Higher-order source-wavefield reconstruction for reverse time migration from stored values in a boundary strip just one point wide

Mulder, Wim

DOI
10.1190/GEO2017-0182.1

Publication date
2018

Document Version
Final published version

Published in
Geophysics

GEOPHYSICS, VOL. 83, NO. 1 (JANUARY-FEBRUARY 2018); P. T31–T38, 4 FIGS., 3 TABLES.
10.1190/GEO2017-0182.1

Higher-order source-wavefield reconstruction for reverse time migration from stored values in a boundary strip just one point wide

Wim A. Mulder[1]

ABSTRACT

One way to deal with the storage problem for the forward source wavefield in reverse time migration and full-waveform inversion is the reconstruction of that wavefield during reverse time stepping along with the receiver wavefield. Apart from the final states of the source wavefield, this requires a strip of boundary values for the whole time range in the presence of absorbing boundaries. The width of the stored boundary strip, positioned in between the interior domain of interest and the absorbing boundary region, usually equals about half that of the finite-difference stencil. The required storage in 3D with high frequencies can still lead to a decrease in computational efficiency, despite the substantial reduction in data volume compared with storing the source wavefields at all or at appropriately subsampled time steps. We have developed a method that requires a boundary strip with a width of just one point and has a negligible loss of accuracy. Stored boundary values over time enable the computation of the second and higher even spatial derivatives normal to the boundary, which together with extrapolation from the interior provides stability and accuracy. Numerical tests show that the use of only the boundary values provides at most fourth-order accuracy for the reconstruction error in the source wavefield. The use of higher even normal derivatives, reconstructed from the stored boundary values, allows for higher orders as numerical examples up to order 26 demonstrate. Subsampling in time is feasible with high-order interpolation and provides even more storage reduction but at a higher computational cost.

INTRODUCTION

Subsurface imaging from seismic data with reverse time migration (RTM) or full-waveform inversion (FWI) requires the correlation of forward-propagated source wavefields and backward-propagated reverse time receiver wavefields (Lailly, 1983; Tarantola, 1984a, 1948b; Whitmore and Lines, 1986). Three-dimensional RTM typically uses higher frequencies than FWI, and storage of source wavefields on disk or solid-state drives or even in main memory may lead to performance bottlenecks, in particular on many-core or GPU hardware. Migration in the frequency domain avoids the storage problem and outperforms migration in the time domain in two dimensions (Marfurt and Shin, 1989; Pratt, 1990; Østmo et al., 2002; Mulder and Plessix, 2004) but not yet in three dimensions (Riyanti et al., 2006; Plessix, 2009; Wang et al., 2010, 2011; Knibbe et al., 2014; Amestoy et al., 2016). For 3D applications, there are several ways to reduce or circumvent the storage of the source wavefields (Dussaud et al., 2008; Nguyen and McMechan, 2015). Data compression with subsampling or with wavelet or Fourier transforms is one approach (Araya-Polo et al., 2011; Sun and Fu, 2013; Knibbe et al., 2014). Checkpointing, storing the wavefield at selected times to recompute the wavefields in small time intervals when needed for correlation with the reverse time computations, is another approach (Griewank, 1992; Symes, 2007). Reliance on the time reversibility of the wave equation requires no storage at all with Dirichlet or Neumann boundary conditions, at the expense of recomputing the source wavefield in reverse time, along with the receiver wavefield. Clapp (2009) proposes reflecting boundaries with nearby random velocity perturbations that generate incoherent backscatter. Time reversal of the source wavefield is exact, and the backscattered waves do not constructively interfere into a subsurface image, although some noise will remain.

With absorbing boundaries, time reversal for source-wavefield reconstruction only works if a strip of boundary values just inside

Manuscript received by the Editor 21 March 2017; revised manuscript received 26 July 2017; published ahead of production 10 October 2017; published online 01 December 2017.
[1]Shell Global Solutions International B.V., Amsterdam, The Netherlands and Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Geoscience & Engineering, Delft, The Netherlands.

the absorbing region with a width of about half the finite-difference stencil is stored (e.g., Dussaud et al., 2008). For typical industrial problems, storage of just these boundary values can still reduce the computational performance, in particular for RTM at higher frequencies. Subsampling in time (Knibbe et al., 2014; Yang et al., 2016) offers a further reduction but can still be insufficient for large-scale many-core or GPU applications. For that reason, several authors have worked on a further size reduction of the stored boundary strip. Vasmel and Robertsson (2016) use a representation theorem (Wapenaar and Fokkema, 2006) and the first-order form of the acoustic wave equation to reduce the width of the stored boundary strip to one point. McGarry et al. (2010) and Feng and Wang (2012) accomplish a width of one point by lowering the order of the scheme toward the boundary strip, leading to some loss of accuracy. Tan and Huang (2014) maintain accuracy using the Lax and Wendroff (1960) idea to translate temporal into spatial derivatives via the wave equation on the boundary together with extrapolation from the interior. They only need a strip of one point wide for lower-order schemes. Stability with higher orders requires additional layers. They consider the first-order formulation of the acoustic wave equation. Liu et al. (2015) consider its second-order formulation and focus on the second spatial derivative of the pressure in the direction normal to the boundary. They construct a dedicated discrete second-derivative operator on the boundary based on exterior and interior values. The partial action of the operator on the exterior values is stored during the forward wavefield computations. During the reverse time computation, its action on the interior values is added to find the normal second derivative on the boundary. With the partial action on the exterior values, they also construct second-derivative operators for the interior points close to the boundary of which stencils would normally require exterior points. Storage requirements for the partial action are the same as a one-point-wide strip. Here, I propose a simpler method based on the Lax-Wendroff idea that maintains accuracy, requires storage of just one point per boundary per time step, and makes subsampling in time straightforward.

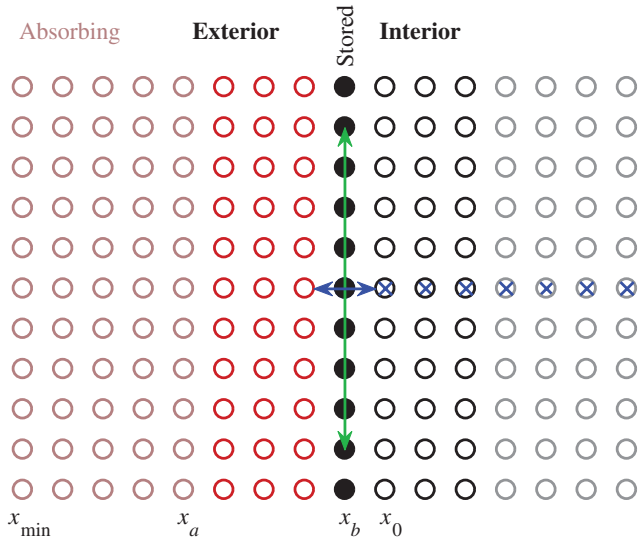

Figure 1. Left boundary: The forward solution is stored at the positions marked by the filled black circles. This example refers to an eighth-order scheme ($M = 8$). The application of equation 2 with $M_t = 1$ to the nine points covered by the green arrow and nine points out of the plane and three time levels provides the second derivative in the direction normal to the boundary, indicated by the blue arrow. Together with $n_i = 7$ interior values, marked by blue crosses, we can extrapolate to $(1/2)M - 1 = 3$ exterior values, drawn as red open circles. The second derivates in x at the $(1/2)M - 1 = 3$ interior points marked as black open circles require $(1/2)M = 4$ points toward the left and right, with some on the left obtained by extrapolation.

Note that the idea of Lax and Wendroff (1960) to compute temporal from spatial derivatives or the other way around, using the governing partial differential equations, also goes by the name of the Cauchy-Kovalevskaya procedure (von Kowalevsky, 1875) and figures in many methods. Examples relevant to seismic modeling include higher-order time stepping (Dablain, 1986; Käser and Dumbser, 2006) and the modified equation approach (Shubin and Bell, 1987).

This paper explores the stability and accuracy of our scheme in an experimental way. The next section describes the method. Subsampling can further reduce the storage requirements but calls for high-order interpolation. Yang et al. (2016) compare three interpolation schemes: Fourier, tapered sinc, and Lagrange polynomial interpolation. They conclude that the discrete Fourier transform provides the most accurate results, but the other two are more efficient. Here, I will only consider barycentric Lagrange interpolation (Berrut and Trefethen, 2004) because it requires less memory than Fourier interpolation. Two-dimensional numerical experiments reveal the stability properties as well as the convergence behavior in terms of the error in the reconstructed source wavefield as a function of grid spacing without and with subsampling.

METHOD

Stabilized extrapolation

To explain the idea, the constant-density acoustic wave equation will suffice:

$$\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = f + \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2}. \tag{1}$$

The pressure $p(t, \mathbf{x})$ depends on time t and position $\mathbf{x} = (x, y, z)$. The velocity is denoted by $c(\mathbf{x})$ and the source term by $f(t, \mathbf{x})$. An equidistant grid has points $x = x_0 + jh_x$ for $j = 0, \ldots, N_x - 1$, and similarly for y and z. Standard second-order time stepping is combined with higher-order second-derivative operators in space that have a stencil width of $M + 1$ points in each coordinate. Expressions for the weights and stability limits can be found elsewhere (Fornberg, 1988; Zhebel et al., 2014). In the context of RTM, all boundaries are assumed to be absorbing. To that end, the domain is enlarged on all sides.

I consider just one boundary, for instance, the lower one at $x = x_0$. The computational domain is enlarged. Around x_0, I introduce $x_{\min} < x_a < x_b \le x_0$, each on a grid point. The absorbing boundary covers the interval $[x_{\min}, x_a]$, as sketched in Figure 1. The point x_a could be closer to x_b, depending on the type of absorbing boundary condition, but I have stayed on the conservative side. For the tests in the next section, sponge boundaries (Cerjan et al., 1985) are applied, which are easy to code up but are not recommended for serious applications. They usually require more points than suggested by Figure 1. Solution values located at x_b are stored during the forward simulation. The velocity is extended outside its original domain by constant extrapolation, making it constant along $x \le x_0$, but not necessarily in the y- and z-directions.

For the reverse time reconstruction of the source wavefield, we can consider the stored boundary values at x_b together with extrapolation from the interior to produce pressure values at positions $x_j < x_b$ needed for the finite differencing in x or, equivalently, we can apply a one-sided difference scheme. Unfortunately, that approach quickly leads to instability when increasing the spatial order of the finite-difference scheme.

Stability can be recovered by computing the even spatial derivatives in the direction normal to the boundary from stored pressure boundary values $p(t, x_b, y, z)$ using the wave equation and assuming sufficient smoothness of the solution. Elimination of $\partial^2 p/\partial x^2$ from equation 1 away from the source and taking its kth power produces

$$\frac{\partial^{2k} p}{\partial x^{2k}}(t, x_b, y, z) = \left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right]^k p(t, x_b, y, z),$$
$$k = 0, 1, \ldots, M_t. \tag{2}$$

The kth power of the differential operator in square brackets means that it should be applied k times. At fixed y and z, these $M_t + 1$ values for the even normal derivatives at x_b together with an additional n_i interior points, located at $x = x_b + jh_x$ ($j = 1, \ldots, n_i$), define an interpolating polynomial of the form $p(x) = \sum_k b_k (x - x_b)^k/k!$ with $M_t + 1$ coefficients $b_0, b_2, \ldots, b_{2M_t}$ and n_i coefficients $b_1, b_3, \ldots, b_{2M_t-1}$ together with $b_{2M_t+1}, b_{2M_t+2}, \ldots$. Note that if $2n_i - 1 < 2M_t - 1$ or $n_i < M_t$, there are not enough interior points to define all derivatives up to $2M_t$. A natural choice for n_i is to match the stencil width, $M + 1$, of the discrete spatial difference operator of order M, implying $n_i + M_t + 1 = M + 1$. The polynomial $p(x)$ determines extrapolated solution values at points outside the domain, at $x = x_b + jh_x$ with $j = -(\frac{1}{2}M - 1), \ldots, -1$, providing sufficient input to evaluate the standard discrete Laplace operator, requiring $\frac{1}{2}M$ neighbors, at the interior points $x = x_b + jh_x$, with $j = 1, \ldots, \frac{1}{2}M - 1$. The reverse time computation now only requires values for $x \geq x_b$, but some extra points are needed in t, y, and z to be able to evaluate the derivatives in equation 2. The same approach should be taken on the other boundary planes.

The combination of extrapolation and standard finite differencing produces a modified, one-sided finite-difference scheme for interior points near the boundary. For $M_t = 0$, such a scheme is known to become unstable at higher orders. Increasing M_t will recover stability as the next section will show. Note that storing $\frac{1}{2}M$ boundary points yields an exact reconstruction of the source wavefield but still involves a fair amount of data in 3D for higher orders.

The main cost of the scheme is a full wavefield simulation, running backward from the final states. In 3D, the repeated application of the second derivates in space and time in a boundary plane, expressed by equation 2, is small. In addition, we can omit the computations for the exterior points and the absorbing boundary strips. It will, however, interrupt the compute flow in a parallel code. Also, the repeated application of the operator on the right-hand side of equation 2 toward the corners of the domain requires additional points, resulting in a larger computational domain during the forward modeling.

Subsampling

The Courant-Friedrichs-Lewy (CFL) condition (Courant et al., 1928) limits the numerical time step Δt. Zhebel et al. (2014), e.g., provide expressions for arbitrary spatial orders with second-order time stepping. The Nyquist-Shannon criterion dictates a maximum sampling interval of $\Delta t_{\max} = 1/(2f_{\max})$, where $f_{\max}$ is the maximum frequency in the data. Usually, $\Delta t_{\max}$ is larger than the time step Δt, allowing for subsampling with $\Delta t_{\text{sub}} = n_{\text{sub}}\Delta t$, with integer $n_{\text{sub}} = \text{ceil}(\Delta t_{\max}/\Delta t) \geq 1$.

If $n_{\text{sub}} > 1$ and we have to take the discrete second time-derivative M_t times, the subsampled data should be interpolated with a sufficiently high order before doing so. If interpolation with a polynomial degree $2M_t$ over $2M_t + 1$ points is applied, the result will be piecewise constant in between the subsampled points. This suggests that the interpolation should be at least of degree $2M_t + n_{\text{sub}}$, using values in $2M_t + n_{\text{sub}} + 1$ points. In general, we can increase (or decrease) this number with M_i and interpolate over a total of $2M_t + n_{\text{sub}} + M_i + 1$ consecutive points, bounded from below by 2 and from above by the total available number of samples. After interpolation to $2M_t + 1$ time values, symmetrically arranged around the current time value, and spaced by the time step of the time-stepping scheme, the application of equation 2 as well as the rest of the procedure can follow in the same way as before.

Barycentric Lagrange interpolation for equidistant data p_j at positions t_j, $j = 0, 1, \ldots, n - 1$, is of the form (Berrut and Trefethen, 2004):

$$p(t) = \frac{\sum_{j=0}^{n-1} \frac{w_j}{t - t_j} p_j}{\sum_{j=0}^{n-1} \frac{w_j}{t - t_j}}, \tag{3a}$$

with weights

$$w_j = (-1)^j \binom{n-1}{j}, \quad j = 0, 1, \ldots, n - 1. \tag{3b}$$

Setting $t_j = t_k + (j - k)n_{\text{sub}}\Delta t$ and $t = t_k + \xi n_{\text{sub}}\Delta t$, with a reference value t_k centered among the data points such that $|\xi| \leq 1/2$, if the start and end time permit, provides

$$p(t) = \sum_{j=0}^{n-1} W_j p_j, \tag{4a}$$

with weights

$$W_j = \frac{\overline{w}_j}{\sum_{j=0}^{n-1} \overline{w}_j}, \qquad \overline{w}_j = \frac{w_j}{\xi - (j - k)}. \tag{4b}$$

Ordering the weights W_j from small to large can further reduce round-off errors during the summation. I have used this sorted summation instead of the relation:

$$\sum_{j=0}^{n-1} \overline{w}_j = -\frac{(n-1)!\Gamma(-(k + \xi))}{\Gamma(n - (k + \xi))}. \tag{5}$$

Note that ξ is a fraction of the form $\xi = \ell/n_{\text{sub}}$ with an integer value of ℓ and that no interpolation is required if ℓ is a multiple of n_{sub}, that is, if ξ itself is an integer.

RESULTS

Stability and accuracy

A series of 2D tests were performed to study the stability and convergence of our scheme for a range of spatial orders M and using only one boundary value with $M_t = 0$, meaning just the pressure value, or using the second ($M_t = 1$) and higher ($M_t > 1$) even normal derivatives computed from the stored values on the boundaries. Figure 2a displays the velocity model for the numerical tests taken from Mulder (2001) and mapped to a finite-difference grid after some smoothing. The source at $x_s = 3000$ m and $z_s = 0$ m has a 15-Hz Ricker wavelet that peaks at zero time. The forward source wavefield is propagated to 2 s at 90% of the maximum time step dictated by the CFL condition. Then, the time-reversed computation starts, back to a time of 1 s. The resulting wavefield is compared with a snapshot at the same time, saved for that purpose during the forward modeling. Figure 2b shows an example of the forward wavefield at 1 s for a set of parameters, and Figure 2c shows its normalized reconstruction error, defined as the difference between the reconstructed and true source wavefield at the same time, divided by the maximum amplitude of the latter.

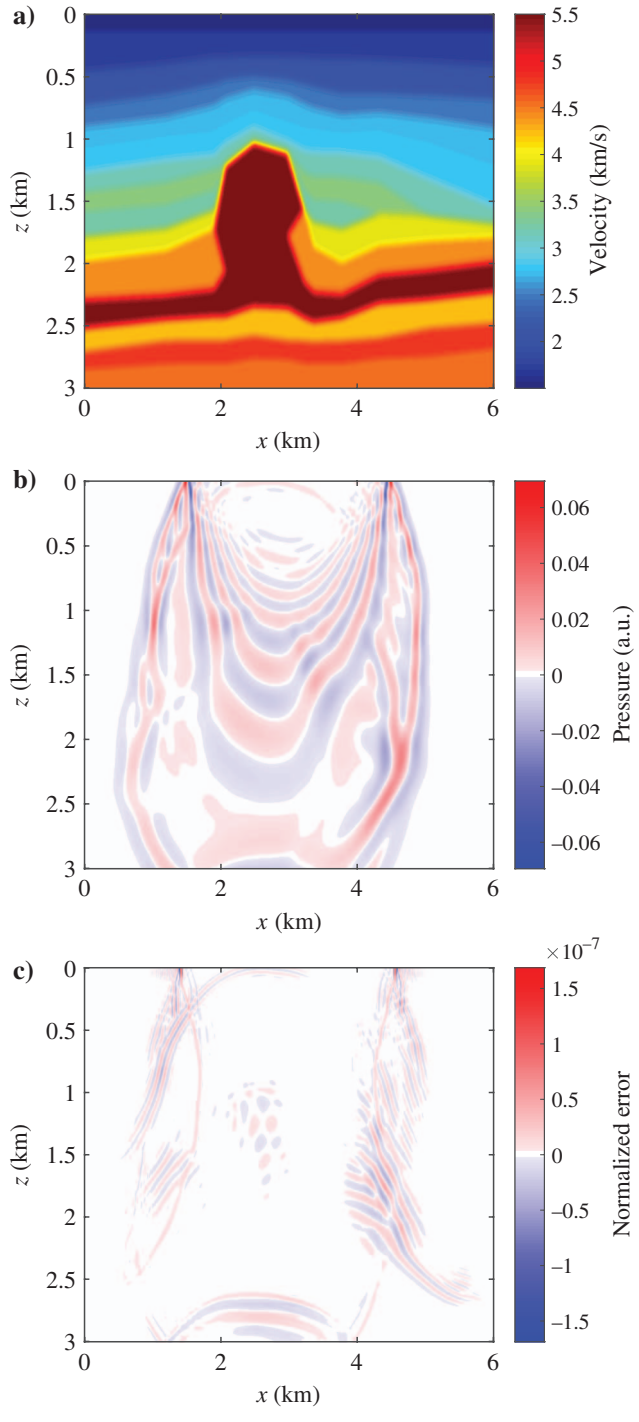

Figure 2. (a) Two-dimensional velocity model, (b) source wavefield at 1 s, and (c) its normalized reconstruction error at the same time, for a scheme with order $M = 8$, $M_t = 1$, and $n_i = 7$ and a 5 m grid spacing.

Table 1 lists parameters for which the method became unstable. The suggested trend is instability for $n_i \geq 6 + 3M_t$. Assuming that the boundary order is at most the interior spatial order or less, the instability affects orders $M \geq n_i + M_t = 6 + 4M_t$. The scheme should then be stable for

$$M_t \geq \max(0, \text{ceil}((M-5)/4)). \tag{6}$$

Table 2 lists measured maximum and root-mean-square (rms) normalized reconstruction errors at 1 s.

With the second-order scheme ($M = 2$), one boundary point already suffices to recover the source wavefield within machine precision. For higher orders, various combinations of even spatial order M, up to $M_t + 1$ even normal boundary derivatives, and n_i interior points were considered. The choice $n_i < M_t$ is allowed but is not sensible because there are not enough interior points to determine all odd derivatives up to degree $2M_t$. Choosing $n_i + M_t < M$ will lower and $n_i + M_t > M$ will raise the order of the boundary scheme. The convergence behavior of the normalized reconstruction error ε as a function of grid spacing h is $\varepsilon \propto h^q$, where the power q can be estimated by a linear fit to a log-log plot.

To illustrate the effect of the parameters, Table 2 explores a lengthy range of options for interior order $M = 4$. The natural choice is $M_t = 0$ and $n_i = M - M_t = 4$, meaning that just the boundary values suffice. For other choices, the estimated power q becomes larger or smaller. Note that measuring q from the maximum error produces more erratic values than from the rms error. The results for $M_t = 1$ and $n_i = 2$ are the same as those for $n_i = 1$ because for $n_i = 2$, the extrapolation weight for the interior point farthest from the boundary happens to be zero. The same is true with $M = 4$ and $M_t = 2$.

For higher orders, only a subset of the results is presented up to order $M = 26$ with M_t taken at the lower bound of equation 6 and

Table 1. Values of n_i as a function of M_t for which the method appeared to be unstable in the 2D test and the affected orders M.

M_t	0	1	2	3	4	5
$n_i \geq$	6	9	12	15	18	21
$M \geq$	6	10	14	18	22	26

$n_i = M - M_t$. An exception is $M = 6$, $M_t = 0$, and $n_i = 5$, which more or less has the correct convergence behavior and also has a large error at a grid spacing of 20 m.

At higher orders, it becomes more difficult to estimate the scheme's order of accuracy because of the smaller range of grid spacings where convergence can be estimated, as illustrated in Figure 3 for order $M = 24$. If the spacing is too large, the region of asymptotic convergence is not yet reached, whereas for too small a spacing, round-off errors start to dominate. For that reason, Table 2 lists estimated powers only up to $M = 12$.

Subsampling

Subsampling in time allows for a further reduction of the stored data volume. With the use of a Ricker wavelet, it is convenient to define $f_{\max}$ as the upper frequency in which the wavelet's amplitude has dropped below a fraction $\alpha_{\max}$ of the peak frequency f_{peak}. The convergence study used an impractically small value of $\alpha_{\max} = 10^{-14}$ and a time step Δt at 90% of its maximum value allowed by the CFL condition. In this way, the subsampling ratio, $n_{\text{sub}} = \text{ceil}(\Delta t_{\text{sub}}/\Delta t)$, depends on the grid spacing.

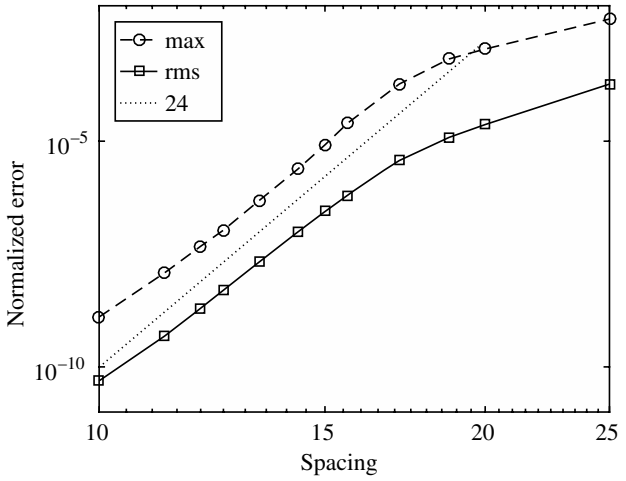

Figure 3. Maximum and rms normalized reconstruction errors as a function of grid spacing for order $M = 24$ with $M_t = 5$ and $n_i = 19$, showing how small the interval is where the expected error slope marked by the dotted line is attained.

Table 2: Maximum and rms errors in the reconstructed source wavefield, divided by its maximum amplitude over the domain, as a function of grid spacing $h_x = h_z = h$ for a finite-difference scheme of order M, using extrapolation with even normal derivatives up to $2M_t$ on the boundary and n_i interior points. Results for orders higher than 6 only include the natural choice, $n_i = M - M_t$. The observed convergence behaves as h^q with order q.

			Maximum error				rms error			
M	M_t	n_i	$h = 20$ m	10 m	5 m	q	$h = 20$ m	10 m	5 m	q
2	0	0	—	—	—	—	—	—	—	—
4	0	0	$3.6\,10^{-2}$	$1.4\,10^{-2}$	$7.5\,10^{-3}$	1.1	$2.3\,10^{-3}$	$9.9\,10^{-4}$	$7.0\,10^{-4}$	0.9
	0	1	$3.2\,10^{-2}$	$7.2\,10^{-3}$	$1.7\,10^{-3}$	2.1	$1.8\,10^{-3}$	$3.7\,10^{-4}$	$1.1\,10^{-4}$	2.0
	0	2	$3.4\,10^{-2}$	$3.8\,10^{-3}$	$6.5\,10^{-4}$	2.9	$1.8\,10^{-3}$	$2.0\,10^{-4}$	$3.6\,10^{-5}$	2.8
	0	3	$4.3\,10^{-2}$	$2.7\,10^{-3}$	$2.2\,10^{-4}$	3.8	$2.3\,10^{-3}$	$1.3\,10^{-4}$	$1.1\,10^{-5}$	3.9
4	0	4	$7.1\,10^{-2}$	$2.1\,10^{-3}$	$1.5\,10^{-4}$	4.4	$3.7\,10^{-3}$	$9.7\,10^{-5}$	$4.3\,10^{-6}$	4.9
	0	5	0.17	$1.7\,10^{-3}$	$5.1\,10^{-5}$	5.8	$9.7\,10^{-3}$	$7.8\,10^{-5}$	$1.7\,10^{-6}$	6.2
4	1	0	$3.2\,10^{-2}$	$1.3\,10^{-2}$	$7.2\,10^{-3}$	1.1	$2.1\,10^{-3}$	$9.7\,10^{-4}$	$6.9\,10^{-4}$	0.8
	1	1,2	$2.7\,10^{-3}$	$1.8\,10^{-4}$	$1.2\,10^{-5}$	3.9	$1.6\,10^{-4}$	$9.1\,10^{-6}$	$7.3\,10^{-7}$	3.9
	1	3	$3.5\,10^{-3}$	$1.4\,10^{-4}$	$9.3\,10^{-6}$	4.3	$2.1\,10^{-4}$	$6.5\,10^{-6}$	$3.0\,10^{-7}$	4.7
	1	4	$4.6\,10^{-3}$	$1.1\,10^{-4}$	$3.3\,10^{-6}$	5.2	$3.0\,10^{-4}$	$5.1\,10^{-6}$	$1.1\,10^{-7}$	5.7
4	2	2,1	$7.3\,10^{-4}$	$1.7\,10^{-5}$	$4.9\,10^{-7}$	5.3	$4.7\,10^{-5}$	$7.5\,10^{-7}$	$1.6\,10^{-8}$	5.7
6	0	5	0.14	$1.6\,10^{-3}$	$4.3\,10^{-5}$	5.8	$7.6\,10^{-3}$	$6.8\,10^{-5}$	$1.3\,10^{-6}$	6.3
6	1	5	$2.9\,10^{-3}$	$7.6\,10^{-5}$	$1.2\,10^{-6}$	5.6	$1.6\,10^{-4}$	$3.2\,10^{-6}$	$4.2\,10^{-8}$	6.0
8	1	7	$8.8\,10^{-3}$	$4.0\,10^{-5}$	$1.7\,10^{-7}$	7.8	$5.1\,10^{-4}$	$1.7\,10^{-6}$	$3.4\,10^{-9}$	8.6
10	2	8	$9.4\,10^{-4}$	$2.5\,10^{-6}$	$2.3\,10^{-9}$	9.3	$5.5\,10^{-5}$	$1.2\,10^{-7}$	$4.3\,10^{-11}$	10
12	2	10	$2.2\,10^{-3}$	$1.1\,10^{-6}$	—	11	$1.6\,10^{-4}$	$5.1\,10^{-8}$	—	12
14	3	11	$4.3\,10^{-4}$	$1.1\,10^{-7}$	—	—	$2.7\,10^{-5}$	$5.2\,10^{-9}$	—	—
16	3	13	$1.3\,10^{-3}$	$5.1\,10^{-8}$	—	—	$7.4\,10^{-5}$	$2.4\,10^{-9}$	—	—
18	4	14	$3.5\,10^{-4}$	$1.2\,10^{-8}$	—	—	$1.7\,10^{-5}$	$4.1\,10^{-10}$	—	—
20	4	16	$1.7\,10^{-3}$	$6.9\,10^{-9}$	—	—	$4.0\,10^{-5}$	$2.6\,10^{-10}$	—	—
22	5	17	$5.2\,10^{-4}$	$1.7\,10^{-9}$	—	—	$1.1\,10^{-5}$	$6.4\,10^{-11}$	—	—
24	5	19	$1.1\,10^{-3}$	$1.3\,10^{-9}$	—	—	$2.3\,10^{-5}$	$4.9\,10^{-11}$	—	—
26	6	20	$1.5\,10^{-4}$	$3.7\,10^{-10}$	—	—	$4.2\,10^{-6}$	$1.6\,10^{-11}$	—	—

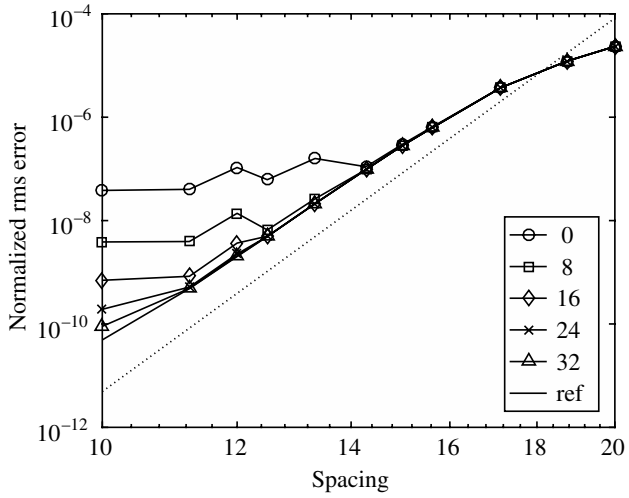

Figure 4. Normalized rms errors for $M = 24$ with subsampling based on $\alpha_{\max} = 10^{-14}$ and running at 90% of the maximum time step. The results correspond to values of M_i of 0, 8, 16, 24, and 32. The legend entry "ref" corresponds to the reference case without subsampling and is identical to the lower graph in Figure 3. The dotted line again marks the expected 24th-order convergence. The subsampled cases approach the reference case for increasing M_i.

Figure 4 shows the rms error as a function of grid spacing for a 24th-order scheme including subsampling. Again, the dotted line indicates the theoretical trend. The graph marked by "ref" is the same as the lower in Figure 3 without subsampling. The subsampling ratio ranges from $n_{\text{sub}} = 5$ for the coarser mesh with a grid spacing of $h = 20$ m to $n_{\text{sub}} = 9$ for the finest with $h = 10$ m. The various curves correspond to different choices of $M_i = 0$, 8, 16, 24, and 32.

A larger M_i clearly improves the result. However, the round-off errors in the barycentric Lagrange interpolation increase for a larger number of interpolation points (Berrut and Trefethen, 2004). Therefore, there is a limit to the best achievable accuracy. For instance, the result at $h = 10$ m for $M_i = 48$, not shown in Figure 4, is less accurate than that for $M_i = 32$.

Table 3 presents errors for the scheme with subsampling and barycentric Lagrange interpolation with $M_i = 0$ extra points, using $2M_t + n_{\text{sub}} + 1$ points in total. The results are nearly the same as those without subsampling in Table 2 at a 20-m grid spacing, for the maximum and the rms error. At a spacing of 10 or 5 m, the errors with subsampling quickly exceed those without subsampling but they are still small enough for practical computations. Figure 4 illustrates how a larger value of M_i can improve the accuracy.

DISCUSSION

The results highlight the numerical properties of our method and suggest stability properties that may serve as a starting point for mathematical proofs. As already remarked, convergence studies become difficult for schemes of very high order: The range of grid spacings between the point where the asymptotic regime starts and levels off because round-off errors becomes small. In practice, we perform computations with numerical parameters that bring us close to where the asymptotic regime of the convergence curve begins and not much further to errors that would be smaller than necessary. Moreover, carrying out the computations in single- rather than double-precision arithmetic may provide a quicker turnaround. On top of that, the spatial discretization of the wave equation and the time-stepping scheme will generate numerical dispersion errors — although there are ways to remove the latter (Stork, 2013; Anderson et al., 2015). Finally, we may be more interested in the effect of our scheme on the final migration image or full-waveform gradient update. I have neglected all that to gain some understanding of the reconstruction error per se. As a result, the listed parameters and their corresponding errors will tend to be too conservative for practical computations.

Our scheme increases the computational effort mainly because of the reverse time recomputation of the forward wavefield, which re-

Table 3: As Table 2, but with subsampling and barycentric interpolation over $2M_t + M_i + n_{\text{sub}} + 1$ points. The observed values correspond to $M_i = 0$. The actual value of the subsampling ratio, n_{sub}, is listed in brackets.

			Maximum error				rms error			
M	M_t	n_i	$h = 20$ m	10 m	5 m	q	$h = 20$ m	10 m	5 m	q
4	0	4	$7.0\,10^{-2}$ (4)	$2.1\,10^{-3}$ (7)	$1.5\,10^{-4}$ (14)	4.4	$3.7\,10^{-3}$ (4)	$9.6\,10^{-5}$ (7)	$4.6\,10^{-6}$ (14)	4.8
6	1	5	$2.8\,10^{-3}$ (4)	$6.8\,10^{-5}$ (8)	$3.0\,10^{-5}$ (16)	3.3	$1.5\,10^{-4}$ (4)	$3.3\,10^{-6}$ (8)	$2.1\,10^{-6}$ (16)	3.1
8	1	7	$8.7\,10^{-3}$ (4)	$4.8\,10^{-5}$ (8)	$1.6\,10^{-5}$ (16)	4.6	$5.1\,10^{-4}$ (4)	$2.1\,10^{-6}$ (8)	$1.1\,10^{-6}$ (16)	4.5
10	2	8	$9.5\,10^{-4}$ (4)	$6.0\,10^{-6}$ (8)	$1.5\,10^{-6}$ (16)	—	$5.4\,10^{-5}$ (4)	$3.0\,10^{-7}$ (8)	$1.2\,10^{-7}$ (16)	—
12	2	10	$2.2\,10^{-3}$ (5)	$1.6\,10^{-5}$ (9)	—	—	$1.6\,10^{-4}$ (5)	$7.8\,10^{-7}$ (9)	—	—
14	3	11	$4.3\,10^{-4}$ (5)	$6.2\,10^{-6}$ (9)	—	—	$2.6\,10^{-5}$ (5)	$3.0\,10^{-7}$ (9)	—	—
16	3	13	$1.3\,10^{-3}$ (5)	$5.8\,10^{-6}$ (9)	—	—	$7.2\,10^{-5}$ (5)	$2.5\,10^{-7}$ (9)	—	—
18	4	14	$3.4\,10^{-4}$ (5)	$2.4\,10^{-6}$ (9)	—	—	$1.6\,10^{-5}$ (5)	$1.1\,10^{-7}$ (9)	—	—
20	4	16	$1.7\,10^{-3}$ (5)	$2.2\,10^{-6}$ (9)	—	—	$3.9\,10^{-5}$ (5)	$9.4\,10^{-8}$ (9)	—	—
22	5	17	$5.0\,10^{-4}$ (5)	$9.8\,10^{-7}$ (9)	—	—	$1.0\,10^{-5}$ (5)	$4.3\,10^{-8}$ (9)	—	—
24	5	19	$1.1\,10^{-3}$ (5)	$8.7\,10^{-7}$ (9)	—	—	$2.4\,10^{-5}$ (5)	$3.8\,10^{-8}$ (9)	—	—
26	6	20	$1.5\,10^{-4}$ (5)	$4.1\,10^{-7}$ (9)	—	—	$4.1\,10^{-6}$ (5)	$1.8\,10^{-8}$ (9)	—	—

quires about the same number of operations as its original forward computation. The additional cost of interpolation and differencing for the determination of the normal derivatives is compensated by the fact that the entire absorbing boundary strip can be emplo ignored during the reverse time stepping.

As an aside, mass-lumped (Zhebel et al., 2014; Mulder and Shamasundar, 2016) or discontinuous Galerkin finite elements (Minisini et al., 2013; Modave et al., 2015) would also only require the equivalent of one point: The solution values on a plane or something close to it on a 3D unstructured mesh.

CONCLUSION

The computation of even spatial derivatives normal to a boundary from stored boundary values with a width of just one point in combination with extrapolation from interior points enables the reconstruction of a source wavefield during reverse time stepping with improved accuracy. Although only fourth-order accuracy is observed when only the boundary values are used, higher orders are obtained if the second and higher even normal derivatives are determined from the stored boundary values. The stability and accuracy were tested up to order 26 for a 2D example, without and with subsampling in time. Generalization to 3D should be straightforward.

ACKNOWLEDGMENT

The author is grateful to E. Duveneck for his comments on the manuscript.

REFERENCES

Amestoy, P., R. Brossier, A. Buttari, J.-Y. L'Excellent, T. Mary, L. Métivier, A. Miniussi, and S. Operto, 2016, Fast 3D frequency-domain full-waveform inversion with a parallel block low-rank multifrontal direct solver: Application to OBC data from the North Sea: Geophysics, **81**, no. 6, R363–R383, doi: 10.1190/geo2016-0052.1.

Anderson, J. E., V. Brytik, and G. Ayeni, 2015, Numerical temporal dispersion corrections for broadband temporal simulation, RTM and FWI: 85th Annual International Meeting, SEG, Expanded Abstracts, 1096–1100, doi: 10.1190/segam2015-5817144.1.

Araya-Polo, M., J. Cabezas, M. Hanzich, M. Pericas, F. Rubio, I. Gelado, M. Shafiq, E. Morancho, N. Navarro, E. Ayguade, J. M. Cela, and M. Valero, 2011, Assessing accelerator-based HPC reverse time migration: IEEE Transactions on Parallel and Distributed Systems, **22**, 147–162, doi: 10.1109/TPDS.2010.144.

Berrut, J.-P., and L. N. Trefethen, 2004, Barycentric Lagrange interpolation: SIAM Review, **46**, 501–517, doi: 10.1137/S0036144502417715.

Cerjan, C., D. Kosloff, R. Kosloff, and M. Reshef, 1985, A nonreflecting boundary condition for discrete acoustic and elastic wave equations: Geophysics, **50**, 705–708, doi: 10.1190/1.1441945.

Clapp, R. G., 2009, Reverse time migration with random boundaries: 79th Annual International Meeting, SEG, Expanded Abstracts, 2809–2813.

Courant, R., K. Friedrichs, and H. Lewy, 1928, Über die partiellen Differenzengleichungen der mathematischen Physik: Mathematische Annalen, **100**, 32–74, doi: 10.1007/BF01448839.

Dablain, M. A., 1986, The application of high-order differencing to the scalar wave equation: Geophysics, **51**, 54–66, doi: 10.1190/1.1442040.

Dussaud, E., W. W. Symes, P. Williamson, L. Lemaistre, P. Singer, B. Denel, and A. Cherrett, 2008, Computational strategies for reverse time migration: 78th Annual International Meeting, SEG, Expanded Abstracts, 2267–2271, doi: 10.1190/1.3059336.

Feng, B., and H. Wang, 2012, Reverse time migration with source wavefield reconstruction strategy: Journal of Geophysics and Engineering, **9**, 691–696, doi: 10.1088/1742-2132/9/6/691.

Fornberg, B., 1988, Generation of finite difference formulas on arbitrarily spaced grids: Mathematics of Computation, **51**, 699–706, doi: 10.1090/S0025-5718-1988-0935077-0.

Griewank, A., 1992, Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation: Optimization Methods and Software, **1**, 35–54, doi: 10.1080/10556789208805505.

Käser, M., and M. Dumbser, 2006, An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes — I. The two-dimensional isotropic case with external source terms: Geophysical Journal International, **166**, 855–877, doi: 10.1111/j.1365-246X.2006.03051.x.

Knibbe, H., W. A. Mulder, C. W. Oosterlee, and C. Vuik, 2014, Closing the performance gap between an iterative frequency-domain solver and an explicit time-domain scheme for 3D migration on parallel architectures: Geophysics, **79**, no. 2, S47–S61, doi: 10.1190/geo2013-0214.1.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Proceedings of the Conference on Inverse Scattering: Theory and Applications, SIAM, 206–220.

Lax, P., and B. Wendroff, 1960, Systems of conservation laws: Communications on Pure and Applied Mathematics, **13**, 217–237, doi: 10.1002/cpa.3160130205.

Liu, S., X. Li, W. Wang, and T. Zhu, 2015, Source wavefield reconstruction using a linear combination of the boundary wavefield in reverse time migration: Geophysics, **80**, no. 6, S203–S212, doi: 10.1190/geo2015-0109.1.

Marfurt, K. J., and C. S. Shin, 1989, The future of iterative modeling in geophysical exploration, *in* E. Eisner, ed., Supercomputers in seismic exploration: Pergamon Press, 203–228.

McGarry, R. G., J. A. Mahovsky, P. P. Moghaddam, D. S. Foltinek, and D. J. Eaton, 2010, Reverse-time depth migration with reduced memory requirements: U.S. Patent Application No. 12/231138.

Minisini, S., E. Zhebel, A. Kononov, and W. A. Mulder, 2013, Local time stepping with the discontinuous Galerkin method for wave propagation in 3D heterogeneous media: Geophysics, **78**, no. 3, T67–T77, doi: 10.1190/geo2012-0252.1.

Modave, A., A. St-Cyr, W. A. Mulder, and T. Warburton, 2015, A nodal discontinuous Galerkin method for reverse-time migration on GPU clusters: Geophysical Journal International, **203**, 1419–1435, doi: 10.1093/gji/ggv380.

Mulder, W. A., 2001, Higher-order mass-lumped finite elements for the wave equation: Journal of Computational Acoustics, **9**, 671–680, doi: 10.1142/S0218396X0100067X.

Mulder, W. A., and R.-É. Plessix, 2004, A comparison between one-way and two-way wave-equation migration: Geophysics, **69**, 1491–1504, doi: 10.1190/1.1836822.

Mulder, W. A., and R. J. Shamasundar, 2016, Performance of continuous mass-lumped tetrahedral elements for elastic wave propagation with and without global assembly: Geophysical Journal International, **207**, 414–421, doi: 10.1093/gji/ggw273.

Nguyen, B. D., and G. A. McMechan, 2015, Five ways to avoid storing source wavefield snapshots in 2D elastic prestack reverse time migration: Geophysics, **80**, no. 1, S1–S18, doi: 10.1190/geo2014-0014.1.

Østmo, S., W. A. Mulder, and R.-É. Plessix, 2002, Finite-difference iterative migration by linearized waveform inversion in the frequency domain: 72nd Annual International Meeting, SEG, Expanded Abstracts, 1384–1387, doi: 10.1190/1.1816917.

Plessix, R.-É., 2009, Three-dimensional frequency-domain full-waveform inversion with an iterative solver: Geophysics, **74**, no. 6, WCC149–WCC157, doi: 10.1190/1.3211198.

Pratt, R. G., 1990, Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging: Geophysics, **55**, 626–632, doi: 10.1190/1.1442874.

Riyanti, C., Y. Erlangga, R.-É. Plessix, W. A. Mulder, and C. V. C. Oosterlee, 2006, A new iterative solver for the time-harmonic wave equation: Geophysics, **71**, no. 5, E57–E63, doi: 10.1190/1.2231109.

Shubin, G. R., and J. B. Bell, 1987, A modified equation approach to constructing fourth order methods for acoustic wave propagation: SIAM Journal on Scientific and Statistical Computing, **8**, 135–151, doi: 10.1137/0908026.

Stork, C., 2013, Eliminating nearly all dispersion error from FD modeling and RTM with minimal cost increase: 75th Annual International Conference and Exhibition, EAGE, Extended Abstracts, Tu 11 07, doi: 10.3997/2214-4609.20130478.

Sun, W., and L.-Y. Fu, 2013, Two effective approaches to reduce data storage in reverse time migration: Computers & Geosciences, **56**, 69–75, doi: 10.1016/j.cageo.2013.03.013.

Symes, W. W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, no. 5, SM213–SM221, doi: 10.1190/1.2742686.

Tan, S., and L. Huang, 2014, Reducing the computer memory requirement for 3D reverse-time migration with a boundary-wavefield extrapolation method: Geophysics, **79**, no. 5, S185–S194, doi: 10.1190/geo2014-0075.1.

Tarantola, A., 1984a, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, doi: 10.1190/1.1441754.

Tarantola, A., 1984b, Linearized inversion of seismic reflection data: Geophysical Prospecting, **32**, 998–1015, doi: 10.1111/j.1365-2478.1984.tb00751.x.

Vasmel, M. L., and J. O. A. Robertsson, 2016, Exact wavefield reconstruction on finite-difference grids with minimal memory requirements:

78th Annual International Conference and Exhibition, EAGE, Extended Abstracts, We SBT5 12, doi: 10.3997/2214-4609.201601175.
von Kowalevsky, S., 1875, Zur Theorie der partiellen Differentialgleichung: Journal für die reine und angewandte Mathematik, **80**, 1–32, doi: 10.1515/crll.1875.80.1.
Wang, S., M. V. de Hoop, and J. Xia, 2010, Acoustic inverse scattering via Helmholtz operator factorization and optimization: Journal of Computational Physics, **229**, 8445–8462, doi: 10.1016/j.jcp.2010.07.027.
Wang, S., M. V. de Hoop, and J. Xia, 2011, On 3D modeling of seismic wave propagation via a structured parallel multifrontal direct Helmholtz solver: Geophysical Prospecting, **59**, 857–873, doi: 10.1111/j.1365-2478.2011.00982.x.
Wapenaar, K., and J. Fokkema, 2006, Green's function representations for seismic interferometry: Geophysics, **71**, no. 4, SI33–SI46, doi: 10.1190/1.2213955.
Whitmore, N. D., and L. R. Lines, 1986, Vertical seismic profiling depth migration of a salt dome flank: Geophysics, **51**, 1087–1109, doi: 10.1190/1.1442164.
Yang, P., R. Brossier, and J. Virieux, 2016, Wavefield reconstruction by interpolating significantly decimated boundaries: Geophysics, **81**, no. 5, T197–T209, doi: 10.1190/geo2015-0711.1.
Zhebel, E., S. Minisini, A. Kononov, and W. A. Mulder, 2014, A comparison of continuous mass-lumped finite elements with finite differences for 3-D wave propagation: Geophysical Prospecting, **62**, 1111–1125, doi: 10.1111/1365-2478.12138.