Geophysical Journal International

Geophys. J. Int. (2011) **184,** 1289–1303 doi: 10.1111/j.1365-246X.2010.04906.x

Theory of multisource crosstalk reduction by phase-encoded statics

G. T. Schuster, X. Wang, Y. Huang, W. Dai and C. Boonyasiriwat

King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. E-mail: chaiwoot.boonyasiriwat@kaust.edu.sa

Accepted 2010 November 29. Received 2010 November 28; in original form 2010 April 20

SUMMARY

Formulas are derived that relate the strength of the crosstalk noise in supergather migration images to the variance of time, amplitude and polarity shifts in encoding functions. A supergather migration image is computed by migrating an encoded supergather, where the supergather is formed by stacking a large number of encoded shot gathers. Analysis reveals that for temporal source static shifts in each shot gather, the crosstalk noise is exponentially reduced with increasing variance of the static shift and the square of source frequency. This is not too surprising because larger time shifts lead to less correlation between traces in different shot gathers, and so should tend to reduce the crosstalk noise. Analysis also reveals that combining both polarity and time statics is a superior encoding strategy compared to using either polarity statics or time statics alone.

Signal-to-noise (SNR) estimates show that $\text{SNR}^{\text{stand.}} = \sqrt{GS}$ for a standard migration image and $\text{SNR}^{\text{super}} = \sqrt{GI}$ for an image computed by migrating a phase-encoded supergather; here, G is the number of traces in a shot gather, I is the number of stacking iterations in the supergather and S is the number of encoded/blended shot gathers that comprise the supergather. If the supergather can be uniformly divided up into Q unique sub-supergathers, then the resulting SNR of the final image is $\text{SNR}^{\text{sub-super}} = \sqrt{QGI}$, which means that we can enhance image quality but at the expense of Q times more cost. The importance of these formulas is that they provide a precise understanding between different phase encoding strategies and image quality.

Finally, we show that iterative migration of phase-encoded supergathers is a special case of passive seismic interferometry. We suggest that the crosstalk noise formulas can be helpful in designing optimal strategies for passive seismic interferometry and efficient extraction of Green's functions from simulated supergathers.

Key words: Probability distributions; Seismic tomography; Computational seismology; Wave propagation.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on April 25, 2016

1 INTRODUCTION

Romero *et al.* (2000) introduced the concept of phase-encoded migration of seismic data, where reverse time migration of blended shot gathers associated with many shot positions could be performed with just two finite-difference (FD) solutions[1] to the wave equation. Blending is performed by stacking N different shot gathers together to form a supergather, where each shot gather is ideally encoded with a unique phase-encoding function ideally uncorrelated with any other phase-encoding function. The potential benefit of this approach, compared to standard migration, is an N-fold reduction in both the size of the input data and in the computation time.

The main problem with supergather migration is crosstalk noise that degrades the quality of the final image. Crosstalk noise is generated by migrating the ith shot gather with the imaging condition *unfortunately* tuned to the position of the jth shot, where $j \neq i$. For example, if $S_i(\mathbf{x})$ is the modelled mono-frequency wavefield at $\mathbf{x}$ for an encoded point source at the ith position, then the multisource modelling with one FD simulation yields $S(\mathbf{x}) = \sum_i S_i(\mathbf{x})$. Similarly, the back-propagated reflections in a supergather are given by $R(\mathbf{x}) = \sum_j R_j(\mathbf{x})$, where $R_j(\mathbf{x})$ represent the extrapolated wavefield at $\mathbf{x}$ associated with the jth source position. The mono-frequency supergather migration image is represented by $m(\mathbf{x}) = S(\mathbf{x})R(\mathbf{x})^* = \sum_i S_i(\mathbf{x})R_i(\mathbf{x})^* + \sum_{j\neq i}\sum_i S_i(\mathbf{x})R_j(\mathbf{x})^*$. The single summation term represents the standard migration image where shot gathers are individually migrated and the images are stacked together, whereas the double summation represents the crosstalk noise that spoils the image quality. Here, the spectral product of any encoding function with its conjugate is equal to 1. The

[1] One solution is used to back-propagate the blended reflections and the other is to model the direct waves from the multiple sources, each encoded with a different encoding function.

© 2011 The Authors
Geophysical Journal International © 2011 RAS

quality of $m(\mathbf{x})$ can be improved by recomputing the supergather migration with different encoding functions and stacking this new image with the original image; but this increases the cost of the procedure.

Adapting the encoding method to waveform inversion was used by Krebs *et al.* (2009) and others (e.g., Dai *et al.* 2009; Virieux & Operto 2009), who showed that multisource waveform inversion of phase-encoded supergathers[2] resulted in images with nearly the same quality as those of standard imaging, but with an order of magnitude less cost. An increase in efficiency was also obtained by Dai & Schuster (2010) in applying least squares migration to supergathers. For the case of 2-D waveform inversion, more than a 100-fold reduction in computation time was achieved in inverting 100-source supergathers. These results were empirically obtained, but did not rigorously reveal the theoretical relationship between the image quality and the encoding functions. Such an understanding could provide clues about how to optimally reduce crosstalk noise in the image from multisource migration or inversion.

In this paper, we derive formulas that show how the quality of the migration image is related to the variance of the user-defined source statics, amplitude statics and polarity statics for multisource imaging. Such formulas can be used to design optimal encoding strategies for multisource imaging; for example analysis reveals that the crosstalk term is exponentially reduced with increasing variance of the static shift and the square of the source frequency. We also provide formulas that relate the number of shot gathers in a superstack to the final signal-to-noise ratio (SNR) of the final image. These insights suggest, for example, that the variance of source statics should increase with coarser grids in multiscale inversion.

The theory section is divided into three subsections. The first subsection describes the equations for iterative stacking migration and iterative least squares migration of encoded supergathers. The next section derives the formula for the SNR as a function of iteration number and the last section derives that for the crosstalk strength term after ensemble averaging over the encoding function. These formulas are accompanied by numerical tests that reveal their relevance to realistic data and migration models.

2 THEORY: PHASE-ENCODED MIGRATION OF SUPERGATHERS

Assume a seismic experiment that consists of S shots, each shooting into the same number G of geophones evenly distributed on the earth's surface. In the frequency domain the sth shot gather is denoted by the $G \times 1$ vector $\mathbf{d}_s$, where each element represents the frequency response at a different geophone. The relationship between the earth model and shot gather $\mathbf{d}_s$ is given by

$$\mathbf{d}_s = \mathbf{L}_s \mathbf{m}, \tag{1}$$

where $\mathbf{m}$ is the $N \times 1$ reflectivity model vector[3] and the $G \times N$ matrix $\mathbf{L}_s$ represents the pre-stack modelling operator for the sth shot gather.

The seismic experiment typically consists of S shots evenly distributed on the earth's surface, and each shot gives rise to a new shot gather $\mathbf{d}_s$. This ensemble of shot gathers can be related to the earth model by the overdetermined system of equations

$$\begin{vmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ . \\ . \\ . \\ \mathbf{d}_S \end{vmatrix} = \begin{bmatrix} \mathbf{L}_1 \\ \mathbf{L}_2 \\ . \\ . \\ . \\ \mathbf{L}_S \end{bmatrix} |\mathbf{m}| \quad \text{or} \quad \mathbf{d} = \mathbf{L}\mathbf{m}, \tag{2}$$

where $\mathbf{L}$ represents the $(S \cdot G) \times N$ overdetermined matrix of pre-stack forward modelling matrices, and $\mathbf{d}$ represents the $(S \cdot G) \times 1$ block vector of shot gathers. The standard migration procedure that approximates the solution to the above system of equations is known as pre-stack migration (Claerbout 1992) and is represented in its simplest form by

$$\hat{\mathbf{m}} = \mathbf{L}^\dagger \mathbf{d} = \sum_{i=1}^{S} \mathbf{L}_i^\dagger \mathbf{d}_i, \tag{3}$$

where $\hat{\mathbf{m}}$ is the standard migration image, and $\mathbf{L}^\dagger$ represents the Hermitian conjugate of the modelling operator, otherwise known as the migration operator. The above equation says that each shot gather is separately migrated and the resulting migration images are summed together to give the standard pre-stack migration image.

The problem with 3-D pre-stack migration is that it is computationally expensive for large models and large numbers of sources with high-frequency content. To reduce this expense, Romero *et al.* (2000) proposed that different shot gathers can be phase encoded, summed together to form supergathers and then migrated all at once with a single finite-difference reverse time migration operation where the imaging condition is tuned to all of the different shot positions. This procedure is given by the following steps:

(i) Phase encode the ith shot gather $\mathbf{d}_i$ by an *orthonormal* phase encoding function $\mathbf{N}_i$ such that $\langle \mathbf{N}_i^* \mathbf{d}_i^* \mathbf{N}_j \mathbf{d}_j \rangle = \mathbf{d}_i^* \mathbf{d}_j \delta_{ij}$, where $\langle\ \rangle$ denotes ensemble averaging over the random phase encoding functions and δ_{ij} is the Kronecker delta function. For the discrete system of modelling equations, phase encoding shot gathers is equivalent to multiplying the ith block of modelling equations by the ith encoding function matrix $\mathbf{N}_i$ to form the encoded $(S \cdot G) \times G$ system of equations:

$$\begin{bmatrix} \mathbf{N}_1 & 0 & 0 \ldots 0 \\ 0 & \mathbf{N}_2 & 0 \ldots 0 \\ & . & \\ & . & \\ & . & \\ 0 & 0 & 0 \ldots \mathbf{N}_S \end{bmatrix} \begin{bmatrix} \mathbf{L}_1 \\ \mathbf{L}_2 \\ . \\ . \\ . \\ \mathbf{L}_S \end{bmatrix} |\mathbf{m}| = \begin{vmatrix} \mathbf{N}_1 \mathbf{d}_1 \\ \mathbf{N}_2 \mathbf{d}_1 \\ . \\ . \\ . \\ \mathbf{N}_S \mathbf{d}_S \end{vmatrix}, \tag{4}$$

or in more compact notation

$$[\mathbf{N}][\mathbf{L}]\mathbf{m} = [\mathbf{N}]\mathbf{d}. \tag{5}$$

Here, $\mathbf{N}_i$ is the $G \times G$ phase-encoding block matrix associated with the $G \times 1$ shot gather $\mathbf{d}_i$, and $[\mathbf{N}]$ represents the $S \cdot G \times G$ encoding matrix that contains all of the block matrices $\mathbf{N}_i$.

For the fixed frequency ω, the block-encoding matrix $\mathbf{N}_i$ can be reduced to the $G \times G$ identity matrix $\mathbf{I}$ multiplied by a scalar factor N_i:

$$\mathbf{N}_i = N_i \mathbf{I}. \tag{6}$$

An example of N_i is the phase shift $e^{i\omega\tau_i}$, where τ_i is a random time shift variable that time shifts the traces in the ith shot gather by the

[2] Krebs *et al.* formed their n-fold supergathers by stacking n shot gathers, each with a random polarity statics for each source.

[3] The earth model is discretized into a regular grid of cells with an unknown reflectivity value m_i in the ith cell. There are N cells and the reflectivity model vector is defined as the $N \times 1$ vector $\mathbf{m}$.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on April 25, 2016

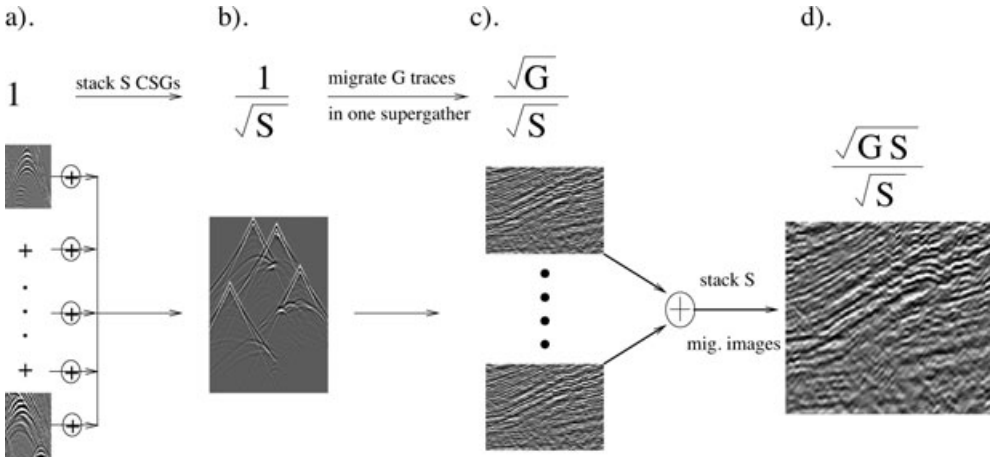

Figure 1. (a) Time-shifted shot gathers, (b) blended supergather created by blending S time-shifted shot gathers, (c) migration images after migrating the supergather for each shot position and (d) final image after summing S migration images. The final SNR enhancement is $\sqrt{G}$.

value τ_i. The important property of the phase-encoding matrices is that they are orthonormal under ensemble averaging; for example for the time shift encoding function we have

$$\left\langle \mathbf{N}_i \mathbf{N}_j^* \right\rangle = \langle e^{i\omega(\tau_j - \tau_i)} \rangle \, \mathbf{I} = \delta_{ij} \, \mathbf{I}, \tag{7}$$

where $\langle \ \rangle$ represents ensemble averaging over the random variables τ_i and τ_j. An example of an encoded supergather is shown in Fig. 1 with (a) time-shifted shot gathers and (b) depicting the supergather created by blending (i.e. adding) the time-shifted shot gathers.

(ii) Summing the encoded shot gathers together can be achieved by multiplying the encoded $(S \cdot G) \times 1$ data vector in eq. (5) by the $G \times (S \cdot G)$ blending matrix

$$\mathcal{I} = |\, \mathbf{I} \quad \mathbf{I} \quad \ldots \quad \mathbf{I} \,| \tag{8}$$

to get

$$\overbrace{[\, \mathbf{N}_1 \mathbf{L}_1 + \mathbf{N}_2 \mathbf{L}_2 + \cdots + \mathbf{N}_S \mathbf{L}_S \,]}^{\mathcal{L}=[\mathcal{I}][\mathbf{N}][\mathbf{L}]} \mathbf{m}$$

$$= \overbrace{\mathbf{N}_1 \mathbf{d}_1 + \mathbf{N}_2 \mathbf{d}_2 + \cdots + \mathbf{N}_S \mathbf{d}_S}^{\mathcal{D}=[\mathcal{I}][\mathbf{N}]\mathbf{d}}; \tag{9}$$

or more compactly

$$\mathcal{L}\mathbf{m} = \mathcal{D}. \tag{10}$$

Here, the encoded $G \times N$ supergather modelling matrix $\mathcal{L}$ and the $G \times 1$ encoded supergather vector $\mathcal{D}$ have their row dimensions reduced by the multiplicative factor of S compared to the original system of eq. (4).

(iii) The phase-encoded migration of supergathers is obtained by applying the adjoint operator $\mathcal{L}^\dagger$ to the encoded supergather $\mathcal{D}$ to get

$$\tilde{\mathbf{m}} = \mathcal{L}^\dagger \mathcal{D} = \sum_{i=1}^{S} \mathbf{L}_i^\dagger \mathbf{N}_i^* \mathbf{N}_i \mathbf{d}_i + \sum_{j \neq i}^{S} \sum_{i=1}^{S} \mathbf{L}_j^\dagger \mathbf{N}_j^* \mathbf{N}_i \mathbf{d}_i$$

$$= \overbrace{\sum_{i=1}^{S} \mathbf{L}_i^\dagger \mathbf{d}_i}^{\text{stand. migration } \hat{\mathbf{m}}} + \overbrace{\sum_{j \neq i}^{S} \sum_{i=1}^{S} \mathbf{L}_j^\dagger \mathbf{N}_j^* \mathbf{N}_i \mathbf{d}_i}^{\text{crosstalk noise } \mathbf{c}}$$

$$= \hat{\mathbf{m}} + \mathbf{c}, \tag{11}$$

which says that the phase-encoded migration image $\tilde{\mathbf{m}}$ consists of the standard migration image $\hat{\mathbf{m}}$ in eq. (3) and the crosstalk term $\mathbf{c}$. In this case it is assumed that $\mathbf{N}_i^* \mathbf{N}_i = \mathbf{I}$ for all integer values of i.

If the norm of the crosstalk vector $\mathbf{c}$ is small then the phase-encoded image $\tilde{\mathbf{m}}$ in eq. (11) (obtained from an S stack supergather) closely approximates the standard pre-stack migration image $\hat{\mathbf{m}}$. Thus, encoded migration of this supergather might result in an S-fold reduction in computation time compared to separately migrating each shot gather.

2.1 Iterative phase-encoded migration of supergathers

The crosstalk vector $\mathbf{c}$ in eq. (11) is typically not small in magnitude for large numbers of shot gathers in the supergather, and so the resulting migration image contains unacceptable noise (Romero *et al.* 2000). To decrease the level of this noise, iterative migrations can be performed and we now discuss two such methods: iterative stacking of encoded migration images and iterative least squares migration of encoded supergathers.

2.1.1 Iterative stacking of encoded migration images

Several encoded migration images can be computed using eq. (11), but each one dynamically uses a new encoding function[4] to form the input supergather. These images can be added (i.e. stacked) to the previous one. If this process is iterated, then the $k+1$ stacked migration image $\tilde{\mathbf{m}}^{(k+1)}$ obtained by stacking k realizations of eq. (11) is given by

$$\tilde{\mathbf{m}}^{(k+1)} = \tilde{\mathbf{m}}^{(k)} + \hat{\mathbf{m}} + \mathbf{c}^{(k+1)}, \tag{12}$$

which, after backward substitution, can be expressed as

$$\tilde{\mathbf{m}}^{(k+1)} = \overbrace{(k+1)\hat{\mathbf{m}}}^{\text{image signal}} + \overbrace{\sum_{k'=1}^{k+1} \mathbf{c}^{(k')}}^{\text{stacked crosstalk noise}}$$

$$= (k+1)\hat{\mathbf{m}} + \sum_{j \neq i}^{S} \sum_{i=1}^{S} \mathbf{L}_j^\dagger \left[\sum_{k'=1}^{k+1} \mathbf{N}_j^{*(k')} \mathbf{N}_i^{(k')} \right] \mathbf{d}_i. \tag{13}$$

[4] The crosstalk term at the k iteration is computed by using the k realization of the random variable τ_i.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on 2016

The block matrix $\mathcal{N}_{ij} = \sum_{k'=1}^{k+1} \mathbf{N}_j^{*(k')}\mathbf{N}_i^{(k')}$ approximates the ensemble averaging of the covariance matrix associated with the encoding functions. Because the phase encoding function $\mathbf{N}_i$ is uncorrelated with $\mathbf{N}_j$ for $j \neq i$, then $\mathcal{N}_{ij}$ grows as $\sqrt{k+1}$ compared to the migration signal $(k+1)\tilde{\mathbf{m}}$ growth of $k+1$. Hence, the SNR of the image associated with eq. (13) should grow as $\sqrt{k}$.

2.1.2 *Iterative least squares migration*

The overdetermined system of equations $\mathbf{Lm} = \mathbf{d}$ is often inconsistent and so an iterative least squares solution can be computed to form what is known as the least squares migration (LSM) image (Nemeth *et al.* 1999; Duquet *et al.* 2000). The LSM image for phase-encoded supergathers is obtained by first forming a data misfit function ϵ

$$\epsilon = \frac{1}{2}(\mathcal{L}\mathbf{m} - \mathcal{D})^{\dagger}(\mathcal{L}\mathbf{m} - \mathcal{D}), \tag{14}$$

so that the iterative steepest descent solution at the $k+1$ iterate is given by

$$\begin{aligned}\mathbf{m}^{(k+1)} &= \mathbf{m}^{(k)} - \alpha^{(k)}\mathcal{L}^{(k)\dagger}(\mathcal{L}^{(k)}\mathbf{m}^{(k)} - \mathcal{D}^{(k)}) \\ &= \mathbf{m}^{(k)} - \alpha^{(k)} \overbrace{\sum_{i=1}^{S} \mathbf{L}_i^{\dagger}\delta\mathbf{d}_i^{(k)}}^{\text{stand. LSM}} \\ &\quad - \alpha^{(k)} \overbrace{\sum_{j\neq i}^{S}\sum_{i=1}^{S} \mathbf{L}_j^{\dagger}\mathbf{N}_j^{*(k)}\mathbf{N}_i^{(k)}\delta\mathbf{d}_i^{(k)}}^{\text{crosstalk noise } \mathbf{c}}.\end{aligned} \tag{15}$$

Here, the data residual at the kth iteration for the ith shot gather is defined as

$$\delta\mathbf{d}_i^{(k)} = \mathbf{L}_i\mathbf{m}^{(k)} - \mathbf{d}_i = \mathbf{L}_i(\mathbf{m}^{(k)} - \mathbf{m}), \tag{16}$$

and $\alpha^{(k)}$ is the scalar step length. Without losing generality but gaining simplicity, we assume that the step length $\alpha^{(k)}$ is constant for all iterations such that $0 < \alpha < 2/\lambda_{\max}$, where $\lambda_{\max}$ is the maximum eigenvalue of the well-conditioned symmetric positive definite (SPD) matrix $\mathcal{L}^{\dagger}\mathcal{L}$. This procedure is sometimes known as the *fixed step size steepest descent method* and is guaranteed to converge for SPD $\mathcal{L}^{\dagger}\mathcal{L}$ (Burden & Faires 2001). Also, we assume that the crosstalk term is small enough so that phase-encoded steepest descent solution converges to a model close to the actual standard LSM model[5] for a reasonable number of iterations.

Substituting eq. (16) into eq. (15), assuming a fixed step size α that satisfies the above convergence condition, and employing backward substitution yields

$$\begin{aligned}\mathbf{m}^{(k+1)} &= \mathbf{m}^{(0)} - \alpha\sum_{i=1}^{S}\mathbf{L}_i^{\dagger}\mathbf{L}_i \overbrace{\left\{\sum_{k'=0}^{k}[\mathbf{m}^{(k')} - \mathbf{m}]\right\}}^{\mathbf{a}} \\ &\quad - \alpha\sum_{j\neq i}^{S}\sum_{i=1}^{S}\mathbf{L}_j^{\dagger}\mathbf{L}_i \overbrace{\left\{\sum_{k'=0}^{k} c_{ij}^{(k')}[\mathbf{m}^{(k')} - \mathbf{m}]\right\}}^{\mathbf{b}}.\end{aligned} \tag{17}$$

[5] Rigorous error bounds for $\delta\mathbf{x}$ and the perturbed system of equations $[\mathbf{A} + \delta\mathbf{A}](\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$ are given in Burden & Faires (2001). Correspondingly, the crosstalk term in eq. (15) can be related to $\delta\mathbf{A}$ while the standard migration term can be related to $\mathbf{A}$. The key point they make is that for well-conditioned $\mathbf{A}$ and a small crosstalk term $\delta\mathbf{A}$, $\delta\mathbf{x} + \mathbf{x}$ is close to $\mathbf{x}$.

Here, $\mathbf{a}$ is the sum of model residuals; $\mathbf{b}$ is the sum of randomized model residuals; $c_{ij}^{(k')}\mathbf{I} = \mathbf{N}_j^{*(k')}\mathbf{N}_i^{(k')}$ and $c_{ij}^{(k')} = e^{i\omega(\tau_i^{(k')} - \tau_j^{(k')})}$ for random time shift encoding.

In eq. (17), the bracketed *sum of model residuals* vector $\mathbf{a}$ should be compared to the last term $\mathbf{b}$ in brackets, which is a randomized sum of model residual vectors. If the sequence of model residuals $\mathbf{m}^{(k')} - \mathbf{m}$ are nearly parallel[6] to one another and are approximately equal to $\delta\mathbf{m}$, then $\mathbf{a}$ and $\mathbf{b}$ in eq. (17) can be approximated as

$$\mathbf{a} \approx (k+1)\delta\mathbf{m}; \quad \mathbf{b} \approx \delta\mathbf{m}\sum_{k'=0}^{k} c_{ij}^{(k')}. \tag{18}$$

Similar to the crosstalk noise in eq. (13), this noise vector $\mathbf{b}$ decreases compared to the signal vector $\mathbf{a}$ as the iteration index k increases. For this example, the iterations in the steepest descent method are a proxy[7] for ensemble averaging seen in eq. (13).

The relative benefit of phase-encoded iterative LSM over iterative stacking of encoded migration images is that, for an accurate velocity model and a well-conditioned and mostly consistent[8] system of equations, it produces images with higher spatial resolution. Empirical results (Dai & Schuster 2010) also suggest that it converges faster than iterative stacking.

The practical implementation of iterative LSM applied to encoded supergathers is the following:

(i) Choose a starting vector $\mathbf{m}^{(k)}$ for $k = 0$.

(ii) Select the kth realization for the values of the random variables $[\tau_1, \tau_2, \ldots, \tau_S]^{(k)}$, and compute the encoded supergather data $\mathcal{D}^{(k)}$.

(iii) Compute the modelled supergather $\mathcal{L}\mathbf{m}^{(k)}$ by one finite-difference simulation with S delayed sources.

(iv) Use the results of the supergather forward modelling to subtract from the phase encoded data $\mathcal{D}^{(k)}$ to get the residual $\mathbf{r}^{(k)} = \mathcal{L}\mathbf{m}^{(k)} - \mathcal{D}^{(k)}$ in eq. (15).

(v) Migrate the residual using $\mathcal{L}^{\dagger}\mathbf{r}^{(k)}$ to update the model with eq. (15).

(vi) Set $k \rightarrow$ k + 1 and repeat steps (ii)–(v) until acceptable convergence.

2.2 SNR versus number of blended shot gathers

Eq. (13) can be used to deduce formulas for SNR enhancement[9] of the migration image as a function of the number I of stacking iterations and the number S of encoded shot gathers in a supergather. For this analysis, we will assume the ideal condition that each trace is

[6] This special case is an ideal gedanken example used to guide, not prove, our intuitive understanding of convergence. Strict convergence criteria for stochastic gradients and approximation are discussed in Kushner & Yin (2003), where at each iteration the gradient is averaged over different realizations of the random encoding functions. This is somewhat similar to the method suggested by eq. (17).

[7] This is similar to the ergodic assumption that the ensemble average of the random variable $x(t)$ at a specific value of t, where $x(t)$ is a stationary white noise series, can be approximated by a weighted summation of $x(t)$ over the time index.

[8] Inconsistency becomes worse with increasing error in the migration velocity model.

[9] SNR is defined as $S/N = \|\mathbf{m}_{\text{ref}}\|/\|\tilde{\mathbf{m}}^{(k+1)} - \mathbf{m}_{\text{ref}}\|$, where $\tilde{\mathbf{m}}^{(k+1)}$ is the $(k+1)$-fold stacked image, and $\mathbf{m}_{\text{ref}}$ is the reference migration image (see Section 2.2.1 for details).

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on April 25, 2016

contaminated[10] with additive white noise and also contains a band-limited signal with amplitude 1. There are G traces/shot gather, S time-shifted shot gathers/supergather and all of the G evenly spaced geophones are fixed in location and record G traces/shot.

Figs 1(a) and (b) illustrates that stacking S time-shifted shot gathers together,[11] will degrade the SNR of a supergather migration image by the factor proportional to $1/\sqrt{S}$; Fig. 1(c) demonstrates that migrating a supergather of G traces for one assumed shot position will enhance the migration image trace to give an SNR proportional to $\sqrt{G}/\sqrt{S}$; and Fig. 1(d) illustrates that stacking S migration images together gives an SNR enhancement of $\sqrt{GS}/\sqrt{S} = \sqrt{G}$. To further enhance the SNR, I stacks of migration images, each with different realizations of different encoding functions, will give the final SNR enhancement proportional to

$$(\text{SNR})^{\text{super}} = \sqrt{GI}. \tag{19}$$

In comparison, the conventional pre-stack migration of S shot gathers (each with G traces) gives an SNR enhancement proportional to

$$(\text{SNR})^{\text{stand.}} = \sqrt{GS}. \tag{20}$$

This assumes that the signal in all of the traces is of the same amplitude and that the geometrical spreading effects have been corrected. Under the ideal conditions stated, phase-encoded migration (eq. 19) requires the same number of iterations as the number of shot gathers S in a supergather to achieve the same SNR level seen in standard pre-stack migration (eq. 20).

An important observation is that $(\text{SNR})^{\text{super}} = \sqrt{GI}$ is independent of the number of shots that comprise the supergather. Therefore, the original supergather can be subdivided into Q sub-supergathers,[12] each one created by blending a unique set of S/Q encoded shot gathers. Iteratively migrating each of these Q sub-supergathers and stacking the images together will lead to an enhanced

$$\text{SNR}^{\text{sub-super}} = \sqrt{QGI}, \tag{21}$$

but at the expense of Q extra supergather migrations. Results shown in Dai & Schuster (2010) and Boonyasiriwat & Schuster (2010) are consistent with this prediction.

2.2.1 *Numerical tests*

To show that the above formula roughly predicts the SNR as a function of the number of encoded image stacks, a 320-source phase-encoded supergather is generated by a Born modelling code for the SEG/EAGE salt model and iteratively migrated and stacked according to eq. (13). The spacing between the evenly spaced sources is 18.3 m, which is about 61 per cent of a wavelength for the 50 Hz Ricker wavelet; the standard deviation of the source static is 0.4 s, which is about 20 times of the source's dominant period. At each iteration of eq. (13), a new realization of the encoding functions is used to time shift each shot gather, followed by blending to form the supergather and its migration. After each iteration, the stacked migration image $\mathbf{m}^{(k)}$ is subtracted from the reference migration image $\mathbf{m}_{\text{ref}}$ to give the noise $\mathbf{m}^{(k)} - \mathbf{m}_{\text{ref}}$, which is then used to calculate SNR $= \|\mathbf{m}_{\text{ref}}\|/\|\mathbf{m}^{(k)} - \mathbf{m}_{\text{ref}}\|$, where $\mathbf{m}_{\text{ref}} = \lim_{p\to\infty} \mathbf{m}^{(p)}$. Note that because there exists a residual discrepancy between $\lim_{p\to\infty} \mathbf{m}^{(p)}$ and the standard migration image $\tilde{\mathbf{m}}$, $\mathbf{m}^{(p)}$ for a large value of p is chosen as $\mathbf{m}_{\text{ref}}$ to single out the effect of the number of stackings on SNR. In Fig. 2, the solid line shows the measured SNR and largely agrees with the dashed line predicted by eq. (19), where the 6500th stacked image is used as $\mathbf{m}_{\text{ref}}$. Note the wiggles in the solid curve confirm the prediction in Appendix B that the estimated SNR has non-zero variance for reasonably large value of $\sigma\omega$.

2.3 Ensemble averaging and crosstalk noise

Under the white-noise assumptions described in the previous section, the averaged crosstalk term in eq. (13) becomes small with increasing number k of stacks, so the phase-encoded image approaches that of the standard migration image $m(\mathbf{x})$. This can be shown by taking the ensemble average of eq. (11):

$$\langle\tilde{\mathbf{m}}\rangle = \overbrace{\sum_{i=1}^{S} \mathbf{L}_i^{\dagger}\mathbf{d}_i}^{\text{stand. migration}} + \overbrace{\sum_{j\neq i}^{S}\sum_{i=1}^{S} \mathbf{L}_j^{\dagger}\mathbf{d}_i \underbrace{\langle N_j^* N_i\rangle}_{R}}^{\text{crosstalk noise}}, \tag{22}$$

where $R = \langle N_j^* N_i\rangle$ is defined as the *crosstalk damping coefficient.*[13] Eq. (22) follows because eq. (6) is henceforth assumed and thus $\mathbf{N}_j^*\,\mathbf{N}_i = N_j^* N_i\,\mathbf{I}$. In this section, we derive the formula of R, which determines the mean value of crosstalk noise, for time-shift statics and for polarity statics. The cases for coupled source–receiver time statics and location statics as encoding functions are relegated to Appendix A; the variances of crosstalk noise for polarity statics and for polarity in combination with time-shift statics are addressed in Appendix B.

As a background, $\Phi(\omega) \stackrel{\text{def}}{=} \langle e^{i\omega\tau}\rangle_\tau$ is known as the *characteristic function* of a random variable τ, and if τ is normally distributed with 0-mean and σ^2-variance, denoted in shorthand as $\tau \sim \mathcal{N}(0,\sigma^2)$, its characteristic function is given by [see, e.g. eq. (8–57) in Papoulis 1991]:

$$\Phi(\omega) = e^{-\omega^2\sigma^2/2}. \tag{23}$$

Recall also the fact that if two random variables u and v are mutually independent, then

$$\langle f(u)g(v)\rangle = \langle f(u)\rangle\langle g(v)\rangle, \tag{24}$$

where $f(\cdot)$ and $g(\cdot)$ are two arbitrary functions that admit well-defined expectation value.

2.3.1 *Source time statics*

$N_s = e^{i\omega\tau_s}$. Assume that $\tau_s \sim \mathcal{N}(0,\sigma^2)$ represents a time shift applied to the sth source, and that the random variables τ_s, $\forall s = 1,\ldots,S$ are independent and identically distributed (i.i.d.). Note that $\langle N_s\rangle$ then happens to be the characteristic function $\Phi(\omega)$ given in eq. (23). Note also

$$N_s^* N_s = 1, \tag{25}$$

[10] We assume a zero-mean white noise series governed by a normal probability density function with a standard deviation of 1.
[11] The time shift is assumed to have a standard deviation greater than the source period.
[12] Here, S/Q is assumed to be an integer number.
[13] The subscripts $_{ji}$ of R are omitted for notational brevity where no ambiguity exists.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on 2016

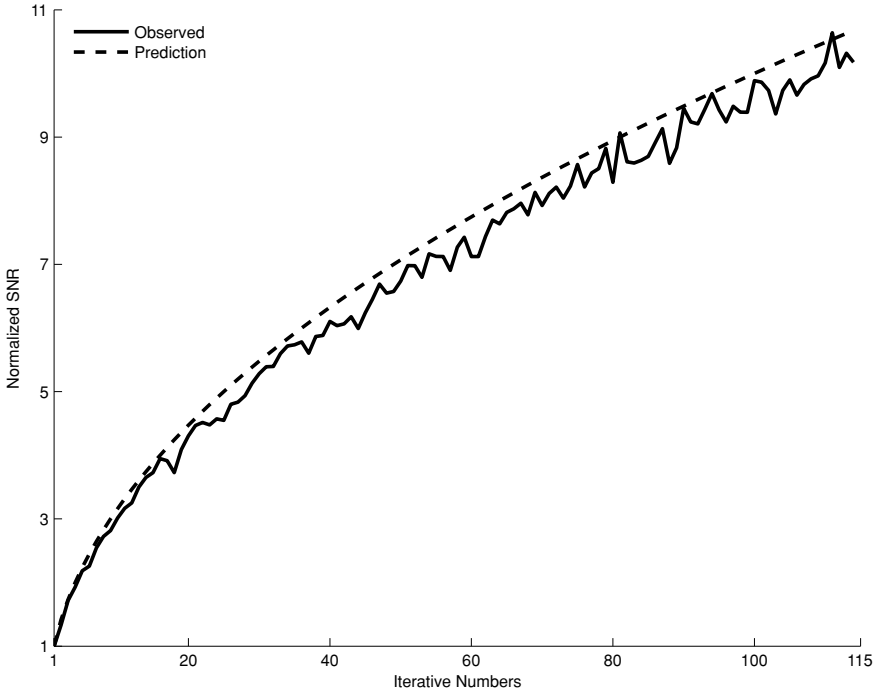

Figure 2. The solid (dotted) line indicates the measured SNR for phase-encoded iterative stacking migration, where the input data set consists of a 320-shot supergather associated with the SEG/EAGE salt model; the dashed line indicates the predicted relation from eq. (19). Both curves have been normalized to the value 1 for the first iterate.

because N_s is of pure phase. The expected cross-correlation of the phase-encoding function can be readily calculated as

$$\begin{aligned} R &= \langle N_s^* \rangle \langle N_{s'} \rangle, \quad \text{if } s \neq s', \\ &= |\langle N_s \rangle|^2 \\ &= \mathrm{e}^{-\omega^2 \sigma^2}, \end{aligned} \tag{26}$$

where the first equality follows from eq. (24), and the last step follows from eq. (23). The foregoing results are summarized as

$$R = \begin{cases} 1 & \text{if } s = s', \\ \mathrm{e}^{-\omega^2 \sigma^2} & \text{if } s \neq s'. \end{cases} \tag{27}$$

This equation resembles a Kronecker-δ function, that is $R \approx \delta_{ss'}$, if $\sigma \gg \omega^{-1}$. Eq. (22) in conjunction with eq. (27) thus says that the crosstalk noise can be reduced if the standard deviation σ of the source static shift is much greater than the period $2\pi/\omega$. Therefore, the variance of the user-specified source statics should increase with coarser grids in a multiscale inversion strategy.

In practice, ensemble averaging is not explicitly performed with supergather migration because this would be too costly. A poor man's proxy for ensemble averaging is the iterative stacking or iterative least squares migration described in the previous sections, or the iterative updates characteristic of iterative waveform inversion (Tarantola 1986; Mora 1987; Krebs *et al.* 2009).

2.3.2 *Numerical tests: encoding by time statics*

Numerical experiments are used to test the validity of eq. (27) for encoded supergathers associated with a single scatterer model as well as the SEG/EAGE salt model. A Born modelling solution to the 2-D wave equation is used to compute the line source responses of the models with both the sources and geophones on the surface. The scatterer is at a depth of 0.5 km in a homogeneous model with the P-wave velocity of 2000 ms^{-1}, and the dominant frequency of the narrow-band Ricker wavelet is 5 Hz; for the salt model the dominant frequency of the narrow-band Ricker source wavelet[14] is 50 Hz. Geophones are evenly distributed on the top surface of each model, with 256 and 320 traces per shot gather for the point scatterer model and the salt model, respectively.

The supergathers consist of 256 and 320 blended shot gathers for the point scatterer model and the salt model, respectively. Each supergather uses a new realization of the source statics and is migrated, and the resulting migration images are stacked together. Fig. 3 shows the standard migration and iterative stacking images with source statics of 0.01, 0.05 and 0.1 s for the single scatterer model. As expected, the migration images become more focused as σ increases because greater time shifts in neighbouring shot gathers reduces their similarity as well as the crosstalk strength in the migration image. This statement is further validated in Fig. 4 which plots both the theoretical (i.e. eq. 27) and numerically calculated crosstalk strength R[15] as a function of σ for the source time statics. There is very close agreement between the theoretical and numerically calculated curves.

The dynamically encoded supergathers associated with the salt model are also migrated and iteratively stacked to give the images shown in Fig. 5. As expected, the images become less noisy with an increase in the number of stacks and (as shown in Fig. 6) the standard deviation of the source statics. It is expected that the

[14] Narrow band sources are used because eq. (27) is only valid for a single frequency source.

[15] What's plotted in Figs 4 and 6 is normalized R, defined as $\hat{R} = \|\mathbf{m}_\sigma^{(k)} - \tilde{\mathbf{m}}\| / \|\mathbf{m}_{\sigma_0}^{(k)} - \tilde{\mathbf{m}}\|$, where $\sigma_0 = 0.01$s and $\tilde{\mathbf{m}}$ is the standard migration image.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on April 25, 2016

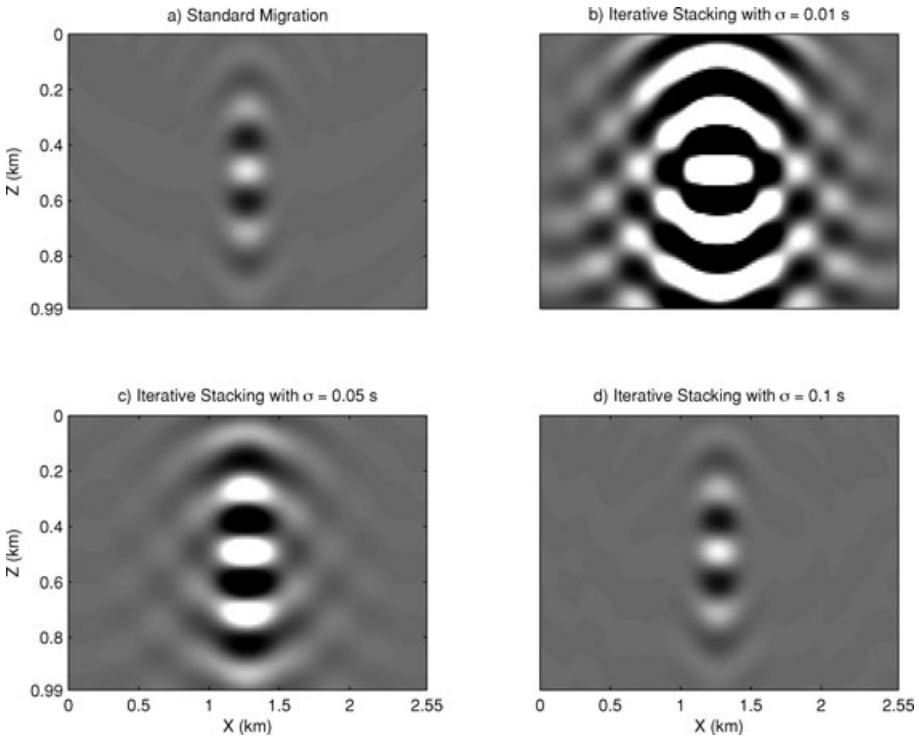

Figure 3. Iterative stacking migration images for data associated with the point scatterer model and the 5 Hz narrow band source wavelet: (a) standard migration, (b–d) show the iterative stacking images for the standard deviation of source static of 0.01, 0.05 and 0.1 s. Only 100 stacks were needed to compute these images.

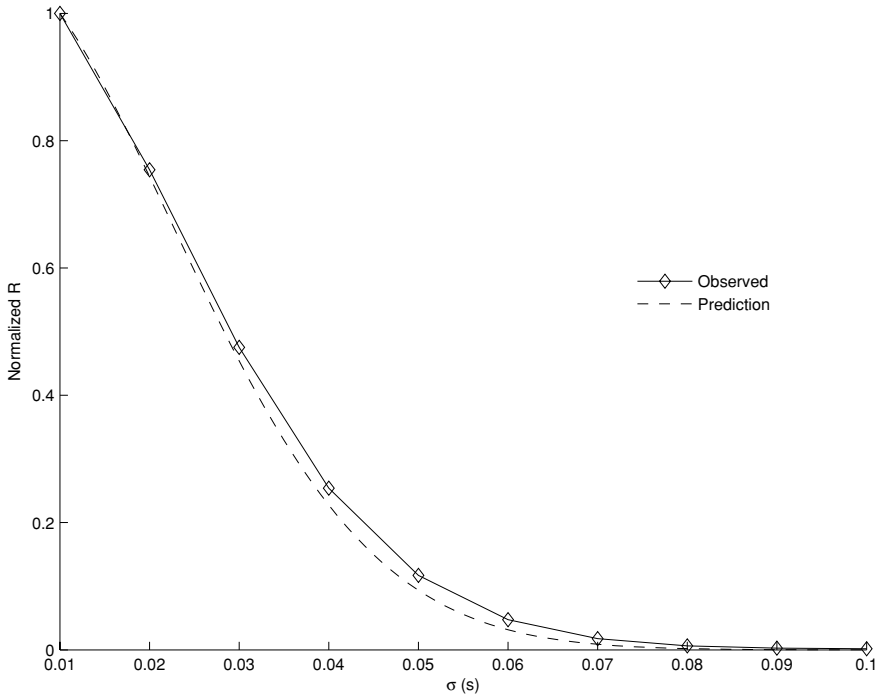

Figure 4. The solid line with diamonds indicates the measured normalized crosstalk damping coefficients for phase-encoded iterative stacking migration, where the input data set consists of a 256-shot supergather associated for a point scatterer model with a homogeneous background velocity of 2000 ms^{-1}; the dashed line indicates the predicted relation from eq. (19), and the input source wavelet is band limited with a dominant frequency of 5 Hz. The observed curve is nearly the same for a range of τ_s standard deviations between 0.01 and 0.1 s.

crosstalk noise in the salt image should be much greater than that for the single scatterer because more scatterers will increase the $\mathbf{L}_i\mathbf{d}_j$ contributions in the crosstalk. This observation might partly explain the curves in Fig. 6, which show that the absolute strength of the crosstalk noise with respect to σ is not well predicted by eq. (27). Better agreement can be reached with a greater number of stacks, but limited computational resources prevented a thorough testing of this claim. Moreover, the theoretical prediction

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on 16

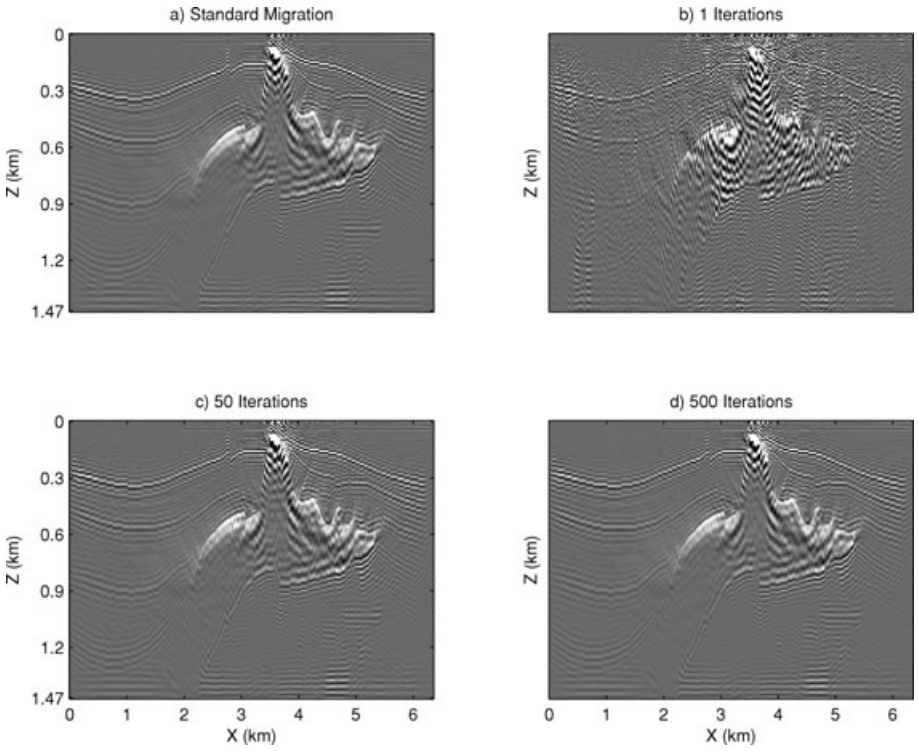

Figure 5. Iterative stacking migration images for the SEG salt model, where the narrow band source wavelet is peaked at 50 Hz: (a) standard migration and (b–d) iterative stacking images after 1, 50 and 500 iterations. Here, the standard deviation of the source statics is 0.05 s.

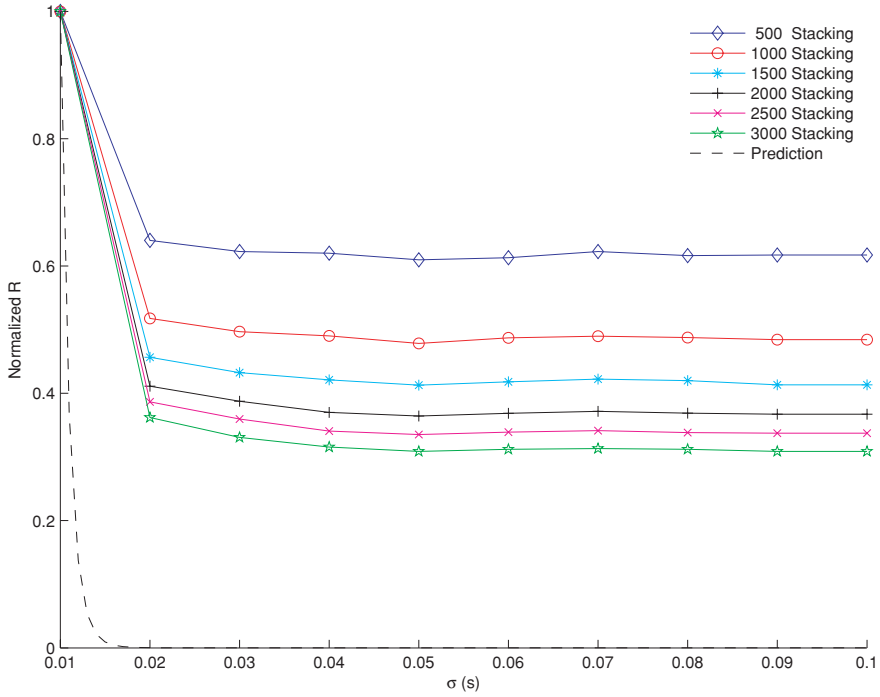

Figure 6. The solid lines with different colours indicate the measured normalized crosstalk damping coefficients for phase-encoded iterative stacking migration after 500, 1000, 1500, 2000 2500 and 3000 stacks, where the input data consists of a 320-shot supergather associated with the SEG salt model. The dashed line indicates the predicted relation from eq. (27), where the input narrow band source wavelet is peaked at 50 Hz.

from eq. (27) is for a single frequency at 50 Hz, whereas the simulations used a narrow-band 50 Hz source with a bandwidth of 26.25 Hz.

The R computed from the salt-model simulations is still higher than the prediction, although there is a trend of increasing agreement as the number of stacks grows. Despite this discrepancy, the multisource migration image after 500 stackings is visually indistinguishable from the single-source counterpart, as evident in Fig. 5. Note the excessive ringing in the images are due to the narrow band nature of source wavelet.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on April 25, 2016

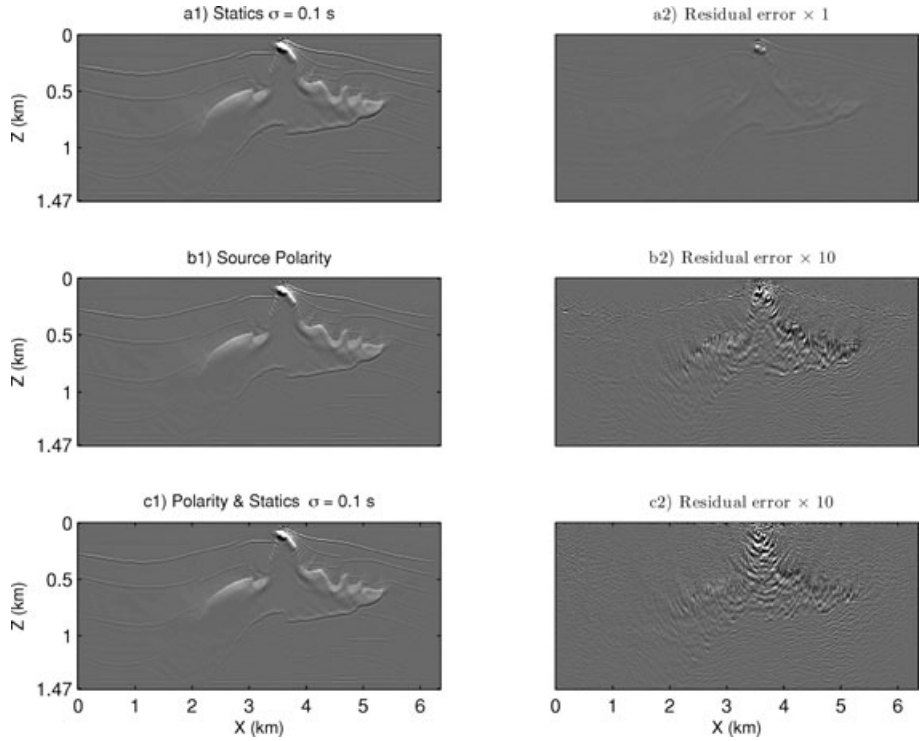

Figure 7. (Left column) iterative migration images $\mathbf{m}_{\text{enc}}$ after 3000 stacks, using encoding techniques such as: (a1) time statics ($\sigma = 0.1$ s), (b1) polarity and (c1) combined time statics ($\sigma = 0.1$ s) and polarity. (Right column) the residual error $\boldsymbol{\delta}$, i.e. the difference between the standard single-source migration $\hat{\mathbf{m}}$ and the ones shown in the left column, with (a2), (b2) and (c2) corresponding to (a1), (b1) and (c1), respectively. The residual error is scaled by a factor of 10, for ease of perception, in (b2) and (c2). The input data set consists of a blended 320-shot supergather associated with the SEG salt model, and the input source wavelet is a Ricker wavelet with peak frequency at 50 Hz.

2.3.3 *Polarity statics*

$N_s = p_s = \pm 1$. Here, the polarity p_s is a binary random variable that is i.i.d., $\forall\, s = 1, \ldots, S$, and p_s takes the values of -1 or 1 with probability 1/2 for each. It can be readily verified that

$$R = \langle p_s p_{s'} \rangle = \delta_{ss'}, \tag{28}$$

which shows that the crosstalk (i.e. when $s \neq s'$) noise is of 0-mean. Compared to the corresponding result of time statics, expressed in eq. (27), eq. (28) is an exact Kronecker-δ function, rather than an approximation as in eq. (27). This means that the performance of polarity encoding should be superior to that of time statics for low frequencies.

That the crosstalk noise has 0-mean is also true if the polarity statics acts in conjunction with other encoding statics as long as they are mutually independent. Namely, if $N_s = p_s h(\tau_s)$, where $h(\tau_s)$ is the other static and p_s is independent of τ_s, then it follows from eq. (24) that the crosstalk damping coefficient for shot gather encoded by both polarity and $h(\tau_s)$ is given by

$$\begin{aligned} R &= \langle p_s p_{s'} \rangle \langle h(\tau_s)^* h(\tau_{s'}) \rangle \\ &= \left\langle |h(\tau_s)|^2 \right\rangle \delta_{ss'}. \end{aligned} \tag{29}$$

This opens up new possibilities for combined encoders such as applying both polarity and time statics to each shot gather, as discussed in Appendix B. Therein the energy (or variance, in the case of 0-mean) of the resulting crosstalk noise can be used to quantify the effectiveness of candidate encoders. This is because a relationship exists between the variance of the crosstalk noise and the required number of stacks to meet a desired criterion (e.g. SNR)—the smaller the former is, the smaller the latter will be (see the paragraph following eq. B8).

The key result (Claim B1) of Appendix B is that the variance of the combined encoder of polarity and time statics is lower-bounded by 50 per cent of the variance of polarity encoding alone. This means that in the ideal case (i.e. when the equality sign in Claim B1 is attained), the former encoding strategy can be twice as good as the latter one.

2.3.4 *Numerical tests: encoding by polarity and by combined polarity and time statics*

The merit of encoding by combined polarity and time statics is substantiated by numerical tests on the salt model. Fig. 7 shows that the combined encoder of polarity and time statics produces a better migration image than computed with either polarity or time statics alone. For example, the residual error associated with the shallow reflector is significantly reduced in panel (c2) as compared to panel (b2).

The SNR[16] curves, as shown in Fig. 8, validate the theoretical prediction in Appendix B that combining both polarity and statics is the superior encoding strategy. In this figure, several features are noteworthy. First, the bottom curve (associated with time statics) appears to be a lot lower than the other two curves. The reason is that the source wavelet contains low-frequency content, which resists the crosstalk damping factor R, because it follows from eq. (26) that

[16] Here, SNR is defined as $\|\hat{\mathbf{m}}\|/\|\boldsymbol{\delta}\|$, where $\hat{\mathbf{m}}$ and $\boldsymbol{\delta}$ are defined in the caption of Fig. 7.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on 2016

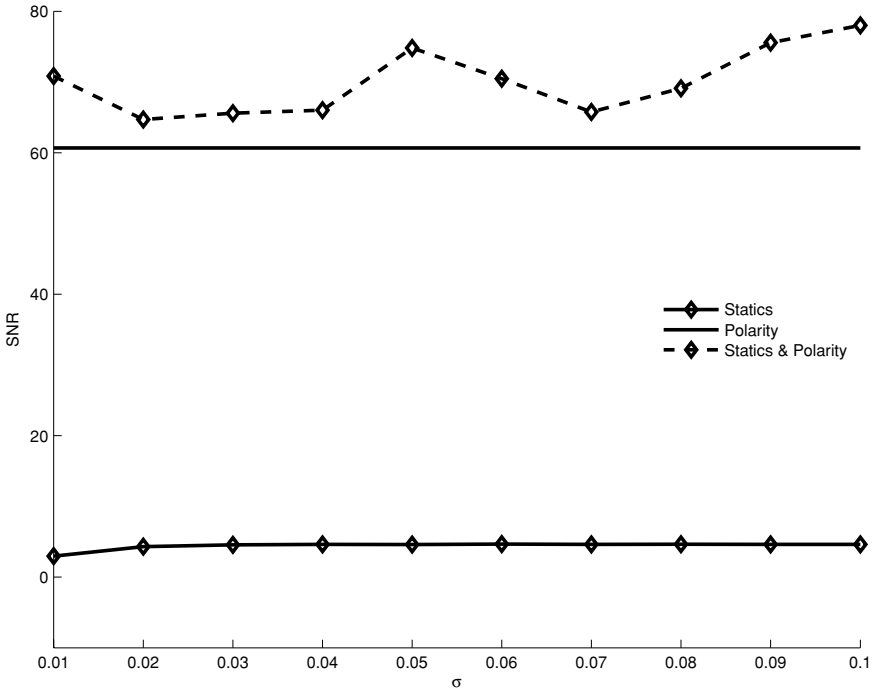

Figure 8. SNR versus the standard deviation of time statics, obtained from iterative migration after 3000 stacks, using encoding techniques such as: (bottom curve, solid, with diamonds) time statics, (middle line, solid) polarity and (top curve, dashed, with diamonds) combined time statics and polarity. The input data set consists of a blended 320-shot supergather associated with the SEG salt model, and the input source wavelet is a Ricker wavelet with peak frequency at 50 Hz.

$\lim_{\omega\to 0} R = 1$. Second, the solid line in the middle is constant with respect to σ, because the polarity encoding is governed by random variables independent of those governing time statics. Third, the apex of the top curve reads $\max \mathrm{SNR}_{pt} \simeq 78$, where the subscript $_{pt}$ denotes combined polarity and time statics. Let SNR_p be the SNR produced with polarity encoding, which is about 60.5 in Fig. 8. Therefore, we have $\max \mathrm{SNR}^2_{pt}/\mathrm{SNR}^2_p \simeq (78/60.5)^2 = 1.66 < 2$, in line with inequality (B30) discussed in Appendix B.

2.3.5 *Dynamic versus static encoding of shot gathers*

Until now, we assumed dynamic encoding schemes that changed the realization of the random variables from iteration to iteration to beat down the crosstalk term. In contrast, static encoding of shot gathers is the only option for field data where all of the shots are excited at nearly the same time and the blended wavefields are recorded as encoded supergathers. The excitation times or polarities of the shots are varied just once in the field trial, but cannot easily be changed after the data are recorded. In this case there should be no reduction of crosstalk by iterative stacking migration because the encoding function cannot be changed after each iteration.

However, eq. (17) suggests that such supergathers can be least squares migrated by an iterative steepest descent method, where the crosstalk strength is reduced by iterations. If the data are not encoded at each iteration then the value of the encoding function $c_{ij}^{(k)}$ in eq. (17) is independent of the iteration k index, but still depends on the i and j indices. This suggests that the summations over i and j can act as proxies for ensemble averaging and, we hypothesize, reduce the crosstalk term under favourable conditions.[17] This reduction in crosstalk noise is expected to be not as great as that for dynamic encoding where the proxy for ensemble averaging is the summation over both the k and the i, j indices.

[17] In contrast, the crosstalk term in the iterative stacking eq. (13) always grows as $O(k)$ for static encoding.

2.3.6 *Numerical tests: dynamic versus static encoding*

The hypothesis about the worse reduction in crosstalk noise by static LSM compared to dynamic LSM is validated by numerical tests on salt model data. Fig. 9 shows that dynamic phase-encoding produces better migration images than the static counterpart for the same number of iterations. At 300 iterations of steepest descent, the migration image with dynamic phase-encoding comes very close to conventional LSM. In terms of residual model error, the final image with dynamic phase-encoding outperforms the static counterpart, as shown in Fig. 10. If dynamic encoding is too expensive, a hybrid strategy is to use static encoding for a large number of iterations, and then reset with a new realization of the encoding functions. This should provide better crosstalk reduction than static LSM, but not as good as dynamic LSM.

3 REMARKS

The results in this paper show that large overdetermined systems of eq. (2) related to seismic imaging can be replaced by a much smaller system of encoded eq. (9). This alternative system of equations with a smaller memory footprint (more suitable to GPUs) can be solved by a dynamically encoded steepest descent method to give nearly the same answer to the original system of equations. Moreover, this algorithm appears to have some similarity to stochastic approximation algorithms (Kushner & Yin 2003). This raises some interesting issues.

(i) What are the properties of a general overdetermined system of equations so that it can be solved by the iterative steepest descent

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on April 25, 2016

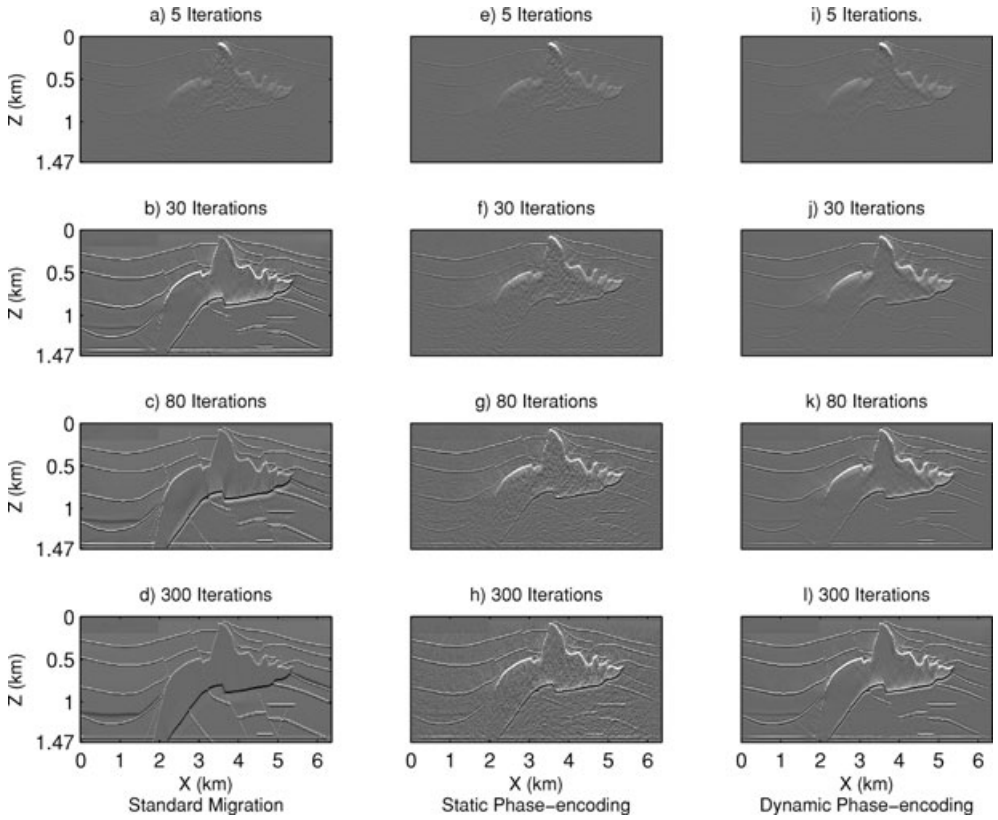

Figure 9. (a–d) Shows the standard LSM images after 5, 30, 80 and 300 iterations; (e–h) illustrate the static phase-encoded LSM images after 5, 30, 80, 300 iterations of a 320-shot supergather; while (i–l) present the dynamic phase encoding LSM images after 5, 30, 80, 300 iterations of a 320-shot supergather. All input data are associated with the SEG/EAGE salt model and a source Ricker wavelet with a peak frequency of 50 Hz.

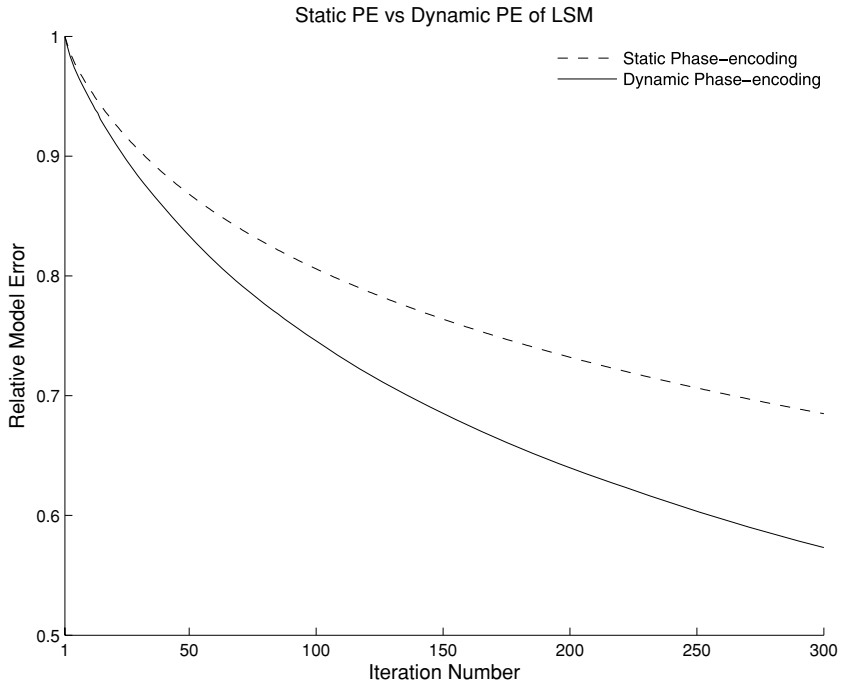

Figure 10. The solid line indicates the measured relative model error of least-squares migration with dynamic phase-encoding source time shift, while the dashed line indicates the one with static phase-encoding source time shift. Input data set consists of a 320-shot supergather associated with the SEG salt model, the input source wavelet is a Ricker wavelet with peak frequency at 50 Hz.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on 16

method with dynamic encoding? One property that seems necessary is that the summations in eq. (9) should not destroy unique information encoded in the data or the matrices. Thus, the encoding matrices should be designed to meet this demand.

(ii) The above algorithm is one that seems ideal for GPUs with large numbers of processors but limited capabilities in data transfer.

(iii) The procedure for iterative stacking migration of dynamically encoded supergathers is a special case of passive interferometry (Wapenaar *et al.* 2005), also known as daylight imaging (Rickett & Claerbout 1999). For daylight imaging, a trace $\mathbf{u}$ at position $\mathbf{x}_u$ and a trace $\mathbf{v}$ at position $\mathbf{x}_v$ are recorded for a random distribution of sources, each with a random encoding function $\mathbf{N}_i$, where

$$\mathbf{v} = \sum_i \mathbf{v}_i \mathbf{N}_i; \quad \mathbf{u} = \sum_i \mathbf{u}_i \mathbf{N}_i. \tag{30}$$

Here $\langle \mathbf{N}_i^* \mathbf{N}_j \rangle = \delta_{ij}$, and $\mathbf{v}_i \mathbf{N}_i$ represents the phase encoded trace due to a point source at $\mathbf{x}_i$ and receiver at $\mathbf{x}_v$. A white noise source spectrum $\mathbf{N}_i$ in the frequency domain serves the role of an encoding function. Taking the product of $\mathbf{v}^*$ with $\mathbf{u}$ and ensemble averaging over the random encoding functions yields

$$\begin{aligned} \langle \mathbf{v}^* \mathbf{u} \rangle &= \left\langle \sum_i \mathbf{v}_i^* \mathbf{N}_i^* \sum_j \mathbf{u}_j \mathbf{N}_j \right\rangle \\ &= \left\langle \sum_i \mathbf{v}_i^* \mathbf{u}_i \right\rangle + \text{crossterms}. \end{aligned} \tag{31}$$

Because of the property $\langle \mathbf{N}_i^* \mathbf{N}_j \rangle = \delta_{ij}$, the crossterms cancel out to give

$$\langle \mathbf{v}^* \mathbf{u} \rangle = \sum_i \mathbf{v}_i^* \mathbf{u}_i, \tag{32}$$

which can be shown to be proportional to the real part of the Green's function for a source at $\mathbf{x}_u$ and receiver at $\mathbf{x}_v$ (Wapenaar *et al.* 2005). Migration of phase-encoded supergathers can be shown to be a special case of eq. (32). To show this we take the following steps.

(i) Consider $\mathbf{u}^*$ to be the downward continued reflections at $\mathbf{x}_u$; the upgoing reflections at $\mathbf{x}_u$ have been extrapolated downwards from an encoded supergather at the earth's surface.

(ii) Interpret $\mathbf{v}$ to be the downgoing source field at $\mathbf{x}_v$ from these encoded sources at the earth's surface.

(iii) If $\mathbf{x}_v \neq \mathbf{x}_u$, that is the position $\mathbf{x}_v$ of the downward continued source field is not equal to that of the continued scattered field, then $\langle \mathbf{v}^* \mathbf{u} \rangle$ approximates the downward extrapolated field in the frequency domain for a shot at $\mathbf{x}_v$ and receiver at $\mathbf{x}_u$; inverse Fourier transforming this product approximates the Green's function for a source at $\mathbf{x}_u$ and receiver at $\mathbf{x}_v$. If $\mathbf{x}_v = \mathbf{x}_u$ and $\langle \mathbf{v}^* \mathbf{u} \rangle$ is summed over all frequencies, then this approximates the migration image obtained by migrating an encoded supergather.

(iv) Repeating this procedure for different supergathers at the earth's surface and stacking the migration images yields a result largely described by the iterative stacking formula (13).

In other words, iterative migration of phase-encoded supergathers is a special case of passive seismic interferometry. This is an important observation because it relates these two fields to one another and so might allow for a fruitful cross-fertilization of knowledge.

As an example, earthquake seismologists (Shapiro *et al.* 2005; Larose *et al.* 2006) record events on a receiver array spread out over a large area. The recorded traces over a specified window of time could be considered as an encoded supergather excited by simultaneous encoded sources that surround the array. Correlating the time windowed traces from two stations and stacking these correlated traces for many different time windows (i.e. encoded supergathers) leads to extraction of the virtual Green's function described earlier. In this case, the number of stacks would correspond to the iteration index I in eq. (21), and the number of recording stations in the array would be G. If the data are filtered according to different azimuthal directions of propagation, then data from a certain range of azimuthal angles can be considered a sub-supergather associated with a unique set of sources located outside the array. If Q sub-supergathers can be assembled and stacked, then eq. (21) suggests a higher SNR compared to processing the entire data set as one supergather. Similarly, these formulas can possibly be used for efficient extraction of shot gathers from simulations of encoded supergathers (Brown & Schuster 2009).

4 CONCLUSIONS

Formulas are presented that relate the strength of the crosstalk term in supergather migration to the covariance of the phase-encoding function. Analysis suggests that crosstalk noise in supergather migration is exponentially reduced with increasing values of static variance and frequencies. Source statics coupled to receiver statics, location statics, and amplitude statics can further reduce the strength of the crosstalk term. These formulas are also relevant in predicting crosstalk reduction for waveform inversion and least squares migration of encoded supergathers because their crosstalk terms are identical to the one in this paper.

The formula for the crosstalk damping coefficient suggests that the variance of the time static should be greater than the dominant period of the source wavelet. Therefore, the statics variance should increase with lower frequencies in a multiscale inversion strategy. For example, an efficient strategy for reducing crosstalk is to apply both temporal and amplitude source statics to the encoded supergather. Both theory and simulations show that on average encoding with both time and polarity statics is superior to encoding with just one of these encoding strategies.

SNR estimates show that $\text{SNR}^{\text{stand.}} = \sqrt{GS}$ for a standard migration image and $\text{SNR}^{\text{super}} = \sqrt{GI}$ for an image computed by migrating a phase-encoded supergather; here, G is the number of traces in a shot gather, I is the number of stacking iterations in the supergather and S is the number of encoded/blended shot gathers that comprise the supergather. The importance of these formulas is that they provide a precise understanding between different phase encoding strategies and image quality.

We also show that iterative migration of phase-encoded supergathers is a special case of passive seismic interferometry. The formulas can possibly be used for designing optimal strategies for passive seismic interferometry and efficient extraction of Green's functions from simulated supergathers.

ACKNOWLEDGMENTS

The authors thank the CSIM members for supporting this research (http://utam.gg.utah.edu/csim).

REFERENCES

Boonyasiriwat, C. & Schuster, G.T., 2010. 3D multisource full-waveform inversion using dynamic quasi-Monte Carlo phase encoding, in *Proceedings of the 80th Annual International Meeting, Society of Exploration Geophysicists,* Expanded Abstracts, pp. 1044–1049.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on April 25, 2016

Brown, S. & Schuster, G.T., 2009. Phase-encoded forward modeling, in *Proceedings of the 71st Conference and Exhibition,* European Association of Geoscientist and Engineers, Expanded Abstracts.
Burden, R. & Faires, J., 2001. *Numerical Analysis,* Brooks Cole.
Claerbout, J., 1992. *Earth Soundings Analysis: Processing versus Inversion,* Blackwell Scientific Publications, Cambridge, MA.
Dai, W. & Schuster, G.T., 2010. Least squares migration of multisource data with a deblurring filter, *Geophysics,* in press.
Dai, W., Zhan, G., Wang, X. & Schuster, G.T., 2009. Multi-source least squares migration, waveform inversion, and MVA analysis, in *Proceedings of the 79th Annual International Meeting,* Society of Exploration Geophysicists, Invited Workshop Talk.
Duquet, B., Marfurt, K. & Dellinger, J., 2000. Kirchhoff modeling, inversion for reflectivity, and subsurface illumination, *Geophysics,* **65**(4), 1195–1209.
Krebs, J.R., Anderson, J.R., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A. & Lacasse, M.D., 2009. Fast full-wavefield seismic inversion using encoded sources, *Geophysics,* **74**(6), WCC177–WCC188.
Kushner, H. & Yin, G., 2003. *Stochastic Approximation and Recursive Algorithms and Applications,* Springer-Verlag, New York.
Larose, E. *et al.,* 2006. Correlation of random fields, *Geophysics,* **71,** SI11–SI21.
Mora, P., 1987. Elastic wavefield inversion, *Ph.D. thesis,* Stanford University.
Nemeth, T., Wu, C. & Schuster, G.T., 1999. Least-squares migration of incomplete reflection data, *Geophysics,* **64**(1), 208–221.
Papoulis, A., 1991. *Probability, Random Variables and Stochastic Processes,* 3rd edn, McGraw-Hill Companies.
Rickett, J. & Claerbout, J., 1999. Acoustic daylight imaging via spectral factorization: helioseismology and reservoir monitoring, *The Leading Edge,* **18,** 957–960.
Romero, L.A., Ghiglia, D.C., Ober, C.C. & Morton, S.A., 2000. Phase encoding of shot records in prestack migration, *Geophysics,* **65,** 426–436.
Shapiro, N., Campillo, M., Stehly, L. & Ritzwoller, M., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science,* **307,** 1615–1618.
Stolt, R. & Benson, A., 1986. *Seismic Migration: Theory and Practice, in Handbook of Geophysical Exploration,* Vol. 5, Geophysical Press, London, UK.
Sun, Y., Schuster, G. & Sikorski, K., 1993. A quasi-Monte Carlo approach to 3-D migration: theory, *Geophysics,* **62,** 918–928.
Tarantola, A., 1986. Linearized inversion of seismic reflection data, *Geophys. Prospect.,* **32,** 998–1015.
Wapenaar, K. & Fokkema, J., 2006. Green's function representations for seismic interferometry, *Geophysics,* **71,** 33–46, doi:10.1190/1.2213955.
Zhan, G. & Schuster, G.T., 2010. Multisource phase-encoded waveform inversion, *J. Explor. Geophys.,* in press.
Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics,* **74,** 105–118.

APPENDIX A: THE MEAN OF CROSSTALK NOISE

The mean value of the crosstalk noise term will be derived for encoding functions associated with receiver time statics, and the effect of randomization of the source locations is discussed.

Case I: coupled source–receiver time statics $N_{gs} = \mathrm{e}^{i\omega\tau_g\tau_s/\tau}$. This means the geophone statics depends on the source position. Assuming τ is a constant and that τ_g is just a geophone static shift but not a random variable and that $\tau_s \sim \mathcal{N}(0, \sigma^2)$, the cross-correlation of the above phase-encoding function becomes, when $s \neq s'$,

$$R = |\langle N_{gs}\rangle|^2 = \mathrm{e}^{-\omega^2\tau_g^2/\tau^2}, \tag{A1}$$

and $R = 1$, when $s = s'$. This is summarized as $R = \delta_{ss'} + (1 - \delta_{ss'})\mathrm{e}^{-\omega^2\tau_g^2/\tau^2}$, implying that the source static coupled to the geophone static has an effective variance of $(\sigma\tau_g/\tau)^2$, which is larger than the source variance σ if $\tau < \tau_g$.

Case II: location statics $N_s = \mathrm{e}^{i\omega\tau_s}$. Random source locations can be roughly incorporated into the data by applying $N_s = \mathrm{e}^{i\omega\tau_s}$ to the shot gathers.[18] Here, τ_s is a random variable that reflects the timing error due to the relocation of the source position. In this case, the averaged correlation function has the same form as that for a source time shift static in eq. (27).

A related strategy for random source and receiver statics was presented by Sun *et al.* (1997) to reduce aliasing artefacts in migration images. Their formula for pre-stack migration says that, after ensemble averaging over the random location variable, the far-field migration image is governed by terms proportional to $\mathrm{e}^{-\omega^2\sigma^2\sin^2\theta/c^2}$, where $\sin\theta$ is inversely proportional to the image point depth in the migration image and c is the velocity in the homogeneous medium. This suggests that the crosstalk artefacts in a multisource migration image might become more prominent for deeper depths in the migration image, which is consistent with the waveform inversion image shown in Fig. A1.

Case III: multisource migration and 3-D statics. The above analysis was presented for 2-D multisource migration. In 3-D multisource migration, an extra summation is included in the migration formula for the extra coordinate in the source variable. This is equivalent to setting g in N_s to be that extra source coordinate, which says that the 3-D damping coefficient R will be desirably reduced by $\mathrm{e}^{-2\omega^2\sigma^2}$ rather than $\mathrm{e}^{-\omega^2\sigma^2}$. However, the crosstalk noise term might grow even larger as the stack number S grows. This suggests that the crosstalk damping coefficient should be tuned to beat down the $\sqrt{S}$ growth of the crosstalk noise.

APPENDIX B: THE VARIANCE OF CROSSTALK NOISE

The variance of the crosstalk noise will be derived for encoding functions associated with polarity and with combined polarity and time statics. We begin with a brief description of the energy of crosstalk noise.

B1 The relation between energy and variance

For a random vector $\mathbf{c}$, such as the crosstalk noise given in eq. (11), the energy is defined as

$$E(\mathbf{c}) \stackrel{\text{def}}{=} \langle\|\mathbf{c}\|^2\rangle, \tag{B1}$$

and the variance is defined as

$$\begin{aligned}\mathrm{Var}(\mathbf{c}) &\stackrel{\text{def}}{=} \langle\|\mathbf{c} - \langle\mathbf{c}\rangle\|^2\rangle \qquad\qquad (B2)\\ &= \langle\mathbf{c}^\dagger\mathbf{c} - \langle\mathbf{c}\rangle^\dagger\mathbf{c} - \mathbf{c}^\dagger\langle\mathbf{c}\rangle + \langle\mathbf{c}\rangle^\dagger\langle\mathbf{c}\rangle\rangle\\ &= \langle\|\mathbf{c}\|^2\rangle - 2\langle\mathbf{c}\rangle^\dagger\langle\mathbf{c}\rangle + \langle\mathbf{c}\rangle^\dagger\langle\mathbf{c}\rangle\\ &= \langle\|\mathbf{c}\|^2\rangle - \|\langle\mathbf{c}\rangle\|^2 = E(\mathbf{c}) - \|\langle\mathbf{c}\rangle\|^2, \qquad (B3)\end{aligned}$$

showing that the energy and the variance are related by

$$E(\mathbf{c}) = \mathrm{Var}(\mathbf{c}) + \|\langle\mathbf{c}\rangle\|^2. \tag{B4}$$

[18] This assumes surface consistency and a weathering layer with random thickness variations.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on 2016

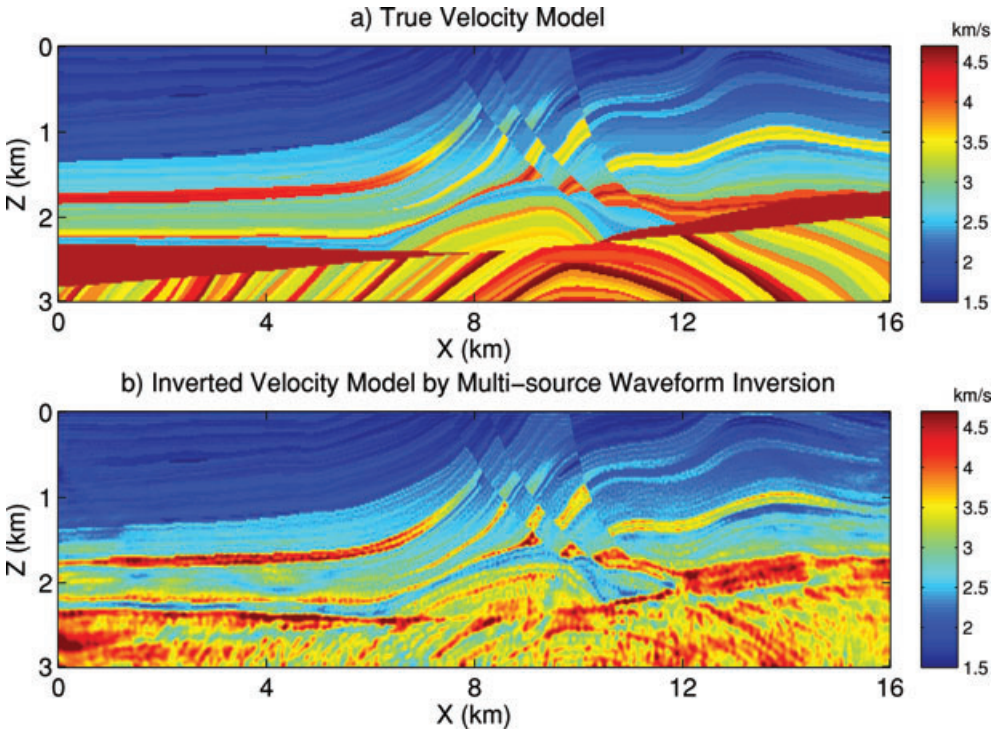

Figure A1. (a) The true velocity model. (b) The multisource waveform inversion result after 300 iterations with a deblurring filter and 100 shots per supergather. Note, the deeper horizons are noisier than the shallow ones (Dai *et al.* 2009; Zhan & Schuster 2010)

Let

$$\bar{\mathbf{c}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{c}_k \tag{B5}$$

be the sample average, such as what is computed in the stacking migration of eq. (13), over K i.i.d. random vectors. It can be verified [see, e.g. eq. (8–22) in Papoulis 1991] that

$$\langle \bar{\mathbf{c}} \rangle = \langle \mathbf{c} \rangle, \tag{B6}$$

$$\mathrm{Var}(\bar{\mathbf{c}}) = \frac{1}{K}\mathrm{Var}(\mathbf{c}), \tag{B7}$$

$$\text{and therefore } E(\bar{\mathbf{c}}) = \frac{1}{K}\mathrm{Var}(\mathbf{c}) + \|\langle \mathbf{c} \rangle\|^2. \tag{B8}$$

Comparing eq. (B8) with eq. (B4) reveals that the variance term in the energy expression gets reduced by sample averaging whereas the mean squared term $\|\langle \mathbf{c} \rangle\|^2$ remains unchanged. In particular, if $\langle \mathbf{c} \rangle = 0$, as is the case with polarity static (*cf*. eq. 29), the terms 'energy' and 'variance' can be used interchangeably, and the algebra in deriving the energy or variance will be greatly simplified. As this is the case studied in the present appendix, henceforth $\langle \mathbf{c} \rangle = 0$ is assumed. Accordingly, eq. (B8) reduces to

$$E(\bar{\mathbf{c}}) = \frac{1}{K}\mathrm{Var}(\mathbf{c}). \tag{B9}$$

This implies that, if we wish to reduce $E(\bar{\mathbf{c}})$ below a certain limit β by K-fold stacking, then the K is derived as $E(\bar{\mathbf{c}}) \le \beta \Rightarrow K \ge \mathrm{Var}(\mathbf{c})/\beta$. That is, the smaller the original $\mathrm{Var}(\mathbf{c})$, the smaller number K of stacks is required. It is for this reason that it is of interest to compute the variance of crosstalk noise.

B2 Combined encoder of polarity and time statics

To this end, let us introduce some notation first. Let $p_s = \pm 1$ be the polarity static as described in Case I of Appendix A. Let $h_s = \mathrm{e}^{i\omega\tau_s}$ be the source time static as described in Section 2.3.1 (h_s is renamed from N_s in Section (23)). Introduce

$$\alpha \overset{\text{def}}{=} \langle h_s^2 \rangle = \langle \mathrm{e}^{2i\omega\tau_s} \rangle$$

$$= \Phi(2\omega) = \mathrm{e}^{-2\omega^2\sigma^2}, \tag{B10}$$

where the last step follows from eq. (23). The combined phase encoder treated in the present appendix is the product of the polarity and the source time statics:

$$N_s \overset{\text{def}}{=} p_s h_s, \tag{B11}$$

where p_s and h_s are mutually independent. Let (i, j), with $i \ne j$, represent an ordered pairs of sources. Define the phase encoders associated with a pair of sources as

$$\psi_{(i,j)} \overset{\text{def}}{=} N_i^* N_j. \tag{B12}$$

Making use of eqs (28) and (B10), the following properties of the combined phase encoder N_s can be verified:

$$\langle N_s \rangle = 0, \tag{B13}$$

$$\langle \psi_{(i,j)} \rangle = \delta_{ij}, \tag{B14}$$

$$\langle N_i N_j \rangle = \langle N_i^* N_j^* \rangle = \alpha\, \delta_{ij}, \tag{B15}$$

$$\langle \psi^*_{(i,j)} \psi_{(k,l)} \rangle = \begin{cases} 1 & \text{if } (i,j) = (k,l), \\ \alpha^2 & \text{if } (i,j) = (l,k), \\ 0 & \text{otherwise.} \end{cases} \tag{B16}$$

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on April 25, 2016

With this preparation, we are now in a position to derive the variance of crosstalk noise. Let $\mathbf{c}_{(i,j)}$ be the crosstalk vector stemming from a pair of sources (i, j), given as

$$\mathbf{c}_{(i,j)} \stackrel{\text{def}}{=} \psi_{(i,j)}\breve{\mathbf{m}}_{(i,j)}, \tag{B17}$$

where $\breve{\mathbf{m}}_{(i,j)} \stackrel{\text{def}}{=} \mathbf{L}_i^\dagger \mathbf{d}_j.$ (B18)

Then eq. (B14) says that $\langle \mathbf{c}_{(i,j)} \rangle = 0$ because $i \neq j$ by the definition given earlier of the pair (i, j). The total crosstalk noise (before ensemble averaging, *cf.* eqs 11 and 22) can be written as:

$$\mathbf{c} = \sum_{\text{valid } (i,j)} \mathbf{c}_{(i,j)}. \tag{B19}$$

Here, 'valid (i, j)' represents *all* valid pairs of sources (i, j). For this example specifically, it represents $i \neq j, i = 1, \ldots, S$, and $j = 1, \ldots, S$. This shorthand is also used in eqs (B21) and (B25). In the latter case it represents all $i < j, i = 1, \ldots, S$ and $j = 1, \ldots, S$.

Because $\mathbf{c}$ is 0-mean, the variance of the total crosstalk noise $\mathbf{c}$ is given by

$$\text{Var}(\mathbf{c}) = \langle \|\mathbf{c}\|^2 \rangle \tag{B20}$$

$$= \sum_{\text{valid } (i,j)(k,l)} \left\langle \mathbf{c}_{(i,j)}^\dagger \mathbf{c}_{(k,l)} \right\rangle \tag{B21}$$

$$= \sum_{\text{valid } (i,j)(k,l)} \left\langle \psi_{(i,j)}^* \psi_{(k,l)} \right\rangle \breve{\mathbf{m}}_{(i,j)}^\dagger \breve{\mathbf{m}}_{(k,l)}. \tag{B22}$$

Because of eq. (B16), all terms in eq. (B22) vanish except for those associated with either $(i, j) = (k, l)$ or $(i, j) = (l, k)$. Introduce

$$l\rho_{ij} \stackrel{\text{def}}{=} \|\breve{\mathbf{m}}_{(i,j)}\|^2 \tag{B23}$$

$$\varrho_{ij} \stackrel{\text{def}}{=} \breve{\mathbf{m}}_{(i,j)}^\dagger \breve{\mathbf{m}}_{(j,i)}. \tag{B24}$$

Note that, while focusing on two source indices $i < j$, the running summation in eq. (B22) involves four surviving terms: $(i, j)(i, j)$, $(j, i)(j, i)$, $(i, j)(j, i)$ and $(j, i)(i, j)$, represented in turn by ρ_{ij}, ρ_{ji}, ϱ_{ij} and ϱ_{ji}. Making use of eq. (B16), the variance of crosstalk noise can finally be expressed as

$$\text{Var}(\mathbf{c}_{pt}) = \sum_{\text{valid } (i<j)} \left[(\rho_{ij} + \rho_{ji}) + \alpha^2 (\varrho_{ij} + \varrho_{ji}) \right]. \tag{B25}$$

$$\therefore \lim_{\alpha \to 0} \text{Var}(\mathbf{c}_{pt}) = \sum_{\text{valid } (i<j)} (\rho_{ij} + \rho_{ji}). \tag{B26}$$

Here, the subscript $_{pt}$ is added to highlight the fact that eqs (B25) and (B26) are obtained while assuming combined polarity and time statics (*cf.* eq. B11).

B3 Polarity encoding alone

If only polarity statics is assumed, that is $N_s \stackrel{\text{def}}{=} p_s$, instead, then all the derivations starting from eq. (B12) leading to eq. (B25) remain valid if α is set to be 1. Accordingly, in this case the variance of crosstalk noise is given as

$$\text{Var}(\mathbf{c}_p) = \sum_{\text{valid } (i<j)} [(\rho_{ij} + \rho_{ji}) + (\varrho_{ij} + \varrho_{ji})], \tag{B27}$$

where the subscript $_p$ represent 'polarity encoding alone'.

Comparing eq. (B27) with eq. (B25) reveals that the effect of the source time static lies in reducing the sum of $(\varrho_{ij} + \varrho_{ji})$. It is of interest to know the relative magnitude of these terms compared to the sum of $(\rho_{ij} + \rho_{ji})$.

B4 Comparison between the combined encoder and polarity encoding alone

Lemma B1 $|\varrho_{ij}| + |\varrho_{ji}| \leq \rho_{ij} + \rho_{ji}$.

Proof. Because of the Cauchy-Schwarz inequality, $|\varrho_{ij}| = |\breve{\mathbf{m}}_{(i,j)}^\dagger \breve{\mathbf{m}}_{(j,i)}| \leq \|\breve{\mathbf{m}}_{(i,j)}\| \cdot \|\breve{\mathbf{m}}_{(j,i)}\| = \sqrt{\rho_{ij}\,\rho_{ji}}$. Similarly, $|\varrho_{ji}| \leq \sqrt{\rho_{ji}\,\rho_{ij}}$. Therefore, $|\varrho_{ij}| + |\varrho_{ji}| \leq 2\sqrt{\rho_{ji}\,\rho_{ij}} \leq \rho_{ij} + \rho_{ji}$, where the last step follows from the fact that $2ab \leq a^2 + b^2$. □

Claim B1 $\lim_{\alpha \to 0} \text{Var}(\mathbf{c}_{pt}) \geq 0.5\,\text{Var}(\mathbf{c}_p)$.

Proof. It follows from Lemma B1 that $\rho_{ij} + \rho_{ji} \geq 0.5(\rho_{ij} + \rho_{ji} + |\varrho_{ij}| + |\varrho_{ji}|) \geq 0.5(\rho_{ij} + \rho_{ji} + \varrho_{ij} + \varrho_{ji})$. Thus,

$$\sum_{\text{valid } i<j} (\rho_{ij} + \rho_{ji}) \geq \sum_{\text{valid } i<j} 0.5[(\rho_{ij} + \rho_{ji}) + (\varrho_{ij} + \varrho_{ji})], \tag{B28}$$

which, after identifying with eqs (B26) and (B27), concludes the proof. □

Claim B1 is the main result of this appendix, suggesting that the combined encoder of polarity and source time statics can yield a crosstalk noise of variance $\text{Var}(\mathbf{c}_{pt})$ that is almost 1/2 that of polarity encoding alone. The fact that combined independent encoders would perform better than any one encoder alone is intuitive because they are able to reduce some high order moments involving encoders pertaining to different sources (*cf.* α^2 in eq. B16).

Now that the variances of two alternative encoding techniques have been related, we show in passing what this implies for SNR, a performance measure commonly used in practice. The SNR in Section 2.3.4 is defined as $\|\hat{\mathbf{m}}\|/\|\boldsymbol{\delta}\|$, where the standard migration $\hat{\mathbf{m}}$ (*cf*. eq. 3) serves as the signal and the residual error $\boldsymbol{\delta}$ between the standard migration and the K-fold averaged phase-encoded migration $\overline{\hat{\mathbf{m}}}$ serves as the noise. Here, $\overline{\hat{\mathbf{m}}}$ is computed by forming the sample-average of eq. (11) over a sample size of K, and therefore $\boldsymbol{\delta}$ is just the sample-average of crosstalk noise $\bar{\mathbf{c}}$ as defined in eq. (B5).

The ratio of the squared SNRs associated with K-fold stacked migration using the two alternative encoding techniques is given as

$$\frac{\text{SNR}_{pt}^2}{\text{SNR}_p^2} = \frac{\|\bar{c}_p\|^2}{\|\bar{c}_{pt}\|^2}$$

$$\simeq \frac{\langle \|\bar{c}_p\|^2 \rangle}{\langle \|\bar{c}_{pt}\|^2 \rangle} = \frac{E(\bar{c}_p)}{E(\bar{c}_{pt})}$$

$$= \frac{\frac{1}{K}\text{Var}(c_p)}{\frac{1}{K}\text{Var}(c_{pt})} = \frac{\text{Var}(c_p)}{\text{Var}(c_{pt})}. \tag{B29}$$

$$\leq 2 \tag{B30}$$

Here, the first line follows because the common factor $\|\hat{\mathbf{m}}\|$ in SNR gets cancelled; the second line follows from the definition in eq. (B1); the third line follows from substituting in eq. (B9) and the last line results from Claim B1. This result suggests that the ratio of variances is reflected in the reciprocal ratio of the squares of SNRs. This reciprocality arises intuitively because the larger the variance of crosstalk noise, the smaller the SNR would be.

Downloaded from http://gji.oxfordjournals.org/ at The University of Western Australia on 2016