STANFORD EXPLORATION PROJECT

Biondo Biondi and Claudio Guerra

Report Number 136, November 2008

SEP136 — TABLE OF CONTENTS

New phase-encoding schemes for wave-equation migration

Claudio Guerra and Biondo Biondi

ABSTRACT

Prestack exploding-reflector modeling aims to synthesize a small dataset comprised of areal shots, while preserving the correct kinematics to be used in iterations of migration velocity analysis. To achieve this goal, the modeled areal data must be combined into sets. However, crosstalk arises during migration due to the correlation of unrelated wavefields. Phase encoding the modeling experiments, can attenuate crosstalk during migration. In the geophysical community, the most used phase-encoding schemes are plane-wave-phase encoding and random-phase encoding. Here, we exploit the application of Gold codes commonly used in wireless communication, radar and medical imaging communities to phase encode data. We also introduce a method to compute random-phase functions by solving an optimization problem ... (more still to come). We show that adequately selecting the Gold codes can potentially shift the crosstalk out of the migration domain, yielding an image free of crosstalk.

INTRODUCTION

Biondi (2006, 2007) introduced the concept of the prestack exploding-reflector modeling. This method synthesizes source and receiver wavefields along the entire survey at the surface, in the form of areal data, starting from a prestack migrated image cube computed with wave-equation migration. For migration velocity analysis, the aim is to generate a considerably smaller dataset than the one used in the initial migration, while maintaining the necessary kinematic information

Conceptually, the synthesized areal data are computed by upward propagating source and receiver wavefields, using as initial conditions subsurface-offset-domain common-image gathers (SODCIG). To decrease the number of experiments to migrate, we take advantage of the linearity of the wave propagation to combine several experiments into a set of composite records. Combining several experiments, though, gives rise to crosstalk during imaging (Biondi, 2006; Guerra and Biondi, 2008). Guerra and Biondi (2008) use pseudo-random-phase encoding (Romero et al., 2000) during the modeling step to attenuate crosstalk. For conciseness reason, from now on we use the term random instead of pseudo-random.

The exploration geophysics community employs pseudo-random codes using intrinsic functions specific to the programing language. These pseudo-random codes present, generally, a uniform distribution and many of the algorithms are based on the *multiplicative congruential algorithm*, invented by Lehmer (1951). Their autocorrelation and cross-correlation functions have no special properties. The autocorrelation function presents nearly periodic side lobes with additive low-amplitude random variations. The peak-to-side lobe ratio is around 30. The cross-correlation function is pseudo-random, and its amplitudes are of

the same order of magnitude as those of the non-zero lags of the autocorrelation function. Herein, theses codes are called conventional random codes.

In wireless communication, specially for systems using Code Division Multiple Access (CDMA), different pseudo-random codes have been widely used (Shi and Schelgel, 2003). Medical imaging (Gran, 2005) and radar communities (Levanon and Mozeson, 2004) also exploit the statistical properties of these pseudo-random codes to increase bandwidth, signal-to-noise ratio and pulse compression. These codes are binary sequences and have unique autocorrelation and cross-correlations properties which make them more suited to achieving the above-mentioned objectives with minimal crosstalk. The autocorrelation function is represented by a large peak, whose amplitude equals the number of samples in the code, and the cross-correlation peaks, at non-zero lags, with the same amplitudes as that of the autocorrelation. Examples of binary pseudo-random codes used by these communities are Golay (Golay, 1961; Tseng, 1972), Kasami (Kasami, 1966) and Gold codes (Gold, 1967). Here, we exploit the properties of the Gold codes to encode seismic data.

In addition to using Gold codes, we introduce a novel method to compute phase functions by solving an optimization problem ...

In the next section we give a brief description of the prestack exploding-reflector modeling. Then we discuss how to compute the Gold codes and the optimized random-phase functions. We illustrate the application of each phase-encoding scheme by migrating phase-encoded areal data synthesized by prestack exploding-reflector modeling.

PRESTACK EXPLODING-REFLECTOR MODELING

Starting from a prestack image obtained by wave-equation migration, areal source and receiver wavefields are modeled at the surface by

$$\begin{aligned} S(x,\omega) &= G(z_\xi, x_\xi - h_\xi; x, z=0, \omega) * I_s(z_\xi, x_\xi, h_\xi), \\ R(x,\omega) &= G(z_\xi, x_\xi + h_\xi; x, z=0, \omega) * I_r(z_\xi, x_\xi, h_\xi), \end{aligned} \tag{1}$$

where $S(x,\omega)$ is the source wavefield at $z=0$; $R(x,\omega)$ is the receiver wavefield at $z=0$; $I_s(z_\xi, x_\xi, h_\xi)$ and $I_r(z_\xi, x_\xi, h_\xi)$ are the prestack images used as initial conditions for the source and receiver wavefield extrapolation, respectively, at a selected position, x_ξ,; $G(z_\xi, x_\xi \pm h_\xi; x, z=0, \omega)$ represents the operator that extrapolates the wavefields from the subsurface to the surface; h_ξ is the subsurface offset; z_ξ is depth; ω is the temporal frequency and x is the spatial coordinate in the data space coinciding with x_ξ. The prestack images used as initial conditions for the source and receiver wavefield extrapolation should be dip-independent gathers computed by changing the dip along the offset direction according to the apparent geological dip (Biondi, 2007). Here we use one-way extrapolators for both modeling and migration of the areal data.

By combining sets of individual modeling experiments into the same areal data, the amount of data input into migration can be significantly decreased. We achieve this by regularly selecting individual experiments and adding them up into their set, after being

upward propagated, according to

$$\begin{aligned}\widetilde{S_n}(x,\omega) &= \sum_{n=1}^{k}\sum_{i=n,k,N} S_i(x,\omega),\\ \widetilde{R_n}(x,\omega) &= \sum_{n=1}^{k}\sum_{i=n,k,N} R_i(x,\omega),\end{aligned} \tag{2}$$

where $\widetilde{S_n}(x,\omega)$, and $\widetilde{R_n}(x,\omega)$ contain k sets of summed areal sources and areal receivers, respectively; N is the number of SODCIGs. Every kth areal datum is selected to compose one set. Pairs of $\widetilde{S_k}(x,\omega)$ and $\widetilde{R_k}(x,\omega)$ are to be used as the areal source and the areal receiver wavefields, respectively, in areal shot migration.

Migration of the combined areal data produces an image by cross-correlating the combined areal source and receiver wavefields,

$$\widetilde{I_m}(z_\xi,x_\xi,h_\xi)=\sum_{\omega}\widetilde{S_m}^{\star}(z_\xi,x_\xi-h_\xi,\omega)\widetilde{R_m}(z_\xi,x_\xi+h_\xi,\omega), \tag{3}$$

where * represents complex conjugation. If $\widetilde{S_m}(x,\omega)$ and $\widetilde{R_m}(x,\omega)$ in equation 2 are comprised by two summed areal shots, the image $\widetilde{I_m}(z_\xi,x_\xi,h_\xi)$ will be given by

$$\begin{aligned}\widetilde{I_m}(z_\xi,x_\xi,h_\xi) = I_1(z_\xi,x_\xi,h_\xi)+I_2(z_\xi,x_\xi,h_\xi)+\\ \sum_{\omega}S_1^{\star}(z_\xi,x_\xi-h_\xi,\omega)R_2(z_\xi,x_\xi+h_\xi,\omega)+\\ \sum_{\omega}S_2^{\star}(z_\xi,x_\xi-h_\xi,\omega)R_1(z_\xi,x_\xi+h_\xi,\omega).\end{aligned} \tag{4}$$

In equation 4, the last two summation terms represent the crosstalk.

Guerra and Biondi (2008) show that in the prestack exploding-reflector method the crosstalk has two distinct origins. The crosstalk that is prominent in the zero-subsurface offset section, which herein we call type-1, is due to the cross-correlation of events in the source wavefield with that in the receiver wavefields modeled by the same SODCIG. Crosstalk in non-zero subsurface offsets, which herein we call type-2, is related to cross-correlation of events in the source wavefields with those in the receiver wavefields modeled by different SODCIGs.

Guerra and Biondi (2008) introduce strategies to attenuate the crosstalk. Migration of (x,z,ω)-random-phase encoded data disperses the crosstalk energy throughout the image as a pseudo-random background noise. By adding more realizations of random-phase encoded areal data, the speckled noise is further attenuated.

OPTIMIZED RANDOM CODES

...

GOLD CODES

Before describing Gold codes it is useful to define maximum length sequences.

Linear feedback shift registers (LFSR) are called state machine,s whose main operations are:to do the following

- compute the input bit according to the tap sequence;
- shift the bit pattern;
- register the output bit; and
- insert the bit computed in the first step into the input bit position.

Maximum length sequences (m-sequences) are composed by the output bits of a LFSR. They are, by definition, the largest codes that can be generated by a LFSR for a given tap sequence.

The tap sequence defines which bits in the current state will be combined to determine the input bit for the next state, generally using module-2 addition (*exclusive or*). Tap sequences can be represented by irreducible polynomials over $\mathbf{GF}(2)$ (*Galois Field* of order 2) – polynomials with coefficients of either 0 or 1, which cannot be represented as the product of two or more polynomials. For example, $x^2 + 1$ is not irreducible over $\mathbf{GF}(2)$ because it can be factored into $x + 1$:

$$(x + 1)(x + 1) = x^2 + 2x + 1 \equiv x^2 + 1. \tag{5}$$

Considering an irreducible polynomial, the corresponding tap sequence is given by the exponents whose polynomial coefficients are 1 (Dinan and Jabbari, 1998). M-sequences are binary pseudo-random sequences of length $(b^n - 1)$, where n is the number of elements of the tap sequence, and $b = 2, 3$ *or* 5.

The autocorrelation function of an m-sequence, $\Phi_{mls}(k)$, is given by

$$\Phi_{mls}(k) = \begin{cases} b^n - 1 & \text{for} \quad k = 0, \\ -1 & \text{for} \quad k \neq 0, \end{cases} \tag{6}$$

where k is the lag of correlation. In spite of the good autocorrelation properties, m-sequences, in general, are not immune to cross-correlation problems, and they may have large and unpredictable cross-correlation values. However, the so-called preferred pairs of m-sequences have cross-correlation functions which might assume the predicted values, -1, $-1 + p$, and $-1 - p$, where $p = 2^{(n+1)/2}$ for n odd or $p = 2^{(n+2)/2}$ for n even. Given a $(2^n - 1)$–length m-sequence, $a(k)$, with $gcd\{n, 4\} = 1$, its preferred pair is the result of decimation computed by applying on $a(k)$ a circular shift of q samples, where $q = 2m + 1$ and $gcd\{m, n\} = 1$. Figure 1 shows the autocorrelation of an m-sequence on the top and its cross-correlation with its preferred pair computed with $m = 5$.

The number of possible preferred pairs of m-sequences is limited, when compared to the requirements of practical applications of wireless communication. Preferred pairs of m-sequences, however, are used to generate Gold codes (Dinan and Jabbari, 1998).

In CDMA, Gold codes are used as chipping sequences that allow several callers to use the same frequency, resulting in less interference and better utilization of the available

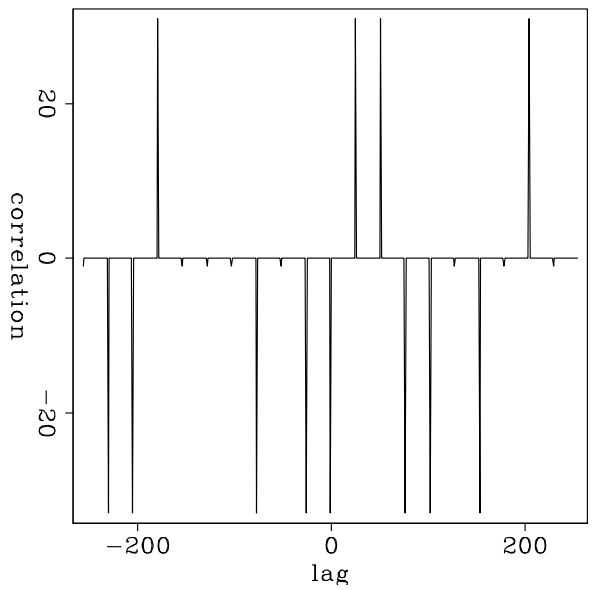


Figure 1: Correlation functions of preferred pairs of m-sequences. Top: autocorrelation of an m-sequence. Bottom: cross-correlation of a preferred pair of m-sequence. claudio1/. prefpairs

bandwidth. As originally proposed by Gold (1967), Gold codes can be computed by module-2 addition (*exclusive or*) of circularly shifted preferred pairs of m-sequences of length $2^n - 1$. The autocorrelation function of a Gold sequence, $\Phi_{gc}(k)$, is given by

$$\Phi_{gc}(k) = \begin{cases} \pm 2^n - 1 & \text{for} \quad k = 0, \\ \pm 1 & \text{for} \quad k \neq 0. \end{cases} \tag{7}$$

More interestingly, the three valued cross-correlation function of Gold sequences, $\Psi_{gc}(k)$, is given by

$$\Psi_{gc}(k) = \begin{cases} \pm(2^n - 1) & \text{for} \quad k = \lambda, \\ \mp 1 & \text{for} \quad k \neq \lambda. \end{cases} \tag{8}$$

where the correlation lag λ is given by the difference between the number of circular shifts applied to the m-sequence to compute the Gold code.

Figure 2 illustrates the correlation properties of the Gold codes. The left part shows the autocorrelation of the Gold code generated with one circular shift of the preferred pair of m-sequence. The right part shows the cross-correlation of the Gold code generated with one circular shift of the preferred pair of m-sequence with that generated with 84 circular shifts. The peak of the cross-correlation occurs at lag 84. For comparison, we show the correlation functions of conventional random codes in Figure 3. The autocorrelation function presents nearly periodic side lobes with additive low-amplitude random variations. The peak-to-side lobe ratio is around 30. The cross-correlation function is pseudo-random, and its amplitudes are of the same order of magnitude as those of the non-zero lags of the autocorrelation function.

After computing Gold codes, we use their phase information to encode the modeling experiments.

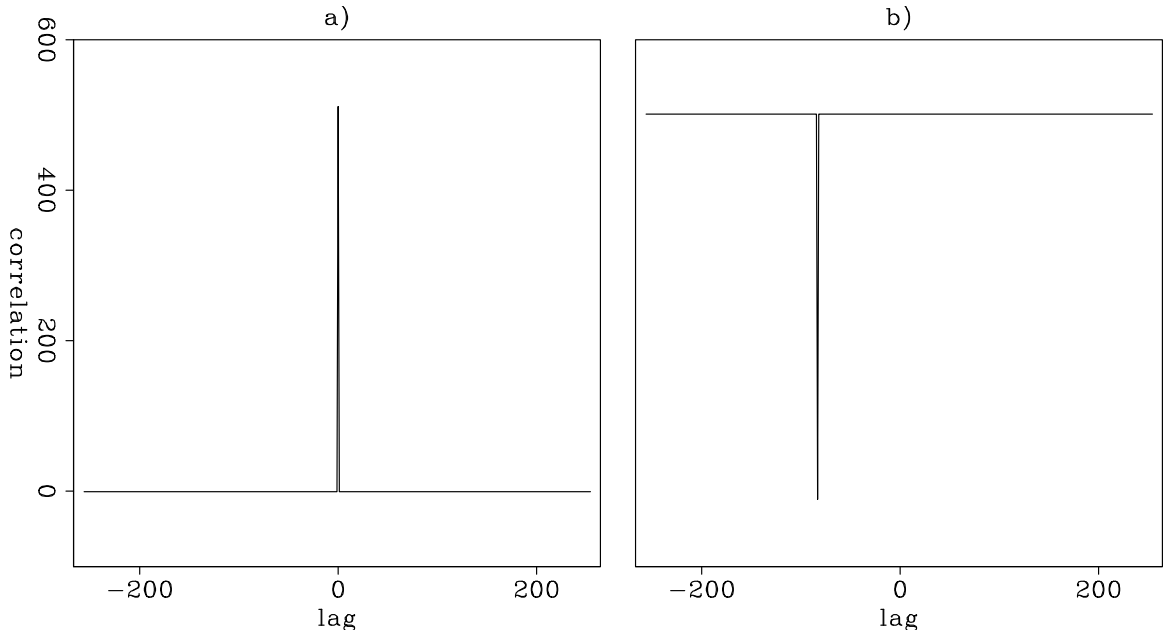


Figure 2: Correlation functions of Gold codes. a) Autocorrelation of Gold code generated with one circular shift of the preferred pair of m-sequence. b) Gold code generated with one circular shift of the preferred pair of m-sequence cross-correlated with that generated with 84 circular shifts. The peak of the cross-correlation occurs at lag 84. claudio1/. gold184

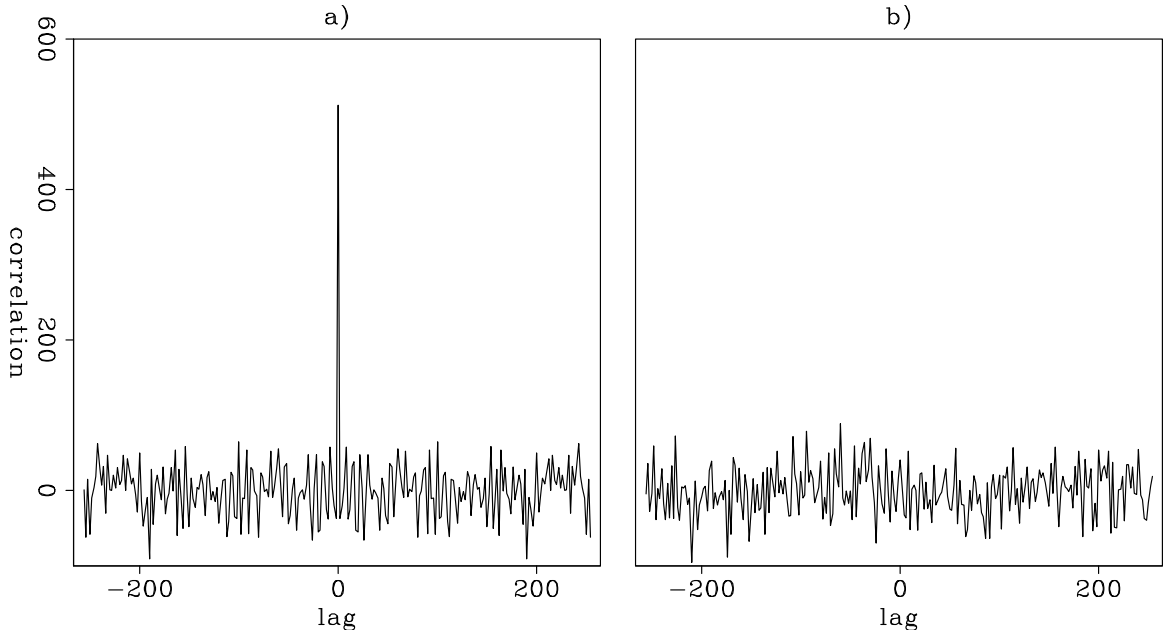


Figure 3: Correlation functions of conventional random codes. a) Autocorrelation of the conventional random code. b) A cross-correlation. claudio1/. rand

EXAMPLES

We illustrate the use of the encoding methods on a simple model of a flat reflector embedded medium with a constant velocity of 2 km/s. The original data is migrated with a 5% slower velocity. We used the same slower velocity to perform the modeling and the areal shot migration. Super-areal data are comprised of the collection of 10 modeling experiments initiated at every 10th CMP coordinate.

Figure 4 shows the areal shot migration of data generated by the prestack exploding-reflector modeling with no phase encoding applied. The panel on the left is the zero-subsurface-offset section, and the panel on the right is a SODCIG. The super-areal data, input to areal shot migration, are comprised of modeling experiments initiated at every tenth SODCIG. The SODCIGs resulting from the areal shot migration, show strong type-2 crosstalk. The crosstalk is periodic with half of the spacing of the modeling experiments in a super-areal shot.

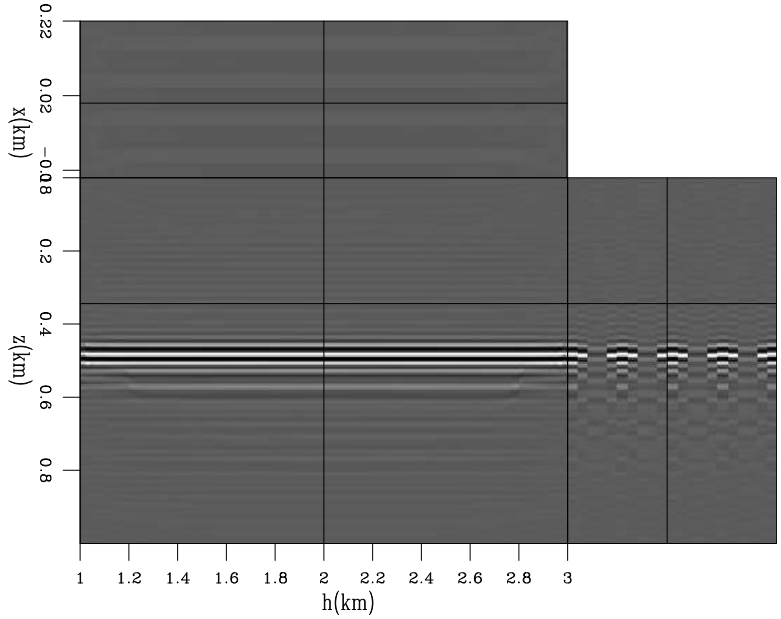


Figure 4: Areal shot migration of synthesized data with no phase encoding applied. The super-areal data comprises 10 modeling experiments. Notice the crosstalk in the SODCIG. claudio1/. perma

Figure 5 shows the areal shot migration of data generated by the prestack exploding-reflector modeling with no combination of modeling experiments into super-areal shots. This result represents the ideal image we would obtain if the crosstalk could be eliminated. Our objective in phase encoding the modeling experiments is to achieve satisfactory crosstalk attenuation in such a way that the moveout information is not altered.

Figure 6 shows areal shot migration of one realization of phase encoding modeling with conventional random codes. The strong crosstalk observed in Figure 5 is now dispersed throughout the image. The dispersed crosstalk can be further attenuated by migrating more random realizations, but this increases the cost of migration. Figure 7 shows the migration of 5 realizations of conventional random encoding modeling.

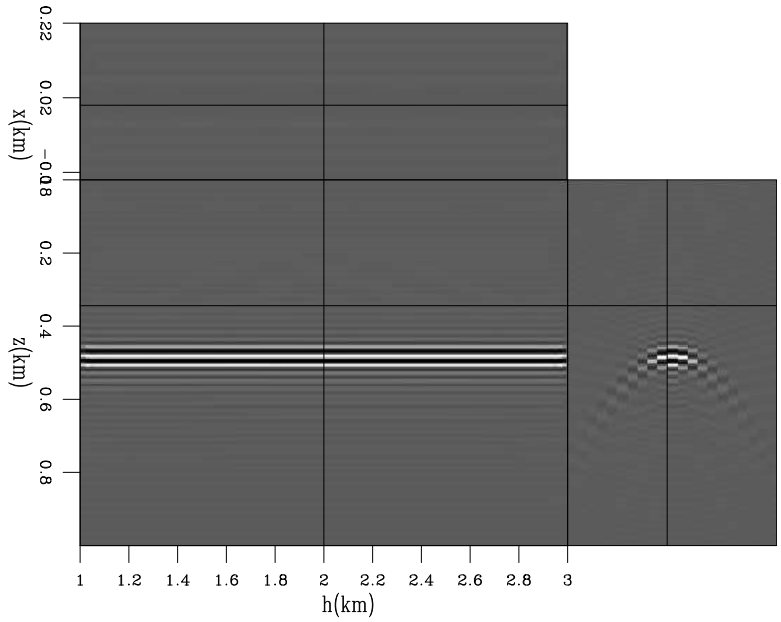


Figure 5: Areal shot migration of synthesized data with no combination of the modeling experiments into super-areal data. claudio1/. perm0

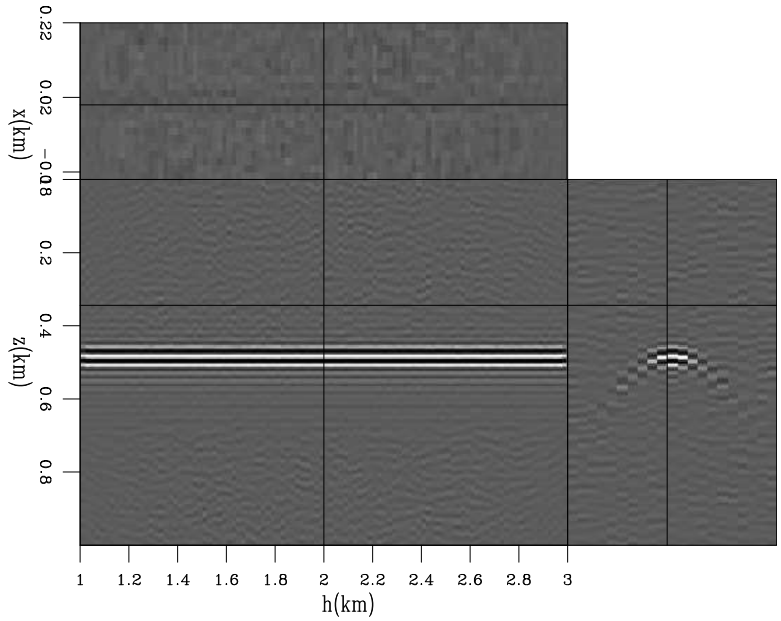


Figure 6: Areal shot migration of one realization of synthesized data with conventional random-phase encoding. claudio1/. conv1r

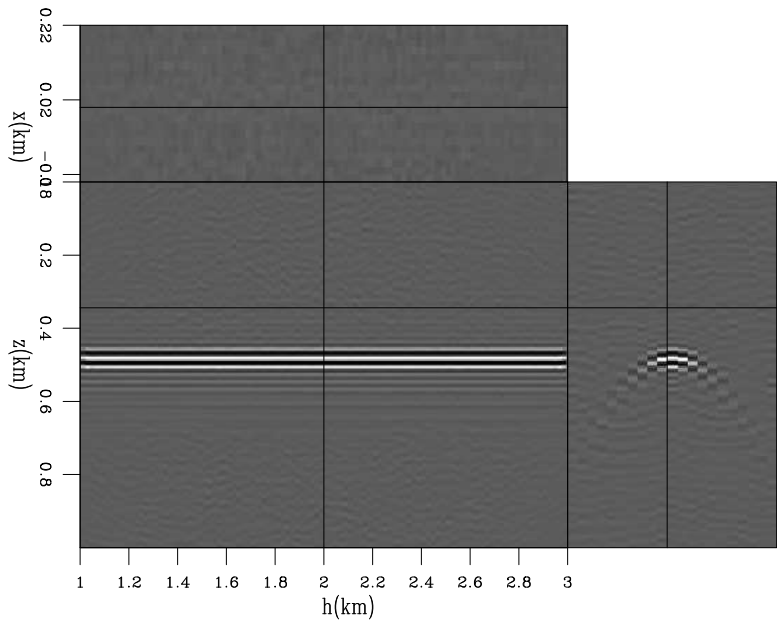


Figure 7: Areal shot migration of five realizations of synthesized data with conventional random-phase encoding. claudio1/. conv5r

Figure 8 shows the migration of one realization of data encoded with codes computed by optimization, and Figure 9 shows the migration of five realizations. No striking difference exists when comparing these results with those of the conventional random codes in Figures 6 and 7, respectively. Both encoding methods disperse the crosstalk energy in a similar manner. (more to come ...)

Crosstalk attenuation is incomplete when using conventional or optimized random codes because their autocorrelations are not perfect spike, nor are their cross-correlations zero everywhere. Gold codes partially satisfy these requirements: the autocorrelation is almost a perfect spike, except for -1's at lags different from zero, and, similarly, the cross-correlations are -1 everywhere except where they peak. Therefore, to obtain good results when using Gold codes, it is critical to select, among the available codes, the ones which provide the best crosstalk attenuation. That is because the cross-correlation functions have peaks with the same magnitude as those of the autocorrelation function.

In Figure 10, the areal shot migration was performed on encoded data for which the Gold codes were selected sequentially. This means that when applying the multi-offset imaging condition, cross-correlation peaks can occur at a distance less than that enclosed by the SODCIG. The crosstalk is present everywhere in the SODCIG and potentially obscures the moveout information for migration velocity analysis. It appears also as 'ghost' reflectors in the zero-subsurface-offset section. It is interesting to notice that the crosstalk in Figure 10 still has its apexes at the same offsets as those present in Figure 5. However, the apexes are displaced in depth. In particular, the crosstalk that occurs close to the edge of the SODCIG in Figure 5 is almost completely shifted out of the SODCIG. The amount of shift is defined by the lag where the cross-correlation peaks. In this example, the super-areal data comprises Gold-encoded modeling experiments in which cross-correlation peaks are

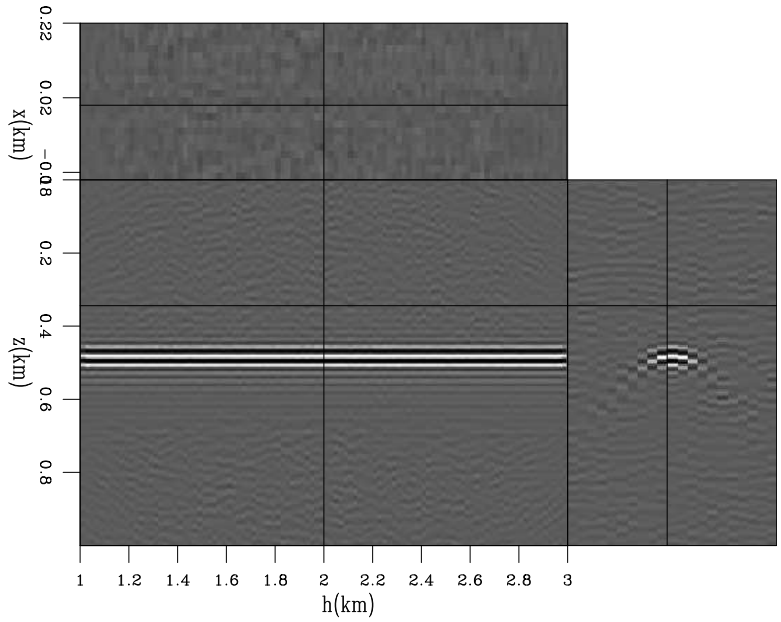


Figure 8: Areal shot migration of one realization of synthesized data with optimized random-phase encoding. claudio1/. opti1

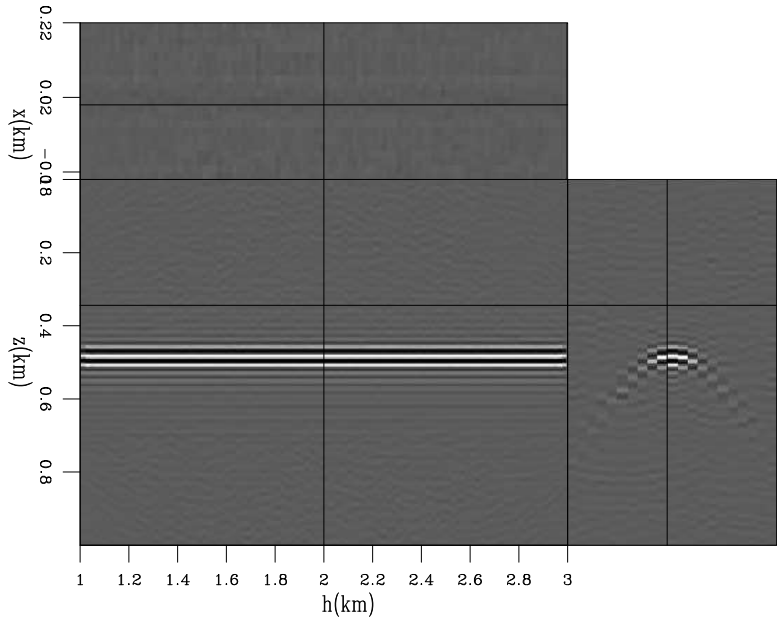


Figure 9: Areal shot migration of five realizations of synthesized data with optimized random-phase encoding. claudio1/. opti5

separated by multiples of 10 lags.

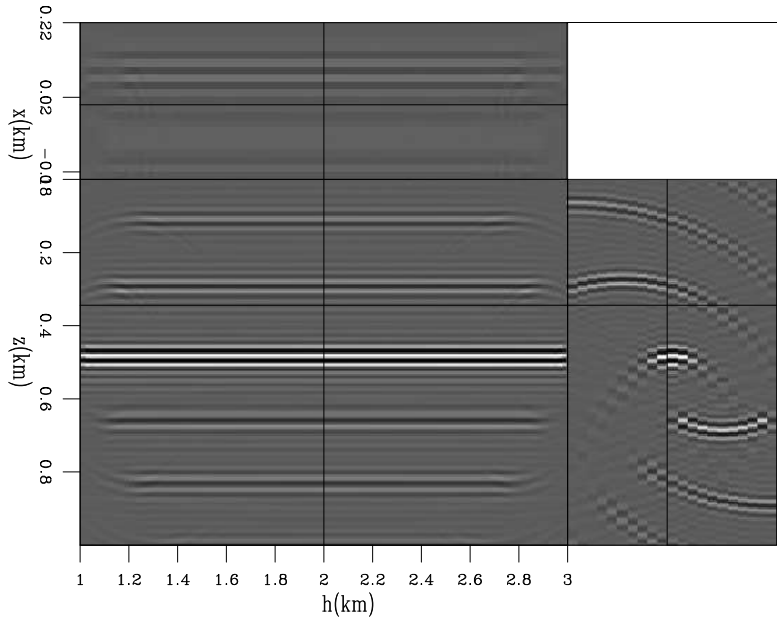


Figure 10: Areal shot migration of one realization of synthesized data with Gold phase encoding. Gold codes are selected sequentially such that the cross-correlation peak separation is multiple of 10. claudio1/. gold1x

For the simple case of constant velocity, the depth shifts, δz, are given by

$$\delta z = \frac{0.5 * v * n_\lambda}{n_\omega * d_\omega} \tag{9}$$

where n_ω is the number of frequencies, d_ω is the frequency interval and n_λ is the lag where the cross-correlation of the Gold codes peaks. For the present example, this amounts to $\delta z = (0.021 * n_\lambda)$km.

One possibility to statistically attenuate the crosstalk is to randomly choose the Gold codes. Figure 11 shows the areal shot migration of one realization of encoded data with randomly selected Gold codes. The crosstalk shows different patterns than that of the sequentially selected Gold codes.

As before, migrating more realizations of randomly selected, Gold-encoded data further attenuates crosstalk. Figure 12 shows the migration of 5 realizations of randomly selected, Gold-encoded data. Comparison with Figure 11 shows that much of the remaining crosstalk energy has been attenuated.

Considering that the crosstalk is shifted in depth, as can be seen in Figure 10, and that the amount of shift is determined by the lag where the cross-correlation peaks (equation 9), one can choose the Gold codes according to a suitable interval that completely shifts the crosstalk out of the SODCIG. This strategy shares similar idea as the linear-phase encoding of (Romero et al., 2000) which aims to shift the crosstalk out of the migration domain by using a linear function of frequency. Figure 13 shows the areal shot migration of data encoded by selecting every 50th Gold code, meaning that the depth shifts are multiples of 1 km. The crosstalk is virtually eliminated. Compare with Figure 5.

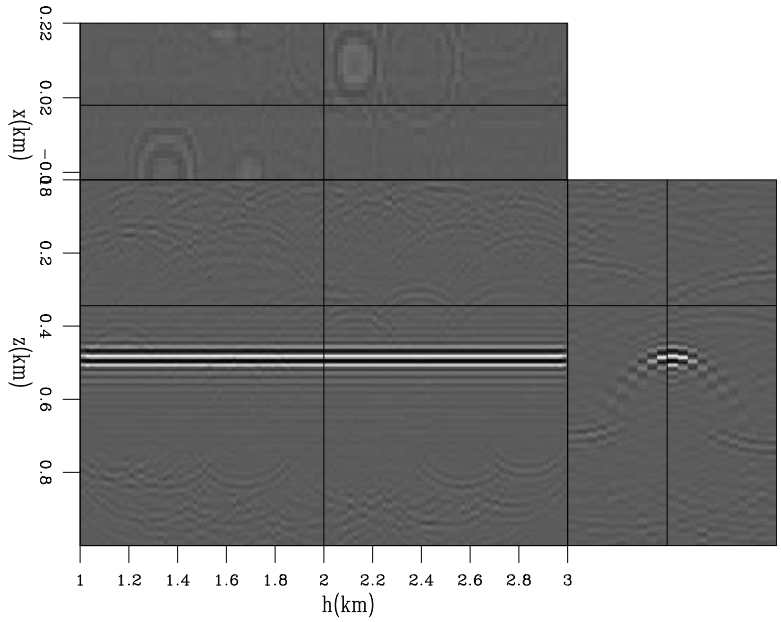


Figure 11: Areal shot migration of one realization of synthesized data with Gold phase encoding. Gold codes are randomly selected. claudio1/. gold1r

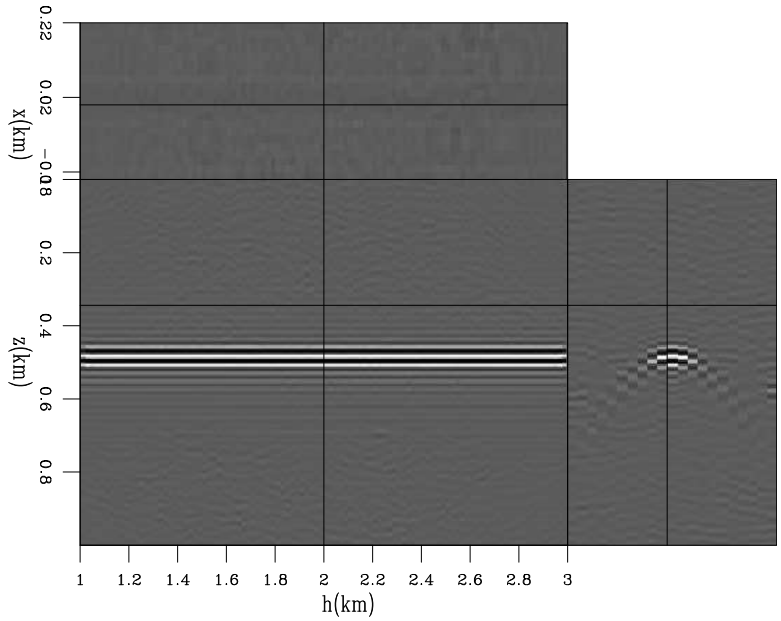


Figure 12: Areal shot migration of five realizations of synthesized data with Gold phase encoding. Gold codes are randomly selected. claudio1/. gold5r

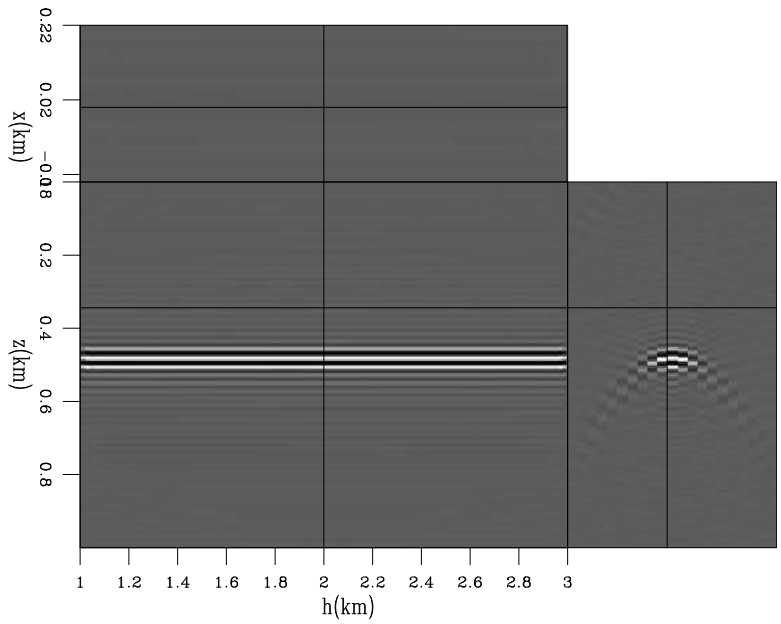


Figure 13: Areal shot migration of one realization of synthesized data with Gold phase encoding. Gold codes are selected such that the crosstalk is shifted out of the migration domain. claudio1/. gold1o

Actually, Gold phase encoding is equivalent to the linear phase encoding introduced by (Romero et al., 2000). The cross-correlation function of Gold codes is given by equation 8. It corresponds to spike of amplitude $\pm 2^n$ at lag λ, which depends on the difference between the number of circular shifts applied to the m-sequences to compute the Gold codes, plus a DC term, ∓ 1. The phase of the cross-correlation function is given by the phase of the spike, $t_\lambda \omega$, and is equal to the phase difference of the input signals. So, if Gold codes have phases $\gamma_1(\omega)$ and $\gamma_2(\omega)$, the phase of their cross-correlation is

$$\gamma_1(\omega) - \gamma_2(\omega) = t_\lambda \omega. \tag{10}$$

Equation 10 is the same as equation 26 of Romero's paper.

This strategy can be less computationally demanding than using conventional random codes, because just one realization is necessary to achieve a good result, while even using more realizations of conventional random encoding does not produce an image with similar quality. Its use in more complex situations, where the shift in depth of the crosstalk is not easily predicted, however, deserves more research.

CONCLUSION

We showed that the use of Gold codes in phase encoding can virtually eliminate crosstalk if the codes are satisfactorily selected. Its advantage over conventional random codes is not only in image quality but also cost. While the elimination of the crosstalk is achieved with only one realization of Gold codes, when using conventional random codes more realizations are needed to obtain reasonable images, increasing the cost of modeling and migration. We also showed the equivalence of Gold phase encoding and linear phase encoding.

We used a simple example with just one reflector to illustrate the encoding techniques. The performance of the Gold codes in the presence of more than one reflector needs to be tested. However, as just one realization is sufficient to drastically reduce the crosstalk, they can be affordable to use in a horizon-based approach to the modeling in which each reflector is modeled separately.

REFERENCES

Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analyis: SEP-Report, **124**, 45–60.

———, 2007, Prestack modeling of image events for migration velocity analysis: SEP-131, 101–118.

Dinan, E. and B. Jabbari, 1998, Spreading codes for direct sequence CDMA and wideband CDMA cellular networks: IEEE Communications Magazine, **36**, 48–54.

Golay, M., 1961, Complementary sequences: IRE Trans. on Info.Theory, **IT-7**, 82–87.

Gold, R., 1967, Optimal binary sequences for spread spectrum multiplexing: IEEE Transactions on Information Theory, **14**, 619–621.

Gran, F., 2005, Spatio-temporal encoding in medical ultrasound imaging: PhD thesis, Technical University of Denmark.

Guerra, C. and B. Biondi, 2008, Prestack exploding reflector modeling: The crosstalk problem: SEP-134, 79–91.

Kasami, T., 1966, Weight distribution formula for some class of cyclic codes: Technical Report R-285.

Lehmer, D. H., 1951, Mathematical methods in large-scale computing units: Technical Report 39.

Levanon, N. and E. Mozeson, 2004, Radar Signals: John Wiley & Sons.

Romero, L., D. Ghiglia, C. Ober, and S. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.

Shi, Z. and C. Schelgel, 2003, Spreading code construction for CDMA: IEEE Communications letters, **1**, 4–6.

Tseng, C., 1972, Complementary sets of sequences: IEEE Trans. on Info.Theory, **IT-18**, 644–652.

New phase-encoding schemes for wave-equation migration

Claudio Guerra and Biondo Biondi

ABSTRACT

Prestack exploding-reflector modeling aims to synthesize a small dataset comprised of areal shots, while preserving the correct kinematics to be used in iterations of migration velocity analysis. To achieve this goal, the modeled areal data must be combined into sets. However, crosstalk arises during migration due to the correlation of unrelated wavefields. Phase encoding the modeling experiments, can attenuate crosstalk during migration. In the geophysical community, the most used phase-encoding schemes are plane-wave-phase encoding and random-phase encoding. Here, we exploit the application of Gold codes commonly used in wireless communication, radar and medical imaging communities to phase encode data. We also introduce a method to compute random-phase functions by solving an optimization problem ... (more still to come). We show that adequately selecting the Gold codes can potentially shift the crosstalk out of the migration domain, yielding an image free of crosstalk.

INTRODUCTION

Biondi (2006, 2007) introduced the concept of the prestack exploding-reflector modeling. This method synthesizes source and receiver wavefields along the entire survey at the surface, in the form of areal data, starting from a prestack migrated image cube computed with wave-equation migration. For migration velocity analysis, the aim is to generate a considerably smaller dataset than the one used in the initial migration, while maintaining the necessary kinematic information

Conceptually, the synthesized areal data are computed by upward propagating source and receiver wavefields, using as initial conditions subsurface-offset-domain common-image gathers (SODCIG). To decrease the number of experiments to migrate, we take advantage of the linearity of the wave propagation to combine several experiments into a set of composite records. Combining several experiments, though, gives rise to crosstalk during imaging (Biondi, 2006; Guerra and Biondi, 2008). Guerra and Biondi (2008) use pseudo-random-phase encoding (Romero et al., 2000) during the modeling step to attenuate crosstalk. For conciseness reason, from now on we use the term random instead of pseudo-random.

The exploration geophysics community employs pseudo-random codes using intrinsic functions specific to the programing language. These pseudo-random codes present, generally, a uniform distribution and many of the algorithms are based on the *multiplicative congruential algorithm*, invented by Lehmer (1951). Their autocorrelation and cross-correlation functions have no special properties. The autocorrelation function presents nearly periodic side lobes with additive low-amplitude random variations. The peak-to-side lobe ratio is around 30. The cross-correlation function is pseudo-random, and its amplitudes are of

the same order of magnitude as those of the non-zero lags of the autocorrelation function. Herein, theses codes are called conventional random codes.

In wireless communication, specially for systems using Code Division Multiple Access (CDMA), different pseudo-random codes have been widely used (Shi and Schelgel, 2003). Medical imaging (Gran, 2005) and radar communities (Levanon and Mozeson, 2004) also exploit the statistical properties of these pseudo-random codes to increase bandwidth, signal-to-noise ratio and pulse compression. These codes are binary sequences and have unique autocorrelation and cross-correlations properties which make them more suited to achieving the above-mentioned objectives with minimal crosstalk. The autocorrelation function is represented by a large peak, whose amplitude equals the number of samples in the code, and the cross-correlation peaks, at non-zero lags, with the same amplitudes as that of the autocorrelation. Examples of binary pseudo-random codes used by these communities are Golay (Golay, 1961; Tseng, 1972), Kasami (Kasami, 1966) and Gold codes (Gold, 1967). Here, we exploit the properties of the Gold codes to encode seismic data.

In addition to using Gold codes, we introduce a novel method to compute phase functions by solving an optimization problem ...

In the next section we give a brief description of the prestack exploding-reflector modeling. Then we discuss how to compute the Gold codes and the optimized random-phase functions. We illustrate the application of each phase-encoding scheme by migrating phase-encoded areal data synthesized by prestack exploding-reflector modeling.

PRESTACK EXPLODING-REFLECTOR MODELING

Starting from a prestack image obtained by wave-equation migration, areal source and receiver wavefields are modeled at the surface by

$$\begin{aligned} S(x,\omega) &= G(z_\xi, x_\xi - h_\xi; x, z=0, \omega) * I_s(z_\xi, x_\xi, h_\xi), \\ R(x,\omega) &= G(z_\xi, x_\xi + h_\xi; x, z=0, \omega) * I_r(z_\xi, x_\xi, h_\xi), \end{aligned} \tag{1}$$

where $S(x,\omega)$ is the source wavefield at $z=0$; $R(x,\omega)$ is the receiver wavefield at $z=0$; $I_s(z_\xi, x_\xi, h_\xi)$ and $I_r(z_\xi, x_\xi, h_\xi)$ are the prestack images used as initial conditions for the source and receiver wavefield extrapolation, respectively, at a selected position, x_ξ,; $G(z_\xi, x_\xi \pm h_\xi; x, z=0, \omega)$ represents the operator that extrapolates the wavefields from the subsurface to the surface; h_ξ is the subsurface offset; z_ξ is depth; ω is the temporal frequency and x is the spatial coordinate in the data space coinciding with x_ξ. The prestack images used as initial conditions for the source and receiver wavefield extrapolation should be dip-independent gathers computed by changing the dip along the offset direction according to the apparent geological dip (Biondi, 2007). Here we use one-way extrapolators for both modeling and migration of the areal data.

By combining sets of individual modeling experiments into the same areal data, the amount of data input into migration can be significantly decreased. We achieve this by regularly selecting individual experiments and adding them up into their set, after being

upward propagated, according to

$$\begin{aligned}\widetilde{S_n}(x,\omega) &= \sum_{n=1}^{k}\sum_{i=n,k,N} S_i(x,\omega),\\ \widetilde{R_n}(x,\omega) &= \sum_{n=1}^{k}\sum_{i=n,k,N} R_i(x,\omega),\end{aligned} \tag{2}$$

where $\widetilde{S_n}(x,\omega)$, and $\widetilde{R_n}(x,\omega)$ contain k sets of summed areal sources and areal receivers, respectively; N is the number of SODCIGs. Every kth areal datum is selected to compose one set. Pairs of $\widetilde{S_k}(x,\omega)$ and $\widetilde{R_k}(x,\omega)$ are to be used as the areal source and the areal receiver wavefields, respectively, in areal shot migration.

Migration of the combined areal data produces an image by cross-correlating the combined areal source and receiver wavefields,

$$\widetilde{I_m}(z_\xi, x_\xi, h_\xi) = \sum_{\omega} \widetilde{S_m}^{\star}(z_\xi, x_\xi - h_\xi, \omega)\widetilde{R_m}(z_\xi, x_\xi + h_\xi, \omega), \tag{3}$$

where * represents complex conjugation. If $\widetilde{S_m}(x,\omega)$ and $\widetilde{R_m}(x,\omega)$ in equation 2 are comprised by two summed areal shots, the image $\widetilde{I_m}(z_\xi, x_\xi, h_\xi)$ will be given by

$$\begin{aligned}\widetilde{I_m}(z_\xi, x_\xi, h_\xi) = I_1(z_\xi, x_\xi, h_\xi) + I_2(z_\xi, x_\xi, h_\xi) + \\ \sum_{\omega} S_1^{\star}(z_\xi, x_\xi - h_\xi, \omega)R_2(z_\xi, x_\xi + h_\xi, \omega) + \\ \sum_{\omega} S_2^{\star}(z_\xi, x_\xi - h_\xi, \omega)R_1(z_\xi, x_\xi + h_\xi, \omega).\end{aligned} \tag{4}$$

In equation 4, the last two summation terms represent the crosstalk.

Guerra and Biondi (2008) show that in the prestack exploding-reflector method the crosstalk has two distinct origins. The crosstalk that is prominent in the zero-subsurface offset section, which herein we call type-1, is due to the cross-correlation of events in the source wavefield with that in the receiver wavefields modeled by the same SODCIG. Crosstalk in non-zero subsurface offsets, which herein we call type-2, is related to cross-correlation of events in the source wavefields with those in the receiver wavefields modeled by different SODCIGs.

Guerra and Biondi (2008) introduce strategies to attenuate the crosstalk. Migration of (x, z, ω)-random-phase encoded data disperses the crosstalk energy throughout the image as a pseudo-random background noise. By adding more realizations of random-phase encoded areal data, the speckled noise is further attenuated.

OPTIMIZED RANDOM CODES

...

GOLD CODES

Before describing Gold codes it is useful to define maximum length sequences.

Linear feedback shift registers (LFSR) are called state machine,s whose main operations are:to do the following

- compute the input bit according to the tap sequence;
- shift the bit pattern;
- register the output bit; and
- insert the bit computed in the first step into the input bit position.

Maximum length sequences (m-sequences) are composed by the output bits of a LFSR. They are, by definition, the largest codes that can be generated by a LFSR for a given tap sequence.

The tap sequence defines which bits in the current state will be combined to determine the input bit for the next state, generally using module-2 addition (*exclusive or*). Tap sequences can be represented by irreducible polynomials over $\mathbf{GF}(2)$ (*Galois Field* of order 2) – polynomials with coefficients of either 0 or 1, which cannot be represented as the product of two or more polynomials. For example, $x^2 + 1$ is not irreducible over $\mathbf{GF}(2)$ because it can be factored into $x + 1$:

$$(x + 1)(x + 1) = x^2 + 2x + 1 \equiv x^2 + 1. \tag{5}$$

Considering an irreducible polynomial, the corresponding tap sequence is given by the exponents whose polynomial coefficients are 1 (Dinan and Jabbari, 1998). M-sequences are binary pseudo-random sequences of length $(b^n - 1)$, where n is the number of elements of the tap sequence, and $b = 2, 3$ *or* 5.

The autocorrelation function of an m-sequence, $\Phi_{mls}(k)$, is given by

$$\Phi_{mls}(k) = \begin{cases} b^n - 1 & \text{for} \quad k = 0, \\ -1 & \text{for} \quad k \neq 0, \end{cases} \tag{6}$$

where k is the lag of correlation. In spite of the good autocorrelation properties, m-sequences, in general, are not immune to cross-correlation problems, and they may have large and unpredictable cross-correlation values. However, the so-called preferred pairs of m-sequences have cross-correlation functions which might assume the predicted values, -1, $-1 + p$, and $-1 - p$, where $p = 2^{(n+1)/2}$ for n odd or $p = 2^{(n+2)/2}$ for n even. Given a $(2^n - 1)$–length m-sequence, $a(k)$, with $gcd\{n, 4\} = 1$, its preferred pair is the result of decimation computed by applying on $a(k)$ a circular shift of q samples, where $q = 2m + 1$ and $gcd\{m, n\} = 1$. Figure 1 shows the autocorrelation of an m-sequence on the top and its cross-correlation with its preferred pair computed with $m = 5$.

The number of possible preferred pairs of m-sequences is limited, when compared to the requirements of practical applications of wireless communication. Preferred pairs of m-sequences, however, are used to generate Gold codes (Dinan and Jabbari, 1998).

In CDMA, Gold codes are used as chipping sequences that allow several callers to use the same frequency, resulting in less interference and better utilization of the available

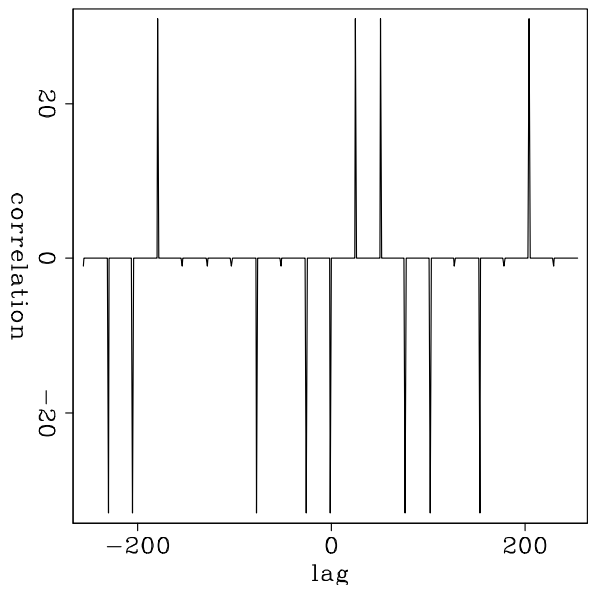


Figure 1: Correlation functions of preferred pairs of m-sequences. Top: autocorrelation of an m-sequence. Bottom: cross-correlation of a preferred pair of m-sequence. claudio1/. prefpairs

bandwidth. As originally proposed by Gold (1967), Gold codes can be computed by module-2 addition (*exclusive or*) of circularly shifted preferred pairs of m-sequences of length $2^n - 1$. The autocorrelation function of a Gold sequence, $\Phi_{gc}(k)$, is given by

$$\Phi_{gc}(k) = \begin{cases} \pm 2^n - 1 & \text{for} \quad k = 0, \\ \pm 1 & \text{for} \quad k \neq 0. \end{cases} \tag{7}$$

More interestingly, the three valued cross-correlation function of Gold sequences, $\Psi_{gc}(k)$, is given by

$$\Psi_{gc}(k) = \begin{cases} \pm(2^n - 1) & \text{for} \quad k = \lambda, \\ \mp 1 & \text{for} \quad k \neq \lambda. \end{cases} \tag{8}$$

where the correlation lag λ is given by the difference between the number of circular shifts applied to the m-sequence to compute the Gold code.

Figure 2 illustrates the correlation properties of the Gold codes. The left part shows the autocorrelation of the Gold code generated with one circular shift of the preferred pair of m-sequence. The right part shows the cross-correlation of the Gold code generated with one circular shift of the preferred pair of m-sequence with that generated with 84 circular shifts. The peak of the cross-correlation occurs at lag 84. For comparison, we show the correlation functions of conventional random codes in Figure 3. The autocorrelation function presents nearly periodic side lobes with additive low-amplitude random variations. The peak-to-side lobe ratio is around 30. The cross-correlation function is pseudo-random, and its amplitudes are of the same order of magnitude as those of the non-zero lags of the autocorrelation function.

After computing Gold codes, we use their phase information to encode the modeling experiments.

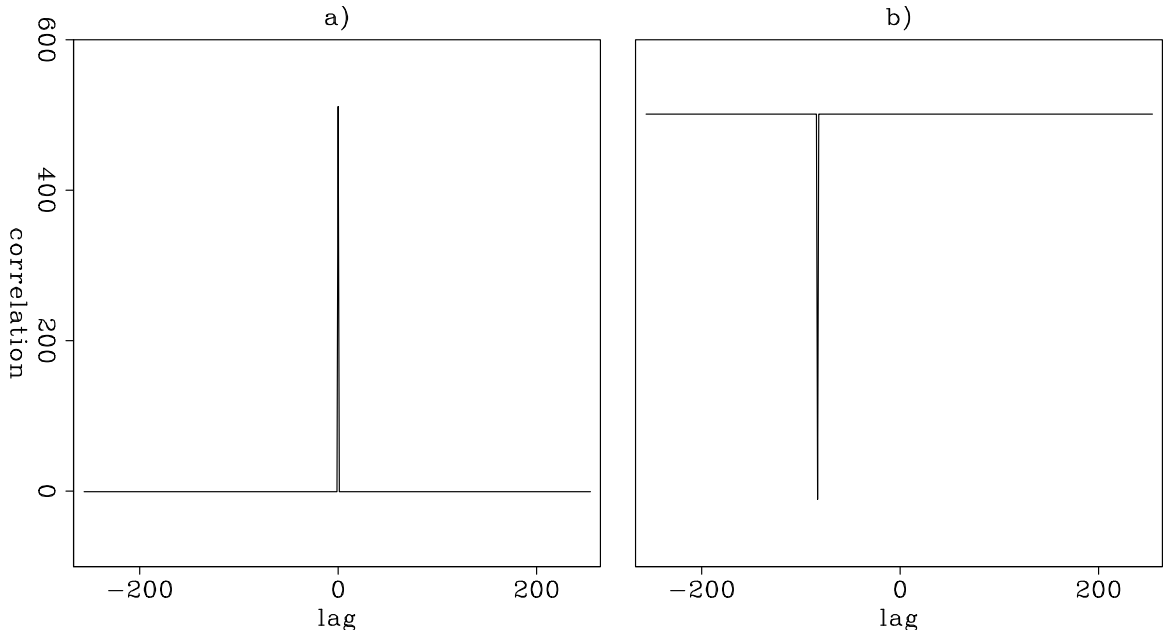


Figure 2: Correlation functions of Gold codes. a) Autocorrelation of Gold code generated with one circular shift of the preferred pair of m-sequence. b) Gold code generated with one circular shift of the preferred pair of m-sequence cross-correlated with that generated with 84 circular shifts. The peak of the cross-correlation occurs at lag 84. claudio1/. gold184

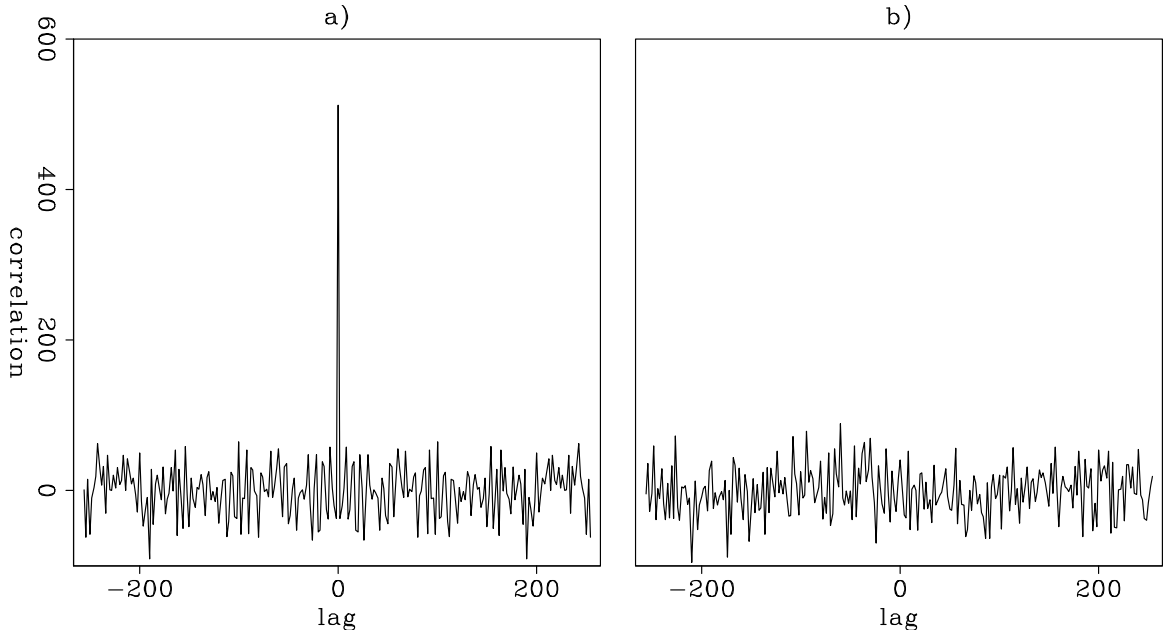


Figure 3: Correlation functions of conventional random codes. a) Autocorrelation of the conventional random code. b) A cross-correlation. claudio1/. rand

EXAMPLES

We illustrate the use of the encoding methods on a simple model of a flat reflector embedded medium with a constant velocity of 2 km/s. The original data is migrated with a 5% slower velocity. We used the same slower velocity to perform the modeling and the areal shot migration. Super-areal data are comprised of the collection of 10 modeling experiments initiated at every 10th CMP coordinate.

Figure 4 shows the areal shot migration of data generated by the prestack exploding-reflector modeling with no phase encoding applied. The panel on the left is the zero-subsurface-offset section, and the panel on the right is a SODCIG. The super-areal data, input to areal shot migration, are comprised of modeling experiments initiated at every tenth SODCIG. The SODCIGs resulting from the areal shot migration, show strong type-2 crosstalk. The crosstalk is periodic with half of the spacing of the modeling experiments in a super-areal shot.

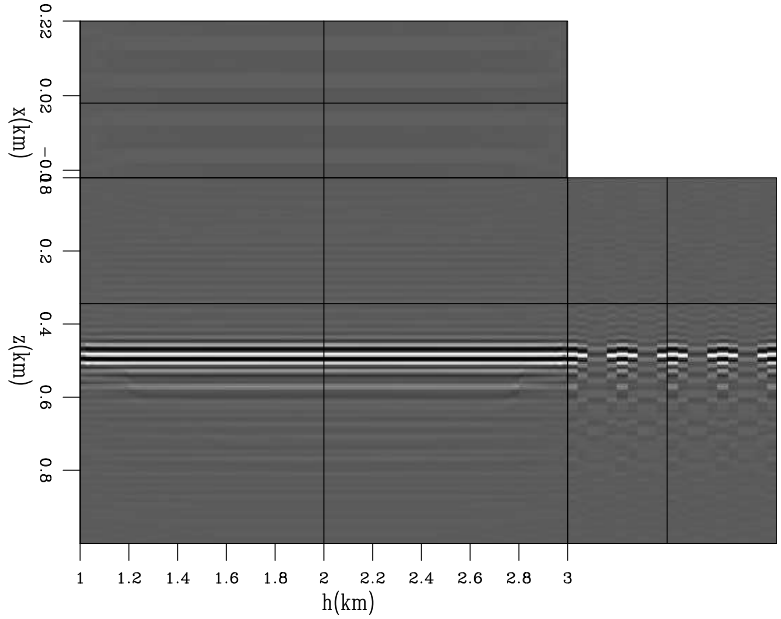


Figure 4: Areal shot migration of synthesized data with no phase encoding applied. The super-areal data comprises 10 modeling experiments. Notice the crosstalk in the SODCIG. claudio1/. perma

Figure 5 shows the areal shot migration of data generated by the prestack exploding-reflector modeling with no combination of modeling experiments into super-areal shots. This result represents the ideal image we would obtain if the crosstalk could be eliminated. Our objective in phase encoding the modeling experiments is to achieve satisfactory crosstalk attenuation in such a way that the moveout information is not altered.

Figure 6 shows areal shot migration of one realization of phase encoding modeling with conventional random codes. The strong crosstalk observed in Figure 5 is now dispersed throughout the image. The dispersed crosstalk can be further attenuated by migrating more random realizations, but this increases the cost of migration. Figure 7 shows the migration of 5 realizations of conventional random encoding modeling.

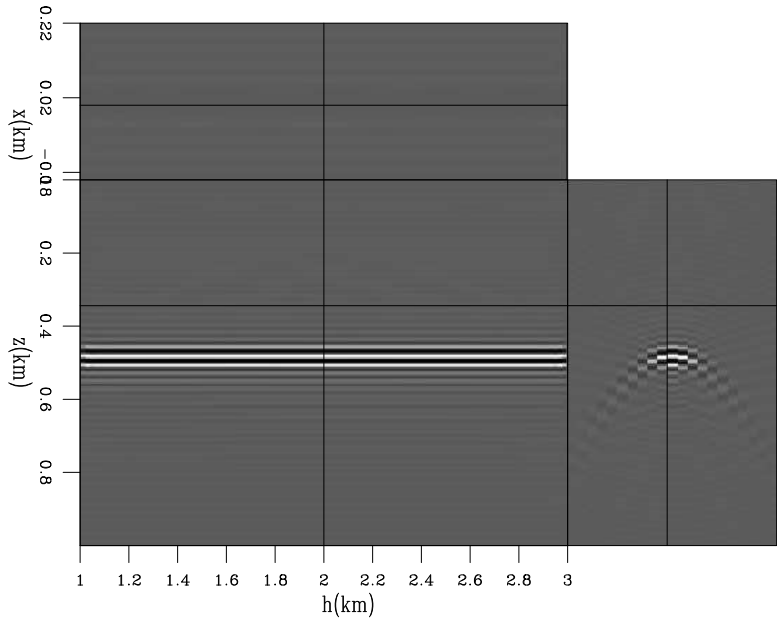


Figure 5: Areal shot migration of synthesized data with no combination of the modeling experiments into super-areal data. claudio1/. perm0

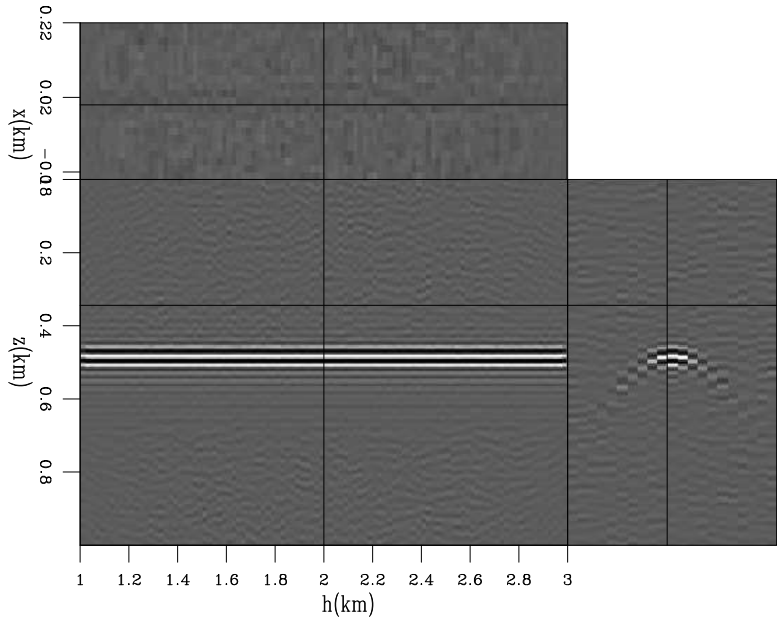


Figure 6: Areal shot migration of one realization of synthesized data with conventional random-phase encoding. claudio1/. conv1r

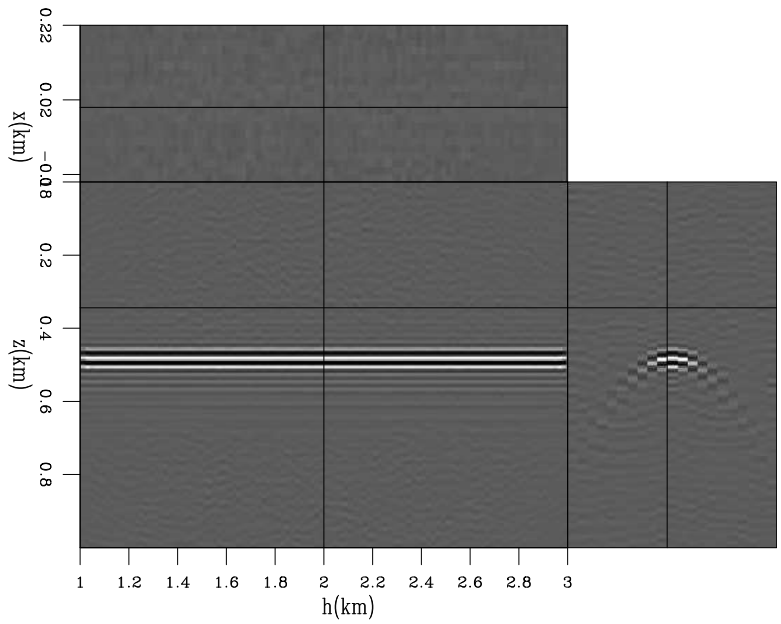


Figure 7: Areal shot migration of five realizations of synthesized data with conventional random-phase encoding. claudio1/. conv5r

Figure 8 shows the migration of one realization of data encoded with codes computed by optimization, and Figure 9 shows the migration of five realizations. No striking difference exists when comparing these results with those of the conventional random codes in Figures 6 and 7, respectively. Both encoding methods disperse the crosstalk energy in a similar manner. (more to come ...)

Crosstalk attenuation is incomplete when using conventional or optimized random codes because their autocorrelations are not perfect spike, nor are their cross-correlations zero everywhere. Gold codes partially satisfy these requirements: the autocorrelation is almost a perfect spike, except for -1's at lags different from zero, and, similarly, the cross-correlations are -1 everywhere except where they peak. Therefore, to obtain good results when using Gold codes, it is critical to select, among the available codes, the ones which provide the best crosstalk attenuation. That is because the cross-correlation functions have peaks with the same magnitude as those of the autocorrelation function.

In Figure 10, the areal shot migration was performed on encoded data for which the Gold codes were selected sequentially. This means that when applying the multi-offset imaging condition, cross-correlation peaks can occur at a distance less than that enclosed by the SODCIG. The crosstalk is present everywhere in the SODCIG and potentially obscures the moveout information for migration velocity analysis. It appears also as 'ghost' reflectors in the zero-subsurface-offset section. It is interesting to notice that the crosstalk in Figure 10 still has its apexes at the same offsets as those present in Figure 5. However, the apexes are displaced in depth. In particular, the crosstalk that occurs close to the edge of the SODCIG in Figure 5 is almost completely shifted out of the SODCIG. The amount of shift is defined by the lag where the cross-correlation peaks. In this example, the super-areal data comprises Gold-encoded modeling experiments in which cross-correlation peaks are

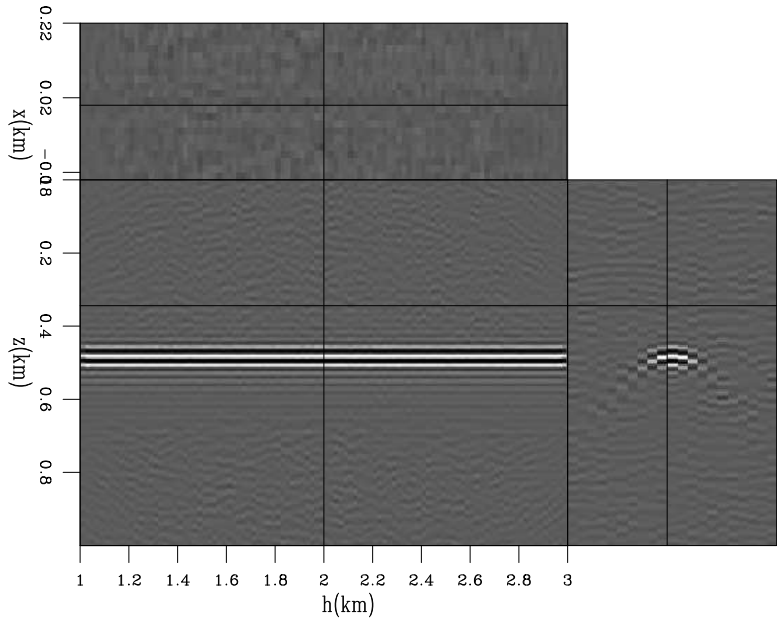


Figure 8: Areal shot migration of one realization of synthesized data with optimized random-phase encoding. claudio1/. opti1

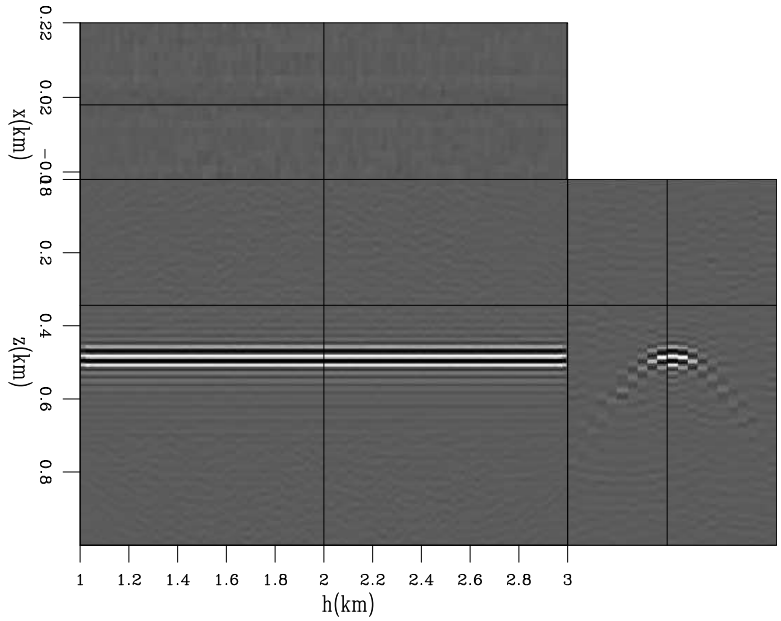


Figure 9: Areal shot migration of five realizations of synthesized data with optimized random-phase encoding. claudio1/. opti5

separated by multiples of 10 lags.

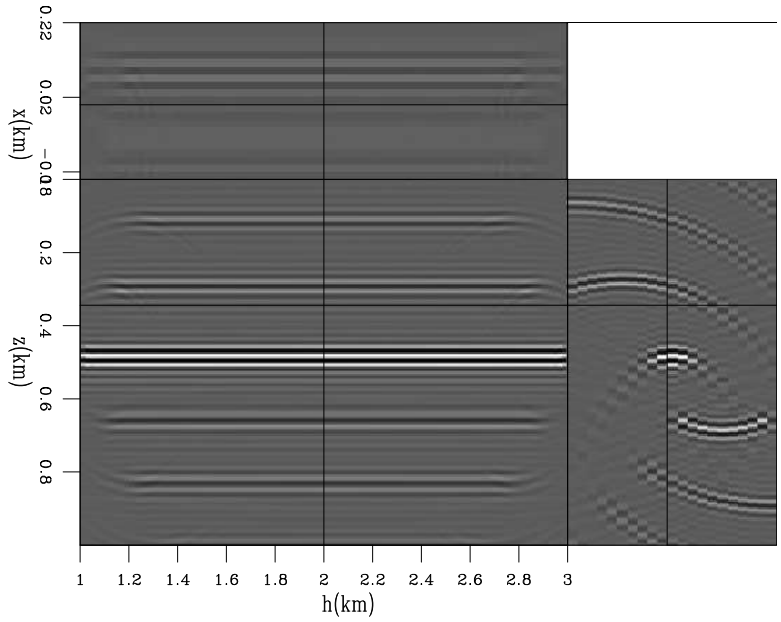


Figure 10: Areal shot migration of one realization of synthesized data with Gold phase encoding. Gold codes are selected sequentially such that the cross-correlation peak separation is multiple of 10. claudio1/. gold1x

For the simple case of constant velocity, the depth shifts, δz, are given by

$$\delta z = \frac{0.5 * v * n_\lambda}{n_\omega * d_\omega} \tag{9}$$

where n_ω is the number of frequencies, d_ω is the frequency interval and n_λ is the lag where the cross-correlation of the Gold codes peaks. For the present example, this amounts to $\delta z = (0.021 * n_\lambda)$km.

One possibility to statistically attenuate the crosstalk is to randomly choose the Gold codes. Figure 11 shows the areal shot migration of one realization of encoded data with randomly selected Gold codes. The crosstalk shows different patterns than that of the sequentially selected Gold codes.

As before, migrating more realizations of randomly selected, Gold-encoded data further attenuates crosstalk. Figure 12 shows the migration of 5 realizations of randomly selected, Gold-encoded data. Comparison with Figure 11 shows that much of the remaining crosstalk energy has been attenuated.

Considering that the crosstalk is shifted in depth, as can be seen in Figure 10, and that the amount of shift is determined by the lag where the cross-correlation peaks (equation 9), one can choose the Gold codes according to a suitable interval that completely shifts the crosstalk out of the SODCIG. This strategy shares similar idea as the linear-phase encoding of (Romero et al., 2000) which aims to shift the crosstalk out of the migration domain by using a linear function of frequency. Figure 13 shows the areal shot migration of data encoded by selecting every 50th Gold code, meaning that the depth shifts are multiples of 1 km. The crosstalk is virtually eliminated. Compare with Figure 5.

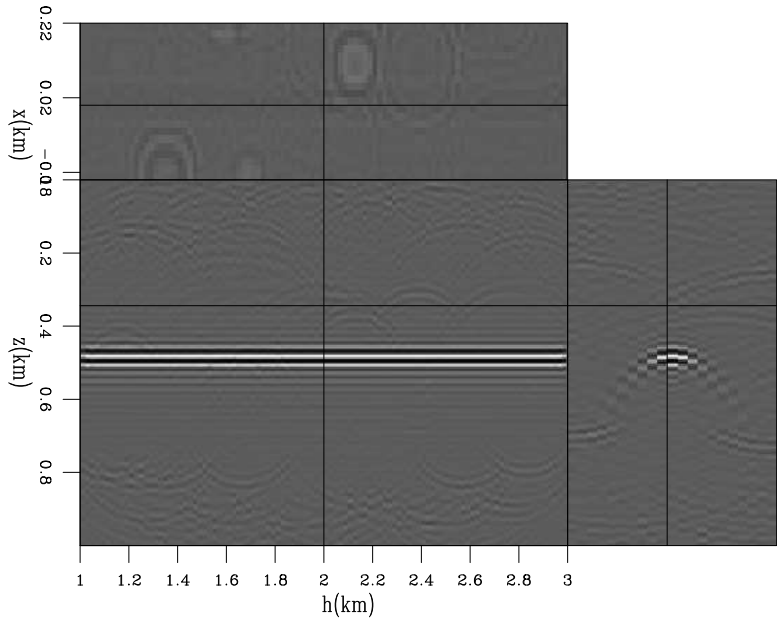


Figure 11: Areal shot migration of one realization of synthesized data with Gold phase encoding. Gold codes are randomly selected. claudio1/. gold1r

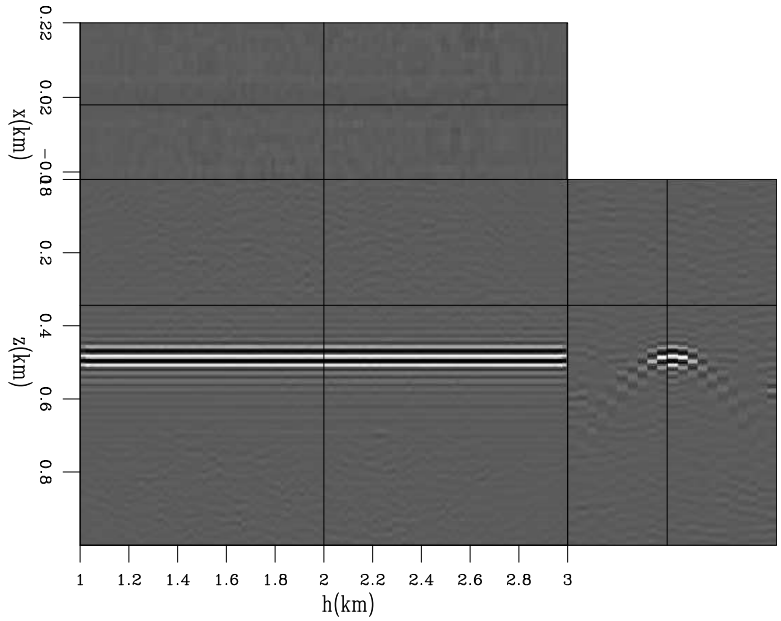


Figure 12: Areal shot migration of five realizations of synthesized data with Gold phase encoding. Gold codes are randomly selected. claudio1/. gold5r

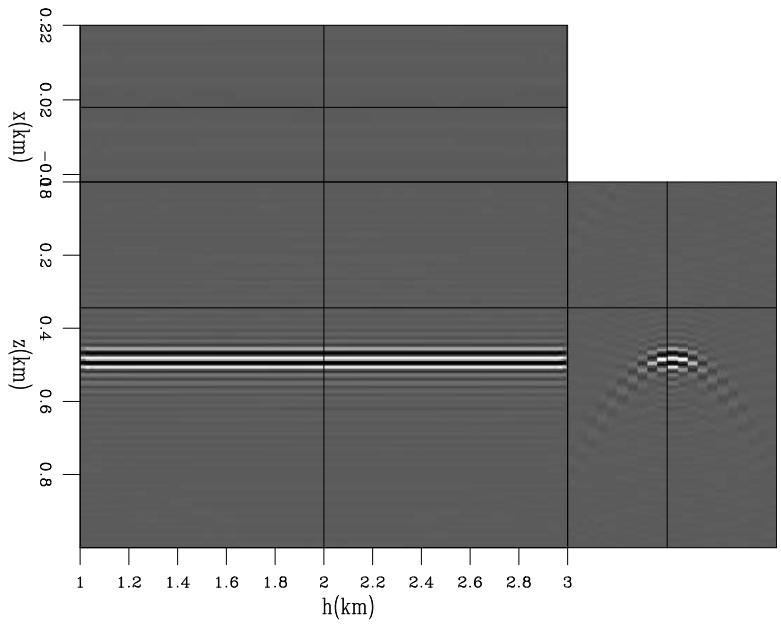


Figure 13: Areal shot migration of one realization of synthesized data with Gold phase encoding. Gold codes are selected such that the crosstalk is shifted out of the migration domain. claudio1/. gold1o

Actually, Gold phase encoding is equivalent to the linear phase encoding introduced by (Romero et al., 2000). The cross-correlation function of Gold codes is given by equation 8. It corresponds to spike of amplitude $\pm 2^n$ at lag λ, which depends on the difference between the number of circular shifts applied to the m-sequences to compute the Gold codes, plus a DC term, ∓ 1. The phase of the cross-correlation function is given by the phase of the spike, $t_\lambda \omega$, and is equal to the phase difference of the input signals. So, if Gold codes have phases $\gamma_1(\omega)$ and $\gamma_2(\omega)$, the phase of their cross-correlation is

$$\gamma_1(\omega) - \gamma_2(\omega) = t_\lambda \omega. \tag{10}$$

Equation 10 is the same as equation 26 of Romero's paper.

This strategy can be less computationally demanding than using conventional random codes, because just one realization is necessary to achieve a good result, while even using more realizations of conventional random encoding does not produce an image with similar quality. Its use in more complex situations, where the shift in depth of the crosstalk is not easily predicted, however, deserves more research.

CONCLUSION

We showed that the use of Gold codes in phase encoding can virtually eliminate crosstalk if the codes are satisfactorily selected. Its advantage over conventional random codes is not only in image quality but also cost. While the elimination of the crosstalk is achieved with only one realization of Gold codes, when using conventional random codes more realizations are needed to obtain reasonable images, increasing the cost of modeling and migration. We also showed the equivalence of Gold phase encoding and linear phase encoding.

We used a simple example with just one reflector to illustrate the encoding techniques. The performance of the Gold codes in the presence of more than one reflector needs to be tested. However, as just one realization is sufficient to drastically reduce the crosstalk, they can be affordable to use in a horizon-based approach to the modeling in which each reflector is modeled separately.

REFERENCES

Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analyis: SEP-Report, **124**, 45–60.

———, 2007, Prestack modeling of image events for migration velocity analysis: SEP-131, 101–118.

Dinan, E. and B. Jabbari, 1998, Spreading codes for direct sequence CDMA and wideband CDMA cellular networks: IEEE Communications Magazine, **36**, 48–54.

Golay, M., 1961, Complementary sequences: IRE Trans. on Info.Theory, **IT-7**, 82–87.

Gold, R., 1967, Optimal binary sequences for spread spectrum multiplexing: IEEE Transactions on Information Theory, **14**, 619–621.

Gran, F., 2005, Spatio-temporal encoding in medical ultrasound imaging: PhD thesis, Technical University of Denmark.

Guerra, C. and B. Biondi, 2008, Prestack exploding reflector modeling: The crosstalk problem: SEP-134, 79–91.

Kasami, T., 1966, Weight distribution formula for some class of cyclic codes: Technical Report R-285.

Lehmer, D. H., 1951, Mathematical methods in large-scale computing units: Technical Report 39.

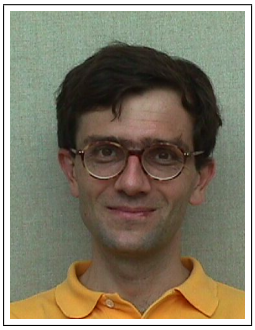
Levanon, N. and E. Mozeson, 2004, Radar Signals: John Wiley & Sons.

Romero, L., D. Ghiglia, C. Ober, and S. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.

Shi, Z. and C. Schelgel, 2003, Spreading code construction for CDMA: IEEE Communications letters, **1**, 4–6.

Tseng, C., 1972, Complementary sets of sequences: IEEE Trans. on Info.Theory, **IT-18**, 644–652.

Research Personnel



Biondo L. Biondi graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. SEG Outstanding Paper award 1994. During 1987, he worked as a Research Geophysicist for TOTAL, Compagnie Francaise des Petroles in Paris. After his Ph.D. at Stanford, Biondo worked for three years with Thinking Machines Co. on the applications of massively parallel computers to seismic processing. After leaving Thinking Machines, Biondo started 3DGeo Development, a software and service company devoted to high-end seismic imaging. Biondo is now Associate Professor (Research) of Geophysics and leads SEP efforts in 3-D imaging. He is a member of SEG and EAGE.

Claudio Guerra received his B.Sc. in Geology from Federal University of Rio de Janeiro, Brazil in 1988 and a M.Sc. from State University of Campinas, Brazil in 1999. Since 1989, he has been working for Petrobras, Brazil. He joined SEP in 2006 and is currently pursuing a Ph.D. in geophysics at Stanford University. He is member of SEG and SBGf.