Acta Geodyn. Geomater., Vol. 4, No. 3 (147), 5-10, 2007

GENERALIZED AVERAGE OF SIGNALS (GAS) - A NEW METHOD FOR DETECTION OF VERY WEAK WAVES IN SEISMOGRAMS

Jiří MÁLEK, Petr KOLÍNSKÝ, Jaroslav ŠTRUNC and Jan VALENTA

Institute of Rock Structure and Mechanics, Academy of Sciences of the Czech Republic, v.v.i.,
V Holešovičkách 41, 182 09 Prague 8, Czech Republic
**Corresponding author's e-mail:* malek@irsm.cas.cz

(*Received* May 2007*, accepted* September 2007)

ABSTRACT

A novel method called Generalized Average of Signals (GAS) for detection of very weak waves in seismograms is described and tested. The general principle of the GAS method is to take advantage of the coherency of the signal, which is extracted. The signals are shaped with moving window and converted to the frequency domain. Then they are non-linearly summed considering their complex representation (amplitudes and phases). The method improves signal-to-noise ratio of coherent seismograms considerably. The GAS method is tested on synthetic seismograms and compared with the PWS method.

KEYWORDS: Generalized average of signals, signal-to-noise ratio, coherent stack, non-linear sum

1. INTRODUCTION

The methods of stacking of similar seismograms are widely used for improvement of signal-to-noise ratio in many seismological problems. One of the examples could be a detection of very weak distant earthquakes or blasts with help of small-aperture seismic arrays (SSA). Originally SSAs were constructed to detect weak distant seismic events, particularly nuclear blasts (Harjes, 1990; Joswig, 1990; Mykkeltveit et al., 1990). Later, they were used for many special seismological tasks (Joswig, 1995; Joswig, 1999; Leonard et al., 1999). Another examples of stacking might be the method of common reflecting point in the exploration seismology (Sheriff et al., 1995), or identification of P-to-S conversions in the P-wave coda of teleseismic earthquakes in the global seismology (Schimmel, 1999).

The general principle of these methods is to take advantage of the coherency of the signal, which is extracted. The distance of the neighbouring seismic stations in the array has to be comparable with wavelength of the amplified signal to ensure strong coherency. The seismic noise has to be non-coherent at the same wavelength. The seismograms in the array are shifted according to expected differences of travel-times of the particular wave phase, which becomes amplified. Examples of elaborated methods for amplifying of weak signals are the f-k transformation (Mykkeltveit et al., 1990) or polarization and the spatial filters by Schimmel and Gallart (2003).

If N is the number of stations in the array, the linear summation yields improvement of the signal-to-noise ratio by the factor $\sqrt{N}$, assuming the absolutely coherent signal and absolutely non-coherent noise. The main goal of this paper is to introduce more powerful method, generalized average, which improves signal-to-noise ratio much more effectively. The basic idea of the method is taken from Schimmell and Paulssen (1997), who proposed the phase-weighted stack method (PWS). Their method is based on the similarity of the instantaneous phases. Our method GAS (which is described bellow) might be considered an extension of the PWS involving also amplitudes of the coherent signals.

2. GENERALIZED AVERAGE OF COMPLEX NUMBERS

First we will introduce the method of generalized average of complex numbers. Let us consider N complex numbers x_j . The arithmetic average y_0 is done by simple formula:

$$y_0\left(x_j\right) = \frac{\sum_{j=1}^{N} x_j}{N} \tag{1}$$

Let us introduce generalized average of order p by the formula

$$y_p\left(x_j\right) = y_0 s^p \tag{2}$$

where

$$s = \frac{\left|y_0\right|}{\sqrt{\frac{\sum_{j=1}^{N}\left|x_j^2\right|}{N}}} = \frac{\left|\sum_{j=1}^{N} x_j\right|}{\sqrt{N\sum_{j=1}^{N}\left|x_j^2\right|}} \tag{3}$$

Order p is a non-negative real number. Parameter s is a real number from the interval $\langle 0,1 \rangle$. The value

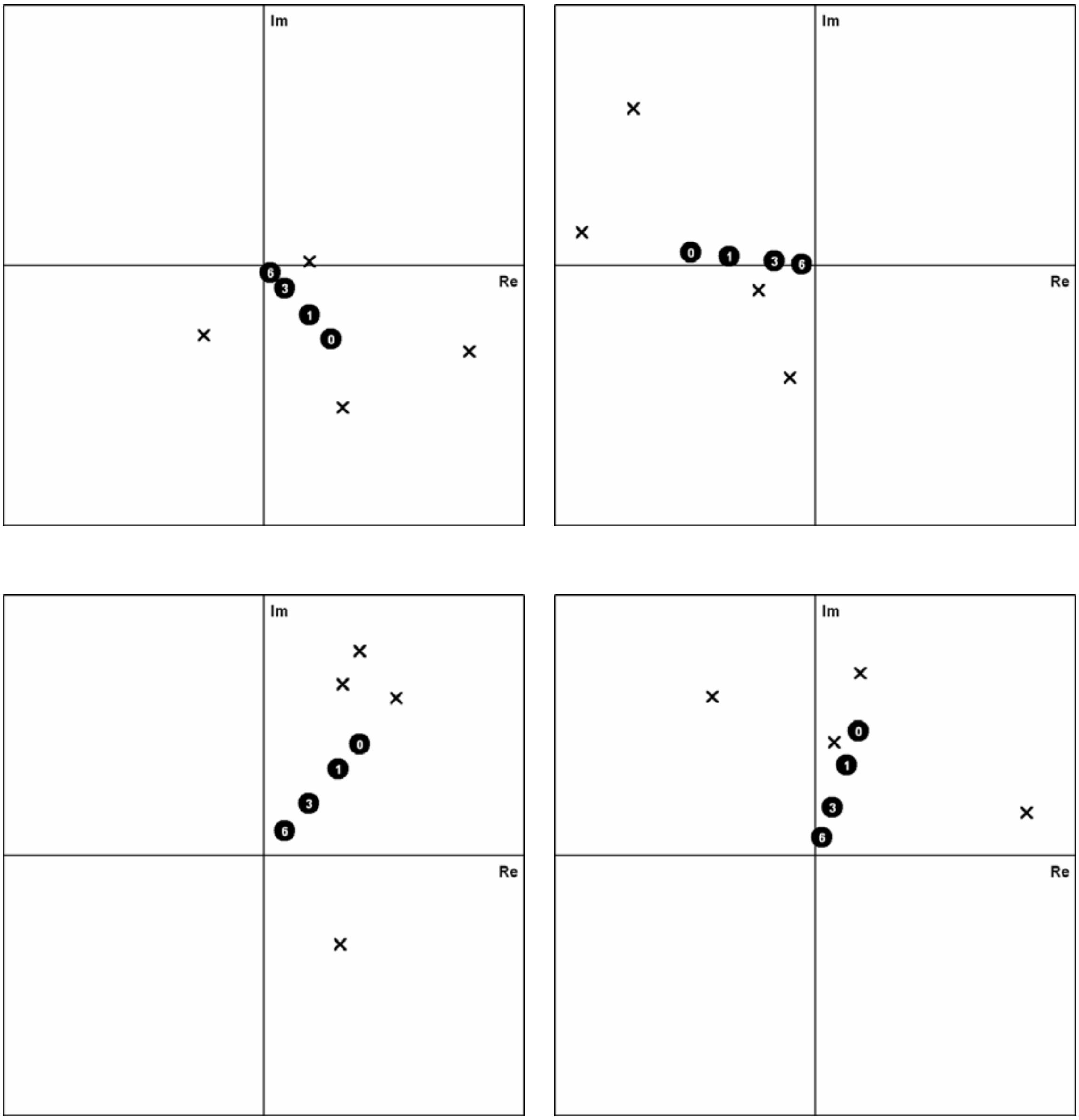

Fig. 1 Examples of the generalized average of complex numbers. Crosses are 4 complex numbers, dots with numbers are their generalized averages of order p = 0, 1, 3, 6.

$s = 1$ is reached only if all numbers x_j are equal. In this case the generalized averages of all orders are equal. We can represent the averages (2) as a complex numbers:

$$y_p = A_p e^{i\varphi} \tag{4}$$

If x_j are not equal, the amplitude A_p decreases with the order p and their limit is 0. The phase φ is independent of p. Four examples of the generalized averages are in Fig. 1.

If x_j represent repeated measurements of some quantity, general averages amplify the measurements with the same results. Higher order of the average p means stronger amplification.

3. GENERALIZED AVERAGE OF SIGNALS (GAS)

Now, we are going to extend generalized average to the situation, when x_j are not complex numbers, but signals $x_j(t)$, i.e. complex time series represented by vectors of complex numbers. Signal $x_j(t)$ with real numbers can represent for instance one component of seismogram from one station in the array. The simplest possibility is to define the resultant seismogram from the array as the averages computed at the fixed time t:

$$\widetilde{G}_p(t) = y_p(x_j(t)) \tag{5}$$

The example with synthetic data is in Fig. 2.

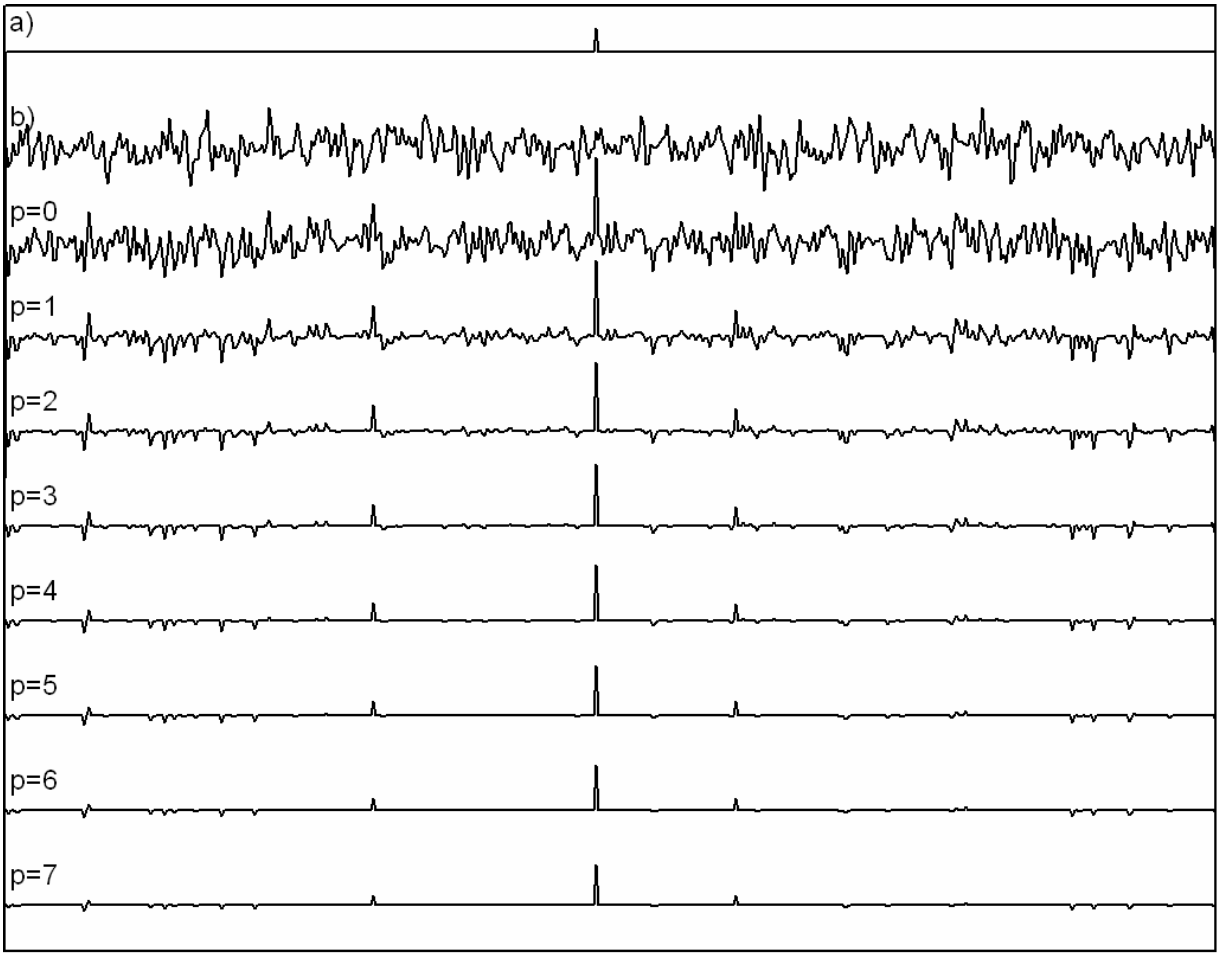

Fig. 2 Generalized average after formula (5), a) the spike signal without noise, b) the same signal with random noise − one of the 9 seismograms, which are averaged. The next traces represent generalized average with order p = 0, 1, ..., 7.

This definition is suitable in some special situations, for instance for detecting peaks in signal, but it does not fit seismograms very well.

We can take advantage from the fact that the spectra of the signal and the noise are different. Therefore, it is convenient to compute the complex spectrum of the seismogram first, make generalized average in the frequency domain and convert it back to the time domain:

$$\hat{G}_p(t) = F^{-1}\left(y_p\left(F\left(x_j(t)\right)\right)\right) \tag{6}$$

where *F()* is a symbol for Fourier transform.

In case of the arithmetic average, for $p = 0$, formulas (5) and (6) give the same result. But for $p > 0$ the results are different. Monochromatic signals, which are the same at all stations, are amplified. Definition according to formula (6) is suitable for long-lasted signals with narrow spectrum, e.g. for volcanic tremors. The example using synthetic data is in Fig. 3.

For body waves, which represent relatively short signals with wide spectrum, we need some compromise between (5) and (6). The first step is the representation of the signals $x_j(t)$ by the sum of the signals shaped by moving Hann (cosine) windows:

$$x_j(t) = \sum_{l=1}^{L} z_j^l(t) \tag{7}$$

$$z_j^l(t) = x_j(t) w^l(t) \tag{8}$$

$$w^l(t) = \frac{1}{2}\left(1 + \cos\left(\pi\left(\frac{t - lh}{h}\right)\right)\right) \text{ for}$$

$$t \in (lh - h, lh + h), \tag{9}$$

$$w^l(t) = 0 \text{ otherwise,}$$

where h is a half-width of the Hann window.

The second step is the application of the generalized average according to (6) to signals $z_j^l(t)$ and the last step is summing the signals back according (7). The full definition of the generalized average is:

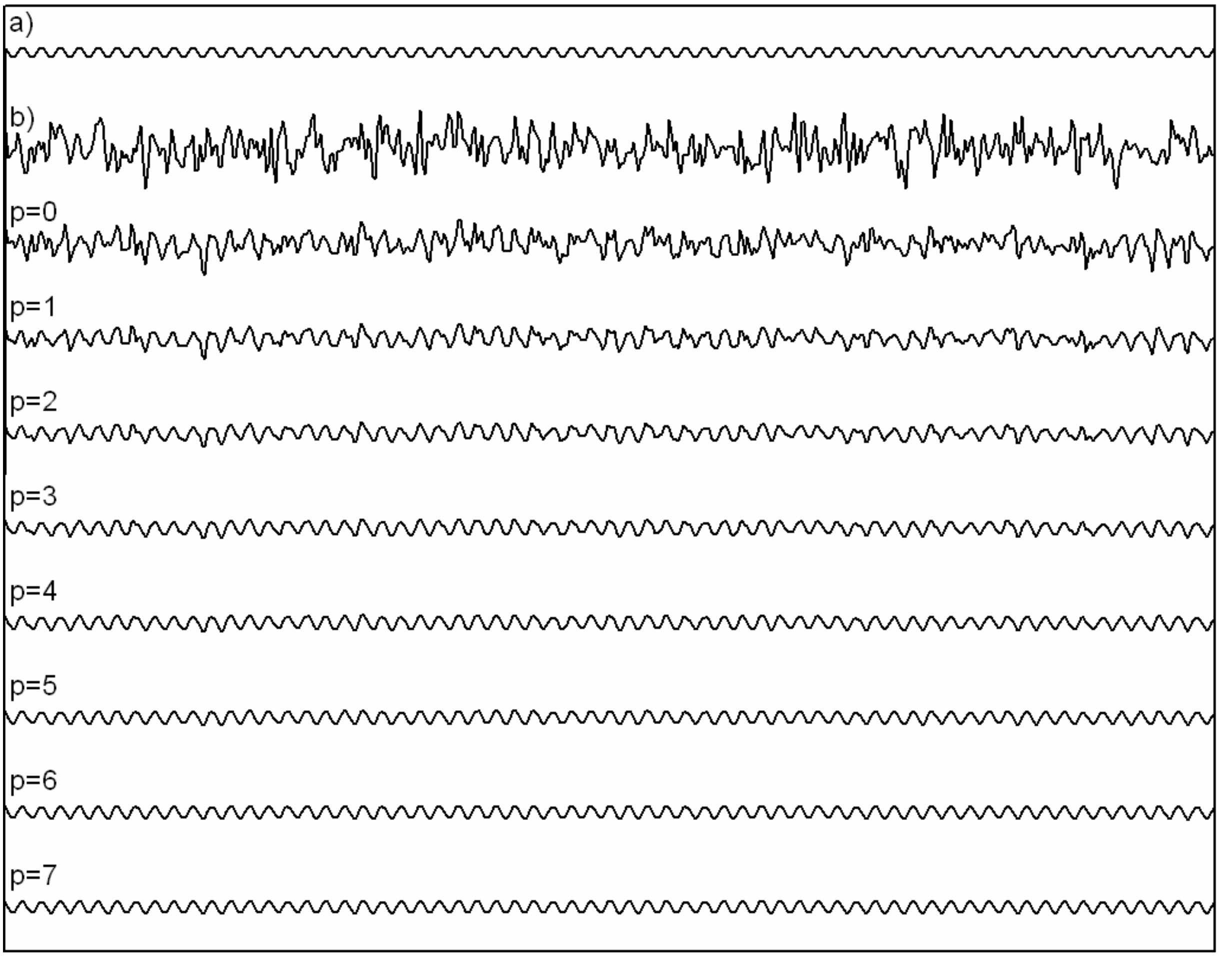

Fig. 3 Generalized average of seismogram using formula (6). a) monochromatic signal without noise, b) the same signal with random noise − one of the 9 seismograms, which are averaged. The next traces represent generalized average with order p = 0, 1, ..., 7.

$$G_p(t) = \sum_{l=1}^{L} F^{-1}\Big(y_p\Big(F\big(x_j(t)w^l(t)\big)\Big)\Big) \tag{10}$$

The formula (10) represents our definition of generalized average of signals (GAS). It is dependant on the half-width of the Hann window *h*. This parameter has to be tuned together with the parameter *p* to obtain optimal results in applications.

In case of $p = 0$ GAS is equal to the arithmetic average of signals (5). If all signals are the same, their GAS is equal to the signals. The example of synthetic data using formula (10) is in Fig. 4. This figure also demonstrates distortion of the signal for higher orders of *p*.

The same data as in Fig. 4 were processed by PWS method developed by Schimmell and Paulssen (1997). This method is defined by formula:

$$g_p(t) = \frac{1}{N}\sum_{j=1}^{N} x_j(t)\left|\frac{1}{N}\sum_{k=1}^{N} \exp\big(i\Phi_k(t)\big)\right|^p \tag{11}$$

where $\Phi_k(t)$ is instantaneous phase of the k-*th* time series, which is computed using Hilbert transform.

The results are in Fig. 5. In this case, both methods (GAS and PWS) give satisfying results but GAS gives better signal-to-noise ratio within the same order of *p*. The shape of the resulting signal is also better using GAS. PWS method is expected to yield better results, if the amplitudes of signals are not the same at all stations as a consequence of local geological conditions. GAS method is more suitable in cases when the stations are situated at the homogeneous block or the differences between sensors are compensated using preliminary filtering.

4. CONCLUSION

The method of Generalized Average of Seismograms (GAS) was developed to improve signal-to-noise ratio of coherent seismograms from seismic arrays. As was demonstrated on synthetic examples, it yields much clearer results than arithmetic average of seismograms. We hope, it could be applied also outside seismology, whenever several coherent signals with noise are available.

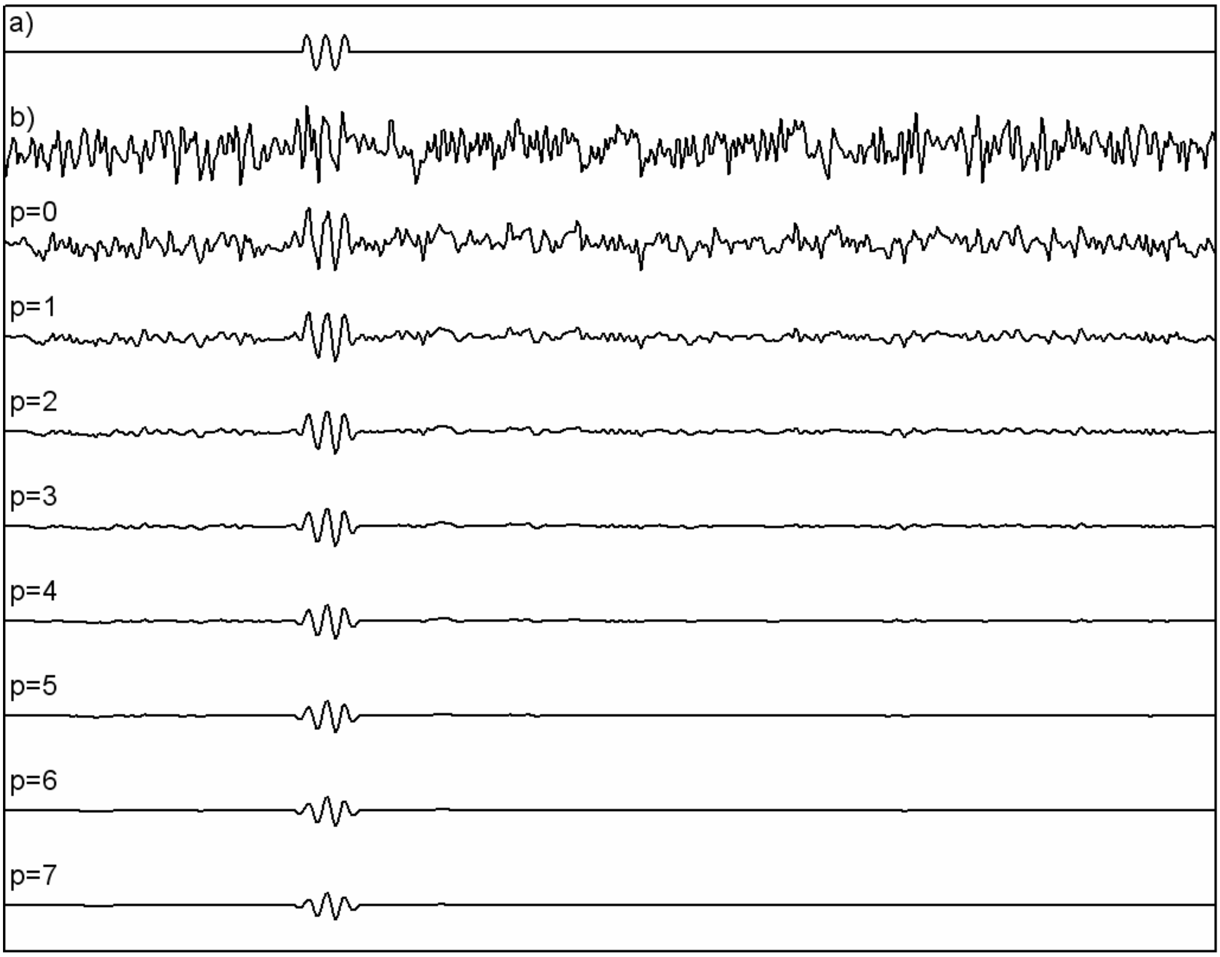

Fig. 4 Generalized average using formula (10). a) the signal without noise, b) the same signal with random noise − one of the 9 seismograms, which are averaged. The next traces represent generalized average with order p = 0, 1, ..., 7.

ACKNOWLEDGEMENT

This work was supported by grant 205/06/1780 of Czech Science Foundation and grant A300460602 of the Grant Agency of the Academy of Sciences of the Czech Republic.

REFERENCES

Harjes, H.P.: 1990, Design and sitting of a new regional array in Central Europe, BSSA, 80(6), 1801-1817.

Joswig, M.: 1990, Pattern recognition for earthquake detection, Bull. Seism. Soc. Am., 80, 170-186.

Joswig, M.: 1995, Automated classification of local earthquake data in the BUG small array, Geophys. J. Int., 120, 262-286.

Joswig, M.: 1999, Automated processing of seismograms by SparseNet. Seism, Res. Letters, 70, 705-711.

Leonard, G., Villagran, M., Joswig, M., Bartal, Y., Rabinowitz, N. and Saya, A.: 1999, Seismic source classification in Israel by signal imaging and rule-based coincidence evaluation, Bull. Seism. Soc. Am., 89, 960-969.

Mykkeltveit, S., Ringdal, F., Kværna, T. and Alewine R.W.: 1990, Application of regional arrays in seismic verification research, Bull. Seism. Soc. Am., 80, 1777-1800.

Schimmel, M. and Paulssen H.: 1997, Noise reduction and detection of weak, coherent signals through phase-weighted stacks, Geophys. J. Int. 130, 497-505.

Schimmel, M.: 1999, Phase cross-correlation: Design, comparisons, and applications, Bull. Seism. Soc. Am., 89, 1366-1378.

Schimmel, M. and Gallart J.: 2003, The use of instantaneous polarization attributes for seismic signal detection and image enhancement, Gephys. J. Int., 155, 653-668.

Sheriff, R.E. and Geldart, L.P.: 1995, Exploration seismology, Cambridge University Press.

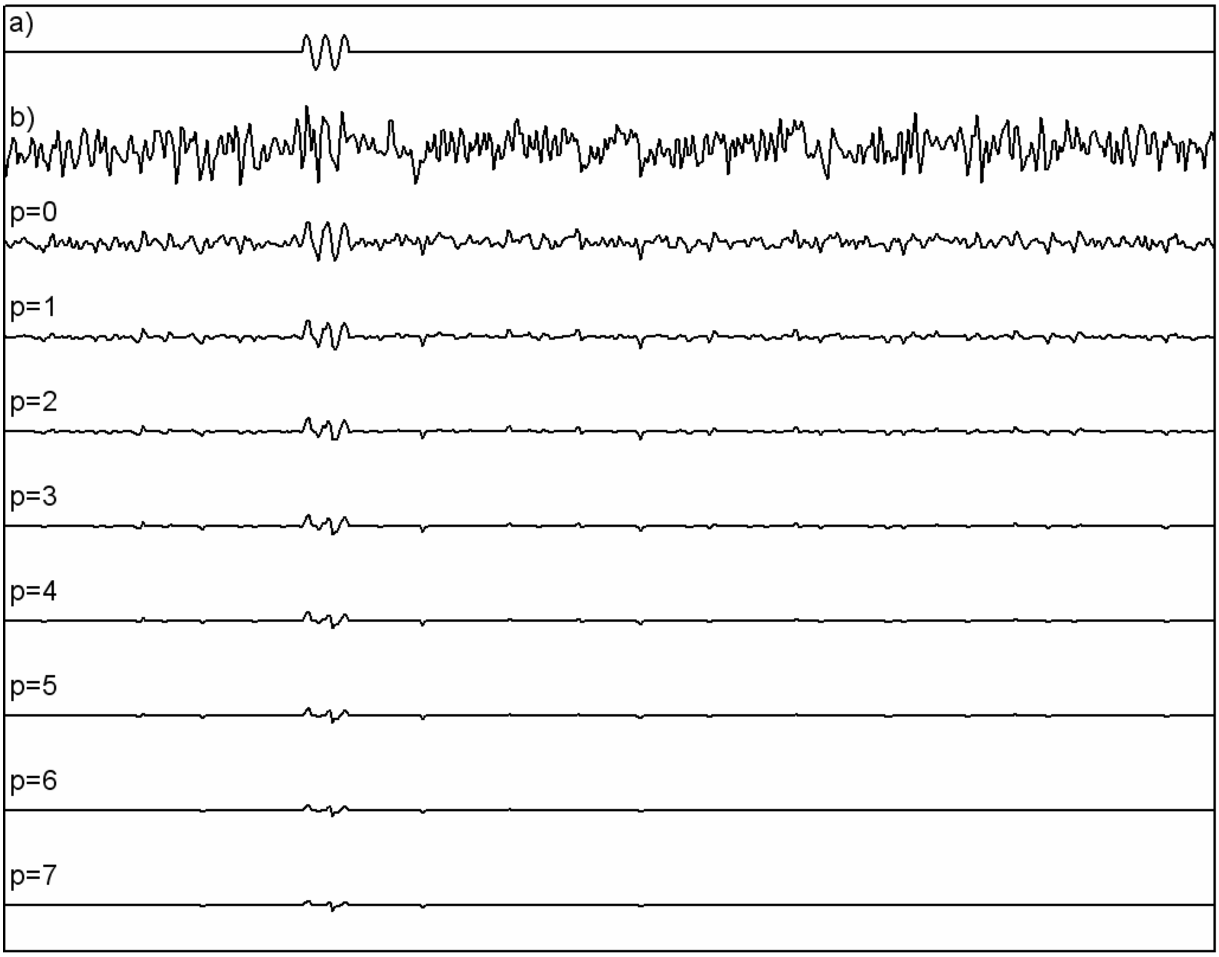

Fig. 5 Phase-weighted stack (Schimmel and Paulssen, 1997) using formula (11). a) the signal without noise, b) the same signal with random noise − one of the 9 seismograms, which are stacked. The next traces represent phase-weighted stacks with order p = 0..7.