-B\mathcal{L}B}^{\text{est}}$ . They are constructed with background wavefields both in the source- and receiver-sides. These contributions cancel each other except at scatterer positions, because in the former the scale factor of the backscattering mechanism is  $\frac{1}{K_0^2}$  and in the latter  $\frac{1}{K_B^2}$ . Since in our example  $K_0 \geq K_B$ , the estimate values in  $\delta K_{B,-B\mathcal{L}B}^{\text{est}}$  are boosted at the scatterer positions which explains the polarity and apparent focussing of the leakage estimate. What we actually see is basically the modified scattering potential  $\mathcal{W}^K$  (see Appendix D).

### 3.3.6 Experiment 6: Lens-shaped background perturbation of Model III, forward decomposition

For the next two experiments we kept the singular part unperturbed and introduced a spatially limited (lens-shaped), relatively strong background perturbation, as shown in Figure 3.26. Again, we have perturbations in both the background and singular wavefields according to eq. (3.2). Source and control receiver distribution in this particularly experiment are the same used in Section 3.3.3.

We are going to analyse the response at both control receivers. Receiver cr154 records almost exclusively backscattered wavefields, because the Fresnel zone of the direct arrival does not cross the lens-shaped perturbation (we say this control receiver is in the *backscattering domain*). Conversely, receiver cr283 records transmitted and forward scattered energy (it is in the *transmission domain*).



### 3.3 Experiment gallery

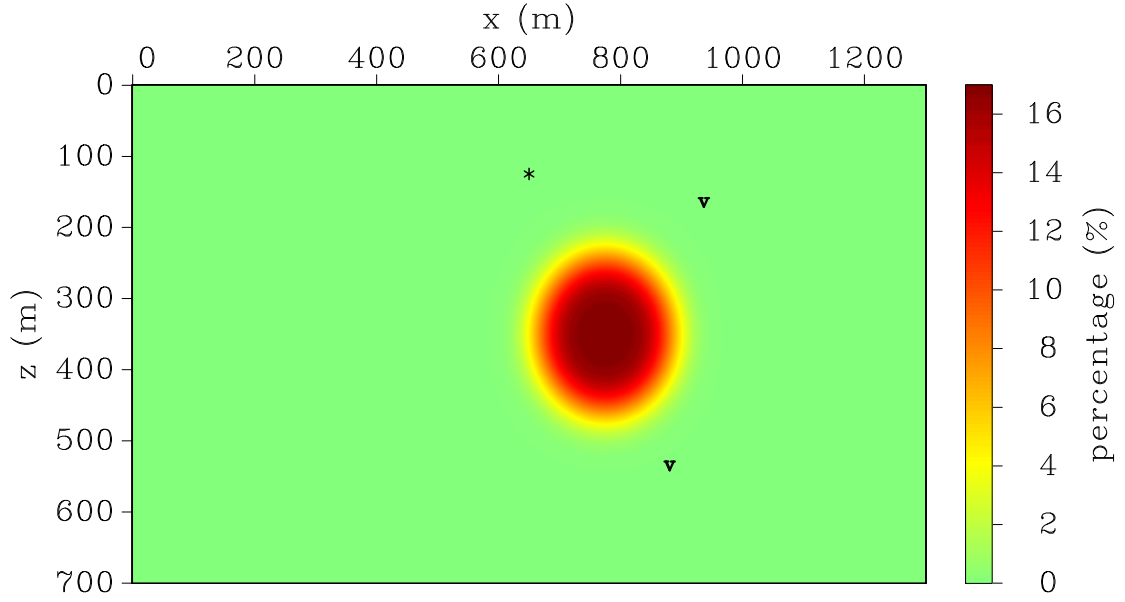


Figure 3.26: EXPERIMENTS 6 AND 7. Perturbation in the bulk-modulus background model,  $\delta K_B$ . The scalebar shows, in percentage, the perturbation magnitude.

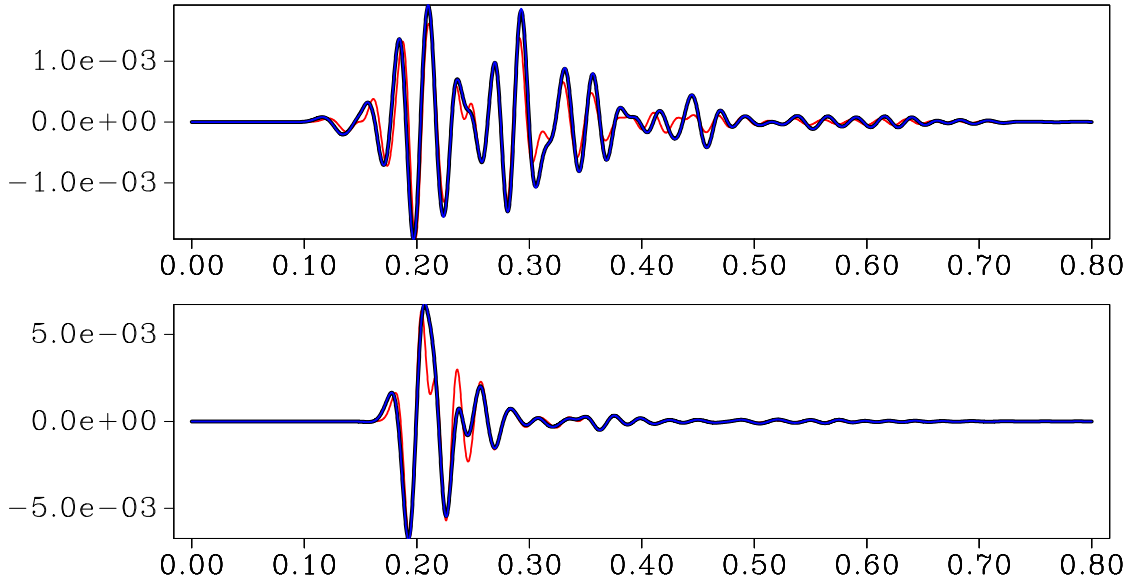


Figure 3.27: EXPERIMENT 6. Comparison between the total wavefield residual,  $\delta p$  (black lines) and sum of background and singular residuals,  $\delta p_B + \delta p_S$  (blue lines). *Top*: Evaluated at control receiver cr154. *Bottom*: Evaluated at control receiver cr283.



cr154		cr283	
Term	$L^2$ -norm ( $\times 10^{-6}$ )	Term	$L^2$ -norm ( $\times 10^{-5}$ )
$\delta p_{S,SBB}$	339.28	$\delta p_{S,SBB}$	147.34
$\delta p_{S,BBS}$	262.27	$\delta p_{S,BBS}$	175.72
$\delta p_{S,SBS}$	214.88	$\delta p_{S,SBS}$	100.91
$\delta p_{S,BBB}$	55.21	$\delta p_{S,-BCLB}$	330.81
$\delta p_{S,-BCLB}$	26.48	$\delta p_{S,BBB}$	294.00
} 53.21		} 38.70	

Table 3.4: Experiment 6:  $L^2$ -norm as a measure of energy at cr154 and cr283. The braces indicate summation of the related traces, the value attached to them being the energy of the resulting trace.

Figure 3.27 compares total linearised wavefield residual (black lines) with the sum of all non-zero contributions of its decomposed counterpart  $\delta p_B + \delta p_S$  (blue lines) and the true residual  $\delta p^{\text{true}}$  (red lines) at control receivers cr154 and cr283. We see, at a first look, that at both control receivers the summation  $\delta p_B + \delta p_S$  and the evaluated residual  $\delta p$  match nicely; and both are a good approximation to the true residual. As expected at a receiver in the backscattering domain, the amplitudes of the first arrivals at cr154 are of the same order of magnitude as coda. On the other hand, at the transmission-domain receiver cr283 the first arrival is one order of magnitude greater then the coda.

To see if Experiment 6 confirms the results of Experiment 3, we compiled Table 3.4 with the wavefield energies for all contributions at both control receivers. As before, the negative interference of the single-scattering terms reduces their combined importance, leaving  $\delta p_{S,SBB}$ ,  $\delta p_{S,BBS}$  and  $\delta p_{S,SBS}$  as the most important contributions. To be noted is the different importance of the single-scattering terms  $\delta p_{S,BBB}$  and  $\delta p_{S,-BCLB}$ :

- At cr154, where only scattering information due to perturbation reaches the receiver, they carry little energy. Moreover, their sum yields a value very close to the the energy carried by  $\delta p_{S,BBB}$  alone. This means that the small energy carried by th contributions is related to the background residual; most of it is related to the interference term  $\delta p_{S,BWB}$  (see also Figure 3.28).
- At cr283, where transmitted energy is recorded, these are the most energetic terms. However, most of the energy, which is related to the background residual, cancels when these terms are summed, leaving only the part related to the interference term which is of the same order of magnitude as at cr154 (see also Figure 3.29).

### 3.3 Experiment gallery

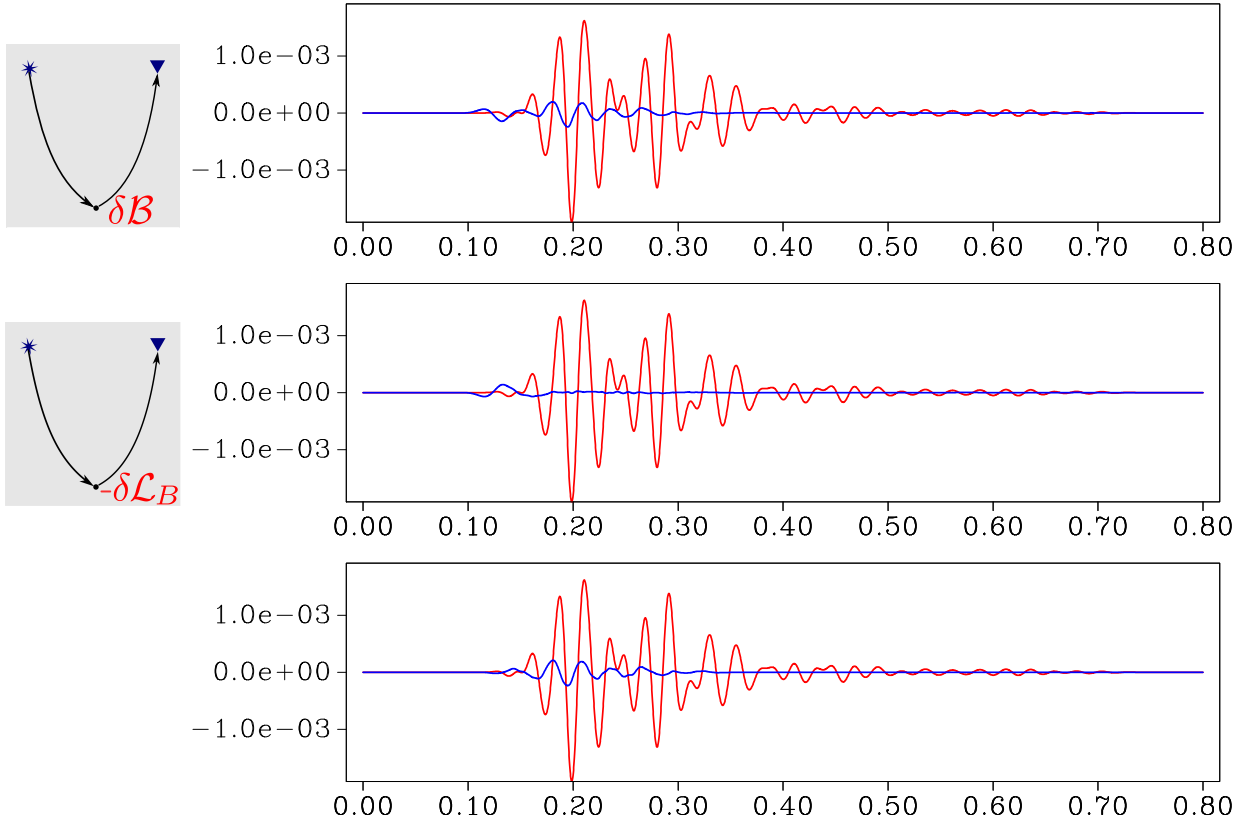


Figure 3.28: EXPERIMENT 6. Single-scattering terms evaluated at control receiver cr154 (blue lines):  $\delta p_{S,BBB}$  (top),  $\delta p_{S,-B L B}$  (centre), and their sum  $\delta p_{S,B W B}$ , compared to the true singular-wavefield residual  $\delta p_S^{\text{true}}$  (red lines). As expected, these terms are very small because the Fresnel zone of the direct wavefield does not intersect with the lens-shaped perturbed area. Compare with Figure 3.29.

In both cases, however, the combined term  $\delta p_{S,B W B}$  is a order of magnitude smaller than the most important terms  $\delta p_{S,SBB}$  and  $\delta p_{S,BBS}$ .

Figures 3.30 and 3.31 compare the most important multiple-scattering contributions  $\delta p_{S,SBB}$ ,  $\delta p_{S,BBS}$  and  $\delta p_{S,SBS}$  (blue lines) individually to  $\delta p_S^{\text{true}}$  (red lines). We see that different parts of the trace are matched by different contributions. However, it is not possible to infer a general rule so as to which contribution will appear at which part of the trace. For instance, at cr154, the first part of the trace is well matched by  $\delta p_{S,SBB}$ ; the central part by  $\delta p_{S,BBS}$ ; and the coda by  $\delta p_{S,SBS}$  (see Figure 3.30). At cr283, different contributions match different parts (see Figure 3.31). The best guess is that the distribution is determined by the general disposition of the scatterers with

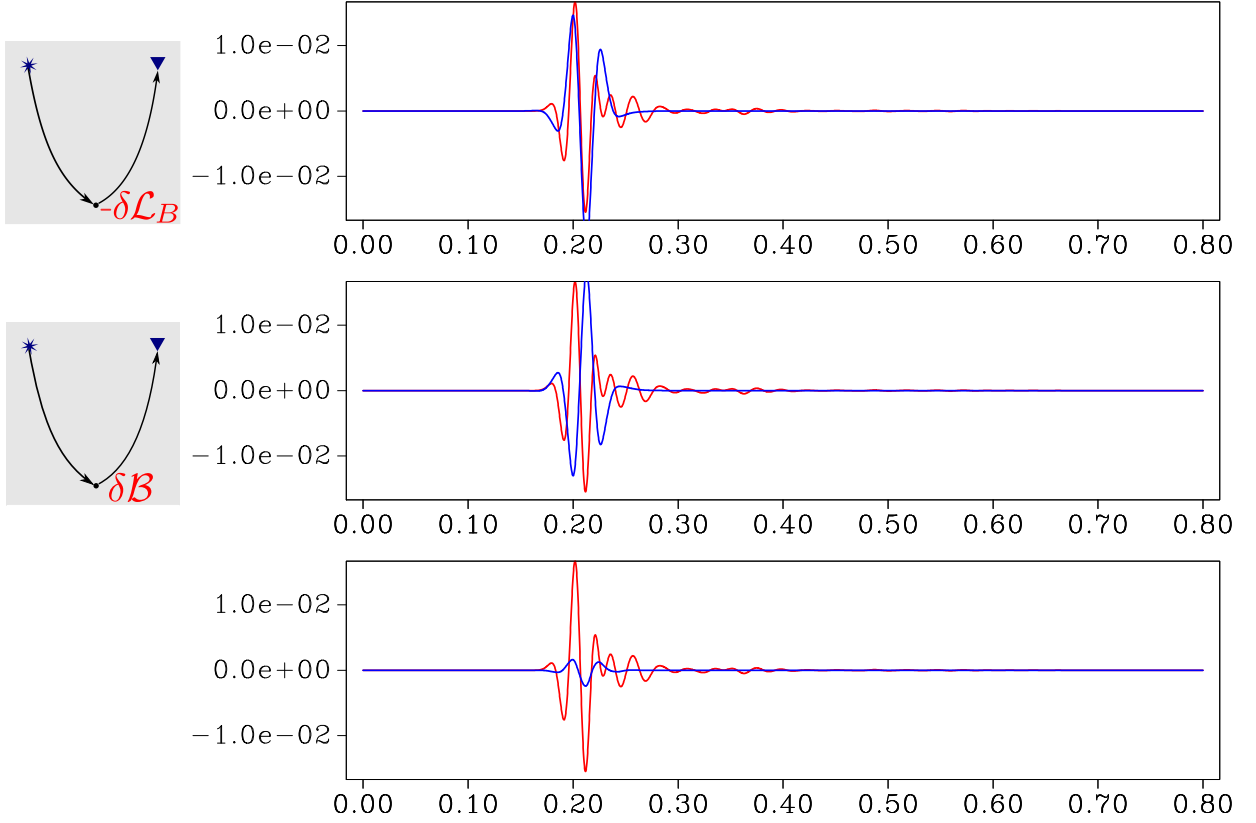


Figure 3.29: EXPERIMENT 6. Single-scattering terms evaluated at control receiver cr283 (blue lines):  $\delta p_{S,BBB}$  (top),  $\delta p_{S,-B\mathcal{L}B}$  (centre), and their sum  $\delta p_{S,BWB}$ , compared to the true singular-wavefield residual  $\delta p_S^{\text{true}}$  (red lines). Each of these terms carries much energy, but their sum contributes very little to the singular-wavefield residual. Compare with Figure 3.28.

respect to the sources and receivers. Due to its strong multiple-scattering nature, the contribution  $\delta p_{S,BBS}$  tends to determine the coda of the trace, while  $\delta p_{S,SBB}$  or  $\delta p_{S,BBS}$  usually define earlier and stronger events.

### 3.3.7 Experiment 7: Lens-shaped background perturbation of Model III, backprojecting residuals

Here, we backproject the true background and singular residuals,  $\delta p_B^{\text{true}}$  and  $\delta p_S^{\text{true}}$  generated in the above forward decomposition experiment with a lens-shaped background perturbation only, i.e.,  $\delta K = \delta K_B$ . As we will see, this perturbation yields the most significant result of our set of

### 3.3 Experiment gallery

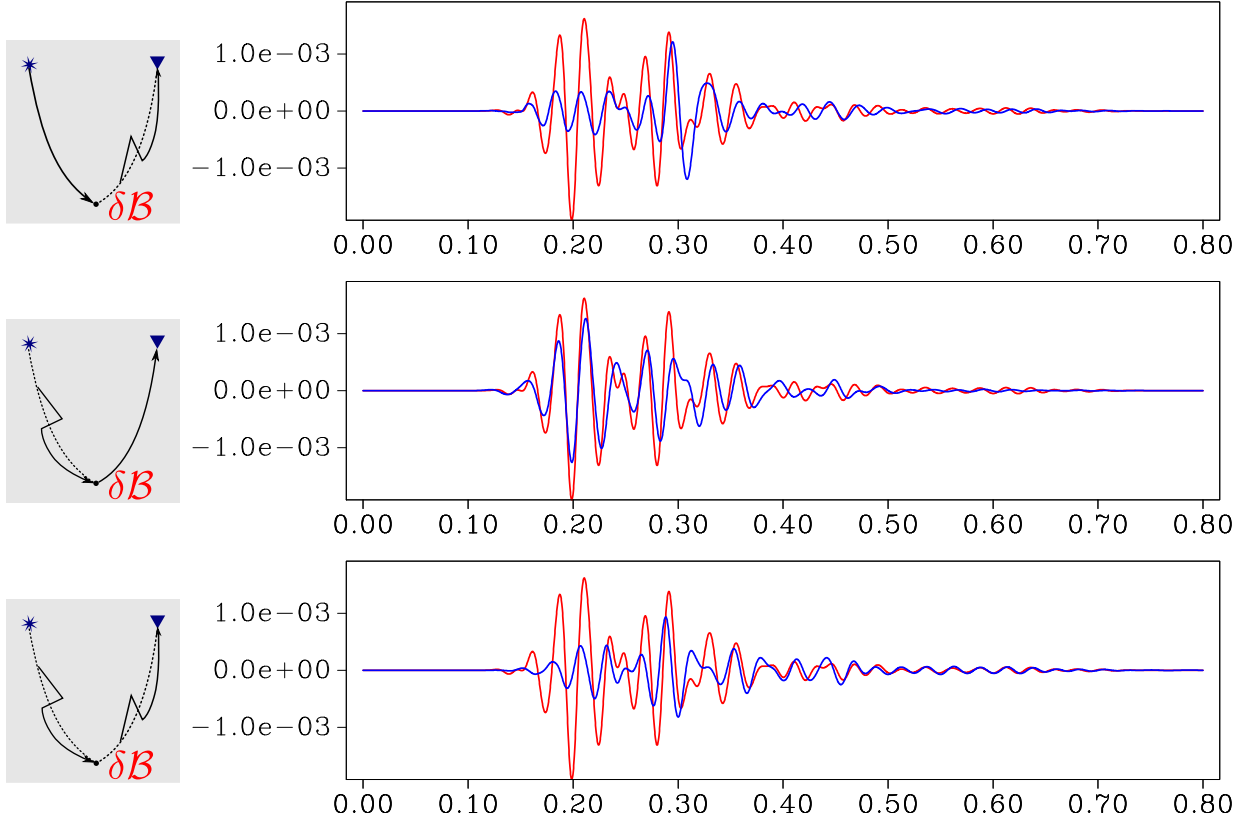


Figure 3.30: EXPERIMENT 6. Terms with the most important contributions to the singular residual evaluated at cr154. *From top to bottom:*  $\delta p_{S,BBS}$ ,  $\delta p_{S,SBB}$  and  $\delta p_{S,SBS}$ .

experiments, namely that meaningful information on the background perturbation can be extracted from the singular-wavefield residual. The perturbation estimates are calculated with the help of equation (3.7).

Figure 3.32 displays, from top to bottom: the total perturbation estimate,  $\delta K^{\text{est}}$  of equation (2.49); the two contributions,  $\delta K_{B,B}^{\text{est}}$  and  $\delta K_{B,S}^{\text{est}}$  of equation (3.7); and finally their sum. The first thing to be noted is that the sum  $\delta K_{B,B}^{\text{est}} + \delta K_{B,S}^{\text{est}}$  does not resembles the total perturbation estimate  $\delta K^{\text{est}}$ . This means that the cross-term contributions  $\delta K_{j(B)}^{\text{est}}$  (first sum of the right hand-side of eq. (3.11)) are not negligible. In Figure 3.33 we can see the sum of these cross-terms and how, when taken into account, they lead to the exact total perturbation estimate  $\delta K^{\text{est}}$ .

Secondly, we see in Figure 3.32 that both  $\delta K_{B,B}^{\text{est}}$  and  $\delta K_{B,S}^{\text{est}}$  fairly account for the true perturbation  $\delta K_B$ . Doubtless, the estimate  $\delta K_{B,B}^{\text{est}}$  from the background residual (second estimate of Figure 3.32) is the one that best represents the true medium perturbation. The inversion done in

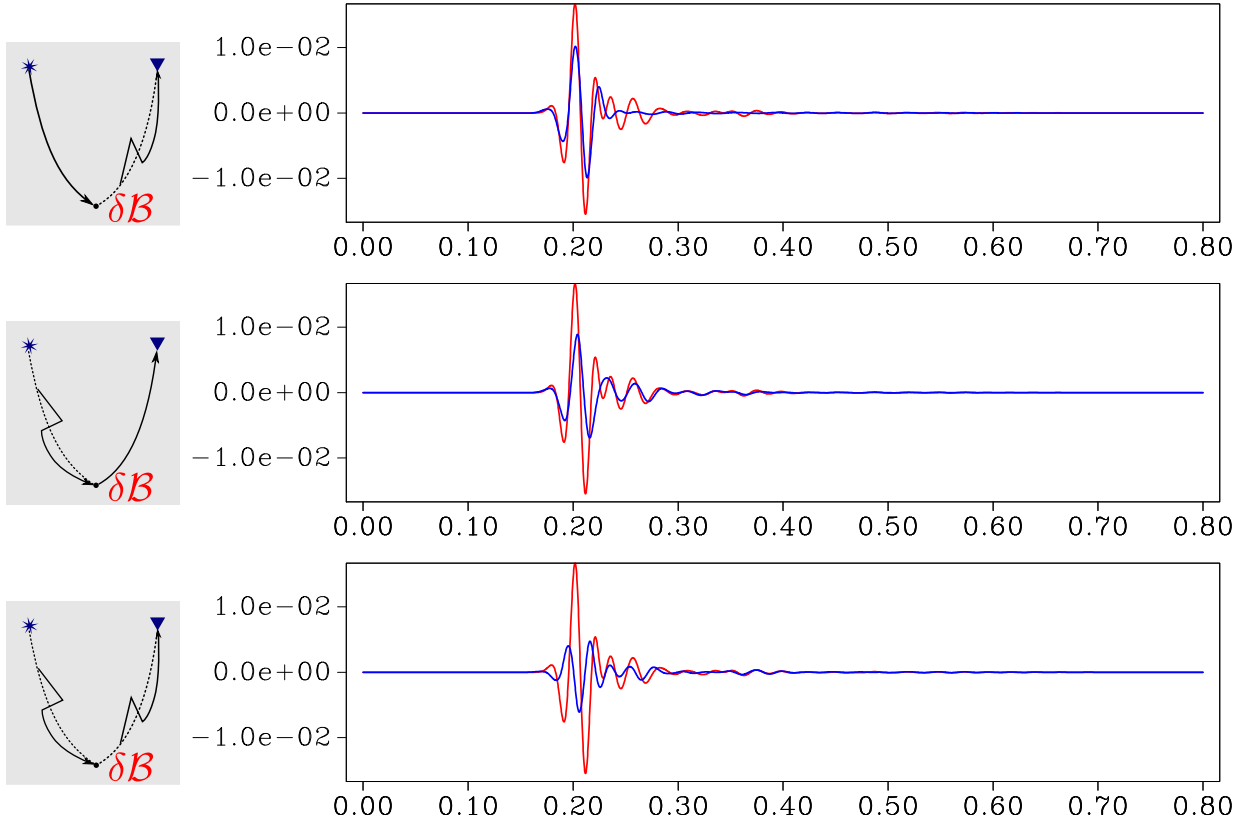


Figure 3.31: EXPERIMENT 6. Terms with the most important contributions to the singular residual evaluated at cr283. *From top to bottom:*  $\delta p_{S,BBS}$ ,  $\delta p_{S,SBB}$  and  $\delta p_{S,SBS}$ .

this case is basically a transmission waveform tomography, which is known to recover very well the low-frequency information, specially when the target region is enclosed by both, sources and receivers (see, e.g., Pratt and Shipp, 1999; Brenders and Pratt, 2007).

But the really important feature here is that a reasonable background perturbation estimate  $\delta K_{B,S}^{\text{est}}$  (Figure 3.32, bottom), even better than the total perturbation estimate  $\delta K^{\text{est}}$ , can be obtained from singular-wavefield residuals. When transmission tomography fails (surface acquisition, for example) the background estimate obtained with the singular residual stands as a possible option (see also Experiment 8, Section 3.3.8).

The most significant contributions to  $\delta K_{B,S}^{\text{est}}$  are  $\delta K_{B,BS}^{\text{est}}$  and  $\delta K_{B,SB}^{\text{est}}$ , displayed in Figure 3.34. Note that, except for the boundary regions close to sources and receivers, the terms  $\delta K_{B,BS}^{\text{est}}$  and  $\delta K_{B,SB}^{\text{est}}$  give practically the same estimate. None of the other contributions alone provides an image of similar quality.

### 3.3 Experiment gallery

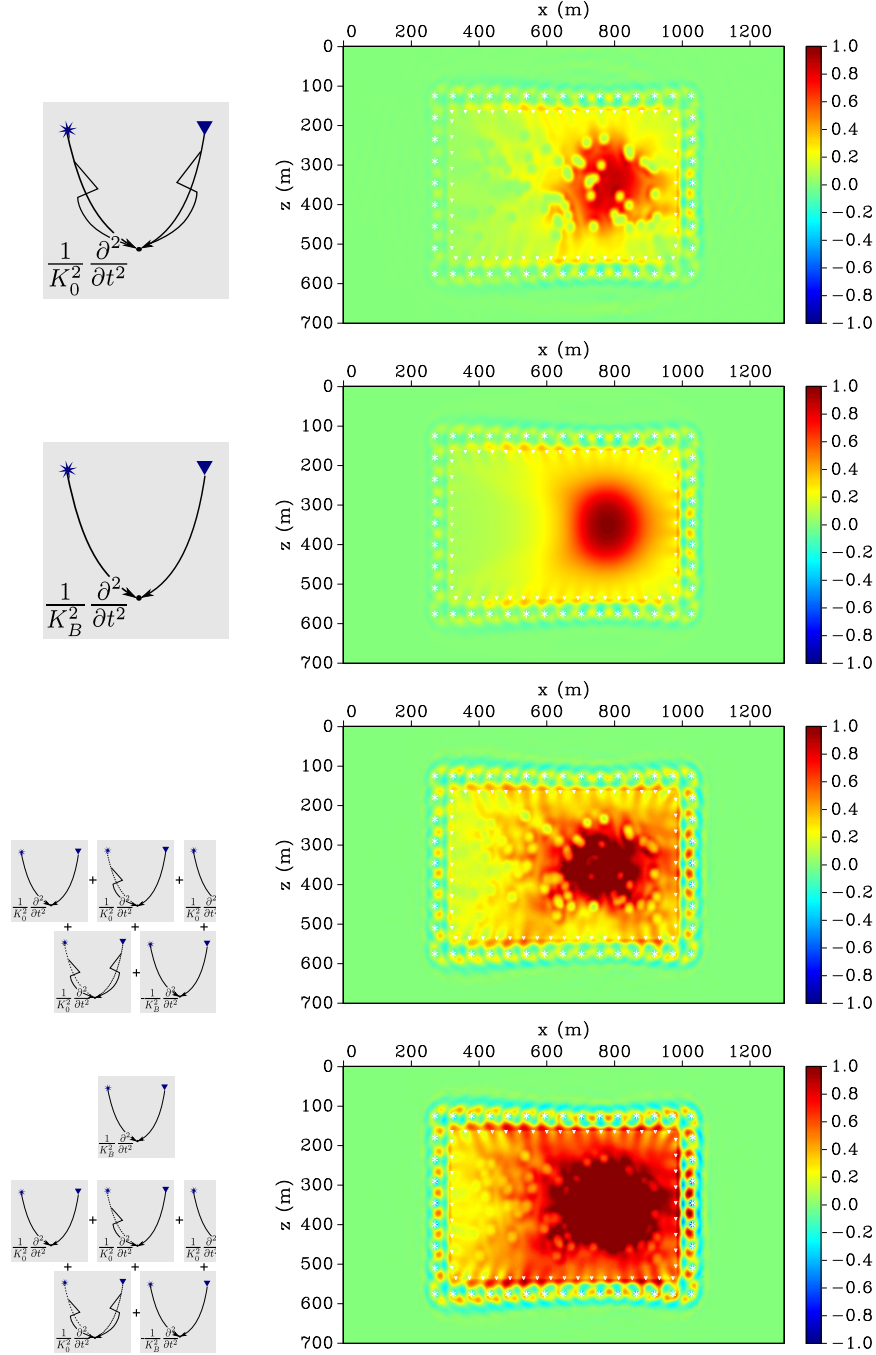


Figure 3.32: EXPERIMENT 7. Backprojecting residuals. *From top to bottom:* Bulk modulus perturbation estimate,  $\delta K^{est}$ , from the total wavefield residual,  $\delta p^{true}$  (equation (2.49)); Background bulk-modulus perturbation estimate,  $\delta K_{B,B}^{est}$ , from background residual,  $\delta p_B^{true}$  (first contribution in first equation of (3.7)); Background bulk-modulus perturbation estimate,  $\delta K_{B,S}^{est}$ , from singular residual,  $\delta p_S^{true}$  (second contribution in first equation of (3.7)); and the sum  $\delta K_{B,B}^{est} + \delta K_{B,S}^{est}$ .

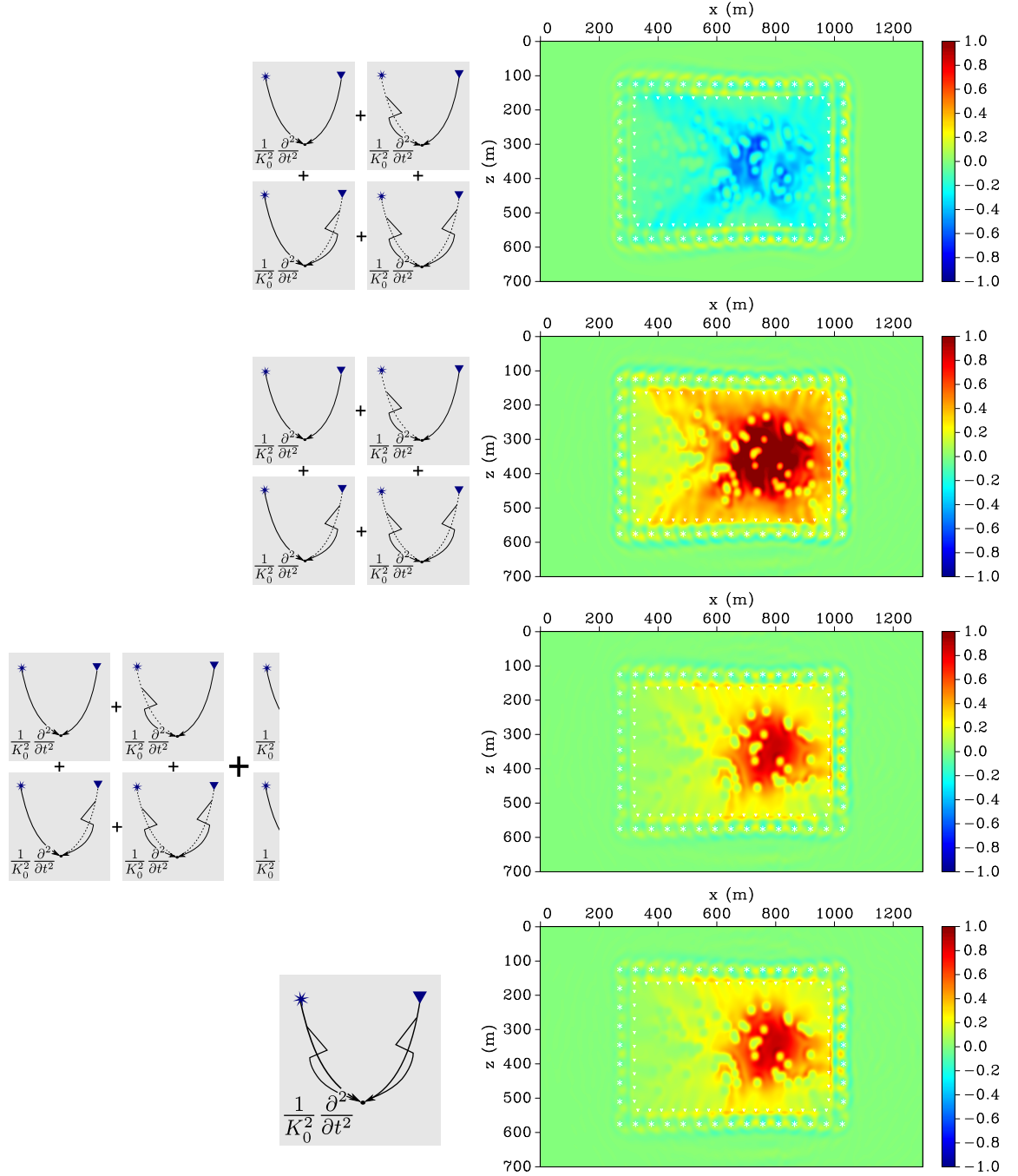


Figure 3.33: EXPERIMENT 7. Backprojecting residuals. Estimates related to eq. (3.11). *From top to bottom:*  $\sum_j \bar{\delta K}_{j(B)}^{\text{est}} = \left( \sum_j \bar{U}_{S,j}^\dagger \right) \hat{\delta p}_B$ ,  $\sum_j \bar{\delta K}_{B,j}^{\text{est}} = \left( \sum_j \bar{U}_{S,j}^\dagger \right) \hat{\delta p}_S$  and their sum. The last is the total perturbation estimate  $\delta K^{\text{est}}$ , put here for comparison.

### 3.3 Experiment gallery

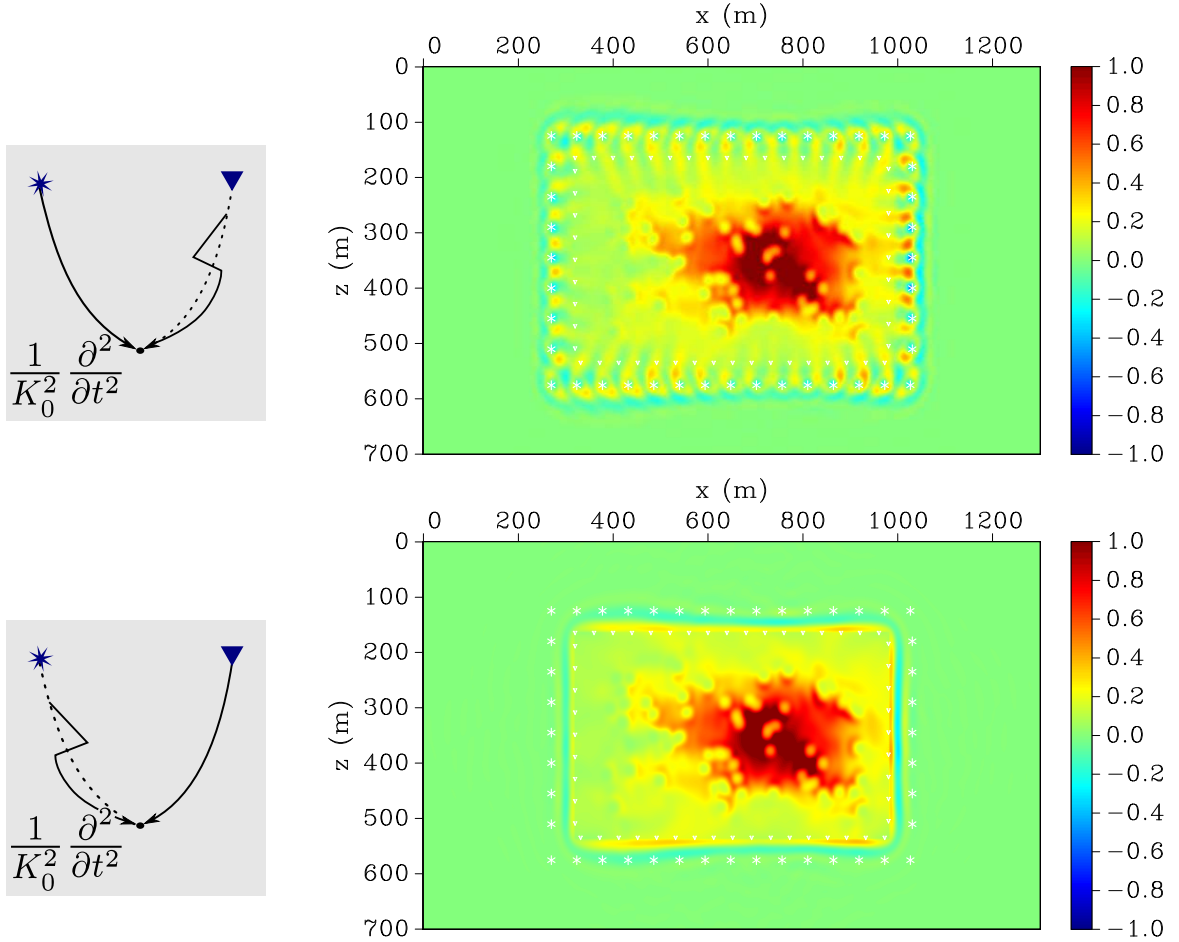


Figure 3.34: EXPERIMENT 7. Backprojecting singular-wavefield residual  $\delta p_S^{\text{true}}$ . *Top*: Subkernel  $\bar{U}_{B,BS}$  used for backprojection. *Bottom*: Subkernel  $\bar{U}_{B,SB}$  used for backprojection. Note the very strong similarity between both estimates apart from the region delimited by the receiver and source lines. It suggests a redundancy that might be explored.

The congruence of the two estimates in Figure 3.34 is a consequence of both the sources and receivers completely enclosing the target region. For a single source-receiver pair, these two contributions are actually complementary (see Figure 3.35). The source where the direct wavefield originates from is marked with a star. The receiver where the residual is observed and backpropagated from is cr283, marked by the triangle in the bottom right corner. In Figure 3.35, top, subkernel  $\bar{U}_{B,BS}$  is used to backproject residuals and the singularities can be seen as acting as receivers, i.e., the residuals seem to be backpropagated from there. On the other hand, in Figure 3.35, bottom, subkernel  $\bar{U}_{B,SB}$  is used. There, the singularities act as sources, i.e., the wavefields seem



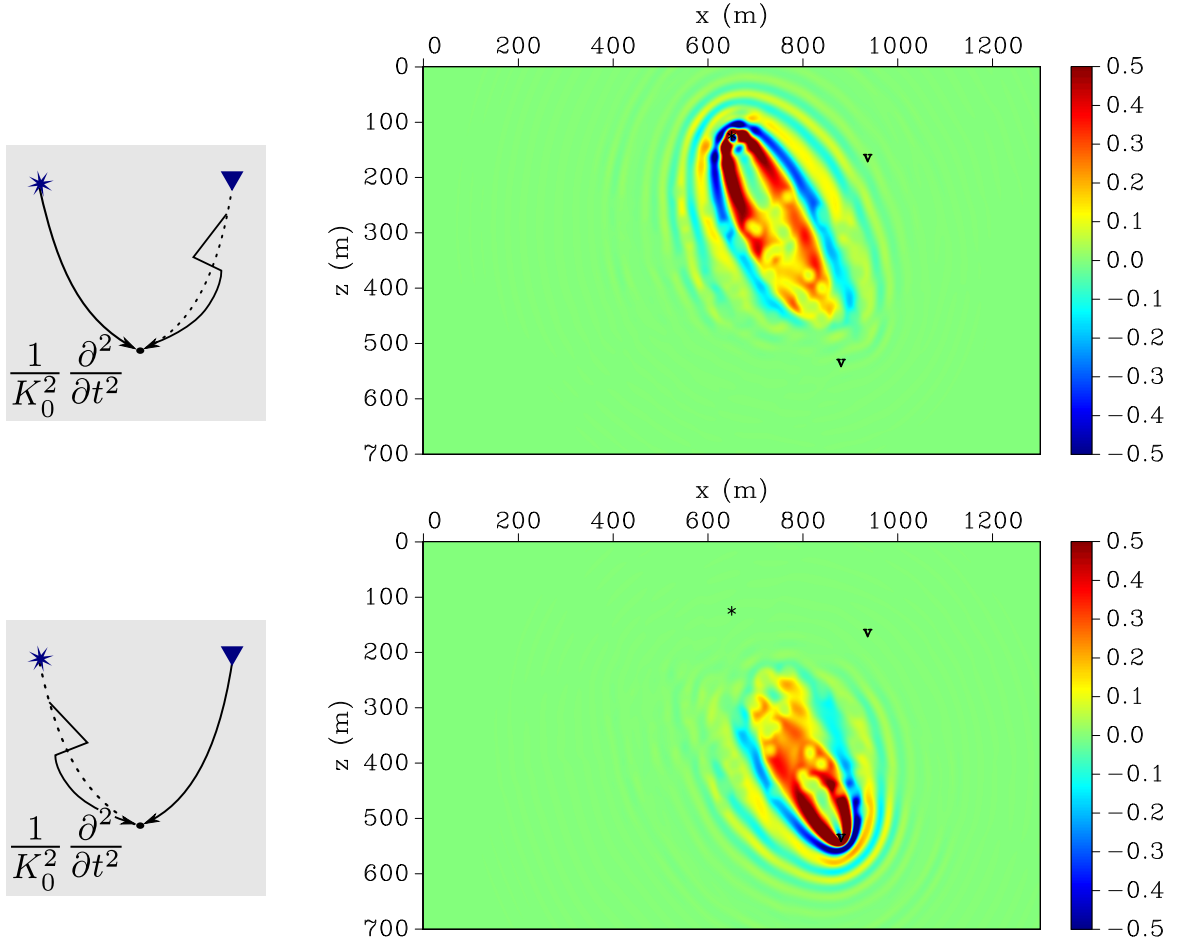


Figure 3.35: EXPERIMENT 7. Backprojecting singular-wavefield residual  $\delta p_S^{\text{true}}$  recorded at cr283 from the source at 1000 m (black star). *Top*: Subkernel  $\bar{U}_{B,BS}$  used for backprojection. The singularities act as receivers. *Bottom*: Subkernel  $\bar{U}_{B,SB}$  used for backprojection. Here, the singularities act as sources. This gives us further insight into the congruence of the parts of Figure 3.34.

to originate from there. Full coverage of sources and receivers around the area of interest and the proximity among them complete the congruence of the estimates shown in Figure 3.34. This analysis indicates that surrounding the target area partly with sources and partly with receivers might be sufficient to recover the perturbations.

### 3.3 Experiment gallery

---

#### 3.3.8 Experiment 8: Lens-shaped background perturbation of Model IV, backprojecting residuals

In this last example we restrict our analysis to the backprojection problem, for the results of the forward decomposition problem resemble the ones obtained for cr154 as reported in the forward decomposition experiment of the Section 3.3.6 (full acquisition geometry). In this experiment, we restrict the acquisition geometry to investigate the role it plays in the inverse problem.

The next experiment uses the same perturbation as Experiments 6 and 7, i.e., lens-shaped background perturbation (see Figure 3.26) and the unperturbed Model IV.

*Experiment 8: backprojecting Residuals.* Figure 3.36 shows the backprojection results together with the acquisition geometry. Source and receivers are marked as white stars and triangles, respectively. From top to bottom, the parts show the undecomposed result  $\delta K^{\text{est}}$  from backprojecting the total residual in the reference model, the estimate  $\delta K_{B,B}^{\text{est}}$  obtained using only the background-wavefield residual, and the estimate  $\delta K_{B,S}^{\text{est}}$  using only the singular-wavefield residual.

The restriction of the acquisition geometry strongly affects the estimate from the total residual,  $\delta K^{\text{est}}$  (Figure 3.36, top), which is not as good as the one in Figure 3.32. But it is still better than the estimate  $\delta K_{B,B}^{\text{est}}$  from the background residual (Figure 3.36, centre) which fails completely. The reason is that this estimate is based on transmission waveform tomography, where only residuals of wavefields that originate at one far side and were recorded at the other far side contribute to the estimate. In other words, in this acquisition geometry, there is not enough far offset for the background residual to yield a good estimate. This is frequently the case in conventional practical inversions that operate with a smooth reference model only. Note, however, that estimate  $\delta K_{B,S}^{\text{est}}$  (Figure 3.36, bottom) from the singular residual provides much better quality even from this geometry with rather short offsets when compared to  $\delta K^{\text{est}}$ . The difference between them is not explained by the cross-terms  $\delta K_{j(B)}^{\text{est}}$  as seen in the previous experiment. Actually, both  $\sum_j \delta K_{j(B)}^{\text{est}} = \left( \sum_j \bar{U}_{S,j}^\dagger \right) \hat{p}_B$  (Figure 3.37) and  $\delta K_{B,B}^{\text{est}}$  (Figure 3.36, centre) contribute almost nothing to  $\delta K^{\text{est}}$  and  $\delta K_B^{\text{est}}$  respectively (see eqs. (3.11) and (3.12)). The contribution that make them differ is  $\delta K_{B,-B\mathcal{L}B}^{\text{est}}$  (see Figure 3.38).

In the same way as in the last experiment, we inspect the contributions  $\delta K_{B,BS}^{\text{est}}$  and  $\delta K_{B,SB}^{\text{est}}$  (Figure 3.39). Although at first sight, they are rather similar to each other, there are some small but visible differences. They do not resemble each other as closely as the corresponding results in Figure 3.34 for Experiment 2, because the spatial coverage of the sources differs slightly from that

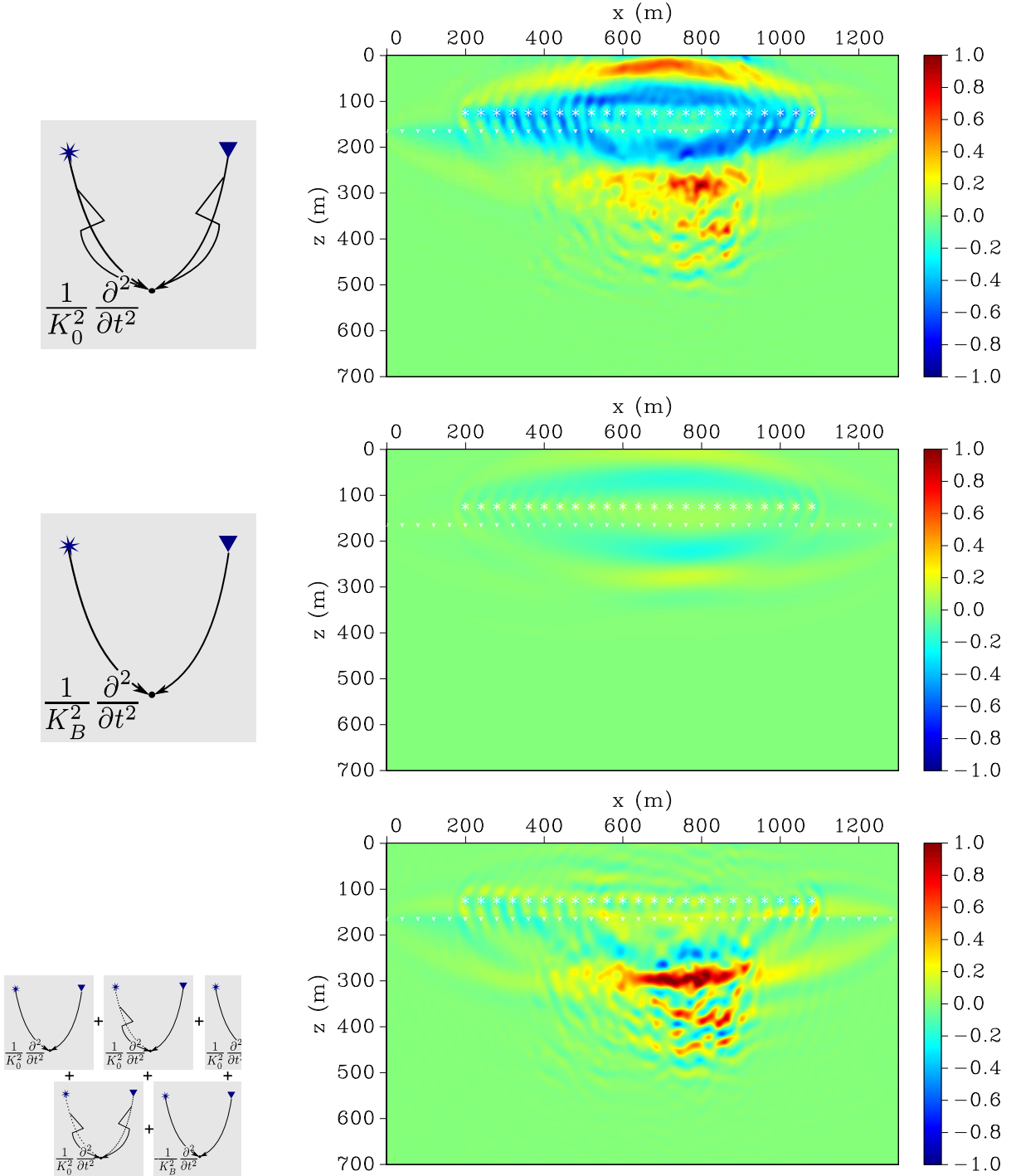


Figure 3.36: EXPERIMENT 8. Backprojecting residuals from “surface” receivers. *Top*: Bulk-modulus perturbation estimate  $\delta K^{\text{est}}$  from total wavefield residual  $\delta p^{\text{true}}$ , *Center*: Bulk-modulus perturbation estimate  $\delta K_{B,B}^{\text{est}}$  from background residual  $\delta p_B^{\text{true}}$ , *Bottom*: Bulk-modulus perturbation estimate  $\delta K_{B,S}^{\text{est}}$  from singular residual  $\delta p_S^{\text{true}}$ .

### 3.3 Experiment gallery

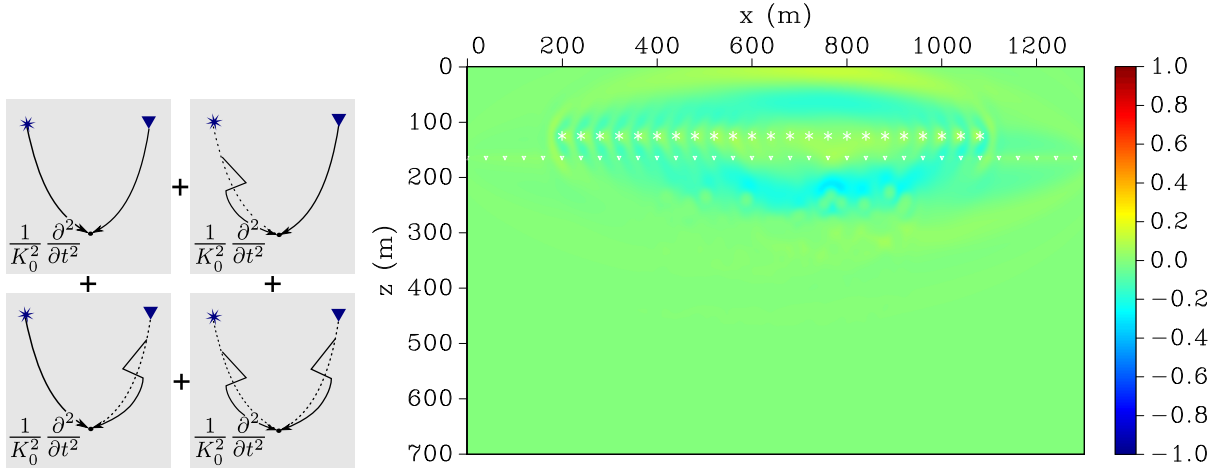


Figure 3.37: EXPERIMENT 8. Cross-term estimate resulting from the sum  $\left(\sum_j \bar{U}_{s,j}^\dagger\right) \hat{\delta p}_B$ .

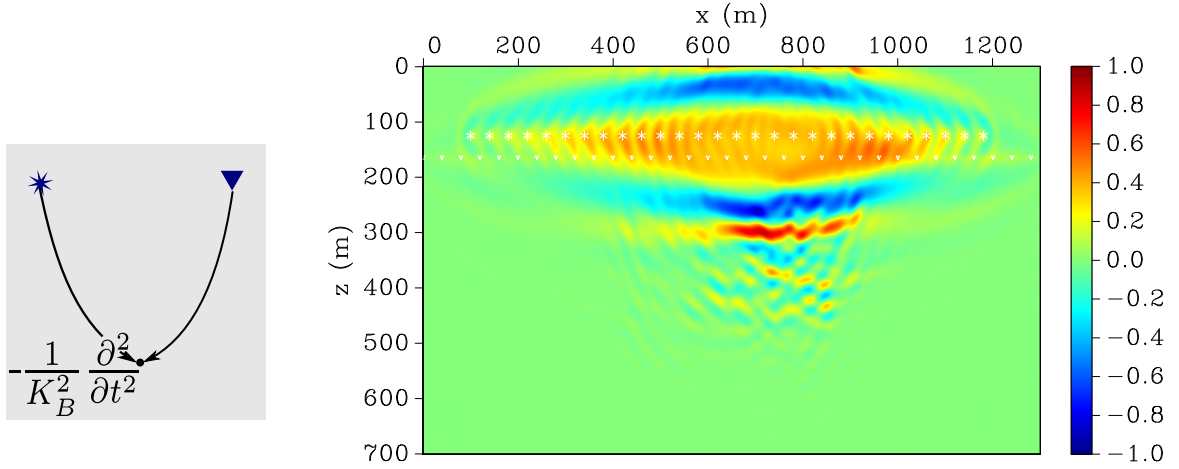


Figure 3.38: EXPERIMENT 8. Single-scattering contribution  $\delta K_{B,-BLB}^{\text{est}}$ .

of the receivers.

However, as before, the contributions  $\delta K_{B,BS}^{\text{est}}$  and  $\delta K_{B,SB}^{\text{est}}$  provide the best image of a single contribution, the last being the best one (see Figure 3.39, top and centre top). We compare this contribution to the single-scattering contribution,  $\delta K_{B,BB}^{\text{est}}$  (see Figure 3.39, centre bottom). This latter contribution is comparable to the result of the conventional approach (see Figure 3.39, bottom), where the complete true residual  $\delta p$  is backpropagated with a Green's function of a smooth background initial model.

It is worthwhile to observe that Experiment 8 is an example where conventional inversion pro-

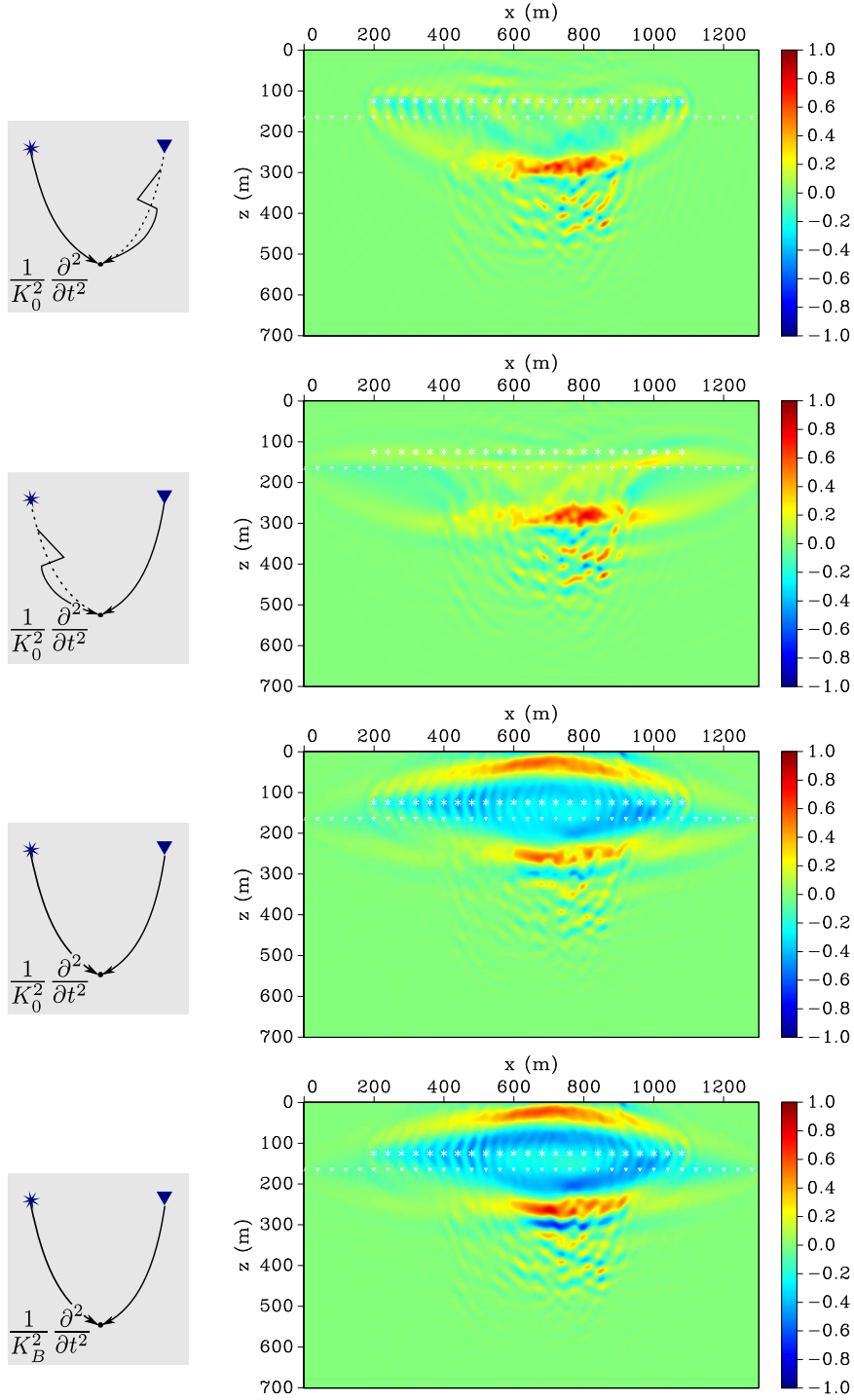


Figure 3.39: EXPERIMENT 8. Bulk-modulus perturbation estimates. *From top to Bottom:*  $\delta K_{B,BS}^{\text{est}}$ ,  $\delta K_{B,SB}^{\text{est}}$ ,  $\delta K_{B,BB}^{\text{est}}$  and conventional estimate  $\delta K_{\text{conv}}^{\text{est}}$  (backpropagation of complete residual  $\delta p$  with only background model).

### 3.3 Experiment gallery

---

duces a fast-varying sensitivity kernel in the sense of Zhu et al. (2009). This happens because the conventional approach is based on the assumption that the background is known and all perturbations happen in the scatterers. Thus, the inversion tries to blame the scatterers for all necessary model updates, even if, as in this example, the error is in the background model.

According to Zhu’s work, when dealing with reflection/scattered data, the occurrence of fast-varying sensitivity kernels is more frequent for narrower offsets (or reflection angles), while slowly-varying ones are more often observed in large-offset acquisitions. This is consistent with the results of our numerical experiments. Fast-varying sensitivity kernels cause slower convergence of the inversion procedure.

Why the multiple-scattering contribution  $\delta K_{B,SB}^{\text{est}}$  (Figure 3.39, centre top) based in a surface acquisition of reflection/scattered data is smoother and closer to the undecomposed estimate  $\delta K^{\text{est}}$  when compared to the conventional estimate (Figure 3.39, bottom)? This is an important observation, because it derives from the very purpose of the decomposition presented we mention before: the singular wavefield when properly handled can provide information on background perturbations. The singular wavefield originates at the singular component of the medium which acts as sources enhancing the illumination of the perturbation. In other words, this experiment allows to conclude that a medium can “illuminate” itself by means of scatterers, so that the need for long offsets is reduced.

Figure 3.40 shows the contributions from a single pair source-receiver (black star and triangle) of the same estimates seen in Figure 3.39. The self-illumination of the medium can be clearly seen when we compare the conventional estimate (Figure 3.40, top) with the other three estimates. While the perturbation estimates are distributed symmetrically around the Fresnel zone of the direct source-to-receiver wavefield, in the other images the estimate perturbations are placed in the scattering area preferably over the perturbed area.

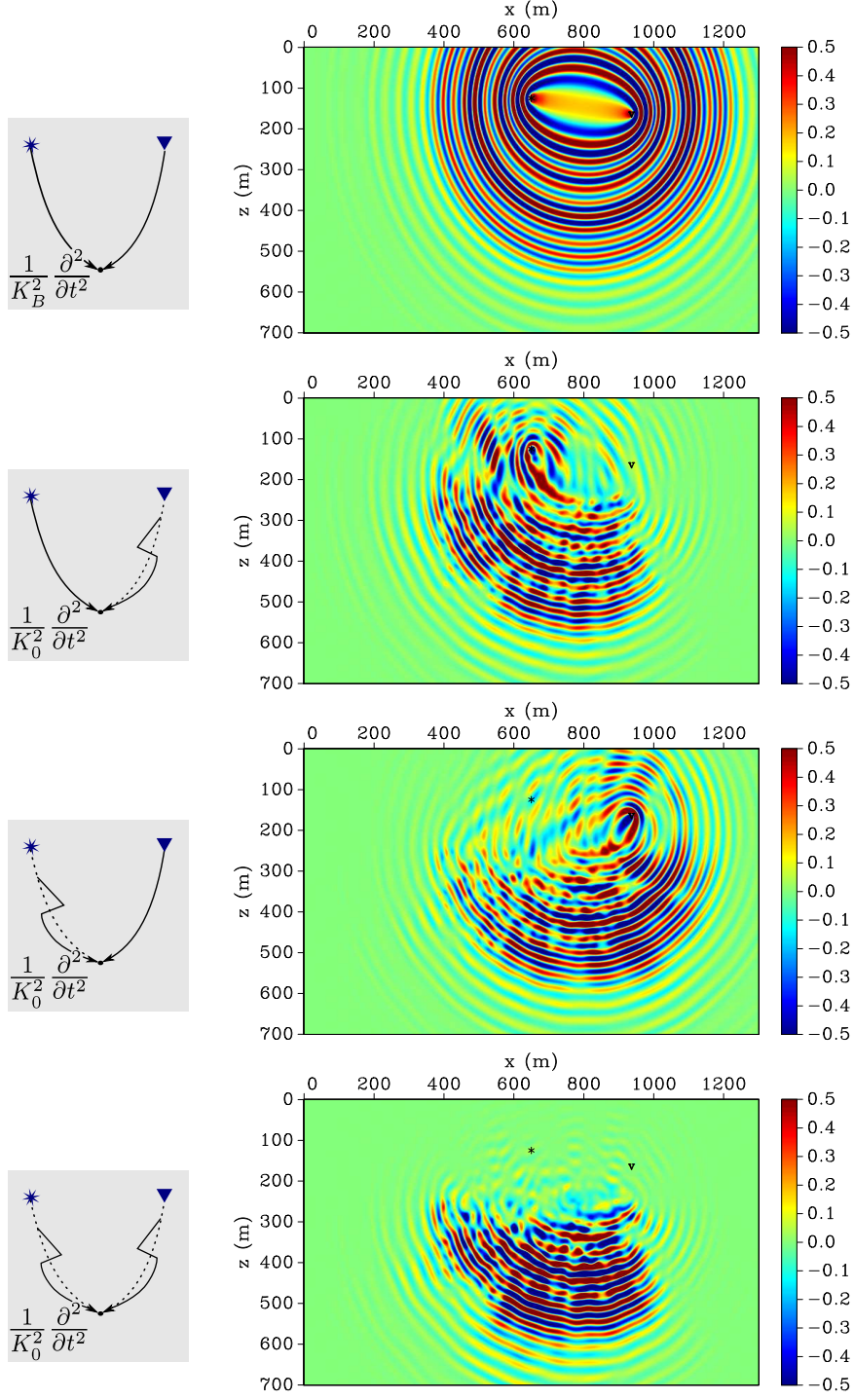


Figure 3.40: EXPERIMENT 8. *From top to bottom:* Single pair source-receiver contributions to conventional estimate  $\delta K_{\text{conv}}^{\text{est}}$ ,  $\delta K_{B,BS}^{\text{est}}$ ,  $\delta K_{B,SB}^{\text{est}}$ , and  $\delta K_{B,BB}^{\text{est}}$ .



## **4. Application: Scattering-based Sensitivity Kernels for Time-lapse Differential Waveform Inversion**

Over the past few years, time-lapse seismic surveys have become a powerful tool used to monitor the fluid-flow in a producing reservoir. Generally speaking, such a survey consists in acquiring and analysing multiple seismic data, repeated at the same site over time in order to look for differences from which one can infer the changes in the reservoir due to production. This is possible because, as fluid saturations and pressures in the reservoir change, the seismic reflection properties change accordingly.

A seismic image contains information on reflections that depends on both the geological structure and its fluid contents. On single-time images these contributions are coupled and difficult to separate. The basic approach in time-lapse surveys is to image the changes in the reservoir by subtracting subsequent time-lapse seismic images of the reservoir from one another. Assuming that geology is time-invariant during production and that repeatability in the seismic data acquisition is, in some way, assured, the image difference from a time-lapse survey would indicate the changes due only to the fluid-flow since, to first order, the geology part subtracts out since it is time invariant (Lumley, 2001).

Recently, full waveform inversion (FWI) has been used as an alternative time-lapse monitoring tool (Queißer and Singh, 2013; Yang et al., 2011; Zheng et al., 2011). FWI allows the reconstruction of high-resolution velocity models of the subsurface through the extraction of the full information content of seismic data (Tarantola, 1984a; Virieux and Operto, 2009). Since the FWI approach delivers high resolution quantitative images of the macro-scale physical parameter, it ought to be a good candidate for monitoring applications to reconstruct the parameter variation through time evolution (Asnaashari et al., 2012).



Asnaashari et al. (2012) studied the robustness of three different FWI methods applied to time-lapse problems: differential method, parallel difference method and the sequential difference method. In the differential method, instead of minimising the difference between the observed and modelled data, the difference of the data between two sets of data is minimised to obtain the time-lapse change estimate (Watanabe et al., 2004). The parallel difference method considers independent inversion of two subsequent data sets, the so called baseline and monitor data-sets, using a similar starting model (Plessix et al., 2010). The sequential difference method uses the recovered baseline model as a starting model for the monitor data inversion.

In this Chapter we apply our scattering-based decomposition to the time-lapse problem, considering the time-lapse change as a perturbation of the singular part of the model, i.e., perturbation of the scattering potential. Under the differential-method framework, we demonstrate that the scattering-based decomposition of the sensitivity kernels allows to take advantage of the illumination of the time-lapse change due to multiple-scattering phenomena in order to improve the perturbation estimates from FWI.

## 4.1 Time-lapse survey as a FWI scattering problem

To apply the decomposition theory developed in this thesis to a time-lapse survey, we think of the baseline model as decomposed into a background and a singular part. We consider both parts of the baseline model as known. The background part could be, for instance, the best smooth velocity model from a conventional inversion technique, and the singular part could be determined by the reflector positions in the corresponding migrated image. Then, the time-lapse changes can be considered as (unknown) perturbations to the singular and/or background parts.

The situation is the simplest, if we consider the time-lapse change as a perturbation of the singular part only. According to equation (2.44), the data difference between baseline and monitor is then the singular-wavefield residual, i.e.,  $\delta p = \delta p_S$ . It means that the data residual  $\delta p^{\text{true}}$  from the difference between the baseline and monitor surveys is going to be backprojected as  $\delta p_S^{\text{true}}$  according to eq. (3.7), that is, the model perturbation estimates are

$$\delta K_S^{\text{est}} = \sum_j \delta K_{S,j}^{\text{est}} = \sum_j \bar{U}_{S,j}^\dagger \widehat{\delta p_S^{\text{true}}}, \quad (4.1)$$

where  $j$  takes the values explained in connection with eq. (3.7).

## 4.2 Numerical experiments

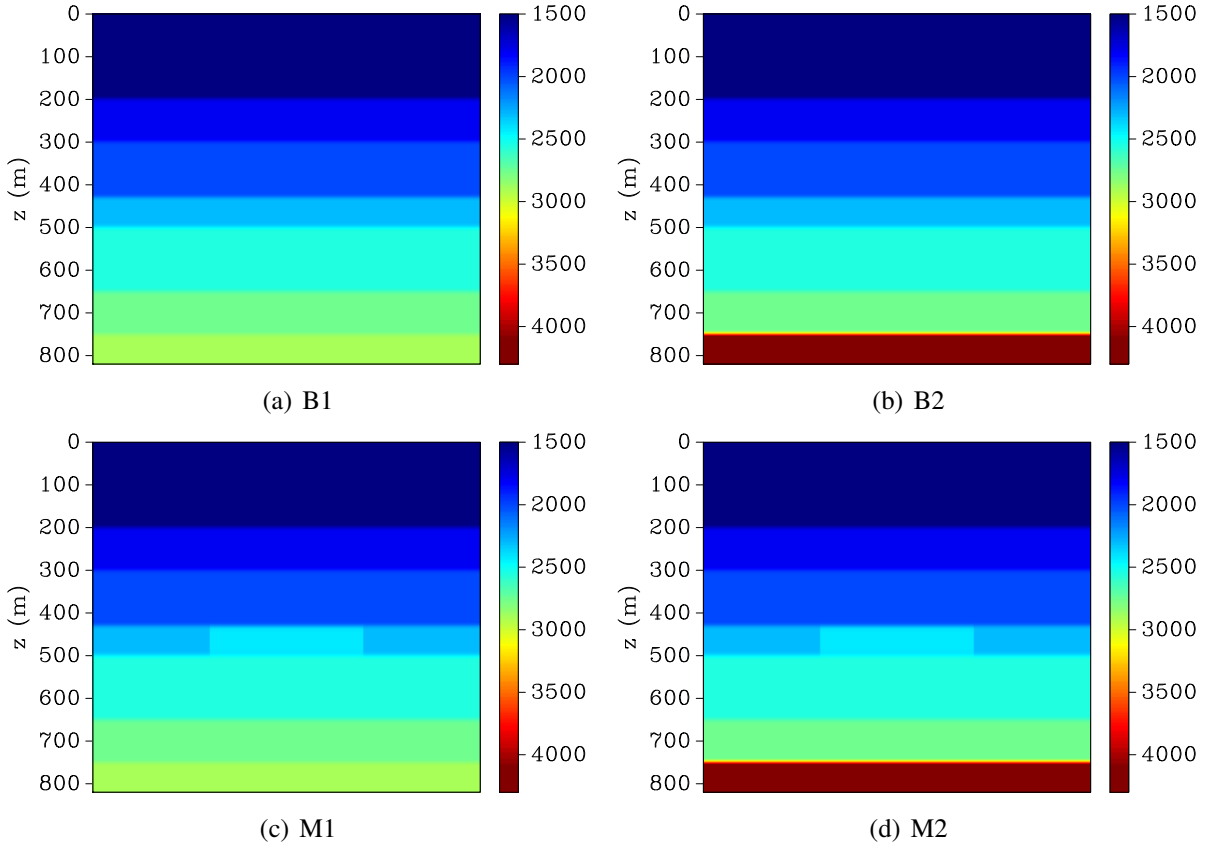


Figure 4.1: Center parts of the baseline (a and b) and monitor (c and d) models for Model Sets 1 (a and c) and 2 (b and d).

## 4.2 Numerical experiments

To test our approach, we have set up four synthetic time-lapse experiments with simple layered models. These experiments used two sets of baseline and monitor models named as Set 1 (Figures 4.1a and 4.1c, respectively) and Set 2 (Figures 4.1b and 4.1d, respectively) and two types of acquisition geometry. The two baseline models are identical except for the velocity contrast at the deepest interface. In the first model, this interface is a weak reflector and in the second model, a strong one. The monitor models differ from the baseline models in both sets by a 5% velocity perturbation in a rectangular lens of 400 m width within the 4<sup>th</sup> layer. The density was considered constant in all models ( $2200 \text{ kg/m}^3$ ). We used two passes of a 20-point triangular moving-average filter on the bulk-modulus baseline models to create the corresponding background models. The singular models are the differences between the complete and background models. For each model

set we performed two experiments with different acquisition geometry at the surface: long and narrow offsets.

In this study we consider both the background and singular parts of the baseline velocity model as fully known.

The numerical experiments consisted of the following basic steps:

1. Simulation of the seismic acquisition by modelling the wavefields in the baseline and monitor models. We used the same modeller as in the numerical experiments of the previous chapter. In the simulations the vertical grid spacing was  $2\text{ m}$ , the horizontal grid spacing was  $4\text{ m}$ , and the time-marching step  $0.2\text{ ms}$ . The recorded wavefields were then resampled to a  $6\text{ ms}$  interval.
2. Computation of the true singular-wavefield residual, which in this case is identical to the true full-wavefield residual, i.e., the difference between the wavefields in the monitor and baseline models.
3. Modelling the wavefield in the smooth background model.
4. Computation of the baseline singular wavefield, i.e., the difference between the wavefields in the baseline and background models.
5. Backpropagation of the true singular-wavefield residual using the receiver-side reference and singular wavefields.
6. Determination of the four contributions to equation (4.1) by crosscorrelation of the back-propagated wavefield residual with the source-side reference and singular wavefields as exemplified in equation (2.51).
7. Stack of the resulting contributions over sources and receivers.

As in Chapter 3, we normalised the total bulk-modulus perturbations estimates to their maximum absolute value. Other estimates, as also individual contributions, are depicted using the same scale factor.

### 4.2.1 Large offset surveys

The first acquisition geometry consisted of 87 shot gather with 726 receivers each, placed at  $10\text{ m}$  depth, spaced at  $4\text{ m}$ , and covering the distance between  $0\text{ m}$  and  $2900\text{ m}$ . Sources were placed at  $40\text{ m}$  depth spaced at  $32\text{ m}$ , and covering the distance between  $80\text{ m}$  and  $2832\text{ m}$  earth's

## 4.2 Numerical experiments

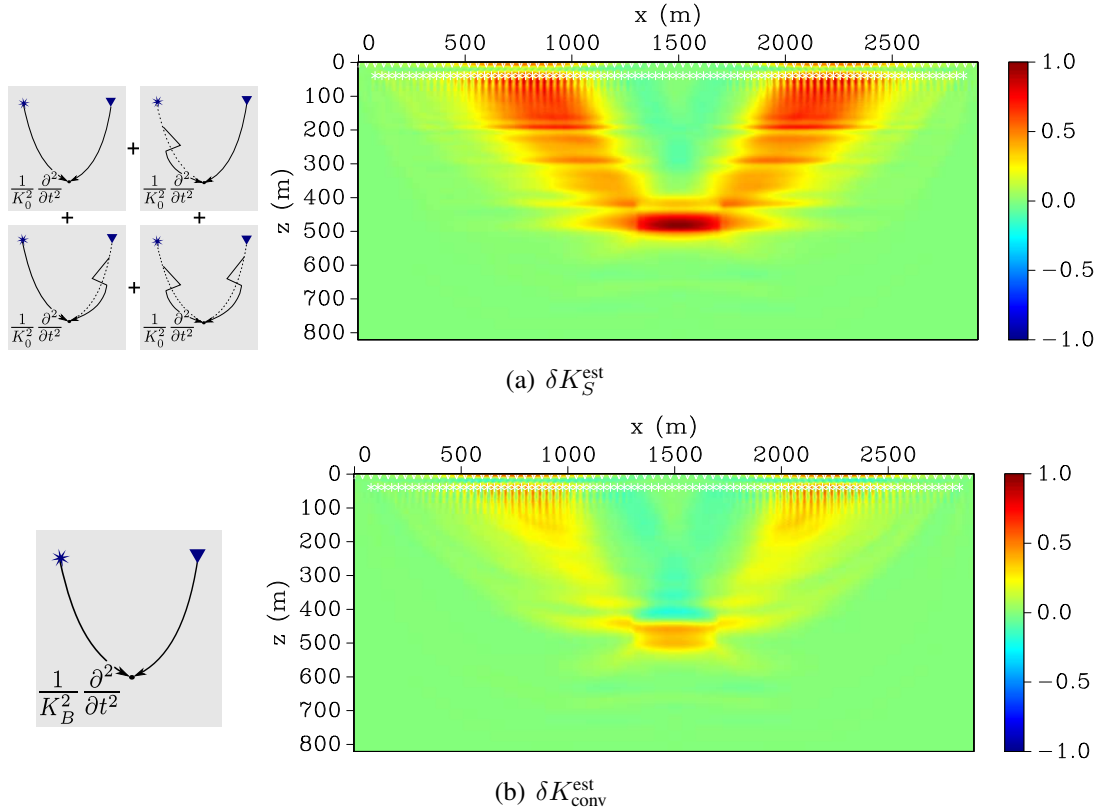


Figure 4.2: SET 1 - LARGE OFFSET. Estimates for bulk-modulus perturbation in the absence of a strong reflector. (a) Singular part estimate according to equation (4.1); (b) Conventional perturbation estimate using wave propagation in the background model. The last is substantially close to contribution  $\delta K_{S,BB}^{\text{est}}$  (see Figure 4.4c, right). Both figures are normalised to their common maximum for comparison

surface. In other words, the survey was positioned almost symmetrically over the rectangular perturbation, centred at 1500 m. Figure 4.2 indicates the source positions by white stars and every 10<sup>th</sup> receiver position by a downward-pointing triangle.

### Set 1

In Figure 4.2a, we show the estimate for the singular-part bulk-modulus perturbation in Model Set 1 with all four contributions according to equation (4.1). For comparison, Figure 4.2b shows the corresponding estimate from the conventional approach as in eq. (3.8), i.e., the result that would be obtained from inversion with a smoothed background model. In Figure 4.2 both estimates are normalised to the same factor which is the maximum value of  $\delta K^{\text{est}}$ . We see, then, that the

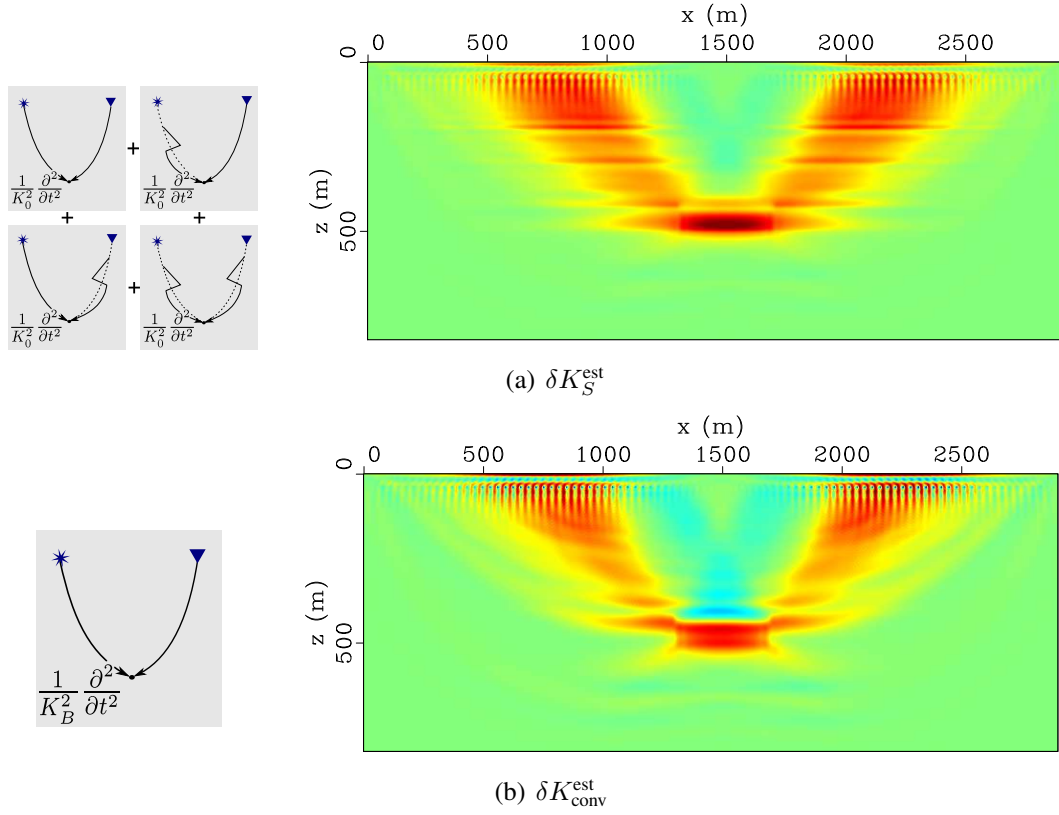


Figure 4.3: SET 1 - LARGE OFFSET. The same estimates seen in Figure 4.2 ( $\delta K_S^{\text{est}}$  and  $\delta K_{\text{conv}}^{\text{est}}$ ) but now each one is normalised to their own maximum.

amplitude the perturbation estimates in  $\delta K_{\text{conv}}^{\text{est}}$  are smaller giving the impression that the artefacts there are weaker.

We see that in the absence of a strong reflector in the bottom of the model, the conventional sensitivity kernel gives a satisfactory estimate. One can also argue that it is even better to the one from the summed contributions. This last one has the time-lapse change better delineated but shows artefacts (“wings”) connecting the acquisition surface to the time-lapse change.

However when we inspect Figure 4.3 we see that the artefacts are equivalent in both conventional and singular part estimates. But in Figure 4.3 we depicted the very same estimates, each one now normalised to their own maximum. In this way, we can see that, proportionally, they are equivalent.

Aside the artefacts (we discuss the further in the text) we can compare the estimates depicted in Figure 4.3. Definitely,  $\delta K_S^{\text{est}}$  is better than  $\delta K_{\text{conv}}^{\text{est}}$ : the perturbation is better delineated, specially at its top, where the conventional one has an non-desired oscillatory behaviour. As a matter of

## 4.2 Numerical experiments

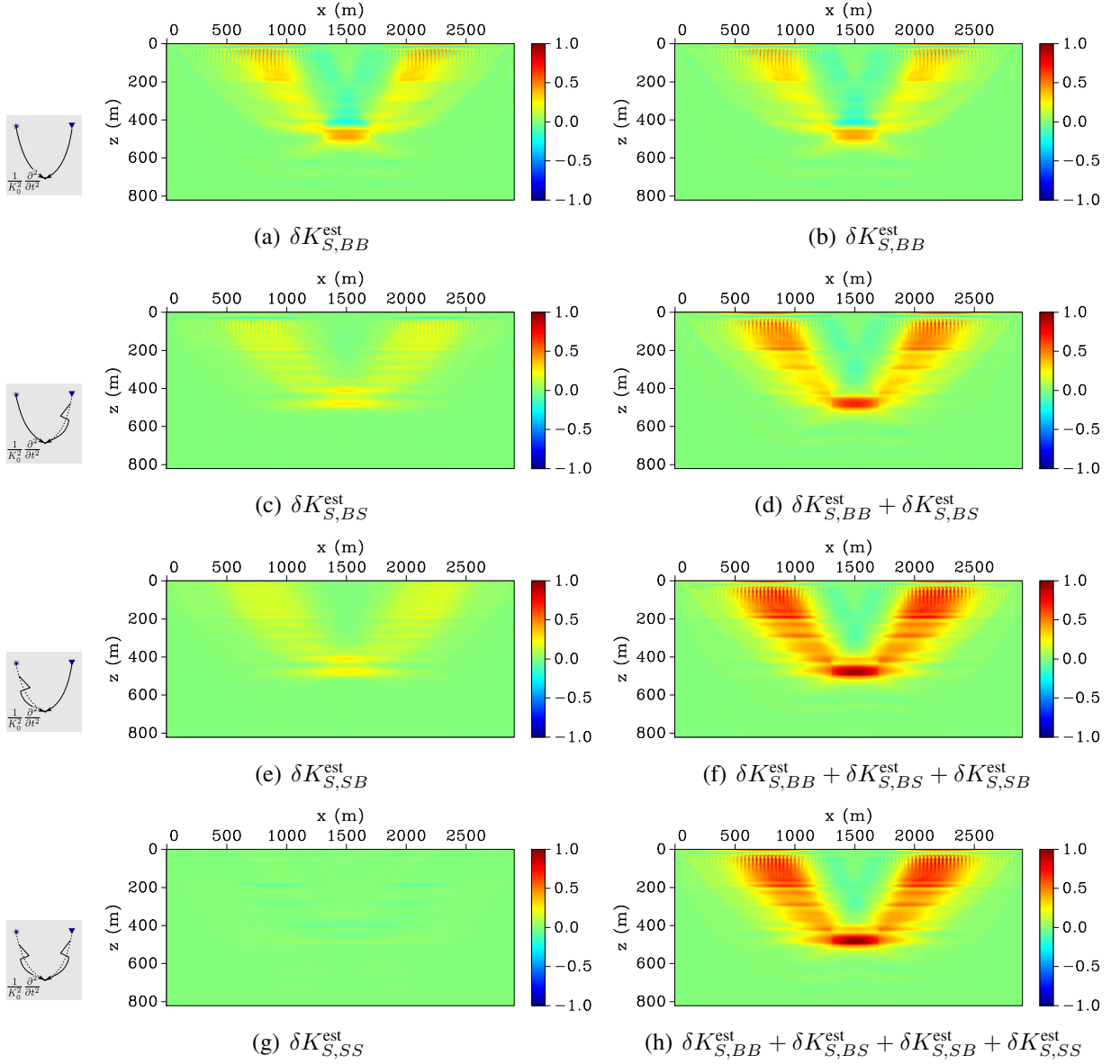


Figure 4.4: SET 1 - LARGE OFFSET. Individual contributions to  $\delta K_S^{\text{est}}$  are shown in the left column. In the right column we see the results of succesively summing up the contributions. Thus, the last one is the result of summing up all four contributions.

fact, the combination of two individual contributions,  $\delta K_{S,BS}^{\text{est}}$  and  $\delta K_{S,BB}^{\text{est}}$  (Figure 4.4c, left) gives, we believe, the best estimate.  $\delta K_{S,BB}^{\text{est}}$  gives nice estimate of perturbation (format of the rectangle correctly recovered) and  $\delta K_{S,BS}^{\text{est}}$  delineates top and bottom, consistent of being a scattering contribution.

The artefacts in both estimates shown in Figure 4.3 are very similar. The actual difference is that in  $\delta K_S^{\text{est}}$  the artefacts show abrupt amplitude changes while in  $\delta K_{\text{conv}}^{\text{est}}$  they are smoother. This is easily explained by the weighting factors of the scattering operator:  $\frac{1}{K_0}$  and  $\frac{1}{K_B}$ . Let us compare the estimates  $\delta K_{\text{conv}}^{\text{est}}$  (Figure 4.2b) and  $\delta K_{S,BB}^{\text{est}}$  (Figure 4.4a). The only difference between them is the different factor used. Comparing the two estimates we can see the imprint of the  $\frac{1}{K_0}$  in  $\delta K_{S,BB}^{\text{est}}$ . As expected, the intensity of the artefacts decreases as we go deeper because the factors get smaller. The factor  $\frac{1}{K_B}$  decreases in a smooth fashion, while  $\frac{1}{K_0}$  contains the full layer information. Note that their magnitude are comparable.

These artefacts are also present in the other contributions. In Figure 4.4 we see each of the individual contributions (left column) and the results of successively adding them up (right column). We can see that these artefacts are comparable in  $\delta K_{S,BS}^{\text{est}}$  and  $\delta K_{S,SB}^{\text{est}}$  (Figures 4.4c and 4.4e) as in the single scattering contribution.

The cause of the artefacts is the uneven illumination due to the combination of surface acquisition and a horizontally stratified medium. In Figure 4.5 we can see the norm of the residual (as a measure of the energy) at the receivers as a function of shots. As expected, due to horizontal stratification of the medium, for shots at the left side of the model (around shot number 20) the most energetic residuals are recorded between 2000 and 2500  $m$ . In the reverse case, most energetic residuals are recorded between 500 and 1000  $m$ . These two ranges are the exact location of the artefacts.

## Set 2

The corresponding results for Model Set 2 with the strong bottom reflector are depicted in Figure 4.6. While the estimate from all contributions (Figure 4.6a) together looks similar to the one in Figure 4.2a, the conventional estimate is visibly different (compare Figure 4.2b and 4.6b), indicating that in this situation, the use of a smooth background model to simulate the backpropagation is insufficient. We conclude that in this case, multiple-scattering contributions carry important information that cannot be neglected. We may think of the scatterers contributing to the inversion by back-illumination. There are two overlapping V-shaped features (see Figure 4.6c) with negative

## 4.2 Numerical experiments

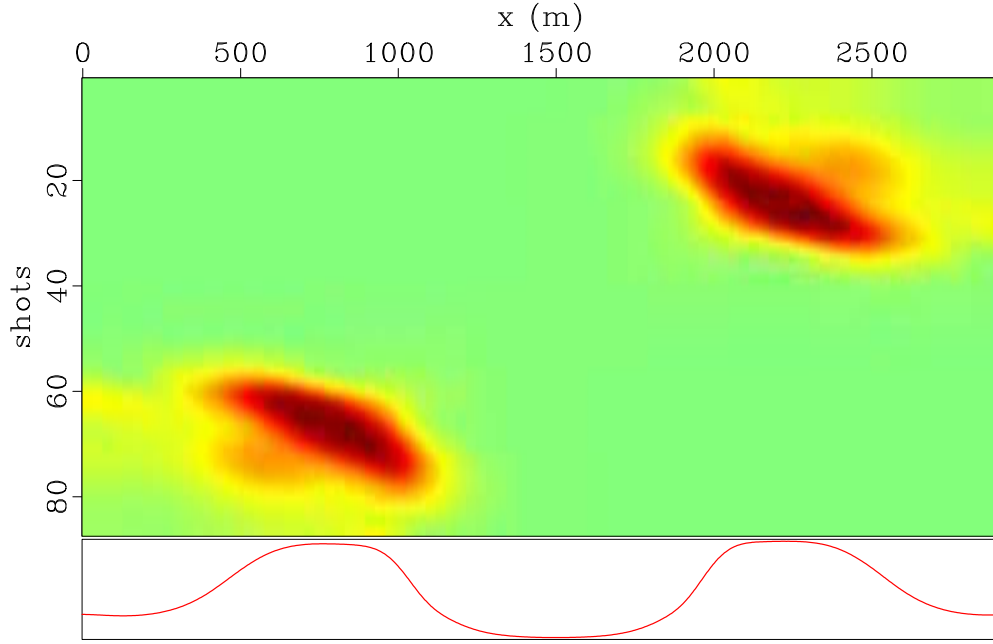


Figure 4.5: SET 1 - LARGE OFFSET. *Top*: the norm of the residual (as a measure of the energy) at the receivers (horizontal axis) as a function of shots (vertical axis). *Bottom*: Result of stacking (integrating) the above function over sources. The energy of the residuals concentrate between 500 to 1000  $m$  and 2000 to 2500  $m$ .

amplitude superimposed in the figure. These features are due to the background wavefield trying to deal with the inverted-phase residual due to the deep reflection at the strong reflector. Since reflections are not accounted for in the smooth background model, they appear in the wrong place, with the wrong sign. Note that they coincide nicely with the ray paths of the main primary reflections (see Figure 4.6c).

These features are also present in the background contribution  $\delta K_{S,BB}^{\text{est}}$  (see Figure 4.7a) which is also single-scattering based. But the multiple-scattered based contributions (see Figures 4.7c, 4.7e and 4.7g) correct this problem. For in the same regions they show positive values which cancel out the negative parts of  $\delta K_{S,BB}^{\text{est}}$ . This become evident when we inspect the right column of Figure 4.7 and see the effect of successive addition.

Here, even more than the previous example, the estimate resulting from the sum of  $\delta K_{S,BB}^{\text{est}}$  and  $\delta K_{S,BS}^{\text{est}}$  seems to give the best estimate as we can verify in Figure 4.7d.

The artefacts are present in the estimates as before. In Figure 4.6a we can see that the width of each wing of the artefact is larger then the previous example. However, the cause is still the same. In Figure 4.8, the perturbation illumination is increased with the aid of the strong bottom reflector



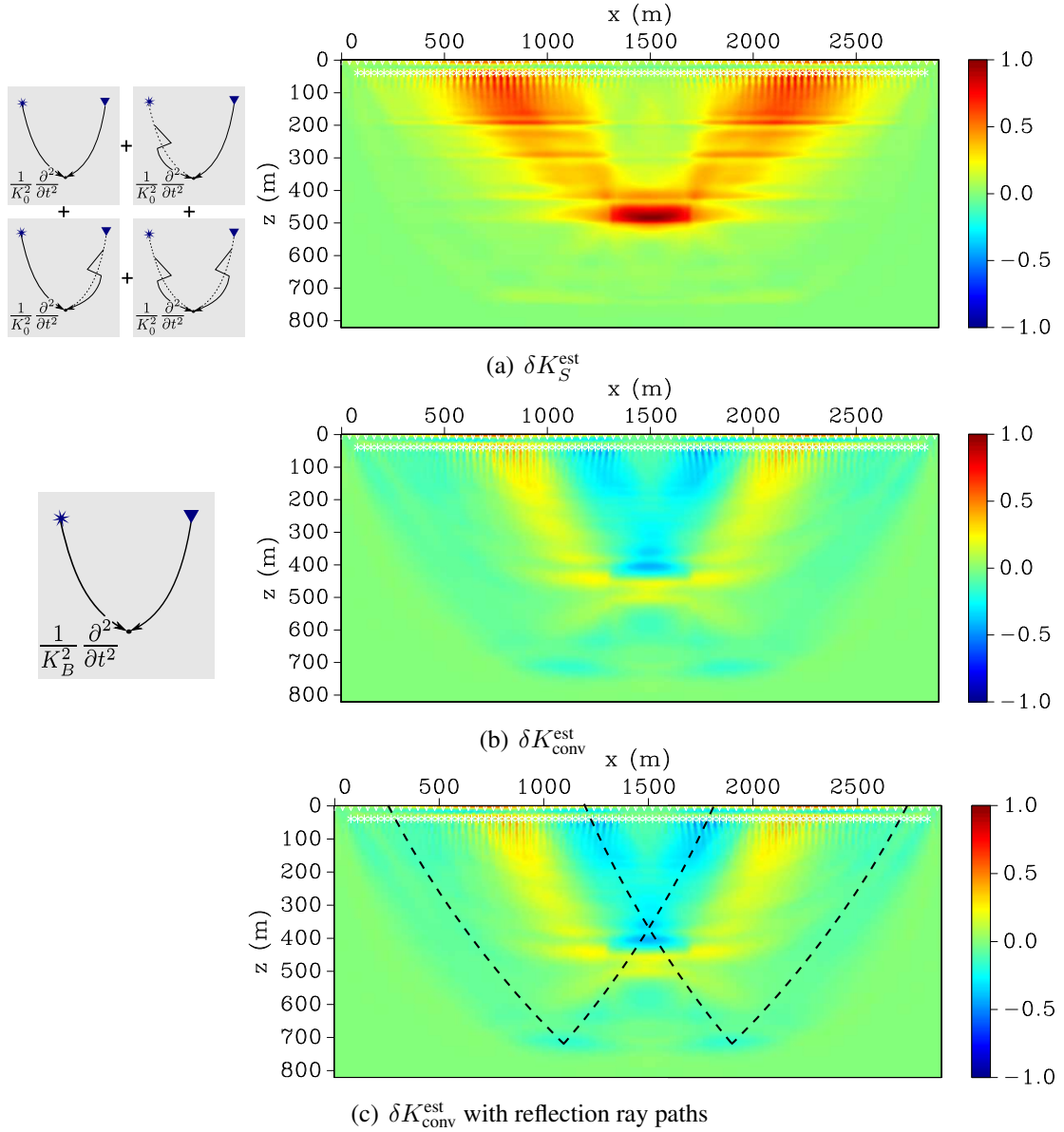


Figure 4.6: SET 2 - LARGE OFFSET. Estimates for bulk-modulus perturbation in the presence of a strong reflector. (a) Singular part estimate according to equation (4.1); (b) Conventional perturbation estimate using wave propagation in the background model. The last is substantially close to contribution  $\delta K_{S,BB}^{est}$  (see Figure 4.7c, right); (c) Same as before, but with dashed ray paths indicating the back-illumination from the strong reflector.

## 4.2 Numerical experiments

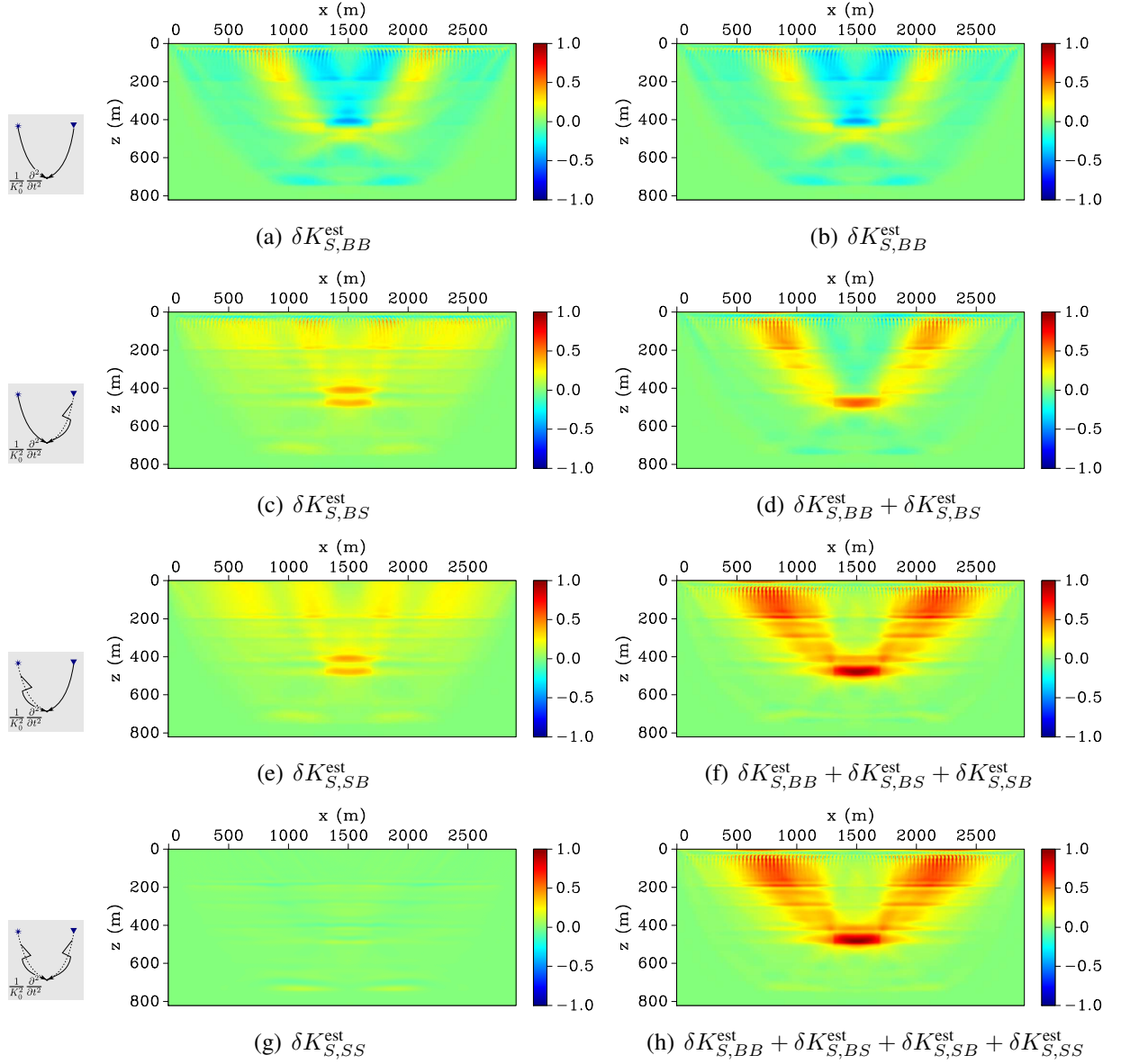


Figure 4.7: SET 1 - LARGE OFFSET. Individual contributions to  $\delta K_S^{\text{est}}$  are shown in the left column. In the right column the result of summing up the current contribution with all the previous ones, which makes, for instance, the first estimate in the left to be no more than  $\delta K_{S,BS}^{\text{est}}$ ; on its turn, the last one is the result of summing up all four contributions.

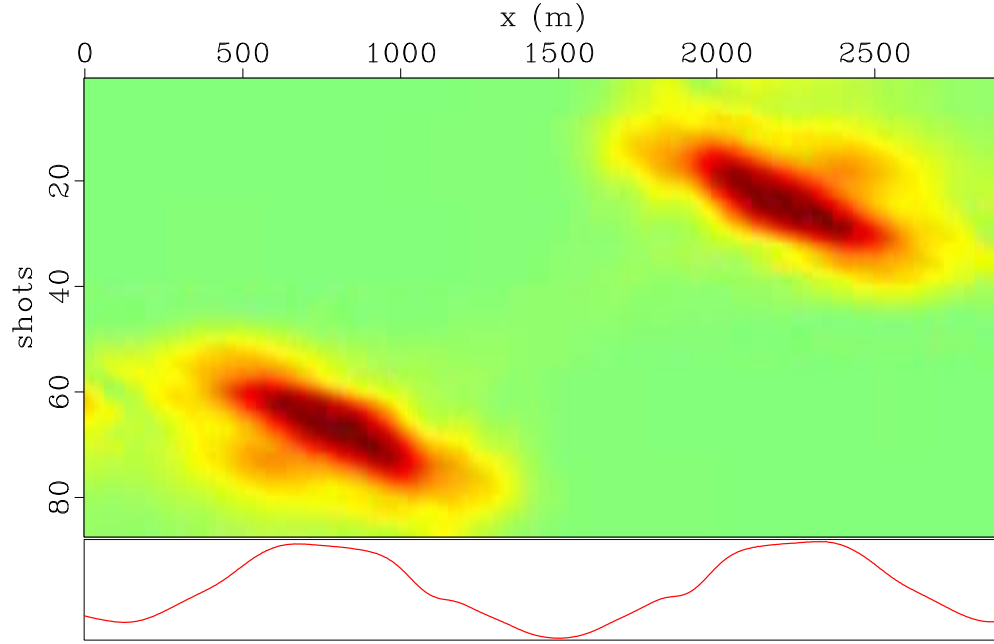


Figure 4.8: SET 2 - LARGE OFFSET. *Top*: the norm of the residual (as a measure of the energy) at the receivers (horizontal axis) as a function of shots (vertical axis). *Bottom*: Result of stacking (integrating) the above function over sources.

indicated by the enlargement of the “bright spots” in the graphic.

#### 4.2.2 Narrow offset surveys

It is known in the literature on FWI that the acquisition geometry has a strong influence on the inversion result. When inversion is done with enclosing source and receiver arrays, it behaves as in transmission tomography, which is known to recover very well the low-frequency information (see, e.g., Pratt and Shipp, 1999; Brenders and Pratt, 2007). On the other hand, according to Zhu et al. (2009), when dealing with reflection/scattered data, narrower offsets (or reflection angles) may lead to fast-varying sensitivity kernels, in opposition to slow-varying ones in large-offset acquisitions. Fast-varying sensitivity kernels cause slower convergence of the inversion procedure.

To investigate this effect in the case of the decomposed inversion, we repeated the above experiments with a shorter survey. The new acquisition geometry consisted of 48 shot gathers with 251 receivers each, and covering the distance between 0 m and 1000 m. Sources were spaced at 20 m and covering the distance between 48 m and 948 m earth’s surface. Thus, the survey was positioned almost symmetrically over the rectangular perturbation, now centred at 500 m. All other

## 4.2 Numerical experiments

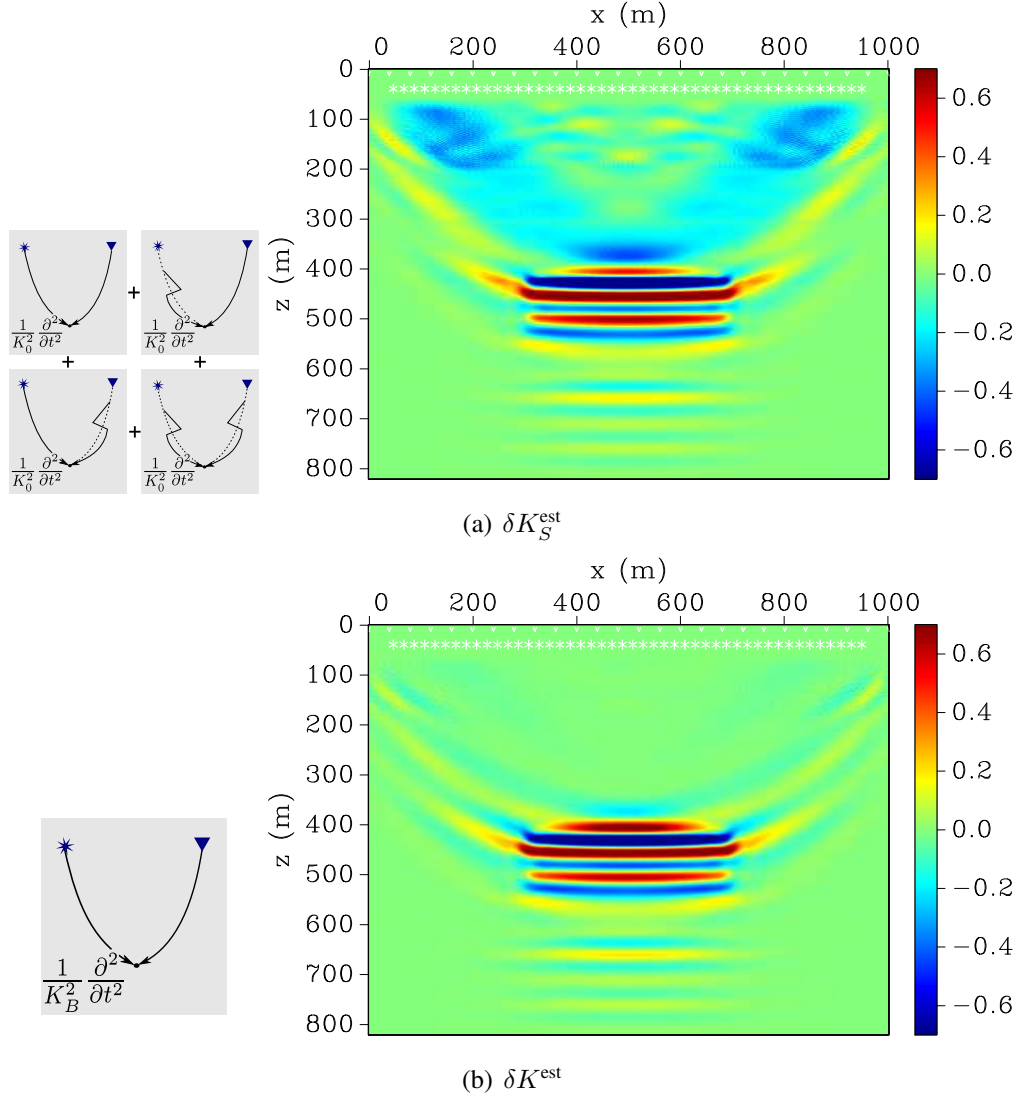


Figure 4.9: SET 1 - NARROW OFFSET. Narrow-offset estimates of the perturbation in the absence of a strong reflector. (a) Singular estimate with all contributions; (b) conventional estimate.

parameters were the same as in the first survey. Figure 4.9 indicates the source positions by white stars and every 10th receiver position by a downward-pointing triangle.

Figure 4.9 shows the resulting perturbation estimates for the model with the weak bottom reflector. To avoid problems with inversion near the sources we muted down to a depth of 90  $m$ .

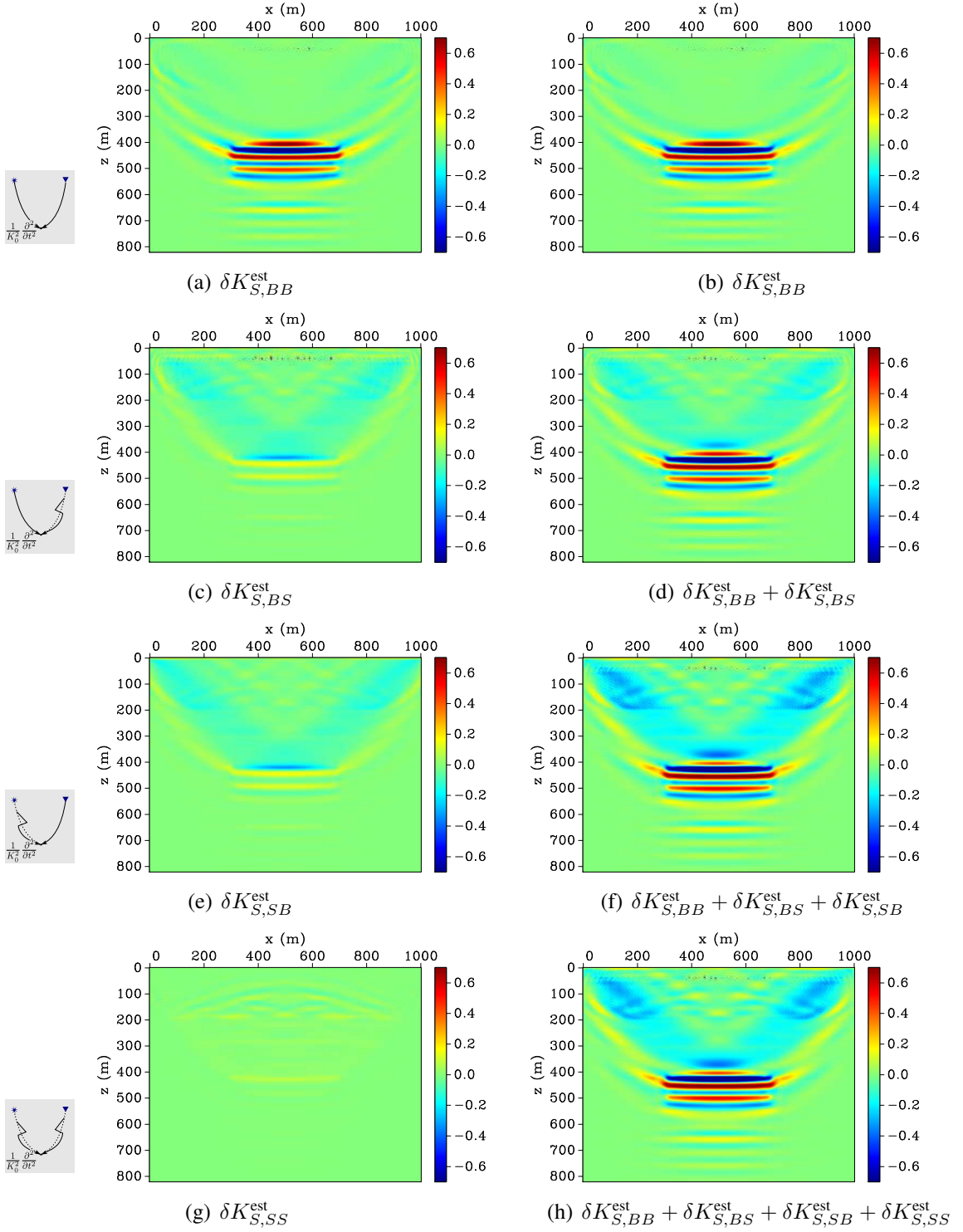


Figure 4.10: SET 1 - NARROW OFFSET. *From top to bottom, left column:* contributions to singular perturbation estimate  $\delta K_{S,BS}^{\text{est}}$ ,  $\delta K_{S,BB}^{\text{est}}$ ,  $\delta K_{S,SS}^{\text{est}}$  and  $\delta K_{S,SB}^{\text{est}}$ . *Right column:* Cumulative summation of individual contributions.

## 4.2 Numerical experiments

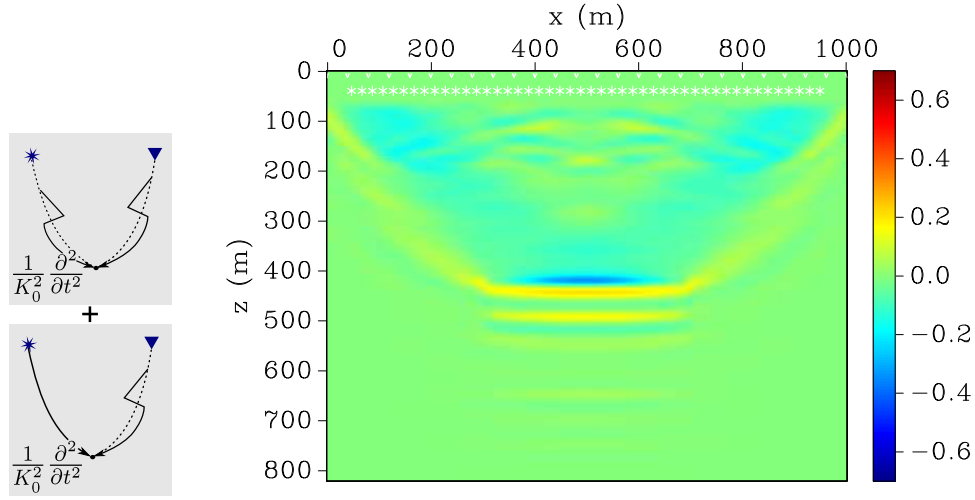


Figure 4.11: SET 1 - NARROW OFFSET. Summation of the multiple-scattering contributions  $\delta K_{S,SS}^{\text{est}}$  and  $\delta K_{S,SB}^{\text{est}}$ .

### Set 1

Figure 4.9a shows the narrow-offset estimate for the singular-part bulk-modulus perturbation in the absence of strong reflector. Figure 4.9b depicts the corresponding conventional estimate. As expected, the figures display fast-varying estimates due to the narrow-offset illumination. In the time-lapse perturbation area, the conventional estimate is rather similar to the four-contributions estimate. Actually, it can be considered a better estimate because it presents fewer artefacts than the four-contribution estimate.

In Figure 4.10 we see in the left column the four individual contributions to the singular estimate,  $\delta K_{S,BS}^{\text{est}}$ ,  $\delta K_{S,BB}^{\text{est}}$ ,  $\delta K_{S,SS}^{\text{est}}$ ,  $\delta K_{S,SB}^{\text{est}}$  and their successive sums (right column). We see that the dominant term is the single-scattering based contribution  $\delta K_{S,BB}^{\text{est}}$ . This characterises the narrow-offset case. We can compare this result with Figure 4.4, left column, where we can see the individual contributions in the long-offset case. There, in contrast, we see that the weight of each contribution is more balanced.

Although much weaker than the dominant contribution, the multiple-scattering based contributions yield estimates that will definitely lead to a better convergence in the inversion. In Figure 4.11, we see the sum of two multiple-scattering based contributions:  $\delta K_{S,SS}^{\text{est}}$  and  $\delta K_{S,SB}^{\text{est}}$ . There we see that the top and bottom of the time-lapse perturbation are nicely delineated. Moreover, the overall estimate does not present the fast-varying behaviour commonly seen in surveys



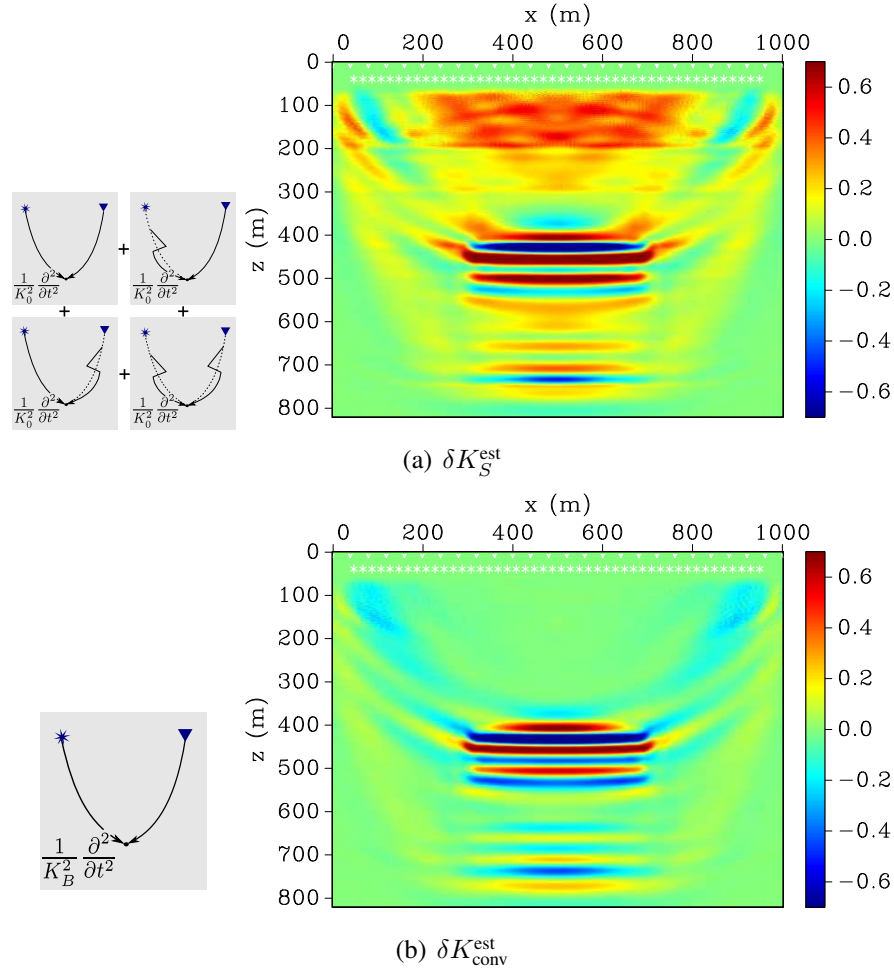


Figure 4.12: SET 2 - NARROW OFFSET. Narrow-offset estimates of the perturbation in the presence of a strong reflector. (a) Singular estimate with all contributions; (b) conventional estimate.

like this.

The drawback of the multiple-scattering contributions is the introduction of artefacts not present in the conventional result.

## Set 2

Figure 4.12a shows the narrow-offset estimate for the singular-part bulk-modulus perturbation for the model with the strong bottom reflector. Figure 4.12b depicts the corresponding conventional estimate. The latter, in the same way as in the previous example, displays fast-varying estimates due to the narrow-offset illumination. However, the four-contributions estimate is a bit better with

## 4.2 Numerical experiments

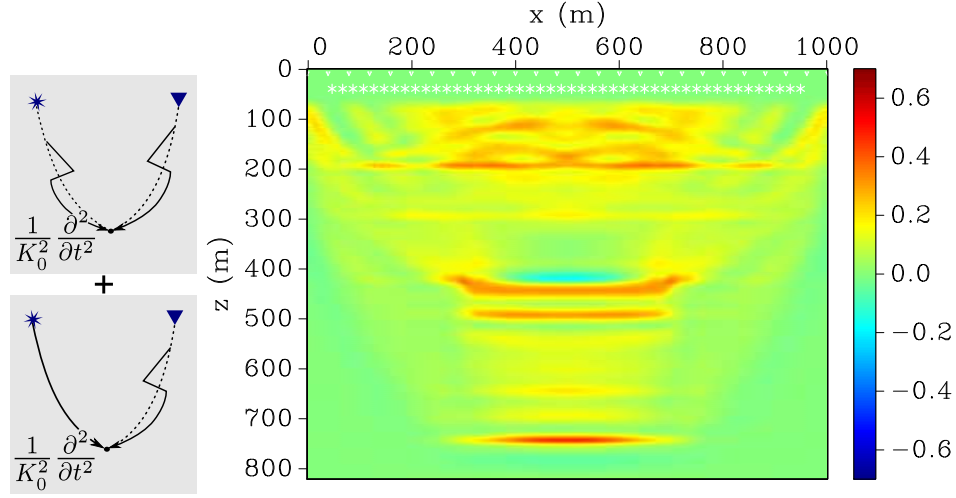


Figure 4.13: SET 2 - NARROW OFFSET. Summation of the multiple-scattering contributions  $\delta K_{S,SS}^{\text{est}}$  and  $\delta K_{S,SB}^{\text{est}}$ .

the strong reflector present, indicating the importance of self-illumination. The multiply-scattered information improves the perturbation estimates, providing a less varying and more delineated response. This indicates that the multiple-scattering contribution perceives both the low- and high-frequency content of the perturbation.

But the best result is seen in Figure 4.13, which shows the summation of the contributions to the singular estimate calculated with the source-side singular wavefield:  $\delta K_{S,SB}^{\text{est}}$  and  $\delta K_{S,SS}^{\text{est}}$ . We see that these contributions alone do a better job than the conventional estimate of Figure 4.12b, nicely delineating the top and bottom of the medium perturbation. The reason is that in this case not the source array alone, but also the singularities contribute to the source-side wavefield in the kernel. Physically, this can be interpreted as the strong reflector illuminating the time-lapse change from bottom to top.

Again, the drawback is the introduction of artefacts that come from the multiple-scattering contributions as can be verified inspecting Figure 4.14.

These narrow-offset experiments demonstrate that it is even more important than for long offset surveys to include the singular part into the model used for backpropagation, since it enhances the self-illumination due to scattering and leads to better perturbation estimates.



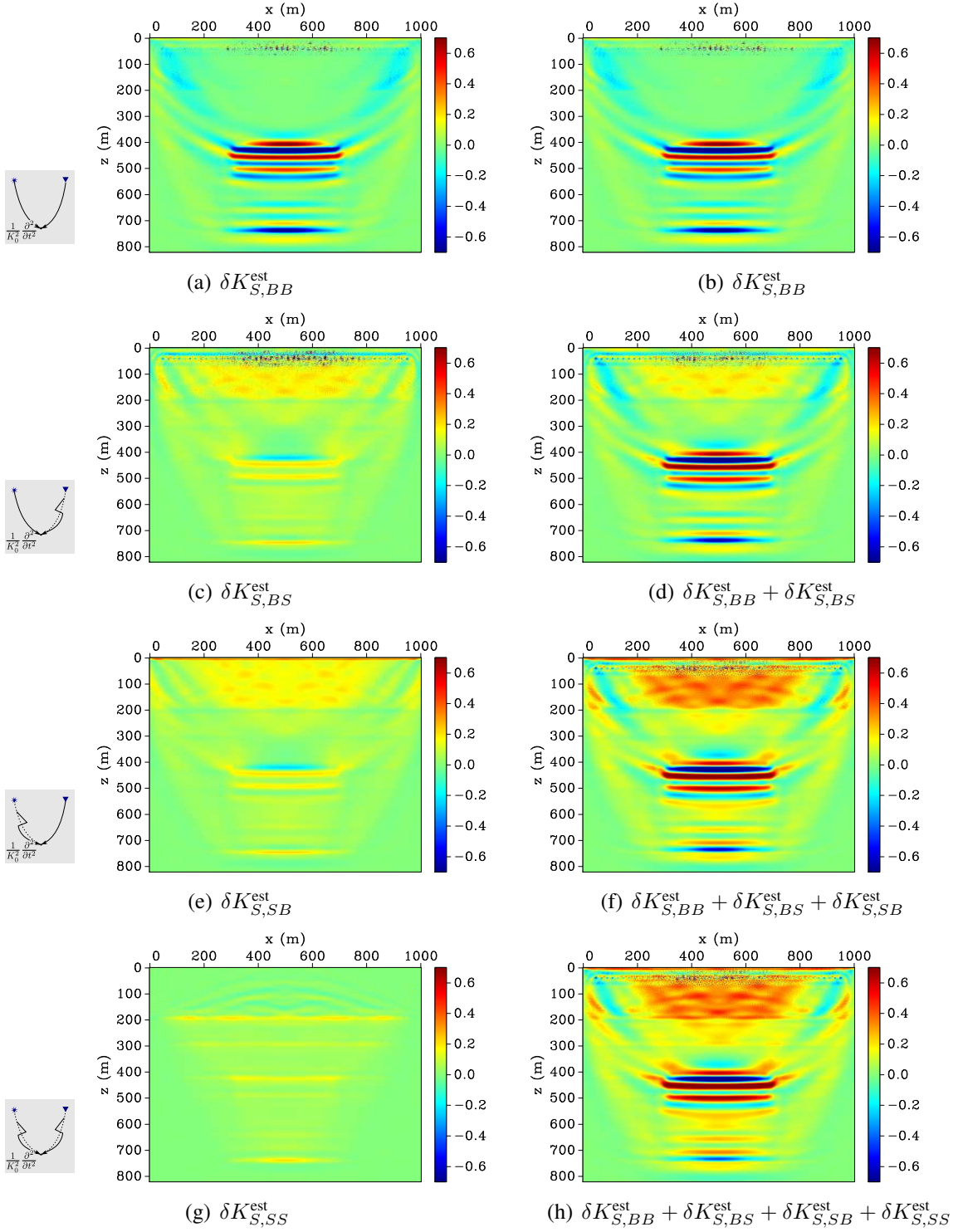


Figure 4.14: SET 2 - NARROW OFFSET. *From top to bottom, right column:* contributions to singular perturbation estimate  $\delta K_{S,BS}^{\text{est}}$ ,  $\delta K_{S,BB}^{\text{est}}$ ,  $\delta K_{S,SS}^{\text{est}}$  and  $\delta K_{S,SB}^{\text{est}}$ . *Left column:* Cumulative summation of individual contributions.

## 5. Final Remarks

We have presented a decomposition of standard full-waveform inversion (FWI) sensitivity kernels into several sub-kernels. We understand as standard FWI the least-square gradient-based nonlinear iterative local optimisation of the full recorded data where the sensitivity kernels are used to evaluate the gradient of the objective function in order to define the descending direction in the minimisation process. For this purpose, we have developed a scattering formulation that decomposes the model parameters into a background and a singular component. We assume that the model decomposition is based on its spatial-frequency content, that is, the smooth background model carries the low-frequency content while the singular component carries the high-frequency information. Both model parts are then perturbed instead of considering the scattering potential as a perturbation of the background and inverting for it. As a consequence, the single sensitivity kernel of the standard formulation of FWI is decomposed into a sum of subkernels.

By including a singular part in the unperturbed reference model, the present approach describes multiple-scattering phenomena. Nonetheless, the resulting subkernels are still based on the Born approximation, in the sense that the residuals are predicted based on single-scattering of the unperturbed wavefields at the model perturbations, which thus must be small compared to the unperturbed model parts. Once the model component perturbations lie in the range of validity of the decomposition approximation, the superposition of all our sub-kernels yields the total Fréchet gradient as in the standard formulations of FWI. The asset of the decomposition is that it uncovers the underlying different levels of interaction between the model components within the data that else are hidden in the sensitivity kernel for the total wavefield, allowing to select the most useful ones depending on the problem and/or model decomposition used. We have used the decomposition to analyse the interplay of the background wavefield and the internally multiple-scattered or reflected wavefield in each subkernel to understand their role in the forward problem.

The forward-problem numerical experiments showed that the linearised total-wavefield residual  $\delta p$  can be accounted as a summation of ten different contributions: one component to the

background wavefield residual and nine components to the singular-wavefield residual from which five depend on background perturbations and four depend on singular-part perturbations. Each contribution is symmetrical to another one or to itself. In our mathematical development of the decomposition we found three non-symmetrical contributions which, latter, turned out to cancel each other.

The importance of each component of the singular-wavefield residual varies from case to case. But, in general, we verified that when the background perturbations are present, the main contributions come from the multiple- scattering term  $\delta p_{S,BBS}$  and  $\delta p_{S,SB B}$ . The single-scattering term  $\delta p_{S,BBB}$  carries information, though not only, on the background-wavefield residual already accounted for by the term  $\delta p_B$ . This information is cancelled by the component  $\delta p_{S,-B L B}$  and the remaining information is carried by the superposition term  $\delta p_{BWB}$ . This last term is relevant only when the background perturbation overlaps the singular component of the model within the Fresnel zone of the total wavefield. When it does, it usually has the same order of magnitude than the strong multiple-scattering term  $\delta p_{S,SB S}$ . When singular-part perturbations are present, there is not a most dominant component in term of energy (as a matter of fact the strong multiple-scattering component carries a little less energy than the others as it might be expected). But the single-scattering component is the single most representative in term of waveform.

The backprojection numerical experiments helped us to understand the role and importance of including an estimate of the singular part of the model in the starting model. The smooth starting model that are usually used in the conventional approach leads to estimates that do not incorporate deep- subsurface information, or at least, not with the desired long wavelength. the shallow part of the subsurface is illuminated by diving waves: the broader the acquisition offset is, the deeper the subsurface is illuminated. The only information from deep region comes from the short-wavelength migration-related contribution. When singular estimates are included in the starting model the back-illumination provided by the scattered/reflected waves allows the estimate to see a broader, deeper area with a larger wavelength. And the estimates evaluated using the subkernels from the decomposition allow us to individualise the most important contribution for a particular problem.

When there are perturbations only in the singular part of the model the summation of the estimates obtained with the adjoint of subkernels correspond to the estimate obtained with the adjoint of the undecomposed kernel. This is not true when there are background perturbations involved. This is so because we are dealing with adjoint and not with inverse operations. The estimate from the undecomposed kernel exhibits cross-terms that are not present in the background

---

perturbation estimate obtained with our formulation.

The backprojection numerical experiments have shown the feasibility of our main claim: the decomposition into subkernels allows to backproject only the singular-wavefield residuals so as to obtain reasonable background-model perturbation estimates. In our numerical experiments with background perturbation, we obtained the single best estimates with multiple-scattering subkernels. This was particularly important in an experiment with a restricted acquisition geometry (reflection data, narrow offset). In this case, the multiple-scattering subkernels take advantage of medium self-illumination provided by the singular wavefields. These results are in clear connection with the ones of the forward problem experiments. In those with background perturbations, the multiple-scattering based terms give the most relevant contribution to the evaluated singular-wavefield residual. This points towards the possibility that FWI can be successfully applied to short-offset data, if a reasonable estimate of the singular part of the model is available and included in the reference model.

On the other hand, we observed that multiple-scattering based subkernels do a poorer job than their undecomposed counterpart when used to update the singular part of the model alone. If the assumptions of undecomposed FWI are satisfied (known smooth reference model, unknown small scatterers), then it is superior to decomposed inversion.

An immediate application of the proposed decomposition of the reference model is as a follow-up of standard model building in connection with imaging. Classically, the initial reference model used in the FWI process is the best, yet smooth, velocity model obtained from some sort of migration velocity analysis and tomography. In the decomposed inversion, this model could play the role of the smooth background part of the initial reference model. However, in conventional FWI, the information provided from the migrated image with this smooth model is not made use of. The positions of key reflectors in this image (possibly with reflectivity estimates from true-amplitude migration) could be taken as the singular part of the initial reference model. Under that framework, our formalism provides an explicit method for jointly using velocity models and migrated seismic images in FWI as well as for understanding their interplay in the model-building process. The analysis we made here indicates that such a procedure should have better convergence properties than the conventional approach.

Another possibility that arises with the discussed decomposition regards the use of selected kernel contributions instead of their complete sum. Synthetic data studies may help to decide which subkernels carry the most significant information for a given data set, so that contributions that carry no or even misleading information can be eliminated from the inversion.

It is important to observe that the initial model decomposition based on the spatial-frequency content done here is not, by all means, the only possible choice. Being completely arbitrary as long as the initial model is fully known, the decomposition allows for a problem-based decision on which features will be part of the background model and which will be considered by means of the singular part. Different decompositions lead to different interpretation of each of the subkernels. This implies a broad range of possible applications of the proposed approach, e.g., frequency-band or layer-stripping-based decompositions.

As an example, we have applied the scattering decomposition of the sensitivity kernels to the time-lapse problem, considering the time-lapse change as a perturbation of the singular part of the model, i.e., perturbation of the scattering potential. Under the differential-method framework, we have demonstrated that the scattering-based decomposition of the sensitivity kernels allows to take advantage of the illumination of the time-lapse change due to multiple-scattering phenomena in order to improve the perturbation estimates from FWI.

We employed the FWI differential method. By considering the baseline model fully known, the data difference between baseline and monitor surveys is the residual to be minimized in order to obtain time-lapse change estimates using the baseline as starting model. Considering the time-lapse change as a perturbation of only the singular part of the model, we circumvented the problem of separating the reference wavefield from the scattered field.

Our numerical experiments demonstrated that it is beneficial to include the singular part into the model used for the necessary back propagations. In the presence of a strong reflector below the time-lapse change, the use of the multiple-scattering-based sensitivity kernels yielded better estimates in both large- and narrow-offset surveys compared to estimates obtained with the conventional sensitivity kernel. In the large-offset case, it better handled the backscattered information coming from the deep reflector which would, otherwise, give rise to spurious estimates. In the narrow-offset case, the scattering-based sensitivity kernels help to bring out the low-frequency spatial information carried by the back-scattered wavefield coming up from the deep reflector. In other words, in the subkernels (e.g.,  $\overline{U}_{S,SB}$ ) where the scattered wavefield acts as source-side wavefield, the singularities or reflectors that originate the scattered wavefield act as sources of illumination, enhancing and broadening the spatial spectral content of the estimate.

### 5.1 Thesis' follow-ups

This thesis helped to understand the role of scattering information within one iteration. A natural next step is to test the effectiveness of the proposed decomposition is a full inversion process. Obviously, it is not practical to use all the subkernels at each iteration. The cost of evaluating each individual contribution is prohibitive, and would overcome any possible gain in the convergence velocity. The idea is to choose one of the scattering-based subkernels and use it instead of the single-scattering conventional subkernel. The choice would be case-dependent but the illumination study done in every survey design could also be used to help the decision of with subkernel would be more suitable for each case.

One case that can benefit from the multiple-scattering subkernels are the complex velocity models with strong shadow areas. In this situation the back-illumination provided by such subkernels could help to individualize the updates in the shadow areas coming from the scattered waves in the reference model that would be otherwise blinded out by the strongest contributions coming from the single-scattering based subkernels.

Another path that can be undertaken is to combine the single-scattering conventional approach with the proposed multiple-scattering decomposition. Ideally, the first iterations should update the long-wavelength component and gradually including higher and higher wavelength information as the process goes on. But as we know, in the conventional approach, the iterations apart from the first generates velocity models components with potential of generating scattered waves. With the use of multiple-scattering subkernels, this drawback can be used in our favour. The first update, may be seen as the singular part for the next iteration. In the second iteration we can use the one of the subkernels proposed in our decomposition to use the scattered waves generated by this components just included in the reference model to update the background model. In the third iteration with the new background component added we can use the conventional kernel to update the singular part. Alternating the updates of the singular part of the model with the ones of background component, we believe better results could be achieved.

The formalism we presented here was based a a twofold decomposition of the reference model. As seen, it is also possible to decompose the model into several parts and consider perturbations to all model and corresponding wavefield contributions. Such a generalisation to a perturbation of all terms to an  $N$ -fold decomposition is conceptually straightforward and should be more explored. The threefold decomposition, as next logical step, could be used to explore the relations between the intermediate wavelength spectral content of the model, to which the seismic waves are rather

insensitive, with the low and high wavelength components of the model.

This study was based in the formalism developed in Tarantola (1984b). We decomposed the kernel resulting from the minimization of a  $L^2$ -norm objective function of the full wavefield. Another way perform the decomposition is to design objective functions with residuals, regularisation and pre-conditioning terms that take into account the reference wavefield decomposition from the start. With these new objective functions the decomposition can be extended the the Hessian kernels that will open the perspective of a resolution analysis and inversion that take into consideration the multiple-scattering information.

Finally, with the recent developments on the acquisition technology, now we can access high quality vector-acoustic data. Extending the decomposition to the vector-acoustic formalism will allow us to take advantage on the natural decoupling of the influence of bulk-modulus and density perturbations have on the pressure and velocity wavefield residuals.

# References

- Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., and Virieux, J. (2012). Time-lapse imaging using regularized FWI: a robustness study. *SEG Technical Program Expanded Abstracts*, pages 1–5.
- Asnaashari, A., Brossier, R., Garambois, S., Ludovic, M., Operto, S., Ribodetti, A., and Virieux, J. (2013). Full waveform inversion : challenges. *SEG FWI workshop*.
- Bae, H. S., Pyun, S., Chung, W., Kang, S.-G., and Shin, C. (2012). Frequency-domain acoustic-elastic coupled waveform inversion using the Gauss-Newton conjugate gradient method. *Geophysical Prospecting*, 60(3):413–432.
- Barnes, C. and Charara, M. (2009). The domain of applicability of acoustic full-waveform inversion for marine seismic data. *Geophysics*, 74(6):WCC91.
- Bednar, J. B., Shin, C., and Pyun, S. (2007). Comparison of waveform inversion, part 2: phase approach. *Geophysical Prospecting*, 55(4):465–475.
- Ben-Hadj-Ali, H., Operto, S., and Virieux, J. (2008). Velocity model building by 3D frequency-domain, full-waveform inversion of wide-aperture seismic data. *Geophysics*, 73(5):VE101–VE117.
- Bleistein, N., Cohen, J. K., and Stockwel, J. W. J. (2001). *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*. Springer Verlag, New York.
- Brenders, A. J. and Pratt, R. G. (2007). Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data. *Geophysical Journal International*, 168(1):152–170.



## REFERENCES

---

- Brossier, R. and Operto, S. (2009). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6).
- Bunks, C., Saleck, F., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457.
- Choi, Y., Min, D.-J., and Shin, C. (2008). Two-dimensional waveform inversion of multi-component data in acoustic-elastic coupled media. *Geophysical Prospecting*, 56(6):863–881.
- Clement, F., Chavent, G., and Gomez, S. (2001). Migration-based traveltime waveform inversion of 2-D simple structures: A synthetic example. *Geophysics*, 66(3):845.
- Cruse, E., Pica, A., Noble, M., McDonald, J. A., and Tarantola, A. (1990). Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics*, 55(5):527–538.
- Fichtner, A. (2010). *Full Seismic Waveform Modelling and Inversion*. Springer Verlag.
- Fichtner, A. and Trampert, J. (2011). Hessian kernels of seismic data functionals based upon adjoint techniques. *Geophysical Journal International*, 185(2):775–798.
- Fletcher, R. and Reeves, C. (1964). Function minimization by conjugate gradients. *The Computer Journal*, 7(2):149–154.
- Gauthier, O., Virieux, J., and Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms: numerical results. *Geophysics*, 51(7):1387–1403.
- Guitton, A. and Symes, W. W. (2003). Robust inversion of seismic data using the Huber norm. *Geophysics*, 68(4):1310–1319.
- Ha, T., Chung, W., and Shin, C. (2009). Waveform inversion using a back-propagation algorithm and a Huber function norm. *Geophysics*, 74(3):R15–R24.
- Hestenes, M. R. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409 – 436.
- Hole, J. A. (1992). Nonlinear high-resolution three-dimensional seismic travel time tomography. *Journal of Geophysical Research: Solid Earth*, 97(B5):6553–6562.

## REFERENCES

---

- Jannane, M., Beydount, W., Crase, E., Cao, D., Koren, Z., Landa, E., Mendes, M., Pica, A., Noble, M., Roeth, G., Singh, S. C., Snieder, R., Tarantola, A., Trezeguet, D., and Xie, M. (1989). Wavelengths of earth structures that can be resolved from seismic reflection data. *Geophysics*, 54(7):906.
- Jeong, W., Lee, H.-Y., and Min, D.-J. (2012). Full waveform inversion strategy for density in the frequency domain. *Geophysical Journal International*, 188(3):1221–1242.
- Kamath, N. and Tsvankin, I. (2013). Full-waveform inversion of multicomponent data for horizontally layered VTI media. *Geophysics*, 78(5):WC113–WC121.
- Lailly, P. (1983). The seismic inverse problem as a sequence of before stack migrations. In *Inverse Scattering Theory and Application*, pages 206–220. Society for Industrial and Applied Mathematics (SIAM).
- Lambaré, G. (2008). Stereotomography. *Geophysics*, 73(5).
- Luemberger, D. G. and Ye, Y. (2010). *Linear and nonlinear programming*. International Series in Operations Research and Management Science. Springer, New York, third edit edition.
- Lumley, D. E. (2001). Time-lapse seismic reservoir monitoring. *Geophysics*, 66(1):50–53.
- Marfurt, K. J. (1984). Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics*, 49(5):533–549.
- Mora, P. (1987). Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, 52(9):1211–1228.
- Mora, P. (1988). Elastic wave-field inversion of reflection and transmission data. *Geophysics*, 53(6):750–759.
- Newton, R. G. (1982). *Scattering Theory of Waves and Particles*. Springer, New York, Heidelberg.
- Plessix, R.-É. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.
- Plessix, R.-É. (2007). A Helmholtz iterative solver for 3D seismic-imaging problems. *Geophysics*, 72(5):SM185–SM194.

## REFERENCES

---

- Plessix, R.-É. (2009). Three-dimensional frequency-domain full-waveform inversion with an iterative solver. *Geophysics*, 74(6).
- Plessix, R.-É. (2012). Waveform Inversion Overview: Where Are We? And What Are the Challenges? *74th EAGE Conference & Exhibition-Workshops*, (June 2012):4–7.
- Plessix, R.-É., Michelet, S., Rynja, H., Kuehl, H., de Maag, J. W., and Hatchell, P. (2010). Some 3D applications of full waveform inversion. *72nd EAGE Conference Extended Abstracts*.
- Polak, G. and Ribiere, E. (1969). Note sur la convergence de méthodes de directions conjuguées. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 3(R1):35–43.
- 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., and Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133:341–362.
- Pratt, R. G. and Shipp, R. M. (1999). Seismic waveform inversion in the frequency domain , Part 2 : Fault delineation in sediments using crosshole data. *Geophysics*, 64(3):901–913.
- Pratt, R. G., Song, Z.-M., Williamson, P., and Warner, M. (1996). Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International*, 124(2):323–340.
- Pratt, R. G. and Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part1: acoustic wave-equation method. *Geophysical Prospecting*, 38(July 1989):287–310.
- Pyun, S., Shin, C., and Bednar, J. B. (2007). Comparison of waveform inversion, part 3: amplitude approach. *Geophysical Prospecting*, 55(4):477–485.
- Queißer, M. and Singh, S. C. (2013). Full waveform inversion in the time lapse mode applied to CO<sub>2</sub> storage at Sleipner. *Geophysical Prospecting*, 61(3):537–555.
- Rickett, J. E. and Sava, P. (2002). Offset and angle-domain common image-point gathers for shot-profile migration. *Geophysics*, 67(3):883.

## REFERENCES

---

- Rodberg, L. S. and Thaler, R. M. (1967). *Introduction to the Quantum Theory of Scattering*. Academic Press, New York.
- Sava, P. and Biondi, B. (2004). Wave-equation migration velocity analysis. I. Theory. *Geophysical Prospecting*, (January):593–606.
- Sava, P. and Fomel, S. (2003). Angle-domain common-image gathers by wavefield continuation methods. *Geophysics*, 68(3):1065.
- Sava, P. and Vasconcelos, I. (2009). Extended common-image-point gathers for wave-equation migration. *71st EAGE Conference & Exhibition*, (June 2009):8–11.
- Sava, P. and Vasconcelos, I. (2011). Extended imaging conditions for wave-equation migration. *Geophysical Prospecting*, 59(1):35–55.
- Sears, T. J., Singh, S. C., and Barton, P. (2008). Elastic full waveform inversion of multi-component OBC seismic data. *Geophysical Prospecting*, 56(6):843–862.
- Shin, C. and Cha, Y. H. (2008). Waveform inversion in the Laplace domain. *Geophysical Journal International*, 173(3):922–931.
- Shin, C. and Ha, W. (2008). A comparison between the behavior of objective functions for waveform inversion in the frequency and Laplace domains. *Geophysics*, 73(5):VE119–VE133.
- Shin, C. and Ho Cha, Y. (2009). Waveform inversion in the Laplace-Fourier domain. *Geophysical Journal International*, 177(3):1067–1079.
- Shin, C. and Min, D.-J. (2006). Waveform inversion using a logarithmic wavefield. *Geophysics*, 71(3):R31–R42.
- Shin, C., Pyun, S., and Bednar, J. B. (2007). Comparison of waveform inversion, part 1: conventional wavefield vs logarithmic wavefield. *Geophysical Prospecting*, 55(4):449–464.
- Shipp, R. M. and Singh, S. C. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, 151(2):325–344.
- Sirgue, L., Etgen, J., and Albertin, U. (2008). 3D frequency domain waveform inversion using time domain finite difference methods. *70th EAGE Conference & Exhibition*, (June 2008):9–12.

## REFERENCES

---

- Sirgue, L. and Pratt, R. G. (2004). Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics*, 69(1):231.
- Snieder, R., Grêt, A., Douma, H., and Scales, J. (2002). Coda wave interferometry for estimating nonlinear behavior in seismic velocity. *Science*, 295:2253.
- Stolt, R. H. and Weglein, A. B. (2012). *Seismic Imaging and Inversion*. Cambridge University Press, Cambridge, UK.
- Symes, W. W. (2008). Migration velocity analysis and waveform inversion. *Geophysical Prospecting*, 56(6):765–790.
- Tarantola, A. (1984a). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.
- Tarantola, A. (1984b). Linearized Inversion of Seismic Reflection Data. *Geophysical prospecting*, 32(6):998–1015.
- Tarantola, A. (1986). A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics*, 51(10):1893–1903.
- Tarantola, A. (1987). Inverse problem theory: Methods for data fitting and model parameter estimation.
- Tromp, J., Tape, C., and Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160(1):195–216.
- Vasconcelos, I., Sava, P., and Douma, H. (2010). Nonlinear extended images via image-domain interferometry. *Geophysics*, 75(6):SA105–SA115.
- Vasconcelos, I., Snieder, R., and Douma, H. (2009). Representation theorems and Green’s function retrieval for scattering in acoustic media. *Physical Review E*, 80(3):1–14.
- Vigh, D., Starr, E. W., and Kapoor, J. (2009). Developing Earth models with full waveform inversion. *The Leading Edge*, pages 432–435.
- Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1.

## REFERENCES

---

- Virieux, J., Operto, S., and Ben-Hadj-Ali, H. (2009). Seismic wave modeling for seismic imaging. *The Leading Edge*, 28:538–544.
- Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Štekl, I., Guasch, L., Win, C., Conroy, G., and Bertrand, A. (2013). Anisotropic 3D full-waveform inversion. *Geophysics*, 78(2):R59–R80.
- Watanabe, T., Shimizu, S., Asakawa, E., and Matsuoka, T. (2004). Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring. *SEG Technical Program Expanded Abstracts*, (October):1–4.
- Woodward, M. J., Nichols, D., Zdraveva, O., Whitfield, P., and Johns, T. (2008). A decade of tomography. *Geophysics*, 73(5):VE5–VE11.
- Wu, R. and Aki, K. (1985). Scattering characteristics of elastic waves by an elastic heterogeneity. *Geophysics*, 50(4):582–595.
- Yang, D., Fehler, M., Malcolm, A., and Huang, L. (2011). Carbon sequestration monitoring with acoustic double-difference waveform inversion: A case study on SACROC walkaway VSP data. *SEG Technical Program Expanded Abstracts*, pages 4273–4277.
- Zheng, Y., Barton, P., and Singh, S. C. (2011). Strategies for elastic full waveform inversion of time-lapse ocean bottom cable (OBC) seismic data. *SEG Technical Program Expanded Abstracts*, pages 4195–4200.
- Zhu, H., Luo, Y., Nissen-Meyer, T., Morency, C., and Tromp, J. (2009). Elastic imaging and time-lapse migration based on adjoint methods. *Geophysics*, 74(6):WCA167.

## REFERENCES

---

# A. Mathematical Definitions

**Fourier Transform.** The fourier transform convention used here is defined as follows.

$$\hat{g}(\omega) = \mathcal{F}[g(t)] = \int_{-\infty}^{\infty} g(t)e^{-i\omega t} dt. \quad (\text{A.1})$$

$$g(t) = \mathcal{F}^{-1}[\hat{g}(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{g}(\omega)e^{i\omega t} d\omega. \quad (\text{A.2})$$

**Scalar and vector quantities.** Bold letters indicate vector or discretized quantities, for instance,  $\mathbf{m}(\mathbf{x})$ . Scalar discretized quantities were denoted by overlined symbols, for instance,  $\overline{\delta K}_B(\mathbf{x}_i)$ . Matrix are denoted by a bold capital letter, for instance,  $\mathbf{M}_1$ .

**Diferential Operators.** Caligrafic letters indicate differential operators, as in  $\mathcal{L}(K(\mathbf{x}'), \rho(\mathbf{x}'))[\cdot]$ . Here the diferential operators are composed by time and spacial derivatives and can be applied onto any wavefield. The spatial derivatives are done with respect to the spacial variable inherited from the model parameters the operator are function of. In this example,  $\mathbf{x}'$ . The time derivatives are done with respect to the time variable inherited from the wavefield the operator are applied onto.

When the derivative is neither with respect to space nor time, the subscript used with operator  $\nabla$  indicates which vector variable the derivative are done with respect to as, for instance, in  $\nabla_{\mathbf{m}}\chi$ .

**Time Convolution.** Time convolution operation is indicated with the symbol '\*' and defined as

$$g * f(t) = \int_{-\infty}^{\infty} g(t - \tau)f(\tau)d\tau \quad (\text{A.3})$$



**Delta function.** The function<sup>1</sup>  $\delta(\mathbf{x} - \mathbf{x}')$  is defined as follows

$$\delta(\mathbf{x} - \mathbf{x}') = 0 \text{ for } \mathbf{x} \neq \mathbf{x}' \quad (\text{A.4})$$

and

$$g(\mathbf{x}') = \int_{\mathbb{R}^3} g(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}') d^3 \mathbf{x}. \quad (\text{A.5})$$

**Adjoint operators, conjugate transpose matrix.** The symbol  $^\dagger$  indicates the adjoint of an operator or the conjugate transpose of a matrix.

---

<sup>1</sup>Actually this is a distribution.

## B. Generalised decomposition and its relation to the proposed decomposition

In this appendix, we generalise to a  $N$ -fold decomposition the one seen in eqs. (2.4) and (2.5) (See Section B.1). In Sections B.2 and B.3, we show that the Expression for the secondary sources in the proposed decomposition presented in this thesis can be derived from applying Born approximation in a fourfold decomposition.

### B.1 Generalised decomposition

The twofold decomposition can be immediately generalised, because eq. (2.4) is of the same type as eq. (2.1) and can be decomposed in the same way into two equations.

Denoting  $p_s$  by  $p_N$  and  $\mathcal{L}_0$  by  $\mathcal{L}_{N-1}$  in eqs. (2.4) and (2.5), we can write

$$\mathcal{L}_{N-1} [p_{0N}(\mathbf{x}, t; \mathbf{x}_s)] = \delta(\mathbf{x} - \mathbf{x}_s) S(t), \quad (\text{B.1})$$

$$\mathcal{L} [p_N(\mathbf{x}, t; \mathbf{x}_s)] = -\mathcal{V}_N [p_{0N}(\mathbf{x}, t; \mathbf{x}_s)], \quad (\text{B.2})$$

where  $\mathcal{V}_N = \mathcal{L} - \mathcal{L}_{N-1}$  and  $p_{0N} = p - p_N$  is the remainder of the wavefield after separation of part  $p_N$ . Decomposing eq. (B.1) again with  $p_s$  denoted by  $p_{N-1}$ ,  $\mathcal{L}$  by  $\mathcal{L}_{N-1}$ , and  $\mathcal{L}_0$  by  $\mathcal{L}_{N-2}$  in eqs. (2.4) and (2.5) leads to

$$\mathcal{L}_{N-2} [p_{0N-1}(\mathbf{x}, t; \mathbf{x}_s)] = \delta(\mathbf{x} - \mathbf{x}_s) S(t), \quad (\text{B.3})$$

$$\mathcal{L}_{N-1} [p_{N-1}(\mathbf{x}, t; \mathbf{x}_s)] = -\mathcal{V}_{N-1} [p_{0N-1}(\mathbf{x}, t; \mathbf{x}_s)], \quad (\text{B.4})$$

where now  $\mathcal{V}_{N-1} = \mathcal{L}_{N-1} - \mathcal{L}_{N-2}$  and  $p_{0N-1} = p_{0N} - p_{N-1} = p - p_N - p_{N-1}$ . In the same way,

eq. (B.3) can be decomposed again to yield the system

$$\mathcal{L}_{N-3} [p_{0N-2}(\mathbf{x}, t; \mathbf{x}_s)] = \delta(\mathbf{x} - \mathbf{x}_s) S(t), \quad (\text{B.5})$$

$$\mathcal{L}_{N-2} [p_{N-2}(\mathbf{x}, t; \mathbf{x}_s)] = -\mathcal{V}_{N-2} [p_{0N-2}(\mathbf{x}, t; \mathbf{x}_s)], \quad (\text{B.6})$$

where  $\mathcal{V}_{N-2} = \mathcal{L}_{N-2} - \mathcal{L}_{N-3}$  and  $p_{0N-2} = p_{0N-1} - p_{N-2} = p - p_N - p_{N-1} - p_{N-2}$ . It is not difficult to observe that this decomposition principle can be applied  $N$  times to finally yield the following system of  $N + 1$  equations,

$$\mathcal{L}_0 [p_0(\mathbf{x}, t; \mathbf{x}_s)] = \delta(\mathbf{x} - \mathbf{x}_s) S(t), \quad (\text{B.7})$$

$$\mathcal{L}_j [p_j(\mathbf{x}, t; \mathbf{x}_s)] = -\mathcal{V}_j \left[ \sum_{i=0}^{j-1} p_i(\mathbf{x}, t; \mathbf{x}_s) \right] \quad (j = 1, \dots, N), \quad (\text{B.8})$$

where  $p = \sum_{i=0}^N p_i$ . The wave operators  $\mathcal{L}_j$  are given by

$$\mathcal{L}_j = \frac{1}{\sum_{i=0}^j K_i(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\sum_{i=0}^j \rho_i(\mathbf{x})} \nabla \right), \quad (\text{B.9})$$

with  $\mathcal{L}_N = \mathcal{L}$ . Moreover, the scattering potentials are

$$\mathcal{V}_j = \mathcal{L}_j - \mathcal{L}_{j-1}; \quad (\text{B.10})$$

and  $K = \sum_{i=0}^N K_i$  and  $\rho = \sum_{i=0}^N \rho_i$ .

Being completely arbitrary, such  $N + 1$ -fold decomposition could be used, for instance, to describe a complex geological model one layer at a time or to decompose it into contributions from several frequency bands.

## **B.2 Four-fold decomposition of a perturbed model and wave-field**

From Section ?? we have the  $N$ -fold decomposition represented by:

$$\mathcal{L}_0 [p_0(\mathbf{x}, t; \mathbf{x}_s)] = \delta(\mathbf{x} - \mathbf{x}_s) S(t), \quad (\text{B.11})$$

## B.2 Four-fold decomposition of a perturbed model and wavefield

---

$$\mathcal{L}_j [p_j(\mathbf{x}, t; \mathbf{x}_s)] = -\mathcal{V}_j \left[ \sum_{i=0}^{j-1} p_i(\mathbf{x}, t; \mathbf{x}_s) \right] \quad (j = 1, \dots, N). \quad (\text{B.12})$$

where

$$K = \sum_{i=0}^N K_i, \quad (\text{B.13})$$

$$\rho = \sum_{i=0}^N \rho_i, \quad (\text{B.14})$$

$$p = \sum_{i=0}^N p_i, \quad (\text{B.15})$$

$$\mathcal{L}_j = \frac{1}{\sum_{i=0}^j K_i(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\sum_{i=0}^j \rho_i(\mathbf{x})} \nabla \right), \quad (\text{B.16})$$

$$\mathcal{V}_j = \mathcal{L}_j - \mathcal{L}_{j-1} \quad (\text{B.17})$$

Note that from (B.17) we can write that

$$\mathcal{L}_j = \mathcal{V}_j + \mathcal{V}_{j-1} + \dots + \mathcal{V}_1 + \mathcal{L}_0. \quad (\text{B.18})$$

Suppose we want to decompose a perturbed model  $K$  (we are considering perturbation only in bulk modulus) in four parts superimposed in such order: at the “bottom” a background model  $K_B$ ; over it the background perturbation  $\delta K_B$ ; on top of it, a singular model  $K_S$ ; and on the top most the corresponding perturbation  $\delta K_S$ , such that

$$K = K_B + \delta K_B + K_S + \delta K_S. \quad (\text{B.19})$$

Comparing eqs. (B.19) and (B.13) we can consider  $K_0 \equiv K_B$ ,  $K_1 \equiv \delta K_B$ ,  $K_2 \equiv K_S$ , and  $K_3 \equiv \delta K_S$  and perform a fourfold decomposition of  $K$ . Setting  $\rho_0 \equiv \rho$  and  $\rho_i \equiv 0$  for  $i = 1, 2, 3$ , we have (dependency of the wavefields on  $(\mathbf{x}, t; \mathbf{x}_s)$  is omitted for simplicity)

$$\mathcal{L}_B [p_B] = \delta(\mathbf{x} - \mathbf{x}_s) S(t), \quad (\text{B.20})$$

$$(\mathcal{L}_B + \mathcal{V}_B) [\delta p'_B] = -\mathcal{V}_B [p_B], \quad (\text{B.21})$$

$$(\mathcal{L}_B + \mathcal{V}_B + \mathcal{V}'_0) [p_S + \delta p'_{S,B}] = -\mathcal{V}'_0 [p_B + \delta p'_B], \quad (\text{B.22})$$

$$(\mathcal{L}_B + \mathcal{V}_B + \mathcal{V}'_0 + \mathcal{V}_S) [\delta p'_{S,S}] = -\mathcal{V}_S [p_B + \delta p'_B + p_S + \delta p'_{S,B}], \quad (\text{B.23})$$

with (omitting the  $\mathbf{x}$ -dependency for simplicity)

$$\mathcal{L}_B = \frac{1}{K_B} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\rho} \nabla \right), \quad (\text{B.24})$$

$$\mathcal{V}_B = \left( \frac{1}{K_B + \delta K_B} - \frac{1}{K_B} \right) \frac{\partial^2}{\partial t^2}, \quad (\text{B.25})$$

$$\mathcal{V}'_0 = \left( \frac{1}{K_B + \delta K_B + K_S} - \frac{1}{K_B + \delta K_B} \right) \frac{\partial^2}{\partial t^2}, \quad (\text{B.26})$$

$$\mathcal{V}_S = \left( \frac{1}{K_B + \delta K_B + K_S + \delta K_S} - \frac{1}{K_B + \delta K_B + K_S} \right) \frac{\partial^2}{\partial t^2}, \quad (\text{B.27})$$

which imply

$$\mathcal{L}_B + \mathcal{V}_B = \frac{1}{K_B + \delta K_B} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\rho} \nabla \right), \quad (\text{B.28})$$

$$\mathcal{L}_B + \mathcal{V}_B + \mathcal{V}'_0 = \frac{1}{K_B + \delta K_B + K_S} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\rho} \nabla \right), \quad (\text{B.29})$$

$$\mathcal{L}_B + \mathcal{V}_B + \mathcal{V}'_0 + \mathcal{V}_S = \frac{1}{K_B + \delta K_B + K_S + \delta K_S} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\rho} \nabla \right), \quad (\text{B.30})$$

It is important to note that no approximation has been made at all, i.e.,  $K_B$ ,  $\delta K_B$ ,  $K_S$ , and  $\delta K_S$  can hold any magnitude relation among them as long they sum up to  $K$ . The prime symbol ' indicates the non-linearised version of the corresponding wavefields and/or operators.

The Physical interpretation of eqs. (B.20) to (B.23) is as plain as the twofold decomposition:

**Eq. (B.20)** A background wavefield  $p_B$  propagates in a background medium  $K_B$  excited by a point-source at  $\mathbf{x}_s$ .

**Eq. (B.21)** A perturbation  $\delta K_B$  in the background medium is introduced; when the background wavefield  $p_B$  runs into the background perturbation it excites a wavefield perturbation  $\delta p'_B$  (a wavefield itself) that propagates in the perturbed background medium  $K_B + \delta K_B$ .

**Eq. (B.22)** A singular component of the reference medium<sup>1</sup>  $K_S$  is introduced in the perturbed

---

<sup>1</sup>The unperturbed reference medium  $K_0$  is defined as  $K_B + K_S$ .

### B.3 Linearisation of the fourfold decomposition

---

background medium; when the perturbed background wavefield  $p_B + \delta p_B$  runs into the singular component it excites a background-perturbed singular wavefield  $p_S + \delta p'_{S,B}$  that propagates in the background-perturbed reference medium  $K_B + \delta K_B + K_S$ .

**Eq. (B.23)** A perturbation  $\delta K_S$  on the singular component of the background-perturbed reference medium is introduced; when the background-perturbed reference wavefield<sup>2</sup>  $p_B + \delta p_B + p_S + \delta p'_{S,B}$  runs into the singular-component perturbation it excites a singular-wavefield perturbation  $\delta p'_{S,S}$  that propagates in the perturbed reference medium  $K_B + \delta K_B + K_S + \delta K_S$ .

### B.3 Linearisation of the fourfold decomposition

Now let us consider that  $\delta K_B, \delta K_S \ll K_B, K_S$ . This allows us to use linear approximations in the set of eqs. (B.20) to (B.23).

From (B.25) we can write

$$\mathcal{V}_B \approx - \left( \frac{\delta K_B}{K_B^2} \right) \frac{\partial^2}{\partial t^2} = \delta \mathcal{L}_B, \quad (\text{B.31})$$

also seen at eq. (2.24). From eq. (B.26) we have

$$\begin{aligned} \mathcal{V}'_0 &\approx \left( \frac{1}{K_B + K_S} - \frac{\delta K_B}{(K_B + K_S)^2} - \frac{1}{K_B} + \frac{\delta K_B}{K_B^2} \right) \frac{\partial^2}{\partial t^2}, \\ &\approx \left( \frac{1}{K_B + K_S} - \frac{1}{K_B} \right) \frac{\partial^2}{\partial t^2} - \delta K_B \left( \frac{1}{(K_B + K_S)^2} - \frac{1}{K_B^2} \right) \frac{\partial^2}{\partial t^2}, \\ &\approx \left( \frac{1}{K_B + K_S} - \frac{1}{K_B} \right) \left[ 1 - \delta K_B \left( \frac{1}{K_B + K_S} + \frac{1}{K_B} \right) \right] \frac{\partial^2}{\partial t^2}, \\ &\approx \mathcal{V}_0 \left[ 1 - \delta K_B \left( \frac{1}{K_B + K_S} + \frac{1}{K_B} \right) \right], \\ &\approx \mathcal{V}_0 - \delta K_B \mathcal{W} \end{aligned} \quad (\text{B.32})$$

where we have used eq. (D.4) with the superscript  $K$  dropped. Finally, from (B.27) we can write

$$\mathcal{V}_S \approx \left( \frac{1}{K_B + K_S} - \frac{\delta K_B + \delta K_S}{(K_B + K_S)^2} - \frac{1}{K_B + K_S} + \frac{\delta K_B}{(K_B + K_S)^2} \right) \frac{\partial^2}{\partial t^2},$$

---

<sup>2</sup>The unperturbed reference wavefield  $p_0$  is defined as  $p_B + p_S$ .

$$\mathcal{V}_S \approx - \left( \frac{\delta K_S}{(K_B + K_S)^2} \right) \frac{\partial^2}{\partial t^2} = \delta \mathcal{S} \quad (\text{B.33})$$

Using eqs. (B.31), (B.32) and (B.33) we can restate eqs. (B.21) to (B.23) as

$$(\mathcal{L}_B + \delta \mathcal{L}_B) [\delta p_B] = -\delta \mathcal{L}_B [p_B], \quad (\text{B.34})$$

$$(\mathcal{L}_B + \delta \mathcal{L}_B + \mathcal{V}_0 - \delta K_B \mathcal{W}) [p_S + \delta p_{S,B}] = -(\mathcal{V}_0 - \delta K_B \mathcal{W}) [p_B + \delta p_B], \quad (\text{B.35})$$

$$(\mathcal{L}_B + \delta \mathcal{L}_B + \mathcal{V}_0 - \delta K_B \mathcal{W} + \delta \mathcal{S}) [\delta p_{S,S}] = -\delta \mathcal{S} [p_B + \delta p_B + p_S + \delta p_{S,B}]. \quad (\text{B.36})$$

Note we dropped the primes on the wavefields denoting that these are their linearised versions.

Hereafter we are going to neglect all second order terms denoted by a perturbation operator, such as  $\delta \mathcal{L}_B$ , applied onto a wavefield perturbation, such as  $\delta p_B$ . Doing so eq. (B.34) becomes

$$\mathcal{L}_B [\delta p_B] = -\delta \mathcal{L}_B [p_B]. \quad (\text{B.37})$$

This equation describes the propagation of the (linearised) background residual. Correspondingly, considering  $\mathcal{L}_B + \mathcal{V}_0 = \mathcal{L}_0$ , eq. (B.35), on its turn, becomes

$$\mathcal{L}_0 [p_S] + \mathcal{L}_0 [\delta p_{S,B}] + \delta \mathcal{L}_B [p_S] - \delta K_B \mathcal{W} [p_S] = -\mathcal{V}_0 [p_B] - \mathcal{V}_0 [\delta p_B] + \delta K_B \mathcal{W} [p_B]. \quad (\text{B.38})$$

In a two-fold decomposition where there were no model perturbations, i.e. ,  $\delta K_B, \delta K_S = 0$  but only the two components of the reference medium such that  $K_0 = K_B + K_S$  the equation that describes the propagation of the singular wavefield  $p_S$  is given by eq. (2.21):

$$\mathcal{L}_0 [p_S] = -\mathcal{V}_0 [p_B]. \quad (\text{B.39})$$

As a Born approximation we can consider the perturbed wavefield as a linear superposition of the unperturbed singular wavefield  $p_S$  and its residuals. This allows us to subtract eq. (B.39) from eq. (B.38) which yields

$$\mathcal{L}_0 [\delta p_{S,B}] = -\delta \mathcal{L}_B [p_S] + \delta K_B \mathcal{W} [p_S] - \mathcal{V}_0 [\delta p_B] + \delta K_B \mathcal{W} [p_B]. \quad (\text{B.40})$$

Finally, eq. (B.36) can be written as

$$\mathcal{L}_0 [\delta p_{S,S}] = -\delta \mathcal{S} [p_B] - \delta \mathcal{S} [p_S] \quad (\text{B.41})$$

### B.3 Linearisation of the fourfold decomposition

---

We know that

$$\begin{aligned}\delta K_B \mathcal{W} - \delta \mathcal{L}_B &= \delta K_B \left[ \left( \frac{1}{(K_B + K_S)^2} - \frac{1}{(K_B)^2} \right) + \frac{1}{K_B^2} \right] \frac{\partial^2}{\partial t^2} \\ &= \left( \frac{\delta K_B}{(K_B + K_S)^2} \right) \frac{\partial^2}{\partial t^2} = -\delta \mathcal{B}\end{aligned}\tag{B.42}$$

according to eq. (2.40). Using this in eq. (B.40) leads us to

$$\mathcal{L}_0 [\delta p_{S,B}] = -\delta \mathcal{B} [p_S] - \mathcal{V}_0 [\delta p_B] + \delta K_B \mathcal{W} [p_B].\tag{B.43}$$

From (B.42) we have  $\delta K_B \mathcal{W} = \delta \mathcal{L}_B - \delta \mathcal{B}$ . So, the previous equation becomes

$$\mathcal{L}_0 [\delta p_{S,B}] = -\mathcal{V}_0 [\delta p_B] - \delta \mathcal{B} [p_B] - \delta \mathcal{B} [p_S] + \delta \mathcal{L}_B [p_B].\tag{B.44}$$

Summing up eqs. (B.41) and (B.44) we arrive at

$$\mathcal{L}_0 [\delta p_{S,B} + \delta p_{S,S}] = -\mathcal{V}_0 [\delta p_B] - (\delta \mathcal{B} + \delta \mathcal{S}) [p_B] - (\delta \mathcal{B} + \delta \mathcal{S}) [p_S] + \delta \mathcal{L}_B [p_B].\tag{B.45}$$

upon doing that  $\delta p_{S,B} + \delta p_{S,S} = \delta p_S$  and  $\delta \mathcal{B} + \delta \mathcal{S} = \delta \mathcal{L}$ .

In summary eq. (B.37) correspond to (2.23) and describes the linearised background-wavefield residual propagation. Eq. (B.45) correspond to eq. (??) and describes the linearised singular-wavefield residual propagation.





## C. Estimates for the Model Perturbations

In this appendix, we derive the expressions for some of the contributions to the model perturbation estimates based on the adjoint subkernel seen in equation 2.50. For each one of the model parameters  $K_B$  and  $\rho_B$ , we choose one estimate based on single-backpropagation ( $\delta K_{B,BS}^{\text{est}}$  and  $\delta \rho_{B,BS}^{\text{est}}$ ) and one based on a double-backpropagation ( $\delta K_{B,b\nabla B}^{\text{est}}$  and  $\delta \rho_{B,b\nabla B}^{\text{est}}$ ).

To arrive at the expression for the other estimate contributions, we have only to perform the proper substitution of the source- and receiver-side wavefields and the potential operator. In all derivations below, the dependency on  $\omega$  is omitted for simplicity, since all quantities in the frequency domain are easily identified by a hat over the symbol.

### C.1 Estimate of the background perturbation $\delta K_{B,BS}^{\text{est}}$

In the frequency domain, the contribution due to perturbation on the background model of the term (2.33.3) is

$$\widehat{\delta p_{S,BBS}}(\mathbf{x}_g; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' \widehat{G}_S(\mathbf{x}'; \mathbf{x}_g) \delta \mathcal{B} [\widehat{p}_B(\mathbf{x}'; \mathbf{x}_s)]. \quad (\text{C.1})$$

Considering only bulk-modulus perturbations, we have

$$\widehat{\delta p_{S,BBS}^K}(\mathbf{x}_g; \mathbf{x}_s) = \int_{\mathbb{V}} d^3 \mathbf{x}' \frac{-\omega^2}{K_0^2(\mathbf{x}')} \widehat{G}_S(\mathbf{x}'; \mathbf{x}_g) \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \delta K_B(\mathbf{x}'), \quad (\text{C.2})$$

which is equation (2.43) in the frequency domain. Discretising the volume integral in this equation

yields

$$\widehat{\delta p}_{S,BBS}^K(\mathbf{x}_g; \mathbf{x}_s) = \left[ \underbrace{\frac{-\omega^2}{K_0^2(\mathbf{x}'_1)} \widehat{G}_S(\mathbf{x}'_1; \mathbf{x}_g) \widehat{p}_B(\mathbf{x}'_1; \mathbf{x}_s)}_{\text{kernel of } \overline{U}_{B,BS}} \quad \cdots \right] \begin{bmatrix} \delta K_B(\mathbf{x}'_1) \\ \vdots \\ \delta K_B(\mathbf{x}'_N) \end{bmatrix}, \quad (\text{C.3})$$

where  $\overline{U}_{B,BS}$  is the one of the components of the Fréchet derivative of the singular-wavefield with respect to the medium perturbations  $\delta K_B$  at  $\mathbf{x}'$ .

Therefore, the adjoint to equation (C.3) constitutes, for each pair  $(\mathbf{x}_g, \mathbf{x}_s)$ , the vector

$$\overline{\delta K}_{B,BS}^{\text{est}}(\mathbf{x}_g, \mathbf{x}_s) = \begin{bmatrix} \delta K_{B,BS}^{\text{est}}(\mathbf{x}'_1 | \mathbf{x}_s, \mathbf{x}_g) \\ \vdots \\ \delta K_{B,BS}^{\text{est}}(\mathbf{x}'_N | \mathbf{x}_s, \mathbf{x}_g) \end{bmatrix} \approx \begin{bmatrix} \frac{-\omega^2}{K_0^2(\mathbf{x}'_1)} \widehat{G}_S^*(\mathbf{x}'_1; \mathbf{x}_g) \widehat{p}_B^*(\mathbf{x}'_1; \mathbf{x}_s) \\ \vdots \\ \frac{-\omega^2}{K_0^2(\mathbf{x}'_N)} \widehat{G}_S^*(\mathbf{x}'_N; \mathbf{x}_g) \widehat{p}_B^*(\mathbf{x}'_N; \mathbf{x}_s) \end{bmatrix} \widehat{\delta p}_S(\mathbf{x}_g; \mathbf{x}_s). \quad (\text{C.4})$$

Finally, summation over all sources and receivers and integration over all frequencies yields, at each point  $\mathbf{x}$  in the target area, the overall estimate of the medium perturbation as represented in equation (2.51).

## C.2 Estimate of the background perturbation $\delta \rho_{B,BS}^{\text{est}}$

Correspondingly, we can also derive the expressions for the density subkernel  $\overline{V}_{B,BS}$  and the perturbation estimate  $\delta \rho_{B,BS}^{\text{est}}(\mathbf{x})$ . The equation corresponding to expression (C.1) for density reads

$$\widehat{\delta p}_{S,BBS}^\rho(\mathbf{x}_g; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' \widehat{G}_S(\mathbf{x}'; \mathbf{x}_g) \nabla \cdot \left( \frac{\delta \rho_B(\mathbf{x}')}{\rho_0^2(\mathbf{x}')} \nabla_{\mathbf{x}'} \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \right). \quad (\text{C.5})$$

Upon the use of the identity

$$\alpha(\mathbf{x}) \nabla \cdot \mathbf{v}(\mathbf{x}) = - [\nabla \alpha(\mathbf{x})] \cdot \mathbf{v}(\mathbf{x}) + \nabla \cdot [\alpha(\mathbf{x}) \mathbf{v}(\mathbf{x})], \quad (\text{C.6})$$

with

$$\alpha(\mathbf{x}) = \widehat{G}_S(\mathbf{x}; \mathbf{x}_g) \quad \text{and} \quad \mathbf{v}(\mathbf{x}) = \frac{\delta \rho_B(\mathbf{x})}{\rho_0^2(\mathbf{x})} \nabla \widehat{p}_B(\mathbf{x}; \mathbf{x}_s), \quad (\text{C.7})$$

## C.2 Estimate of the background perturbation $\delta\rho_{B,BS}^{\text{est}}$

---

equation (C.5) can be recast into the form

$$\begin{aligned}\widehat{p}_{S,BBS}^\rho(\mathbf{x}_g; \mathbf{x}_s) &= \int_{\mathbb{V}} d^3\mathbf{x}' \left[ \nabla_{\mathbf{x}'} \widehat{G}_S(\mathbf{x}'; \mathbf{x}_g) \cdot \nabla_{\mathbf{x}'} \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \right] \frac{\delta\rho_B(\mathbf{x}')}{\rho_0^2(\mathbf{x}')} \\ &\quad - \int_{\mathbb{V}} d^3\mathbf{x}' \nabla_{\mathbf{x}'} \cdot \left[ \widehat{G}_S(\mathbf{x}'; \mathbf{x}_g) \nabla_{\mathbf{x}'} \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \frac{\delta\rho_B(\mathbf{x}')}{\rho_0^2(\mathbf{x}')} \right].\end{aligned}\quad (\text{C.8})$$

Using Gauss' Theorem, the last term can be written as an surface integral over the earth. Considering homogeneous boundary conditions this term becomes

$$\begin{aligned}\int_{\mathbb{V}} d^3\mathbf{x}' \nabla_{\mathbf{x}'} \cdot \left[ \widehat{G}_S(\mathbf{x}'; \mathbf{x}_g) \nabla_{\mathbf{x}'} \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \frac{\delta\rho_B(\mathbf{x}')}{\rho_0^2(\mathbf{x}')} \right] &= \\ \oint_{\mathbb{S}} d^2\mathbf{x}' \left[ \widehat{G}_S(\mathbf{x}'; \mathbf{x}_g) \nabla_{\mathbf{x}'} \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \frac{\delta\rho_B(\mathbf{x}')}{\rho_0^2(\mathbf{x}')} \right] \cdot \hat{\mathbf{n}} &= 0,\end{aligned}\quad (\text{C.9})$$

where  $\hat{\mathbf{n}}$  is the unitary normal vector to the surface  $\mathbb{S}$  pointing in the outward direction. Therefore

$$\widehat{p}_{S,BBS}^\rho(\mathbf{x}_g; \mathbf{x}_s) = \int_{\mathbb{V}} d^3\mathbf{x}' \left[ \nabla_{\mathbf{x}'} \widehat{G}_S(\mathbf{x}'; \mathbf{x}_g) \cdot \nabla_{\mathbf{x}'} \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \right] \frac{\delta\rho_B(\mathbf{x}')}{\rho_0^2(\mathbf{x}')}.\quad (\text{C.10})$$

After discretisation of the volume integral, we obtain

$$\widehat{p}_{S,BBS}^\rho(\mathbf{x}_g; \mathbf{x}_s) = \left[ \underbrace{\frac{1}{\rho_0^2(\mathbf{x}'_1)} \nabla_{\mathbf{x}'} \widehat{G}_S(\mathbf{x}'_1; \mathbf{x}_g) \cdot \nabla_{\mathbf{x}'} \widehat{p}_B(\mathbf{x}'_1; \mathbf{x}_s)}_{\text{kernel of } \overline{V}_{B,BS}} \quad \dots \right] \begin{bmatrix} \delta\rho_B(\mathbf{x}'_1) \\ \vdots \\ \delta\rho_B(\mathbf{x}'_N) \end{bmatrix}.\quad (\text{C.11})$$

Thus, the adjoint provides, for each pair  $(\mathbf{x}_g, \mathbf{x}_s)$ , the vector

$$\begin{aligned}\overline{\delta\rho_{B,BS}}^{\text{est}}(\mathbf{x}_g, \mathbf{x}_s) &= \begin{bmatrix} \delta\rho_{B,BS}^{\text{est}}(\mathbf{x}'_1 | \mathbf{x}_s, \mathbf{x}_g) \\ \vdots \\ \delta\rho_{B,BS}^{\text{est}}(\mathbf{x}'_N | \mathbf{x}_s, \mathbf{x}_g) \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\rho_0^2(\mathbf{x}'_1)} \nabla_{\mathbf{x}'} \widehat{G}_S^*(\mathbf{x}'_1; \mathbf{x}_g) \cdot \nabla_{\mathbf{x}'} \widehat{p}_B^*(\mathbf{x}'_1; \mathbf{x}_s) \\ \vdots \\ \frac{1}{\rho_0^2(\mathbf{x}'_N)} \nabla_{\mathbf{x}'} \widehat{G}_S^*(\mathbf{x}'_N; \mathbf{x}_g) \cdot \nabla_{\mathbf{x}'} \widehat{p}_B^*(\mathbf{x}'_N; \mathbf{x}_s) \end{bmatrix} \widehat{p}_S(\mathbf{x}_g; \mathbf{x}_s).\end{aligned}\quad (\text{C.12})$$

Finally, summation over all sources and receivers and integration over all frequencies provides, at each point  $\mathbf{x}$  in the target area, the density-perturbation estimate contribution

$$\delta\rho_{B,BS}^{\text{est}}(\mathbf{x}) = \sum_s \sum_g \int d\omega \frac{1}{\rho_0^2(\mathbf{x})} \overbrace{\nabla \widehat{p}_B^*(\mathbf{x}, \omega; \mathbf{x}_s)}^{\text{source-side wavefield}} \cdot \underbrace{\nabla \widehat{G}_S^*(\mathbf{x}, \omega; \mathbf{x}_g) \widehat{p}_S(\mathbf{x}_g, \omega; \mathbf{x}_s)}_{\text{backpropagation of } \widehat{p}_S}. \quad (\text{C.13})$$

### C.3 Estimate of the background perturbation $\delta K_{B,b\mathcal{V}B}^{\text{est}}$

Since the wave phenomena for  $\widehat{p}_{s,b\mathcal{V}0}$  involve two different scattering mechanisms, the derivation of the expression for the contribution  $\delta K_{B,b\mathcal{V}B}^{\text{est}}$  and the corresponding subkernel  $\overline{U}_{B,b\mathcal{B}B}$  is slightly more complicated. Substituting equation (2.25) in expression (2.33.2), we have

$$\widehat{p}_{S,b\mathcal{V}B}(\mathbf{x}_g; \mathbf{x}_s) = \int_{\mathbb{V}} d^3\mathbf{x}' \widehat{G}_B(\mathbf{x}'; \mathbf{x}_g) \mathcal{V}_0 \left[ \int_{\mathbb{V}} d^3\mathbf{x}'' \widehat{G}_B(\mathbf{x}''; \mathbf{x}') \delta \mathcal{L}_B[\widehat{p}_B(\mathbf{x}''; \mathbf{x}_s)] \right], \quad (\text{C.14})$$

the bulk-modulus contribution of which reads

$$\widehat{p}_{S,b\mathcal{V}B}^K(\mathbf{x}_g; \mathbf{x}_s) = \int_{\mathbb{V}} d^3\mathbf{x}' \widehat{G}_B(\mathbf{x}'; \mathbf{x}_g) \mathcal{V}_0^K \left[ \int_{\mathbb{V}} d^3\mathbf{x}'' \frac{\omega^2}{K_B^2(\mathbf{x}'')} \widehat{G}_B(\mathbf{x}''; \mathbf{x}') \widehat{p}_B(\mathbf{x}''; \mathbf{x}_s) \delta K_B(\mathbf{x}'') \right], \quad (\text{C.15})$$

Here, the bulk-modulus part of the scattering potential is

$$\mathcal{V}_0^K = -\omega^2 \left( \frac{1}{K_0(\mathbf{x})} - \frac{1}{K_B(\mathbf{x})} \right). \quad (\text{C.16})$$

Discretising the bulk-modulus part of equation (2.33.2) yields

$$\widehat{p}_{S,b\mathcal{V}B}^K(\mathbf{x}_g; \mathbf{x}_s) = \underbrace{\left[ \widehat{G}_B(\mathbf{x}'_1; \mathbf{x}_g) \mathcal{V}_0^K \quad \cdots \quad \widehat{G}_B(\mathbf{x}'_N; \mathbf{x}_g) \mathcal{V}_0^K \right]}_{\mathbf{M}_1} \begin{bmatrix} \widehat{p}_B(\mathbf{x}'_1; \mathbf{x}_s) \\ \vdots \\ \widehat{p}_B(\mathbf{x}'_N; \mathbf{x}_s) \end{bmatrix}. \quad (\text{C.17})$$

### C.3 Estimate of the background perturbation $\delta K_{B,b\nu B}^{\text{est}}$

---

Correspondingly, discretisation of equation (C.15) yields

$$\widehat{p}_{S,b\nu B}^K(\mathbf{x}_g; \mathbf{x}_s) = \underbrace{\mathbf{M}_1 \mathbf{M}_2}_{\overline{U}_{B,b\nu B}} \begin{bmatrix} \delta K_B(\mathbf{x}_1'') \\ \vdots \\ \delta K_B(\mathbf{x}_N'') \end{bmatrix}, \quad (\text{C.18})$$

in which

$$\mathbf{M}_2 = \begin{bmatrix} \frac{\omega^2}{K_B^2(\mathbf{x}_1'')} \widehat{G}_B(\mathbf{x}_1''; \mathbf{x}_1') \widehat{p}_B(\mathbf{x}_1''; \mathbf{x}_s) & \cdots & \frac{\omega^2}{K_B^2(\mathbf{x}_N'')} \widehat{G}_B(\mathbf{x}_N''; \mathbf{x}_1') \widehat{p}_B(\mathbf{x}_N''; \mathbf{x}_s) \\ \vdots & \ddots & \vdots \\ \frac{\omega^2}{K_B^2(\mathbf{x}_1'')} \widehat{G}_B(\mathbf{x}_1''; \mathbf{x}_N') \widehat{p}_B(\mathbf{x}_1''; \mathbf{x}_s) & \cdots & \frac{\omega^2}{K_B^2(\mathbf{x}_N'')} \widehat{G}_B(\mathbf{x}_N''; \mathbf{x}_N') \widehat{p}_B(\mathbf{x}_N''; \mathbf{x}_s) \end{bmatrix}. \quad (\text{C.19})$$

From equation (C.18) follows its adjoint vector

$$\overline{\delta K}_{B,b\nu B}^{\text{est}}(\mathbf{x}_g, \mathbf{x}_s) = \begin{bmatrix} \delta K_{B,b\nu B}^{\text{est}}(\mathbf{x}_1'' | \mathbf{x}_s, \mathbf{x}_g) \\ \vdots \\ \delta K_{B,b\nu B}^{\text{est}}(\mathbf{x}_N'' | \mathbf{x}_s, \mathbf{x}_g) \end{bmatrix} = \mathbf{M}_2^\dagger \mathbf{M}_1^\dagger \widehat{p}_S(\mathbf{x}_g; \mathbf{x}_s), \quad (\text{C.20})$$

in which

$$\mathbf{M}_2^\dagger = \begin{bmatrix} \frac{\omega^2}{K_B^2(\mathbf{x}_1'')} \widehat{G}_B^*(\mathbf{x}_1''; \mathbf{x}_1') \widehat{p}_B^*(\mathbf{x}_1''; \mathbf{x}_s) & \cdots & \frac{\omega^2}{K_B^2(\mathbf{x}_1'')} \widehat{G}_B^*(\mathbf{x}_1''; \mathbf{x}_N') \widehat{p}_B^*(\mathbf{x}_1''; \mathbf{x}_s) \\ \vdots & \ddots & \vdots \\ \frac{\omega^2}{K_B^2(\mathbf{x}_N'')} \widehat{G}_B^*(\mathbf{x}_N''; \mathbf{x}_1') \widehat{p}_B^*(\mathbf{x}_N''; \mathbf{x}_s) & \cdots & \frac{\omega^2}{K_B^2(\mathbf{x}_N'')} \widehat{G}_B^*(\mathbf{x}_N''; \mathbf{x}_N') \widehat{p}_B^*(\mathbf{x}_N''; \mathbf{x}_s) \end{bmatrix}. \quad (\text{C.21})$$

and

$$\mathbf{M}_1^\dagger = \begin{bmatrix} \widehat{G}_B^*(\mathbf{x}_1'; \mathbf{x}_g) \mathcal{V}_0^K \\ \vdots \\ \widehat{G}_B^*(\mathbf{x}_N'; \mathbf{x}_g) \mathcal{V}_0^K \end{bmatrix}. \quad (\text{C.22})$$

Therefore, the complete contribution for this Fréchet subkernel reads

$$\begin{aligned} \delta K_{B,b\mathcal{V}B}^{\text{est}}(\mathbf{x}) = & - \sum_s \sum_g \int d\omega \frac{-\omega^2}{K_B^2(\mathbf{x})} \overbrace{\widehat{p}_B^*(\mathbf{x}, \omega; \mathbf{x}_s)}^{\text{direct wavefield}} \times \\ & \int_{\mathbb{V}} d^3\mathbf{x}' \underbrace{\widehat{G}_B^*(\mathbf{x}, \omega; \mathbf{x}')}_{\text{2nd back-prop.}} \overbrace{-\omega^2 \left( \frac{1}{K_0(\mathbf{x}')} - \frac{1}{K_B(\mathbf{x}')} \right)}^{\text{scattering potential at } \mathbf{x}'} \underbrace{\widehat{G}_B^*(\mathbf{x}', \omega; \mathbf{x}_g) \widehat{p}_S(\mathbf{x}_g, \omega; \mathbf{x}_s)}_{\text{1st back-propagation of } \widehat{p}_S \text{ to } \mathbf{x}'} . \end{aligned} \quad (\text{C.23})$$

This result is shown in equation (2.52).

## C.4 Estimate of the background density perturbation $\delta\rho_{B,b\mathcal{V}B}^{\text{est}}$

Finally, we derive the corresponding expression for the background density perturbation  $\delta\rho_{B,b\mathcal{V}B}^{\text{est}}$  and the associated sub-kernel  $\overline{V}_{B,b\mathcal{B}B}$ . Analogously to equation (C.16), we, considering only perturbations in the density model, we define

$$\mathcal{V}_0^\rho = -\nabla \cdot \left[ \left( \frac{1}{\rho_0(\mathbf{x})} - \frac{1}{\rho_B(\mathbf{x})} \right) \nabla \right]. \quad (\text{C.24})$$

With this notation, we have from equation (C.14)

$$\begin{aligned} \widehat{p}_{S,b\mathcal{V}B}^\rho(\mathbf{x}_g; \mathbf{x}_s) = & - \int_{\mathbb{V}} d^3\mathbf{x}' \widehat{G}_B(\mathbf{x}'; \mathbf{x}_g) \mathcal{V}_0^\rho \left[ \int_{\mathbb{V}} d^3\mathbf{x}'' \frac{1}{\rho_B^2(\mathbf{x}'')} \left[ \nabla_{\mathbf{x}''} \widehat{G}_B(\mathbf{x}''; \mathbf{x}') \cdot \nabla_{\mathbf{x}''} \widehat{p}_B(\mathbf{x}''; \mathbf{x}_s) \right] \delta\rho_B(\mathbf{x}'') \right]. \end{aligned} \quad (\text{C.25})$$

where we substituted  $\widehat{p}_B(\mathbf{x}'; \mathbf{x}_s)$  by the expressing corresponding to equation (C.10). Discretisation of the volume integrals in equation (C.25) leads to

$$\widehat{p}_{S,b\mathcal{V}0}^\rho(\mathbf{x}_g; \mathbf{x}_s) = \underbrace{\mathbf{M}_3 \mathbf{M}_4}_{\overline{V}_{B,b\mathcal{B}B}} \begin{bmatrix} \delta\rho_B(\mathbf{x}_1'') \\ \vdots \\ \delta\rho_B(\mathbf{x}_N'') \end{bmatrix}, \quad (\text{C.26})$$

#### C.4 Estimate of the background density perturbation $\delta\rho_{B,b\mathcal{V}B}^{\text{est}}$

---

where

$$\mathbf{M}_3 = \begin{bmatrix} \widehat{G}_B(\mathbf{x}'_1; \mathbf{x}_g) \mathcal{V}_0^\rho & \cdots & \widehat{G}_B(\mathbf{x}'_N; \mathbf{x}_g) \mathcal{V}_0^\rho \end{bmatrix} \quad (\text{C.27})$$

and

$$\mathbf{M}_4 = \begin{bmatrix} -\frac{1}{\rho_B^2(\mathbf{x}_1'')} \nabla_{\mathbf{x}''} \widehat{G}_B(\mathbf{x}_1''; \mathbf{x}_1') \cdot \nabla_{\mathbf{x}''} \widehat{p}_B(\mathbf{x}_1''; \mathbf{x}_s) & \cdots & -\frac{1}{\rho_B^2(\mathbf{x}_N'')} \nabla_{\mathbf{x}''} \widehat{G}_B(\mathbf{x}_N''; \mathbf{x}_1') \cdot \nabla_{\mathbf{x}''} \widehat{p}_B(\mathbf{x}_N''; \mathbf{x}_s) \\ \vdots & \ddots & \vdots \\ -\frac{1}{\rho_B^2(\mathbf{x}_1'')} \nabla_{\mathbf{x}''} \widehat{G}_B(\mathbf{x}_1''; \mathbf{x}_N') \cdot \nabla_{\mathbf{x}''} \widehat{p}_B(\mathbf{x}_1''; \mathbf{x}_s) & \cdots & -\frac{1}{\rho_B^2(\mathbf{x}_N'')} \nabla_{\mathbf{x}''} \widehat{G}_B(\mathbf{x}_N''; \mathbf{x}_N') \cdot \nabla_{\mathbf{x}''} \widehat{p}_B(\mathbf{x}_N''; \mathbf{x}_s) \end{bmatrix}. \quad (\text{C.28})$$

Thus, the adjoint operation to equation (C.26) provides

$$\overline{\delta\rho_{B,b\mathcal{V}B}^{\text{est}}}(\mathbf{x}_g; \mathbf{x}_s) = \begin{bmatrix} \delta\rho_{B,b\mathcal{V}B}^{\text{est}}(\mathbf{x}_1'' | \mathbf{x}_s, \mathbf{x}_g) \\ \vdots \\ \delta\rho_{B,b\mathcal{V}B}^{\text{est}}(\mathbf{x}_N'' | \mathbf{x}_s, \mathbf{x}_g) \end{bmatrix} = \mathbf{M}_4^\dagger \mathbf{M}_3^\dagger \widehat{p}_S(\mathbf{x}_g; \mathbf{x}_s), \quad (\text{C.29})$$

which, after summation over all sources and receivers and integration over all frequencies, yields

$$\delta\rho_{B,b\mathcal{V}0}^{\text{est}}(\mathbf{x}) = - \sum_s \sum_g \int d\omega \int_{\mathbb{V}} d^3\mathbf{x}' \frac{1}{\rho_B^2(\mathbf{x})} \overbrace{\nabla \widehat{p}_B^*(\mathbf{x}; \mathbf{x}_s)}^{\text{direct wavefield}} \cdot \underbrace{\nabla \widehat{G}_B^*(\mathbf{x}; \mathbf{x}')}_{\text{2nd back-prop.}} \mathcal{V}_0^\rho \underbrace{\left[ \widehat{G}_B^*(\mathbf{x}'; \mathbf{x}_g) \widehat{p}_S(\mathbf{x}_g; \mathbf{x}_s) \right]}_{\text{1st back-propagation of } \widehat{p}_S \text{ to } \mathbf{x}'}. \quad (\text{C.30})$$





## D. Superposition of Subkernels of the Singular-wavefield Residual

In this appendix, we derive and analyse the superposition of the contributions  $\delta p_{S,BBB}$  and  $\delta p_{S,-B\mathcal{L}B}$  to the singular-wavefield residual. This is motivated by the destructive interference between these terms observed in the Experiments 3 and 6. For simplicity, we will carry out the analysis in the frequency domain for the bulk-modulus estimate components  $\widehat{\delta p}_{S,BBB}^K$  and  $\widehat{\delta p}_{S,-B\mathcal{L}B}^K$ .

The two contributions in question are given by the background part of equation (2.33.4) and by equation (2.33.8). Their bulk-modulus parts read in the frequency domain

$$\widehat{\delta p}_{S,BBB}^K(\mathbf{x}_g; \mathbf{x}_s) = - \int_{\mathbb{V}} d^3 \mathbf{x}' \frac{\omega^2}{K_0^2(\mathbf{x}')} \widehat{G}_B(\mathbf{x}'; \mathbf{x}_g) \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \delta K_B(\mathbf{x}') \quad (\text{D.1})$$

and

$$\widehat{\delta p}_{S,-B\mathcal{L}B}^K(\mathbf{x}_g; \mathbf{x}_s) = \int_{\mathbb{V}} d^3 \mathbf{x}' \frac{\omega^2}{K_B^2(\mathbf{x}')} \widehat{G}_B(\mathbf{x}'; \mathbf{x}_g) \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \delta K_B(\mathbf{x}') . \quad (\text{D.2})$$

Thus, their sum defines the combined contribution

$$\begin{aligned} \widehat{\delta p}_{S,BWB}^K(\mathbf{x}_g; \mathbf{x}_s) &= \widehat{\delta p}_{S,BBB}^K(\mathbf{x}_g; \mathbf{x}_s) + \widehat{\delta p}_{S,-B\mathcal{L}B}^K(\mathbf{x}_g; \mathbf{x}_s) \\ &= - \int_{\mathbb{V}} d^3 \mathbf{x}' \omega^2 \left( \frac{1}{K_0^2(\mathbf{x}')} - \frac{1}{K_B^2(\mathbf{x}')} \right) \widehat{G}_B(\mathbf{x}'; \mathbf{x}_g) \widehat{p}_B(\mathbf{x}'; \mathbf{x}_s) \delta K_B(\mathbf{x}') \\ &= \int_{\mathbb{V}} d^3 \mathbf{x}' \widehat{G}_B(\mathbf{x}'; \mathbf{x}_g) \mathcal{W}^K [\widehat{p}_B(\mathbf{x}'; \mathbf{x}_s)] \delta K_B(\mathbf{x}') , \end{aligned} \quad (\text{D.3})$$

where  $\mathcal{W}^K$  is the bulk-modulus part of a modified scattering potential. It is defined as

$$\mathcal{W}^K = -\omega^2 \left( \frac{1}{K_0^2(\mathbf{x}')} - \frac{1}{K_B^2(\mathbf{x}')} \right) = \mathcal{V}_0^K \left( \frac{1}{K_0(\mathbf{x}')} + \frac{1}{K_B(\mathbf{x}')} \right) , \quad (\text{D.4})$$

where  $\mathcal{V}_0^K$  is given by equation (C.16). Analogous considerations lead to the definitions of the combined contributions  $\widehat{\delta p}_{S,BWB}^\rho(\mathbf{x}_g; \mathbf{x}_s)$  as well as  $\widehat{\delta p}_{S,BWS}^K(\mathbf{x}_g; \mathbf{x}_s)$  and  $\widehat{\delta p}_{S,BWS}^\rho(\mathbf{x}_g; \mathbf{x}_s)$ .

Equation (D.3) demonstrates that these combined contributions carry scattered energy only from that part of  $\mathbb{V}$  where background perturbations overlap the singularities within the reference model. Consequently, we can conclude from this property that the background perturbation estimate obtained with the adjoint kernel of equation (D.3) will be non-zero only at the singularities in the reference model. In other words, if the reference model is chosen as the smooth background, these parts do not contribute at all.