AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE
10.1002/2016JB013049

This article is a companion to *Afonso et al.* [2013] doi:10.1002/jgrb.50124 and *Afonso et al.* [2013] doi:10.1002/jgrb.50123.

Key Points:
- We present a novel multiobservable probabilistic tomography method to image the thermochemical structure of the lithosphere and upper mantle
- We apply the new method to image study the lithosphere-asthenosphere system beneath the Colorado Plateau and surrounding areas
- A complex and vigorous lithosphere-asthenosphere system is imaged, with multiscale feedback mechanisms driving surface processes

Supporting Information:
- Supporting Information S1

Correspondence to:
J. C. Afonso,
juan.afonso@mq.edu.au

Citation:
Afonso, J. C., N. Rawlinson, Y. Yang, D. L. Schutt, A. G. Jones, J. Fullea, and W. L. Griffin (2016), 3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle: III. Thermochemical tomography in the Western-Central U.S., *J. Geophys. Res. Solid Earth*, *121*, 7337–7370, doi:10.1002/2016JB013049.

Received 3 APR 2016
Accepted 14 SEP 2016
Accepted article online 20 SEP 2016
Published online 18 OCT 2016

3-D multiobservable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle: III. Thermochemical tomography in the Western-Central U.S.

Juan Carlos Afonso[1], Nicholas Rawlinson[2], Yingjie Yang[1], Derek L. Schutt[3], Alan G. Jones[4,5], Javier Fullea[5], and William L. Griffin[1]

[1]CCFS, Department of Earth and Planetary Sciences, Macquarie University, Sydney, New South Wales, Australia, [2]School of Geosciences, University of Aberdeen, Aberdeen, UK, [3]Department of Geosciences, Colorado State University, Fort Collins, Colorado, USA, [4]Complete MT Solutions, Manotick, Ottawa, Canada, [5]Dublin Institute for Advanced Studies, Dublin, Ireland

Abstract We apply a novel 3-D multiobservable probabilistic tomography method that we have recently developed and benchmarked, to directly image the thermochemical structure of the Colorado Plateau and surrounding areas by jointly inverting *P* wave and *S* wave teleseismic arrival times, Rayleigh wave dispersion data, Bouguer anomalies, satellite-derived gravity gradients, geoid height, absolute (local and dynamic) elevation, and surface heat flow data. The temperature and compositional structures recovered by our inversion reveal a high level of correlation between recent basaltic magmatism and zones of high temperature and low Mg# (i.e., refertilized mantle) in the lithosphere, consistent with independent geochemical data. However, the lithospheric mantle is overall characterized by a highly heterogeneous thermochemical structure, with only some features correlating well with either Proterozoic and/or Cenozoic crustal structures. This suggests that most of the present-day deep lithospheric architecture reflects the superposition of numerous geodynamic events of different scale and nature to those that created major crustal structures. This is consistent with the complex lithosphere-asthenosphere system that we image, which exhibits a variety of multiscale feedback mechanisms (e.g., small-scale convection, magmatic intrusion, delamination, etc.) driving surface processes. Our results also suggest that most of the present-day elevation in the Colorado Plateau and surrounding regions is the result of thermochemical buoyancy sources within the lithosphere, with dynamic effects (from sublithospheric mantle flow) contributing only locally up to ~15–35%.

1. Introduction

The western U.S. has undergone extensive tectonomagmatic activity since the Laramide orogeny [80–45 Ma], including episodes of flat subduction, crustal shortening, lithospheric delamination/alteration, crustal extension, voluminous volcanism, and epeirogenesis [cf. *Humphreys et al.*, 2003; *Humphreys and Coblentz*, 2007; *Karlstrom et al.*, 2012]. Despite being one of the better studied regions in the world, a number of key questions regarding the tectonic evolution of Western-Central U.S. remain contentious [cf. *Humphreys and Coblentz*, 2007; *Karlstrom et al.*, 2012]. Foremost among these are the evolution and present-day structure of the lithosphere-asthenosphere boundary (LAB) and the thermochemical structure of the lithosphere, including its role in the rise and tectonic stability of the Colorado Plateau (CP). Numerous studies have addressed these issues, yet often with contrasting results. For instance, lithospheric thickness is commonly estimated either from seismic proxies, xenolith thermobarometry, gravity studies, magnetotelluric studies, or thermal/isostatic modeling. There is no a priori reason, however, why the results from these diverse methods should be strictly comparable, consistent, or compatible, despite the fact that they all sample the same structure (see discussion in *Afonso et al.* [2016]). In the Western-Central U.S., there are only a handful of studies that explicitly attempt to estimate lithospheric thickness/geometry, and their predictions differ considerably in scale and magnitude [*Yuan and Romanowicz*, 2010; *Levander et al.*, 2011; *Kumar et al.*, 2012; *Hansen et al.*, 2013; *Becker et al.*, 2014; *Pasyanos et al.*, 2014; *Lekic and Fisher*, 2014].

Significant controversy also exists regarding the causes for the contrasting tectonomagmatic activity and epirogenesis experienced by the CP and surrounding provinces. In the case of the CP high topography and stability, most early, and some recent, ideas favor lithospheric features/processes as the main cause [e.g., *Spencer*, 1996; *McQuarrie and Chase*, 2000; *Humphreys et al.*, 2003; *Roy et al.*, 2009; *Jones et al.*, 2015; *Molnar et al.*, 2015]. In these studies, specific inferred features of either the crust (e.g., strong lower crust, hydration) or the whole lithosphere (e.g., metasomatism of a depleted mantle, widespread heating) are thought to have controlled the post-Laramide evolution of topography and tectonic stability of the CP. In contrast, deeper dynamic or mantle convection mechanisms have become increasingly invoked [e.g., *van Wijk et al.*, 2010; *Liu and Gurnis*, 2010; *Karlstrom et al.*, 2012; *Flament et al.*, 2013; *Becker et al.*, 2014; *Liu*, 2015]. While such convection models have the advantage of explicitly considering the dynamic state of the lithosphere-mantle system, they typically greatly simplify lithospheric structure, thus implicitly favoring dynamic causes to explain most or all of the topography. This conclusion is at odds with abundant evidence in Western U.S. and worldwide in support of large compositional, seismic, and structural heterogeneities within the lithosphere (either in its crustal or mantle component) correlating with the location of topographic anomalies, major tectonic boundaries, foci of magma intrusion, and major ore deposits [e.g., *Doin et al.*, 1996; *Griffin et al.*, 2009, 2013; *Begg et al.*, 2009; *Crow et al.*, 2011; *Hasterok and Chapman*, 2007; *Levandowski et al.*, 2014; *Jones et al.*, 2015; *Molnar et al.*, 2015]; these suggest that the thermochemical structure of the lithosphere is a key factor in modulating or controlling surface processes. This view is also consistent with the global distribution of free air, geoid and isostatic anomalies [e.g., *Haxby and Turcotte*, 1978; *Molnar et al.*, 2015], and our current understanding of cold thermal boundary layers in planetary interiors (the bathymetry of oceanic plates being the prime example) [e.g., *Parsons and Sclater*, 1977; *Parsons and Daly*, 1983].

At least part of the controversy lies in the fact that reliable estimates of the 3-D temperature distribution and compositional structure beneath continents, strictly necessary to assess the roles of intralithospheric versus deeper processes, are difficult to obtain and surrounded by methodological controversy. For instance, it is well known that in order to isolate temperature from compositional effects in the mantle, data sets sensitive to both absolute seismic velocities (ideally both Vs and Vp) and bulk density need to be combined with an appropriate description of temperature-dependent anelastic effects [e.g., *Forte and Perry*, 2000; *Karato and Karki*, 2001; *Deschamps et al.*, 2002; *Lee*, 2003; *Jackson and Faul*, 2010; *Afonso et al.*, 2010, 2013a]. However, the formal combination of these data sets (not models!) is far from straightforward and rarely performed. Furthermore, *Afonso et al.* [2013a, 2016] clearly demonstrated that within the peridotitic compositional space, the solution to the joint inversion for bulk density and seismic velocities is highly nonunique, and special care needs to be taken in selecting appropriate and representative a priori compositional information.

By overlooking the above issues, we seriously limit our ability to achieve a deeper understanding of important features and processes in the Earth's interior and their surface expressions (see review by *Afonso et al.* [2016]). The problem is aggravated by the fact that vastly different and competing mechanisms can explain some observations equally well. For instance, the late Cenozoic uplift of the edges of the plateau and the associated volcanism can be explained either by the effects of small-scale circulation along the edges of the plateau [e.g., *van Wijk et al.*, 2010] or by allowing simple conductive heating of a variable-thickness lithospheric plate [*Roy et al.*, 2009]. Likewise, the high elevation of the interior of the plateau can be explained by compositional anomalies within the lithosphere, by crustal effects, or by deep dynamic support from sublithospheric mantle flow (see references above).

All of these issues raise two crucial questions: (i) how can we obtain reliable models of the first-order thermochemical structure of the lithosphere (and thus isolate its contribution from deeper sublithospheric processes)? and (ii) to what extent are different tectonic models and processes actually supported or required by independent geophysical, geological, and geochemical evidence? The first question was addressed in our previous contributions [*Afonso et al.*, 2013a, 2013b, Parts I and II], which, in addition to presenting sensitivity and resolution tests, highlighted the need to combine multiple observations with complementary sensitivities into a thermodynamically and internally consistent framework. Here we focus on the second question and introduce the concept of "thermochemical tomography" or "multiobservable probabilistic tomography," which builds upon the multiobservable probabilistic inversion method presented in *Afonso et al.* [2013a, 2013b]. We use this method to produce a 3-D thermochemical model of the Western-Central U.S. by jointly inverting a number of complementary data sets to directly image the complex feedbacks between structure and processes. Working within this internally and thermodynamically consistent framework [cf. *Afonso et al.*, 2013a, 2013b] allows us to move beyond traditional inversion approaches of one or two geophysical

data sets. Specifically, we jointly invert (i) P and S teleseismic travel times, (ii) Rayleigh wave dispersion curves (from ambient noise and teleseismic surface wave tomography), (iii) geoid height (filtered to remove effects from features deeper than ~500 km), (iv) Bouguer gravity anomalies, (v) diagonal components of the Marussi tensor (gravity gradients), (vi) surface heat flow, and (vii) absolute elevation (including both local and dynamic isostasy). These data sets are specifically chosen due to their complementary sensitivities to compositional versus thermal anomalies [e.g., *Karato*, 1993; *Deschamps et al.*, 2002; *Afonso et al.*, 2008a, 2010, 2013a; *Fullea et al.*, 2009]. The nonlinearity of the forward problems and dynamic effects (i.e., sublithospheric flow arising from subsurface density variations) are also fully accounted for (see section 2). In addition, the compositional parameter space (CaO-FeO-MgO-Al_2O_3-SiO_2 or CFMAS) used in the upper mantle is not based on restrictive ad hoc ranges but covers the entire observed range in mantle peridotites, from fertile lherzolites to highly depleted harzburgites and dunites [*Griffin et al.*, 2009; *Afonso et al.*, 2013a]. These are all decisive advantages that allow us to explore the 3-D temperature and compositional structure of the Earth's lithosphere and sublithospheric upper mantle with unprecedented detail and minimum input of subjective information.

2. Data Sets and Their Uncertainties

2.1. Nonseismic Data Sets

Bouguer anomalies were obtained from the continent-wide compilation by *Kucks* [1999] and elevation data from the ETOPO2v2 Global Database (http://www.ngdc.noaa.gov/mgg/fliers/06mgg01.html). To minimize incompatibilities arising from differences between the shallow inverted densities ρ_I and the assumed reduction density $\rho_r = 2670$ kg m^{-3} in the original gravity data, a standard correction $\Delta Bc = 2\pi Gh(\rho_I - \rho_r)$ is applied to the original Bouguer anomalies during the inversion.

Gravity gradients were taken from the satellite mission GOCE [*Pail et al.*, 2011]. In particular, we used the satellite-only global Earth model GOCO03S (http://www.goco.eu/) and computed the gradients up to degree and order 220 at a height of 100 km using a spherical harmonics synthesis code [*Fullea et al.*, 2015]. At this height, the diagonal components of the Marussi tensor (especially the vertical component) have good signal-to-noise ratio and sensitivity to crustal and uppermost mantle structure [e.g., *Bouman et al.*, 2013].

Geoid heights were taken from the global Earth model EGM2008, which includes spherical harmonic coefficients up to degree and order 2190 [*Pavlis et al.*, 2012]. The total geoid signal was filtered to remove long wavelengths associated with density anomalies deeper than ~400 km (i.e., degrees 2–9 removed).

Surface heat flow values were obtained by interpolating available data points (measurements) included in the Geothermal Map of North America (http://smu.edu/geothermal/2004NAMap/2004NAmap.htm). We ignored a few outliers in the original database with values >200 or <25 mW m^2, as they clearly reflect local conditions such as shallow underground water flow, not representative of the deeper thermal structure.

Uncertainties for elevation, gravity anomalies, gravity gradients, and geoid height were computed as follows: for each column used to discretize the 3-D volume (Figure 2a and next section), we read all available values in the data set located within the area covered by the column. We then compute the mean and variance of all values. The means represent the input data for each column, while the variances are used to obtain the associated standard deviation σ. These are not true observational uncertainties, but rather a measure of the natural variability of the fields within each column's surface (observational uncertainties are much smaller and thus ignored here). A similar procedure was used for surface heat flow where enough data points were available. In regions where the original coverage was not dense enough, we assigned a minimum uncertainty of 15%.

2.2. Seismic Data Sets

2.2.1. Rayleigh Wave Dispersion Data

Rayleigh wave dispersion curves at periods from 8 to 150 s were obtained from the combination of ambient noise tomography and teleseismic surface wave tomography. At periods from 8 s to 40 s, Rayleigh wave dispersion curves are taken from the phase velocity maps of *Shen et al.* [2013]. These authors used ambient noise tomography and eikonal tomography to obtain dispersion velocity maps and associated uncertainties. This method constructs a phase travel time map centered at each station using phase travel times measured from cross correlations of ambient noise between the central station and other stations in a seismic array [*Lin et al.*, 2010]. The phase velocity map centered at each station is obtained by converting the phase travel time map using the eikonal equation [cf. *Shearer*, 2009]. The final phase velocity maps are obtained by averaging each

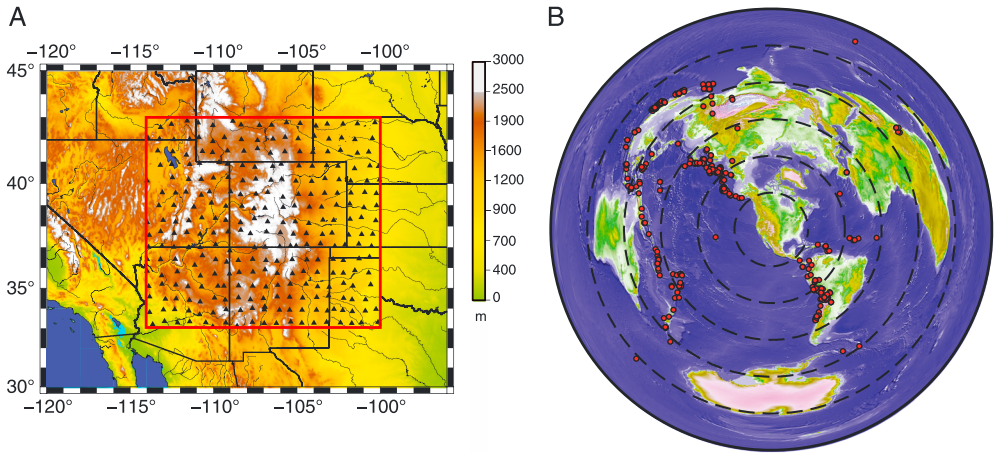

Figure 1. (a) Location of the EarthScope USArray stations (black triangles) used in this study. Background color map indicates elevation. (b) Distribution of the teleseismic events (red circles) from which a total of 22,220 teleseismic travel times (both S and P sources) were selected for this study. Phases used include P, PP, PKiKP, PcP, S, and ScS phases.

station-centered phase velocity map, and the associated uncertainties are evaluated by the variations of the individual phase velocity maps [*Lin et al.*, 2010].

For periods between 40 and 150 s, Rayleigh wave dispersion curves were obtained by inverting teleseismic surface waves using two-plane-wave tomography [e.g., *Yang and Forsyth*, 2006]. We selected about 300 TA stations within our study area and collected about 450 teleseismic events recorded between 2007 and 2009 with magnitudes >5.5 and epicentral distances longer than 3000 km. Because our study region is large, with an aperture >1000 km, which exceeded the limitation of plane wave assumption in Cartesian and spherical coordinates, we divide the whole area into four subregions with nearly equal area. The interference of two plane waves is used to model an incoming surface wavefield propagating in each subregion separately with a total of eight plane waves to represent the full incoming wavefield. Surface wave tomography is performed jointly over the whole region by calculating the sensitivities of phase and amplitude of each plane wave to phase velocity maps using 2-D finite-frequency sensitivity kernels [*Zhou et al.*, 2006]. Uncertainties in phase velocities are estimated from the a posteriori model covariance matrix [*Yang and Forsyth*, 2006] based on the propagation of data uncertainties, which are evaluated using an average data misfit for each event. More details on these procedures can be found in *Yang et al.* [2008].

2.2.2. Body Wave Data Set

The body wave data used in the inversion was taken from the EarthScope/U.S. Transportable Array (Figure 1a) and is composed of 15,568 teleseismic travel times from 236 P sources and 6652 travel times from 100 S sources (Figure 1b). Phases used include P, PP, PKiKP, PcP, S, and ScS phases from sources with magnitudes >5.8. To improve the coverage from the north, additional *P* wave arrivals with north directed back azimuths were selected from high-quality magnitude 5.2–5.6 events. P traces were filtered between 0.1 and 2 Hz, whereas S traces were filtered between 0.01 and 1 Hz, then aligned through multichannel cross correlation [*Van Decar and Crosson*, 1990]. Uncertainties associated with each travel time were estimated from the variances obtained during the least squares solution of the systems of linear equations for relative arrival times [*Van Decar and Crosson*, 1990].

3. Probabilistic Inversion

As explained in *Afonso et al.* [2013b], the entire inversion is based on a Markov chain Monte Carlo (MCMC) approach and involves two main stages: In the first exploratory stage, we subdivide the 3-D volume into individual 1-D rectangular columns (140 1° × 1° columns in this study) and invert a subset of the entire data vector in each column using those observables that are most sensitive to the 1-D subsurface structure (i.e., surface-wave dispersion curves, surface heat flow, geoid height under a 1-D approximation, and absolute elevation assuming lithospheric isostasy). The vertical parameterization of each column is illustrated in Figure 2b and described in the next section. Relatively large theoretical uncertainties are assigned to the forward solvers to compensate for the 1-D approximation and avoid overrestricting the parameter space during

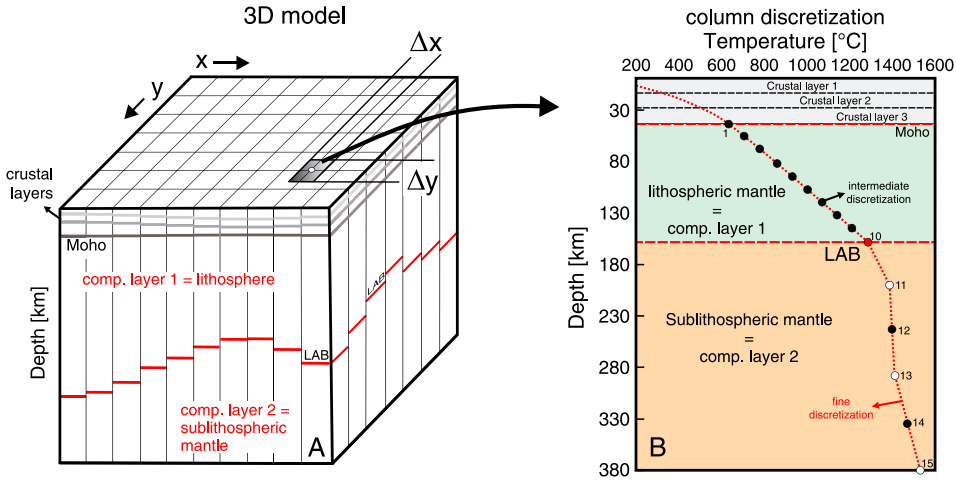

Figure 2. Discretization of the 3-D volume into its constituent columns (modified from *Afonso et al.* [2013b]). Each column has three discretization scales used in solving the forward problems. The finest scale (represented as red points along the geotherm in Figure 2b) is used to perform all computations related to the forward problems (e.g., geotherm computation, dispersion curves, etc); the internodal distance of this mesh is typically $\leqslant$2 km. The intermediate scale is represented by the so-called "thermodynamic nodes" (circles along the geotherm), at which the energy minimization problem is solved; the properties obtained at these nodes are linearly interpolated to the nodes of the fine scale mesh when solving the forward problems. The coarsest scale is represented by the mantle compositional layers and the crustal layers [see *Afonso et al.*, 2013b]. Sublithospheric thermodynamic nodes in white (11, 13, and 15) are nodes at which the temperature is a free parameter retrieved by the inversion; temperatures at intermediate nodes (i.e., 12 and 14) are computed by linear interpolation from the values at these nodes. Note that as the LAB depth is itself a free parameter, the depth of the thermodynamic nodes (and compositional layers) changes in each iteration of the MCMC inversion.

this exploratory stage [cf. *Afonso et al.*, 2013b]. All free parameters included in, and retrieved by, the inversion are listed in Table 1.

We set the initial parameter space large enough to include all plausible parameter configurations (see next section; Table 1). Since most of the observables used in this stage are strongly sensitive to the main model parameters (e.g., temperature structure, Moho depth, etc.), the initial parameter space is greatly reduced once the first inversion stage is completed. During this stage, we use the Delayed Rejection Adaptive Metropolis algorithm [*Haario et al.*, 2006] to obtain a population of acceptable models (i.e., realizations) that are broadly consistent with the aforementioned data. Numerous tests indicate that, at this stage, ~300,000 realizations per column are enough to reach stationarity of the chain. To be conservative, a total of 400,000 realizations were run for each column; a typical acceptance rate is ~35–40%.

Once the first stage is completed, we use the resulting population of acceptable models as priors in the second full 3-D stage ("refinement stage" of *Afonso et al.* [2013b]), where the entire data vector (gravity anomalies, 3-D geoid anomalies, gravity gradients, body wave travel times, etc.) is inverted using a standard Metropolis-Hastings algorithm. The initial 3-D model (i.e., starting point of the simulation chain) is constructed by assembling the maximum likelihood models of each 1-D column. New models are obtained by randomly choosing a column and sampling from its prior distribution one parameter at a time to guarantee a satisfactory acceptance rate. Effectively, this process "updates" our prior belief (the previous posterior) by adding new information to the problem [cf. *Kitanidis*, 2011; *Biegler et al.*, 2011; *Afonso et al.*, 2013b]. The final posterior distribution (i.e., a joint probability function in the parameter and data space) obtained after this second stage therefore represents our best "state of knowledge" and constitutes the most general solution to our inverse problem. A step-by-step description of the entire procedure is included in Figure 3 and details are given in *Afonso et al.* [2013b]. The total number of realizations needed to achieve stationarity during this stage is $\sim 7 \times 10^6$, with a typical acceptance rate of ~30–35%.

In contrast to traditional tomography techniques, the primary variables in our inversion are temperature, major-element composition, and pressure rather than seismic velocity, electrical conductivity, or density. These "secondary parameters" are obtained for each specific realization of the primary variables (i.e., a specific thermochemical Earth model) by solving a Gibbs free-energy minimization problem

Table 1. Parameters and Priors in Each Column Making Up the 3-D Domain[a]

Parameter[b]	Minimum	Maximum	Type
LAB (km)	45	330	U
$Al_2O_3^{lith}$ (wt %)	0.5	4.5	U
FeO^{lith} (wt %)	6.0	10.0	U
$Al_2O_3^{sublith}$ (wt %)	1.5	4.5	U
$FeO^{sublith}$ (wt %)	6.0	10.0	U
T_{410} (°C)	1330	1660	U
T_{int} (°C)	1220	1660	U
$T_{sublith}$ (°C)	1200	1500	U
STP density[1] (kg m^{-3})	2000	2700	U
STP density[2] (kg m^{-3})	2400	2820	U
STP density[3] (kg m^{-3})	2650	3150	U
Vp/Vs[1]	1.65	2.2	U
Vp/Vs[2]	1.65	1.83	U
Vp/Vs[3]	1.65	1.83	U
Δh^1 (km)	−0.5	0.5	G
Δh^2 (km)	−6.0	6.0	G
Δh^3 (km)	see notes	see notes	G
RHP (μW m^{-3})	0.4	1.8	U

[a]Listed minimum and maximum values are true limits for prior distributions of type U (uniform) and 1σ for distributions of type G (Gaussian). Parameters relevant to the three-layer crust are separated by a horizontal line.

[b]The compositional priors for major elements MgO and CaO are identical to those presented in Afonso et al. (2013a), which are functions of Al_2O_3 content; SiO_2 is computed as 100 - ($MgO+Al_2O_3$ +CaO+FeO). T_{410}, T_{int}, and $T_{sublith}$ are the temperatures of thermodynamic nodes 15 (bottom of model), 13 (intermediate sublithospheric node), and 11 (first sublithospheric node) in Figure 2. STP density^i are the bulk densities for each crustal layer at standard PT conditions (i.e., surface). These values can be easily associated with specific compositions. The actual vertical density structure used in each forward model is a function of the current geotherm and pressure distribution according to the assumed coefficient of thermal expansion (CTE) and compressibility (β) for each crustal layer i; these are listed in Table 2. Δh^i are the thickness variations allowed for each crustal layer i (i=1 is upper crust/sediments, i=2 is middle crust and i=3 is lower crust). The initial sediment cover and total crustal thickness is taken from the work of Shen et al. (2013). In the case of Δh^3, we use the uncertainties reported in Shen et al. (2013). RHP is the mean volumetric radiogenic heat production of crustal layers 1 and 2 (for layer 3 we assume a constant value of 5×10^{-7} W m^{-3})

[cf. *Connolly*, 2009; *Afonso et al.*, 2013a, 2013b; *Khan et al.*, 2011, 2016; *Kuskov et al.*, 2006]. Therefore, traditional tomographic images (e.g., *S* wave velocity structure) become a ubiquitous by-product of our inversion rather than the main result, with the important difference that our tomographic images explain a large number of independent observations. Note also that since we use a Bayesian approach, the solution to the inversion problem is represented by a posterior probability density function (PDF) rather than a single model. The codes to perform the entire inversion are available from the first author upon request.

3.1. Model Discretization and A Priori Information

We retain the same model discretization as in our previous work [*Afonso et al.*, 2013b] but with a new parameterization for the crust. The 3-D volume (and each constituent 1-D column) is discretized at three different scales. The finest discretization scale makes up the finite-difference mesh (computation nodes) used in the numerical solution of the forward problems, both in 1-D and 3-D. The resolution of this scale is typically set to $\leq$2 km. The intermediate discretization scale (thermodynamic nodes) is used to calculate equilibrium assemblages and associated thermophysical properties (e.g., seismic velocities, compressibility, bulk density, etc.) by free-energy minimization at specific P-T-C conditions. In this study we use 15 thermodynamic nodes; 10 in the lithosphere and 5 in the sublithospheric mantle. All thermophysical properties computed at the

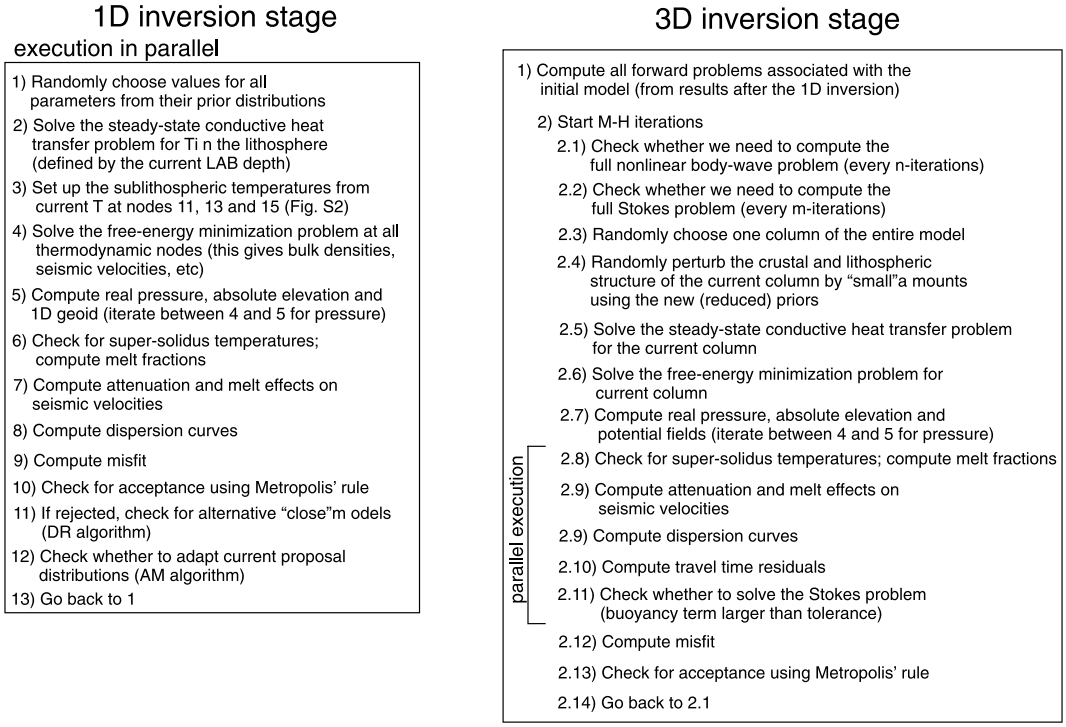

Figure 3. Summary of the steps involved in the computational implementation of the two stages of the multi-observable probabilistic inversion. In the 1-D stage, each column is computed on a separate processor. DR = Delayed Rejection, AM = Adaptive Metropolis, M-H = Metropolis-Hastings.

thermodynamic nodes are linearly interpolated to the computation nodes (fine mesh) during the solution of the forward problems. The third and coarsest discretization scale (compositional layers) refers to the actual number of layers with independent and constant bulk composition (i.e., specific CFMAS compositions) in the model. In this work we use two layers to represent the bulk compositions of the lithospheric and sublithospheric mantle independently (Figure 2b).

The parameters used to define a particular realization are the thermal thickness of the lithosphere (i.e., depth to the lithosphere-asthenosphere boundary or LAB), the major-oxide compositions of each of the two compositional layers (Figure 2), and a number of crustal parameters that we describe below. We use the depth to the LAB as an independent parameter because it controls the first-order thermal structure of the column, which in turn controls the main signatures of all observables [*Afonso et al.*, 2013a, 2013b]. Note that the four compositional parameters (i.e., SiO_2 is computed as 100-CaO-MgO-FeO-Al_2O_3) needed to describe compositions within the five-oxide CFMAS system are not all independent. According to observations in natural samples, CaO and MgO are correlated with Al_2O_3 in peridotitic rocks and, therefore, the only "truly" independent oxides are Al_2O_3 and FeO. These correlations are well known to mantle petrologists and geochemists and have been summarized in *Afonso et al.* [2013a] and *Afonso et al.* [2016] in the context of probabilistic inversions. We refer the reader to these works for more details on the sampling strategy.

In choosing the optimal parameterization for the crust, we adhere to two basic principles: (a) to ensure that the parameterization is as general and as realistic as possible, and (b) to seek parameterizations that describe the vertical structure with the minimum number of unknowns needed to explain the data. As a compromise to these two principles, we subdivide the crust into three layers (upper crust/sediments, middle crust, and lower crust; see Figure 2), each with its own set of fundamental parameters: coefficient of thermal expansion (CTE), isothermal compressibility (β), thermal conductivity, bulk density at surface conditions $\rho(P_o; T_o)$, volumetric radiogenic heat production (RHP), thickness, and Vp/Vs ratio. During the inversion, only the last four are treated as unknowns. This is justified on the grounds that model predictions are highly sensitive to changes in these last four parameters, but only mildly sensitive to realistic changes in the other three (i.e., varying these parameters results in changes that are within the uncertainty of the posterior PDFs). Table 2 lists the values of CTE, compressibility, and thermal conductivity used in our inversion. We therefore do not solve the energy

Table 2. Crustal Parameters With Constant Values[a]

Parameter	Value
CTE^1 ($°C^{-1}$)	2.7×10^{-5}
CTE^2 ($°C^{-1}$)	2.6×10^{-5}
CTE^3 ($°C^{-1}$)	2.3×10^{-5}
k^1 ($W\,m^{-1}°C^{-1}$)	2.8
k^2 ($W\,m^{-1}°C^{-1}$)	2.6
k^3 ($W\,m^{-1}°C^{-1}$)	2.3
β^1 (Pa^{-1})	5.0×10^{-11}
β^2 (Pa^{-1})	1.5×10^{-11}
β^3 (Pa^{-1})	1.2×10^{-11}

[a]CTE = coefficient of thermal expansion; k = thermal conductivity; β = isothermal compressibility.

minimization problem within crustal layers. Instead, we relate their densities, Vp and Vs, assuming Birch's law of correspondent states [cf. *Karato*, 2008].

The PT-dependent bulk density at each node within the crust is obtained as

$$\rho_{(P,T)} = \rho_{(P_o,T_o)}\,[1 - \alpha(T - T_o) + \beta(P - P_o)] \quad (1)$$

The corresponding Vp can be computed assuming the relation

$$\mathrm{Vp\,(km/s)} = 39.128\rho - 63.064\rho^2 + 37.083\rho^3 - 9.1819\rho^4 + 0.8215\rho^5 \quad (2)$$

This is the relation proposed by *Brocher* [2005] with the last term modified to better represent crystalline mafic rocks. Finally, Vs in the crust is obtained from the computed Vp (equation (2)) and the current Vp/Vs sampled from the prior by the MCMC algorithm. Note that since the isothermal compressibility is fixed, one could obtain Vs from equations (1) and (2) only (i.e., no need for treating Vp/Vs as a variable) by assuming that the isothermal and adiabatic compressibilities are identical. Although this is usually a good approximation for Earth materials, this approach will cause our results to be entirely controlled by equation (2). Since real rocks do not follow equation (2) perfectly, but exhibit a relatively large variability around it, we prefer to keep Vp/Vs as a separate independent variable to account for the observed natural variability.

Priors for all parameters (for each individual column) are listed in Table 1. These priors are large enough to cover most of the expected variability in model parameters. In particular, the priors related to the compositional and thermal structure cover more than 2σ of the variability observed globally [*Griffin et al.*, 2009; *Afonso et al.*, 2013a].

3.2. Forward Problems

All forward problems solved during the inversion (heat transfer, potential fields, surface wave velocities, isostasy, free-energy minimization, etc.) have been described in detail in parts I and II [*Afonso et al.*, 2013a, 2013b] and in *Fullea et al.* [2015], except for the body wave tomography and Stokes flow problems. Therefore, in what follows, we only provide a brief summary of previously described forward problems and focus on the new implementations of body wave tomography and Stokes flow into the probabilistic inversion framework.

3.2.1. Summary of Main Forward Problems

Once a particular realization of LAB depth is chosen by sampling its prior PDF, the conductive steady state heat transfer equation is solved within the lithopsheric domain subject to Dirichlet boundary conditions at the surface (typically $T_s = 10°C$) and at the LAB ($T_{LAB} = 1250°C$). Given the close relationship between the mechanical strength of rocks and their temperature, we define the LAB as the depth to the 1250°C isotherm [cf. *Afonso et al.*, 2016]. This is consistent with results from numerous numerical simulations of mantle-like fluids with realistic viscosity laws [e.g., *van Wijk et al.*, 2010; *Ballmer et al.*, 2011; *Afonso et al.*, 2008b; *Zlotnik et al.*, 2008; *Gerya*, 2010] as well as with results obtained in this study (see section 5). While the original parameterization of *Afonso et al.* [2013b] allows for perturbations to steady state geotherms, we do not consider this case here. In the sublithospheric mantle, temperature is free to vary (according to data fits) in three of the five thermodynamic nodes (white circles in Figure 2b): the topmost (11), the intermediate (13), and the bottom node (15). A continuous geotherm over the fine mesh is then obtained by linear interpolation of these three values and the node at the LAB (10).

With the temperature structure already computed, the Gibbs free-energy minimization problem is solved at the thermodynamic nodes using an initial "master" pressure profile $P(z)$ and the corresponding compositions assigned for each node. This, however, is a nonlinear problem as the actual pressure profile depends on the density structure obtained from the energy minimization problem. Therefore, we use an iterative scheme in which the pressure profile is updated at each iteration using the density structure from the previous iteration. The density-pressure iterations are stopped when the difference in calculated pressure between two

successive iterations becomes less than 50 MPa in all nodes, which is equivalent to an uncertainty of <2 km in the node depth.

The computation of 3-D geoid height, gravity anomalies, gravity gradients, and absolute elevation have been described in detail in several of our previous works [e.g., *Afonso et al.*, 2008a, 2013b; *Fullea et al.*, 2009, 2015], and therefore, we do not repeat it here. We have only slightly changed the 1-D geoid approximation (needed in the first stage of the inversion) with respect to our previous works; we discuss these modifications in Appendix A. We also neglect flexural effects in this study given their small contribution at the current discretization scale [*Turcotte and Schubert*, 1982].

Anharmonic elastic moduli (or seismic velocities) are corrected for anelasticity and melt content (see sections 3.3 and 3.4) before computing dispersion curves and travel times. Surface wave dispersion curves are computed with a modified version of the *disp*96 code [*Herrmann*, 2002], using a layered structure based on the fine mesh. Explicitly, we subdivide each node of the fine mesh into 10 sublayers with constant properties obtained by interpolation between two consecutive nodes of the fine mesh. Therefore, a typical discretization scale for the calculation of dispersion curves is $\lesssim$200 m. For depths below the 410 km discontinuity (limit of our physical/numerical domain) and down to 2000 km, we use the parameters from the AK135 model [*Kennett et al.*, 1995].

3.2.2. Body Wave Tomography

Given the time involved in computing teleseismic travel times in a 3-D volume (a nonlinear problem), the implementation of grid-based eikonal solvers within probabilistic inversions is challenging, particularly when solving it together with other observables such as gravity potential or surface wave data. We tackle this problem with a strategy that, albeit simple, provides reliable results.

The key idea is not to solve the full nonlinear problem at every step of the MCMC random walk, but only every n steps instead. In between nonlinear solutions, a simpler and much faster linear approximation is adopted, in which travel time residuals are computed by a simple integration of slowness along each fixed ray path. The nonlinear problem, which outputs both ray paths and theoretical travel time residuals for a given velocity structure, is solved with a modified version of the Fast Marching Method [*Rawlinson et al.*, 2006]. Outside the local (target) volume, the Earth is assumed to be spherically symmetric, which allows rapid computation of *ak*135 travel times from the sources to the boundaries of the volume. More details can be found in *Rawlinson et al.* [2006].

Although our locally linear approximation assumes that the geometry of ray paths is independent of the velocity structure, this is reasonable for teleseismic paths, which tend not to deviate significantly when low-amplitude lateral structure is present. A formal proof that only second-order errors in travel times are introduced by this assumption follows directly from Fermat's principle [e.g., *Nolet*, 1987]. Moreover, by updating the paths every n steps as the inversion proceeds, the nonlinear behavior of the system is ultimately explicitly considered. However, the success of this approach relies on the fulfillment of two crucial requirements. First, the linear approximation must compute travel times along the ray paths computed in a previous full nonlinear computation. Second, and most importantly, the model must not change significantly between each nonlinear step, in order to guarantee that the ray paths computed in the last nonlinear step are still a valid approximation for the current model. Our method is specifically designed to satisfy these two requirements, and our results show that the inversion always converges to good solutions. The actual value of n needs to be chosen based on preliminary tests [*Afonso et al.*, 2013b], as it depends on the physical features of the region under study. Here we used $n = 35,000$, which means that we solve the full nonlinear problem >40 times during the second stage of the inversion.

3.2.3. Stokes Problem

The rationale used here to estimate first-order dynamic effects on surface observables relies on the well-known premise that in a highly viscous fluid with a strongly temperature-dependent viscosity (i.e., mantle-like fluid), its surface topography is primarily controlled by the density structure within the upper, conductive, thermal boundary layer (i.e., thermal lithosphere) [e.g., *Torrance and Turcotte*, 1971; *Jarvis and Peltier*, 1982; *Parsons and Daly*, 1983; *Buck*, 1986; *Buck and Parmentier*, 1986]; the success of thermal isostasy in explaining the topography of rifted areas and of the ocean floor is testament to the validity of this premise. Density anomalies below the "rigid" (and conductive) lithosphere contribute to the surface topography only via normal stresses associated with viscous flow and transferred to the lithosphere via viscous coupling. We therefore explicitly separate the contribution of sublithospheric viscous flow (truly dynamic effect) from that

related to the thermochemical buoyancy of the lithosphere [*Torrance and Turcotte*, 1971]. The logic borrows from what happens when a block of dense solid wood of nonnegligible thickness is placed in a tank containing a viscous convecting medium of higher density (e.g., by heating one side of the tank). The largest (first-order) topographic adjustment (sinking of the block) is controlled by the purely isostatic component associated with the local buoyancy difference between the wood and the substratum. The underlying convecting pattern represents a perturbation to the isostatic component not related to the intrinsic buoyancy of the block.

The implementation of dynamic effects arising from sublithospheric mantle flow into our inversion scheme follows a similar logic to that used in the travel time tomography problem (i.e., full solution is computed every n simplified solutions). We start by computing the instantaneous flow corresponding to the initial 3-D thermochemical structure at the beginning of the second stage (i.e., the maximum likelihood model after the 1-D inversion stage). This is done by first extending the computational domain along the three Cartesian axes (to minimize boundary effects) and then solving the discrete (finite differences) matrix system using an iterative multigrid method for variable viscosity [*Gerya*, 2010].

Dynamic topography induced purely by sublithospheric flow is calculated by assuming zero vertical velocity (i.e., free slip) at all nodes above a critical conductive geotherm and computing the vertical components of the stress tensor at the deepest of such nodes across the entire model. The rationale behind this assumption is that above a critical geotherm, temperatures are low enough (or viscosities high enough) to make flow negligible compared to that taking place in the underlying mantle. We note that this approach is similar to that of *Buck* [1986], *Buck and Parmentier* [1986], and *Marquart and Schmeling* [1989], among others, and has been shown to be a good approximation to more sophisticated approaches at times longer than the isostatic relaxation time [*Crameri et al.*, 2012]. Here we take 1150°C as the critical conductive geotherm, consistent with results from numerous studies [e.g., *Buck and Parmentier*, 1986; *van Wijk et al.*, 2010; *Ballmer et al.*, 2011; *Afonso et al.*, 2008b; *van Hunen et al.*, 2005; *Kaislaniemi and van Hunen*, 2014]. For the convecting sublithospheric material, we assume Newtonian rheology with a preexponent factor $A = 5.3 \times 10^{15}$, an activation energy $E = 250$ kJ mol^{1} and an activation volume $V = 4 \times 10^{-6}$ m^3 mol^{1}. Similar values have been extensively used in earlier studies [e.g., *van Wijk et al.*, 2010; *Ballmer et al.*, 2011; *Afonso et al.*, 2008b; *van Hunen et al.*, 2005].

The step described above represents the full Stokes problem that we *do not* want to solve at every iteration, as it would render our probabilistic method impractical (i.e., too computationally expensive). Therefore, after the first complete flow field and boundary stresses are computed, the simpler and faster *Stokeslets* method [e.g., *Morgan*, 1965; *Harper*, 1984] is adopted to compute the dynamic effect of subsequent random perturbations of the 3-D model. This is an acceptable approximation, as the random perturbations sampled during the MCMC simulation are small and local by design [*Afonso et al.*, 2013b], and therefore, they create small departures $\Delta\rho$ from the background density field. Departures $\Delta\rho$ are converted to equivalent point forces and their associated normal stresses at the critical nodes are computed using the method of images [e.g., *Blake and Chwang*, 1974; *Morgan*, 1965; *Harper*, 1984] and added to those from the last accepted 3-D solution. In practice, however, many perturbations result in flow variations below the actual numerical tolerance used to solve the discrete 3-D Stokes problem (i.e., velocities and stresses from the last accepted model are very similar to the current ones). Therefore, we only update stresses if the random perturbation produces a buoyancy change larger than a prescribed tolerance. As with the body wave problem, a full 3-D problem is recomputed every n of these fast updates. Numerous trial tests indicate that values of $4000 < n < 25,000$ give satisfactory results; here we used $n = 4000$. We are currently working on a more accurate and efficient way of solving the Stokes problem based on Reduced Basis Modeling [cf. *Quarteroni et al.*, 2016], and results will be presented in a forthcoming publication.

Although we have not formally inverted for viscosity in this work (this is part of a current study), we did run preliminary tests using higher and lower activation energies (i.e., more or less viscous sublithospheric mantle) and found that acceptable joint fits to all the observables cannot be achieved with values of E that result in average asthenospheric viscosities $\gtrsim 2 \times 10^{20}$ Pa s. This range is well within commonly accepted values (e.g., *van Wijk et al.* [2010], *Ballmer et al.* [2011], *Afonso et al.* [2008b], and *Kaislaniemi et al.* [2014], among others).

3.3. Melting Model

Melting in the mantle and crust as functions of inverted temperatures are estimated by defining the following solidus and liquidus temperatures (in °C):

$$T_S^C = 920 + 156P \tag{3}$$

$$T_L^C = 1120 + 156P \tag{4}$$

$$T_S^M = 1080 + 134.2P - 6.581P^2 + 0.1054P^3 \tag{5}$$

$$T_L^M = 1762 + 57.46P - 3.487P^2 + 0.077P^3 \tag{6}$$

where T_S^C is the solidus of crustal rocks, T_L^C the liquidus of crustal rocks, T_S^M the solidus of mantle rocks, T_L^M the liquidus of mantle rocks, and P is pressure in GPa. These solidus temperatures are representative of dry granitic rocks and dry peridotites [*Hirschmann*, 2000; *Winter*, 2010], and therefore, partial melting predicted by our model should be taken as indicative only. Despite this rough estimate, we will show (section 5) that our predictions agree remarkably well with results from other studies and evidence of recent magmatism.

With equations (3)–(6), melt fractions can be computed as

$$M = \frac{(T - T_S)}{(T_L - T_S)} \tag{7}$$

where T is the actual temperature of models drawn during the inversion. We neglect the effect of melt on bulk density reduction, as it is always small compared to the uncertainties in data sets constraining this property. The effects of melts on seismic velocities, on the other hand, are significant and we estimate them based on the results of *Hammond and Humphreys* [2000]. In the absence of more detailed information, we chose the average of the minimum and maximum values proposed by these authors for Vs and Vp velocity reductions for melt fractions ≤1%: (−5.3 = dlnVs/%melt and −2.4 = dlnVp/%melt). Over 99% of our predicted melt fractions are below 1% (see section 5), so these derivatives are considered adequate.

3.4. Anelastic Attenuation

Any attempt to relate information about seismic velocities to the physical state of the Earth's interior must explicitly include the effects of anelasticity [e.g., *Karato*, 1993, 2008; *Goes et al*., 2000; *Priestley and McKenzie*, 2006; *Jackson and Faul*, 2010; *Afonso et al*., 2010; *Khan et al*., 2011]. Anelastic effects (e.g., seismic attenuation and velocity dispersion) are commonly incorporated in tomography studies by using the so-called Q factor. Typically, the latter is considered to be either constant (i.e., 1-D constant profile from a global model) or laterally variable but independent of frequency [cf. *Kennett*, 2002]. Laboratory experiments with olivine samples, however, predict complicated behaviors of Q with frequency, grain size, fugacity, fluid content, pressure, and temperature [cf. *Karato*, 2008; *Jackson and Faul*, 2010]. Unfortunately, although the temperature effect on attenuation in olivine-rich samples is relatively well understood, large uncertainties remain regarding the effects of other factors. Moreover, it is difficult (impossible in some cases) to use available seismic data to validate some of these effects.

Partly due to these complications, there have been a number of different approaches in the literature to infer the physical state of the upper mantle from regional and global seismological data/models. Two are most relevant to our discussion (others include, e.g., *Goes et al.* [2000] and *Cammarano et al.* [2003]). One approach adheres strictly to laboratory-based models of attenuation, which are supported by sound physical arguments, while assuming constant "realistic" values for some parameters that are not well constrained [e.g., *Afonso et al.*, 2008a; *Jackson and Faul*, 2010; *Khan et al.*, 2011; *Fullea et al.*, 2012]. The second approach is based on semiempirical correlations between seismological observations and thermal models of the lithosphere [e.g., *Priestley and McKenzie*, 2006]. Both approaches have their strengths and limitations. The laboratory-based approach has the advantage of relying on solid physical theories that can, in principle, be tested with laboratory experiments. However, it is unlikely that the restricted physicochemical conditions at which experiments are performed in the laboratory are representative of the wide range of expected

conditions within the Earth's interior. The semiempirical approach has the advantage of relaxing some conditions/parameters to better reproduce seismological observations. However, there is no a priori reason why relationships biased toward thermal *models* of oceanic mantle and stable cratonic areas should apply worldwide. In fact, there is evidence that attenuation and physical dispersion of seismic waves behave differently depending on the tectonic environment and lithospheric structure [e.g., *Dalton and Faul*, 2010].

In this work, we adopted a somewhat intermediate approach in which we adhere to a microphysical model validated by laboratory experiments at conditions pertaining to the upper mantle and seismic frequencies, but we allowed the multiple data sets to determine the value of some poorly constrained parameters. Specifically, we compute attenuation effects with the extended Burgers formalism of *Jackson and Faul* [2010] while simultaneously inverting for the peak of elastically accommodated grain boundary sliding, ΔP, and bulk activation volume. Inverting for ΔP is of particular importance, as it exerts a major control over the actual reduction of seismic velocities predicted by the model, yet it is a parameter subject to considerable experimental uncertainties. We discuss this further in section 5.2.

Since *Jackson and Faul*'s [2010] model is frequency dependent, we account for the different resultant $Q(f)$ factors and observational frequencies when computing dispersion curves and body wave travel times. The average grain size is kept constant and equal to 4 mm during the inversion, consistent with many previous estimates [cf. *Pearson et al.*, 2003; *Karato*, 2008; *Faul and Jackson*, 2005; *Afonso et al.*, 2008a].

3.5. Misfit Functions

Data sets used in this work are sensitive to different aspects of the problem and have different units, uncertainties, and data points. The simultaneous inversion of these data requires an adequate weighting of each data set to avoid solutions that are dominated by nonphysical factors, such as the total number of data points or large absolute values of specific data sets. However, the question of how to weight the different data sets in joint inversions, and how to set up a representative misfit function, does not have a straightforward answer. Strictly, multiobservable inversion is a multiobjective or vector optimization problem [cf. *Marler and Arora*, 2004]. Unfortunately, at present there are no satisfying/practical multiobjective theories in the context of multiobservable probabilistic inversions; characterizing the Pareto front and solving the statistical inference problem are still treated as separate tasks. Moreover, in multiobservable probabilistic inversions, the final misfit function depends not only on the statistics of the data sets (e.g., number of data points, uncertainties, etc.) but also on the actual goals of the inversion (e.g., which data set is more sensitive to the sought parameters).

In lieu of a more general approach, we choose to adopt the well-known weighted-sum approach. We use two different but related misfit functions in the two stages of the inversion described in section 3. In the first stage, we invert elevation, dispersion curves, geoid height (under the 1-D approximation), and surface heat flow. Therefore, we adopt the following misfit function:

$$E_{1D} = \sum_i \left[\frac{(d_i - c_i)}{\sigma_i}\right]^2 + \frac{1}{(n_j - W)} \sum_{j=1}^{j=n_j} \left[\frac{(r_j - r_j^c)}{\sigma_j}\right]^2 \tag{8}$$

where d_i is observed data and c_i is predicted data for i = elevation, geoid, surface heat flow, and σ_i are their respective observational uncertainties (section 2). The parameter n_j is the total number of data points (20 in this study) used to sample the Rayleigh dispersion curves at specific periods, r_j^c is the predicted phase velocity at period j, r_j the observed phase velocity at period j, and σ_j their respective uncertainties. For each observable i, the first term in equation (8) will be $\lesssim 1$ for optimal fits and no additional normalization is needed (their uncertainties act as "natural" normalization terms). For dispersion curves, however, optimal values resulting from adding the individual misfits at each period j will amount to a value close to n_j; thus the need to equalize this term with a factor $1/n_j$. We note that such equalization is not strictly necessary in probabilistic approaches. In practice, however, it can sometimes help in improving the efficiency of the MCMC algorithm and its statistical estimators. The additional parameter W is a subjective weighting factor that allows us to better control and test the overall dominance of dispersion data over other data sets with flexibility and simplicity. The resulting combined weighting factor $1/(n_j\text{-}W)$ is only one of many possible options to achieve this and is adopted here on convenience grounds only. Exploratory tests indicated that values $2 < W < 10$ give satisfactory results in terms of global misfit minimization. In the present study we use $W = 4$.

During the second refinement stage of the inversion (full 3-D), we adopt the following global misfit function

$$E_{3D} = \sum_{m=1}^{m=N_c} \left\{ \sum_i \left[\frac{(d_i - c_i)}{\sigma_i} \right]^2 + \frac{1}{(n_j - W)} \sum_{j=1}^{j=n_j} \left[\frac{(r_j - r_j^c)}{\sigma_j} \right]^2 \right\} + \frac{\mathrm{bw}}{N_s} \sum_{k=1}^{k=N_s} \left[\frac{(\delta ts_k - \delta ts_k^c)}{\sigma_{sk}} \right]^2 + \frac{(1 - \mathrm{bw})}{N_p} \sum_{l=1}^{l=N_p} \left[\frac{(\delta tp_l - \delta tp_l^c)}{\sigma_{pl}} \right]^2 \quad (9)$$

where N_c is the total number of columns making up the 3-D volume, d_i and c_i are as in equation (8), but now i = Bouguer anomalies, gravity gradients, elevation, geoid height, surface heat flow, and σ_i are their respective uncertainties; N_s and N_p are the total number of relative travel time residuals for S and P phases, respectively; δts_k = observed residuals for S phases; δts_k^c = predicted residuals for S phases; δtp_l = observed residuals for P phases; δtp_l^c = predicted residuals for P phases. Parameter bw is a weighting factor ranging from 0 to 1 that permits control over the relative contributions of each body wave data set (*P* and *S*) to the overall misfit function. Since body wave data is implicitly considered a *single type* of data, in addition to normalizing the misfits by the total number of residuals (N_s and N_p) we also enforce that their individual weights sum to 1. Based on numerous trial tests for our specific data sets, we chose bw = 0.3 for this study.

4. Sensitivity and Data Fits

4.1. Sensitivity of Data to Different Parameters of the Models

The main motivations for combining different complementary data sets in our inversion is to minimize the range of acceptable models consistent with available data and to increase the sensitivity to temperature and compositional heterogeneities. These two thermodynamic variables control the bulk properties of rocks such as elastic moduli and bulk density; pressure is less important for lateral anomalies. Therefore, the internally consistent use of independent data sets differentially sensitive to density and seismic velocity anomalies is one of the most reliable approaches to separate temperature from bulk compositional effects. This is commonly exemplified by the behavior of the ratio $\partial ln\rho/\partial lnVs$, which is negative (positive) when bulk composition (temperature) is the main cause for changes in Vs and ρ [e.g., *Forte and Perry*, 2000; *Karato and Karki*, 2001; *Deschamps et al.*, 2002; *Afonso et al.*, 2010]. For peridotitic systems, this is true regardless of whether spinel or garnet are the stable Al-rich phase [*Afonso et al.*, 2010] or whether melt depletion has been isobaric or polybaric [*Afonso and Schutt*, 2012]. When anelasticity is properly accounted for, the ratios $\partial ln\rho/\partial lnVp$ and $\partial lnVp/\partial Vs$ also offer complementary sensitivities [e.g., *Karato and Karki*, 2001; *Lee*, 2003; *Afonso et al.*, 2010]. We emphasize, however, that since both density and elastic moduli depend on bulk composition, any attempt to make inferences about the thermochemical state of the Earth's interior ultimately requires a framework that ensures an internally and thermodynamically consistent coupling between pressure, temperature, and composition.

In our inversion, gravity anomalies, gravity gradients, and geoid height all provide complementary sensitivities to density anomalies at different wavelengths and depths [cf. *Afonso et al.*, 2013a; *Panet et al.*, 2014; *Bouman et al.*, 2013; *Fullea et al.*, 2015]. Similarly, the combination of fundamental mode surface wave data and teleseismic travel times provides crucial complementary sensitivities to shallow and deep anomalies while at the same time allowing the useof absolute seismic velocities [e.g., *West et al.*, 2004; *Rawlinson and Fishwick*, 2011; *Obrebski et al.*, 2011; *Afonso et al.*, 2016]. The latter point is of paramount importance when using seismic information to assess the thermochemical structure of the Earth [cf. *Afonso et al.*, 2016]. Surface heat flow and elevation provide complementary, yet less stringent, constraints on the average thermal and density structure of the lithosphere.

To illustrate the impact of individual data sets on the thermochemical structure of the lithosphere and upper mantle in the study region, we invert four different combinations of data sets: (i) nonseismic data sets only, (ii) Rayleigh wave dispersion curves only, (iii) *P* wave relative travel time residuals only, and (iv) Rayleigh wave dispersion curves together with *P* wave travel time residuals. Additional sensitivity tests to model parameters can be found in *Afonso et al.* [2013b] and *Shan et al.* [2014]. These four inversions are run with identical priors, misfit functions, data uncertainties, and parameterizations as those used in the real full inversion described in section 5. However, we do not solve for either Stokes flow, gravity anomalies, or lateral components of the

Marussi tensor in combination (i) to keep computational time to a minimum while still providing instructive results. Note that given the model-based nature of our inversion scheme, in which the absolute velocity distribution is always known, we can relate the model's absolute velocities to relative travel time residuals and thus invert them directly for the thermochemical structure of the model. This would not be generally possible in traditional travel time tomography, where only velocity anomalies are recovered.

Figure 4 summarizes the results from these four inversions for LAB depth, lithospheric composition, and temperature structure. The individual data fits in all four inversions are excellent and generally better than when inverting all data sets simultaneously. The average standard deviations for each field are also included for comparison. Three main conclusions can be drawn from these results. First, nonseismic data sets alone struggle to provide a well-defined LAB that is consistent with seismic data. Moreover, the deep thermal structure is, as expected, highly uncertain (i.e., large σ). Note also that the average composition of the lithospheric mantle is significantly different from those retrieved by the other three inversions and subject to large uncertainties (i.e., flat posterior PDFs). These results highlight the well-known ambiguity of nonseismic data in constraining the deep thermochemical structure of the Earth. Although some improvements are expected when using the full tensor of gravity gradients, gravity anomalies, and a reduced number of parameters, the sensitivity to deep thermochemical anomalies will remain relatively low.

Second, the inversion of fundamental mode Rayleigh wave dispersion curves results in a significant improvement in the recovered LAB structure as well as the deep thermal structure (Figures 4e–4h). In particular, a significant variance reduction is achieved in LAB estimates and thermal structure. The same is not true, however, for the average composition of the lithosphere. In this case, not only is the pattern significantly different from that obtained by the inversion of nonseismic data but it also remains poorly constrained (i.e., large σ). This issue is aggravated when inverting *P* wave travel times alone (Figure 4 I-L), which results in a lower sensitivity to shallow thermal structure and almost no sensitivity to lithospheric composition. This can be understood, at least in part, by the low sensitivity of travel time tomography at depths shallower than the interstation distance (~70 km in this case) due to the poor ray crossover at these depths [cf. *Evans and Achauer*, 1993; *Rawlinson et al.*, 2006]. The fact that a large portion of the region seems to be characterized by a shallow LAB clearly exacerbates this problem. Conversely, the sensitivity to deeper temperature structure seems to be slightly better with *P* wave travel times than for normal mode surface wave data alone.

Lastly, the combination of surface wave data and *P* wave teleseismic travel times (with anelasticity taken into account) results in a considerable uncertainty reduction in all three fields plotted in Figure 4. Moreover, the LAB pattern is remarkably close to our final best estimate in section 5. Here we also begin to see a lithospheric compositional pattern which is more consistent with xenolith evidence and bears a close resemblance to our best final model (section 5). Note that while temperature variability at shallow depths is not reduced to any significant extent compared to the surface wave-only inversion, the deeper temperature structure is significantly more constrained in the joint inversion case. All these improvements can be understood in light of the higher sensitivity of Vp/Vs to compositional anomalies than Vp or Vs separately [e.g., *Karato and Karki*, 2001; *Lee*, 2003; *Afonso et al.*, 2010] and the increased sensitivity to deep anomalies gained by including body wave data.

The results summarized in this section illustrate well some important characteristics of surface wave and body wave travel time data inversion for the thermochemical structure of the lithosphere and upper mantle. On the one hand, surface wave data has high sensitivity to temperature structure, particularly at shallow depths, but no sensitivity to bulk composition (see also *Shan et al.* [2014]). *P* wave travel times alone, on the other hand, offer a slightly better sensitivity to deeper temperature anomalies, but cannot constrain composition or temperature structure at shallow depths. Taking advantage of the complementarity of both data sets in joint inversions delivers clear benefits. However, including additional data sensitive to bulk density remains strictly necessary to increase the sensitivity to compositional anomalies and allow for a more robust discrimination of temperature versus compositional signatures. We discuss this further when presenting the results from the full thermochemical tomography in section 5.

4.1.1. Global Misfits and Posterior PDFs

It is important to discuss some general features of the global misfit function and posterior PDFs before focusing on the main results. As expected, the quality of the simultaneous fit to all data sets (*P* and *S* teleseismic travel times, Rayleigh wave dispersion curves, geoid height, Bouguer gravity anomalies, diagonal components of the Marussi tensor, surface heat flow, and absolute elevation) varies spatially. A map of normalized

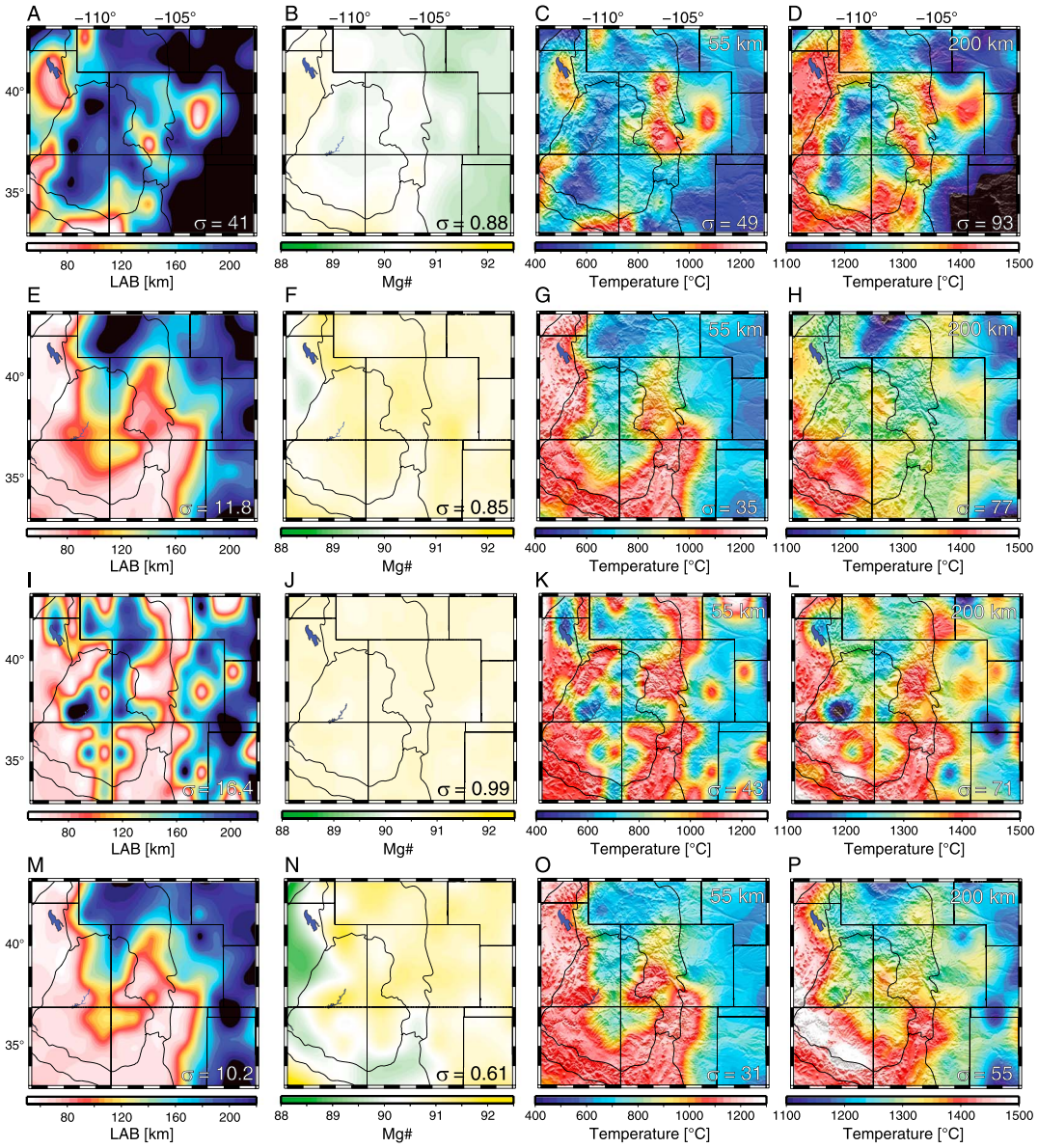

Figure 4. LAB depth, average lithospheric composition (as bulk Mg#), and temperature structure obtained from inverting specific subsets of the complete data vector used in this study. (a–d) Results from inverting geoid height, vertical component of the gravity gradient tensor, surface heat flow, and isostatic elevation. (e–h) Results from inverting normal mode Rayleigh wave dispersion curves for periods from 8 to 150 s. (i–l) Results from inverting *P* wave relative travel time residuals. (m–p) Results from jointly inverting normal mode Rayleigh wave dispersion curves and *P* wave travel time residuals. The average standard deviation (σ) for each field is included for comparison. The temperature field is shown for two depths (55 and 200 km depth). Figures 4m and 4n should be compared with Figures 7a and 7d.

global misfits (equation (9)) associated with the results presented in the next section is shown in Figure 5 as the median of the entire posterior PDF (the actual PDF is always non-Gaussian, but positively skewed with a tail toward larger misfit values). The absolute global misfits are normalized with respect to the minimum misfit value in the study area, which is found at 40.5°N–103.5°W. At this location, predictions given by our best models are within 1.7 σ of the observed data *for all* inverted data sets. Similar excellent fits are reached in all regions depicted as white in Figure 5. We note that misfits related to dispersion data tend to be the most important contribution to the global misfit, particularly in those areas close to the boundaries of the study region. This occurs primarily due to irregularities or "jumps" in the overlapping region of the dispersion curves (Figure 5b and also Figure 9f), where data from ambient noise and earthquakes are combined to obtain a single dispersion curve [*Yang et al.*, 2008; *Shan et al.*, 2014]. While in most cases both data sets produce smooth

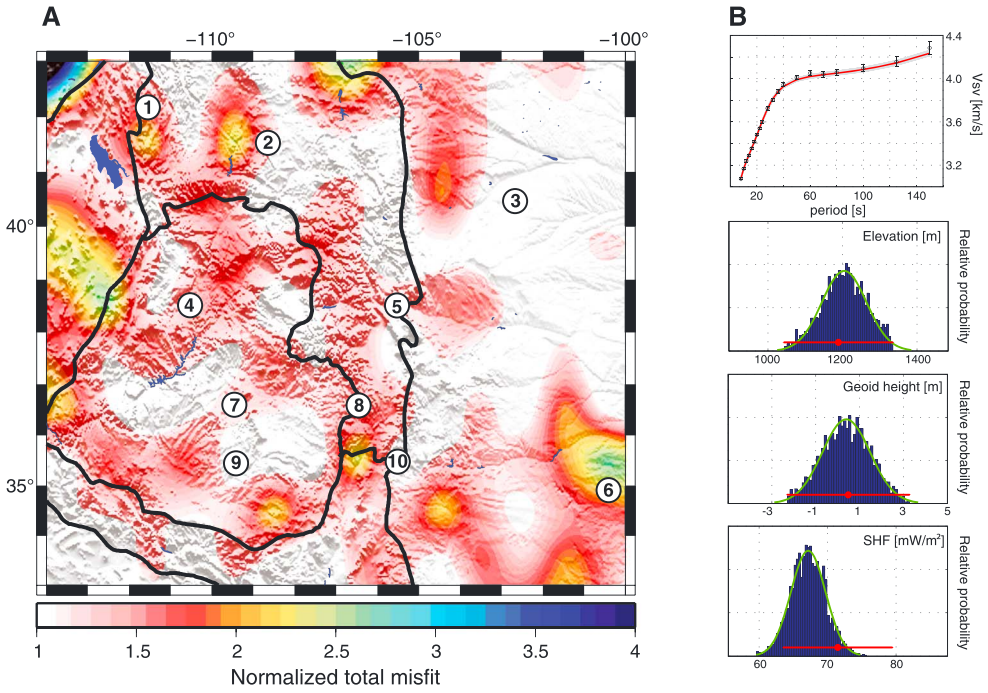

Figure 5. (a) Total misfit normalized to the minimum misfit found during the inversion. Circled numbers indicate the locations corresponding to the three examples of the full posterior PDF shown in Figure 6 and the additional seven examples in the supporting information. These locations were selected to illustrate the range of results afforded by the inversion. (b) Example of joint data fit for Rayleigh dispersion curve, elevation, geoid height, and surface heat flow at location 3 in Figure 5a. Red line in top panel is the mean of all models (in gray) in the posterior PDF. Blue histograms are posterior PDFs. Red dots and lines denote, respectively, the means and 2σ of observed data. Green lines are Gaussian distributions fitted to the posterior PDFs.

curves over the entire periods range, in some cases the overlapping region exhibits artificial "jumps" that the model cannot fit. Therefore, the actual misfit in Figure 5 may be considered an upper limit.

Figure 6 depicts (as 2-D joint distributions) three representative examples of the full posterior PDF for a subset of relevant model parameters. The locations of these PDFs are indicated in Figure 5. These examples were chosen because they depict the whole range of possible posteriors, from localized (i.e., well constrained) to very broad distributions (i.e., high uncertainty); additional examples are included in the supporting information. One can see that while most parameters are relatively well constrained (e.g., LAB depth, crustal density, mantle temperature, and lithospheric Mg#), others are more uncertain. In particular, the Vp/Vs ratio of the first crustal layer (Vp/Vs_1) is poorly constrained in most cases. This is not surprising, as although its properties can be highly heterogeneous, this first crustal layer is typically very thin (sediment cover in most columns), and therefore, its contribution to the overall fit of the observables is typically minor. Likewise, the bulk density of the first crustal layer is also poorly constrained. Future work implementing receiver function data will help improve the resolution of crustal structure (Appendix A).

5. Main Results and Some Implications

5.1. Thermochemical Structure of the Western-Central U.S.

Figures 7a and 8 show the depth to the thermal LAB (median of the posterior PDF) and the overall 3-D temperature structure of the study region, respectively. These figures reveal a highly heterogeneous lithospheric structure beneath Western-Central U.S. (several depth slices of the temperature structure are given in the supporting information). Shallow LAB depths (<100 km) are imaged beneath the Northern Basin and Range (NBR), Colorado Plateau/Basin and Range Transition Zone (CP-BR-TZ), Rio Grande Rift (RGR), and eastern Rocky Mountains (RM). The thin and hot thermal lithosphere imaged beneath the RM agrees well with the high temperatures suggested recently based on an independent thermal modeling of mantle seismic velocities [*Hansen et al.*, 2013]. A thick lithospheric reentrant or wedge-like structure, connected to the RM and Wyoming Craton to the north, dominates the central part of the CP. Note, however, that the average lithospheric

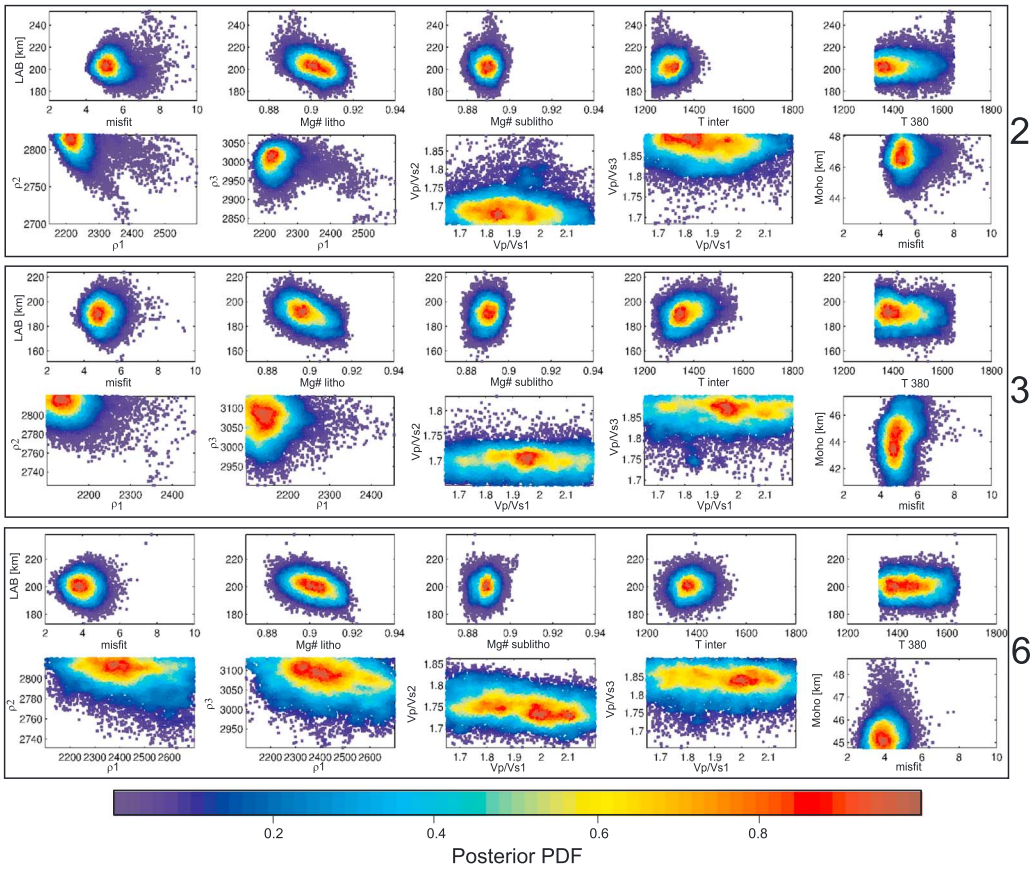

Figure 6. Examples of the full posterior PDF, plotted as joint probabilities between a few relevant parameters for three locations indicated in Figure 5. Mg# litho = average lithospheric Mg#; Mg# sublitho = average sublithospheric Mg#; T_{inter} = temperature of intermediate sublithospheric node (node 13 in Figure 2); T_{380} = temperature at 380 km depth; ρ_i = bulk density of crustal layer *i*; $Vp/Vs_i = Vp/Vs$ ratio of crustal layer *i*; Moho = Moho depth; misfit = total misfit (arbitrary relative scale). Additional examples can be found in the supporting information.

thickness in the interior of the CP does not exceed ~120–130 km, except in its northern boundary. Although we have not included MT data in this work, we note that recent conductivity models agree well with our results (e.g., compare Figure 2 in *Meqbel et al.* [2014] with Figure 8d in this study).

Large horizontal temperature gradients in the mantle (mostly due to steps in lithospheric structure) are found throughout the region, especially along the transitions from (i) the Wyoming Craton into the NBR (Figure 8d) and (ii) the GP into the RM and CP-BR-TZ (Figures 8c and 8e). Such step changes in lithospheric structure constitute areas prone to generating sublithospheric small-scale and/or edge-driven convective instabilities [e.g., *van Wijk et al.*, 2010; *Ballmer et al.*, 2011; *Afonso et al.*, 2008b; *Zlotnik et al.*, 2008; *Davies and Rawlinson*, 2014]. Our results indicate that this is actually the case, as the retrieved sublithospheric flow and associated isotherms exhibit a clear nonconductive nature reminiscent of small-scale convection (SSC) patterns (Figure 8).

It is important to emphasize here that the thermal (and compositional) structure in our model are driven entirely by the constraining geophysical data sets. Thus, although the 3-D momentum equations are solved to obtain the instantaneous mantle flow and dynamic pressures arising from specific thermochemical structures required by data fits, SSC patterns emerge naturally from the inversion of data *only*, and not as a result of flow computations. It is notable then that the location and dominant wavelengths (200–300 km) of the imaged convection cells (particularly beneath the GP region, Figure 8) agree well with those expected from numerical simulations of upper mantle convection [e.g., *Solomatov and Moresi*, 2000; *van Wijk et al.*, 2010].

The sublithospheric flow structure imaged here has implications for LAB depth estimates based purely on seismic data, as typical indicators such as high-velocity anomalies or maximum negative gradients in Vs [*Eaton et al.*, 2009] may be affected by local convective features (e.g., SSC or downwellings). In other words,

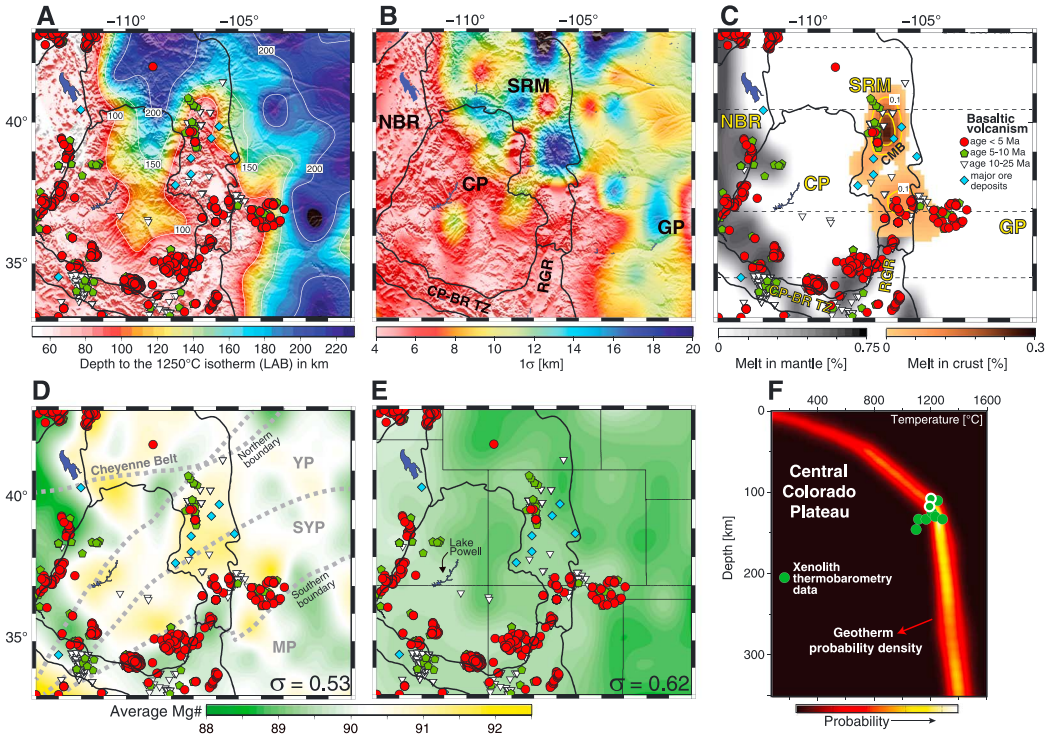

Figure 7. (a) Depth to LAB. (b) Uncertainty in LAB depth as 1σ. (c) Locations where predicted temperatures are higher than the mantle and crustal solidi. Dashed lines indicate the location of transects shown in Figure 8. Yellow contours indicate percent of melt in the crust. (d) Average lithospheric Mg#. Dotted lines represent proposed locations of major sutures, including the Proterozoic suture between the Yavapai (YP) and Mazatzal (MP) crustal provinces [*Whitmeyer and Karlstrom*, 2007]. (e) Average sublithospheric Mg#. In Figures 7a–7e all color maps are averages of the entire posterior PDF. Basaltic volcanism is shown by red dots (age <5 Ma), green pentagons (age 5–10 Ma), and white triangles (age 10–25 Ma) from the NAVDAT database (http://www.navdat.org). Black lines indicate physiographic provinces (defined in the main text). Cyan diamonds show the locations of major ore deposits along the Colorado Mineral Belt (CMB). (f) Posterior probability distribution for the present day geotherm beneath central CP obtained in this study along with P-T estimates from xenolith data [*Li et al.*, 2008].

anomalously cold sublithospheric mantle, such as that associated with SSC drips, may be mistakenly interpreted as belonging to the lithosphere when seismic velocities alone are used as a proxy. This seems to be the case, e.g., in the western CP (Lake Powell area), where previous estimates based on seismic data reported LAB depths similar to those in the thick CP interior [*Levander et al.*, 2011], whereas our multiobservable tomography yields estimates <100 km, but with an anomalously cold mantle immediately beneath the LAB (Figure 8). This discrepancy is not related to the use of different seismic data sets, as we use similar data and also recover localized high-velocity anomalies in this region (Figure 9). The reason is that our inversion favors a sublithospheric dynamic feature (i.e., downwelling with $T \gtrsim 1250°C$) instead of a thick conductive LAB in order to be able to simultaneously fit all constraining data sets. However, this is a complex region where lithospheric and sublithospheric mantle are clearly interacting and in a dynamic state (Figure 8) and even the combined sensitivity of all data sets cannot entirely rule out LAB depths of ~100 km in this region.

The imaged lithospheric structure correlates well with both the location of recent volcanism and the pattern of shear-wave seismic velocities at depths of ~100 km (Figures 7a and 9a) (see also, e.g., *Levander et al.* [2011], *Liu et al.* [2011], and *Shen et al.* [2013]). There is a less obvious correlation between thermal lithospheric thickness (Figure 7a) and mapped Precambrian crustal provinces based on surface features (Figure 7d). This likely indicates that most of the original deep lithospheric structure created during these ancient tectonic episodes has been affected/modified by later deep events. Similarly, although there seems to be a slightly better correlation between thermal structure and Cenozoic crustal provinces (Figure 7a), our results suggest that the original deep structure associated with the creation of these provinces has been also modified at later stages. Indeed, it still is being modified today (Figure 8). An interesting observation is that the lithospheric thinning affecting the southernmost part of the RM seems to stop at, or near, the location of the Proterozoic Cheyenne

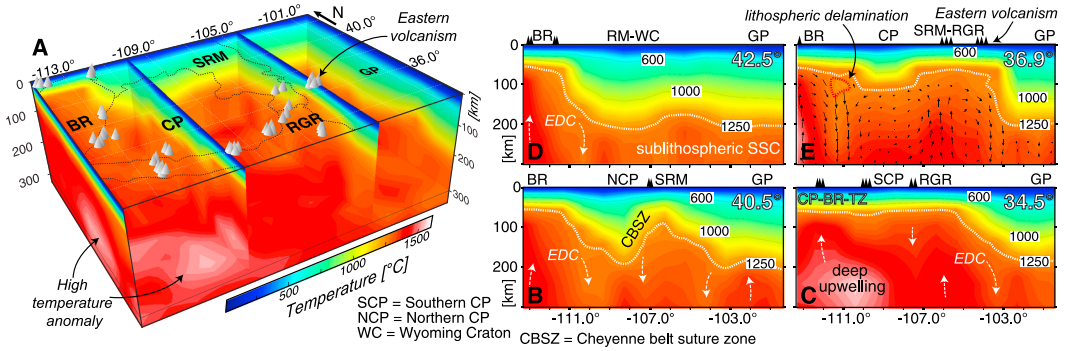

Figure 8. (a) The 3-D rendering of the thermal structure beneath Western-Central U.S. White cones indicate major volcanic localities. A large high temperature anomaly is visible at depths 200–350 km beneath the western and southern edges of the CP and CP-BR-TZ. Note that the relatively isolated eastern volcanism is located above a step in the lithospheric structure with relatively high sublithospheric temperatures. (b–e) Temperature structure along four transects shown in Figure 7b. Clear indications of both edge-driven (EDC) and sublithospheric small-scale (SSC) convection are visible in these transects. Overall local circulation is indicated by white arrows; Figure 8c includes a 2-D section of the full instantaneous 3-D sublithospheric flow (as transparency) after removing the long-wavelength 3-D circulation trend to highlight the local flow pattern; a downwelling is imaged beneath the western edge of the CP. All color scales are averages from the entire posterior PDF.

belt suture zone (Figures 7a and 7d). This region is also associated with a discontinuity in the lithospheric bulk composition, as observed in Figure 7d. Regardless of its actual origin (Proterozoic suture, accretion of ocean plateau, etc.), it seems that this zone plays a major role in the tectonothermal evolution of the region. We discuss this further below.

The correlation between recent volcanism and LAB structure is not an unexpected feature and emphasizes the role of both high temperature anomalies in the sublithospheric mantle and a shallow LAB in generating mafic melts and facilitating its migration to shallow levels. Indeed, despite our simple implementation of mantle melting (section 3.3), regions where we obtain temperatures above the mantle solidi coincide remarkably well with the location of recent basaltic volcanism (Figure 7c) and regions of highly fertile lithospheric mantle (low Mg# in Figure 7d). We do not regard this as coincidental, as it is well known that melt refertilization of the lithospheric mantle occurs in regions that have experienced pervasive basaltic volcanism and/or significant tectonism [e.g., *Porreca et al.*, 2006; *Griffin et al.*, 2009]. Melt extraction from the sublithospheric mantle would deplete this region (visible in Figure 7e) and transfer the least refractory components to lithospheric depths. The fact that our inversion intrinsically identifies these effects in areas where there is independent and abundant evidence (xenoliths) for mantle refertilization by melt percolation [e.g., *McGuire and Mukasa*, 1997; *Smith*, 2000; *Lee et al.*, 2001; *Porreca et al.*, 2006] is not only reassuring but also evidence of the validity and effectiveness of thermochemical tomography (see also results in *Shan et al.* [2014] and *Guo et al.* [2016]). Similarly, large areas of the lithospheric mantle of the CP are characterized by more depleted compositions, in agreement with independent observations from mantle xenoliths (Figures 7d and 7f, see section 6.2). However, the actual distribution of depleted lithosphere within the CP and surroundings is less regular than usually assumed and does not conform to the Cenozoic boundaries of the CP (consistent with geochemical signatures of basaltic lavas in the region [see, e.g., *Crow et al.*, 2011]). Similarly, although the average Mg# beneath the GP is higher than in the west, it also exhibits significant variability (albeit of smaller amplitude). For instance, the relatively fertile (note that the actual amplitude is small) SE-NW domain in the lower right corner (Texas) of Figure 7d is a robust feature of our model and supplementary tests (not shown) indicate that it extends well into Texas. Although it is highly speculative to assign an origin to this anomaly based on surface information only, we note that this region coincides with the terminal parts the pre-Cambrian southern Oklahoma aulacogen. In principle, the lithospheric thinning associated with this event could have affected the lithospheric mantle in this region and produced small amounts of partial melt that refertilized the lower lithosphere. Alternatively, subsequent thermal relaxation could have incorporated asthenospheric material into the thermal lithosphere. Both cases could explain the observed anomaly.

We end this section by noting that while the compositional field obtained from the full thermochemical tomography (Figures 7d and 7e) is similar to that obtained by jointly inverting Rayleigh wave dispersion curves and *P* wave teleseismic travel time residuals (Figure 4n), the actual amplitudes and associated standard

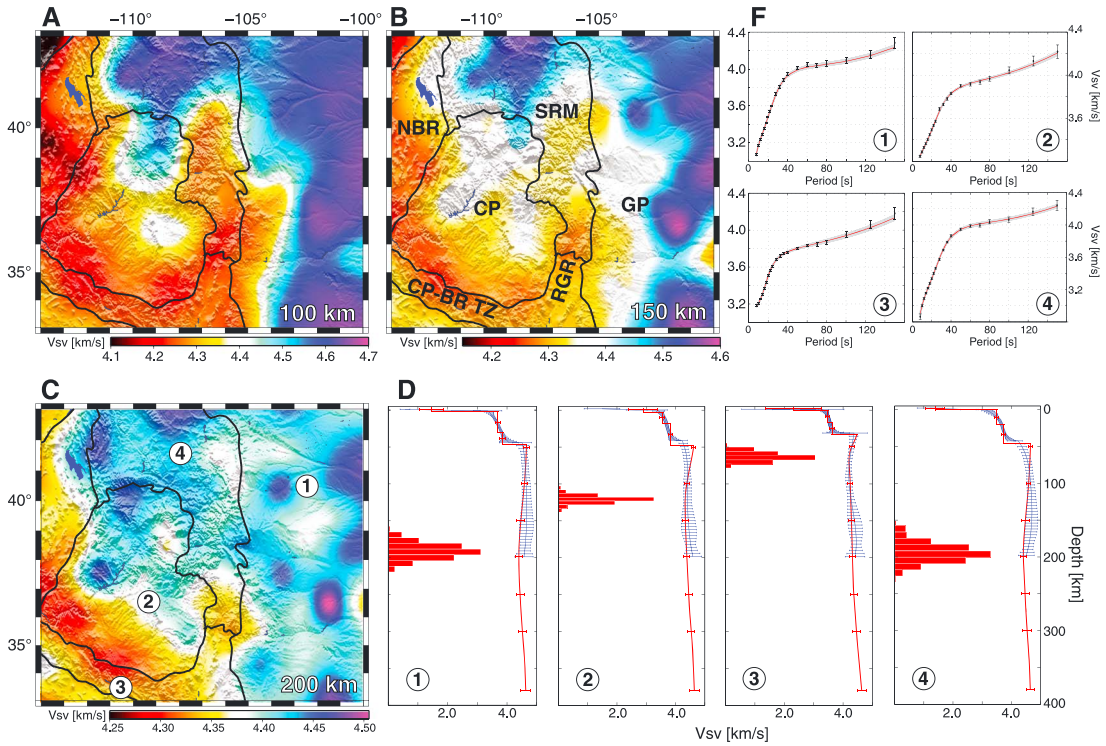

Figure 9. (a–c) Vsv structure as the median of the posterior PDF at three depths: 100, 150, and 200 km. (d) Examples of 1-D Vsv structure. The locations of these 1-D profiles are indicated in Figure 9c; red histograms denote posterior PDFs for LAB depth at each location. Blue lines are the results of *Shen et al.* [2013] with error bars denoting 2σ, as reported by these authors. Red lines are results from the present study with error bars denoting 2σ. Note that while *Shen et al.* [2013] used continuous splines for the crust, we use a three-layer parameterization with constant properties in each layer. (f) Dispersion curve fits for the four 1-D profiles shown in Figure 9d. Red lines = mean of all posterior models (in gray).

deviations are different. As explained in section 4.1, this is due to the addition of data sets sensitive to bulk density, which results in a much higher sensitivity to compositional anomalies.

5.2. Associated Seismic Structure

As previously stated, traditional tomographic images (seismic structure) are a by-product of our inversion scheme. Figure 9 shows depth slices of the velocity structure associated with the thermochemical structure presented above. The maps show the median of the posterior PDF. Our results compare very well with those recently presented by a number of different groups [e.g., *Yang et al.*, 2008; *Pollitz and Snoke*, 2010; *Moschetti et al.*, 2010a, 2010b; *Liu et al.*, 2011; *Schmandt and Humphreys*, 2010; *Levander et al.*, 2011; *Shen et al.*, 2013, *Hansen et al.*, 2013, among many others]; all major features discussed in these works are visible in our probabilistic tomography as well. In particular, our results are very similar to those of *Shen et al.* [2013]. This is somewhat expected, as the Rayleigh dispersion curves for periods <40 s used in this study are the same as those used by these authors. There are, however, a few significant local differences in the vertical patterns of Vsv (Figure 9d). For instance, *Shen et al.* [2013] reports significantly higher velocities in the range 100–200 km depth beneath the south tip of the Wyoming Craton (~ −108°, 42°) but significantly lower velocities between the Moho and 80 km depth (Figure 9d, column 4). While our vertical profile exhibits a more linear trend of decreasing velocity down to the LAB, theirs has a marked convex shape where velocities first increase with depth with a steep gradient and then decrease down to LAB depths. While the fit to the dispersion data is excellent in both studies (Figure 9f), supplementary tests indicate that it is extremely hard to obtain a convex velocity profile of that magnitude if the geotherm is to be approximately conductive (see also *Shapiro and Ritzwoller* [2004] and *Jones et al.* [2015]) and all other data sets are to be fitted simultaneously, particularly the vertical component of gravity gradients. The local discrepancies between the two models are therefore simply explained by the use of different physical parameterizations and constraining data sets. This also explains the smaller standard deviations associated with our profiles (i.e., there is less room for variability if all observations are to be explained simultaneously).

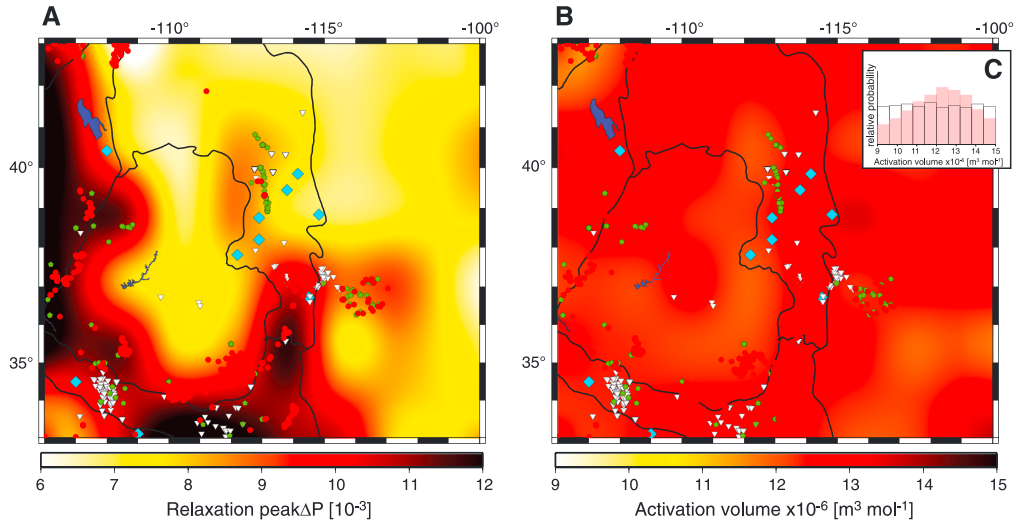

Figure 10. (a) Values for the peak of elastically accommodated grain-boundary sliding retrieved by our multiobservable inversion. The color scale indicates the actual range allowed during the inversion. Note the close correlation of high values of ΔP (i.e., high attenuation) with regions of hot mantle (compare with Figures 4a and 5 in the main text), which suggests the presence of additional factors that increase the attenuation of seismic waves in these regions but that are not included in the current laboratory-based attenuation model. (b) Recovered values for the activation volume (as the mean of the posterior PDF) associated with the extended Burgers model. This parameter shows a more homogeneous spatial distribution (average value $\sim 12.5 \times 10^{-6}$ m^3 mol^{-1}) but its posterior PDF is relatively broad. (c) Prior (white) and posterior (pink) distributions for activation volume.

A comparison of Figures 7a and 9a and 9b shows the strong correlation between Vs structure and subsurface temperature. This is yet another example of the well-known sensitivity of shear waves to temperature structure. Consistent with the results in section 4.1, here we find that the incorporation of independent data related to the Vp structure (i.e., travel times) adds important additional constraints, especially for isolating compositional effects, but it is not truly critical for obtaining first-order estimates of the LAB and shallow thermal structure. In other words, our results indicate that robust LAB estimates can be obtained by inverting surface wave data only and that the resulting temperature structure is generally a good fit to Vp data as well. The opposite, however, is not always true.

5.2.1. Attenuation Parameters

The results for the inverted attenuation parameters are depicted in Figure 10. We note that, in general, the extended Burgers formalism of *Jackson and Faul* [2010] can well explain most of the variability observed in the data, even when a fixed ΔP close to that reported in *Jackson and Faul* [2010] is used; a point already noted by other researchers [*Dalton and Faul*, 2010]. However, when ΔP is included as a free parameter in the multiobservable inversion, results show that hot mantle regions require higher attenuation (i.e., larger values of relaxation strength) than colder regions for optimum data fit (Figure 10a). We interpret this as clear evidence that natural phenomena (e.g., near-solidus bulk and grain boundary effects, inadequate correction for melt, etc.) not yet included in current laboratory-based attenuation models (or in our parameterization) contribute to the bulk attenuation in hot mantle regions. We emphasize, however, that the final PDFs obtained by including ΔP as a free parameter and those obtained by using an average value of ~0.085 for ΔP are similar. The main difference is that the mean of the PDF for LAB depths (defined by the temperature structure) is slightly displaced toward shallower depths. However, this difference is always <5%.

The recovered values (as means of the posterior PDF) for the bulk activation volume when allowed to vary within the large experimental uncertainties are more homogeneous than for ΔP (Figure 10b), and therefore less affected by the local variability of the mantle's thermochemical structure. However, the posterior PDF is rather broad, suggesting that the actual value of the bulk activation volume is not well constrained. This is consistent with the results of numerous supplementary tests that demonstrate that using different (constant) values for the activation volume in the inversion has no significant effect on the final thermochemical structure of the lithosphere. It does, however, affect the structure of the mantle at depths $\gtrsim$300 km, where the combined sensitivity of our data sets is at its lowest.

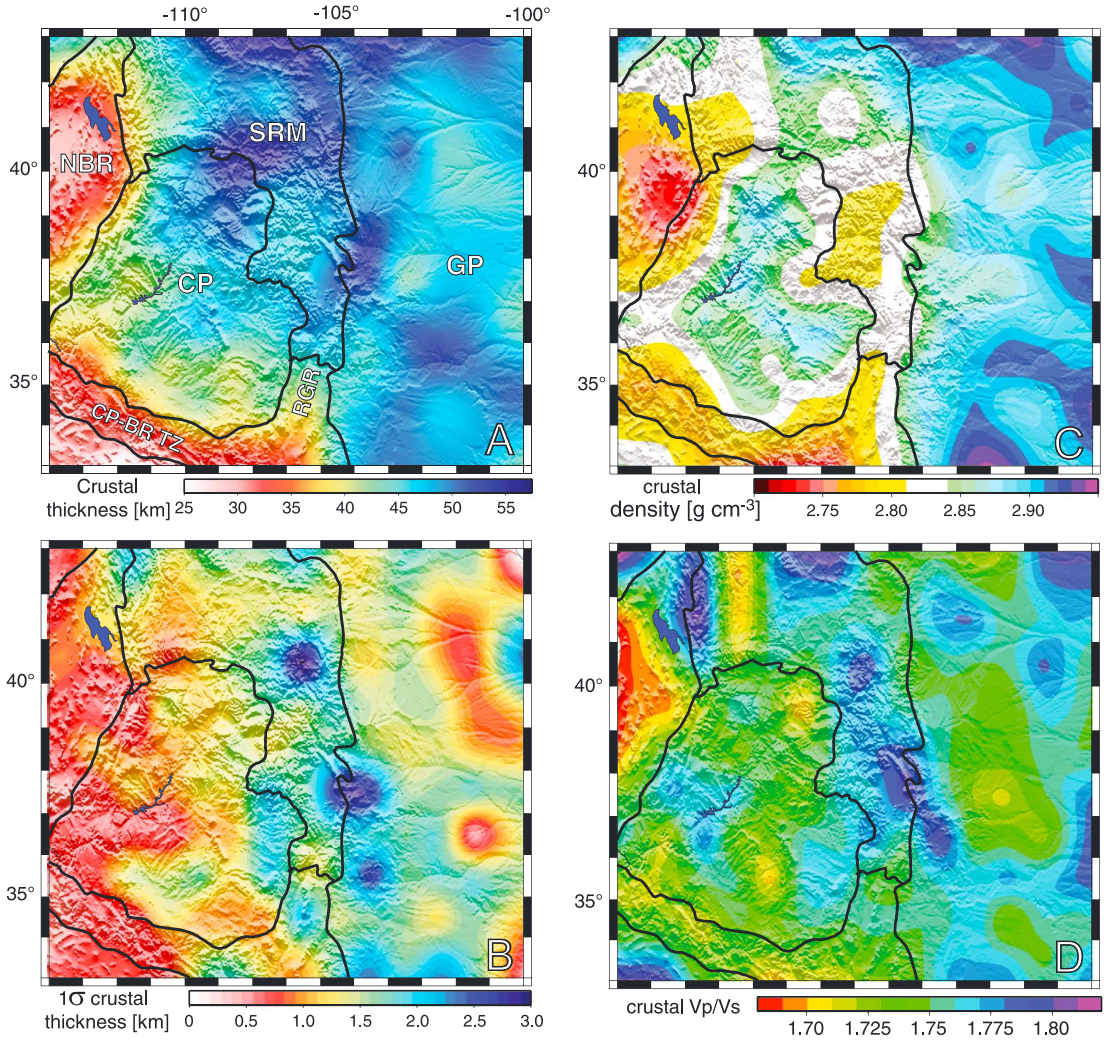

Figure 11. (a) Crustal thickness obtained in this study. (b) Uncertainty in estimated crustal thickness as 1σ from the full posterior PDF. (c) Average crustal density. (d) Average crustal Vp/Vs ratio. All maps except Figure 11b plot the median of the full posterior PDF.

5.3. Crustal Structure

While the main focus of this work is the deeper lithospheric structure, here we summarize the main features of the crustal structure obtained from our inversion (section 3.1). Figure 11 shows the recovered crustal thickness, average density and average Vp/Vs ratio (as the median of the actual PDF). Within uncertainties, our results are comparable to those obtained by *Gilbert* [2012], *Lowry and Perez-Gussinye* [2011], and *Shen et al.* [2013]. While such agreement provides confidence in our knowledge of the crust in this region, there are a few noteworthy differences between our results and those from the aforementioned works. First, while *Gilbert* [2012], *Lowry and Perez-Gussinye* [2011], and *Shen et al.* [2013] report a thick and relatively homogenous crust beneath and around the Colorado Mineral Belt region, our model predicts a shallower Moho, in better agreement with recent detailed seismic studies [e.g., *Rumpfhuber and Keller*, 2009; *Hansen et al.*, 2013]. We also obtain a relatively low velocity/low density lower crust beneath this region (Figure 11c); an observation that has been pointed out by other authors as well based on different methods/data [e.g., *Hansen et al.*, 2013; *Levandowski et al.*, 2014; *Jones et al.*, 2015].

Second, we do not obtain anomalously thick crust in the SW region of the CP, as reported in *Gilbert* [2012] and *Lowry and Perez-Gussinye* [2011]. The Moho in these two works is mainly constrained by receiver function analyses, which are known be complex in this region [e.g., *Levander et al.*, 2011]. Since we have not included receiver functions in the present study, we cannot draw a direct comparison at this point. However, we note

that the results obtained by *Shen et al.* [2013], who jointly inverted Rayleigh dispersion curves and receiver functions, do not show a thick crust in this area.

Third, and related to the last point, our average Vp/Vs values for the crust (Figure 11d) beneath the RM are more similar to those of *Hansen et al.* [2013] and *Gilbert and Sheehan* [2004] than to the low values reported in *Lowry and Perez-Gussinye* [2011]. In our models, the reason for obtaining relatively high average Vp/Vs values for the crust beneath this region is related mainly to the high predicted temperatures (i.e., shallow LAB) in the mid and lower crust, for which there is abundant independent evidence [e.g., *Decker*, 1995; *Hansen et al.*, 2013]. High crustal temperatures are expected to result in high Vp/Vs values, as (i) Vs is more strongly affected by temperature than Vp, (ii) partial melting and/or release of fluids is promoted and these increase the Vp/Vs ratio. It is also problematic to justify large amounts of felsic (quartz-rich) rocks at mid-lower crustal levels when temperatures are so high, as partial melting of the felsic components will occur (Figure 7). Moreover, we have conducted preliminary tests (Appendix B) including *P* wave receiver functions into the inversion and obtained similar Vp/Vs values as those associated with the model in Figure 11.

While our crustal model is consistent with all the inverted data sets, it is clear from the above discussion that more detailed investigations, including receiver function analyses and thermodynamic modeling, are needed to elucidate the thermochemical nature of the crust in this region and the root causes for current discrepancies among different studies. We are currently working on this [e.g., *Afonso et al.*, 2015], and the results will be presented in a forthcoming publication.

6. Discussion

6.1. Dynamic Versus Static Contributions to Elevation

The present-day elevation and uplift history of the CP and surrounding regions are the subject of long-standing debate [cf. *Humphreys and Coblentz*, 2007; *Karlstrom et al.*, 2012]. A plethora of hypotheses have been put forward, covering the entire spectrum of possibilities, ranging from purely intralithospheric causes to purely deep dynamic processes [e.g., *Spencer*, 1996; *McQuarrie and Chase*, 2000; *Humphreys et al.*, 2003; *Roy et al.*, 2009; *Liu and Gurnis*, 2010; *Karlstrom et al.*, 2012; *Becker et al.*, 2014; *Molnar et al.*, 2015]. One of the main reasons for this controversy lies in the simple fact that available data cannot be used to distinguish unequivocally between different hypotheses. Several alternatives can be used to explain the same observation. For instance, the late Cenozoic uplift of the edges of the plateau can be explained equally well by the effects of small-scale circulation along the edges of the plateau [e.g., *van Wijk et al.*, 2010] or by allowing simple conductive heating of a variable-thickness lithospheric plate [*Roy et al.*, 2009]. The latter mechanism is well known; if a step in the thermal structure of the lithosphere is produced, subsequent thermal relaxation toward a quasi steady state structure will create local uplift close to the original step. If lateral compositional anomalies are also included, this effect can be magnified [*Afonso et al.*, 2008a; *Gradmann et al.*, 2013]. The magnitude of the uplift produced by such "static" mechanisms is of the same order as that generated by edge-driven convection [*van Wijk et al.*, 2010]. Likewise, the high elevation of the interior of the plateau can be explained by compositional anomalies within the lithosphere, by crustal effects, or by deep dynamic support from sublithospheric mantle flow (see references above).

The approach used in this work separates static (lithospheric isostasy) from dynamic contributions (sublithospheric dynamic loads) as described in section 3.2.3. Although it is neither the only nor the most sophisticated approach, it does provide a practical link between the two sources of absolute topography, allowing us to quantify their individual contributions. A critical advantage of the method presented in this paper is that it does not only include all major potential contributions to present-day elevation in a thermodynamically and internally consistent manner, but it naturally finds the most probable combination of them (in a best fitting sense) as required by all the constraining data. In other words, it is the actual data (jointly inverted) that balance the contributions with no stringent a priori requirements other than well-accepted physical theories (i.e., the forward problems solved during the inversion).

The resulting dynamic contribution (in the sense of *Molnar et al.* [2015]) to present-day topography obtained from our inversion is shown in Figure 12. While our results are comparable to those obtained by *Becker et al.* [2014], the actual magnitudes are generally smaller in our model. We obtain moderate and localized dynamic contributions to present-day topography, amounting to ~15–35% of the absolute elevation in some regions, mainly around the edges of the CP (Figure 12). Therefore, most of the observed elevation in this region seems to be supported by intralithospheric sources. In this context, it must be acknowledged that according to the

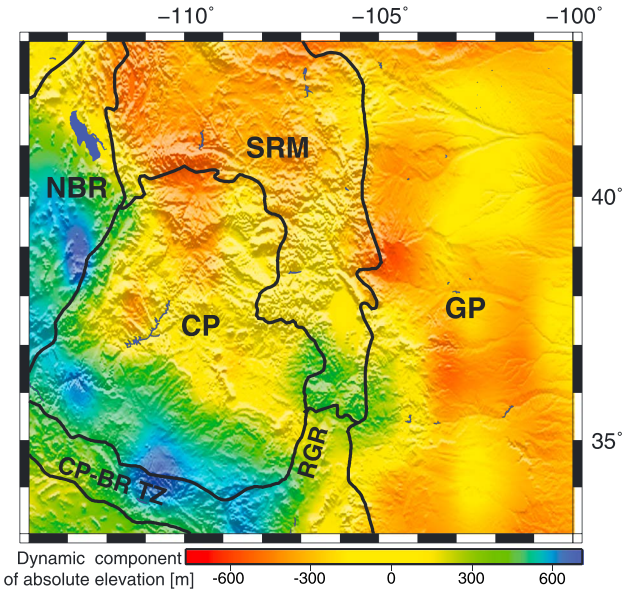

Figure 12. Dynamic component of absolute elevation (exclusively from sublithospheric upper mantle flow). The latter is a response to the present-day sublithospheric density variations required to jointly fit all data sets used in this study.

lithospheric isostasy principle, elevation changes of the order of ~500–600 m can readily be obtained by only minor modifications (well within their individual uncertainties) of Moho depth, crustal density, average lithospheric mantle composition, and/or LAB depth. This illustrates the fundamental conundrum in separating static from dynamic contributions to topography and explains why dynamic contributions can be easily masked in studies relying on either crustal or lithospheric isostasy arguments [e.g., *Spencer*, 1996; *McQuarrie and Chase*, 2000]. We are currently working on a more efficient and accurate coupling of detailed lithospheric imaging (in the spirit of the present study) and mantle flow simulations that we hope would provide a more complete understanding of the individual roles of lithospheric versus sublithospheric contributions to the evolution of topography in continents.

A notable caveat in our current implementation is the exclusion of long-wavelength deep structure (e.g., thermal anomalies at >400 km depth) which, in principle, could affect the elevation of the modeled area through a long-wavelength dynamic contribution. Whether or not the explicit inclusion of such structures can significantly affect the results presented in this work remains to be tested. However, at least two arguments can be advanced that suggest that our results should not be significantly affected by lower mantle structures. First, it is well known that lithospheric structure correlates well with surface tectonics and elevation worldwide; the best examples being passive margins, orogens, and oceanic lithosphere. Studies that estimate large values ($\gtrsim$1000 m) of "dynamic topography" from deep sources based on global tomography studies commonly ignore the contribution/modulation due to the internal thermochemical/rheological structure of the lithosphere. Therefore, such intralithospheric effects get lumped together with the true dynamic components arising exclusively from mantle flow, and thus, the estimated dynamic contribution would typically be unrealistically high (see also discussion in *Steinberger* [2016]). Second, long-wavelength deep anomalies (>400 km depth), as imaged by global tomography, will contribute to surface elevation at wavelengths comparable to, and likely larger than, that of the deep anomaly due to the actual large-scale peripheral flow produced by such anomalies (see examples in *Crameri et al.* [2012]). Also, these anomalies are typically subject to large uncertainties in their dimensions due to limitations imposed by the data and tomography methods. It is therefore hard to envision that the surface uplift caused by a lower mantle anomaly will differentially affect individual provinces delimited by surface tectonics (e.g., the CP). More likely is the possibility of a lower mantle anomaly providing a long-wavelength "background contribution" to elevation and the lithospheric/upper mantle structure controlling the shorter-wavelength features (recall that our geoid is filtered to remove long-wavelength contributions). Even in this case, the actual 3-D patterns and thermochemical structure given by our model will not change; only small changes to the absolute magnitudes would be produced.

6.2. Xenolith Data

There is abundant information in the literature about mantle xenoliths in and around the CP region. We have collected thermobarometric and chemical information from well-studied mantle xenoliths (a total of 160 samples) for the central CP, Grand Canyon area, southeast CP/RGR area, and BR/CP transition zone from a number of references [*Laughlin et al.*, 1971; *Kudo et al.*, 1972; *Smith and Levy*, 1976; *Ehrenberg and Griffin*, 1979; *Smith*, 1979; *Aoki*, 1981; *Ehrenberg*, 1982; *Wilshire et al.*, 1988; *Alibert*, 1994; *Smith and Riter*, 1997; *Riter and Smith*, 1996; *Smith et al.*, 1999; *Smith*, 2000; *Li et al.*, 2008]. In particular, the two green circles with thick white borders in Figure 7f represent two recent estimates from well-preserved garnet-bearing Navajo xenoliths reported in

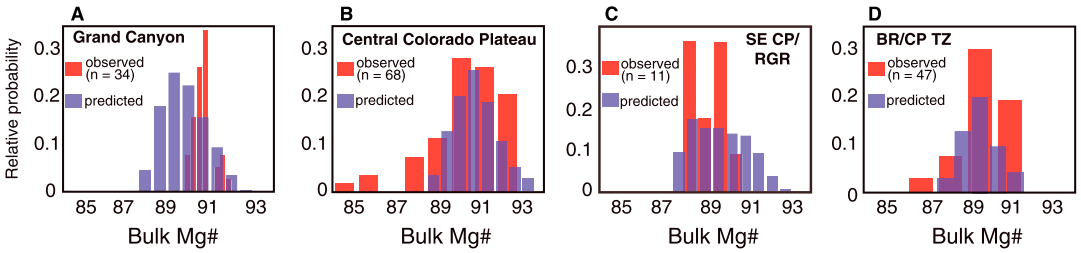

Figure 13. Comparison between observed and predicted lithospheric bulk Mg# in (a–d) four different regions. Observed distributions (red histograms) are obtained from local xenolith information (references in main text). Predicted distributions (blue) are the final posterior PDF from our inversion. The number of xenolith samples used to obtain the observed distributions are also indicated in the panels.

Li et al. [2008], from which temperatures and pressures of equilibration were simultaneously obtained with two-pyroxene thermometers and barometers. Although this compilation is far from exhaustive when compared to the bulk of available studies, it is large enough to provide some representative comparisons with our predictions. The population of xenoliths from the Rio Puerco volcanic field in New Mexico is an exception, as it is composed of only 11 samples [*Porreca et al.*, 2006], and therefore, it should not be considered statistically robust or representative of the entire lithospheric mantle in this area.

Figure 13 shows a comparison between predicted (from our inversion) and empirical (from xenolith data) probability distributions for lithospheric bulk Mg#. Considering the local nature of mantle samples, their different ages, and the uncertainties they carry in terms of representativeness of large portions of the lithospheric mantle, the agreement between predicted and empirical distributions is remarkable. We note, however, that comparing xenolith data with results from geophysical imaging techniques is not straightforward and caution should always be exercised when using these two sources of information together. In principle, xenoliths are a localized (horizontally and vertically), potentially unrepresentative, sample of the mantle at the time of extrusion (typically several tens of Ma before present). By contrast, geophysical imaging is a larger-scale snapshot of the present-day structure. Geodynamic events such as lithospheric delamination, magma percolation, and small-scale convection, which can significantly alter the thermochemical structure of the lithosphere, are known to be relatively fast, taking place over timescales of a few Ma [e.g., *van Wijk et al.*, 2010; *Burov and Gerya*, 2014]. Therefore, there is no strong a priori reason why xenoliths included in volcanic rocks that erupted tens of Ma ago should be taken as accurate representations of the lithospheric mantle today. However, if no significant modification of the lithosphere occurred after xenoliths were exhumed, we would expect to see a correlation between the bulk composition derived from mantle xenoliths and that predicted by thermochemical tomography. This seems to be the case in the central CP. As mentioned previously, our results suggest that the central and northern CP are the least modified regions, preserving a thicker lithospheric mantle than surrounding domains. This is also the region with the "oldest" xenoliths ($\gtrsim$25 Ma) in our database, which record a geotherm similar to that predicted by our inversion (Figure 7e).

A curious case mentioned above is the relatively fertile domain close to the Cheyenne Belt (Figure 7c). This is a robust feature of the model in the sense that it is always present in all valid models given by the inversion. This anomaly, indicating a compositionally denser and faster portion of lithospheric mantle, also coincides with the change from shallow to thick LAB in the NE end of the CP. Although mantle xenoliths have been recovered from volcanic rocks in and around this area [e.g., *Eggler et al.*, 1988; *Carlson et al.*, 2004; *Schulze et al.*, 2008; *Ashchepkov et al.*, 2013], they are rare, of Paleozoic age, highly heterogenous, and therefore provide information on only a limited portion of the deep mantle at the time of extrusion. We note, however, that fertile lherzolites and Fe-rich eclogites are common lithologies in these xenolith suites [e.g., *Schulze et al.*, 2008; *Ashchepkov et al.*, 2013], and therefore potential candidates for explaining the relatively dense and fast nature of this mantle region. However, the questions of how and when this portion of mantle acquired its present-day fertile characteristics, as well as the role of Laramide and post-Laramide processes, remain open to debate.

6.3. Hydration Effects

As a first-order approximation at the scale of interest, we assumed dry compositions for the computation of equilibrium mantle mineral assemblages via free energy minimization. This is justified based on the following two arguments. First, the forward problems used in this study are not affected by the small amounts of water commonly observed in mantle samples (<300 ppm wt) [*Yu et al.*, 2011, and references therein], as the actual

modes and phase compositions are insensitive to such water contents. Second, water-rich domains in the mantle are typically restricted to local veins and localized reaction zones [cf. *Coltorti and Gregorie*, 2008; *O'Reilly and Griffin*, 2013], which cannot be captured by our current parameterization.

Small amounts of water, however, can affect bulk viscosity and, highly likely, attenuation of seismic waves [e.g., *Karato*, 1986, 2008, 2012, *Hirth and Kohlstedt*, 1996; *Kohlstedt*, 2006; *Jones et al.*, 2012]. However, in the case of viscosity, we do not strictly invert for it in this work but assume it to be a function of temperature and pressure only via appropriate calibration of activation energies and volumes (Appendix A). The effect of water can therefore be thought of as being implicitly absorbed by the values of activation energies used in the inversion. Several test runs using higher or lower activation energy values indicated that acceptable joint fits to all the observables were not possible using activation energies that would result in average asthenospheric viscosities $\gtrsim 2 \times 10^{20}$ Pa s (but lower values are acceptable). This range is well within values used for slightly hydrated mantle (e.g., *van Wijk et al.* [2010], *Ballmer et al.* [2011], *Afonso et al.* [2008b], and *Kaislaniemi et al.* [2014], among others). In the case of attenuation, the main effect of small amounts of water seems to be an increase in the effect of the dissipation peak due to grain boundary sliding [*Karato*, 2012]; this hypothesis is consistent with a limited number of experimental results [e.g., *Aizawa et al.*, 2008]. Since we treat the dissipation peak of elastically accommodated grain boundary sliding (ΔP) as an unknown (section 5.2), we are actually implicitly considering additional factors affecting attenuation (e.g., water content) that are not explicitly modeled/controlled by inversion parameters.

A completely different situation would arise if average water content reached levels >1 wt % ($>1 \times 10^4$ ppm wt) in extended portions of the lithospheric mantle. This would not only modify the modes and compositions of the typical mantle phases but introduce variable amounts of the so-called hydrous phases such as amphiboles, micas, serpentinite, etc. (i.e., modal metasomatism). These hydrous phases tend to have significantly different physical properties (e.g., less dense, lower velocities) from common anhydrous phases; if their combined volumetric proportion becomes >15% (note that this requires water contents >1 wt %) in large portions of the lithospheric mantle, they can affect the computation of predicted observables to a significant extent and thus modify inversion results. However, this is a highly unlikely situation, as average water contents $\gtrsim$1 wt % in large portions of the lithospheric mantle are extremely difficult to justify based on either current knowledge from xenoliths (including those from the CP, BR, and RM areas [e.g., *Peslier and Luhr*, 2006; *Grant et al.*, 2007; *Li et al.*, 2008; *Yu et al.*, 2011]) or laboratory experiments [e.g., *Skogby*, 2006; *Beran and Libowitzky*, 2006; *Frost*, 2006]. Therefore, our assumption of a dry lithospheric mantle for modeling phase assemblages should be a good first-order approximation.

6.4. Implications for the Recent Tectonic Evolution

Our results do not require major dynamic components from sublithospheric mantle flow to explain either the present-day elevation of the CP or any other geophysical observation included in our inversion. This suggests that lithospheric structure (e.g., crustal thickness and density, temperature, and mantle composition) is the main cause for the high elevation of the plateau relative to its surroundings. It also highlights the importance of shallow lithospheric structure, and its interaction with a dynamic sublithospheric mantle, in controlling surface topography and volcanism.

We therefore suggest that the present-day thermochemical structure of the region is the result of two main interacting geodynamic processes. First, during the transition from Mesozoic flat subduction to a Cenozoic extensional regime, pervasive alteration of the lithospheric mantle and heterogeneous lithospheric thinning created a highly irregular LAB geometry. In this scenario, the more depleted and less extended CP lithosphere remained as a relatively thick lithospheric keel surrounded by reactivated regions of thinned, hot, and metasomatized lithosphere in the west, south, and east. The addition of fluid-rich phases [*Humphreys et al.*, 2003; *Crossey et al.*, 2009] at the base and edges of the CP lithosphere could have generated some intralithospheric buoyancy and decreased the viscosity of these mantle regions. All of this activity promoted the second process, namely, widespread heating of the remaining CP lithosphere (with addition of extra thermal buoyancy [*Roy et al.*, 2009]) and the development of SSC cells at the edges of thick keels and beneath regions of thick lithosphere (e.g., GP). These downwellings, including episodic delaminations, eroded the weakened margins of the CP and brought up asthenospheric material through forced return flows (Figure 8). This scenario is consistent with geochemical signatures of basaltic lavas that show that Neogene volcanism not only migrated toward the center of the CP but also became increasingly more asthenospheric [*Crow et al.*, 2011]. It is also supported by mantle degassing estimates from CO_2 and He analyses in groundwater samples

[*Crossey et al.*, 2009]. Regions affected by forced upwellings produced basaltic melts that percolated through and reacted with the overlying lithosphere, imparting the fertile signature (i.e., refertilized mantle, Figures 7c and 7d). The associated increase in compositional density is, however, counteracted by a gain in thermal buoyancy and a dynamic contribution from the upwellings (Figures 8 and 12).

We directly image the complex dynamics of sublithospheric convection and corroborate the view that the edges of the CP are being eroded by delamination and/or edge-driven downwellings, as previously suggested based on indirect proxies [e.g., *Schmandt and Humphreys*, 2010; *Levander et al.*, 2011; *van Wijk et al.*, 2010; *Crow et al.*, 2011]. These downwellings transfer material from the lower parts of the lithospheric mantle to the sublithospheric mantle, changing its mean temperature and composition (Figures 7d and 7e). The removal of these less depleted portions of the lithosphere leaves behind the original ultradepleted shallow mantle cores beneath the CP crust [*Lee et al.*, 2001; *Afonso et al.*, 2008a; *Griffin et al.*, 2009], thus increasing the average Mg# of the remaining lithospheric mantle. This explains why our results show highly depleted lithospheric mantle in regions where the lower lithosphere of the CP has been thinned or removed by downwellings (Figures 7d and 8). In other words, the shallow depleted portions of the lithospheric mantle that have not been delaminated and/or pervasively refertilized by melt metasomatism seem to have retained their original depleted nature (Figures 7d, 7f, and 13). Conversely, regions that have been significantly affected by melt metasomatism (e.g., CP-BR-TZ, RGR, NBR) exhibit an anomalous fertile nature (Figures 7d and 13).

Although more intriguing, a somewhat similar explanation could be given for the relatively depleted lithospheric mantle beneath the eastern RM (Figure 7d). According to our results, the lithosphere in this region retains only ~40 km of the shallowest mantle, but it has not experienced pervasive basaltic metasomatism (i.e., refertilization) of a magnitude comparable to that in, for example, the CP-BR-TZ. This is consistent with the gap in recent basaltic activity (see also *Crow et al.* [2011]), its slightly lower sub-Moho (higher) temperatures (wave speeds) than the northern and southern adjacent regions, its slightly shallower Moho, and the high average elevation. This region also has an anomalously low-density (hydrous?) lower crust (Figure 11c) [*Jones et al.*, 2015], which may suggest low amounts of mafic underplating and/or intrusions coming from the mantle. It also coincides with the location of massive ore deposits (i.e., Colorado Mineral Belt, Figure 7d) typically found above ancient and depleted mantle that experienced metasomatism by aqueous fluids, not by voluminous mafic melts [e.g., *Lorand et al.*, 2000]. We also note that *Griffin et al.* [2013] and *Hansen et al.* [2013] recently argued in favor of an anomalously buoyant lithosphere (with remnants of Archean mantle) beneath this region using independent geochemical and geophysical data sets. Therefore, it seems possible that while the lower portions of the lithospheric mantle beneath the eastern RM was removed, its old, shallower, and buoyant Archean mantle escaped pervasive melt metasomatism.

7. Conclusions

The results reported here (and in parts I and II) demonstrate that multiobservable thermochemical tomography offers a sound method to characterize the thermochemical structure of the lithosphere and upper mantle and opens new opportunities for deep-Earth imaging. In this method, all physical and chemical parameters defining an Earth model are linked together by fundamental thermodynamic relations, rather than by ad hoc empirical assumptions. This allows us to directly invert for the fundamental variables defining the physical state of the Earth's interior, namely, temperature, pressure, and major-element composition using a multitude of data sets with complementary strengths: body wave teleseismic data, surface wave phase dispersion data, gravity anomalies, long-wavelength gravity gradients, geoid height, absolute elevation, and surface heat flow data. In this probabilistic inversion scheme, traditional tomographic images of physical parameters such as 3-D seismic velocity become a "free" by-product. However, our tomographic images are, by design, also thermodynamically compatible with all the other inverted observables instead of satisfying one type of data set only. This is important, as any model deemed representative of the real physical state of the Earth's interior should pass the test of explaining other geophysical data sets as well.

This study shows that the present-day thermochemical structure of the lithosphere beneath Western-Central U.S. is highly heterogeneous, preserving the imprints of both ancient and recent geodynamic processes across multiple scales. This contrasts with the usual simplistic view of compositionally homogenous lithospheric domains correlating with surface crustal provinces. Indeed, our results demonstrate that the deep present-day lithospheric structure correlates weakly with Proterozoic and Cenozoic provinces defined by surface features. Vigorous small-scale convection cells in the sublithospheric mantle are evident, some of

which seem to develop in response to the highly irregular lithospheric structure (i.e., edge-driven convection). Widespread melt metasomatism (i.e., refertilization) of the lithospheric mantle is also identified in areas where asthenospheric partial melting is predicted, pointing to complex interactions and thermochemical feedbacks between asthenosphere and lithosphere. This is in agreement with the location of recent volcanism and independent xenolith and geochemical evidence. The present-day thermal and compositional structure (i.e., density) obtained from the joint inversion of all data sets indicates that most of the current surface elevation in the Western-Central U.S. is supported by intralithospheric features. Dynamic contribution to topography arising from sublithospheric convection is only modest (~15–35%) and localized.

In summary, the intricate lithospheric thermochemical structure imaged beneath Western-Central U.S. seems to be linked more closely to physicochemical interactions with the sublithospheric mantle via small-scale convection, delamination, melt production, and metasomatism rather than to physiographic or surface structures, which may be "frozen-in" structures not representative of the deep subsurface structure. Such physicochemical interactions (perhaps triggered by larger geodynamic episodes) include episodic transfer of mass and energy from the asthenosphere to the base of the lithosphere and vise versa, which can significantly modify the properties of this complex transitional region (lithosphere-asthenosphere transition zone) in space and time and are likely to be key global and recurrent factors controlling the tectonic evolution of continental interiors and intraplate epeirogenesis.

Appendix A: The 1-D Geoid Approximation

During the 1-D exploratory stage of the inversion, the geoid height is computed considering contributions from both isostatic (i.e., within the lithosphere) and nonisostatic (i.e., sublithospheric) density anomalies. The isostatic lithospheric contribution is computed with the well-known 1-D formulation for shallow, long-wavelength isostatic density distributions [e.g., *Turcotte and Schubert*, 1982]

$$\Delta N = \frac{-2\pi G}{g_0} \int_0^H z\, \Delta\rho(z)\, \mathrm{d}z \tag{A1}$$

where H is the total thickness of the model, G the universal gravitational constant, and g_0 the reference acceleration of gravity on the reference geoid. For all practical purposes, it suffices to take $g_0 = 9.81$ m s^{-2} everywhere. $\Delta\rho$ refers to the density difference between a problem column and a reference density column [cf. *Turcotte and Schubert*, 1982; *Afonso et al.*, 2008a]. The contribution form sublithospheric density distributions that are not isostatically compensated is estimated by solving the gravitational potential equation for a prism [e.g., *Nagy et al.*, 2000; *Fullea et al.*, 2009]

$$\Delta N = \frac{\rho G}{g_0} \left|\left| xy\ln(z+r) + yz\ln(x+r) + zx\ln(y+r) - P\Big|_{x_1}^{x_2}\Big|_{y_1}^{y_2}\Big|_{z_1}^{z_2} \right.\right. \tag{A2}$$

where

$$r = \sqrt{x^2 + y^2 + z^2} \tag{A3}$$

$$P = \frac{z^2}{2}\arctan\left(\frac{xy}{zr}\right) + \frac{x^2}{2}\arctan\left(\frac{yz}{xr}\right) + \frac{y^2}{2}\arctan\left(\frac{xz}{yr}\right) \tag{A4}$$

The Cartesian coordinates ($x_1 < x_2$, $y_1 < y_2$, $z_1 < z_2$) define the dimension of the prism and are referred to a coordinate system with its origin coinciding with the point where the geoid height is computed. An alternative to equation (A2) is the simpler gravitational potential equation for a cylinder with equal height h and volume V as the prism (with radius $R = \sqrt{V/\pi h}$)

$$\Delta N = \frac{2\pi G}{g_0} \int_0^H \Delta\rho(z)\, [(R^2 + z^2)^{1/2} - z]\, \mathrm{d}z \tag{A5}$$

These two expressions differ by less than 3% for typical prism dimensions used in our inversions. Both equations (A1) and (A2) are numerically integrated (trapezoidal rule) using the same fine finite-difference mesh as for the temperature calculation (i.e., computation nodes).

When solving equation (A2) (or its equivalent equation (A5)) in each 1-D column making up the 3-D volume, one has some liberty in choosing the x and y dimensions of the prism (or equivalently R in equation (A5)). The most intuitive choice would be to use the dimensions of the 1-D columns used in the discretization of the 3-D volume (section 3.1, Figure 2). This, however, would significantly underestimate the contribution from sublithopsheric anomalies of larger lateral extension, as the geoid signal does not decay with lateral distance as fast as with gravity anomalies (see also *Afonso et al.* [2013b]). This underestimation would be compensated inside the current column by assigning it an incorrect sublithospheric density structure. Similarly, using much larger x and y values in equation (A2) would overestimate the contribution from relatively focused sublithospheric density anomalies. Choosing adequate values for the surface area in equation (A2) is therefore important, as otherwise the density structure (i.e., thermochemical structure) constrained by the 1-D stage may not represent an acceptable initial model for the 3-D stage, thus compromising the performance and reliability of the 3-D stage. Although there is no simple rule of thumb to determine a priori the surface dimension of the prisms, in practice it is straightforward to find adequate values by running synthetic tests and/or preliminary inversions. Numerous tests indicated that for the study area values of 2.5×10^4 -5.7×10^4 km^2 provide good results.

Appendix B: Inversion Including Receiver Functions

In order to check the consistency of our crustal model with receiver function (RF) data in the RM region, we included local P wave RF information (Station S22A, long: −106.8, lat: 37.7) into the inversion. In this example we use four crustal layers instead of three in order to improve the fit to RF. The results of joint fits to dispersion and RF data is shown in Figure B1; fits to all other data sets are similar to those obtained in the previous

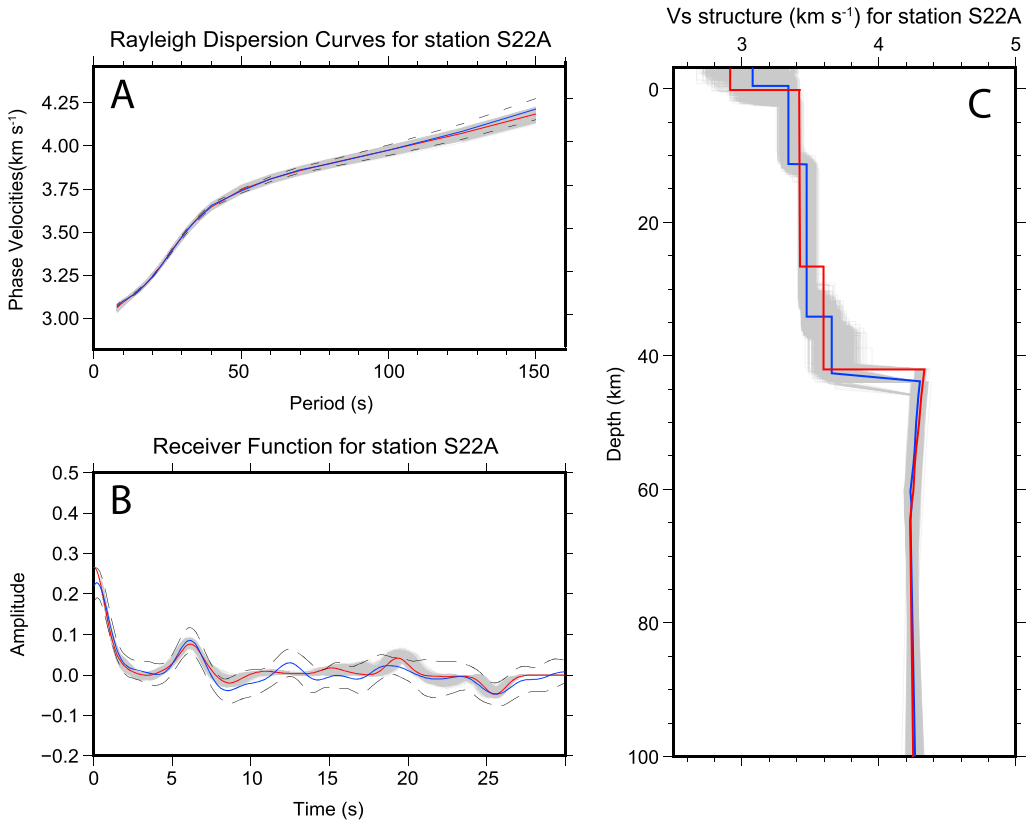

Figure B1. Results from a joint inversion including local receiver function (RF) data. (a) Fit to dispersion data. Gray lines represent the last 5000 models of the entire MCMC chain obtained in the inversion; red line is the mean of the 5000 models; blue line is the mean of the data; dashed lines are 1σ associated with the data. (b) Fit to receiver function data for station S22A (located at Long: −106.8, Lat: 37.7); gray, blue, dashed and red lines as in Figure B1a. (c) Crustal model in terms of its Vs structure. Blue line is the mean of the last 5000 models in the MCMC chain obtained by including RF data; red line is the mean from the full posterior PDF obtained without RF data and described in the main text. Both crustal models (with and without RF) predict similar average Vp/Vs ratios (~1.78–1.79).

inversion without RF. Considering the difference in the number of crustal layers, both crustal models (with and without RF data) are almost identical. The one constrained by RF shows a slightly deeper Moho (Figure B1) and produces a good fit to RF data. Nevertheless, the corresponding Vp/Vs average crustal value obtained when including RF remains similar to the one obtained without RF and reported in the text (Vp/Vs ~ 1.78–1.79). We are currently working on a full joint inversion including RF in the entire region. We expect that this study will help better constrain Vp/Vs values in the crust.

Acknowledgments

We are indebted to F. Darbyshire and J. von Hunen for useful comments on earlier versions of this work. This manuscript benefited from thorough and constructive reviews by W. Levandowski and an anonymous reviewer. We also thank J. Connolly, M. Sambridge, B. Kennett, S. Lebedev, B. Shan, U. Faul, and M. Qashqai for insightful discussions about, and contributions to, some of the concepts presented in this paper. The work of J.C.A. has been supported by two Australian Research Council Discovery grants (DP120102372 and DP110104145). Seismic data are from the IRIS DMS. D.L.S. acknowledges support from NSF grant EAR-135866. This is contribution 848 from the ARC Centre of Excellence for Core to Crust Fluid Systems (http://www.ccfs.mq.edu.au) and 1106 in the GEMOC Key Centre (http://www.gemoc.mq.edu.au).

References

Afonso, J. C., and D. Schutt (2012), The effects of polybaric partial melting on the density and seismic velocities of mantle restites, *Lithos*, *134–135*, 289–303.

Afonso, J. C., M. Fernàndez, G. Ranalli, W. L. Griffin, and J. A. D. Connolly (2008a), Integrated geophysical-petrological modelling of the lithospheric-sublithospheric upper mantle: Methodology and applications, *Geochem. Geophys. Geosyst.*, *9*, Q05008, doi:10.1029/2007GC001834.

Afonso, J. C., S. Zlotnik, and M. Fernandez (2008b), The effects of compositional and rheological stratifications on small-scale convection under the oceans: Implications for the thickness of oceanic lithosphere and seafloor flattening, *Geophys. Res. Lett.*, *35*, L20308, doi:10.1029/2008GL035419.

Afonso, J. C., G. Ranalli, M. Fernandez, W. L. Griffin, S. Y. O'Reilly, and U. Faul (2010), On the Vp/Vs - Mg# correlation in mantle peridotites: Implications for the identification of thermal and compositional anomalies in the upper mantle, *Earth Planet. Sci. Lett.*, *289*, 606–618.

Afonso, J. C., J. Fullea, W. L. Griffin, Y. Yang, A. G. Jones, J. A. D. Connolly, and S. Y. O'Reilly (2013a), 3-D multi-observable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle. I: A priori petrological information and geophysical observables, *J. Geophys. Res. Solid Earth*, *118*, 2586–2617, doi:10.1002/jgrb.50124.

Afonso, J. C., J. Fullea, Y. Yang, J. A. D. Connolly, and A. G. Jones (2013b), 3-D multi-observable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle. II: General methodology and resolution analysis, *J. Geophys. Res. Solid Earth*, *118*, 1650–1676, doi:10.1002/jgrb.50123.

Afonso, J. C., S. Zlotnik, and P. Diez (2015), An efficient and general approach for implementing thermodynamic phase equilibria information in geophysical and geodynamic studies, *Geochem. Geophys. Geosyst.*, *16*, 3767–3777, doi:10.1002/2015GC006031.

Afonso, J. C., M. Moorkamp, and J. Fullea (2016), Imaging the lithosphere and upper mantle, in *Integrated Imaging of the Earth: Theory and Applications*, vol. 218, edited by M. Moorkamp et al., pp. 191–218, Wiley, Hoboken, N. J.

Ashchepkov, I. V., H. Downes, R. Mitchell, and N. V. Vladykin (2013), Wyoming craton mantle lithosphere: Reconstructions based on xenocrysts from Sloan and Kelsey Lake Kimberlites, in *Proceedings of 10th International Kimberlite Conference*, vol 1, edited by D. G. Pearson et al., Springer, India, doi:10.1007/978-81-322-1170-9_2.

Alibert, C. (1994), Peridotite xenoliths from western Grand Canyon and the Thumb: A probe into the subcontinental mantle of the Colorado Plateau, *J. Geophys. Res.*, *99*, 21,605–21,620.

Aizawa, Y., A. Barnhoorn, U. H. Faul, J. D. Fitz Gerald, I. Jackson, and I. Kovács (2008), Seismic properties of Anita Bay Dunite: An exploratory study of the influence of water, *J. Petrol.*, *49*, 841–855.

Aoki, K. (1981), Major element geochemistry of chromian spinel peridotite xenoliths in the Green Knobs Kimberlite, New Mexico. Science Reports, Tohoku University, Series 3, *Mineral. Petrol. Econ. Geol.*, *15*(1), 127–130.

Ballmer, M. D., G. Ito, J. van Hunen, and P. J. Tackley (2011), Small-scale convection induces spatio-temporal variability in Hawaiian volcanism, *Nat. Geosci.*, *4*, 457–460.

Becker, T. W., C. Faccenna, E. D. Humphreys, A. R. Lowry, and M. S. Miller (2014), Static and dynamic support of western United States topography, *Earth Planet. Sci. Lett.*, *401*, 234–246.

Begg, G. C., et al. (2009), The lithospheric architecture of Africa: Seismic tomography, mantle petrology and tectonic evolution, *Geosphere*, *5*, 23–50.

Beran, A., and E. Libowitzky (2006), Water in natural mantle minerals II: Olivine, garnet and accessory minerals, *62*, 169–191

Biegler, L., et al. (Eds.) (2011), *Large-Scale Inverse Problems and Quantification of Uncertainty*, Wiley.

Blake, J. R., and A. T. Chwang (1974), Fundamental singularities of viscous flow, *J. Eng. Math.*, *8*(1), 23–29, doi:10.1007/bf02353701.

Brocher, T. M. (2005), Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. Seismol. Soc. Am.*, *95*, 2081.

Bouman, J., et al. (2013), GOCE gravity gradient data for lithospheric modeling, *Int. J. Appl. Earth Observ. Geoinform.*, *35*, 16–30, doi:10.1016/j.jag.2013.11.001.

Buck, W. R. (1986), Small-scale convection induced by passive rifting: the cause of uplift of rift shoulders, *Earth Planet. Sci. Lett.*, *77*, 362–372.

Buck, W. R., and E. M. Parmentier (1986), Convection beneath young oceanic lithosphere: Implications for thermal structure and gravity, *J. Geophys. Res.*, *91*(B2), 1961–1974, doi:10.1029/JB091iB02p01961.

Burov, E., and T. Gerya (2014), Asymmetric three-dimensional topography over mantle plumes, *Nature*, *513*, 85–89.

Carlson, R. W., A. J. Irving, D. J. Schulze, and B. C. Hearn Jr. (2004), Timing of Precambrian melt depletion and Phanerozoic refertilization events in the lithospheric mantle of the Wyoming Craton and adjacent Central Plains Orogen, *Lithos*, *77*, 453–472.

Cammarano, F., S. Goes, P. Vacher, and D. Giardini (2003), Inferring upper mantle temperatures from seismic velocities, *Phys. Earth planet. Inter.*, *139*, 197–222.

Coltorti, M., and G. Gregoire (Eds.) (2008), *Metasomatism in Oceanic and Continental Lithospheric Mantle: Geological Society Special Publication*, 361 pp., vol. 293, Geol. Soc. Spec. Publ., London.

Connolly, J. A. D. (2009), The geodynamic equation of state: What and how, *Geochem. Geophys. Geosyst.*, *10*, Q10014, doi:10.1029/2009GC002540.

Crameri, F., H. Schmeling, G. J. Golabek, T. Duretz, R. Orendt, S. J. H. Buiter, D. A. May, B. J. P. Kaus, T. V. Gerya, and P. J. Tackley (2012), A comparison of numerical surface topography calculations in geodynamic modelling: An evaluation of the sticky air method, *Geophys. J. Int.*, *189*, 38–54, doi:10.1111/j.1365-246X.2012.05388.x.

Crow, R., et al. (2011), Shrinking of the Colorado Plateau via lithospheric mantle erosion: Evidence from Nd and Sr isotopes and geochronology of Neogene basalts, *Geology*, *39*, 27–31.

Crossey, L. J., et al. (2009), Degassing of mantle-derived CO_2 and He from springs in the southern Colorado Plateau region, *Geol. Soc. Am. Bull.*, *121*, 1034–1053.

Dalton, C., and U. H. Faul (2010), The oceanic and cratonic upper mantle: Clues from joint interpretation of global velocity and attenuation models, *Lithos*, *120*, 160–172.

Davies, D. R., and N. Rawlinson (2014), On the origin of recent intra-plate volcanism in Australia, *Geology*, *42*, 1031–1034, doi:10.1130/G36093.1.
Decker, E. R. (1995), Thermal regimes of the southern Rocky Mountains and Wyoming Basin in Colorado and Wyoming in the United States, *Tectonophysics*, *244*, 85–106.
Deschamps, F., J. Trampert, and R. Snieder (2002), Anomalies of temperature and iron in the uppermost mantle inferred from gravity data and tomographic models, Phys, *Earth planet. Inter.*, *129*, 245–264.
Doin, M.-P., L. Fleitout, and D. McKenzie (1996), Geoid anomalies and the structure of continental and oceanic lithospheres, *J. Geophys. Res.*, *101*(B7), 16,119–16,135, doi:10.1029/96JB00640.
Eaton, D. W., F. Darbyshire, R. L. Evans, H. Grutter, A. G. Jones, and X. Yuan (2009), The elusive lithosphere-asthenosphere boundary (LAB) beneath cratons, *Lithos*, *109*, 1–22.
Eggler, D. H., J. K. Meen, F. Welt, F. O. Dudas, K. P. Furlong, M. E. McCallum, and R. W. Carlson (1988), Tectonomagmatism of the Wyoming Province, in *Colorado Volcanism*, edited by J. Drexler and E. E. Larson, pp. 25–40, Colorado School of Mines Quarterly, Boulder, Colo.
Ehrenberg, S. N., and W. L. Griffin (1979), Garnet granulite and associated xenoliths in minette and serpentinite diatremes of the Colorado Plateau, *Geology*, *7*, 483–487.
Ehrenberg, S. N. (1982), Petrogenesis of garnet lherzolite and megacrystalline nodules from the Thumb, Navajo volcanic field, *J. Petrol.*, *23*(4), 507–547, doi:10.1093/petrology/23.4.507.
Evans, J. R., and U. Achauer (1993), Teleseismic velocity tomography using the ACH method: Theory and application to continental-scale studies, in *Seismic Tomography*, edited by H. M. Iyer and K. Hirahara, pp. 319–360, Chapman and Hall, London.
Faul, U. H., and I. Jackson (2005), The seismological signature of temperature and grain size variation in the upper mantle, *Earth Planet. Sci. Lett.*, *234*, 119–134.
Flament, N., M. Gurnis, and R. D. Muller (2013), A review of observations and models of dynamic topography, *Lithosphere*, *5*, 189–210.
Forte, A. M., and H. K. C. Perry (2000), Geodynamic evidence for a chemically depleted continental tectosphere, *Science*, *290*, 1940–1944.
Frost, D. J. (2006), The stability of hydrous mantle phases, *Rev. Min. Geochem.*, *62*, 243–271.
Fullea, J., J. C. Afonso, J. A. D. Connolly, M. Fernàndez, D. Garcia-Castellanos, and H. Zeyen (2009), LitMod3D: An interactive 3-D software to model the thermal, compositional, density, seismological, and rheological structure of the lithosphere and sublithosheric upper mantle, *Geochem. Geophys. Geosyst.*, *10*, Q08019, doi:10.1029/2009GC002391.
Fullea, J., S. Lebedev, M. R. Agius, A. G. Jones, and J. C. Afonso (2012), Lithospheric structure in the Baikal-central Mongolia region from integrated geophysical-petrological inversion of surface-wave data and topographic elevation, *Geochem. Geophys. Geosyst.*, *13*, Q0AK09, doi:10.1029/2012GC004138.
Fullea, J., J. Rodríguez-González, M. Charco, Z. Martinec, A. Negredo, and A. Villaseñor (2015), Perturbing effects of sub-lithospheric mass anomalies in GOCE gravity gradient and other gravity data modelling: Application to the Atlantic-Mediterranean transition zone, *Int. J. App. Earth Obs. Geoinf.*, *35*, 54–69.
Gerya, T. (2010), *Introduction to Numerical Geodynamic Modelling*, Cambridge Univ. Press, Cambridge, U. K.
Gilbert, H. J., and A. F. Sheehan (2004), Images of crustal variations in the intermountain west, *J. Geophys. Res.*, *109*, B03306, doi:10.1029/2003JB002730.
Gilbert, H. (2012), Crustal structure and signatures of recent tectonism as influenced by ancient terranes in the western United States, *Geosphere*, *8*, 141–157, doi:10.1130/GES00720.1.
Goes, S., R. Govers, and P. Vacher (2000), Shallow mantle temperatures under Europe from P and S wave tomography, *J. Geophys. Res.*, *105*(B5), 11,153–11,169, doi:10.1029/1999JB900300.
Gradmann, S., J. Ebbing, and J. Fullea (2013), Integrated geophysical modelling of a lateral transition zone in the lithospheric mantle under Norway and Sweden, *Geophys. J. Int.*, *194*(3), 1358–1373.
Grant, K. J., J. Ingrin, J. P. Lorland, and P. Dumas (2007), Water partition between mantle minerals from peridotite xenoliths, *Cont. Min. Pet.*, *154*, 15–34.
Griffin, W. L., G. C. Begg, and S. Y. O'Reilly (2013), Continental-root control on the genesis of magmatic ore deposits, *Nat. Geosci.*, *6*, 905–910.
Griffin, W. L., S. Y. O'Reilly, J. C. Afonso, and G. Begg (2009), The composition and evolution of lithospheric mantle: A re-evaluation and its tectonic implications, *J. Petrol.*, *50*, 1185–1204.
Guo, Z., J. C. Afonso, M. T. Qashqai, Y. Yang, and Y. J. Chen (2016), Thermochemical structure of the North China Craton from multi-observable probabilistic inversion: Extent and causes of cratonic lithosphere modification, *Gondwana Res.*, *37*, 252–265, doi:10.1016/j.gr.2016.07.002.
Haario, H., M. Laine, A. Mira, and E. Saksman (2006), DRAM: Efficient adaptive MCMC, *Stat. Comp.*, *16*(4), 339–354.
Hammond, W. C., and E. D. Humphreys (2000), Upper mantle seismic wave attenuation: Effects of realistic partial melt distribution, *J. Geophys. Res.*, *105*, 10,987–10,999.
Hansen, S. M., K. G. Dueker, J. C. Stachnik, R. C. Aster, and K. E. Karlstrom (2013), A rootless rockies—Support and lithospheric structure of the Colorado Rocky Mountains inferred from CREST and TA seismic data, *Geochem. Geophys. Geosyst.*, *14*, 2670–2695, doi:10.1002/ggge.20143.
Harper, J. F. (1984), Mantle flow due to internal vertical forces, *Phys. Earth Planet. Inter.*, *36*, 285–290, doi:10.1016/0031-9201(84)90052-9.
Hasterok, D., and D. S. Chapman (2007), Continental thermal isostasy: 2. Application to North America, *J. Geophys. Res.*, *112*, B06415, doi:10.1029/2006JB004664.
Haxby, W. F., and D. L. Turcotte (1978), On isostatic geoid anomalies, *J. Geophys. Res.*, *83*(B11), 5473–5478, doi:10.1029/JB083iB11p05473.
Herrmann, R. B. (2002), Computer Program in Seismology, Version 3.3, Saint Louis University. [Available at http://www.eas.slu.edu/eqc/eqcsoftware.html.]
Hirschmann, M. M. (2000), Mantle solidus: Experimental constraints and the effects of peridotite composition, *Geochem. Geophys. Geosyst.*, *1*, 1042, doi:10.1029/2000GC000070.
Hirth, G., and D. L. Kohlstedt (1996), Water in the oceanic upper mantle: Implications for rheology, melt extraction and the evolution of the lithosphere, *Earth Planet. Sci. Lett.*, *144*, 93–108.
Humphreys, E., and D. Coblentz (2007), North American dynamics and western US tectonics, *Rev. Geophys.*, *45*, RG3001, doi:10.1029/2005RG000181.
Humphreys, E., et al. (2003), How laramide-age hydration of North American lithosphere by the farallon slab controlled subsequent activity in the Western United States, *Int. Geol. Rev.*, *45*, 575–595.
Jackson, I., and U. H. Faul (2010), Grainsize-sensitive viscoelastic relaxation in olivine: Towards a robust laboratory-based model for seismological application, *Phys. Earth Planet. Int.*, *183*, 151–163.
Jarvis, G. T., and W. R. Peltier (1982), Mantle convection as a boundary layer phenomenon, *Geophys. J. R. Astron. Soc.*, *68*, 389–427, doi:10.1111/j.1365-246X.1982.tb04907.x.

Jones, A. G., J. Fullea, R. L. Evans, and M. R. Muller (2012), Water in cratonic lithosphere: Calibrating laboratory-determined models of electrical conductivity of mantle minerals using geophysical and petrological observations, *Geochem., Geophys., Geosys.*, *13*, Q06010, doi:10.1029/2012GC004055.
Jones, C. H., K. H. Mahan, L. A. Butcher, W. B. Levandowski, and G. L. Farmer (2015), Continental uplift through crustal hydration, *Geology*, *43*, 355–358, doi:10.1130/G36509.1.
Kaislaniemi, L., and J. van Hunen (2014), Dynamics of lithospheric thinning and mantle melting by edge-driven convection: Application to Moroccan Atlas mountains, *Geochem. Geophys. Geosyst.*, *15*(8), 3175–3189, doi:10.1002/2014GC005414.
Kaislaniemi, L., J. van Hunen, M. B. Allen, and I. Neill (2014), Sublithospheric small-scale convection: A mechanism for collision zone magmatism, *Geology*, *42*, 291–294.
Karato, S.-I. (1986), Does partial melting reduce the creep strengh of the upper mantle?, *Nature*, *319*, 309–310.
Karato, S.-I. (1993), Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, *20*, 1623–1626.
Karato, S.-I. (2008), *Deformation of Earth Materials: An Introduction to the Rheology of the Solid Earth*, Cambridge Univ. Press, Cambridge, U. K.
Karato, S-I. (2012), On the origin of the asthenosphere, *Earth Planet. Sci. Lett.*, *321*, 95–103.
Karato, S.-I., and B. B. Karki (2001), Origin of lateral variation of seismic wave velocities and density in the deep mantle, *J. Geophys. Res.*, *106*(21), 21,771–21,783.
Kennett, B. (2002), *The Seismic Wavefield*, vol. 1, Cambridge Univ. Press, Cambridge, U. K.
Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.*, *122*, 108–124.
Khan, A., L. Boschi, and J. A. D. Connolly (2011), Mapping the Earth's thermochemical and anisotropic structure using global surface wave data, *J. Geophys. Res.*, *116*, B01301, doi:10.1029/2010JB007828.
Khan, A. (2016), On Earth's mantle constitution and structure from joint analysis of geophysical and laboratory-based data: An example, *Surv. Geophys.*, *37*(1), 149–189, doi:10.1007/s10712-015-9353-z.
Karlstrom, K. E., et al. (2012), Mantle-driven dynamic uplift of the Rocky Mountains and Colorado Plateau and its surface response: Toward a unified hypothesis, *Lithosphere*, *4*, 3–22.
Kitanidis, P. K. (2011), Bayesian and geostatistical approaches to inverse problems, in *Large-Scale Inverse Problems and Quantification of Uncertainty*, edited by L. Biegler et al., pp. 71–85, Wiley, Chichester, U. K.
Kohlstedt, D. L. (2006), The role of water in high-temperature rock deformation (in water in nominally anhydrous minerals), *Rev. Mineral. Geochem.*, *62*, 377–396.
Kucks, R. P. (1999), Bouguer gravity anomaly data grid for the conterminous US. U.S.Geological Survey, National geophysical data grids; gamma-ray, gravity, magnetic, and opographic data for the conterminous United States. USGS Digital Data Series.
Kudo, A. M., D. G. Brookins, and A. W. Laughlin (1972), Sr isotopic disequilibrium in lherzolites from the Puerco Necks, New Mexico, *Earth Planet. Sci. Lett.*, *15*, 291–295.
Kumar, P., X. Yuan, R. Kind, and J. Mechie (2012), The lithosphere-asthenosphere boundary observed with USArray receiver functions, *Solid Earth*, *3*, 149–159.
Kuskov, O. L, V. A. Kronrod, and H. Annersten (2006), Inferring upper-mantle temperatures from seismic and geochemical constraints: Implications for Kaapvaal craton, *Earth Planet. Sci. Lett.*, *244*, 133–154.
Laughlin, A. W., D. G. Brookings, A. M. Kudo, and J. D. Causey (1971), Chemical and strontium isotopic investigations of ultramafic inclusions and basalt, Bandera Crate, New Mexico, *Geochem. Cosmochem. Acta*, *35*, 107–113.
Lee, C.-T. (2003), Compositional variation of density and seismic velocities in natural peridotites at STP conditions: Implications for seismic imaging of compositional heterogeneities in the upper mantle, *J. Geophys. Res.*, *108*, 2441, doi:10.1029/2003JB002413.
Lee, C-T., Q. Yin, R. L. Rudnick, and S. B. Jacobsen (2001), Preservation of ancient and fertile lithospheric mantle beneath the southwestern United States, *Nature*, *411*, 69–73.
Lekic, V., and K. Fisher (2014), Contrasting lithospheric signatures across the western United States revealed by Sp receiver functions, *Earth Planet. Sci. Lett.*, *402*, 90–98.
Levander, A., et al. (2011), Continuing Colorado plateau uplift by delamination-style convective lithospheric downwelling, *Nature*, *472*, 461–465.
Levandowski, W., C. H. Jones, W. Shen, M. H. Ritzwoller, and V. Schulte-Pelkum (2014), Origins of topography in the western U.S.: Mapping crustal and upper mantle density variations using a uniform seismic velocity model, *J. Geophys. Res. Solid Earth*, *119*, 2375–2396, doi:10.1002/2013JB010607.
Li, Z.-X., et al. (2008), Water contents in mantle xenoliths from the Colorado Plateau and vicinity: Implications for the mantle rheology and hydration-induced thinning of continental lithosphere, *J. Geophys. Res.*, *113*, B09210, doi:10.1029/2007JB005540.
Lin, F-C., M. H. Ritzwoller, Y. Yang, M. P. Moschetti, and M. J. Fouch (2010), Complex and variable crustal and uppermost mantle seismic anisotropy in the Western United States, *Nat. Geosci.*, *4*(1), 55–61, doi:10.1038/ngeo1036.
Liu, K., A. Levander, F. Niu, and M. S. Miller (2011), Imaging crustal and upper mantle structure beneath the Colorado Plateau using finite frequency Rayleigh wave tomography, *Geochem. Geophys. Geosyst.*, *12*, Q07001, doi:10.1029/2011GC003611.
Liu, L. (2015), The ups and downs of North America: Evaluating the role of mantle dynamic topography since the Mesozoic, *Rev. Geophys.*, *53*, 1022–1049, doi:10.1002/2015RG000489.
Liu, L., and M. Gurnis (2010), Dynamic subsidence and uplift of the Colorado Plateau, *Geology*, *38*, 663–666.
Lorand, J. P., G. Schmidt, H. Palme, and K. L Kratz (2000), Highly siderophile element geochemistry of the Earth's mantle: New data for the Lanzo (Italy) and Ronda (Spain) orogenic peridotite bodies, *Lithos*, *53*, 149–164.
Lowry, A. R., and M. Perez-Gussinye (2011), The role of crustal quartz in controlling Cordilleran deformation, *Nature*, *471*(7338), 353–357, doi:10.1038/nature09912.
Marler, R. T., and J. S. Arora (2004), Survey of multi-objective optimization methods for engineering, *Struct. Multidisc. Optim.*, *26*, 369–395.
Marquart, G., and H. Schmeling (1989), Topography and geoid undulations caused by small scale convection beneath continental lithosphere of variable elastic thickness, *Geophys. J. Int.*, *97*, 511–527, doi:10.1111/j.1365-246X.1989.tb00520.X.
Meqbel, N. M., et al. (2014), Deep electrical resistivity structure of the northwestern US derived from 3D inversion of USArray magnetotelluric data, *Earth Planet. Sci. Lett.*, *402*, 290–304.
Molnar, P., P. C. England, and C. H. Jones (2015), Mantle dynamics, isostasy, and the support of high terrain, *J. Geophys. Res. Solid Earth*, *120*, 1932–1957, doi:10.1002/2014JB011724.
Morgan, W. (1965), Gravity anomalies and convection currents, *J. Geophys. Res.*, *70*, 6175–6187.
McQuarrie, N, and C. G. Chase (2000), Raising the Colorado Plateau, *Geology*, *28*, 91–94.
McGuire, A. V., and S. B. Mukasa (1997), Magmatic modification of the uppermost mantle beneath the basin and range to Colorado Plateau Transition Zone; Evidence from xenoliths, Wikieup, *Arizona. Cont. Min. Petr.*, *128*(1), 52–65.

Moschetti, M. P., M. H. Ritzwoller, F. Lin, and Y. Yang (2010a), Seismic evidence for widespread western-US deep-crustal deformation caused by extension, *Nature*, *464*(7290), 885–889, doi:10.1038/nature08951.
Moschetti, M. P., M. H. Ritzwoller, F.-C. Lin, and Y. Yang (2010b), Crustal shear wave velocity structure of the Western United States inferred from ambient seismic noise and earthquake data, *J. Geophys. Res.*, *115*, B10306, doi:10.1029/2010JB007448.
Nagy, D., G. Papp, and J. Benedek (2000), The gravitational potential and its derivatives for the prism, *J. Geod.*, *74*, 552–560, doi:10.1007/s001900000116.
Nolet, G. (1987), Seismic wave propagation and seismic tomography, in *Seismic Tomography*, edited by G. Nolet, pp. 1–23, Reidel, Dordrecht, Netherlands.
Obrebski, M., R. M. Allen, F. Pollitz, and S. H. Hung (2011), Lithosphere-asthenosphere interaction beneath the western United States from the joint inversion of body wave traveltimes and surface wave phase velocities, *Geophys. J. Int.*, *185*(2), 1003–1021, doi:10.1111/j.1365-246X.2011.04990.x.
O'Reilly, S. Y., and W. L. Griffin (2013), Mantle metasomatism, in *Metasomatism and the Chemical Transformation of Rock*, pp. 471–533, Springer, Berlin.
Pail, R., et al. (2011), First GOCE gravity field models derived by three different approaches, *Journal of Geodesy*, *85*(11), 819–843, doi:10.1007/s00190-011-0467-x.
Panet, I., G. Pajot-Metivier, M. Greff-Lefftz, L. Metivier, M. Diament, and M. Mandea (2014), Mapping the mass distribution of Earth's mantle using satellite-derived gravity gradients, *Nature Geosci.*, *7*(2), 131–135, doi:10.1038/ngeo2063.
Parsons, B., and J. G. Sclater (1977), An analysis of the variation of ocean floor bathymetry and heat flow with age, *J. Geophys. Res.*, *82*(5), 803–827, doi:10.1029/JB082i005p00803.
Parsons, B., and S. Daly (1983), The relationship between surface topography, gravity anomalies, and temperature structure of convection, *J. Geophys. Res.*, *88*(B2), 1129–1144, doi:10.1029/JB088iB02p01129.
Pasyanos, M. E., T. G. Masters, G. Laske, and Z. Ma (2014), LITHO1.0: An updated crust and lithospheric model of the Earth, *J. Geophys. Res. Solid Earth*, *119*, 2153–2173, doi:10.1002/2013JB010626.
Pavlis, N. K., S. A. Holmes, S. C. Kenyon, and J. K. Factor (2012), The development and evaluation of the Earth Gravitational Model 2008 (EGM2008), *J. Geophys. Res.*, *117*, B04406, doi:10.1029/2011JB008916.
Pearson, D. G., D. Canil, and S. B. Shirey (2003), Mantle samples included in volcanic rocks: Xenoliths and diamonds, in *The Mantle and Core (R. W. Carlson, ed.), vol. 2, Treatise on Geochemistry*, edited by H. D. Holland and K. K. Turkeian, pp. 169–253, Elsevier-Pergamon, Oxford, U. K.
Peslier, A. H., and J. F. Luhr (2006), Hydrogen loss from olivines in mantle xenoliths from Simcoe (USA) and Mexico: Mafic alkalic magma ascent rates and water budget of the sub-continental lithosphere, *Earth Planet. Sci. Let.*, *242*, 302–319.
Pollitz, F. F., and J. A. Snoke (2010), Rayleigh-wave phase-velocity maps and three-dimensional shear velocity structure of the western US from local non-plane surface wave tomography, *Geophys. J. Int.*, *180*, 1153–1169, doi:10.1111/j.1365-246X.2009.04441.x.
Porreca, C., J. Selverstone, and K. Samuels (2006), Pyroxenite xenoliths from the Rio Puerco volcanic field, New Mexico: Melt metasomatism at the margin of the Rio Grande rift, *Geosphere*, *2*, 333–351.
Priestley, K., and D. McKenzie (2006), The thermal structure of the lithosphere from shear wave velocities, *Earth Planet. Sci. Lett.*, *244*, 285–301.
Quarteroni, A., A. Manzoni, and F. Negri (2016), *Reduced Basis Methods for Partial Differential Equations: An Introduction*, 296 pp., Springer, Berlin.
Rawlinson, N., and S. Fishwick (2011), Seismic structure of the southeast Australian lithosphere from surface and body wave tomography, *Tectonophysics*, *572–573*, 111–122.
Rawlinson, N., A. M. Reading and B. L. N. Kennett (2006), Lithospheric structure of Tasmania from a novel form of teleseismic tomography, *J. Geophys. Res.*, *111*, B02301, doi:10.1029/2005JB003803.
Riter, J. C. A., and D. Smith (1996), Xenolith contraints on the thermal history of the mantle below the Colorado Plateau, *Geology*, *24*, 267–270.
Roy, M., T. H. Jordan, and J. Pedersen (2009), Colorado Plateau magmatism and uplift by warming of heterogeneous lithosphere, *Nature*, *459*, 978–982.
Rumpfhuber, E.-M., and G. R. Keller (2009), An integrated analysis of controlled and passive source seismic data across an Archean-Proterozoic suture zone in the Rocky Mountains, *J. Geophys. Res.*, *114*, B08305, doi:10.1029/2008JB005886.
Shan, B., J. C. Afonso, Y. Yang, C. J. Grose, Y. Zheng, X. Xiong, and L. Zhou (2014), The thermochemical structure of the lithosphere and upper mantle beneath south China: Results from multiobservable probabilistic inversion, *J. Geophys. Res. Solid Earth*, *119*, 8417–8441, doi:10.1002/2014JB011412.
Shen, W., M. Ritzwoller, and V. Schulte-Pelkum (2013), A 3-D model of the crust and uppermost mantle beneath the Central and Western US by joint inversion of receiver function and surface wave dispersion, *J. Geophys. Res. Solid Earth*, *118*, 1–15, doi:10.1029/2012JB009602.
Skogby, H. (2006), Water in natural mantle minerals I: Pyroxenes, *Rev. Min. Geochem.*, *62*, 155–167.
Smith, D. (1979), Hydrous minerals and carbonates in peridotite inclusions from the Green Knobs and Buell Park kimberlitic diatremes on the Colorado Plateau, in *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics*, edited by F. R. Boyd and M. H. O.A, pp. 345–356, AGU, Washington, D. C.
Smith, D. (2000), Insights into the evolution of the uppermost continental mantle from xenolith localities on and near the Colorado Plateau and regional comparisons, *J. Geophys. Res.*, *105*, 16,769–16,781.
Smith, D., and S. Levy (1976), Petrology of Green Knobs diatreme, New Mexico, and implications for the mantle below the Colorado Plateau, *Earth Planet. Sci. Lett.*, *19*, 107–125.
Smith, D., and J. C. A Riter (1997), Genesis and evolution of low-Al orthophroxene in spinel peridotite xenoliths, Grand Canyon field, Arizona, USA, *Cont. Min. Petrol.*, *127*, 391–404.
Smith, D., J. C. A. Riter, and S. A. Mertzman (1999), Water-rock interactions, orthopyroxene growth, and Si-enrichment in the mantle: Evidence in xenoliths from the Colorado Plateau, southwestern United States, *Earth Planet. Sci. Lett.*, *165*, 45–54.
Schmandt, B., and E. Humphreys (2010), Complex subduction and small-scale convection revealed by body-wave tomography of the western United States upper mantle, *Earth Planet. Sci. Lett.*, *297*, 435–445.
Shapiro, N. M., and M. H. Ritzwoller (2004), Thermodynamic constraints on seismic inversions, *Geophys. J. Int.*, *157*, doi:10.1111/j.1365-246X.2004.02254.x.
Shearer, P. M. (2009), *Introduction to Seismology*, 2nd ed., Cambridge Univ. Press, Cambridge, U. K.
Schulze, D. J., H. G. Coppersmith, B. Harte, and L.-A. Pizzolato (2008), Mineral inclusions in diamonds from the Kelsey Lake Mine, Colorado, USA: Depleted Archean mantle beneath the Proterozoic Yavapai province, *Geochim. Cosmochim. Acta*, *72*, 1685–1695, doi:10.1016/j.gca.2007.09.037.

Solomatov, V. S., and L. N. Moresi (2000), Scaling of time-dependent stagnant lid convection: Application to small-scale convection on Earth and other terrestrial planets, *J. Geophys. Res.*, *105*, 21,795–21,817.
Spencer, J. E. (1996), Uplift of the Colorado Plateau due to lithosphere attenuation during Laramide lowangle subduction, *J. Geophys. Res.*, *101*, 13,595–13,609.
Steinberger, B. (2016), Topography caused by mantle density variations: Observation-based estimates and models derived from tomography and lithosphere thickness, *Geophys. J. Int.*, *205*, 604–621, doi:10.1093/gji/ggw040.
Torrance, K. E., and D. L. Turcotte (1971), Structure of convection cells in the mantle, *J. Geophys. Res.*, *76*(5), 1154–1161, doi:10.1029/JB076i005p01154.
Turcotte, D. L., and G. Schubert (1982), *Geodynamics: Applications of Continuum Physics to Geological Problems*, John Wiley, New York.
Van Decar, J. C., and R. S. Crosson (1990), Determination of teleseismic relative phase arrival time using multi-channel cross-correlation and least squares, *Bull. Seismol. Soc. Am.*, *80*, 150–169.
van Hunen, J., S. Zhong, N. M. Shapiro, and M. H. Ritzwoller (2005), New evidence for dislocation creep from 3-D geodynamic modelling of the Pacific upper mantle structure, *Earth Planet. Sci. Lett.*, *238*(1–2), 146–155
van Wijk, J. W., et al. (2010), Small-scale convection at the edge of the Colorado Plateau: Implications for topography, magmatism, and evolution of Proterozoic lithosphere, *Geology*, *38*, 611–614.
West, M., W. Gao, and S. Grand (2004), A simple approach to the joint inversion of seismic body and surface waves applied to the southwest US, *Geophys. Res. Lett.*, *31*, L15615, doi:10.1029/2004GL020373.
Whitmeyer, S. J., and K. E. Karlstrom (2007), Tectonic model for the Proterozoic growth of North America, *Geosphere*, *3*, 220–259.
Winter, J. D. (2010), *An Introduction to Igneous and Metamorphic Petrology*, Prentice Hall, New York.
Wilshire, H. G., C. E. Meyer, J. K. Nakata, L. C. Calk, J. W. Shervais, J. E. Nielson, and E. C. Schwarzman (1988), Mafic and ultramafic xenoliths from volcanic rocks of the Western United States, *U.S. Geol. Surv. Open File Rep., 85-139*, Denver, Colo.
Yang, Y., and D. W. Forsyth (2006), Regional tomographic inversion of amplitude and phase of Rayleigh waves with 2-D sensitivity kernels, *Geophys. J. Int.*, *166*, 1148–1160.
Yang, Y., M. H. Ritzwoller, F.-C. Lin, M. P. Moschetti, and N. M. Shapiro (2008), Structure of the crust and uppermost mantle beneath the western United States revealed by ambient noise and earthquake tomography, *J. Geophys. Res.*, *113*, B12310, doi:10.1029/2008JB005833.
Yu, Y., X. S. Xu, W. L. Griffin, S. Y. O'Reilly, and Q. E. Xia (2011), H_2O contents and their modification in the Cenozoic subcontinental lithospheric mantle beneath the Cathaysia block, SE China, *Lithos*, *126*, 182–197.
Yuan, H., and B. Romanowicz (2010), Lithospheric layering in the North American craton, *Nature*, *466*, 1063–1068, doi:10.1038/nature09332.
Zhou, Y., G. Nolet, F. A. Dahlen, and G. Laske (2006), Global upper-mantle structure from finite-frequency surface wave tomography, *J. Geophys. Res.*, *111*, B04304, doi:10.1029/2005JB003677.
Zlotnik, S., J. C. Afonso, P. Diez, and M. Fernandez (2008), Small-scale gravitational instabilities under the oceans: Implications for the evolution of oceanic lithosphere and its expression in geophysical observables, *Philos. Mag. A*, *88*, 3197–3217, doi:10.1080/14786430802464248.