Simultaneous multi-source acquisition using m-sequences

Joe Wong

ABSTRACT

Maximal length sequences, or m-sequences, are periodic mathematical periodic entities with values of -1 and 1. A single m-sequence can be used to construct multiple shifted m-sequences or Gold codes. Within the set of shifted m-sequences or derived Gold codes, each member has an autocorrelation that closely approximates a delta function but is correlated weakly to other members. These sequences and codes are known generally as pseudorandom binary sequences (PRBSs), and being weakly correlated to each other, they are suitable for driving multiple seismic vibrators simultaneously with little crosstalk between the vibrator sources. Operating multiple vibrators simultaneously in real-world seismic surveys increases acquisition productivity dramatically. I use numerical examples to show how signals from multiple vibrators simultaneously driven by weakly-correlated PRBSs (in particular, shifted m-sequences) can be separated into individual seismograms unique to each vibrator with minimal crosstalk.

INTRODUCTION

Recently, Pecholcs et al. (2010) reported on using 24 vibrators operating simultaneously in close proximity to acquire data for a land 3D seismic survey. The use of 24 simultaneous sources resulted in more than 40,000 shot points per 24-hour period and in an impressive increase in acquisition productivity. The multiple vibrators were controlled by a set of modified Gold codes. Sallas et al. (2011) and Sallas and Gibson (2008) have described using modified Gold codes as pilot signals to drive multiple vibrator sources simultaneously for surface reflection surveys. Maxwell et al. (2008) have presented details on the signal characteristics of large seismic vibrators driven by Gold codes. They claimed that an important advantage of using Gold codes instead of frequency sweeps as pilots to drive vibrators was the increase in effective low-frequency power.

Maximal-length sequences, or m-sequences, are well-defined mathematical periodic entities with values of -1 and +1. They simulate random noise in that they have autocorrelations that approximate the delta function. For this reason, they fall into a category called pseudorandom binary sequences (PRBSs). The mathematical theory behind m-sequences can be found in Watson (1962) and in Golomb and Gong (2005).

Gold codes are closely related to m-sequences. A set of related Gold codes can be derived from a pair of optimal m-sequences using the theory published by Gold (1967). Like m-sequences, each member of the set is periodic and binary-valued with values of -1 and 1. All members have autocorrelations that approximate the delta function, and so they are a variety of PRBS. They are weakly coupled under cross-correlation, i.e., the maxima and minima of the cross-correlations are not zero, but, as is shown in Appendix A, have predictable absolute values less than the autocorrelation peaks. In this sense, the Gold codes can be considered to form a quasi-orthogonal set (each member in a quasi-orthogonal set has an autocorrelation that approximates a delta function, but its cross-correlation with members produces values significantly less than the autocorrelation peak value).

Sets of m-sequences also can be constructed to be quasi-orthogonal under certain conditions. Quasi-orthogonality makes both types of PRBS suitable for controlling multiple vibrator sources simultaneously. It allows the combined signals from the simultaneous vibrators to be separated by cross-correlation into traces with little crosstalk between the signals produced by the individual sources.

We note that correlation involving periodic m-sequences and Gold codes is circular in nature, and involves wrap-around if only one period or cycle is referenced. In this report, correlation, convolution, and shifting are assumed to be circular.

Both m-sequences and Gold codes are used in diverse fields of science and engineering. Examples of applications are the estimation of impulse responses of linear systems, wireless communication, data encryption, and GPS/GNSS technology (Holmes, 2005). Gold codes are popular for wireless communications because thousands or even millions of weakly-correlated forms can be easily produced from different preferred (or optimal) pairs of m-sequences. In seismic acquisition, m-sequences have been used successfully in crosswell applications to drive piezoelectric vibrators (Hurley, 1983; Wong et al. 1983, 1987; Yamamoto, 1994; Wong, 2000). Information on the properties of m-sequences and Gold codes and generating algorithms is presented in a companion CREWS Research Report (Wong, 2012).

M-SEQUENCES FOR DATA ACQUISITION

In the introduction, it was pointed out that Gold codes have been used to drive multiple vibrators simultaneously for land-based seismic data acquisition. In some situations involving simultaneous-source seismic acquisition, where the number of simultaneous sources is limited to 32 or less, and where there is need to detect very weak signals, the simpler m-sequences may be superior alternatives to Gold codes, because, within limited ranges of correlation lags, their autocorrelations are better approximations of the delta function than those of Gold codes of similar length (see Appendix A). In addition, the crosstalk between m-sequences designed driving for multiple simultaneous sources can be much less than that produced by Gold codes. I present examples and results showing how m-sequence PRBSs can be used effectively to drive multiple vibrators simultaneously for high-efficiency seismic data acquisition.

Defining shifted m-sequences

Mathematically, an m-sequence is characterized by its degree *m*, its fundamental length$L$, and its base period t_b. The fundamental length L is related to the degree m bye Equation 1:

$$L = 2^m - 1\,, \tag{1}$$

where *m* is an integer. The exponential form of Equation 1 means that the fundamental length grows very quickly with the degree. For practical applications, we limit m to values of 5 to 20. If *m* = 11, then *L* = 2047; when *m* = 15, *L* = 8191.

In terms of real time, the m-sequence is periodic with period equal to

$$T_m = L \cdot t_b \; . \tag{2}$$

For practical application, we include two further parameters: the sample time t_s and the shift time t_{shift}. A time series version of the m-sequence is obtained when it is sampled at the sample time t_s. The ratio of the base period to sample time length of the sampled m-sequence is the up-sample ratio:

$$r = t_b / t_s \;\; , \tag{3}$$

where r typically have integer values of 1, 2, 4, 8, or 16.

After constructing a sampled m-sequence S_1 with these parameters, multiple shifted versions are easily produced by systematically applying a shift time t_{shift}. If we want a set of N shifted m-sequences, we define t_{shift} by

$$t_{shift} \leq T_m / N \; . \tag{4}$$

For example, if T_m is 8188ms and N is four, then t_{shift} can be given a value of 2040ms. The set $\{S_1, \; S_2, S_3, \; S_4\}$ is obtained by applying time shifts of 0ms, 2040ms, 4080ms, and 6120ms. We note that N is also the number of simultaneous vibrator sources we wish to use in our seismic survey.

Auto- and cross-correlations of shifted m-sequences

The autocorrelations of the m-sequences S_1 to S_N are periodic and have the appearance of a series of triangular peaks. The peak values on the triangles have a value of rL, and the bases of the triangles rest on a constant DC level of $-r$. Removing the factor r, we obtain scaled peak and DC values of L and -1, respectively. The base of each triangular peak has width equal to $2t_b$. The base width compared to t_{shift} and the period T_m is very narrow when the degree m is greater than 11 and N is less than 32. In effect, each triangle on the autocorrelations approximates the delta function. The two examples with degrees m = 5 and m = 7 on Figure 1 show how the scaled triangular peaks are increasingly better approximations of the delta function as m increase.

Within a correlation window restricted to lag times less than t_{shift}, all the scaled values of cross-correlations between different S_i are equal to -1; they are nearly zero compared to the scaled peak value L of the autocorrelations. It is this quasi-orthogonal property under correlation that makes the shifted m-sequences

$$\{ \; S_j \; , \; j = 1, 2, \ldots N \; \}$$

suitable for use as pilots for controlling multiple vibrators operating simultaneously.

Driving multiple vibrators with quasi-orthogonal shifted m-sequences

The m-sequences shown on Figure 1 are too short for practical seismic acquisition, and sequences with longer fundamental lengths L must be used. For example, for the crosswell seismic scanning instrumentation described by Wong (2000), an m-sequence

PRBS with m = 11, L = 2047, t_b = 0.10ms, and t_s = 0.05ms was used for driving a piezoelectric vibrator source.

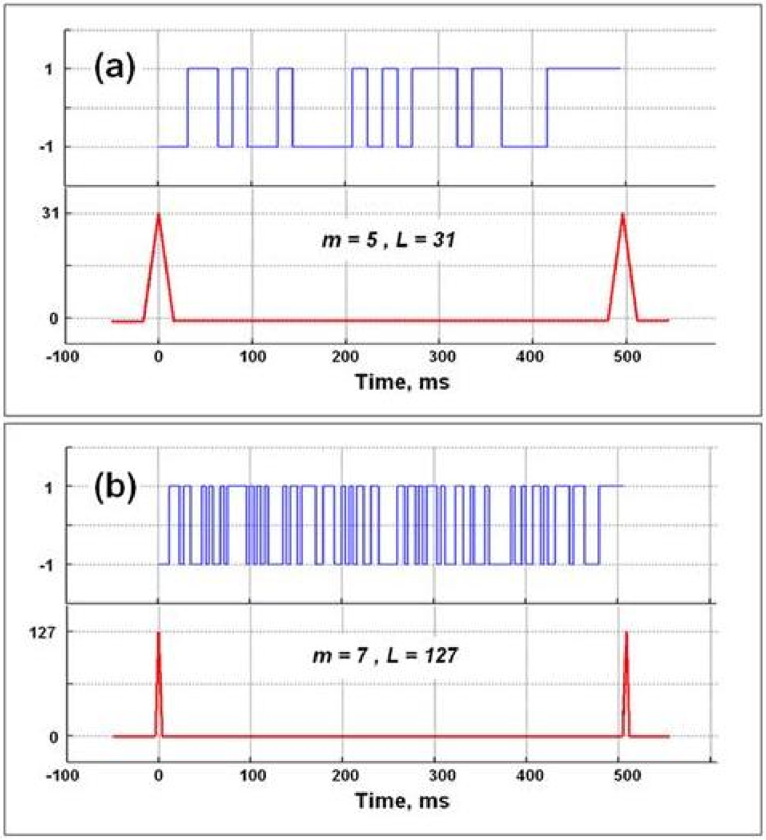

FIG. 1: (a) An m-sequence with m=5, L=31, t_b=16ms, and t_s=1ms (plotted in blue), and its autocorrelation (plotted in red). Both are periodic with period T = 488ms. (b) An m-sequence with m=7, L=127, t_b=4ms, and t_s=1ms (plotted in blue), and its autocorrelation (plotted in red). Both are periodic with period T = 508ms.

To advance the discussion on using shifted m-sequences for simultaneous-source acquisition, let us analyze a specific case where there are four vibrators. We start with an m-sequence with m = 11, L = 2047, base period t_b = 4ms, and sampling time t_s = 1ms. The sampled sequence will be 8188ms long. From this starting sequence S_1, we can construct three other sequences S_2, S_3, and S_4 by systematically shifting S_1 (with wrap-around) by 2040ms, 4080ms, and 6120ms. The initial 2040ms of the four shifted sequences S_1 to S_4 are shown on Figure 2.

The important characteristic of these four sequences is that, when restricted to a correlation window with lags between 0ms and 2040ms, their autocorrelations look like delta functions with peak values of 8188 while the cross-correlations between two different sequences have a constant value of -4. In other words, *within the window of restricted lags*, the shifted m-sequences are almost perfectly orthogonal under correlation. The auto- and cross-correlations for S_1 to S_4 for lag times 0ms to 2040ms are displayed on Figure 3.

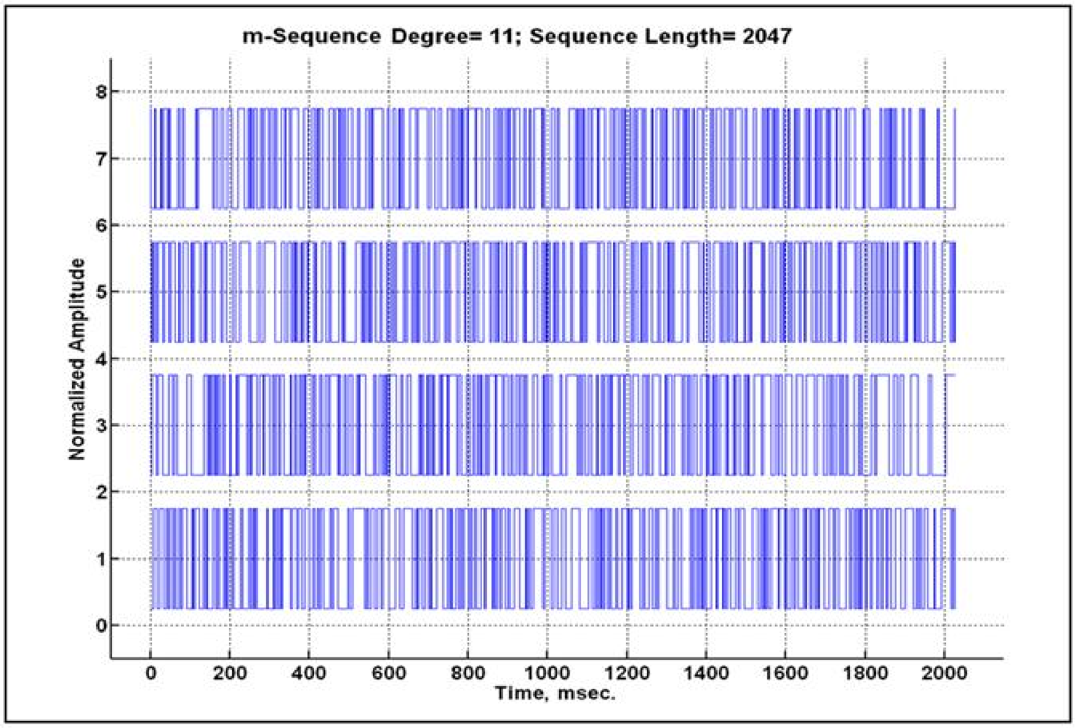

FIG. 2: Four shifted m-sequences (m=11, L=2047, t_b=4ms, t_s=1ms). Time shifts = 0ms, 2040ms, 4080ms, and 6120ms. Full cycle lengths = 8188ms; only the first 2040ms are displayed.

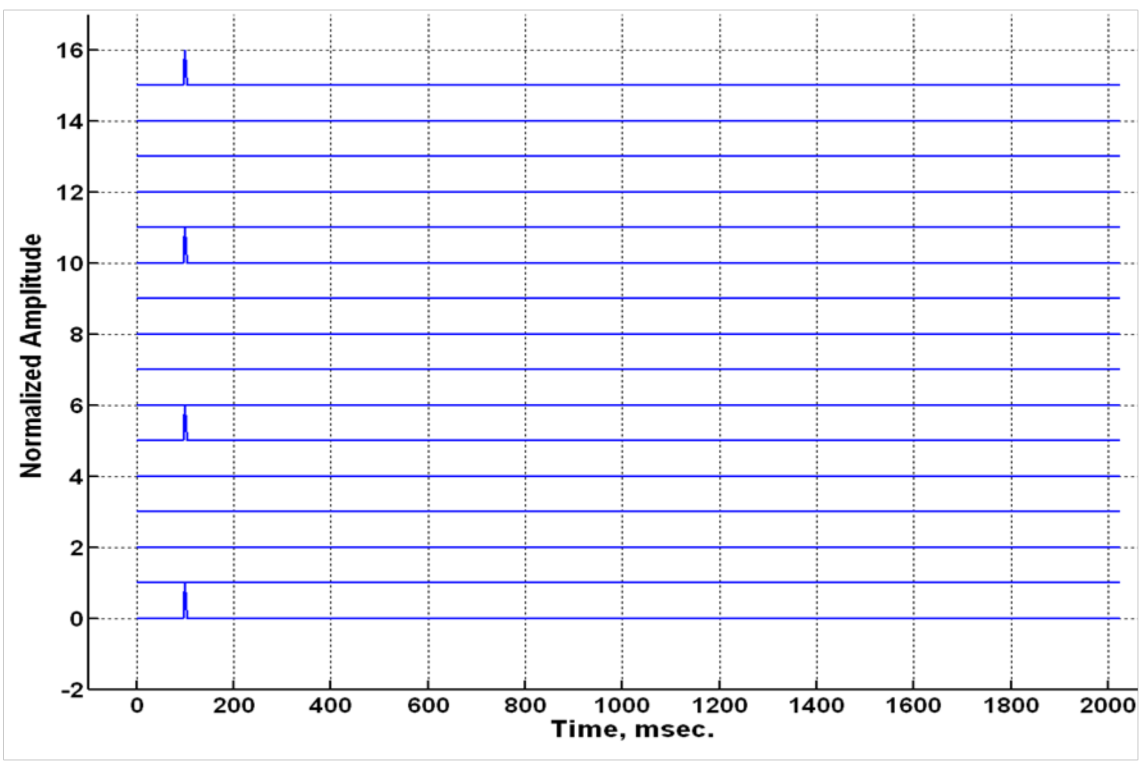

FIG. 3: Auto- and cross-correlations of the four shifted m-sequences of Figure 3. The traces have been shifted forward by 100ms to show the triangular form of the correlation peaks.

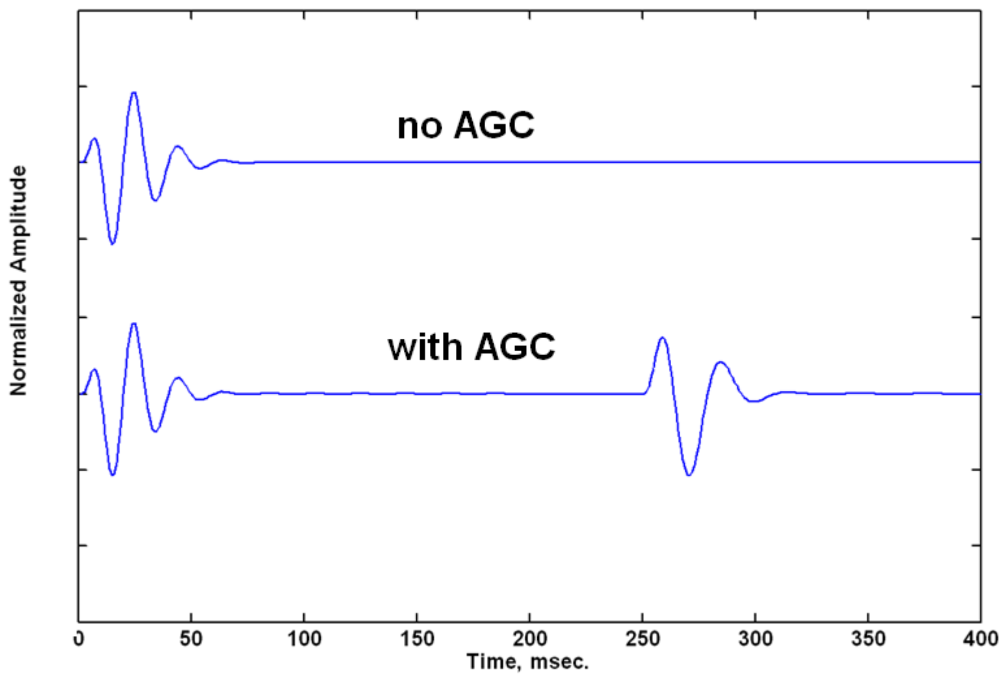

FIG. 4: Source function with a strong first event and a very weak second event. The amplitude of the weak event is 10^{-4} times that of the strong event.

Now, if each of the four vibrators in our survey is controlled by by a pilot equal to a different S_j, and if the four vibrators are driven simultaneously, the combined signal received at any geophone can be separated almost perfectly into four traces simply by cross-correlation with each of the S_i. The signal on each separated trace is associated solely with the source being driven by the appropriate S_i. There will be little if any crosstalk from the other sources.

Figure 4 shows plots of two wavelets used to represent seismic arrivals in the impulse response of the earth to a seismic source. The second wavelet is very weak; it has an amplitude 10^{-4} times that of the first wavelet. Only the strong wavelet is visible on the normalized plot of Figure 4(a); both wavelets are visible on the AGC plot of Figure 4(b).

On Figure 5, we plot the convolutions of the wavelets with each of the shifted m-sequences. The convolutions yield received signals R_1 to R_4 that are filtered versions of the original m-sequence S_1 to S_4. For each received signal, we have specified different time moveouts for the strong and the weak arrivals. Each received signal has also been successively delayed by 100ms to represent increasing arrival times due to increasing separations between the four sources and a single receiver. The initial times of the R_j have zero values because the receiver sees no energy until vibrations have travelled the distance between the sources and the receiver. At the single receiver, the detected signal will be the sum R_T since the four sources are operating simultaneously.

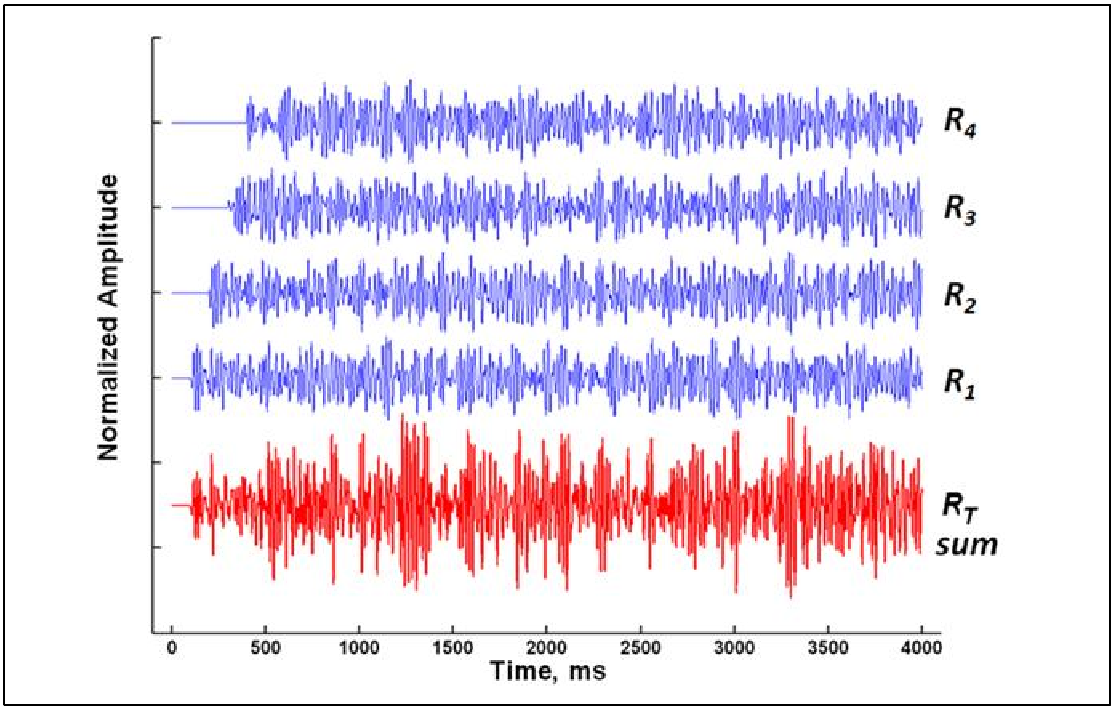

FIG. 5: Convolutions R_1 to R_4 of the source function with each of the m-sequence pilots on Figure 2. R_1 to R_4 are shifted forward in time to simulate increasing arrival times due to increasing source-receiver separations. The sum R_T is the total signal received at a single receiver.

Complete and incomplete correlations

To obtain the theoretical peak value of L and constant value of -1 (after scaling), auto- and cross-correlation must involve complete m-sequences and complete filtered m-sequences. If we drive a vibrator with only one complete m-sequence and listen for a time somewhat longer than the time T_m, then inevitably there will be zero values in the received signal for a short time before the source energy arrives at the receiver, and possibly also for a short time before the end of the listen time. Cross-correlating the m-sequence pilots with a received signal with zero values unconnected with the original m-sequences causes incomplete correlation and produces correlation noise. To avoid this problem of incomplete correlation, we activate each vibrator for two complete PRBS cycles, and also listen for two complete PRBS cycles. Then, for times corresponding to the second cycle of the activating PRBS, the signals R_1 to R_4 will always have the values of one complete (but delayed) cycle of the filtered m-sequences.

Figure 6 shows the traces separated by cross-correlation with the individual pilots S_1 to S_4 using values of R_T beginning at zero time. The correlations are incomplete because of the zero values that are present at the beginnings of the signals R_1 to R_4. Even though the correlation is incomplete, the strong arrivals appear quite clearly. However, close examination of the traces reveal a low level of correlation noise. On the AGC plot of Figure 7, there is no hint at all that the weak arrivals exist; they are completely hidden by the correlation noise.

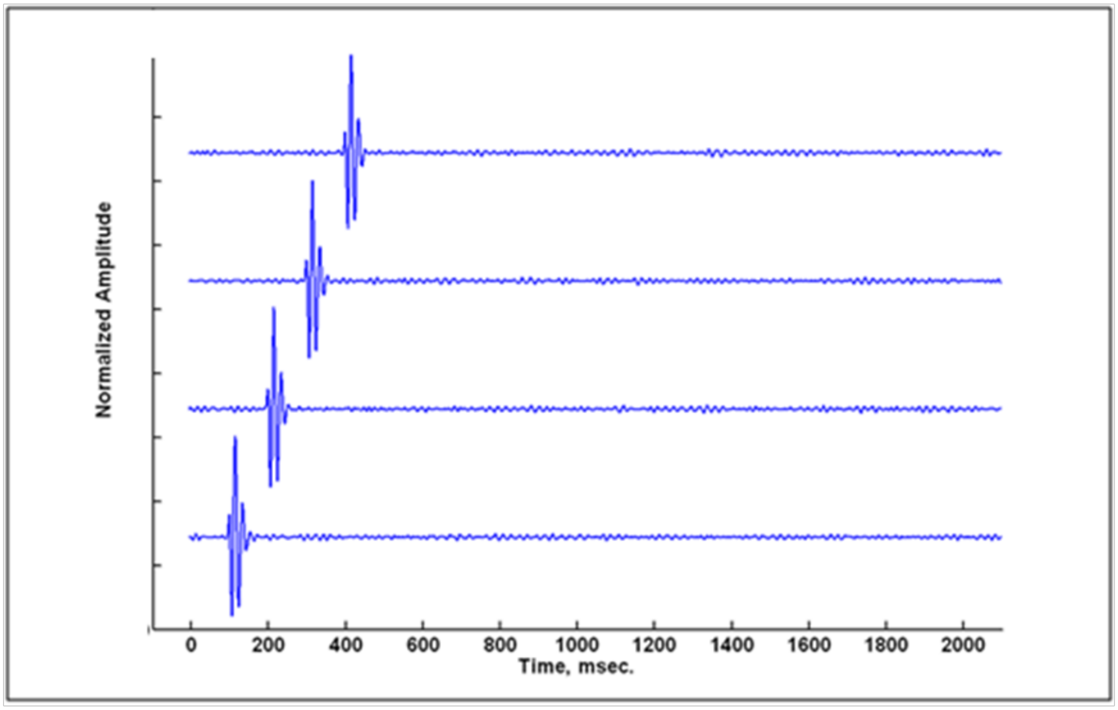

FIG. 6: Cross-correlation of first cycle of the received signal R_T with the shifted m-sequences of Figure 2, yielding separated seismic traces. Each trace has been normalized before plotting.

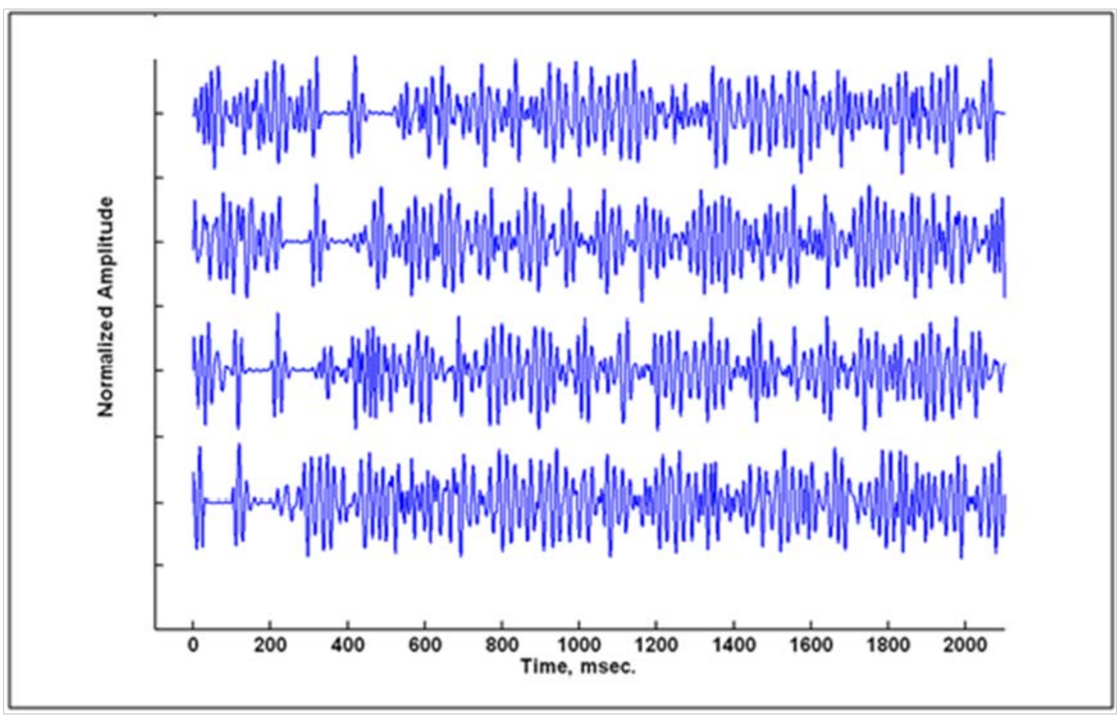

FIG. 7: AGC plot of the traces on Figure 6. The weak arrivals are hidden in the correlation noise.

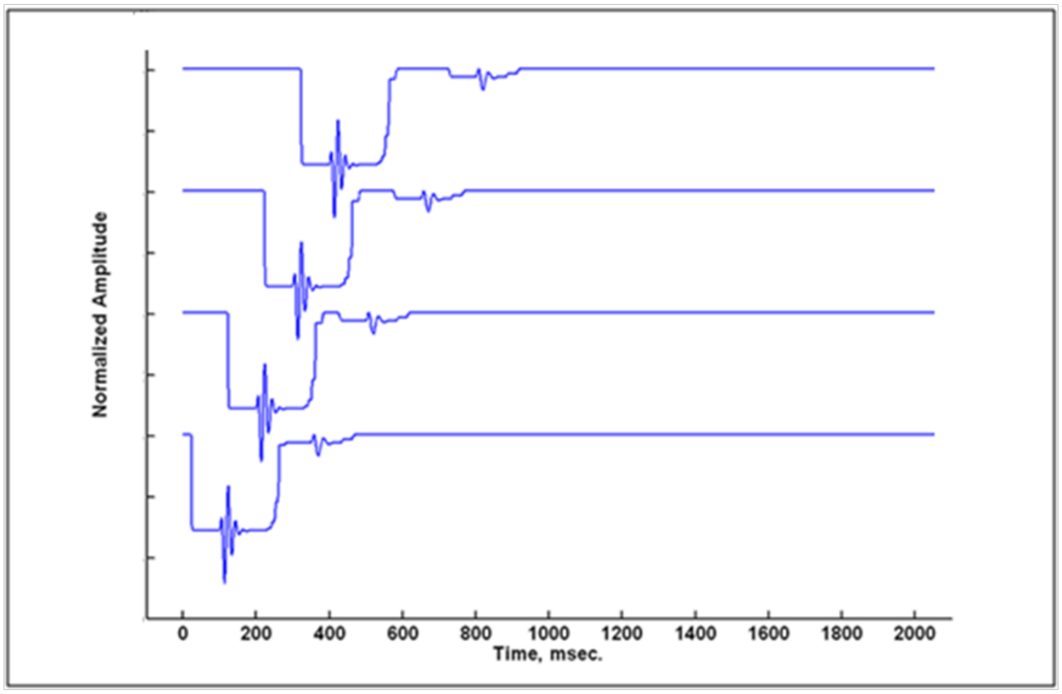

FIG. 8: Cross-correlation of second cycle of the received signal R_T with the shifted m-sequences of Figure 2, yielding separated seismic traces with the absence of correlation noise. On this AGC plot, the weak arrivals are visible, but are distorted by the small DC level inherent to the autocorrelations of m-sequences.

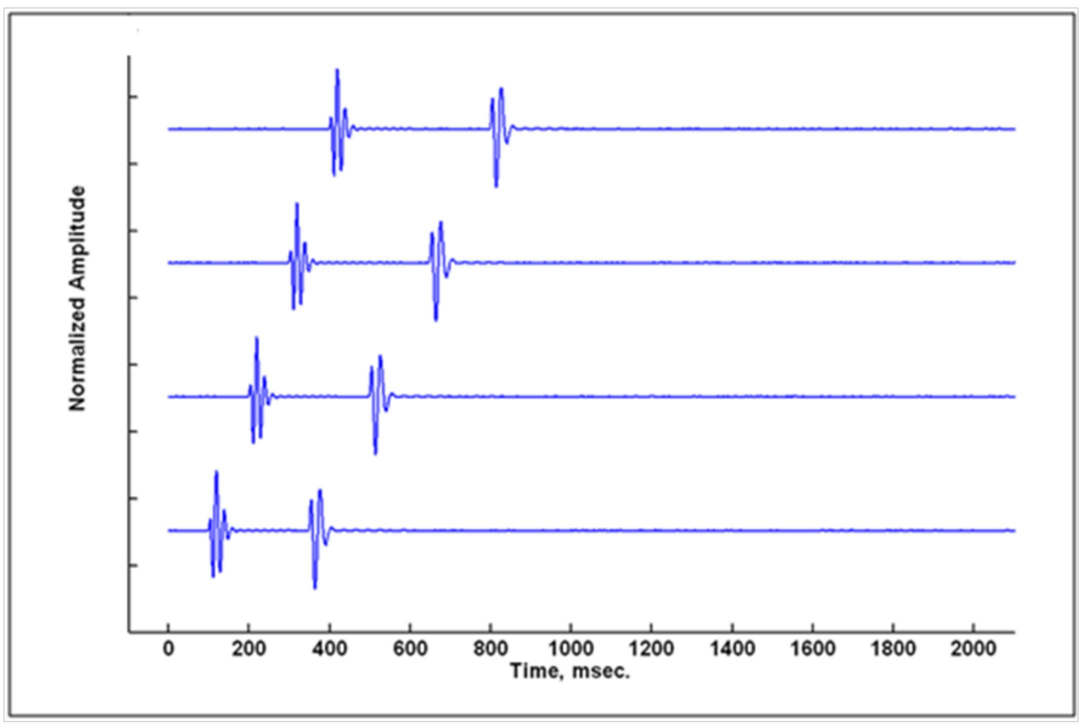

FIG. 9: AGC plot of the traces on Figure 8 after removal of the DC level on each trace. Even though the amplitude of the weak arrival is 80 dB down from the amplitude of the strong arrival, the weak arrival is very clear.

The values of R_T for times coincident with the second activating PRBS cycle will be the sum of complete filtered (but delayed) m-sequences. None of these values will be zero. Consequently, cross-correlation with the m-sequence pilots using the second half of R_T will be complete, and there will be no correlation noise. Traces separated by cross-correlating with the second half of R_T are displayed with AGC on Figure 8, and both the strong and the weak arrivals appear quite clearly. The weak arrivals ride top of a DC level that is the consequence of the scaled off-peak values of -1 in the m-sequence autocorrelations. On Figure 9, the DC level has been removed before AGC plotting, and the weak arrivals are very distinct.

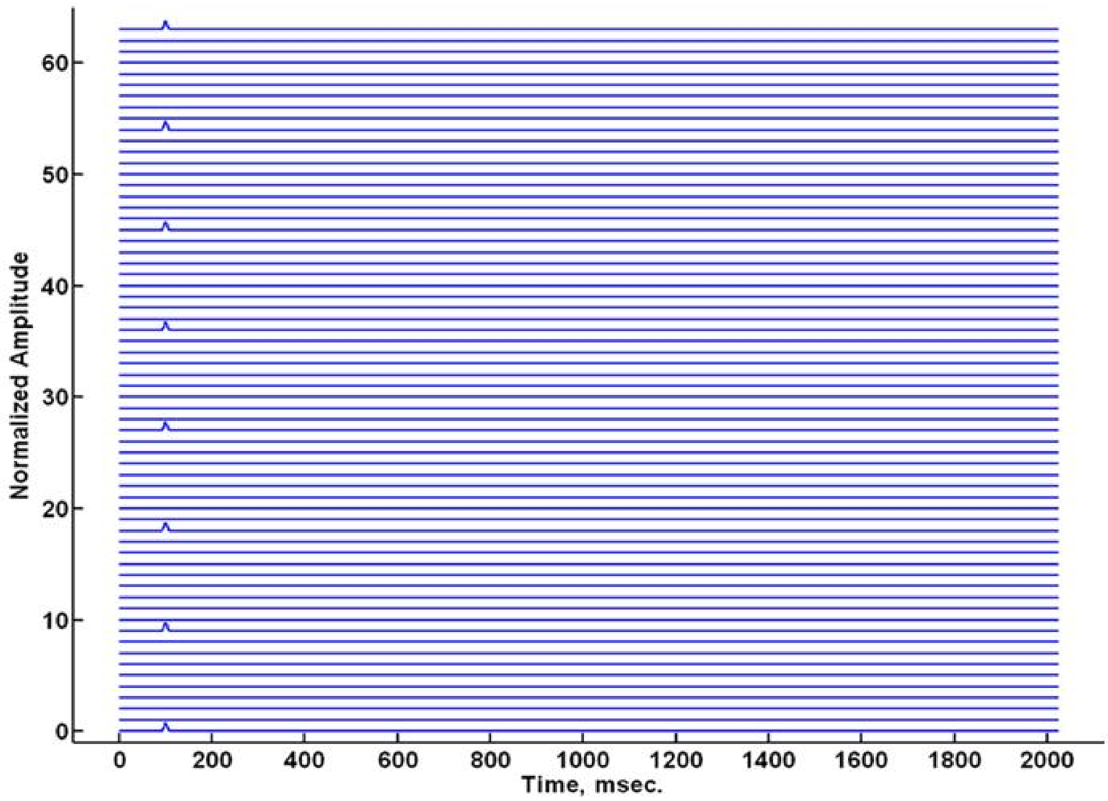

FIG. 10: Auto/cross correlations of 8 shifted m-sequences, showing quasi-orthogonal characteristics. Degree m = 11, L = 2047, t_b = 8ms, t_s =1ms, shift = 2040ms. .

Figure 10 displays the auto- and cross-correlations of a set of eight shifted m-sequences constructed using the parameters listed in the caption. The correlations confirm that the m-sequences are quasi-orthogonal in the time window 0ms to 2040ms, and hence are suitable for acting as pilot signals for controlling eight vibrator sources operating simultaneously.

Figure 11(a) shows the individual convolutions R_1 to R_8 of the wavelets of Figure 4 with the eight shifted m-sequence pilots. On Figure 11(b), the seismic traces obtained by complete correlation of the sum signal R_T with the pilots are displayed with AGC after DC level removal. Both the strong and the weak arrivals are very distinct.

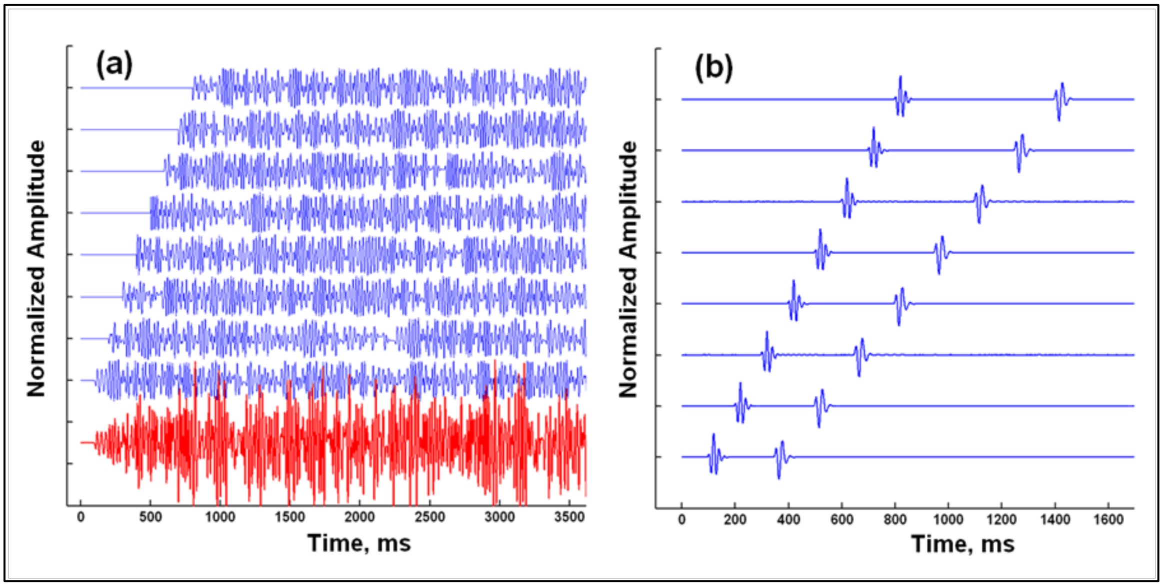

FIG 11: (a) Convolutions of the source wavelets with eight shifted m-sequences, delayed by arrival times between eight sources and one receiver, (R_1 to R_8, shown in blue). The red trace is R_T, the sum of R_1 to R_8, detected at a single receiver. (b) Separated seismic traces obtained by complete correlation of R_T with the individual shifted m-sequences. AGC plot with DC level removed.

SUMMARY AND DISCUSSION

Numerical simulation has shown that shifted m-sequences can be used in place of Gold sequences for simultaneous multiple-source seismic acquisition. In Appendix A, we present evidence that shifted m-sequence pilots may be a better choice than Gold codes, because the use of m-sequences leads to recovered seismograms with much lower levels of correlation noise.

We can design a set of shifted m-sequence pilots for driving multiple vibrators simultaneously by following the steps listed below:

1. Choose number of independent vibrator sources, e.g., $n = 8$.
2. Choose the digital sample rate, e.g., $t_s = 1$ms.
3. Choose the base period of the m-sequence PRBS, e.g., $t_b = 8$ms.
4. Choose the degree and fundamental length of the m-sequence, e.g., $m = 11$, $L = 2047$.
5. Create one period of the m-sequence with over-sampling by a factor of $r = t_b/t_s$. In the example, we get an initial sequence S_1 with $N = (8\text{x}2047) = 16{,}396$ digital points.
6. Choose n_{shift} the number of points for shifting S_1. This number should be an integer less than N/n. In the example, $N/n = 2047$, so $n_{shift} = 2040$ is suitable.

7. Create versions of S_1 by systematically applying the shift with wrap-around. In the example, $S_2 = S_1$ shifted forward by 2040 points, $S_3 = S_2$ shifted forward by 2040 points, … $S_8 = S_7$ shifted forward by 2040 points.

To avoid using zero values in the received signals and so ensure complete correlation, activate each vibrator with two complete cycles of the pilots, listen for two complete cycles, and do circular correlation using the values of the summed received signal corresponding to the times of the second cycle.

ACKNOWLEDGEMENT

The contents of this report have been contributed in part by JODEX Limited. CREWES is supported financially by NSERC and its industrial sponsors.

REFERENCES

Gold, R., 1967. Optimal binary sequences for spread spectrum multiplexing: IEEE Trans. Inform. Theory, **13**(4), 619-621.

Golomb, S.W., and Gong, G., 2005. Signal design for good correlation: for wireless communication, cryptography, and radar: ISBN-0-521-82104-5.

Holmes, J.K., 2007. Spread spectrum systems for GNSS and Wireless Communications, Artech House, Norwood, ISBN-9787-59693-083-4.

Hurley, P., 1983. The development and evaluation of a crosshole seismic system for crystalline rock environments: M.Sc. thesis, University of Toronto.

Pecholcs, P., Lafon, S. K., Al-Ghamdi, T., Al-Shammery, H., Kelamis, P. G., Huo, S. X., Winter, O., Kerboul, J.B., and Klein, T., 2010. Over 40,000 vibrator points per day with real-time quality control: opportunities and challenge: SEG Exp. Abstracts, **29**, 111-115.

Maxwell, P., Gibson, J., Egreteau, A., Lin,,F.,Guido Baeten, G.,and Sallas, J., 2010. Extending low frequency bandwidth using pseudorandom sweeps, SEG Exp. Abstracts, **29**, 101-105.

Sallas, J., Gibson, J., Maxwell, P., and Lin,,F., 2011 . Pseudorandom sweeps for simultaneous sourcing acquisition and low-frequency generation, The Leading Edge, **30**, 1162-1172.

Sallas, J. J., Gibson, J. B., Lin, F., Winter, O., Montgomery, R., and Nagarajappa, P., 2008: Broadband vibroseis using simultaneous pseudorandom sweeps: SEG Exp. Abstracts., **27**, 100-104.

Sallas, J. J., and Gibson, J.B., 2008. Efficient seismic data acquisition with source separation, U.S. Patent No. 7,859,945.

Watson, E.J., 1962. Primitive Polynomials (Mod 2): Math. Comp., **16**, 368, 1962.

Wong, J., 2012. Spread spectrum techniques for seismic data acquisition: CREWES Research Report **24**, this volume.

Wong, J., 2000. Crosshole seismic imaging for sulfide orebody delineation near Sudbury, Ontario, Canada: Geophysics, **65**, 1900-1907.

Wong, J., Hurley, P., and West, G.F., 1987, Cross-hole seismic scanning and tomography: The Leading Edge, **6**, 31-34.

Wong, J., Hurley, P., and West, G.F., 1983, Crosshole seismology in crystalline rocks: Geophys. Res. Lett., **10**, 686-689.

Yamamoto, T., Nye, T., Kuru, M., 1994, Porosity, permeability, shear strength: crosswell tomography beneath an iron foundry: Geophysics, **59**, 1530-1541.

APPENDIX A: COMPARING m-SEQUENCES AND GOLD CODES

Gold codes have autocorrelations that approximate the delta function less well than the autocorrelations of m-sequences. However, in many applications such as wireless communications, where thousands or even millions of (practically) non-interfering channels are needed, Gold sequences are used instead of the simpler m-sequences because large numbers of weakly-coupled Gold sequences can be produced from just a few pairs of optimal m-sequences. For practical use as vibrator pilot signals for simultaneous-source seismic acquisition, where the number of vibrators is limited to 32 or less, m-sequences may be a superior alternative to Gold codes.

The scaled autocorrelations of both m-sequences and Gold codes of degree m have peak values equal to

$$L = 2^m - 1\,. \tag{A1}$$

The off-peak values of the scaled autocorrelations of m-sequences are always equal to a constant value of -1 independent of both lag and m. On the other hand, the off-peak values of the scaled autocorrelations of Gold codes are dependent on both lag and m. These off-peak values oscillate between three quantities:

$$-t(m),\ \ -1,\ \ t(m) - 2\,, \tag{A2}$$

$$t(m) = \ 2^{(m+1)/2} + 1\,,\ odd\ m\,, \tag{A3}$$

$$t(m) = \ 2^{(m+2)/2} + 1\,,\ \ even\, m\,. \tag{A4}$$

These quantities also define the values found in the cross-correlations of Gold codes defined by the same optimal pair of m-sequences (Gold, 1967). For large m, the ratio of the peak autocorrelation values to off-peak values for m-sequences is approximately equal to 2^m, while the same ratio for Gold codes is approximately equal to $2^{m/2}$. It is because of this fact that we can say that the autocorrelations of m-sequencesare better approximations to the delta function.

The quantities in Equations A3 and A4 means that more correlation noise and crosstalk exist between different Gold codes than exist between members of a set of shifted m-sequences. The increased correlation noise has deleterious effects when it is necessary to detect very weak signals such as those that arise from deep reflections.

Figures A1 and A2 show a single m-sequence and a single Gold code (both of degree m = 9) and their normalized autocorrelations. Although it is difficult to see any significant differences between the two types of PRBS, their autocorrelations are crucially different. For the m-sequence autocorrelation, the normalized off-peak values are constant and hardly discernible (they are equal to -1/511). For the Gold code autocorrelation, the off-peak values are much more visible, and they oscillate in a seemly random fashion between -33/511, -1/511, and 31/511, in accordance with Equations A1 to A4. For m = 9, the Gold code correlation noise is on the order of 24 dB, and the noise is unavoidable. For comparison, both Pecholcs et al. (2010) and Sallas et al. (2011) quote SNR values of about 20 dB for Vibroseis field data collected used Gold code pilots.

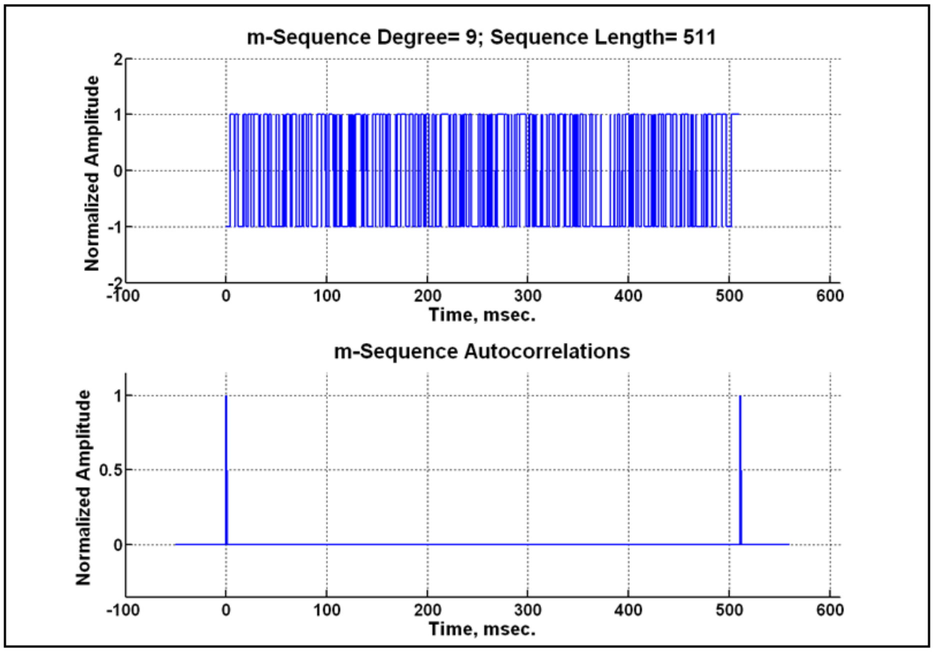

FIG. A1: An m-sequence (m=9, L=511, t_b=1ms, t_s=1ms) and its autocorrelation

.

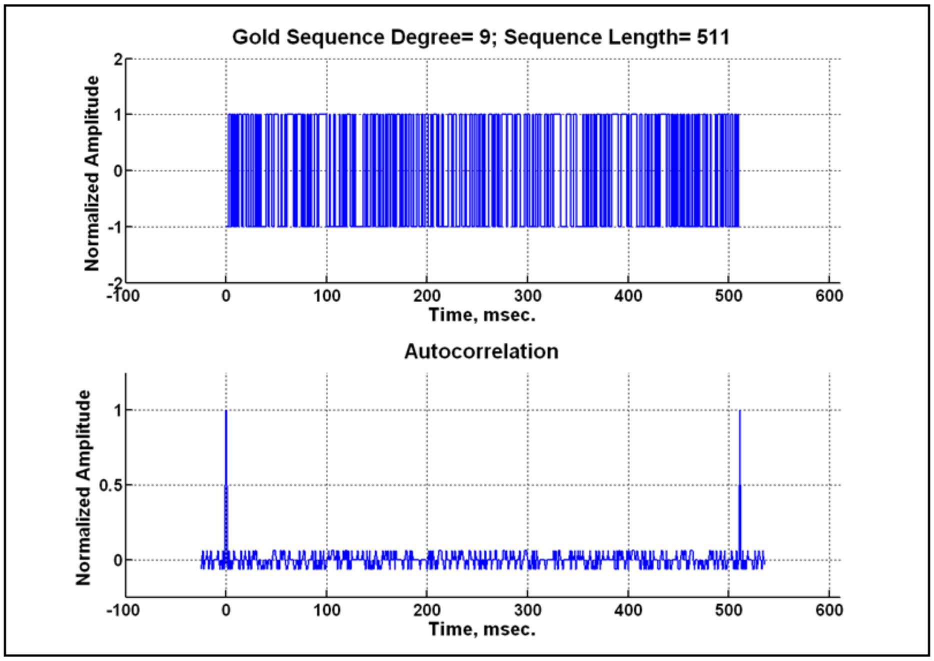

FIG. A2: A Gold code (m=9, L=511, t_b=1ms, t_s=1ms) and its autocorrelation.

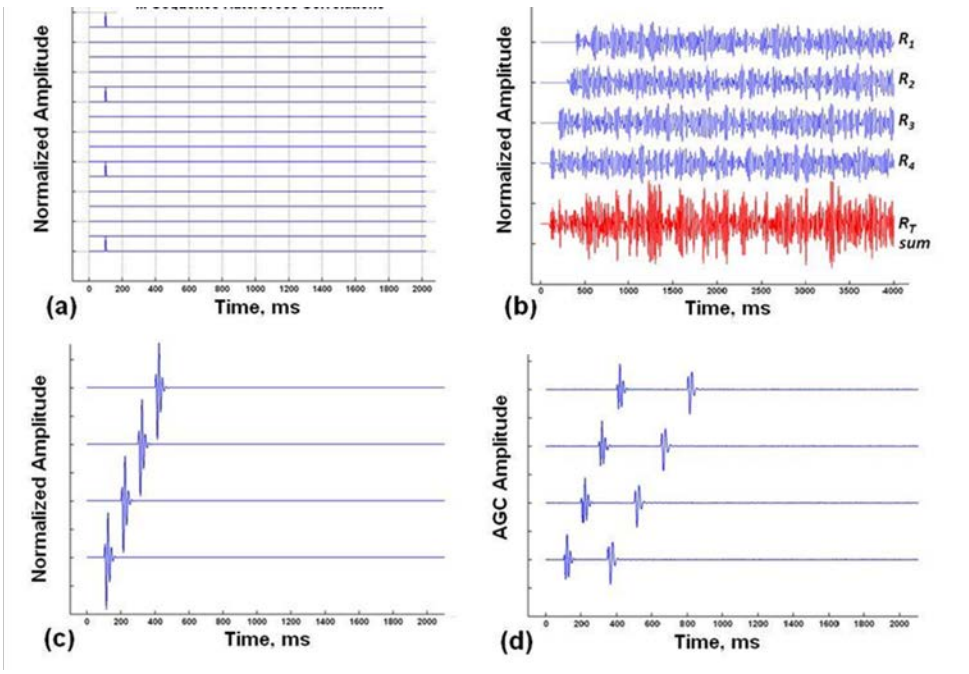

FIG: A3: (a) Auto/cross correlations of four shifted m-sequences. (b) Convolutions of wavelet with four shifted m-sequences, delayed by arrival time between four sources and one receiver (R_1 to R_4); R_T (in red) is the sum of R_1 to R_4 detected at the receiver. (c) Seismic traces obtained by complete cross-correlation of R_T with each of the shifted m-sequences. (d) AGC plot of traces in (c) after removal of DC level.

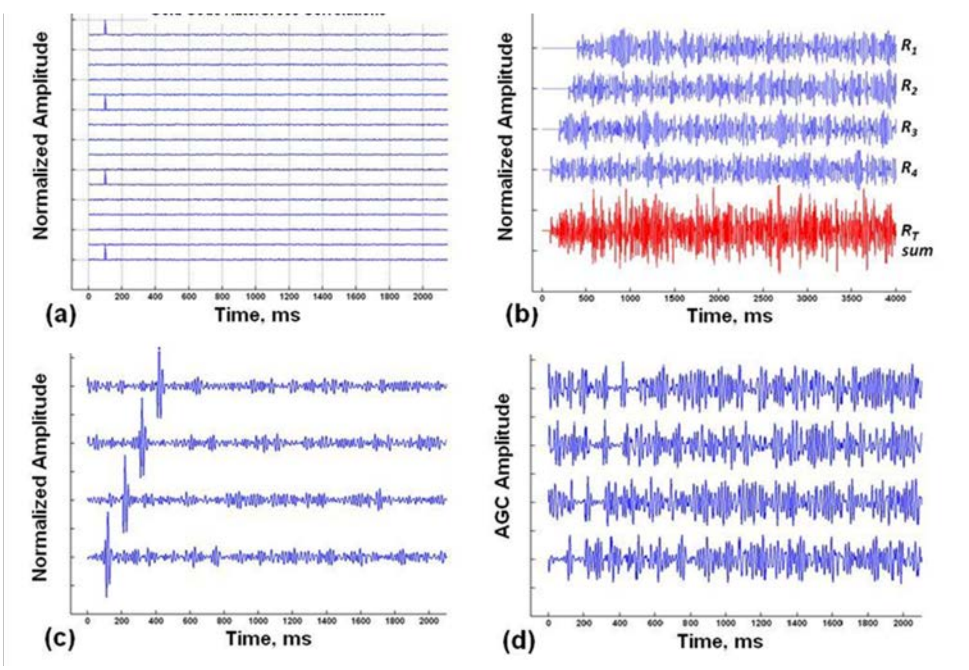

FIG: A4: (a) Auto/cross correlations of four Gold codes. (b) Convolutions of wavelet with four Gold codes, delayed by arrival time between four sources and one receiver (R_1 to R_4); R_T (in red) is the sum of R_1 to R_4 detected at the receiver. (c) Seismic traces obtained by complete correlation of R_T with each of the Gold codes. (d) AGC plot of traces in (c).

Figures A3 and A4 emphasize the difference in using shifted m-sequences versus using Gold codes as pilot signals to control multiple seismic sources operating simultaneously. The shifted m-sequences and Gold codes used to produce the data on the figures have identical parameters: m = 11, L = 2047, t_b = 4ms, t_s = 1ms. The period of the shifted m-sequences is 8188ms, and the time shift between adjacent members is 2040ms. The wavelets used in the convolutions are those on Figure 4, where the weak wavelet has an amplitude that is 10^{-4} times that of the strong wavelet.

For our comparisons of autocorrelations and recovered seismic traces on Figures A3 and A4, we must restrict our attention to the time interval 0ms to 2040ms where the shifted m-sequences are “orthogonal”. We see that the correlation noise for the m-sequences is a constant value, and it is smaller than the oscillating correlation noise for the Gold codes (by a factor of about 1/64). On the traces produced with m-sequence pilots, both the strong and the weak arrivals are clearly visible with AGC display. On the traces produced with Gold codes pilots, the weak arrival is completely lost in the correlation noise.

The evidence shown on Figures A3 and A4 strongly suggests that, in general, m-sequences may be more effective than Gold codes for driving vibrating seismic sources when the goal is to detect weak reflection events.