Gravity inversion using wavelet-based compression on parallel hybrid CPU/GPU systems: application to southwest Ghana

Roland Martin, Vadim Monteiller, Dimitri Komatitsch, Stéphane Perrouty, Mark Jessell, Sylvain Bonvalot, Mark Lindsay

▶ **To cite this version:**

Roland Martin, Vadim Monteiller, Dimitri Komatitsch, Stéphane Perrouty, Mark Jessell, et al.. Gravity inversion using wavelet-based compression on parallel hybrid CPU/GPU systems: application to southwest Ghana. Geophysical Journal International, 2013, 195 (3), pp.1594-1619. 10.1093/gji/ggt334 . hal-00941851

HAL Id: hal-00941851
https://inria.hal.science/hal-00941851

Submitted on 18 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Geophysical Journal International

Geophys. J. Int. (2013) **195,** 1594–1619
Advance Access publication 2013 September 18

doi: 10.1093/gji/ggt334

Gravity inversion using wavelet-based compression on parallel hybrid CPU/GPU systems: application to southwest Ghana

Roland Martin,[1] Vadim Monteiller,[2,3] Dimitri Komatitsch,[2] Stéphane Perrouty,[1] Mark Jessell,[1] Sylvain Bonvalot[1] and Mark Lindsay[1,4]

[1]*Laboratoire GET, Université Toulouse* 3 *Paul Sabatier, IRD, CNRS UMR* 5563, *Observatoire Midi-Pyrénées,* 31400 *Toulouse, France. E-mail: roland.martin@get.obs-mip.fr*
[2]*Laboratory of Mechanics and Acoustics, CNRS UPR* 7051, *Aix-Marseille Univ, Centrale Marseille,* 13402 *Marseille Cedex* 20, *France*
[3]*Équipe-projet Magique3D, INRIA Bordeaux Sud-Ouest, Université de Pau et des Pays de l'Adour,* 64013 *Pau, France*
[4]*School of Geosciences, Monash University, P.O. Box* 28E, *Victoria,* 3800, *Australia*

Accepted 2013 August 20. Received 2013 July 12; in original form 2012 December 7

SUMMARY

We solve the 3-D gravity inverse problem using a massively parallel voxel (or finite element) implementation on a hybrid multi-CPU/multi-GPU (graphics processing units/GPUs) cluster. This allows us to obtain information on density distributions in heterogeneous media with an efficient computational time. In a new software package called TOMOFAST3D, the inversion is solved with an iterative least-square or a gradient technique, which minimizes a hybrid L^1-/L^2-norm–based misfit function. It is drastically accelerated using either Haar or fourth-order Daubechies wavelet compression operators, which are applied to the sensitivity matrix kernels involved in the misfit minimization. The compression process behaves like a preconditioning of the huge linear system to be solved and a reduction of two or three orders of magnitude of the computational time can be obtained for a given number of CPU processor cores. The memory storage required is also significantly reduced by a similar factor. Finally, we show how this CPU parallel inversion code can be accelerated further by a factor between 3.5 and 10 using GPU computing. Performance levels are given for an application to Ghana, and physical information obtained after 3-D inversion using a sensitivity matrix with around 5.37 trillion elements is discussed. Using compression the whole inversion process can last from a few minutes to less than an hour for a given number of processor cores instead of tens of hours for a similar number of processor cores when compression is not used.

Key words: Wavelet transform; Inverse theory; Numerical approximations and analysis; Satellite gravity; Gravity anomalies and Earth structure.

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

1 INTRODUCTION

Potential methods provide a useful way of investigating the Earth's interior as they can be used, for instance, as a complementary method to seismic imaging in geophysical exploration. Density contrasts inferred from gravity techniques can be correlated with seismic wave speeds and densities retrieved by seismic techniques. Gravimetry, when coupled with seismic tomography or ultrasounds acoustic methods, thus improves velocity–density correlations and above all constrains the density distribution of the geological structures under study. In the long term the objective is to obtain better images by correlating properties and performing joint inversions such as seismic/gravity inversions (Vermeesch *et al.* 2009; Gallardo & Meju 2011; Bailey *et al.* 2012) or gravity/magnetic inversions (Li & Oldenburg 2003; Fullagar *et al.* 2004, 2008; Gallardo 2007; Guillen *et al.* 2008; Commer 2011; Gallardo & Meju 2011; Moorkamp *et al.* 2011). The solution of the inverse problem depends on the formulation and discretization of the 3-D forward problem. Semi-analytical techniques are widely used in potential-field data inversion approaches for: regional-scale studies (Battacharyya 1980; Pilkington *et al.* 1994; Blakely 1995; Garcia-Abdeslem 2000; Garcia-Abdeslem *et al.* 2001; Zhang *et al.* 2004; Zhdanov 2009; Zhdanov *et al.* 2011; Cuma *et al.* 2012), general linear methods (Battacharyya 1980; Pilkington *et al.* 1994; Pilkington & Hildebrand 2000; Vermeesch *et al.* 2009; Morgan *et al.* 2011; Zhdanov *et al.* 2011), non-linear inversion techniques such as genetic algorithms or simulated annealing (Hildebrand *et al.* 1998; Garcia-Abdeslem 2000; Garcia-Abdeslem *et al.* 2001; Garcia-Abdeslem 2008) or the analytic signal method (Ortiz-Aleman & Urrutia-Fucugauchi 2010) in the case of magnetic-field inversion. Low-order finite-element methods as in Zhang *et al.* (2004), Agarwal & Srivastava (2010) and Zhdanov *et al.* (2011) are also used to compute the gravitational anomalies by replacing the voxel formulae of each block contribution by Gauss point-based finite-element

© The Authors 2013. Published by Oxford University Press on behalf of The Royal Astronomical Society.

formulations or by solving the full linear system related to the Laplace equation using a weak formulation. Zhang *et al.* (2004) pointed out some of the difficulties associated with the computation of gravity anomalies for large-density distributions in 3-D as it can be computationally demanding. Bottlenecks for inversion can occur, particularly if parallel genetic algorithms are used with prohibitive number of arrays stored in memory at each iteration of the process. The availability of increasing computation facilities and supercomputing resources now provides new perspectives for handling large data systems and significantly improving the efficiency of potential-field data inversion.

Wavelet analysis is a technique that is commonly used to capture relevant information from signal series. For instance in the field of potential methods, Poisson kernel wavelets have been used to localize sources responsible for measured potential fields in gravimetry, resistivity or magnetism (Mauri *et al.* 2010, 2011). Using a few Poisson kernel wavelets, multiscale wavelet tomography can been applied to study the impact of changes of sources in space and time on the measured potential field signals. Also, following a similar idea, 2-D analysis of wavelet-based inversion has been performed in Hornby *et al.* (1998) and Boschetti *et al.* (2001) by using appropriate wavelets defined by the given physics under study. It allows one to define depth extents of major geological units. Here, we aim at introducing another wavelet basis to process the massive inverse problem and to recover the density anomalies distribution. To achieve this, we propose a double acceleration technique for potential-field data inversion using both a multi–graphics processing unit (GPU) parallel implementation of L^2-based Newton–Raphson optimized least-square method and a high-order compression of the sensitivity matrices. The compression of sensitivity matrices allows us to reformulate the misfit function in the wavelet domain and to solve the inverse problem faster. Compression of the sensitivity matrix consists of applying wavelet operators to it, drastically reducing by this means the product of the transformed sensitivity matrix with a solution vector in the wavelet domain. This reduction of the cost of the matrix–vector product is possible thanks to a thresholding process applied to the transformed sensitivity matrix. High compression ratios and related sparseness up to two or three orders of magnitude can be reached. Compression is thus advantageous in terms of both memory storage and computational speed. This double acceleration, due to a hybrid CPU/GPU multiprocessor approach and matrix compression, allows us not only to accelerate the inversions but also to process data sets one or two orders of magnitude larger than classical data volumes. We invert gravity data from southwest Ghana and perform inversions of the main geological unit geometries for a given *a priori* density model obtained from samples collected from the field. We focus our study on fast iterative solvers for 3-D inversion algorithms to compute the density distribution of the subsurface down to 10 km. Below this depth, gravity inversion does not give well-constrained results. Of course the technique can be also applied to any inversion process involving sensitivity matrices, as encountered in seismic or magnetic inversion among many other problems.

The paper is organized as follows: In Section 2, we describe how recent supercomputing technology such as hybrid CPU/GPU clusters lead to new computational challenges in 3-D inversion using potential and/or joint techniques and a multiresolution approach. In Section 3, the physical problem is posed and the space discretization formulations are given. We present the mathematical background of the gravity problem and we show the discretization of the problem with a voxel method or finite elements. In Section 4, the wavelet compression process is described and we show how the compression wavelets behave as pre-conditioners of the system and can handle massive data volumes. We will show how the number of wavelet coefficients, that describe the main signal content of the sensitivity kernels, can be reduced by up to two or three orders of magnitude according to the choice of a given threshold. which is advantageous in terms of memory storage and computational speed. In Section 5, the inversion process is described using both L^1 or L^2 norm misfit functions. In Section 6, we show application results of data inversion in Ghana for original uncompressed sensitivity kernels of around several trillion elements and for compressed sensitivity kernels of a few billion elements. Accuracy, convergence and stability are addressed. In Section 7, parallel performances of our code TOMOFAST3D (tomographic imaging using very fast inversion technique in 3-D) using multi-CPU or multi-GPU are shown and we discuss (i) how the computational time of the inversion code scales with the number of processor cores for pure multi-CPU clusters or hybrid multi-CPU/multi-GPU clusters and (ii) how inversion can take less than 1 hr or even a few minutes in some cases on both huge data and parameter sets. We conclude in Section 8.

2 NEW CHALLENGES AND IMPROVEMENTS FOR MASSIVELY PARALLEL MODELLING

Realistic high-resolution 3-D gravity or magnetic modelling is a task made difficult by the requirement of large computational memory allocation resources and CPU time consumption. One of the big challenges of recent years has been to deal with huge anomaly data sets and huge sets of density or magnetic parameters. Matrix compression methods based on wavelet transform can be a complementary way to deal with such huge inverse problems by taking advantage of the strong decay of potential fields (Pilkington & Hildebrand 2000; Li & Oldenburg 2003) with increasing block-to-source distance and of the smoothness of gravity or magnetic sensitivity kernel matrices. Indeed, the strong decay of potential fields as $1/r^2$ in gravimetry and $1/r^3$ in magnetism aids the construction of smooth gravity or magnetic kernels. As a result pre-conditioning of the linear inverse problem system can be built according to both depth weighting and global behaviour of the potential field to accelerate the whole inversion process.

To solve the inversion of large data sets, it is crucial to have fast solvers of the direct problem applied to large parameter volumes. Indeed, low-order, sequential or non-massively parallel finite elements (Agarwal & Srivastava 2010; Zhang *et al.* 2004) on the one hand and sequential (Talwani *et al.* 1959; Battacharyya 1980; Blakely 1995; Garcia-Abdeslem 2000; Garcia-Abdeslem *et al.* 2001) or parallel (Commer 2011; Moorkamp *et al.* 2011) voxel-based semi-analytical solutions on the other hand are generally used to solve 3-D forward gravity problems. Several articles on 3-D massively parallel inverse problem resolution using purely CPU supercomputing platforms with gradient methods based on L^2 misfit function minimization have been published in the last 2 yr (Commer 2011; Wilson *et al.* 2011; Zhdanov *et al.* 2011; Cuma *et al.* 2012). Stochastic techniques such as genetic algorithms (Zhang *et al.* 2004) are also used on these platforms but are still too greedy in terms of both processor usage and memory storage of increasingly large matrices. Recently, to achieve further acceleration some authors have solved the forward problem using multithreaded implementations on GPUs supercomputing platforms in which multi-CPU platforms are connected to graphics card processor units. This has been done for instance for seismic wave propagation (Komatitsch *et al.* 2010; Michéa &

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

Komatitsch 2010; Komatitsch 2011) and for gravimetry (Moorkamp *et al.* 2010, 2011). Gradient methods based on L^2 misfit function are commonly used on the multi-CPU. More recently a probability function approach (Liu *et al.* 2012) has been proposed on multi-GPU platforms for magnetic inverse problems. These authors took advantage of the high performance of GPUs to achieve drastic acceleration of voxel algorithms. For magnetic inversion, Zhdanov *et al.* (2011) have tested gradient methods on a few hundred million parameters and around 500 000 data using multi-CPU architectures and obtained interesting weak scaling on up to 576 processor cores. However, solutions are obtained after several tens of hours. In the field of forward gravity problems, Moorkamp *et al.* (2010) have used far fewer voxel parameter distributions (around one million at the most).

Since we are going to also handle large numerical grids, we decide to make use of hybrid computing using many GPUs (Owens *et al.* 2008; Fatahalian & Houston 2008; Kirk & Hwu 2010) in parallel. The goal of using GPU technology is to significantly accelerate the calculations because in recent years GPUs have quickly become an important and powerful way to carry out scientific computations in the case of algorithms that lend themselves well to parallel computing. Current GPUs can be seen as hardware implementations of a Single Instruction Multiple Data (SIMD) programming model in which a large number of elementary processor cores as well as a hardware scheduler maintain thousands of elementary lightweight threads active simultaneously by effectively suspending tasks that are waiting for memory transactions and switching to other tasks that are ready to compute. The CUDA language programming has been used to port the Fortran 90 version with careful rearrangement of the calculations and the GPU memory occupancy. In addition, we use message passing based on the Message Passing Interface (MPI, Gropp *et al.* 1994; Pacheco 1997)) to exchange information between the different compute nodes that carry the different GPU boards of the computer cluster.

The main issues related to depth-weighting, wavelet compression, data inversion process, porting of the code on multicore or multi-GPU clusters and code performance study will be addressed in the forthcoming sections of this paper because these ingredients are essential to drastically accelerate the data inversion procedure. We will also show how we obtain double acceleration of the inversion algorithm by computing on hybrid multi-CPU/multi-GPU supercomputing platforms and by compressing the sensitivity matrix in the inversion procedure using wavelet operators.

3 PHYSICAL FORMULATION OF THE GRAVITY PROBLEM AND ITS DISCRETIZATION: CALCULATION OF POTENTIALS AND SENSITIVITY KERNELS

3.1 Gravity equation

The gravity force is the resultant of the gravitational force and the centrifugal force. However, in this paper, we consider only regional scale, therefore we do not take into account the effects related to the rotation of the Earth. In the absence of centrifugal forces, the gravitational potential at a given position $\mathbf{r}$ caused by an arbitrary density distribution ρ is given by

$$\Phi(\mathbf{r}) = G \int_{\Omega} \frac{\rho(\mathbf{r}')}{||\mathbf{r} - \mathbf{r}'||} \mathrm{d}\mathbf{r}', \tag{1}$$

where $\mathbf{r}'$ represents one point position within the density distribution and G the universal constant of gravity equal to

$$G = 6.672 \times 10^{-11}\ \mathrm{m^3\,kg^{-1}\,s^{-2}}. \tag{2}$$

Let us now show that the gravity field can be computed by the linear relation

$$g = S\rho, \tag{3}$$

where g is the distribution of the gravity anomaly at a given set of locations $\mathbf{r} = (x, y, z)$, S the gravity sensitivity matrix that can also be seen as a global gravity mass matrix and ρ the global density distribution.

Let us then discretize the Earth domain into voxels that can be equidimensional blocks. We use a set of prisms of rectangular section and we compute the sensitivity kernel using formulae given by Blakely (1995) as follows. This semi-analytical gravity approach is based on the global sum of all the partial analytical gravity contributions of each voxel that compose the earth model under study. A semi-analytical solution can be seen as a hybrid analytic and spatially discretized solution as the one we describe hereafter for the gravity problem. The gravity effect described by vertical gravity force of a mass distribution located in coordinates (ξ, η, ζ) on surface points (x, y, z) can be computed along the z direction as

$$g(x, y, z) = \frac{\partial \Phi}{\partial z}$$

$$= G \int\int_{v}\int \rho \frac{(z - \zeta)}{[(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2]^{3/2}} \mathrm{d}\xi\, \mathrm{d}\eta\, \mathrm{d}\zeta. \tag{4}$$

The gravity field can therefore be computed by the linear relation

$$g(x, y, z) = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} S^{i,j,k}\rho_{i,j,k}, \tag{4}$$

where

$$S^{i,j,k}(x, y, z) = G\mu_{i,j,k}\left(z_k tg^{-1}\left(\frac{x_i y_j}{z_k R_{i,j,k}}\right) - x_i \log(R_{i,j,k} + y_j)\right.$$

$$\left. - y_j \log(R_{i,j,k} + x_i)\right), \tag{5}$$

with $R_{i,j,k} = \sqrt{x_i^2 + y_j^2 + z_k^2}$, $\mu_{i,j,k} = (-1)^i(-1)^j(-1)^k$, $x_i = x - \xi_i$, $y_j = y - \eta_j$, $z_k = z - \zeta_k$, $i, j, k = 1$ or 2 and (ξ_i, η_j, ζ_k) being the coordinates of the elemental cell vertices. g is the gravity anomaly computed at the observation point. If we want to be more accurate by taking into account more easily topography or the distorted interfaces between geological structures we can introduce finite-element integration in space. The topography can be obtained by extending the upper blocks to the level of the topography. Furthermore, to reduce the computational cost of the logarithm and arctangent calculations, Gauss quadrature weights could be introduced to integrate eq. (4) as is commonly performed in finite-element integration rules. This could be performed without loss of accuracy compared to the semi-analytic integration and without using further logarithm or arctangent functions. In the elements close to topography, this integration becomes

$$g(x, y, z)$$

$$= G \int\int_{v}\int \rho \frac{(z - \zeta)}{[(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2]^{3/2}} J\, \mathrm{d}\alpha\, \mathrm{d}\beta\, \mathrm{d}\psi, \tag{6}$$

where J is the local Jacobian of the bijective mapping of the reference cube defined by $(\alpha, \beta, \psi) \in [0,1] \times [0,1] \times [0,1]$ to a

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

given deformed cube that is defined by $(\xi, \eta, \zeta) \in \Omega$ and is located close to the topography. Then, after discretization and integration on elementary deformed cubes, the gravity anomaly is computed as:

$$g(x, y, z) = G \sum_{l=1,N} \sum_{i=1,N_\alpha} \sum_{j=1,N_\beta} \sum_{k=1,N_\psi} \rho_{ijk}^{l} \times \frac{\left(z - \zeta_{ijk}^{l}\right)}{\left[\left(x - \xi_{ijk}^{l}\right)^2 + \left(y - \eta_{ijk}^{l}\right)^2 + \left(z - \zeta_{ijk}^{l}\right)^2\right]^{3/2}} \times J_{ijk}^{l} \omega_i \omega_j \omega_k \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\psi, \quad (7)$$

where $N = N_x \times N_y \times N_z$ is the number of elementary cell volumes, N_α, N_β and N_ψ are the numbers of collocation or Gauss points associated to each cell in each direction α, β or ψ, and ω_l (with $l = i, j$ or k) are the quadrature weights of integration over each cell. If single point Gauss quadrature integration is chosen we have $\omega_l = 1$ for $l = i, j$ or k and $N_\alpha = N_\beta = N_\psi = 1$, $\mathrm{d}\xi = \mathrm{d}\eta = \mathrm{d}\gamma = 1$. The effect of density in block l on the gravity potential at data position (i_D, j_D, k_D) located at an anomaly point (x,y,z) of the surface can be defined as:

$$g_l^{i_D, j_D, k_D} = G \sum_{i=1,N_\alpha} \sum_{j=1,N_\beta} \sum_{k=1,N_\psi} \rho_{ijk}^{l} \times \frac{\left(z - \zeta_{ijk}^{l}\right)}{\left[\left(x - \xi_{ijk}^{l}\right)^2 + \left(y - \eta_{ijk}^{l}\right)^2 + \left(z - \zeta_{ijk}^{l}\right)^2\right]^{3/2}} \times J_{ijk}^{l} \omega_i \omega_j \omega_k \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\psi. \quad (8)$$

We can then write in a more compact form that

$$g(x, y, z) = \sum_{l=1,N} g_l^{i_D, j_D, k_D} = G \sum_{l=1,N} M^l \rho_l, \quad (9)$$

where M^l behaves as a local gravity mass matrix restricted to the contribution of each cell, ρ_l being the density distribution local to each cell.

Finally the gravity field can be computed by the linear relation

$$g = S\rho, \quad (10)$$

where $S = GM$ is the global gravity sensitivity matrix.

4 COMPRESSION OF THE KERNELS WITH ORTHOGONAL OR BI-ORTHOGONAL WAVELETS

Recovering large-scale 3-D geological models requires many density model parameters and a vast amount of data to be inverted for. Unfortunately high-resolution 3-D sensitivity kernels, which are full arrays, can require huge amounts of memory storage and their use for solving both forward and subsequently the inverse problems can thus be prohibitive. However, for each data row, the amplitude of the coefficients of the related subarray strongly decays as a function of the distance separating the data location from a parameter mapped on a block or a mesh point, depending on the semi-analytical blocky integration or the numerical integration that has been chosen. It is then reasonable to define an amplitude threshold as a way to create sparse versions of the sensitivity matrix. We apply such an amplitude threshold criterion to convert the originally full arrays into sparse versions with several sizes. To construct such sparse compressed arrays we apply different thresholds as percentage of maximum amplitude in each data row-wise subarray. These thresholds are defined as follows. As the kernels show continuous and smooth behaviour, it is well known that the application of wavelet operators to the matrices will provide very efficient compression ratios. Indeed, we will demonstrate that sensitivity kernels and density models have compact or sparse representations in the wavelet domain. This is the basic motivation for formulating the gravity inverse problem using a discrete basis of orthogonal or bi-orthogonal wavelets. We use two different compression wavelet operators, namely second-order Haar wavelets and fourth-order wavelets of Daubechies type (D4) for the reconstruction, and show that Daubechies wavelets are more adapted to the continuous sensitivity kernels that we use. Indeed, they are able to recover, respectively, the linear or quadratic information contained in the input signal, given here by the gravity sensitivity kernels. It is also important to show how these two kinds of wavelets behave in terms of compression efficiency. We thus show hereafter how these wavelets behave and why Daubechies wavelets are more efficient in terms of compression of the sensitivity matrices.

Lifting fourth-order Daubechies wavelets (CDF4) (Cohen *et al.* 1992) could also be considered [see for instance one of our articles (Chevrot *et al.* 2012)] but they are not orthogonal (they are bi-orthogonal) and do not preserve energy. We have also used them as well as D4 wavelets but even if it appears that they can be more efficient than more classical orthogonal wavelets, we will not discuss them here for sake of clarity on how given compression errors can define the thresholding of relevant wavelet coefficients. We test here our method for different compression ratios related to different applications of wavelet operators in a multiresolution perspective. Comparisons of their performance are established according to speed of computation and accuracy level.

Let us summarize the principal lines of the compression procedures and also briefly summarize the calculation of the wavelets. More details can be found on orthogonal Daubechies wavelets in Daubechies (1988, 1992) and on bi-orthogonal lifting Daubechies wavelets in Cohen *et al.* (1992) or Daubechies & Sweldens (1998). The reader is also referred to Uytterhoeven *et al.* (1997) for higher-order lifting wavelets up to sixth order and the explicit description of the coefficients involved. For a given set V of elements defining a function in an interval [0,1], let V^0 be the vector space of constant functions in this interval and V^j the vector space of piecewise constant functions over each 2^j equal subintervals. We can define a basis function defined on each vector space V^j, which verifies the following property:

$$V^0 \subset V^1 \subset \ldots \subset V^j \subset V^{j+1} \subset \ldots. \quad (11)$$

A basis function can be defined on each vector space V^j with the following scaling functions:

$$\phi_i^j = \phi(2^j x - i), \; i = 1, \ldots, 2^j - 1, \quad (12)$$

where

$$\phi(x) = 1, \quad \text{for} \quad 0 \leq x \quad \text{and} \quad x < 1$$
$$= 0, \quad \text{for} \quad 0 > x \quad \text{and} \quad x > 1. \quad (13)$$

The wavelets are the functions in V^{j+1} that are orthogonal to all functions in V^{j+1}. They can describe all the details of the functions in V^{j+1} that can not be described in V^j. They belong to the orthogonal subspace W^j, which verifies $V^{j+1} = V^j \oplus W^j$. The Haar wavelets are the piecewise flat shaped functions given by

$$\psi_i^j = \psi(2^j x - i), \; i = 1, \ldots, 2^j - 1, \quad (14)$$

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

where

$$\psi(x) = 1 \quad \text{for} \quad 0 \le x \quad \text{and} \quad x < \frac{1}{2}$$
$$= 0 \quad \text{otherwise.} \tag{15}$$

In three dimensions, the Haar wavelet functions can be described by a tensor product of 1-D basis functions since the V^j spaces are orthogonal. The 3-D scaling functions are given by

$$\Phi^j_{l,m,n}(x, y, z) = \phi^j_l(x)\,\phi^j_m(y)\,\phi^j_n(z), \tag{16}$$

and the 3-D Haar wavelets by

$$\Psi^j_{1,l,m,n}(x, y, z) = 2^{\frac{3j}{2}}\,\phi^j_l(x)\,\phi^j_m(y)\,\psi^j_n(z),$$
$$\Psi^j_{2,l,m,n}(x, y, z) = 2^{\frac{3j}{2}}\,\phi^j_l(x)\,\psi^j_m(y)\,\phi^j_n(z),$$
$$\Psi^j_{3,l,m,n}(x, y, z) = 2^{\frac{3j}{2}}\,\psi^j_l(x)\,\phi^j_m(y)\,\phi^j_n(z),$$
$$\Psi^j_{4,l,m,n}(x, y, z) = 2^{\frac{3j}{2}}\,\phi^j_l(x)\,\psi^j_m(y)\,\psi^j_n(z),$$
$$\Psi^j_{5,l,m,n}(x, y, z) = 2^{\frac{3j}{2}}\,\psi^j_l(x)\,\phi^j_m(y)\,\psi^j_n(z),$$
$$\Psi^j_{6,l,m,n}(x, y, z) = 2^{\frac{3j}{2}}\,\psi^j_l(x)\,\psi^j_m(y)\,\phi^j_n(z),$$
$$\Psi^j_{7,l,m,n}(x, y, z) = 2^{\frac{3j}{2}}\,\psi^j_l(x)\,\psi^j_m(y)\,\psi^j_n(z), \tag{17}$$

which defines an orthonormal basis of V^j. The wavelet decomposition of a function f in the V^j space is then

$$f = \sum_{j=0}^{j\max} \sum_{l,m,n} \left[\langle f \mid \Phi^j_{l,m,n} \rangle \Phi^j_{l,m,n} + \sum_{i=1}^{7} \langle f \mid \Psi^j_{i,l,m,n} \rangle \Psi^j_{i,l,m,n} \right], \tag{18}$$

where $\langle f \mid g \rangle$ defines the L^2 product in the V^j subspace. The function f can then take the form $f = \sum_M a_M f_M$, where a_M are the coefficients of the wavelet decomposition and f_M are the basis functions.

Algorithmically, the compressed function f denoted by f_l at the compression scale level l can be retrieved by using the following compression matrix W^l as

$$f_l = (\Pi_{i=1,l} W_i) f, l = 1 \text{ to max scale level,} \tag{19}$$

where $(\Pi_{i=1,l} W_i)$ is the matrix product of the compression operators at each scale level i of the space V^j. An interesting feature of the compression matrix is that the reconstruction matrix is exactly its transpose because the basis function is orthogonal. This means that no inverse matrix has to be computed, which is very useful for reducing memory storage and for performing easy computations. Further information on the implementation of the construction of Haar and fourth-order Daubechies wavelet compression can be found in the Appendices.

A wavelet transform leads to a representation of a function with a number of coefficients that is equal to the initial number of data describing this function. However in most cases a large number of these wavelet coefficients are very small and can thus be neglected without losing any significant amount of information. The essential question is then how to deal with many wavelet coefficients that are very small. We introduce an error threshold after applying successive wavelet operators to each row and over a fixed number of scales. Then, removing all the coefficients having an absolute value below this threshold leads to only small errors in the reconstructed functions. For a given scale level (i.e. a given number of successive applications of the wavelets in each spatial direction), we must define a relevant way to define a reconstruction error between an initial function f, given here by the sensitivity matrix S, and a reconstructed function f_R that is given by the sensitivity matrix S_R. This error is:

$$E = \frac{||S - S_R||^2_L}{||S||_{L^2}}. \tag{20}$$

With orthogonal wavelets such as those of Haar or Daubechies, the wavelet transform preserves energy. As a result of the orthogonality of the wavelet basis, for an input signal (sensitivity) matrix S of length N this property can be written:

$$\sum_{n=1}^{N} S_n^2 = \sum_{n=1}^{N} W_n^2, \tag{21}$$

where the W_n are the coefficients of the wavelet expansion of signal S in the wavelet basis. This property is very convenient for quantitatively defining the compression in the wavelet domain. The compression algorithm using the ℓ_2 norm involves three steps (Stollnitz *et al.* 1995): computing the wavelet coefficients representing the input signal in the wavelet basis, sorting these coefficients in order of decreasing absolute values and finally determining the error as a function of the number of wavelet coefficients kept for the reconstruction, taking the coefficients sorted by decreasing absolute values. If we denote by M_i the number of wavelet coefficients kept for the reconstruction for each row i (i.e. for each data i) of the sensitivity kernel, the reconstruction error E for each row is simply given by:

$$E(M_i) = \sum_{n=M_i+1}^{N} W_n^2. \tag{22}$$

After normalization it can be defined as

$$E(M_i) = \sum_{n=M_i+1}^{N} W_n^2 \Big/ \sum_{n=1}^{N} W_n^2 = \sum_{n=M_i+1}^{N} W_n^2 \Big/ ||S||_{L^2}. \tag{23}$$

Let us compute now this error for different compression scale level choices and different wavelets (Haar or D4). For an original kernel of $890 \times 890 \times 68$ parameter elements corresponding to a single datum, we have tested Haar or D4 wavelets for three different compression scales. We observe in Fig. 1 that for a three-scale level, D4 compresses the kernel more efficiently by more than an order of magnitude than Haar wavelets. For the same reconstruction error a four-scale level D4 compression is more efficient because it reduces the number of retained wavelet coefficients and increases the compression ratio compared to both three-scale-level Haar and D4: this compression ratio is defined as the ratio of the number of elements of the original matrix over the number of wavelet elements retained after thresholding and reconstruction of the wavelet-compressed matrix. Haar wavelets give higher errors and much worse compression ratios at the fourth level (not shown here) than D4. This is due to the fact that the sensitivity matrix is continuous while the Haar wavelets are more adapted to discontinuous functions. The five-scale level D4 reduces the number of coefficients by another factor around 2 and can also be chosen as a good candidate as four-scale D4. We have thus chosen to take four levels of compression in all the inversions to study the accuracy of the reconstruction using D4 wavelets for 0.001 per cent, 0.005 per cent, 0.01 per cent and 0.05 per cent and 0.1 per cent given kernel reconstruction errors. Compression ratios of, respectively, 0.239 per cent, 0.1166 per cent, 0.085664 per cent, 0.04344 per cent and 0.03029 per cent have been obtained. In Fig. 2,

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

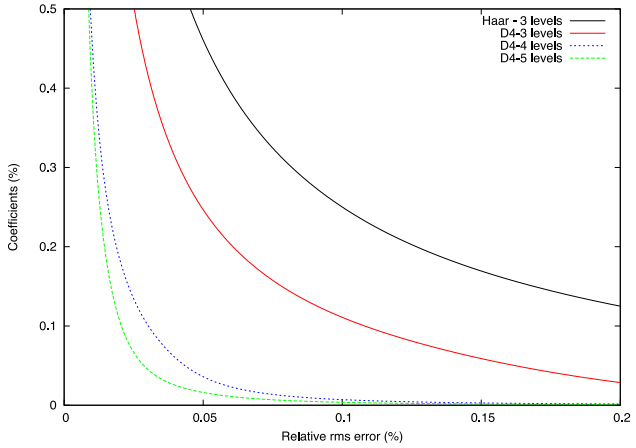

Figure 1. Number of coefficients selected for a given thresholding error. The number of selected wavelet coefficients decreases with the compression scale: five (green), four (blue) and three (red) compression scales are used for the Daubechies wavelets. Three-scale Haar wavelets (black curve) are less efficient than three Daubechies wavelets with more coefficients kept.

we show gravity kernels with these different accuracies and compressions. We can see that the kernel is not well reconstructed with a lack of coherence for errors higher than 0.01 per cent, while for errors lower than 0.01 per cent the kernels are still coherent, so we can say that three levels at the most can be sufficient to describe the kernels. Furthermore an optimized compression ratio and reconstruction error can be obtained automatically by minimizing the logarithm $\log_{10}(E)$ of the error curve. In the case of Fig. 2, the minimum of $|\log_{10}(E)\mathrm{d}M_i|$ is reached for $M_i = 0.1245$ per cent of the total number of coefficients, with a related reconstruction error $E(M_i) = 0.004314$ per cent. $M_i = 0.1245$ per cent lies between the previously given values 0.1166 per cent and 0.239 per cent, which we have shown to result in good reconstruction.

By resorting to compression we thus introduce a natural smoothing filter and we can quantify exactly the reconstruction error as well as the amount of signal lost after compression. Compression ratios as high as 400 can be reached and are higher than those obtained by Li & Oldenburg (2003) (magnetic inversion case with ratios of 20 to 50). This is due to the fact that the compression efficiency depends on the model and on the chosen spatial discretization as well as on the physics of the problem and on the decay behaviour of the potential field from the source (gravity field decay in our case and magnetic field decay in the case of Li and Oldenburg). This also depends on whether the spatial grid is oversampled or not and if successive wavelets are applied as in our application: oversampled matrices can be subjected to more efficient compression. We then take the reconstruction error into account implicitly in the covariance matrices involved in the inversion algorithm and in the reconstruction of the density model. In practice the reconstruction error is hidden in the regularization parameter of the misfit function defined by the L^2-curve as discussed in the forthcoming section about the inverse problem.

In addition, the sparseness of the sensitivity matrix increases with compression and the parallel code improves in efficiency by factors that are proportional to the sparseness of the linear system. As a conclusion, if we apply successively four times the fourth-order Daubechies wavelet operators to the gravity sensitivity matrix then compression is more efficient than Haar compression and also compression ratios as high as around 400 can be obtained for a chosen reconstruction error of 0.239 per cent and for the specific matrix built for southwest Ghana region.

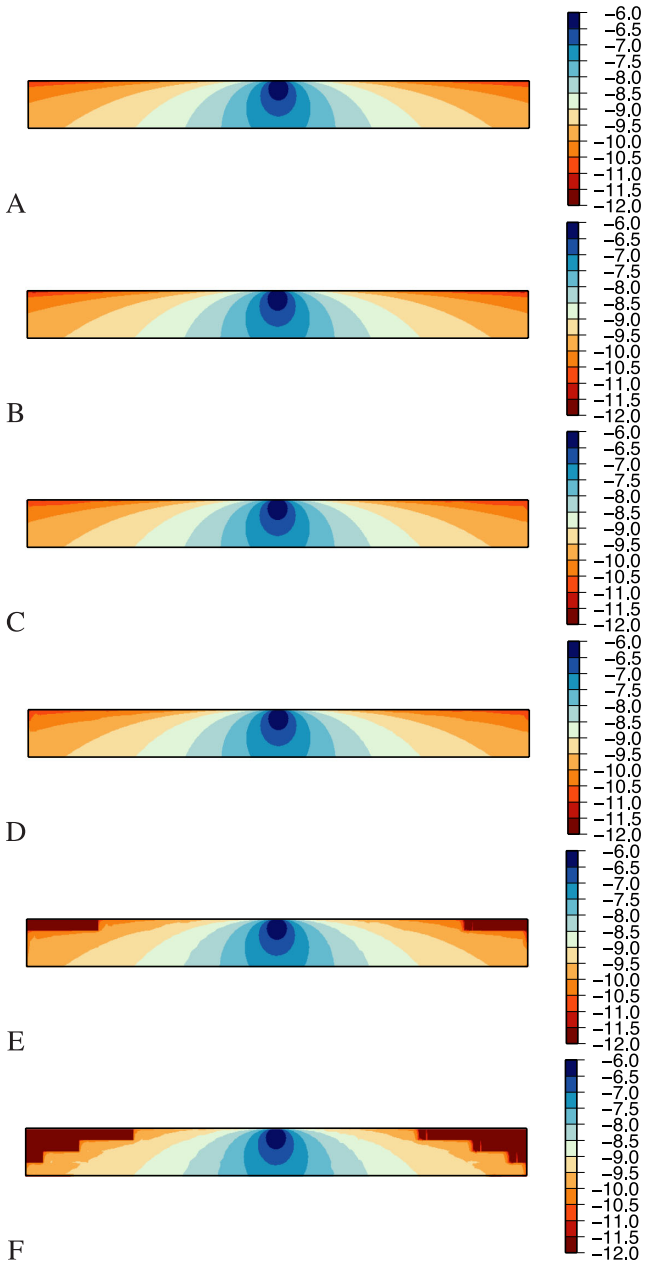

Figure 2. (A) Cross-section of an original non-compressed kernel for a given data located in the centre of the computational domain. Compressed kernels using fourth-order Daubechies wavelets and a fourth compression scale are shown from (B) to (F) for reconstruction errors of, respectively, 0.001, 0.005, 0.01, 0.05 and 0.1 per cent. The corresponding compression ratios are, respectively, 0.239, 0.1166, 0.085664, 0.04344 and 0.03029 per cent. We can then observe that the best reconstructions are obtained with compression ratios of 0.239 per cent and 0.1166 per cent (B and C) with little distortion of the kernel and little missing information when compared to the non-compressed kernel. Kernels with stronger compression (D to F) introduce signal distortion, spurious artefacts close to the boundaries and lack of information that will produce inaccurate solutions during the inversion process. Optimized threshold errors applied to retain an optimized number of compression wavelet coefficients are then located around the inflection point of the logarithm of the error curve given in Fig. 1. This inflection point corresponds to the minimum of the derivative of the logarithmic reconstruction error and to kernel reconstruction errors lower than 0.01 per cent and compression ratios higher than 0.12 per cent.

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

5 INVERSE PROBLEM SOLVER USING L^1- OR L^2- NORM

5.1 Inversion algorithm

Let us apply forward and inverse gravity modelling in three dimensions to compute the Bouguer and free-air gravity anomalies for a complex realistic model describing the main structural elements of southwest Ghana taken as a validation target of our inversion technique. For the resolution of the inverse problem we use an assemblage of 53 862 800 prismatic bodies ($890 \times 890 \times 68$ cuboidal prisms). Each prismatic body has a dimension of 200 m in the x, y and z directions. Let us now describe the inverse problem. At each compression scale l, we must find a set ρ^l of pixels ρ_k^l such that gravity data honours the equation

$$g = S^l \rho_l^T, \tag{24}$$

with $S^l = Sw^l$ and $\rho_l = \rho w^l$, where w^l is a wavelet operator applied in the three directions. We can observe that the inverse problem can be formulated in the wavelet domain, the parameters being reconstructed by applying the transpose wavelet operator to the wavelet coefficients ρ_l at the scale l. Compression and reconstruction are operated through the algorithms given in (A1) for the Haar wavelet and in (A2) for the fourth-order Daubechies wavelet. Gauss–Newton Raphson, Tikhonov or classical gradient-based solver inversion techniques or non-linear stochastic algorithms like simulated annealing (Ortiz-Aleman & Martin 2005; Garcia-Abdeslem 2008) or genetic algorithms for the global optimization can be used to minimize L^2 or L^1 misfit functions. For the sake of simplicity we choose to use the least-square root (LSQR) inversion technique (Chapman & Pratt 1992; Pratt & Chapman 1992; Pratt *et al.* 1998; Pratt 1999; Sirgue & Pratt 2004; Loris *et al.* 2007, 2010). The principal steps of the whole sequential LSQR algorithm are depicted in Fig. 3 and its implementation is given in the Appendices in pseudo-code form.

Given the S sensitivity matrix and g a data vector we aim to find a model ρ that minimizes the misfit function:

$$J(\rho) = ||S\rho - g||^2_{L^2} . \tag{25}$$

One of the main problems encountered when solving an inverse problem is the non-uniqueness of the solution (Koren *et al.* 1991; Parker 1994; Mosegaard & Tarantola 1995, 2002), particularly in very unconstrained problems such as data inversion based on potential methods. This is why it is important to pre-condition the linear system and converge to a realistic solution in less iterations, and why we apply weighting functions to the matrix S. Following (Farquharson & Oldenburg 2004; Pilkington 2009) they can be built as a diagonal matrix P with elements varying according to a z^m power law, z being the depth of a block according to the data row under consideration. Exponent m can be taken as 1/2, 1 or 3/2. Here we choose to take $m = 1$ because with $m = 1/2$ the structures are shifted towards the top of the model and with $m = 3/2$ towards the bottom of the model, which is not very realistic. As an example, in Fig. 4 we compute a synthetic gravity response at the top of a medium consisting of a 6-km-long and 6-km-wide cuboidal structure located at 3-km depth and a 2-km sphere in diameter located at 2-km depth. Densities are 2.67, 3.07 and 3.22 kg m^{-3} for the background, the cuboidal structure and the sphere, respectively (corresponding to density contrasts of 0, +0.4 and +0.55 with respect to the reference density value). The size of the computational medium is 20 km wide and long and 5 km deep. We try different z^m-law pre-conditioners. After inversion, all solutions give very similar gravity responses which all fit almost exactly the synthetic response. We can see that with $m = 1$, the depth of the top of the structures is better defined than with $m = 1/2$ or $m = 3/2$. Furthermore this value also better modifies the middle part of the model as shown on the synthetic case. In most geophysical applications, the *a priori* model is believed to be better constrained in its upper part by geology and the data inversion aims to improve the mid to lower parts of the model. In our application to real data (Section 6), we will finally arbitrarily retain the pre-conditioning value $m = 1$ to illustrate how the constraints on the upper model given by geology may be kept as strong constraints during the inversion process. The shape of the structures is not well retrieved because of the non-uniqueness of the solution but the purpose of our work here is not to define the shape but to provide, without being too intrusive to the initial *a priori* models, a very fast inversion tool to define the excess or lack of mass compared with a given *a priori* model. If we wanted to be more intrusive, other depth weighting pre-conditioners P (Li & Oldenburg 2003; Farquharson & Oldenburg 2004; Pilkington 2009; Commer 2011) could be applied on gradient $\nabla J(\rho) = S^T(S\rho - g)$ such that:

$$\nabla J(\rho) = P^T S^T (S\rho - g). \tag{26}$$

This weighting could be applied because the sensitivity kernels are different by a few orders of magnitude in the vicinity of measured points compared to points located at depth and because the solution given by linear methods is only perturbed in the first few layers of the model. To avoid this, down-weighting could be applied to these layers, which would allow us to restrict the inversion to remain inside a given depth range (Commer 2011). In our model, we have chosen a 13.6-km depth range with 68 blocks of 200 m each due to the given *a priori* geological model defined at the resolution of the measured gravity data (1 point every 200 m at the surface). Below this depth, geologists are not able to provide physically relevant density models. Minimum and maximum values of the model could be fixed between these two depths to avoid overshoots or undershoots during the iterative inversion process. We decided not to apply this type of pre-conditioning as it would too strongly limit the inversion process and would force the method to assume that the new possible model is very close to the initial model. Even if this possible preconditioner were implemented, here we made the choice of the z law pre-conditioning for the sake of clarity and to avoid such strong intrusion in the initial *a priori* model. Besides, to avoid boundary effects, that is, effects coming from the finite size of the grid that we defined, we have extended the values of the model located at the bottom boundary into an additional outer box of 10 km thickness. We have also extended the four lateral boundaries of the model by the values of a vertically averaged profile of the model. This has allowed us to drastically damp boundary effects. We thus solve the new problem:

$$\text{Minimize } ||S_1 m_1 - g||_{L^2} \tag{27}$$

with $S_1 = SP$ and $\rho = Pm_1$. We then use a wavelet transform to compress the pre-conditioned sensitivity matrix S_1. We apply the wavelet transform W and let $A = S_1 W^{-1}$. Then we have $S_1 m_1 = (S_1 W^{-1})(W m_1) = A m_w$ with $W^{-1} = W^T$ because the wavelet function basis is orthogonal. This last property of the wavelet inverse operator is extremely important in the sense that no linear system has to be solved strictly speaking and the application of the related transpose operator is performed naturally. We then minimize $||Am_w - g||_{L^2}$ and we compute the solution in the physical parameter space by trivial inverse-wavelet transform: $m_1 = W^{-1} m_w = W^T m_w$. To solve the minimization problem we add

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

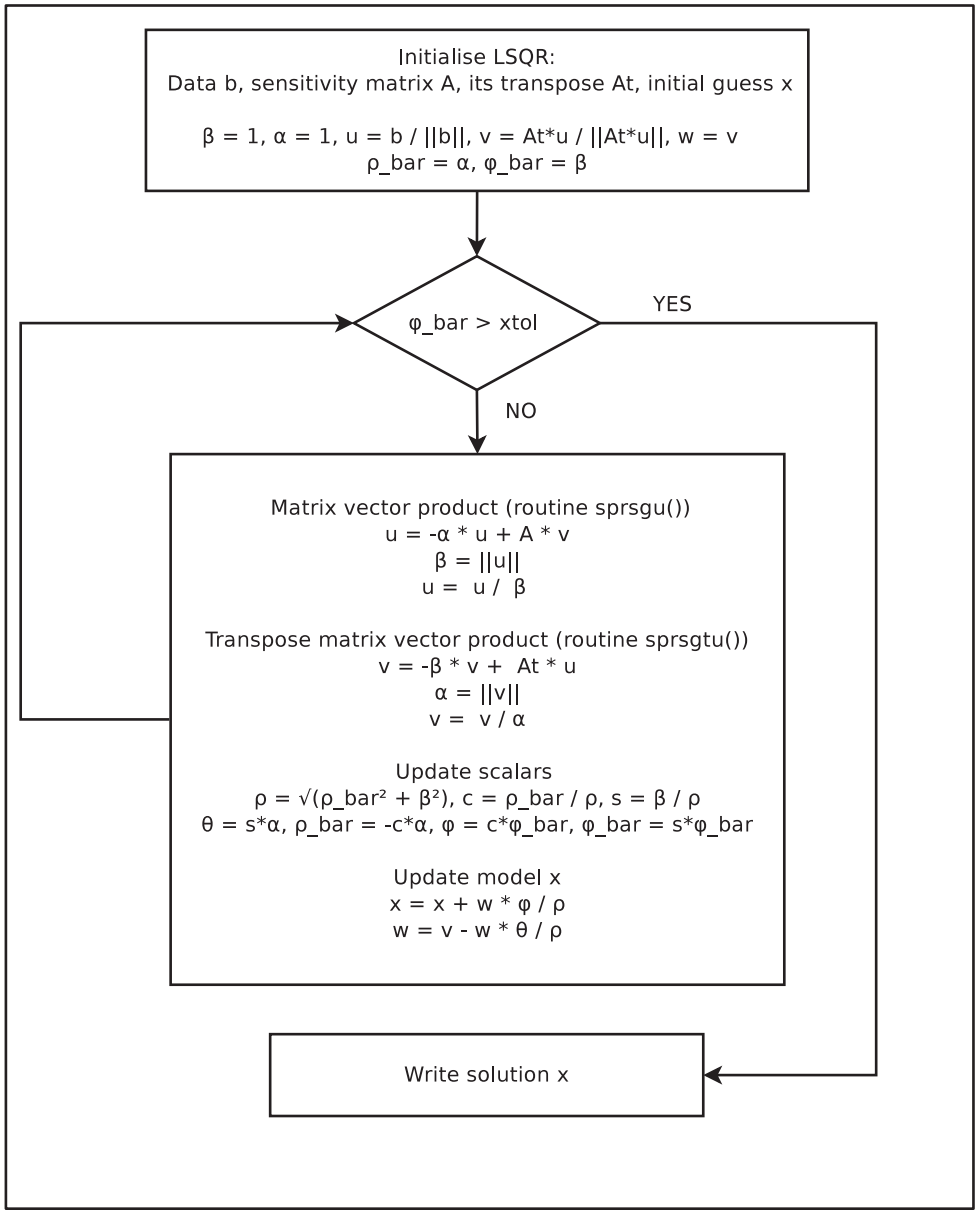

Figure 3. Flow diagram of L^2-norm–based inversion corresponding to the algorithm given in the Appendix.

a stabilizing penalty function on the parameters using a L^2 norm or a L^1 norm as follows:

$$||Am_w - g||^2_{L^2} + \lambda\ ||m_w||^2_{L^2}, \tag{28}$$

or

$$||Am_w - g||^2_{L^2} + \mu\ ||m_w||_{L^1}\ . \tag{29}$$

The effect of the L^2 damping function on the parameter solution is smoother than the effect of the L^1 norm function, which provides sharper parameter contrasts and discontinuities. If expression (28) is considered then a least-square (LSQR) algorithm is used to minimize $||Am_w - g||^2_{L^2} + \lambda\ ||m_w||^2_{L^2}$ according to the algorithm described in the Appendices. If expression (29) is considered then a more complex gradient algorithm is used to minimize $||Am_w - g||^2_{L^2} + \mu\ ||m_w||^2_{L^1}$ according to a different algorithm given in the Appendices, which has been published by Loris *et al.* (2007, 2010) and improved by Beck & Teboulle (2009) in the FISTA (Fast Iterative Shrinkage-Thresholding Algorithm) inversion code. The efficiency of this second L^1-/L^2-norm–based technique allows us to obtain sharper models but does not show better results than those obtained using purely L^2 norm misfit functions. We performed L^2/L^1 hybrid inversions in the synthetic case for different values of the damping parameter μ (inverse of the weight applied to the gradient) that appears in the misfit function (eq. 29). For instance, in Fig. 5 we show three different inverted models obtained after 100 000 iterations of the inversion process for a z-law pre-conditioner, for values of the damping parameter μ of 10^4, 10^3 or 10^2 and for a constant gradient step value of 0.05. Tuning these two parameters is not easy because it depends on the problem under study, on trial and error convergence tests and on the behaviour of the misfit function (expression 29) to be minimized. We can observe that the main structures (the cuboidal structure and the sphere) are recovered and have shapes similar to the LSQR reconstructed solutions but more than 100 000 iterations should then be needed to better reconstruct

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

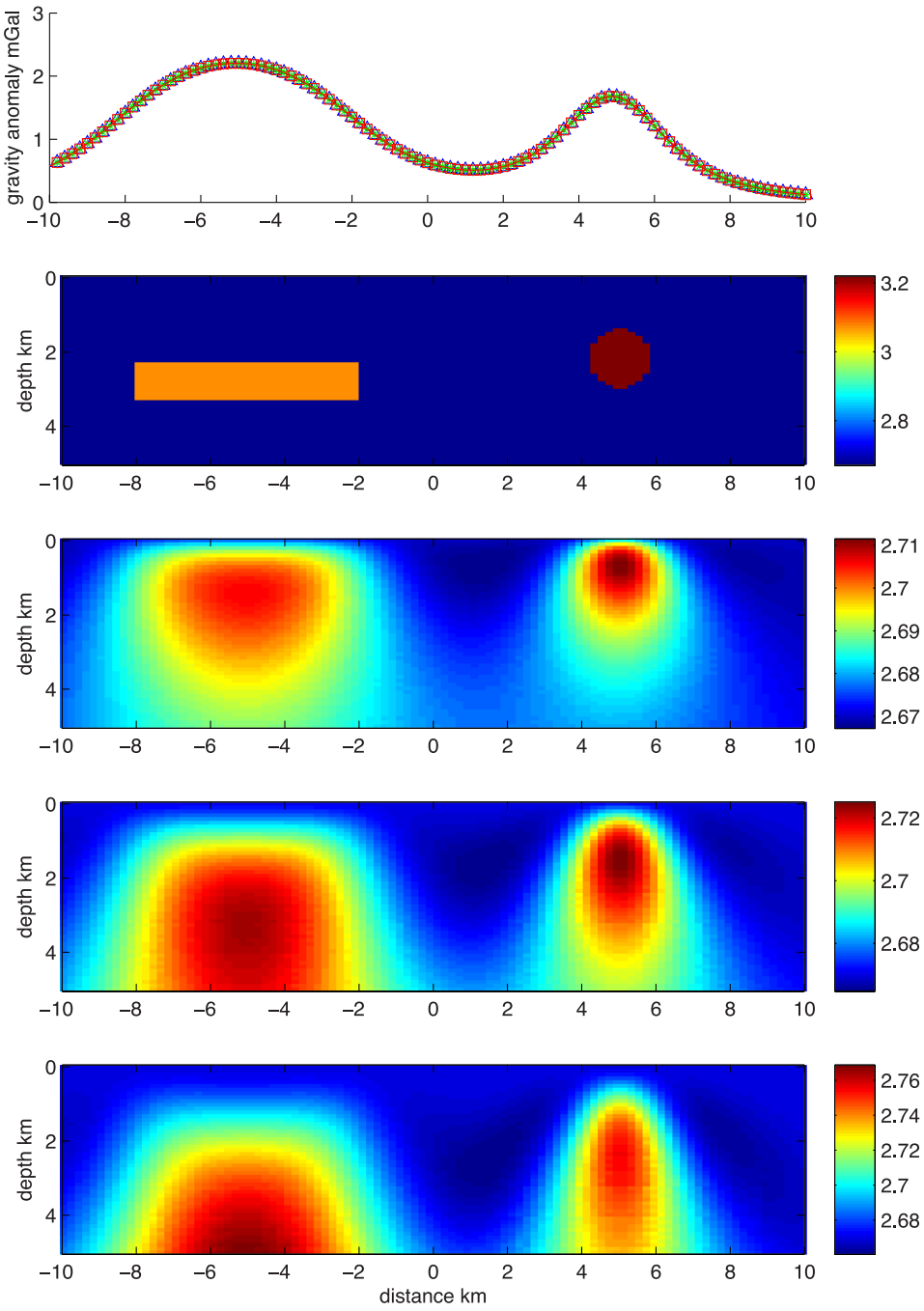

Figure 4. (Top panel) Synthetic response and inverted gravity response using LSQR (L^2 norm) inversion technique, and (top to bottom) true and inverted models for different pre-conditioners $\sqrt{z}$ ($m = 1/2$), z ($m = 1$), $z\sqrt{z}$ ($m = 3/2$). The z pre-conditioner locates better anomalies than the other ones.

all features of the model, which is prohibitive in practice. In Fig. 6, we used the same configuration as in this L^2-/L^1-norm–based inversion but with a $z^{3/2}$-law pre-conditioner. In both cases (z and $z^{3/2}$ laws), solutions show sharp profiles with a strong discontinuity at a depth corresponding to the bottom side of the cuboidal structure. However, in the $z^{3/2}$-law pre-conditioner case, solutions are shifted towards the bottom of the computational domain and are worse than solutions provided using the z-law pre-conditioner. What is interesting here is that L^2-/L^1-norm–based algorithm is able to capture the main discontinuities and can be a complementary tool to L^2-based techniques depending on the problem to be solved. The LSQR technique is much faster, with only 300 iterations and with less computations for a same relative error criterion of 0.01 applied to the exit of the inversion procedure. If the damping factor is decreased then solutions get sharper but the misfit function increases and inversely if it is increased solutions are smoother but the misfit decreases and the technique converges to solutions that are similar to solutions obtained with LSQR-/L^2-norm–based inversion techniques. For these two reasons we will show essentially results for L^2-norm–based algorithms like LSQR. The reasonable choice of the regularization parameter is problem dependent; this will be discussed in the next section on the results obtained on real data from Ghana in western Africa. More optimized L^2-/L^1-norm–based inversion algorithms could be used or developed to accelerate the

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

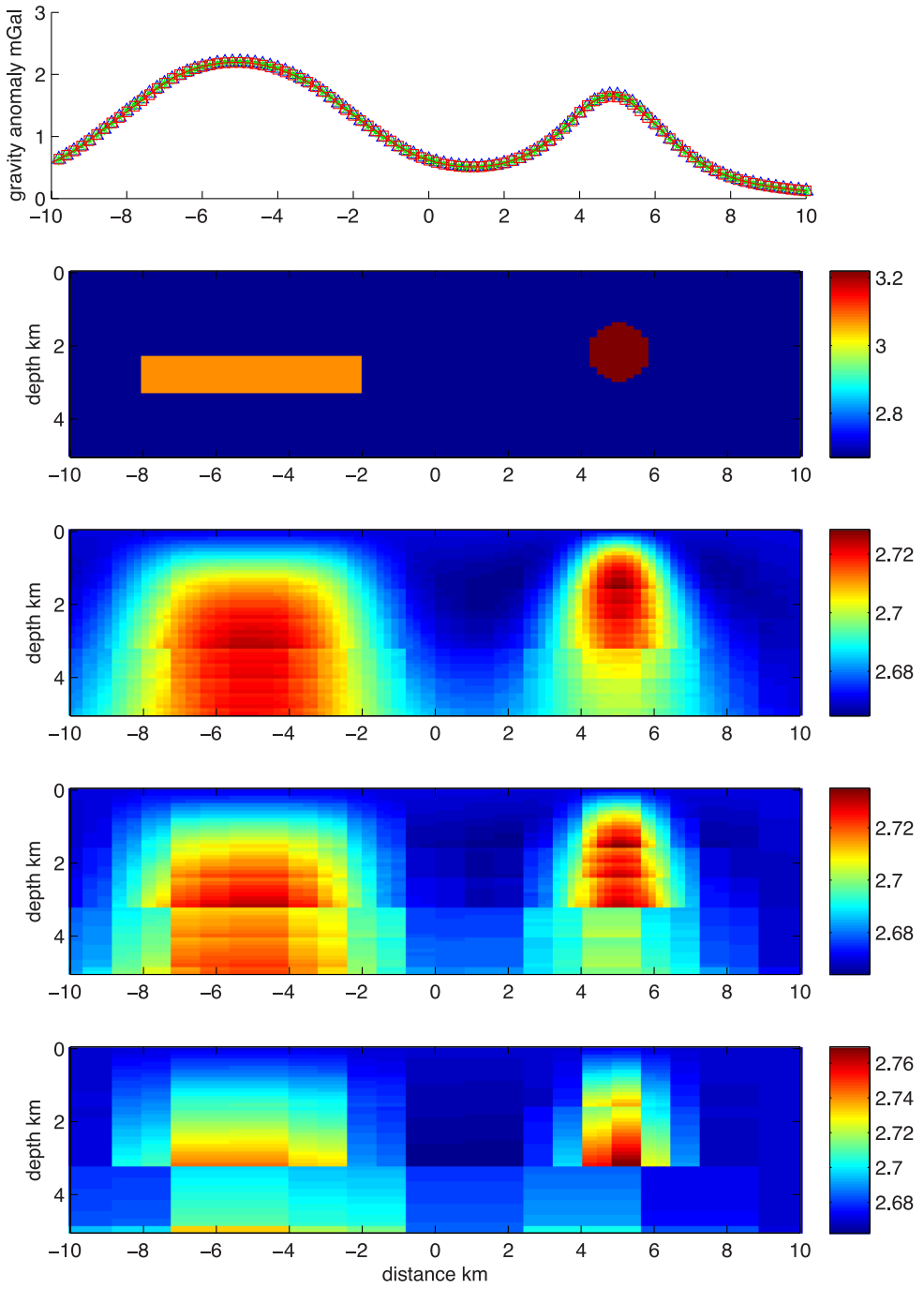

Figure 5. (Top panel) Synthetic response and inverted gravity response using a z-law pre-conditioner and the L^1 norm on the parameters in the same synthetic case as in Fig. 4. (Top to bottom) true and inverted models for different damping factors μ (10^4, 10^3 and 10^2), which is the inverse of the weight factor applied to the gradient, and a gradient step value of 0.05. The solutions are obtained after 100 000 iterations. We can observe that the cuboidal structure and the sphere are reconstructed and have shapes similar to the LSQR reconstructed solutions, but a sharp discontinuity can be observed at a depth corresponding to the base of the true cuboidal structure. This hybrid technique can then be a complement to the L^2-norm–based technique. However, the LSQR/L^2 technique is much faster, with 300 iterations and with less computations to produce the solutions.

convergence of the inversion technique. To achieve this, FISTA-like algorithms or better inversion codes could be introduced in a near future. Such techniques could handle media that contain piece-wise constant physical property distributions if they were introduced in non-linear inverse problems based on interface detections such as in seismics or in joint cross-gradient–based magnetism and gravimetry.

6 APPLICATION TO A REAL CASE: REGIONAL GRAVITY INVERSION OF SOUTHWEST GHANA

To test our inversion code we have chosen southwest Ghana in West Africa where many geological and geophysical studies have been conducted, thus providing us with useful constraints. Our knowledge

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

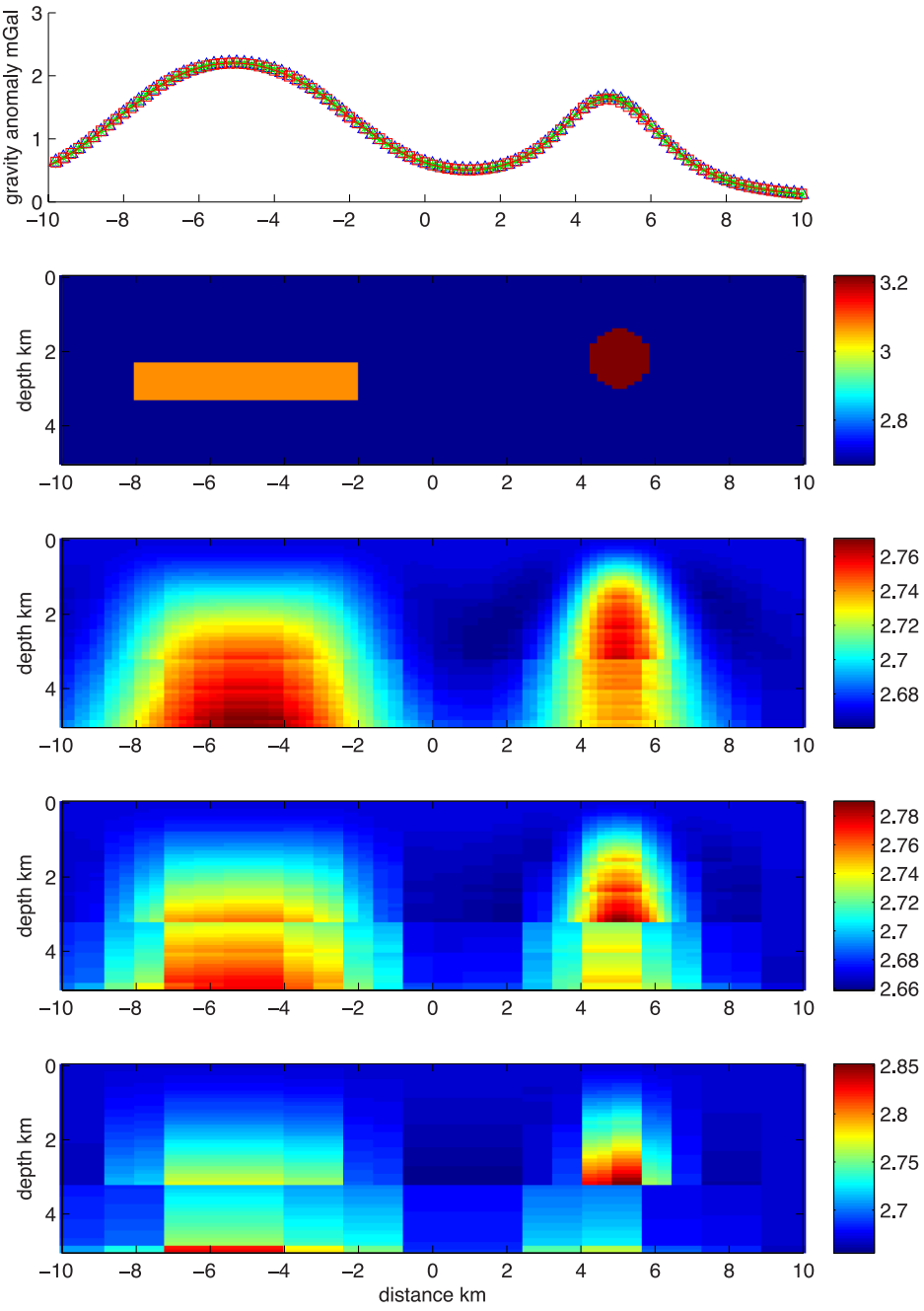

Figure 6. Same L^2-/L^1-norm–based inversions as in Fig. 5 but with a $z^{3/2}$-law pre-conditioner. We can observe again that a discontinuity is still present at the depth corresponding to the lower side of the true cuboidal structure even if all the reconstructed features are shifted towards the bottom of the computational domain. The L^2-/L^1-norm–based algorithm seems to be capable of capturing strong discontinuities but not clearly the whole main features such as the sphere in this particular case. This depends on the problem under study and on the linearity of the problem. For more constrained models and for non-linear inversion techniques based on reflections and surface capturing (seismic or joint cross-gradient–based magnetics/gravimetry inversions), piecewise constant models could be better retrieved using such hybrid misfit functions.

of the geology of this area, including *in situ* density measurements (Perrouty *et al.* 2012) and other geophysical observations, allows us to propose an acceptable prior model for the gravity data inversion. In addition, the gravity data available for this region provide a homogeneous data set with a significant number of data that can be inverted to improve the geological modelling of subsurface crustal structures. In this paper we have chosen to assign constant density values to each lithostratigraphic unit to simplify the task of visualizing the differences between prior and inverted models. The ability to vary input densities according to a given statistical law is straightforward and would clearly be made part of a complete inversion workflow but would not really add anything to our proof-of-concept tests of the inversion scheme.

6.1 Geological setting and gravity pattern of southwest Ghana

The area considered for this study is located in the southeast part of the West African craton, in Ghana (Fig. 7). This craton mostly consists of a series of greenstone belts oriented N–S or NE–SW,

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

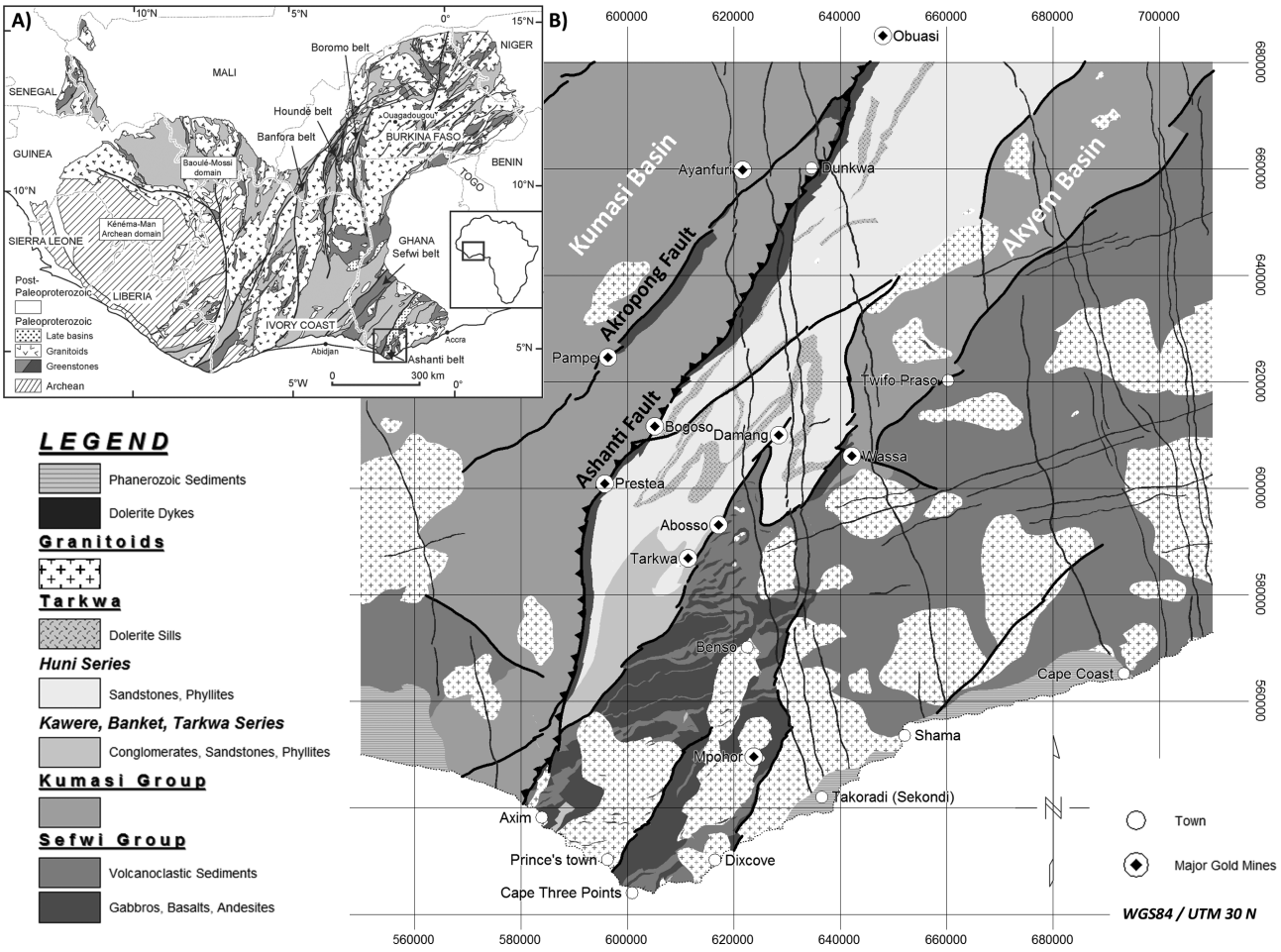

Figure 7. Map of southwest Ghana with inset map showing the southern part of the West African Craton (WAC). Northern and southeastern parts of the craton are formed by series of palaeoproterozoic greenstone belts, such as the Ashanti Belt in southwest Ghana (black square) (modified after Milési, J.P., Feybesse, J.L., Pinna, P., Deschamps, Y., Kampunzu, H., Muhongo, S., Lescuyer, J.L., Le Goff, E., Delor, C., Billa, M., Ralay, F., Heinry, C., 2004. Geological Map of Africa 1:10 000 000. SIG Afrique. Project. http://www.sigafrique.net).

such as the Ashanti Belt in Ghana, which were formed during the Eburnean orogeny: a major Palaeoproterozoic tectonic event (Bonhomme 1962). The area presents a significant number of world-class gold deposits often located along faulted or sheared contacts (including the Ashanti Fault) between the major stratigraphic groups. 3-D forward modelling of the Ashanti Belt using gravity data provides a mean of exploring the subsurface geometries of the gold-rich faulted contacts and which represent an aid to mineral exploration and aid in the interpretation of the tectonic evolution of the area.

The geology of the Ashanti Belt is dominated by three palaeoproterozoic stratigraphic groups and by granitoids intrusions, overlain along the coast by Phanerozoic sediments. The Sefwi Group (Adadey *et al.* 2009) consists of metamorphosed Birimian (2.2–2.1 Ga) volcanic rocks interbedded with volcanoclastic sediments and micaschists in the Cape Coast area. The Kumasi Group consists of Birimian metamorphosed sedimentary rocks (mainly phyllites) deposited in the Kumasi Basin and in the Akyem Basin (Fig. 8). The Tarkwa Group is a metasedimentary basin located on top of the Sefwi Group. It represents the bulk of the Ashanti Belt. Granitoid intrusions occurred in two stages called Eoeburnean (2190–2150 Ma) and Eburnean (2125–2000 Ma) by Perrouty *et al.* (2012), following the terminology of De Kock *et al.* (2011). These granitoids have varying compositions with tonalite and leucogranite endmembers. The distributions of measured densities for each lithology are given in Perrouty (2012).

The gravity surveys carried out in this region within the last decades provide useful information on the gravity anomalies that enable us to delineate the broad gravity patterns of this area and relate them to the main geological units. Fig. 8 shows the free-air gravity anomaly of southwest Ghana derived by the Bureau Gravimétrique International (BGI; International Association of Geodesy 2012) from the $2.5' \times 2.5'$ global gravity model EGM08 (Pavlis *et al.* 2008). It shows two main zones separated by the major Ashanti Fault. To the east, a series of gravity highs (amplitude up to 50 mGal) correlate with the presence of thick high-density metavolcanic domains (Sefwi Group) within the Ashanti Belt. To the west, a relative gravity low (amplitude within 15 to 30 mGal) is associated with the presence of Kumasi Basin metasediments. The major gravity gradient between these two domains correlates with the Ashanti Fault marking the eastern border of the Ashanti Belt. Moreover, most of the gravity patterns observed in this region are elongated along the main tectonic lineaments and geological units, indicating that density contrasts between geological structures are significant and that gravity data inversion is an appropriate tool to investigate the deep structures.

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

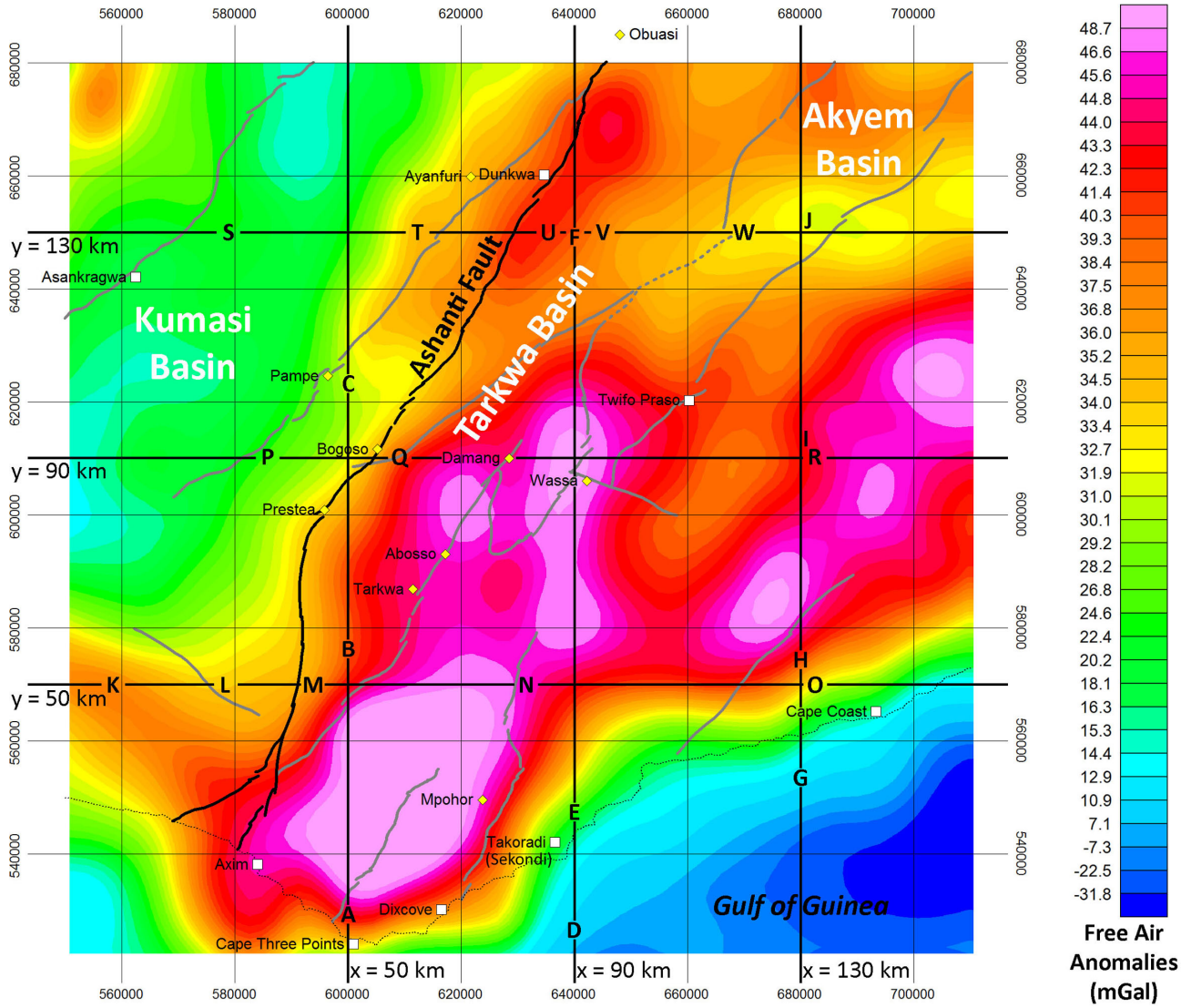

Figure 8. Free-air gravity anomaly map of southwest Ghana derived from surface gravity measurements (source: Bureau Gravimétrique International). Coordinates are WGS84/UTM30N. Grey lines represent major faults. Sections and locations highlighted in Figs 11–13 are also displayed. The Ashanti Fault (in black) limits two major gravity domains: the Sefwi Group to the east and the Kumasi Group to the west (see text and Fig. 7 for details).

The remarkable density contrast between the two sides of the fault as modelled in 2-D sections by Barritt & Kuma (1998) and Hastings (1982) led to their suggesting the presence of a tectonic contact with massive post-sedimentation vertical movement. However, kilometric scale movement along this major tectonic fault is contradicted by the absence of any metamorphic gap between the two sides of the fault and by structural evidence (Allibone *et al.* 2002). On the eastern side of the Ashanti Fault, the Tarkwaian Basin formed on top of metavolcanic rocks from the Sefwi Group, which causes a slight reduction of the gravity anomaly, mostly perceptible on forward-modelled 2-D cross-sections published by Barritt & Kuma (1998) and Hastings (1982).

6.2 Gravity data set and prior density model

Using the gravity data sources available in the BGI database from land surveys for Ghana and the surrounding countries (Ivory Coast, Burkina Faso, Togo, Benin) as well as from marine surveys for the oceanic part (Pontoise *et al.* 1990), we have generated a regular grid of the free-air gravity anomaly projected into UTM (WGS 84, Zone 30N) and interpolated with a spatial resolution of 5 km. We have extended the data set beyond the area of study to avoid edge effects. We thus obtained a regular gravity grid of 316 × 316 interpolated data (i.e. 99 856) that we used as the constraining gravity observations for the inversion process.

As mentioned earlier, our knowledge of the geology of this area allows us to propose a 3-D density model for the whole studied area constrained by geophysical observations and by *in situ* density measurements (Perrouty *et al.* 2012). The 3-D geological model of the southern Ashanti Belt was built in several steps, using field structural data, geological contacts from maps and cross-sections constrained by geophysical and petrophysical data (a–c in Fig. 9). A hybrid methodology combining two modelling systems was used to build the Ashanti Belt prior 3-D model: Geomodeller (www.geomodeller.com) allowed us to calculate implicit interfaces (Fig. 9d) between geological units with orientations constrained by structural measurements using a co-kriging approach (Lajaunie *et al.* 1997; Calcagno *et al.* 2008). Calculated surfaces

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

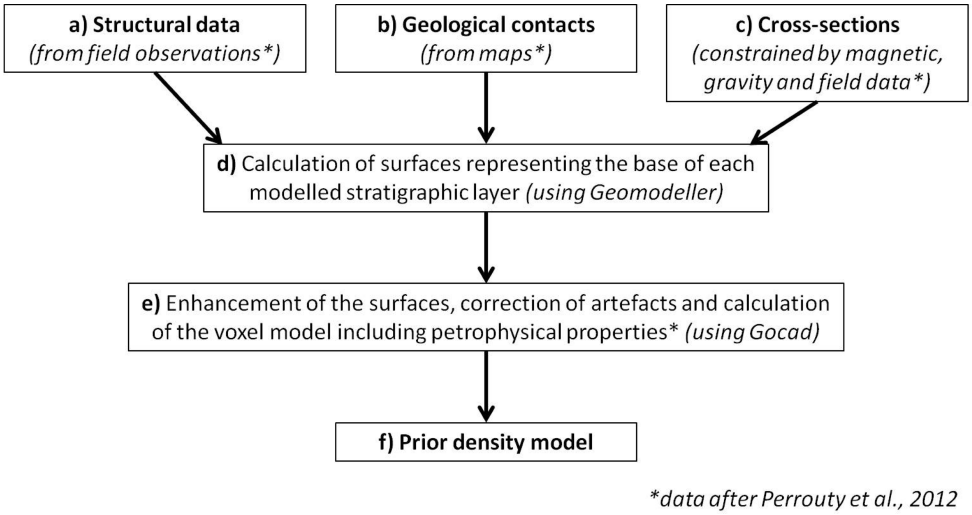

Figure 9. Simplified diagram illustrating the steps used to build the Ashanti Belt prior model: (a) field structural data (Perrouty *et al.* 2012), (b) geological contacts from maps and (c) cross-sections constrained by geophysical and petrophysical data were used to calculate the interfaces between the geological units (d) with Geomodeller. These surfaces were then tidied up with Gocad and a voxel model was generated (e), which corresponds to the prior density model (f).

were then imported into Gocad (http://www.pdgm.com) and defined by node locations [discrete smooth interpolation/DSI method of Mallet (1992)]. Each surface consisting of triangular facets linking neighbour nodes could then be modified to enhance the model by correcting artefacts (Fig. 9e) and by adding geological detail that was not modelled with the Geomodeller software, such as the dolerite dykes (assuming a thickness close to 200 m). We wanted to include the dolerite dykes in the model given their very strong signature in the aeromagnetic data, although the low resolution of the gravity data makes them unimportant for gravity calculations. The Ashanti Belt model is composed of 19 stratigraphic surfaces [based on the stratigraphy proposed by Perrouty *et al.* (2012)], six surfaces for the intrusive rocks, including two dyke generations and nine major faults. Polyhedra defined by the intersecting surfaces bordering geological units and faults were used to define the voxel model, generated with Gocad, and corresponding to the prior density model (Fig. 9f). A constant density value was assigned to each voxel within a single lithostratigraphic unit, based on representative density measurements of Perrouty *et al.* (2012) and Hastings (1978, 1982).

To build the final model, we thus obtained a density distribution model composed of 890 × 890 (horizontally) by 68 cubes (vertically) with a resolution of 200 m. The whole sensitivity matrix has around 5.37 trillion elements (the number of data multiplied by the number of density parameters) and this inversion problem suits very well the huge computing power of multi-CPUs or multi-GPUs. For an error of around 0.5 per cent and a compression ratio of around 99.5 per cent, the inversion is performed in less than 50 min on 32 CPU processes and less than 5 min on 192 GPU processes.

To smooth some high-frequency variations in the gravity signal, which cannot be resolved by our geological model, we have continued the interpolated gravity anomaly upwards to 510 m above sea level, corresponding to the highest elevation of the area. This also enabled us to evaluate the effect of topography in the inversion process. We performed inversion tests taking into account (or not) the topography but the differences in the results of the inversion in terms of density distributions were not significant. This is probably due to the relatively flat topography of the region (mean elevation and standard deviation are, respectively, equal to 183 and 127 m and are lower than the horizontal spacing of the grid and of the order of two voxels in the vertical direction) and the smoothing effects of the upward continuation of the data set as well as the filtering of the model and topography by wavelet compression.

6.3 Choice of the regularization parameter

An important issue is how to choose the regularization parameter λ in the misfit function (28). To define a reasonably optimized value we have performed several inversions with different values and we have built a L^2-type curve (Hansen 2001) as in Fig. 10.

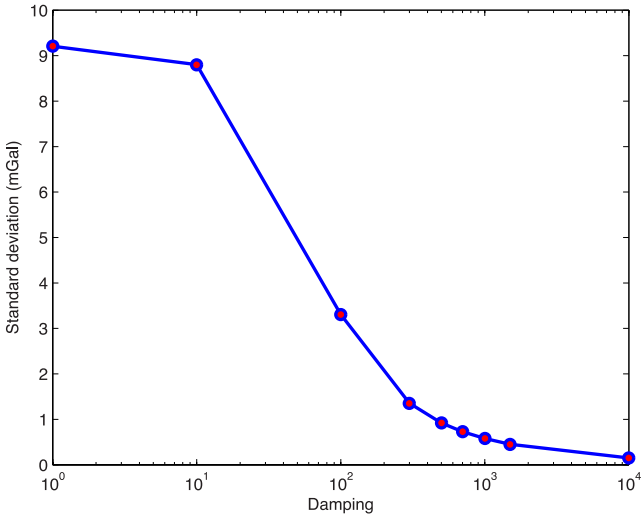

Figure 10. L^2 curve to select the regularization parameter λ. We represent the variance error on the data obtained after different inversion runs for different λ damping values from 1 to 10^4. As discussed in the text, the relevant models are obtained for the computed anomaly data, which have both low data variance error values and physical meaning. These models correspond to λ values lying in the vicinity of the inflection point of the L^2-type curve (here between values of 100 and 10 000), that is, the zero of its second derivative. In the range of λ values between 500 and 1000, the corresponding value of the second derivative is divided by more than 50 with respect to the second derivative related to a λ value of 100. Values of λ typically between 500 and 1000 are thus acceptable.

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

This curve gives the variance between observed and calculated data according to the damping parameter. Optimized values are located in the vicinity of the inflection of the L^2-type curve, somewhere between $\lambda = 100$ and $\lambda = 10\,000$. The absolute value of the second derivative of the L^2-type curve decreases quickly from 1.77×10^{-4} for $\lambda = 100$ to 3.08×10^{-6}, 8.97×10^{-7} and 3.044×10^{-7} for $\lambda = 500$, 700 and 1000, respectively. For values of λ of 500, 700 and 1000, the ratio of the corresponding second-order derivative value with respect to $\lambda = 100$ is 57.4, 197.23 and 581.23. These ratios of the second-order derivative being low enough, optimized values of λ can be selected in a range of 500 and higher. An upper bound of the λ values should be chosen based on physical arguments that we discuss later.

In Fig. 11, we show the observed gravity anomaly and the reconstructed one obtained using a four-scale fourth-order Daubechies compression operator, a 0.1 per cent compression error on the kernels and three different λ regularization parameters corresponding to the L^2 curve. The *a priori* model has been built using the VpMg inversion software package, which inverts anomaly data by solving the main geological unit geometries according to initial densities collected in the field. We then apply our inversion procedure to that new model. For a value of 10 000, the damping λ is too strong even if the residual anomalies are very low (smaller than around 1 per cent). Spurious ripples appear in the anomaly and these solutions are therefore not satisfactory. For values of the regularization factor as low as 100, solutions produce anomaly residuals, which are too large with maximal values around 10 per cent and the inversion has not converged adequately, with misfit far from zero. For the reasons given, λ values in the range between 500 and 1000 will not produce much difference in the solutions. In particular, when arbitrarily choosing a value of $\lambda = 1000$, we observe that solutions are more relevant with more physical meaning and with few maximum residual anomaly values (around 5 per cent in absolute value). Indeed, for this value an important part of the residuals corresponds to structures presenting sharp discontinuities such as faults or coastal lines, which are geologically relevant. We have therefore used a regularization parameter $\lambda = 1000$ for all the subsequent inversions.

6.4 Results of the inversion process

Based on a previously built 3-D model of the southern Ashanti Belt, we extracted six cross-sections at distances of 50, 90 and 130 km from the southwestern corner of the study area, parallel to the N–S and E–W axes (Figs 8, 12 and 13). These cross-sections are produced after inversion of the gravity data set and for parameters discussed in the previous two subsections. They highlight some difference between the density and geometry of the prior model and the result after the new inversions. Changes of the model by the inversion process reflect uncertainties on our 3-D knowledge of the area and due to the inherent ambiguity in the gravity data, we can interchangeably reinterpret the differences in terms of changes in model geometry or as density variations within geological units.

6.4.1 Modification of density or geometry within the Sefwi Group

As described in the Introduction, the 3-D model, which includes the Sefwi metavolcanic group was initially modelled using the Geomodeller software package using contacts and bedding orientations observed in the field. The modelling process extrapolates the expected location of the layers from these data (Calcagno *et al.* 2008). Inversion of gravity data allows us to better constrain the predictions produced by this modelling process. In many areas (see Figs 12 and 13, labels C, F, K, M, T and V), metavolcanics or metavolcanosedimentary layers modelled within the Sefwi group effectively disappear after inversion. In other cases, modelled geological unit densities have been overestimated, or underestimated, such as in B or Q, respectively. Particularly, in B, the density of the metabasalts layer clearly decreases after inversion although the orientation of the densest area (in red) is preserved.

6.4.2 Granitoid intrusions

Granitoids intrusions were modelled in the prior model using their border locations and associated dip orientation (again using the Geomodeller package). Field observations do not allow us to constrain their depth precisely. Consequently most of the intrusion depths can be revised relative to the prior model using the information shown on the inverted sections. For example, the depth of intrusions A and E have been overestimated although intrusion H seems to be significantly deeper than previous estimates.

In other areas (D, G, I and J) dense bodies appear and may correspond to buried granitoid intrusions, using an average granitoid density in southwest Ghana around $2.8\,\mathrm{g\,cm^{-3}}$, or smaller even denser bodies such as gabbros. A particularly interesting result of this gravity inversion concerns location N where the massive Sekondi tonalite shows a significant dip to the west of the profile after inversion, although this intrusion was previously supposed to be subvertical in the prior model. This new proposed orientation indicates that the possibly gold-mineralized Dixcove Fault, which represents the eastern border of this intrusion is also dipping to the west. Modifications of the shape of granitoid intrusion are also shown in R where the initial orientation and depth seem to be preserved on the west, in contrast with the east that is less deep. Note that the presence of high-density ($3.0\,\mathrm{g\,cm^{-3}}$) but thin dolerite dykes also disturb the result of inversion in this area, but are really too fine a scale to be usefully reinterpreted following inversion.

6.4.3 (Meta)sedimentary basins

The Kumasi Basin (P and S), the Tarkwa Basin (U) and the Akyem Basin (W) densities and shapes are little modified by inversions. This agrees with field observation indicating a stratigraphy characterized by alternating phyllites and sandstones with little variation in density. In L, a moderate-density body appears at the bottom of the Kumasi Basin, close to the contact with the highly gold-mineralized Ashanti Fault. This body may correspond to a Sefwi group lithology (metavolcanosediment?), thought to constitute the basement of the basin (Perrouty *et al.* 2012). Phanerozoic sediments are present in few areas in southwest Ghana and mostly occurred offshore. Their density is not constrained by measurement and was only estimated from the presumed lithologies to be around $2.5\,\mathrm{g\,cm^{-3}}$. Gravity inversion seems to agree with this density in most of the cross-section. However, the presence of a dense body below these sediments in O suggests a compensation of underestimated density or the presence of granitoids intrusions below the basin.

7 IMPLEMENTATION AND PERFORMANCE OF THE PARALLEL INVERSE PROBLEM SOLVER

To study the performance of the code, we performed different tests on the hybrid multi-CPUs/multi-GPUs cluster 'Titane' of the French supercomputing centre GENCI/CCRT/CEA, which consists of 1596

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

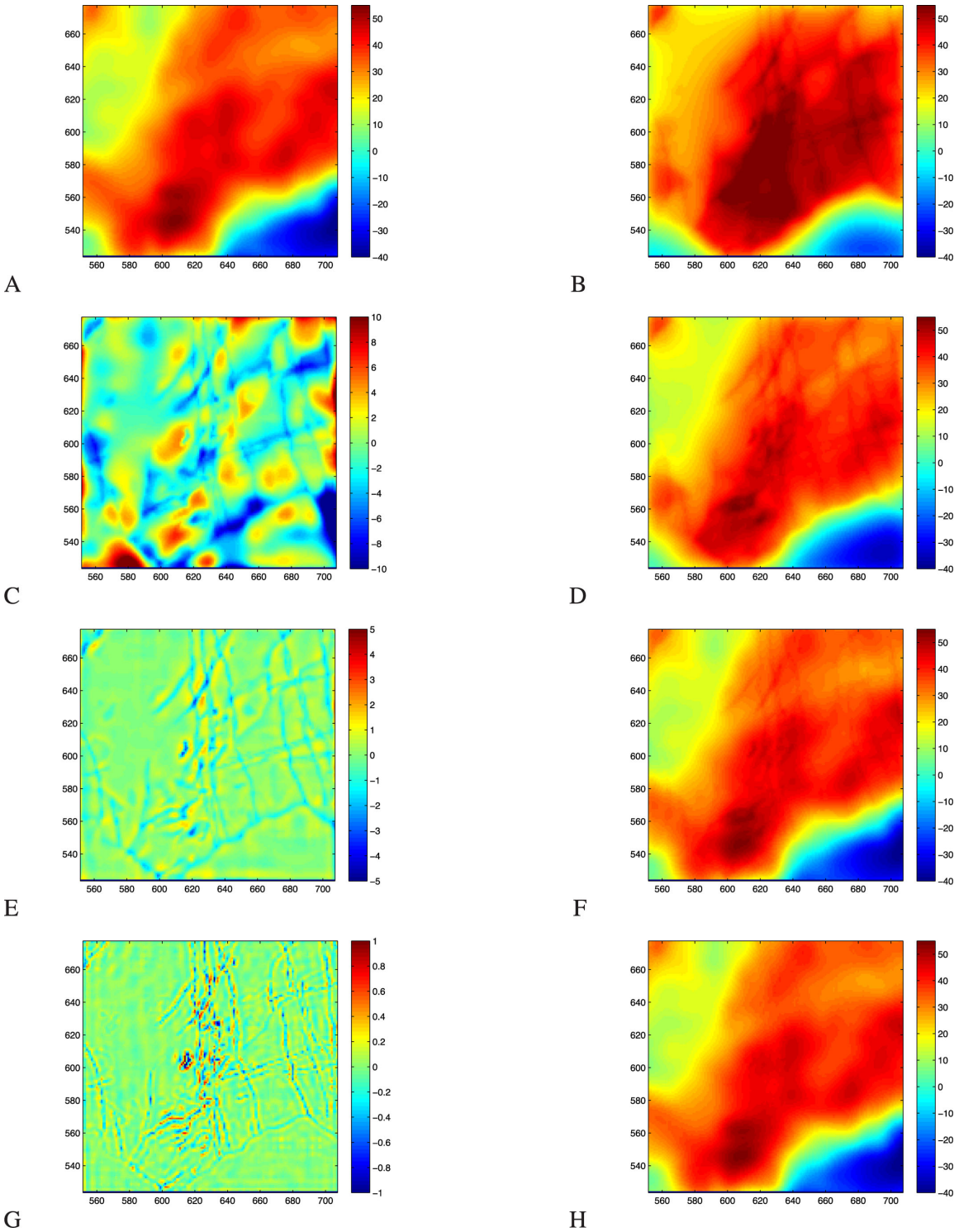

Figure 11. Observed gravity anomaly (A), anomaly for the *a priori* model built with Geomodeller and VpMg (B), anomaly errors and computed anomaly data for $\lambda = 100$ (C and D), $\lambda = 1000$ (E and F) or $\lambda = 10\,000$ (G and H) after inversion process. As discussed in the text, errors are too high for a 100 value (around 10 per cent). For a 10 000 value, data errors are lower than 1 per cent but too many high-frequency ripples appear and are widespread at the surface. For a 1000 value, errors are lower than 5 per cent and maximum absolute errors are located mainly along coastlines and along faults and principal material discontinuities. A λ value of 1000 is therefore a good compromise between low variance data error, low misfit function and physical computed anomaly data. This value of λ lies in the vicinity of the L^2-type inflection point.

Bull Novascale R422 servers with an Infiniband network. Each server has two Intel Xeon 5570 nodes with 2.93 GHz quadri-cores each and a total memory of 24 Gbytes. For the GPU part of the machine, 96 CPU servers are connected to 48 GPU servers by PCI Express-2 buses. Each NVIDIA Tesla S1070 GPU consists of four GPUs with 4 Gbytes of memory on each. We will use from 32 to 192 GPUs in one series of runs and from 32 to 2096 processor cores in another series of pure CPU runs.

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

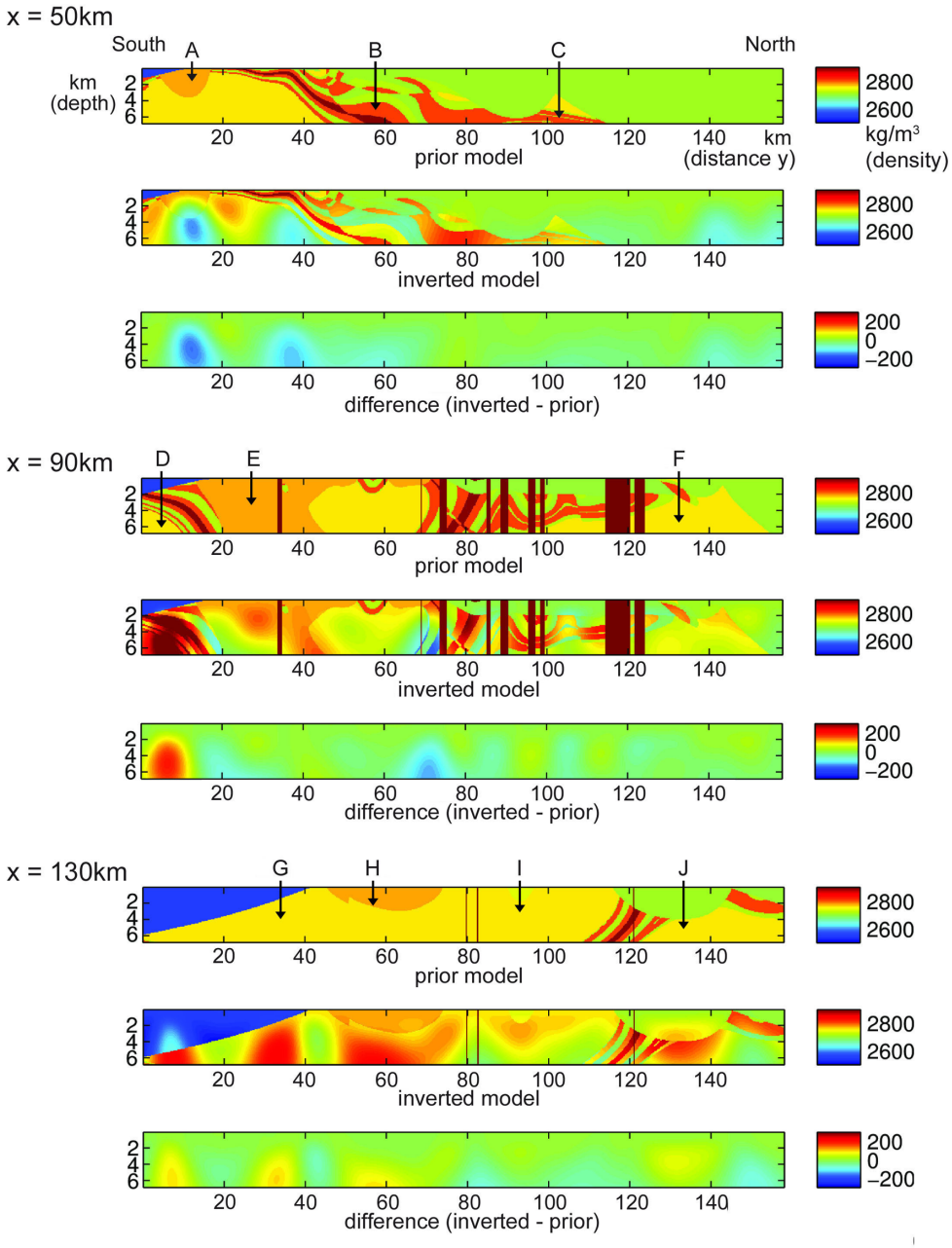

Figure 12. Sections along the *Y* axis from the prior and inverted density model. Labels A to J highlight areas discussed in the text. See Fig. 8 to locate the sections.

7.1 Multi-CPU clusters with message passing

To accelerate the inversion algorithm further (see Fig. 3), a parallel version is implemented by dividing the gravity kernel according to data rows over *Nproc* processes using an MPI Cartesian communication topology. Each process contains $Ndata/Nproc \times NX \times NY \times NZ$ kernel submatrices. The entire density grid is loaded on each process by reading once and for all the properties of the *a priori* model of southwest Ghana in sparse format. Each node then computes the partial set of anomalies using a matrix–vector product for each local set of data row-wise submatrix. In this way, the compressed sensitivity kernels can then be computed, split and stored in each process. Then, for each row the sub-kernel matrices are compressed independently according to the parameter columns and according to the thresholding process discussed in Section 4. As can be seen in Fig. 14, these kernel calculations and compressions can take more than 50 per cent of the CPU time. The least-square inversion process then begins. An initialization process needs therefore to set initial values to a variable u and an orthogonal variable v by performing a normalization on u and a transpose matrix–vector $A^t u$ product followed by a normalization on v. Normalization needs global reduction operations over all the processes. At each iteration of the inversion norms of u and v as well as Av and $A^{\mathrm{T}}u$ products are computed in parallel and using an MPI reduction operation. Finally the whole density anomaly solution is computed and stored on the master node by calling a global communication MPI_GATHERV function. The iterative loop of the inversion process then stops once a relative error criterion is reached.

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

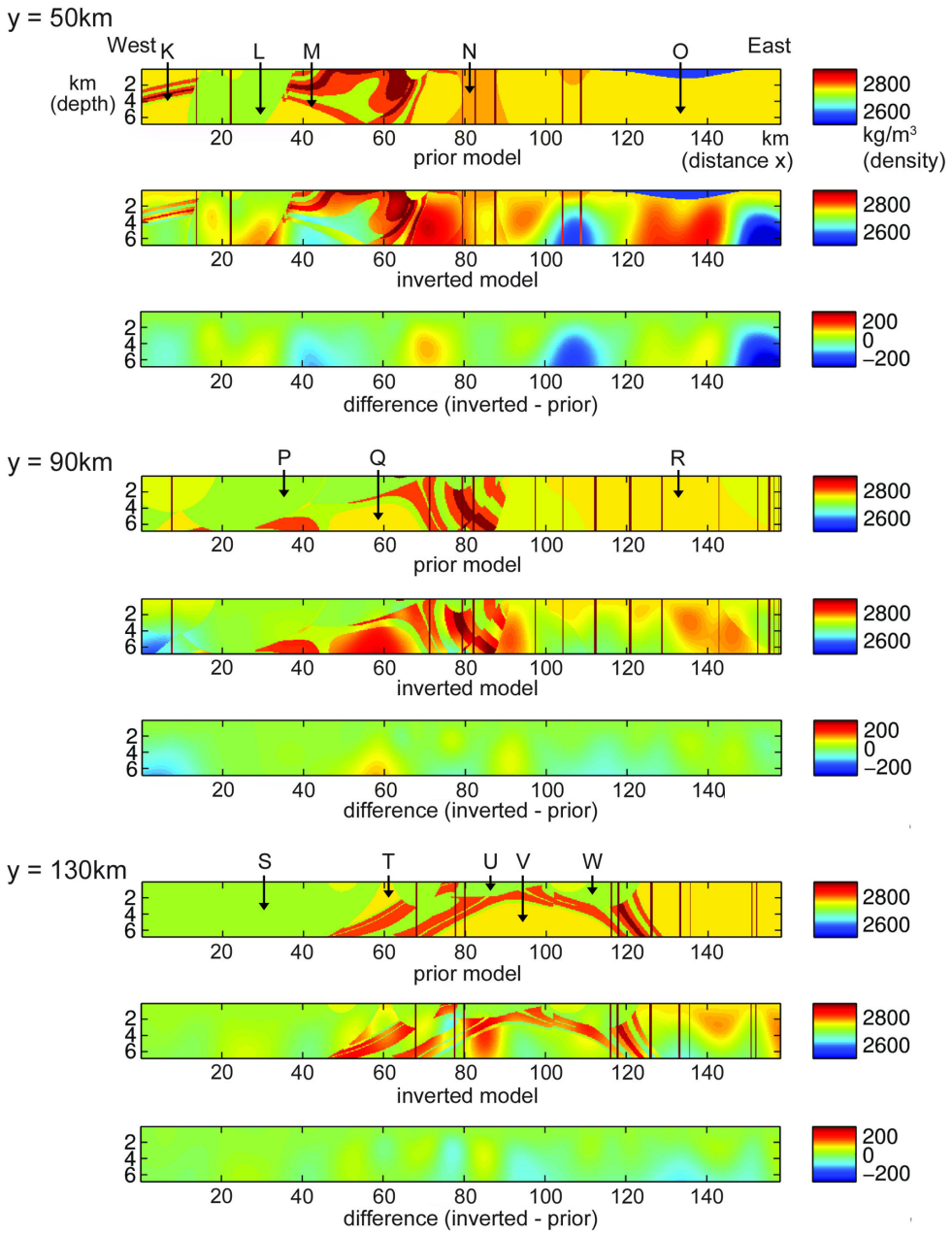

Figure 13. Sections along the X axis from the prior and inverted density model. Labels K to W highlight areas discussed in the text. See Fig. 8 to locate the sections.

One of the major challenges we have faced is the treatment of the transposed matrix A^{T} over all the processes because at each iteration of the inversion process all the processes need to communicate, which reduces the efficiency of the parallel algorithm. We thus decided to build the transposed matrix A^{T} once for all before starting the inversion to reduce the communication bottleneck. Another challenge has been to reduce the initial load imbalance between processes, owing to the fact that the non-zero matrix coefficients obtained after compression are not equally distributed on all the processes. To overcome this issue, in a second pre-processing step we redispatch some coefficients over all the processes to obtain similar number of elements and to perform similar numbers of floating operations such as matrix–vector products. In addition, we perform a renumbering of sensitivity matrix elements to address them as linearly as possible in the global memory of each processor, thus trying to improve pre-fetching in the cache of the processors.

Doing so, we can solve 3-D models with as many as 54 million parameters and 99 856 data in less than 50 min on 32 CPU cores and less than 6 min on 256 CPU cores (see A, C and E in Fig. 14) for the whole inversion process (reading the data, global sensitivity kernel construction, four level compression and inversion). Calculation of kernels and their compression take around 35 min and LSQR inversion process around 14 min on 32 processes while they take, respectively, around 5 and 2 min on 256 processes. CPU timings are therefore divided by a factor around 7.1 when using 256 processes instead of 32. The scalability can be considered as very good for such a problem that is generally very difficult to scale, as can be seen in Fig. 14(E) where scalability lies between 95 and 101 per cent. Values higher than 100 per cent are due to

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

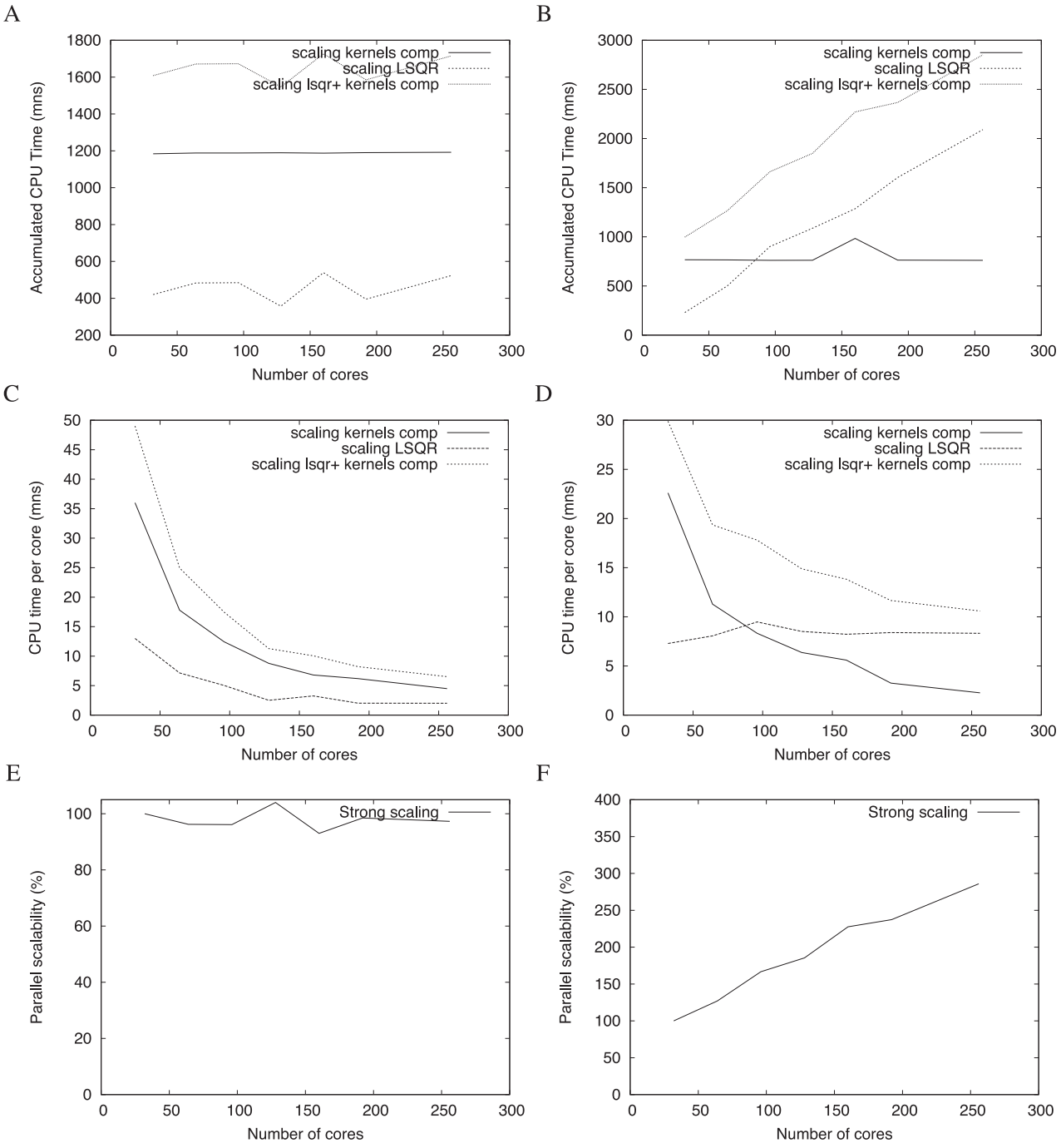

Figure 14. Performances of huge (left-hand side) (890 × 890 × 68 parameters and 99 856 anomaly data) and medium (right-hand side) (178 × 178 × 68 parameters and 1493 data) sensitivity matrix inversion using L^2 least-square algorithm on a pure CPU cluster. In the huge case strong scaling is almost perfect up to 256 processes. For 32 processes, the cumulated CPU time per core is around 49 min while it is around 6 min for 256 processes. For eight times more processes we compute the same problem around 8.3 times faster (C). Conversely, for the medium-size case, scaling becomes worse with an increased number of processes (B and F). Indeed, in (D) the CPU time per core becomes almost unchanged and reaches a near constant value due to the fact that the inversion is dominated by global communications. In the huge case (left-hand side), calculations are so big that the time spent in communications is much less than the time spent in kernel calculations and inversion: communications then represent less than 7 per cent of the total.

errors lower than around 2 per cent, which can be induced by system noise or particular hardware multiprocessor exploitation. Accumulated times are more significant from that point of view than times for each process since compression kernel calculations are almost constant for all processes with little fluctuation around an almost constant average value of 400 min for the inversion process (see Fig. 14 A).

We now present another test with smaller kernels and show performance levels in Fig. 14 (sketches B, D and F). We solve a 3-D model with around 54 million parameters and 1493 data in less than 30 min on 32 CPU cores and less than 10 min on 256 CPU cores (see B, D and F in Fig. 14). We observe that in this case performances are very poor because there are not enough calculations per node compared to MPI communications between CPU cores. The

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

inversion process takes between 7 and 8 min whatever the number of processes because MPI_REDUCE-like operations and related global interprocess communications are getting longer and longer in time with the increasing number of processes. Conversely the construction and compression of the sensitivity matrices are made in each process without significant global communications. Global communications are no longer overlapped by calculations during the LSQR inversion process, the number of floating operations being too small in each process. The only parts that scale well are the calculation of the sensitivity kernels and their compression.

7.2 Multi-GPU clusters

To port the main compute-intensive part of the code to GPUs, we make heavy use of the CUBLAS and CUSPARSE libraries (NVIDIA Corporation 2011). CUBLAS is a free implementation of Basic Linear Algebra Subprograms (BLAS, Dongarra *et al.* 1988) built by NVIDIA (Santa Clara, CA, USA), the manufacturer of the GPU boards that we use in this study, on top of their NVIDIA CUDA programming language. CUSPARSE contains a set of basic linear algebra subroutines that are designed to handle sparse matrices and that are also implemented by NVIDIA on top of their CUDA programming language. Since the main part of our algorithm almost entirely consists of vector–vector operations such as the computation of the norm of a vector, or products and/or sums between vectors, followed by sparse matrix–vector operations, porting the code turned out to be relatively straightforward. Testing was simple as well, since BLAS and sparse matrix libraries containing the same standard subroutines were available for our original Fortran code as well, and thus it was easy to verify that the original Fortran code, the Fortran code converted to BLAS and sparse-matrix library calls, as well as the GPU/CUDA code based on CUBLAS and CUSPARSE all gave the same results down to roundoff accuracy. Regarding optimizing the MPI message-passing operations, we resorted to using the so-called memory pinning option of CUDA (NVIDIA Corporation 2011) on the compute nodes hosting the GPU boards to accelerate memory copies of the MPI messages between the CPUs and the GPUs (both ways) by preventing memory page swapping in the Linux system of these compute nodes.

Regarding the level of performance obtained, in Figs 15 and 16 we show the timings of the main parts of the code for different numbers of CPUs and GPUs. We obtain acceleration factors around 5.7 on 32 GPUs (280 cumulated minutes) and 5.5 on 160 GPUs (323 min) when compared to 32 and 160 CPU cores for the whole code and for the first kernel test (Fig. 15). We obtain acceleration factors of around 4.9 on the kernel calculation and compression. In the inversion process we reach factors of 14 for 32 GPUs and 7.1 for 160 GPUs. The decrease of the time with the number of processes is not linear because MPI_REDUCE communications between all the CPU hosts increase. In addition, to perform these global reduction operations, host-to-device and device-to-host copies of some arrays must be performed and have also a time cost between 400 and 700 clock cycles. The reductions, performed by a device-to-host copy, a sum based on a MPI_ALL_REDUCE global communication between CPUs and then a host-to-device copy are difficult to overlap by computations. As in the first test of Fig. 15, the second kernel test gives an acceleration factor of almost five for the whole code when multi-GPUs are used, compared to the case of multi-CPUs. However, as can be seen in Fig. 16, scalability deteriorates with increasing numbers of processes owing to predominant communications. In this second kernel test, very poor performances are very clear from that point of view. For multi-GPUs as well as for multi-CPUs, scalability decreases almost linearly with the number of processes and time cost of the inversion reaches an almost constant value for any set of CPUs or GPUs due to the predominance of global communications. As often, there is thus a trade-off to be looked for between using more processes and having to handle more costly global communications.

In Fig. 17(A) we show weak scaling of the inversion process obtained on a pure CPU cluster. Weak scaling is obtained when increasing the problem size and the number of processor cores by the same factor. For multi-CPUs we use 1500, 3000, 6000, 12 000, 24 000, 48 000 and 99 856 data on, respectively, 32, 64, 128, 256, 512, 1024 and 2048 processor cores. We observe that scaling is good because it has a behaviour close to constant in terms of CPU time per data volume and per core. In Figs 17(B) and (C) we compare the weak scaling of multi-CPUs to multi-GPUs, respectively, for the whole code (kernels calculations, compression and inversion) and for the inversion phase only. We test 1500, 3000, 6000 and 9000 data for 32, 64, 128 and 192 CPU cores or GPUs. Again we observe good weak scaling for multi-GPUs, with only small fluctuations. GPUs accelerate the code by a factor between 4 and 5 for the whole code and around 3 in the inversion phase. For the largest case, that is, for the highest number of parameters and data, the acceleration of the code using GPUs is higher and can reach acceleration factors up to 14 as said before in the 'strong' scaling analysis when fewer processes (around 32) are used. Conversely, we reach lower accelerations around 4 or 5 when 192 GPUs are used, which means that GPUs are more efficiently used when they have a sufficient amount of work to do.

8 CONCLUSIONS AND FUTURE WORK

In this paper, we propose a new method to perform fast gravity data inversion using wavelet compression as well as hybrid GPU/CPU computing. The introduction of both gravity sensitivity kernel compression and depth-weighting pre-conditioners has led to significant acceleration of forward modelling and data inversion of around one to two orders of magnitude on CPU clusters. We have evaluated the performance of the proposed approach with an application to the gravity modelling of southwest Ghana, for which the available geological and geophysical information have allowed us to test the inversion procedure with a realistic prior geological 3-D model. Most of the main structural elements of this region and the observed lithologies were accounted for in order to create density distribution that allowed us to calculate, by forward gravity modelling, a constrained geological model of the upper crust of southwest Ghana. We discretized this model into approximately 53.8 million prismatic cells and we used an observed gravity data set of almost 100 000 points in our inversion procedure. We developed an efficient parallel algorithm together with the multiscale compression of gravity sensitivity kernels for the calculation of gravity anomalies using computer clusters. When applying four-scale levels of the fourth-order Daubechies compression operator we obtained compression ratios of around 500 to 1000 (from 0.239 per cent to 0.116 per cent of all the wavelet coefficients retained) for several tens of millions of discretization density blocks and also obtained a good scaling of the code. We have also been able to define an automatic criterion, which determines an optimal compression ratio by minimizing the derivative of the logarithm of the error due to compression. Fourth-order Daubechies wavelet operators have proved to be more efficient than Haar wavelets, which only compress sensitivity matrices with a sparseness reaching lower than

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

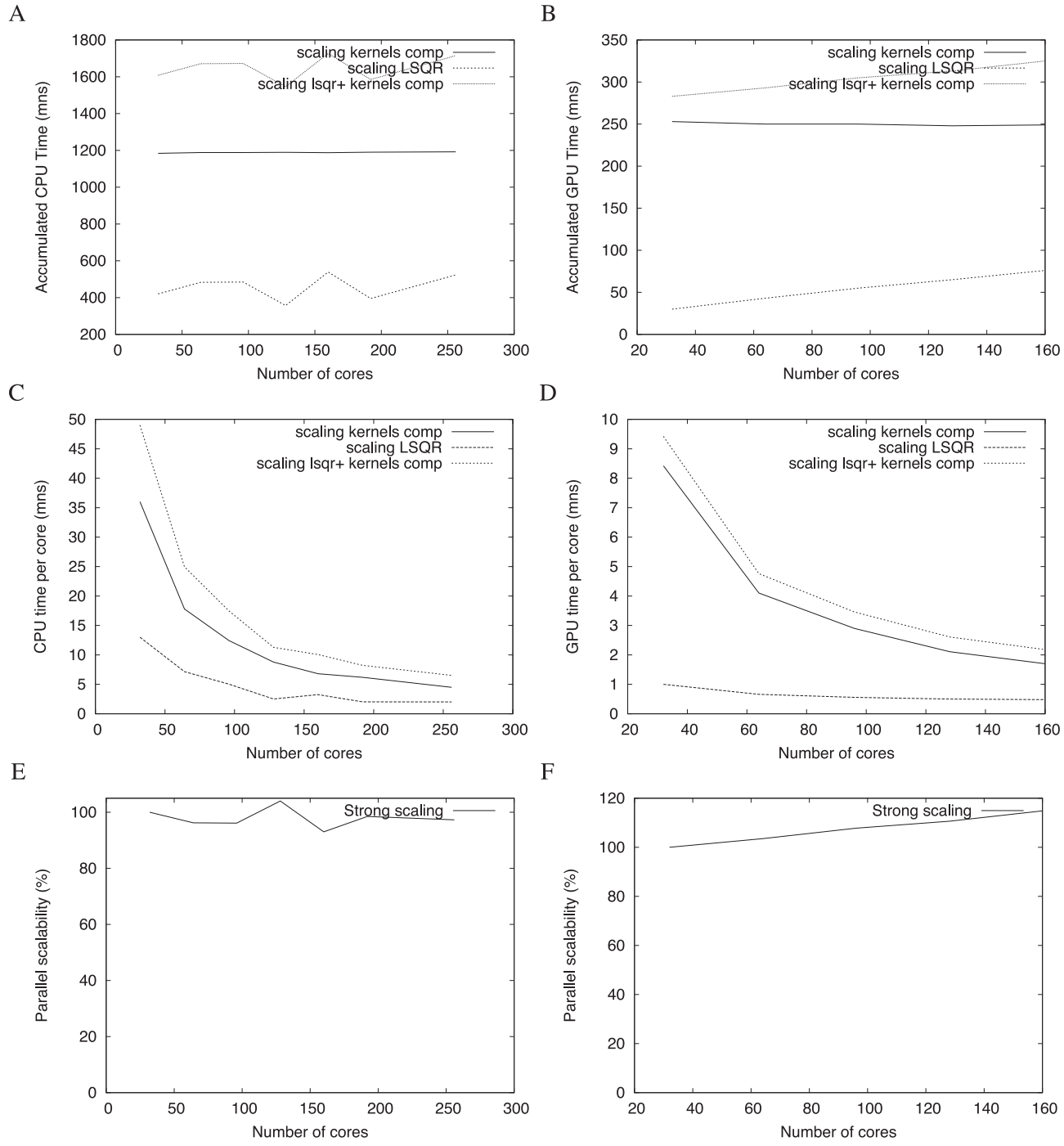

Figure 15. Comparisons between CPU (left-hand side) and GPU (right-hand side) cases for the huge sensitivity matrix case (890 × 890 × 68 parameters and 99 856 anomaly data). Acceleration factors around 5.7 on 32 GPUs (280 accumulated minutes) and 5.5 on 160 GPUs (323 min) are obtained compared to using 32 and 160 CPU cores for the whole code. Acceleration factors of around 4.9 are obtained for the kernel calculation and compression. In the inversion process factors of 14 for 32 GPUs and 7.1 for 160 GPUs can be reached. Reasonable scalability is obtained, which slightly deteriorates with an increasing number of processes (see text for more explanations).

90 per cent in the cases treated here. In summary, an approximate speedup factor between 500 and 1000 has been obtained based on compression due to the great smoothness of the gravity kernels in the Section 4. Inversions have been performed on up to 2048 CPU cores. Supplementary acceleration factors of around 28 for both kernel calculation and the compression process, and 3.5 and 7 for the inversion part can be obtained on hybrid multi-CPU/multi-GPU clusters. This double speedup has allowed us to invert gravity anomalies of the Ghana region in western Africa for huge observed data and parameter sets in less than 15 min instead of several hours or tens of hours.

In future work, further parallelization of fourth-order Daubechies discrete wavelets or Cohen–Daubechies–Fauveau lifting scheme wavelet compression, already used in Chevrot *et al.* (2012) could be improved on GPUs following Shahbahrami (2012) or using the split-and-merge method of Argüello *et al.* (2012).

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

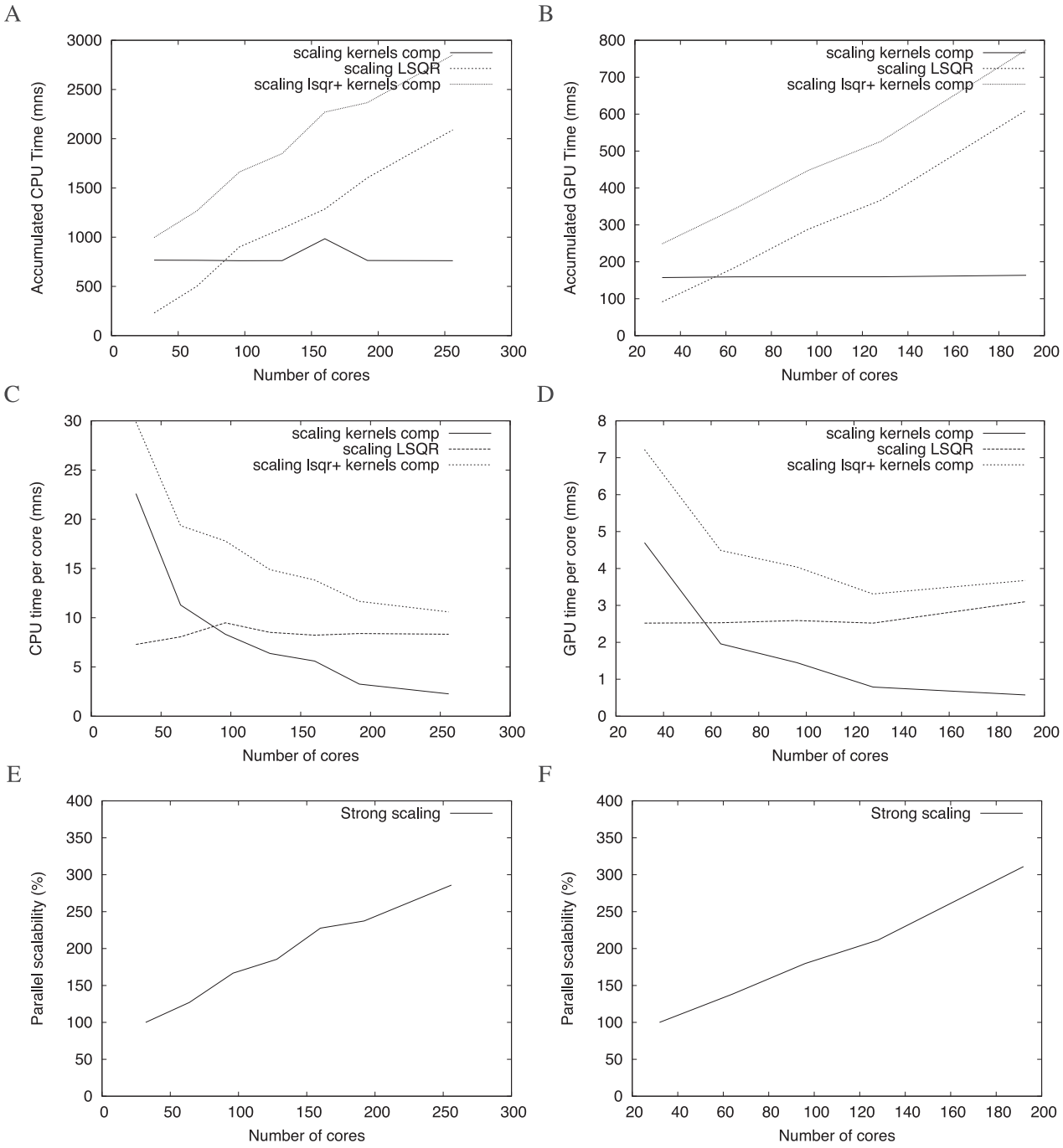

Figure 16. Comparisons for the high-frequency case between CPU (left-hand side) and GPU (right-hand side) calculations. As in the first test of Fig. 15, we have an acceleration factor of almost five between multi-CPU and multi-GPU computations for the whole code. However, scalability deteriorates gradually with increasing number of processes owing to predominant CPU communications and/or CPU/GPU memory copies.

The implementation of high-performance computing methods in gravity inversion is of increasing importance in geophysical interpretation. Inverse modelling combining seismic and gravity data for instance has been proposed to provide useful constraints for regional modelling as shown by recent studies (Vermeesch *et al.* 2009; Gallardo & Meju 2011; Moorkamp *et al.* 2011; Morgan *et al.* 2011). We aim at constraining gravity with high-order finite difference (Martin & Komatitsch 2009; Martin *et al.* 2010) or higher order finite-element (Peter *et al.* 2011) techniques to better image the geological interfaces at high seismic source frequencies. This could be applied in the near future to different regions at different scales using computational domains that can be efficiently truncated with, now popular, optimized perfectly matched layers (Martin *et al.* 2008). At larger scales, the use of high-performance computing will help to calculate more accurate models of the Earth's interior taking advantage of the increasing resolution and accuracy of regional and global data sets from widely distributed seismic networks [e.g. in Africa, Durrheim & Nyblade (2009)] and from recent global models of potential field data. The recent GOCE space gravity mission as well as high-resolution global gravity products combining ground and space measurements (Pavlis *et al.* 2008; Andersen 2010; Balmino *et al.* 2012; Bonvalot *et al.* 2012) now provide

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

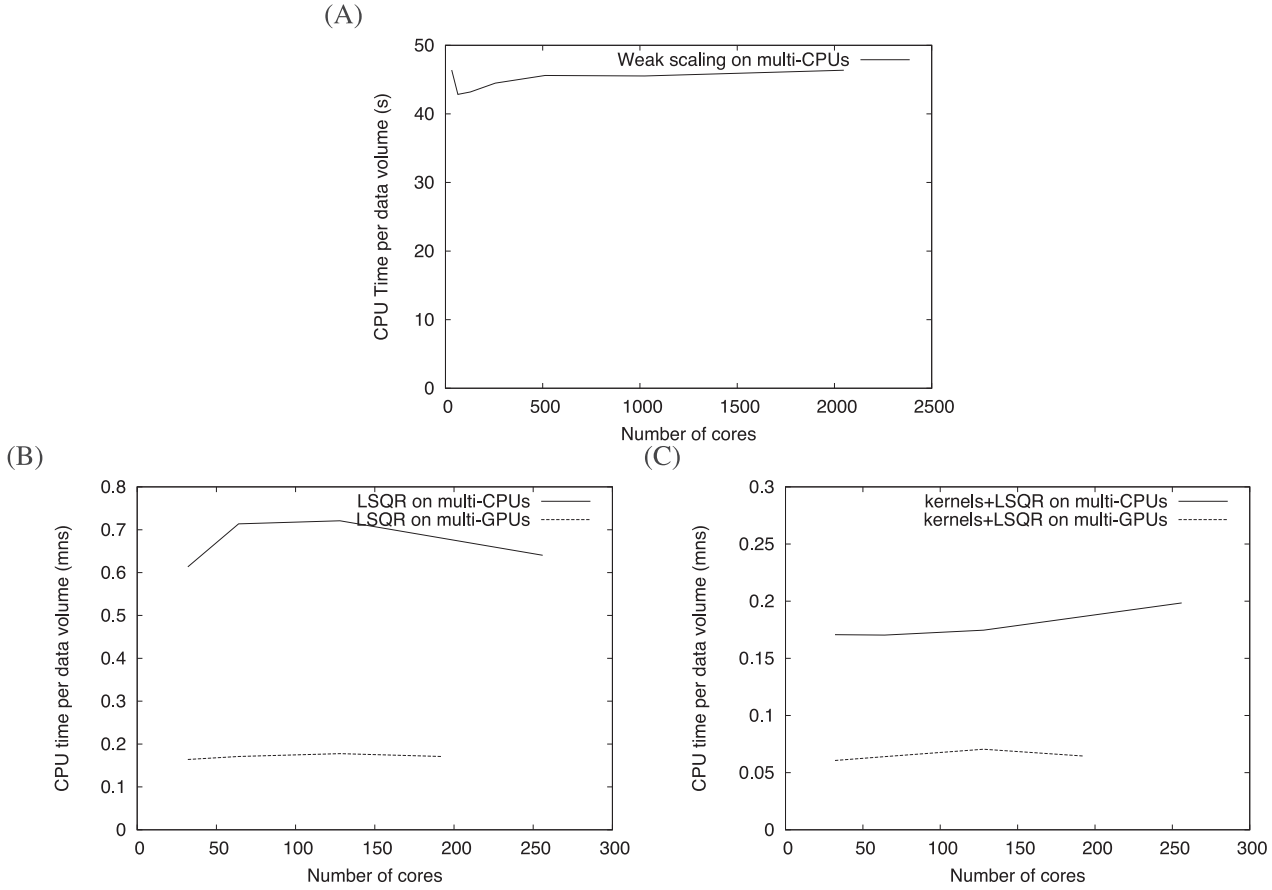

Figure 17. Weak scaling of the inversion process obtained on the multi-CPU (A) cluster. For multi-CPUs we used 1500, 3000, 6000, 12 000, 24 000, 48 000 and 99 856 data on, respectively, 32, 64, 128, 256, 512, 1024 and 2048 processes. We observe that scaling has a behaviour close to constant with few fluctuations. In (B) and (C) we compare the weak scaling of multi-CPUs to that of multi-GPUs for the whole code (kernels calculations, compression and inversion) as well as for the inversion phase only. We test 1500, 3000, 6000 and 9000 data for 32, 64, 128 and 192 for both CPU cores and GPUs. We also observe good weak scaling for multi-GPUs. GPUs accelerate the code by a factor between 4 and 5 for the whole code and around 3 for the inversion phase only.

unprecedented data sources for the gravity field of the Earth to be used for high-performance combined geophysical inversion at regional or global scales.

ACKNOWLEDGEMENTS

This work was granted access to the high-performance computing resources of the French supercomputing centre, CCRT, under allocation #2012-046351 awarded by GENCI (Grand Équipement National de Calcul Intensif) and of the regional computing centre, CALMIP of Toulouse, France, under allocation #P1138. We also thank INRIA for ANR AHPI grant. The authors thank Dominik Göddeke for discussion on sparse solvers for GPU computers and Sébastien Chevrot for discussions on wavelet compression.

REFERENCES

Adadey, K., Clarke, B., Théveniaut, H., Urien, P., Delor, C., Roig, J.Y. & Feybesse, J.L., 2009. Geological map explanation: map sheet 0503 B (1:100 000), CGS/BRGM/Geoman, Tech. Rep, Geological Survey Department of Ghana (GSD). MSSP/2005/GSD/5a.

Agarwal, B.N.P. & Srivastava, S., 2010. A FORTRAN program to implement the method of finite elements to compute regional and residual anomalies from gravity data, *Comput. Geosci.*, **36,** 848–852.

Allibone, A., McCuaig, T., Harris, D., Etheridge, M., Munroe, S. & Byrne, D., 2002. Structural controls on gold mineralization at the Ashanti Gold Deposit, Obuasi, Ghana, *Soc. Econ. Geologist.*, **9,** 65–93. [Special publication]

Andersen, O., 2010. The DTU10 global gravity field and mean sea surface: improvements in the Arctic, in *Proceedings of the 2nd IGFS Meeting*, Fairbanks, Alaska, USA.

Argüello, F., Heras, D.B., Bóo, M. & Lamas-Rodríguez, J., 2012. The split-and-merge method in general purpose computation on GPUs, *Parallel Comput.*, **38,** 277–288.

Bailey, I.W., Miller, M.S., Liu, K. & Levander, A., 2012. V_s and density structure beneath the Colorado Plateau constrained by gravity anomalies and joint inversions of receiver function and phase velocity data, *J. geophys. Res.*, **117,** B02313, doi:10.1029/2011JB008522.

Balmino, G., Vales, N., Bonvalot, S. & Briais, A., 2012. Spherical harmonic modelling to ultrahigh degree of Bouguer and isostatic anomalies, *J. Geod.*, **86**(7), 499–520.

Barritt, S. & Kuma, J., 1998. Constrained gravity models and structural evolution of the Ashanti Belt, southwest Ghana, *J. Afr. Earth Sci.*, **26**(4), 539–550.

Battacharyya, B.K., 1980. A generalized multibody model for inversion of magnetic anomalies, *Geophysics*, **45,** 255–270.

Beck, A. & Teboulle, M., 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imag. Sci.*, **2**(1), 183–202.

Blakely, R.J., 1995. *Potential Theory in Gravity and Magnetic Applications*, Cambridge Univ. Press.

Bonhomme, M., 1962. Contribution à l'étude géochronologique de la plateforme de l'Ouest Africain, *Ann. Fac. Sci., Université de Clermont-Ferrand, France, Géologie, Minéralogie*, **5,** 1–62.

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

Bonvalot, S. *et al.* Eds., 2012. *World Gravity Map*, BGI/CGMW/CNES/IRD.

Boschetti, F., Hornby, P. & Horowitz, F.G., 2001. Wavelet based inversion of gravity data, *Explor. Geophys.*, **32,** 48–55.

Calcagno, P., Chilès, J., Courrioux, G. & Guillen, A., 2008. Geological modelling from field data and geological knowledge. Part I. Modelling method coupling 3D potentiel-field interpolation and geological rules, *Phys. Earth planet. Inter.*, **171,** 147–157.

Chapman, C. & Pratt, R.G., 1992. Traveltime tomography in anisotropic media—I: theory, *Geophys. J. Int.*, **109,** 1–19.

Chevrot, S., Martin, R. & Komatitsch, D., 2012. Optimized discrete wavelet transforms in the cubed-sphere with the lifting scheme: implications for global finite-frequency tomography, *Geophys. J. Int.*, **191**(3), 1391–1402.

Cohen, A., Daubechies, I. & Fauveau, J., 1992. Biorthogonal bases of compactly supported wavelets, *Commun. Pure appl. Math.*, **45**(5), 485–560.

Commer, M., 2011. Three-dimensional gravity modelling and focusing inversion using rectangular meshes, *Geophys. Prospect.*, **59,** 966–979.

Cuma, M., Wilson, G. & Zhdanov, M., 2012. Large-scale 3D inversion of potential data, *Geophys. Prospect.*, **60,** 1186–1199.

Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets, *Commun. Pure appl. Math.*, **41**(7), 909–996.

Daubechies, I., 1992. Ten lectures on wavelets, in *Proceedings of the CBMS-NSF Regional Conference Series in Applied Mathematics*, Vol. 61, SIAM.

Daubechies, I. & Sweldens, W., 1998. Factoring wavelet transforms into lifting steps, *J. Fourier Anal. Appl.*, **4**(3), 247–269.

De Kock, G., Armstrong, R., Siegfried, H. & Thomas, E., 2011. Geochronology of the Birim Supergroup of the West African craton in the Wa-bolé region of west-central Ghana: implications for the stratigraphic framework, *J. Afr. Earth Sci.*, **59**(1), 1–40.

Dongarra, J.J., Du Croz, J., Hammarling, S. & Hanson, R.J., 1988. An extended set of FORTRAN basic linear algebra subprograms, *ACM Trans. Math. Softw.*, **14**(1), 1–17.

Durrheim, R. & Nyblade, A., 2009. Africa array special volume: introduction, *S. Afr. J. Geol.*, **112,** 209–212.

Farquharson, C. & Oldenburg, D., 2004. A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems, *Geophys. J. Int.*, **156,** 411–425.

Fatahalian, K. & Houston, M., 2008. A closer look at GPUs, *Commun. ACM*, **51**(10), 50–57.

Fullagar, P., Pears, G., Hutton, D. & Thmpson, A., 2004. 3D gravity and aeromagnetic inversion for MVT lead-zinc exploration at pillara, Western Australia, *Explor. Geophys.*, **35**(2), 142–146.

Fullagar, P., Pears, G. & McMonnies, B., 2008. Constrained inversion of geologic surfaces-pushing the boundaries, *Leading Edge*, **27**(1), 98–105.

Gallardo, L., 2007. Multiple cross-gradient joint inversion for geospectral imaging, *Geophys. Res. Lett.*, **34,** L19301, doi:10.1029/2007GL030409.

Gallardo, L. & Meju, M., 2011. Structure-coupled multiphysics imaging in geophysical sciences, *Rev. Geophys.*, **49,** 1–20.

Garcia-Abdeslem, J., 2000. Nonlinear 3-D inversion of gravity data over a sulfide ore body, *Geofisica Internacional*, **39**(2), 179–188.

Garcia-Abdeslem, J., 2008. 3D forward and inverse modelling of total-field magnetic anomalies caused by a uniformly magnetized layer defined by a linear combination of 2D Gaussian functions, *Geophysics*, **73**(1), L11–L18.

Garcia-Abdeslem, J., Espinosa-Cardea, J., Mungua-Orozco, L., Wong-Ortega, V. & Ramrez-Hernndez, J., 2001. Crustal structure from 2-D gravity and magnetic data modeling, magnetic power spectrum inversion, and seismotectonics in the Laguna Salada Basin, northern Baja California, Mexico, *Geofis. Int.*, **40**(2), 67–87.

Gropp, W., Lusk, E. & Skjellum, A., 1994. *Using MPI, Portable Parallel Programming with the Message-Passing Interface*, MIT Press.

Guillen, A., Calcagno, P., Courrioux, G., Joly, A. & Ledru, P., 2008. Geological modelling from field data and geological knowledge, part II: modelling validation using gravity and magnetic data inversion, *Phys. Earth planet. Inter.*, **171,** 158–169.

Hansen, P., 2001. The L-curve and its use in the numerical treatment of inverse problems, in *Advances in Computational Bioengineering, Computational Inverse Problems in Electrocardiology*, Vol. **5**, pp. 119–142, ed. Johnston, P., WIT Press.

Hastings, D., 1978. The southwest Ghana gravity survey, with West African regional correlations and implications. Tech. Rep, Unpublished Report, Ghana Geological Survey.

Hastings, D., 1982. On the tectonics and metallogenesis of West Africa: a model incorporating new geophysical data, *Geoexploration*, **20,** 295–327.

Hildebrand, A., Pilkington, M., Ortiz-Aleman, C., Chavez, R., Urrutia-Fucugauchi, J., Connors, M., Graniel-Castro, E. & Niehaus, D., 1998. Mapping Chicxulub crater structure with gravity and seismic reflection data, meteorites: flux with time and impact effects, *Geol. Soc. Lond. Spec. Publ.*, **140,** 155–176.

Hornby, P., Boschetti, F. & Horowitz, F.G., 1998. Analysis of potential field data in the wavelet domain, *Geophys. J. Int.*, **137,** 175–196.

Kirk, D.B. & Hwu, W.-M.W., 2010. *Programming Massively Parallel Processors: A Hands-on Approach*, Morgan Kaufmann.

Komatitsch, D., 2011. Fluid-solid coupling on a cluster of GPU graphics cards for seismic wave propagation, *C. R. Acad. Sci., Ser. IIb Mec.*, **339,** 125–135.

Komatitsch, D., Göddeke, D., Erlebacher, G. & Michéa, D., 2010. Modeling the propagation of elastic waves using spectral elements on a cluster of 192 GPUs, *Comp. Sci. Res. Dev.*, **25**(1–2), 75–82.

Koren, Z., Mosegaard, K., Landa, E., Thore, P. & Tarantola, A., 1991. Monte Carlo estimation and resolution analysis of seismic background velocities, *J. geophys. Res.*, **96**(B12), 20 289–20 299.

Lajaunie, C., Courrioux, G. & Manuel, L., 1997. foliation fields and 3D cartography in geology: principles of a method based on potential interpolation, *Math. Geol.*, **29**(4), 571–584.

Li, Y. & Oldenburg, D.W., 2003. Fast inversion of large-scale magnetic data using wavelet transforms and a logarithmic barrier method, *Geophys. J. Int.*, **152,** 251–265.

Liu, G., Meng, X. & Chen, Z., 2012. 3D magnetic inversion based on probability tomography and its GPU implement, *Comput. Geosci.*, **48,** 86–92.

Loris, I., Nolet, G., Daubechies, I. & Dahlen, F., 2007. Tomographic inversion using l_1-norm regularization of wavelet coefficients, *Geophys. J. Int.*, **170,** 359–370.

Loris, I., Douma, H., Nolet, G., Daubechies, I. & Regone, C., 2010. Nonlinear regularization techniques for seismic tomography, *J. Comput. Phys.*, **229,** 890–905.

Mallet, J.L., 1992. Discrete smooth interpolation in geometric modeling, *Comput.-Aided Des.*, **24,** 178–191.

Martin, R. & Komatitsch, D., 2009. An unsplit convolutional perfectly matched layer technique improved at grazing incidence for the viscoelastic wave equation, *Geophys. J. Int.*, **179**(1), 333–344.

Martin, R., Komatitsch, D. & Gedney, S.D., 2008. A variational formulation of a stabilized unsplit convolutional perfectly matched layer for the isotropic or anisotropic seismic wave equation, *Comput. Model. Eng. Sci.*, **37**(3), 274–304.

Martin, R., Komatitsch, D., Gedney, S.D. & Bruthiaux, E., 2010. A high-order time and space formulation of the unsplit perfectly matched layer for the seismic wave equation using Auxiliary Differential Equations (ADE-PML), *Comput. Model. Eng. Sci.*, **56**(1), 17–42.

Mauri, G., Williams-Jones, G. & Saracco, G., 2010. Depth determinations of shallow hydrothermal systems by self-potential and multi-scale wavelet tomography, *J. Volc. Geotherm. Res.*, **191**(3–4), 233–244.

Mauri, G., Williams-Jones, G. & Saracco, G., 2011. MWTmat application of multiscale wavelet tomography on potential fields, *Comput. Geosci.*, **37**(11), 1825–1835.

Michéa, D. & Komatitsch, D., 2010. Accelerating a 3D finite-difference wave propagation code using GPU graphics cards, *Geophys. J. Int.*, **182**(1), 389–402.

Moorkamp, M., Jegen, M., Roberts, A. & Hobbs, R., 2010. Massively parallel forward modeling of scalar and tensor gravimetry data, *Comput. Geosci.*, **36**(36), 680–686.

Moorkamp, M., Heincke, B., Jegen, M., Roberts, A. & Hobbs, R., 2011. A framework for 3-D joint inversion of MT, gravity and seismic refraction data, *Geophys. J. Int.*, **184**(1), 477–493.

Morgan, J.V., Warner, M.R., Collins, G.S., Grieve, R.A.F., Christeson, G.L., Gulick, S.P.S. & Barton, P.J., 2011. Full waveform tomographic images

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

of the peak ring at the Chicxulub impact crater, *J. geophys. Res.*, **116,** B06303, doi:10.1029/2010JB008015.
Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.*, **10**(B7), 12 431–12 447.
Mosegaard, K. & Tarantola, A., 2002. *Probabilistic Approach to Inverse Problems*, Academic Press.
NVIDIA Corporation, 2011. *NVIDIA CUDA toolkit documentation version 4. NVIDIA Corporation*, Santa Clara, California, USA.
Ortiz-Aleman, C. & Martin, R., 2005. Two-phase oil-gas pipe flow imaging by simulated annealing, *J. geophys. Eng.,* **2**(1), 32–37.
Ortiz-Aleman, C. & Urrutia-Fucugauchi, J., 2010. Aeromagnetic anomaly modeling of central zone structure and magnetic sources in the Chicxulub crater, *Phys. Earth planet. Inter.*, **179**(3–4), 127–138.
Owens, J.D., Houston, M., Luebke, D.P., Green, S., Stone, J.E. & Phillips, J.C., 2008. GPU computing, *Proc. IEEE*, **96**(5), 879–899.
Pacheco, P.S., 1997. *Parallel Programming with MPI*, Morgan Kaufmann.
Parker, R., 1994. *Geophysical Inverse Theory*, Princeton Univ. Press.
Pavlis, N.K., Holmes, S.A., Kenyon, S.C. & Factor, J.K., 2008. An earth gravitational model to degree 2160: EGM2008, Tech. Rep, General Assembly of the European Geosciences Union, Vienna, Austria, 2008 April 13–18.
Perrouty, S., 2012. Évolution Structurale de la Ceinture Minéralisée d'Ashanti, Sud-Ouest Ghana, *PhD thesis*, Université de Toulouse III - Paul Sabatier, France.
Perrouty, S., Aillères, L., Jessell, M., Baratoux, L. & Bourassa, Y., 2012. Revised Eburnean geodynamic evolution of the gold-rich southern Ashanti Belt, Ghana, with new field and geophysical evidence of pre-Tarkwaian deformations, *Precambrian Res.*, **204–205,** 12–39.
Peter, D. *et al.*, 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.*, **186**(2), 721–739.
Pilkington, M., 2009. 3D magnetic data-space inversion with sparseness constraints, *Geophysics*, **74**(1), L7–L15.
Pilkington, M. & Hildebrand, A., 2000. Three-dimensional magnetic imaging of the Chicxulub crater, *J. geophys. Res.*, **105,** 23 479–23 491.
Pilkington, M., Hildebrand, A. & Ortiz-Aleman, C., 1994. Gravity and magnetic field modeling and structure of the Chicxulub crater, Mexico, *J. geophys. Res.*, **99**(E6), 13 147–13 162.
Pontoise, B., Bonvalot, S., Mascle, J. & Basile, C., 1990. Structure crustale de la marge transformante de Côte d'Ivoire-Ghana, *C.R. Académie des Sciences de Paris*, **310**(2), 527–534.
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. & Chapman, C., 1992. Traveltime tomography in anisotropic media -II: application, *Geophys. J. Int.*, **109,** 20–37.
Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophysics*, **133**(2), 341–362.
Shahbahrami, A., 2012. Algorithms and architectures for 2D discrete wavelet transform, *J. Supercomput.*, **62,** 1045–1064.
Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics*, **69**(1), 231–248.
Stollnitz, E., deRose, T. & Salesin, D., 1995. Wavelets for computer graphics: a primer, part 1, *IEEE Comput. Graph. Appl.*, **15**(3), 76–86.
Talwani, M., Worzel, J.L. & Landisman, M., 1959. Rapid computation for two-dimensional bodies with application to the Mendocino submarine fracture zone, *J. geophys. Res.*, **64,** 49–59.
The International Association of Geodesy, 2012. The International Gravimetric Bureau, *J. Geod.*, **86**(10), 946–949.
Uytterhoeven, G., Roose, D. & Bultheel, A., 1997. Wavelet transforms using the lifting scheme. Tech. Rep, Report ITA-Wavelets- WP1.1 (Revised Version).
Vermeesch, P.M., Morgan, J.V., Christeson, G.L., Barton, P.J. & Surendra, A., 2009. Three-dimensional joint inversion of traveltime and gravity data across the Chicxulub impact crater, *J. geophys. Res.*, **114**, B02105, doi:10.1029/2008JB005776.
Wilson, G., Cuma, M. & Zhdanov, M., 2011. Massively parallel 3D inversion of gravity and gradiometry data, *Preview,* **152,** 29–34.
Zhang, J., Wang, C., Shiz, Y., Cai, Y., Chi, W., Dreger, D., Cheng, W. & Yuan, Y., 2004. Three-dimensional crustal structure in central Taiwan from gravity inversion with a parallel genetic algorithm, *Geophysics*, **69**(4), 917–924.
Zhdanov, M., 2009. New advances in regularized inversion of gravity and electromagnetic data, *Geophys. Prospect.*, **57,** 463–478.
Zhdanov, M., Liu, X., Wilson, G. & Wan, L., 2011. Potential field migration for rapid imaging of gravity gradiometry data, *Geophys. Prospect.*, **59**(6), 1052–1071.

APPENDIX A: HAAR AND DAUBECHIES COMPRESSION IMPLEMENTATION

Using Haar wavelet compression, the procedures applied in each direction of the space are as follows:

Predict

$$D_{2i+1} = S_{2i} - S_{2i+1}, \qquad \text{for } i = 1, (N+1)/2 - 1.$$

Update

$$D_{2i} = S_{2i} - (1/2)S_{2i+1}, \quad \text{for } i = 1, (N+1)/2 - 1.$$

Normalization

$$D_{2i} = D_{2i}\sqrt{2},$$
$$D_{2i+1} = D_{2i+1}/\sqrt{2}, \tag{A1}$$

where vector D represents the coefficients of the dyadic decomposition after wavelet compression of a given 1-D matrix S. For the lifting fourth-order Daubechies wavelet, the wavelets are built as follows:

Update

$$D_{2i+1} = S_{2i+1} + \sqrt{3}S_{2i}, \quad \text{for } i = 0, (N+1)/2 - 1.$$

Predict

$$D_{2i} = S_{2i} - (\sqrt{3}/4)D_{2i+1} - ((\sqrt{(3)} - 2)/4)D_{2i+3}, \quad \text{for } i = 1, (N+1)/2 - 1.$$

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021

Update

$$D_{2i} = D_{2i+1} - D_{2i-2}, \quad \text{for } i = 0, (N+1)/2 - 1.$$

Normalization

$$\begin{aligned} D_{2i} &= D_{2i}(\sqrt{3}+1)/\sqrt{2}, \\ D_{2i+1} &= D_{2i+1}(\sqrt{3}-1)/\sqrt{2}. \end{aligned} \tag{A2}$$

APPENDIX B: INVERSION ALGORITHMS USING L^1 AND L^2 NORMS

The least-square algorithm is used to minimize the L^2-norm misfit function (28) as follows:

Set a starting model $x_0 = 0$ or an a priori model

(1) Initialization

$\beta_0 = ||b||, \alpha_0 = ||A^{\mathrm{T}}b||, \phi_0^1 = ||b||, \rho_0^1 = \alpha_0, v_0 = A^{\mathrm{T}}(b), w_0 = v_0, u_0 = b$

(2) For $i = 0$, Niter repeat steps 3 to 6

(3) Update using matrix-vector and transpose matrix-vector products

$u_{i+1} = -\alpha_i u_i + Av_i$

$\beta_{i+1} = ||u_{i+1}||, u_{i+1} = u_{i+1}/\beta_{i+1}$

$v_{i+1} = -\beta v_i + A^{\mathrm{T}}u_{i+1}$

$\alpha_{i+1} = ||v_{i+1}||, v_{i+1} = v_{i+1}/\alpha_{i+1}$

(4) Update coefficients

$\rho_i = \left(\rho_i^{1\,2} + \beta_{i+1}^2\right)^{1/2}, c_i = \rho_i^1/\rho_i, s_i = \beta_{i+1}/\rho_i, \theta_{i+1} = s_i\alpha_{i+1}, \phi_i = c_i\phi_i^1, \phi_{i+1}^1 = s_i\phi_i^1$

(5) Update current model

$x_i = x_{i-1} + (\phi_i/\rho_i)w_i$

$w_{i+1} = v_{i+1} - (\theta_{i+1}/\phi_i)w_i$

(6) Test for convergence

if $\phi_{i+1}^1 < \varepsilon\phi_1^1$ stop (B1)

Whereas, a gradient method is used to minimize the L^2-/L^1-norm misfit function (29) as follows:

(1) Initialization

Set a starting model given an a priori model x_0

(2) For $i = 0$, Niter repeat steps 3 to 5

(3) Update

$x_{n+1} = T(x_n)$

(4) with function

$T(x) = S_\mu(x + \alpha A^{\mathrm{T}}(b - Ax))$

(5) Test for convergence

if $|x_{i+1} - x_i| < \varepsilon |x_1|$ stop (B2)

where the soft-thresholding function S_μ is given by

$S_\mu = u - \mu$ if $u \geq \mu$

$S_\mu = 0$ if $|u| \leq \mu$

$S_\mu = u + \mu$ if $u \leq -\mu$ (B3)

Downloaded from https://academic.oup.com/gji/article/195/3/1594/623670 by guest on 18 June 2021