保几何特性的全波形层析成像及其在龙门山地区的应用

董兴朋^{1,2}, 杨顶辉^{2*}, 祝贺君³, 陈赟⁴

1. 中国地震局地质研究所, 北京 100029;

2. 清华大学数学科学系, 北京 100084;

3. Department of Geosciences, The University of Texas at Dallas, Dallas 75080, USA;

4. 中国科学院地质与地球物理研究所岩石圈演化国家重点实验室, 北京 100029

* 通讯作者, E-mail: ydh@tsinghua.edu.cn

收稿日期: 2021-01-19; 收修改稿日期: 2021-09-07; 接受日期: 2021-09-22; 网络版发表日期: 2021-12-17

国家重点研发计划项目(编号: 2017YFC1500301)和国家自然科学基金项目(批准号: U1839206、42004077)资助

摘要 经典的基于L2范数的波形反演难以克服周期跳跃问题, 导致迭代反演往往陷入局部极小值点, 无法给出正确的成像结果。根据最优传输理论, 文章采用一种新的基于Wasserstein度量的目标函数, 该度量能保持分布的几何特性, 提升反问题的稳定性与凸性。数值实验表明, 基于Wasserstein度量的全波形成像具有更大的收敛半径、更快的收敛速率并能有效克服周期跳跃问题。将此方法应用到龙门山地区, 获得了该区域可靠的岩石圈速度模型。成像结果显示, 四川盆地结晶地壳楔入高原块体内部, 青藏高原东部中下地壳表现出均匀的低S波速度特征, 表明中下地壳韧性弱物质流和块体间强相互作用可能共同主导了龙门山的隆升; 此外, 发现该区域内强震(如汶川、芦山地震)不仅发生在高、低速交界处, 还处于正、负径向各向异性过渡带上, 加深了对该区域强震孕育机制的理解和约束。

关键词 最优传输理论, 保几何特性, 周期跳跃问题, 龙门山造山带, 下地壳流模型

1 引言

青藏高原的隆升和向东扩展机制是当代构造地质学研究的热点。作为青藏高原东缘与扬子克拉通分界的主体(Copley和Mckenzie, 2007), 龙门山断裂带有着青藏高原周缘最陡峭的地形梯度, 在30km的水平距离内其海拔由500m左右急剧上升到5000m左右(王二七和孟庆任, 2008)。龙门山推覆构造带同时还位于中国南北地震带中段, 长约500km, 宽30~60km, 2008年汶

川M_s8.0级地震就发生在此, 是中国最显著的强震活动区域之一(何祥丽等, 2020)。龙门山地区地壳横向变形缓慢且缺失典型的前陆盆地(Burchfiel等, 1995), 使得其隆升机制难以用传统的断层和褶皱造山理论来解释。Tapponnier等(2001)和Royden等(1997)分别提出脆性地壳增厚假说和韧性中下地壳流假说来解释龙门山的形成, 这两种假说争论的本质在于龙门山地区的变形机制究竟是受控于浅部地壳(逆冲断层)还是深部地壳(下地壳蠕滑)。浅部地壳变形可以通过大地测量或

中文引用格式: 董兴朋, 杨顶辉, 祝贺君, 陈赟. 2022. 保几何特性的全波形层析成像及其在龙门山地区的应用. 中国科学: 地球科学, 52(3): 450–461, doi: 10.1360/SSTe-2021-0016

英文引用格式: Dong X, Yang D, Zhu H, Chen Y. 2022. Geometry-preserving full-waveform tomography and its application in the Longmen Shan area. Science China Earth Sciences, 65(3): 437–448, <https://doi.org/10.1007/s11430-021-9849-5>

地表地质调查来获得, 而要获取深部地壳的变形数据则并不容易.

自20世纪70年代中期Aki等提出地震层析成像方法至今(Aki和Lee, 1976), 已经历半个世纪的发展, 现已成为探测地球内部结构和深部过程的主要手段. 迄今为止, 基于射线理论进行走时层析成像在地球内部结构重建中仍占据主流地位. 但经典的射线理论是基于地震波是无限高频(波长无限短)假设提出的, 即地震波的旅行时仅受无限窄射线路径上的介质结构影响, 而与射线路径以外的介质变化无关. 但实际上, 从地震波的激发、传播到接收全过程, 无限高频假设均不具有物理可实现性, 即地震波必然是有限频的(Dahlen等, 2000; Hung等, 2004; 黄雪源等, 2021). 在有限频条件下, 强的小尺度非均匀体引起的散射和波前愈合效应, 在震相到时上是无法区分的, 即使在现有的计算资源下, 模型网格不断加密, 射线走时层析成像分辨率仍然有限(Tong等, 2011).

全波形反演(Full Waveform Inversion, FWI)是一种尽量使用波形记录全部特征信息的反演方法, 一般是基于波动方程进行全波场正演模拟, 包括了全部震相的相位、频率、振幅等波形中所蕴涵的特征信息, 从而跳出了经典走时成像在异常体尺度、扰动幅度、反演参数等方面的局限. 不同于传统的射线层析成像, 全波形层析成像考虑了波前愈合效应, 并且利用包含运动学和动力学信息在内的完整波形记录. 从数学上讲, 它是一个带偏微分方程约束的优化问题(Lailly, 1983; Virieux和Operto, 2009; Engquist和Froese, 2014), 通过不断减小预测数据和实际数据之间的误差来迭代更新预估模型. 近年来, 得益于计算能力的提高和现代数值计算方法的进步(如谱元法, Komatitsch和Tromp, 1999; Tromp等, 2008), 基于波动方程的全波形层析成像已经被应用到勘探、区域/大陆和全球尺度地震学研究中(Tarantola, 1984; Pratt和Worthington, 1990; Marquering等, 1999; Tromp等, 2005; Tape等, 2009; Fichtner等, 2009; Chen等, 2015; Bozdağ等, 2016; Jing等, 2016; Wang等, 2019).

作为衡量观测波形和合成波形近似程度的指标, 目标函数在一定程度上决定了反演的成败(Tromp, 2020). 经典的基于L2范数的全波形反演通过逐点拟合观测数据和合成数据来约束地下结构, 该类方法难以克服周期跳跃问题, 即当合成波形与观测波形之间时

移大于半个周期时, 优化目标函数中会存在许多局部极小值, 导致反演难以收敛到全局最优解(Virieux和Operto, 2009; Métivier等, 2016). 因此改善优化问题目标函数, 避免反演陷入局部极小值是全波形层析成像研究中的热点和难点. 最优传输理论是几何、概率论和偏微分方程等多学科理论的交叉、前沿研究领域, 近年来已在工程和医疗领域得到应用(Villani, 2003). Wasserstein度量是基于最优传输理论发展起来的一种新的度量理论, 它可以衡量两个概率密度分布在重排意义下的最优距离. 研究表明, 二次Wasserstein度量(W2范数)具有对时移周期单调的特性, 是解决周期跳跃问题的理想工具(Engquist和Froese, 2014; Yang等, 2018). 目前, W2范数已被应用于地震勘探领域的全波形反演中, 取得了很好的效果(Yang等, 2018). 基于W2范数的被动源伴随成像最近由Dong等(2019)提出, 研究表明该方法相较于传统的L2范数具有更大的收敛域和更快的收敛速度, 能够极大提升全波形层析成像的计算效率. 此外, 针对实际地震波形信号可能存在的强噪声干扰等问题, 提出多窗Wasserstein距离这一全新概念来尽可能降低噪声信号对成像结果的影响, 提高了成像结果的可靠性(Dong等, 2019; Dong和Yang, 2020).

然而, 上述研究中, W2度量的保几何特性以及它与Monge-Ampère方程的关系并没有被揭示出来, 且只给出了伴随震源的离散形式, 并没有给出连续形式的表达式(Dong等, 2019). 在本文中, 我们首先给出Wasserstein度量与Monge-Ampère方程之间的联系, 并详细阐述其保几何映射的特性, 该性质使得基于Wasserstein度量的目标函数具有更好的凸性. 然后, 我们从变分求导的定义出发, 基于泰勒级数展开给出了伴随震源表达式的连续形式, 加深了对基于W2度量伴随成像的理解. 虽然龙门山及其邻近地区已经有大量的地震层析成像结果, 揭示了四川盆地与青藏高原东部之间的速度结构变化(Lei和Zhao, 2009; 韦伟等, 2010; 李志伟等, 2011; 陈运泰等, 2013; 张风雪等, 2018; Pei等, 2019), 但该区域岩石圈结构仍然没有得到很好的约束. 前人的研究大多基于射线理论, 其优化问题计算成本小, 但仅在地震波长远小于非均匀体尺度时才有效(Rawlinson等, 2010). 我们将最新发展的基于W2范数的全波形地震层析成像应用到龙门山及其邻近区域, 获得了可靠的岩石圈成像结果, 揭示了龙门山及其邻近区域速度结构和岩石圈变形的细节.

2 基于Wasserstein距离的全波形成像

2.1 Wasserstein距离与Monge-Ampère方程

假定有两个总测度相等的概率分布 μ 和 ν , 我们寻找一个区域到自身的同胚映射, 该映射需要同时满足两个条件: (1) 保测度映射, (2) 两个概率测度之间的传输代价最小, 满足以上条件的映射我们称为这两个概率分布之间的最优传输映射(Villani, 2003).

对于任何有界的Borel函数 $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$, 映射 S 将概率分布 μ 映射成 ν 且满足:

$$\int_{\mathbb{R}^n} \varphi(y) d\nu(y) = \int_{\mathbb{R}^n} \varphi(S(x)) d\mu(x). \quad (1)$$

考虑到将一个点从 x 移动到 y 所需要的二次运输成本是 $|x-y|^2$, 则传输映射 $S: \mu \rightarrow \nu$ 对应的二次传输成本定义为

$$E(S) := \int_{\mathbb{R}^n} |x - S(x)|^2 d\mu(x). \quad (2)$$

在所有保测度映射组成的集合 M 中, 传输代价最小的映射被称为最优传输映射, 其对应的传输成本被称为概率分布 μ 和 ν 之间的Wasserstein距离(Figalli, 2017):

$$W^2(\mu, \nu) = \inf_{S \in M} \int_{\mathbb{R}^n} |x - S_{\mu, \nu}(x)|^2 d\mu(x). \quad (3)$$

Brenier(1991)证明在 μ 绝对连续且 μ 和 ν 均满足有限二阶矩情况下, 最优传输映射存在并唯一, 它可以表示成一个凸函数的梯度.

Brenier's定理: f 和 g 是 $\mathbb{R}^n$ 上的两个概率密度函数, 满足 $\mu=f dx$ 和 $\nu=g dx$. 如果 μ 在勒贝格测度意义下绝对连续, 则

(1) 存在唯一的最优传输映射 S ;

(2) 存在一个凸函数 $u: \mathbb{R}^n \rightarrow \mathbb{R}$, 最优传输映射满足 $S = \nabla u$.

我们将上述结论带入方程(1), 对于任意的 $\varphi \in C_c^\infty(\mathbb{R}^n)$, 有

$$\begin{aligned} & \int_{\mathbb{R}^n} \varphi(S(x)) f(x) dx \\ &= \int_{\mathbb{R}^n} \varphi(y) g(y) dy \\ &= \int_{\mathbb{R}^n} \varphi(S(x)) g(S(x)) |\det \nabla S(x)| dx. \end{aligned} \quad (4)$$

则我们可得

$$f(x) = g(S(x)) |\det \nabla S(x)|. \quad (5)$$

由条件 $S = \nabla u$ 可知 u 满足Monge-Ampère方程:

$$\det D^2 u = \frac{f}{g \circ \nabla u}. \quad (6)$$

Monge-Ampère方程是完全非线性的、退化的椭圆型偏微分方程, 最早由法国数学家Monge(1784)和Ampère(1820)提出. 近年来, Monge-Ampère方程已经在最优传输理论和天体物理学领域得到了重要应用(Frisch等, 2002; Brittany, 2009).

由上述可知, 要获得最优传输映射, 则必须求解Monge-Ampère方程. 然而, 高维的Monge-Ampère方程没有解析解, 数值求解亦十分复杂, 但在一维情况下, 传统的直方图均衡化就是最优传输映射(Gu等, 2016; Soheil等, 2016). 假如我们有两个连续的概率分布 μ 和 ν , 其对应的累积分布函数(Cumulative Distribution Function, CDF)为

$$F(x) = \int_{-\infty}^x d\mu, \quad G(y) = \int_{-\infty}^y d\nu. \quad (7)$$

那么直方图均衡化映射就是最优传输映射: $S(x) := G^{-1} \circ F(x)$.

实际地震信号是一维的时间序列, 故记录时长为 T_0 观测波形和合成波形之间的Wasserstein距离具有精确表达式(Yang等, 2018; Wu等, 2018; Dong等, 2019):

$$W^2(f, g) = \int_0^{T_0} |t - G^{-1}(F(t))|^2 f(t) dt. \quad (8)$$

2.2 Wasserstein距离的优势

在用 W_2 距离刻画的度量空间中, 原始信号 $d(t)$ 和时移信号 $d(t+s)$ 之间的测地线精准描述了时间偏移的过程, 而 L_2 距离刻画的欧式空间中, 两个信号之间的测地线与时间偏移之间没有物理联系, 导致了非线性反问题的多解性. 与 L_2 距离相比, Wasserstein距离具有以下两点优势(Soheil等, 2017): (1) 在保持分布自身几何特性的情况下将一个概率分布转换成另外一个; (2) 同时给出了两个概率分布之间的Wasserstein度量和映射方案. L_2 范数通过逐点比较两个概率分布之间的不同来计算其距离, 忽略了概率分布内在的几何特性. 图1中有两个分布 f 和 g , 在不同时移情况下, 它们之间的 L_2 范数距离是相等的, 可见 L_2 范数在某些情况下会

图 1 两个分布 f (虚线)和 g (实线)在不同时移(dt)下的L2范数与W2范数距离

“失效”，无法真实反映两个分布之间的远近，而W2范数则不存在此问题。保几何特性使得Wasserstein距离相较于L2距离有着更好的凸性，这里我们给出一个简单的数值算例来说明此问题。图2中， $f(t)$ 和 $g(t)$ 为两个一维的Ricker子波，它们振幅相同，两者只相差一个时移 dt ，可以看到两者L2距离函数在全局最小值附近存在两个局部极小值，而它们的W2距离函数则是一个严格凸函数。因此，在利用梯度类方法寻找最优解时，在L2范数下极易陷入局部极小值，而在W2范数下则可准确地收敛到全局最小值点。

此外，相较于L2距离，Wasserstein距离对噪声不敏感。Yang等(2018)证明，对于 N 个数据点的一维信号 $f(t)$ 和 $g(t)$ ，均匀噪声分布 $U[-c, c]$ 对L2距离的影响为 $O(1)$ 而对W2距离的影响则是 $O(1/N)$ 。这里给出一个数值算例来说明此问题，假定两个一维波形信号分别是合成波形 $f(t)$ 与观测波形 $g(t)$ ，在观测波形 $g(t)$ 上加均匀随机噪声得到 $g(t)+\text{noise}$ ，我们分别计算 $f(t)$ 与 $g(t)$ 、 $f(t)$ 与 $g(t)+\text{noise}$ 在L2范数和W2范数下的距离(图3)，可以看到噪声对L2距离影响较大，而对W2距离则几乎没有影响。

2.3 基于Wasserstein度量的伴随震源

全波形反演本质上是一个带有偏微分方程(Partial Differential Equation, PDE)约束的优化问题(Engquist

图 2 一维时间序列 f 与 g 分别在W2和L2范数意义下的距离
(a) Ricker子波/ f 和 g ; (b) 它们之间在不同时移 dt 下的W2和L2距离

和Froese, 2014)，通过不断减小合成信号和观测信号之间的误差来迭代更新模型。计算目标函数的Fréchet导数一般采用伴随状态法，即通过正向传播的地震波场与时间反转的伴随波场做相关来给出(Lailly, 1983; Tarantola, 1984)，因此Fréchet梯度和伴随震源的推导

图3 不同波形(a)及其在L2和W2范数下的距离(b)

$f(t)$, 合成波形; $g(t)$, 观测波形; $g(t)+\text{噪声}$, 添加均匀随机噪声后观测波形

十分重要, 因为Fréchet梯度给出了模型更新的方向, 而伴随震源是梯度计算的前提. 基于变分原理, 我们推导了W2度量下目标泛函梯度方向的严格数学形式. 假定 $f(t)$ 和 $g(t)$ 分别表示理论波形和观测波形, $h(t)$ 表示理论波形的扰动, ε 为小常数, 我们可以获得如下的伴随源:

$$\frac{\delta W^2[f(t)]}{\delta f(t)} = (t - G^{-1}(F(t)))^2 - 2 \int_{t_0}^{T_0} (t - G^{-1}(F(\tau))) \times \frac{f(\tau)}{g(G^{-1} \circ F(\tau))} d\tau. \quad (9)$$

其推导和离散化详见网络版附录A(<http://earthcn.scichina.com>).

2.4 数值算例

上述讨论表明, 与传统的L2度量相比, 保几何特性W2范数具有更好的凸性, 尤其适合使用梯度类的优化方法进行迭代反演. 在小扰动情况下, 基于L2度量和基于W2度量的波形反演均能给出很好的结果, 而在大扰动情况下(初始模型偏离真实模型超过半个波长), L2反演会由于周期跳跃问题而陷入局部极值, 而W2范数由于其对时移周期单调的特性, 反演仍能收敛到全局最优解. 为了说明此性质, 我们给出了一个数值算例. 速度为 4000 m s^{-1} 的矩形区域位于 3000 m s^{-1} 的背景

模型中(速度扰动超过30%, 图4a), 反演的初始模型为 3000 m s^{-1} 的均匀模型. 20个震源均匀排列在模型100m的深度层上, 而101个接收器则等间距排列在模型1900m的深度层上. 地震子波为主频10Hz的Ricker子波, 采用L-BFGS优化方法进行迭代反演(Nocedal, 1980). 图4b和4c分别为基于L2和W2范数的模型恢复结果, 误差收敛曲线见图4d. 可以看到, 基于W2范数的全波形反演收敛速度更快, 且很好地重建了模型, 而基于L2范数的反演则陷入了由周期跳跃产生的局部极小值中, 只进行了17次迭代反演便终止.

3 龙门山及邻近区域保几何全波形层析成像

龙门山及邻近地区的地形和地质构造如图4所示, 该区域是青藏高原东部松潘-甘孜地块和四川盆地的过渡地带, 除龙门山断裂带之外, 区内还分布有龙日坝断裂、龙泉山断裂和石棉断裂等. 前人的研究显示地壳和地幔中的各向异性与其介质(如云母、橄榄石等)晶体的晶格优势排列有直接成因联系, 因此地震波的各向异性分布能够为地壳/岩石圈变形机制提供重要约束(Savage, 1999; Karato等, 2008). 特别地, 地球深部正径向各向异性($V_{\text{SH}} > V_{\text{SV}}$)的低剪切波速层可能是由于岩石圈内较低强度的弱物质沿水平向流动导致矿物晶体定向排列产生的(Weiss等, 1999; Shapiro等, 2004; Huang等, 2010).

本研究一共使用了21个区域地震事件(震源参数见网络版附表1)和43个固定地震台站接收到的波形数据(郑秀芬等, 2009)(图5). 反演涉及四个径向各向异性参数, 分别为纵波速度(V_p), 水平极化横波速度(V_{SH}), 垂直极化横波速度(V_{SV})和密度(ρ). 对于单个波形窗口的W2距离, 我们可以定义如下的误差函数 χ_i :

$$\chi_i = \int_{T_0}^{T_1} |t - G^{-1}(F(t))|^2 f(t) dt, \quad (10)$$

其中 T_0 和 T_1 分别表示窗口的起始和结束时间. 误差函数的变分可以表示为模型扰动Fréchet敏感核的体积分(Tromp等, 2005; Zhu等, 2015):

$$\delta \chi = \int (K_{V_p} \delta \ln V_p + K_{V_{\text{SV}}} \delta \ln V_{\text{SV}} + K_{V_{\text{SH}}} \delta \ln V_{\text{SH}} + K_{\rho} \delta \ln \rho) d^3x, \quad (11)$$

K_{V_p} 、 $K_{V_{\text{SV}}}$ 、 $K_{V_{\text{SH}}}$ 和 K_{ρ} 分别对应P波、SV波、SH波和密

图 4 真实模型(a)、基于L2范数(b)和W2范数(c)目标函数的反演结果及误差收敛曲线(d)

(a)中五角星代表震源, 三角形代表接收器

度的敏感核. 我们利用谱元法(Komatitsch和Tromp, 1999; Fichtner等, 2009)模拟地震波场的传播并数值计算Fréchet导数. 对原始波形数据, 我们首先对其进行去仪器响应、去均值、去倾斜等处理, 然后将波形数据滤波至20~40s频带, 初始模型为FEWA18(Tao等, 2018). 在谱元法网格剖分中, 为了保证数值解的精度, 一般模拟的最短波长内至少要包含两个谱单元(Fichtner等, 2013). 在本研究中, 为了保证数值模拟的高精度, 我们对空间网格进行了加密处理, 保证最短波长内至少包含四个谱单元, 尽可能地提高波场模拟精度. 我们利用自动选取走时窗软件FLEXWIN(Maggi等, 2009)对特定的震相(如P、S震相等)进行框定, 进而计算该波形窗口的误差和伴随震源. 对于叠加于有效信号之上的噪音, 一方面, Wasserstein距离对噪声不敏感, 这也是其相较于L2距离的优势(Yang等, 2018); 另一方面, 我们通过对目标震相窗口施加一系列的先决条件来减低噪音的影响, 如目标震相窗口需满足: 互相关系数 $CC \geq 0.70$, 信噪比 $SNR \geq 3.0$ 等. 灵活的时间窗选取是全波形反演方法的一大优点, 有利于最大化使用波形数据获得高精度三维速度模型. 在本研究中,

体波和面波数据均参与了反演.

所有选定窗口的W2误差函数可以写成如下形式:

$$\chi(m) = \frac{1}{N_{\omega}} \sum_{e=1}^E \sum_{i=1}^{N_{\omega}^e} \int_{T_{i, \text{start}}}^{T_{i, \text{end}}} |t - G^{-1}(F(t))|^2 f(t) dt, \quad (12)$$

式中, N_{ω}^e 表示单个地震事件 e 的波形窗口数, E 表示地震事件总数, $N_{\omega} = \sum_{e=1}^E N_{\omega}^e$ 表示所有选定的时间窗. 所有敏感核求和即为梯度, 它指示了模型更新的方向. 我们采用共轭梯度法进行反演, 步长搜索策略采用Tape等(2007)给出的两点二次插值法, 对于梯度预处理, 我们遵循Fichtner等(2009)提出的方法, 该预处理因子能有效增强震源和台站之间波传播路径上的梯度值, 并降低震源和台站处的梯度值. 在反演过程中, 上一次迭代的模型作为下一次迭代的初始模型, 直至当前与上一次迭代之间误差减少不明显时, 即认为迭代达到收敛而终止反演过程. 需要指出的是, 虽然Wasserstein度量降低了目标函数的非凸性, 但地震波形信号与速度结构之间的强非线性仍可能会导致反演收敛速度缓慢, 故在反演迭代时若基于W2范数误差下降不明显

时, 我们则在此步迭代采用波形走时目标函数来降低反演非线性程度, 加快整个反演过程的收敛速度。

在全波形反演中, 执行传统的检测板测试非常耗时, 因为它需要与实际结构反演相同的计算资源(Zhu等, 2012)。作为替代方案, 我们采用Fichtner和Trampert (2011)提出的点扩散函数测试(Point-Spread Function test, PSF)来评估模型的成像质量。为了评估我们方法的水平分辨率以及不同反演参数之间的耦合情况, 我们在模型的左上部深度15km处设置一水平尺度为 $0.5^{\circ} \times 0.5^{\circ}$ 的低SV波速异常, 然后分别计算SV波、SH波、P波和密度的点扩散函数, 结果表明可以较好地恢复SV波扰动图案的主要特征, 且SV波与SH波、P波和密度之间的折中(trade-off)可以忽略(图6)。反演后走

时残差呈正态分布, 走时差均值从反演前的 -1.60s 减少到 -0.06s , 且波形窗的数量也从反演前的150增加到415, 说明反演后模型相较初始模型变得更好(图7)。为了进一步验证反演模型的可靠性, 我们给出了某地震事件在台站SC.AXI处反演前后的合成波形与观测波形三分量对比图, 可以看到, 反演后合成波形在相位和振幅上都得到了明显的提升, 与观测数据更为接近(图8)。

4 结果和讨论

图9给出了从地表2km到地下90km不同深度处龙门山及邻近地区3D岩石圈各向同性S波速度($v_s^{\text{iso}} = 2v_{\text{SH}}/3 + v_{\text{SV}}/3$)分布图。与前人层析成像结果相比, 我们模型揭示了该区域不同地块之间更强的速度对比, 给出了高速结晶地壳详细分布并确认了四川盆地所具有的稳定克拉通特性(Lei和Zhao, 2009; 李志伟等, 2011; 张风雪等, 2018)。在浅地表(2km), 由于巨厚沉积物的存在, 四川盆地以低剪切波速为主, 而表面

图5 龙门山及邻近地区地形及地质构造背景

黑线表示断层; 沙滩球表示反演用到的地震事件的震源机制; 黑色三角表示用到的台站位置; 大、小五角星应分别代表了汶川、芦山地震(震源参数信息来自于IRIS网站, <https://www.iris.edu/hq/>); 蓝色虚线对应图10所示剖面位置。

图6 研究区域左上15km深度处低SV波速区点扩散函数(PSF)测试

(a) 密度 ρ 的PSF, (b) P波的PSF, (c) SH波的PSF, (d) SV波的PSF。色标单位 $1 \times 10^{-3} \text{s}^2 \text{m}^{-4}$

图7 初始模型(a)和最终模型(b)的走时残差直方图
 右上角标出了两个模型走时差均值、方差以及波形窗口数量

图8 台站SC.AXI接收到的某地震事件反演前后观测波形和合成波形三分量对比图

该地震事件发震时间2014年, 震源深度24.6km; 黑线表示观测波形, 绿线和红线分别表示反演前、后合成波形

被三叠纪复理石所覆盖的松潘-甘孜地块则呈现高剪切波速(图9a)。在15和20km深度, 四川盆地整体呈现高剪切波速特征, 表明它具有连续的结晶基底分布, 而汶川地震和芦山地震均发生在高、低剪切波速区域交界处; 值得注意的是高剪切波速区跨过龙门山从四川盆地一直延伸至松潘-甘孜地块(图9b和9c), 接近龙日坝断裂带, 表明四川盆地的结晶地壳已经楔入青藏高原内部, 这与接收函数和深地震探测结果相一致(Wang等, 2018)。在45km深度, 相对高速的四川盆地被周围低速地块所包围; 在60和90km深度, 龙门山底部由高

剪切波速区转为中低剪切波速区(图9d~9f)。

为了进一步说明不同地体间的深部接触关系和变形机制, 图10给出了一条穿过松潘-甘孜地块、龙门山断层和四川盆地的垂直剖面A-A'(初始模型见网络版附录B)。SV波速度变化的基本特征是松潘-甘孜地块中下地壳表现出大范围的低速, 而稳定的四川盆地中下地壳则为高速, 上部覆盖着大量低速沉积物。我们定义如下的径向各向异性参数: $\zeta = (V_{SH} - V_{SV})/V_S$ 。龙门山底部中上地壳呈现正的径向各向异性, 结合中、下地壳弱物质的塑性特征和莫霍面深度的空间变化, 表

图9 不同深度层(2~90km)各向同性S波速度分布图

大、小五角星分别代表了汶川、芦山地震

明中、下地壳弱物质可能在龙门山底部聚集, 并沿莫霍面倾斜向上涌动(Dong等, 2019). 上述现象结合四川盆地结晶地壳楔入高原内部的特征, 我们认为不同地体之间在边界处的强烈相互作用和塑性地壳流共同控制了龙门山的隆起.

垂向剖面中一个显著的特征是强地震(例如汶川和芦山地震)发震位置不仅位于高速和低速区域的交界处, 而且还位于径向各向异性由正到负的过渡带上. 这表明这两次大地震的发生与沉积层、上地壳和中下地壳的变形有关. 汶川地震前, 龙门山断裂的走滑速率仅 1.5mm a^{-1} (Tang等, 2004), 断层处于高度闭锁状态. 在印度板块的推挤和高原的重力作用下, 中下地壳中的韧性物质沿莫霍面蠕变. 这种蠕变使上部地壳中的水平应力在青藏高原低速上地壳、四川盆地的沉积地层和四川盆地结晶地壳这三者交界处不断积累. 当这种应变累积到一定程度时, 就会产生失衡, 瞬间释放

出大量能量, 从而引发汶川地震. 然而, 芦山地震的发震机理可能与汶川地震不同. 芦山地震震源位于松潘-甘孜地块上、下地壳交界处, 其发震机制可能是中下地壳韧性弱物质材料的顶部与脆性上地壳底部之间的摩擦失稳. 前人关于汶川和芦山地震震源破裂过程反演显示两者断层破裂宽度相差十几倍, 平均滑动距离前者是后者的数倍(张勇等, 2008; 2013; 陈运泰等, 2013). 基于我们的成像结果, 我们认为汶川地震和芦山地震的发震机制不同, 使得两者在断层强度、规模和滑动距离方面存在差异, 加之由于中下地壳韧性材料的强度低, 故芦山地震无法达到汶川地震那样的规模.

5 结论

本文详细阐述了Wasserstein度量能保持概率分布

图 10 沿剖面A-A'的地表高程、剪切波速度和径向各向异性分布图

(a) 地表高程, (b) SV波速度, (c) 径向各向异性. PA, 正径向各向异性; NA, 负径向各向异性; HV, 高速; LV, 低速. 大、小五角星分别代表了汶川、芦山地震. 剖面位置见图5

自身几何特征的优良性质, 该特性使得基于Wasserstein度量的目标函数具有更好的凸性, 在一定程度上很好地解决了传统的基于L2范数的全波形反演难以克服的周期跳跃问题和非线性反演结果解非唯一性的本质问题. 然后, 利用变分原理导出了基于二次Wasserstein度量的Fréchet导数和伴随震源的表达式. 数值算例表明, 基于W2范数的全波形层析成像具有更大的收敛域, 对初始模型不敏感, 且迭代反演误差下降更快, 能够精确重建速度变化剧烈的地下模型, 极大地提高了全波形反演成像结果的可靠性. 进一步将本研究提出的基于W2范数的全波形地震层析成像新技术应用于龙门山及其邻近地区, 获得了该区域可靠的岩石圈速度模型. 结果显示, 龙门山区域的地壳形变同时受中下地壳的塑性流动和边界上不同块体之间强烈相互作用所控制, 同时发现该区域大地震通常发生在高、低速区域的交界处以及正、负径向各向异性区域之间的过渡带. 但是, 汶川地震和芦山地震的发震机制不同, 汶川地震是由于沿龙门山断裂带的逆冲推覆, 挤

压四川盆地的结晶地壳和沉积物造成的, 而芦山地震则是由韧性的中下地壳与脆性上地壳之间的摩擦失稳引起的.

致谢 本文绝大部分图件利用GMT程序包(Wessel等, 2013)绘制. 文中使用的地震波形数据来自中国地震局地球物理研究所国家测震台网数据备份中心(<http://www.seisdmc.ac.cn/>), 特此感谢.

参考文献

- 陈运泰, 杨智娴, 张勇, 刘超. 2013. 从汶川地震到芦山地震. 中国科学: 地球科学, 43: 1064–1072
- 何祥丽, 李海兵, 王焕, 张蕾, 孙知明, 司家亮. 2020. 蠕滑断裂带岩石组成和构造特征分析: 以龙门山灌县-安县断裂带为例. 岩石学报, 36: 3209–3224
- 黄雪源, 杨顶辉, 童平, 高原, 石玉涛, 吴昊. 2021. 首都圈地壳结构的拟波形层析成像研究. 中国科学: 地球科学, 51: 103–119
- 李志伟, 胥颐, 黄润秋, 郝天姚, 徐亚, 刘劲松, 刘建华. 2011. 龙门山地区的P波速度结构与汶川地震的深部构造特征. 中国科学: 地球科学, 41: 283–290
- 王二七, 孟庆任. 2008. 对龙门山中生代和新生代构造演化的讨论. 中国科学D辑: 地球科学, 38: 1221–1233
- 韦伟, 孙若昧, 石耀霖. 2010. 青藏高原东南缘地震层析成像及汶川地震成因探讨. 中国科学: 地球科学, 40: 831–839
- 张风雪, 吴庆举, 丁志峰. 2018. 青藏高原东部P波速度结构及其对高原隆升的启示. 科学通报, 63: 1949–1961
- 郑秀芬, 欧阳飏, 张东宁, 姚志祥, 梁建宏, 郑洁. 2009. “国家数字测震台网数据备份中心”技术系统建设及其对汶川大地震研究的数据支撑. 地球物理学报, 52: 1412–1417
- 张勇, 冯万鹏, 许力生, 周成虎, 陈运泰. 2008. 2008年汶川地震的时空破裂过程. 中国科学D辑: 地球科学, 38: 1186–1194
- 张勇, 许力生, 陈运泰. 2013. 芦山4·20地震破裂过程及其致灾特征初步分析. 地球物理学报, 56: 1408–1411
- Aki K, Lee W H K. 1976. Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model. *J Geophys Res*, 81: 4381–4399
- Ampère A M. 1820. Mémoire contenant l'application de la théorie. *Journal de l'École Poly technique*
- Bijwaard H, Spakman W, Engdahl E R. 1998. Closing the gap between regional and global travel time tomography. *J Geophys Res*, 103: 30055–30078
- Bozdag E, Peter D, Lefebvre M, Komatitsch D, Tromp J, Hill J, Podhorszki N, Pugmire D. 2016. Global adjoint tomography: first-

- generation model. *Geophys J Int*, 207: 1739–1766
- Brenier Y. 1991. Polar factorization and monotone rearrangement of vector-valued functions. *Comm Pure Appl Math*, 44: 375–417
- Brittany D F. 2009. Numerical methods for two second order elliptic equations. Dissertation for Master's Degree. Vancouver: Simon Fraser University
- Burchfiel B C, Chen Z, Liu Y, Royden L H. 1995. Tectonics of the Longmen Shan and adjacent regions, central China. *Int Geol Rev*, 37: 661–735
- Chen M, Niu F L, Liu Q, Tromp J, Zheng X F. 2015. Multiparameter adjoint tomography of the crust and upper mantle beneath East Asia: 1. Model construction and comparisons. *J Geophys Res-Solid Earth*, 120: 1762–1786
- Copley A, Mckenzie D. 2007. Models of crustal flow in the India-Asia collision zone. *Geophys J Int*, 169: 683–698
- Dahlen F A, Hung S H, Nolet G. 2000. Fréchet kernels for finite-frequency traveltimes-I. Theory. *Geophys J Int*, 141: 157–174
- Dong X P, Yang D H. 2020. Reanimating lithosphere of North China: New insights from full-waveform seismic tomography. *Sci Bull*, 65: 775–777
- Dong X P, Yang D H, Niu F L. 2019. Passive adjoint tomography of the crustal and upper mantle beneath eastern Tibet with a W2-norm misfit function. *Geophys Res Lett*, 46: 12986–12995
- Engquist B, Froese B D. 2014. Application of the Wasserstein metric to seismic signals. *Commun Math Sci*, 12: 979–988
- Fichtner A, Kennett B L N, Igel H, Bunge H P. 2009. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophys J Int*, 179: 1703–1725
- Fichtner A, Saygin E, Taymaz T, Cupillard P, Capdeville Y, Trampert J. 2013. The deep structure of the north Anatolian fault zone. *Earth Planet Sci Lett*, 373: 109–117
- Fichtner A, Trampert J. 2011. Hessian kernels of seismic data functionals based upon adjoint techniques. *Geophys J Int*, 185: 775–798
- Fichtner A, Villaseñor A. 2015. Crust and upper mantle of the western Mediterranean-Constraints from full-waveform inversion. *Earth Planet Sci Lett*, 428: 52–62
- Figalli A. 2017. The Monge-Ampère Equation and Its Application. Zürich: European Mathematical Society (EMS). 200
- Frisch U, Matarrese S, Mohayae R, Sobolevski A. 2002. A reconstruction of the initial conditions of the universe by optimal mass transportation. *Nature*, 417: 260–262
- Gu X F, Luo F, Sun J, Yau S T. 2016. Variational principles for Minkowski type problems, discrete optimal transport, and discrete Monge-Ampère equations. *Asian J Math*, 20: 383–398
- Huang H, Yao H, van der Hilst R D. 2010. Radial anisotropy in the crust of SE Tibet and SW China from ambient noise interferometry. *Geophys Res Lett*, 37: L21310
- Hung S, Shen Y, Chiao L. 2004. Imaging seismic velocity structure beneath the Iceland hot spot: A finite frequency approach. *J Geophys Res*, 109: B08305
- Jing H, Yang D H, Wu H. 2016. Dynamic inversion method based on the time-staggered stereo-modeling scheme and its acceleration. *Geophys J Int*, 207: 1675–1687
- Karato S I, Jung H, Katayama I, Skemer P. 2008. Geodynamic significance of seismic anisotropy of the upper mantle: New insights from laboratory studies. *Annu Rev Earth Planet Sci*, 36: 59–95
- Komatitsch D, Tromp J. 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophys J Int*, 139: 806–822
- Lailly P. 1983. The seismic inverse problem as a sequence of before stack migrations. In: Proceedings of the Conference on Inverse Scattering Theory and Application. Society for Industrial and Applied Mathematics. Philadelphia, 206–220
- Lei J, Zhao D. 2009. Structural heterogeneity of the Longmenshan fault zone and the mechanism of the 2008 Wenchuan earthquake (Ms 8.0). *Geochem Geophys Geosyst*, 10: Q10010
- Maggi A, Tape C, Chen M, Chao D, Tromp J. 2009. An automated time-window selection algorithm for seismic tomography. *Geophys J Int*, 178: 257–281
- Marquering H, Dahlen F A, Nolet G. 1999. Three-dimensional sensitivity kernels for finite-frequency traveltimes: The banana-doughnut paradox. *Geophys J Int*, 137: 805–815
- Métivier L, Brossier R, Mérogot Q, Oudet E, Virieux J. 2016. Increasing the robustness and applicability of full-waveform inversion: An optimal transport distance strategy. *Leading Edge*, 35: 1060–1067
- Monge G. 1784. Sur le calcul intégral des équations aux différences partielles. Mémoires del'Académie des Sciences
- Nocedal J. 1980. Updating quasi-Newton matrices with limited storage. *Math Comp*, 35: 773
- Paul T, Xu Z, Francoise R, Bertrand M, Nicolas A, Gerard W, Yang J. 2001. Oblique stepwise rise and growth of the Tibet plateau. *Science*, 294: 1671–1677
- Pei S P, Niu F L, Yehuda B Z, Sun Q, Liu Y B, Xue X T, Su J R, Shao Z G. 2019. Seismic velocity reduction and accelerated recovery due to earthquakes on the Longmenshan fault. *Nat Geosci*, 12: 387–392
- Pratt R G, Worthington M H. 1990. Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method. *Geophys Prospect*, 38: 287–310
- Rawlinson N, Davies D R, Pilia S. 2017. The mechanisms underpinning Cenozoic intraplate volcanism in eastern Australia: Insights from seismic tomography and geodynamic modeling. *Geophys Res Lett*,

- 44: 9681–9690
- Rawlinson N, Pozgay S, Fishwick S. 2010. Seismic tomography: A window into deep Earth. *Phys Earth Planet Inter*, 178: 101–135
- Royden L H, Burchfiel B C, King R W, Wang E, Chen Z, Shen F, Liu Y. 1997. Surface deformation and lower crustal flow in eastern Tibet. *Science*, 276: 788–790
- Savage M K. 1999. Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting? *Rev Geophys*, 37: 65–106
- Shapiro N M, Ritzwoller M H, Molnar P, Levin V. 2004. Thinning and flow of Tibetan crust constrained by seismic anisotropy. *Science*, 305: 233–236
- Soheil K, Park S, Thorpe M, Slepcev D, Rohde G K. 2017. Optimal mass transport signal processing and machine-learning applications. *IEEE Signal Process Magaz*, 34: 43–59
- Soheil K, Zou Y, Gustavo K R. 2016. Sliced wasserstein kernels for probability distributions. *IEEE Signal Process Magaz*, doi: 10.1109/CVPR.2016.568
- Tang W Q, Liu Y P, Chen Z L, Zhang Q Z, Zhao J X. 2004. GPS study on longmenshan fault zone. *J Geodesy Geodyna*, 24: 57–59
- Tao K, Grand S P, Niu F L. 2017. Full-waveform inversion of triplicated data using a normalized-correlation-coefficient-based misfit function. *Geophys J Int*, 210: 1517–1524
- Tao K, Grand S P, Niu F L. 2018. Seismic structure of the upper mantle beneath eastern Asia from full waveform seismic tomography. *Geochem Geophys Geosyst*, 19: 2732–2763
- Tape C, Liu Q, Maggi A, Tromp J. 2009. Adjoint tomography of the southern California crust. *Science*, 325: 988–992
- Tape C, Liu Q, Tromp J. 2007. Finite-frequency tomography using adjoint methods-Methodology and examples using membrane surface waves. *Geophys J Int*, 168: 1105–1129
- Tapponnier P, Xu Z, Roger F. 2001. Oblique stepwise rise and growth of the Tibet Plateau. *Science*, 294: 1671–1678
- Tarantola A. 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49: 1259–1266
- Tong P, Zhao D, Yang D. 2011. Tomography of the 1995 Kobe earthquake area: Comparison of finite-frequency and ray approaches. *Geophys J Int*, 187: 278–302
- Tromp J. 2020. Seismic wavefield imaging of Earth's interior across scales. *Nat Rev Earth Environ*, 1: 40–53
- Tromp J, Komatitsch D, Liu Q Y. 2008. Spectral-element and adjoint methods in seismology. *Commun Comput Phys*, 3: 1–32
- Tromp J, Tape C, Liu Q Y. 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophys J Int*, 160: 195–216
- Villani C. 2003. Topics in Optimal Transportation. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence
- Virieux J, Operto S. 2009. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74: WCC1–WCC26
- Wang J, Yang D H, Jing H, Wu H. 2019. Full waveform inversion based on the ensemble Kalman filter method using uniform sampling without replacement. *Sci Bull*, 64: 321–330
- Wang X, Chen L, Ai Y S, Xu T, Jiang M, Ling Y, Gao Y. 2018. Crustal structure and deformation beneath eastern and northeastern Tibet revealed by P-wave receiver functions. *Earth Planet Sci Lett*, 497: 69–79
- Weiss T, Siegesmund S, Rabbel W, Bohlen T, Pohl M. 1999. Seismic velocities and anisotropy of the lower continental crust: A review. *Pure Appl Geophys*, 156: 97–122
- Wessel P, Smith W H F, Scharroo R, Luis J, Wobbe F. 2013. Generic mapping tools: Improved version released. *Eos Trans AGU*, 94: 409–410
- Wu H, Chen J, Huang X, Yang D. 2018. A new earthquake location method based on the waveform inversion. *Commun Comput Phys*, 23: 118–141
- Yang Y, Engquist B, Sun J, Hamfeldt B F. 2018. Application of optimal transport and the quadratic Wasserstein metric to full-waveform inversion. *Geophysics*, 83: R43–R62
- Zhang G Q. 2011. Lectures of Variational Methods. Beijing: Higher Education Press (HEP)
- Zhu H, Bozdağ E, Peter D, Tromp J. 2012. Structure of the European upper mantle revealed by adjoint tomography. *Nat Geosci*, 5: 493–498
- Zhu H, Bozdağ E, Tromp J. 2015. Seismic structure of the European upper mantle based on adjoint tomography. *Geophys J Int*, 201: 18–52

(责任编辑: 陈凌)