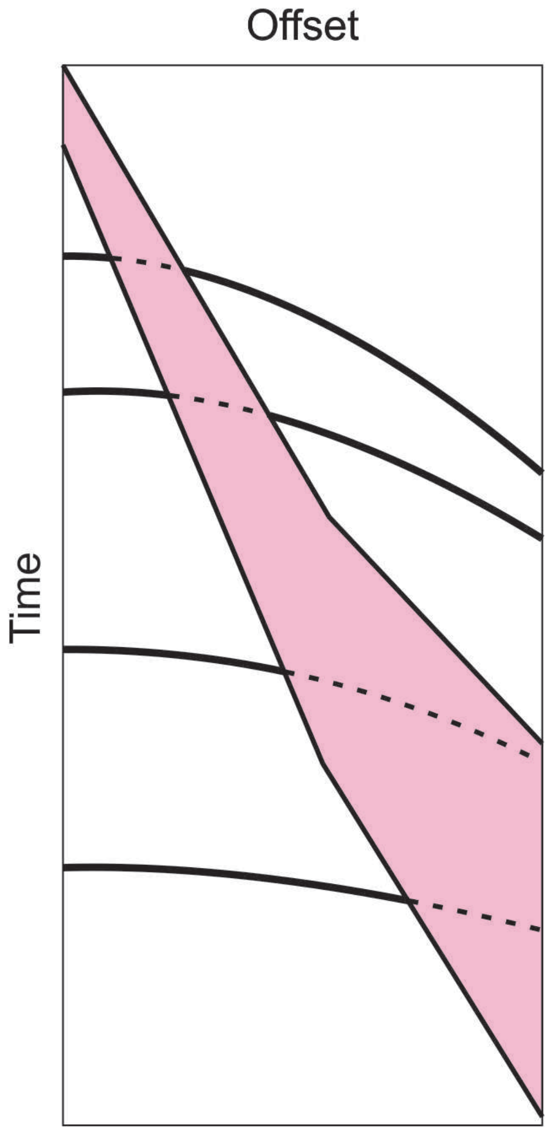
Ground-roll subtraction from seismic record with significant trace-to-trace variations in the energy of random noise

Olena Tiapkina (NTNU) , **Martin Landrø**(NTNU),

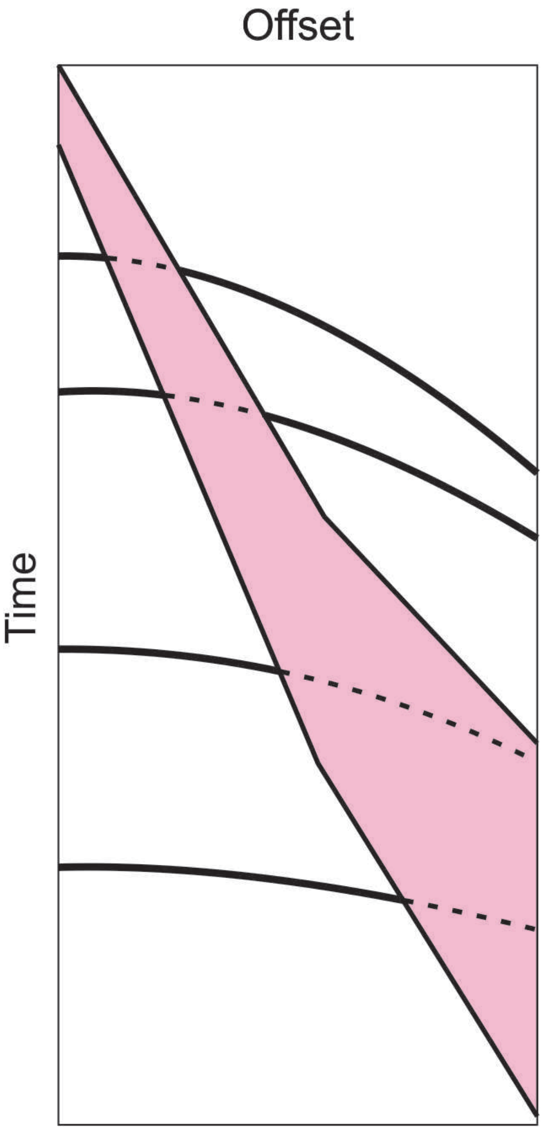
and **Yuriy Tyapkin** (SE Naukanaftogaz)

Introduction



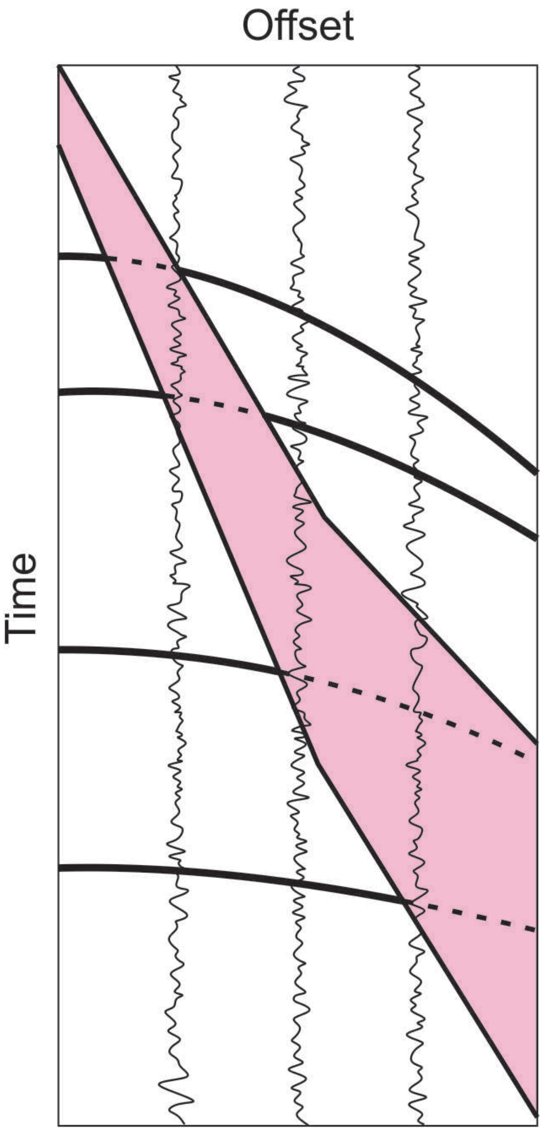
- f-k filtering

Introduction



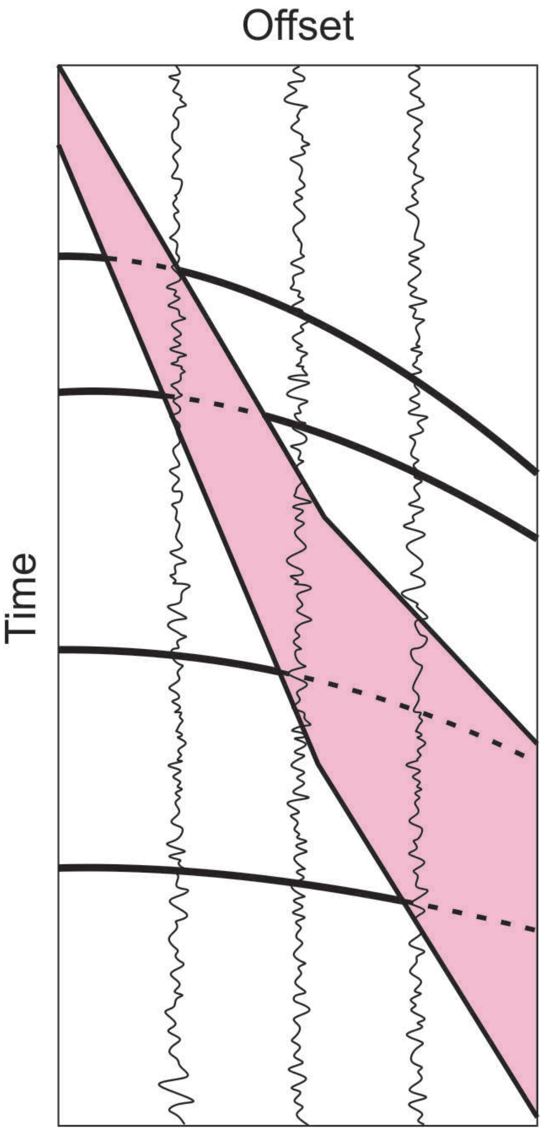
- f-k filtering
- Singular value decomposition (SVD)

Introduction



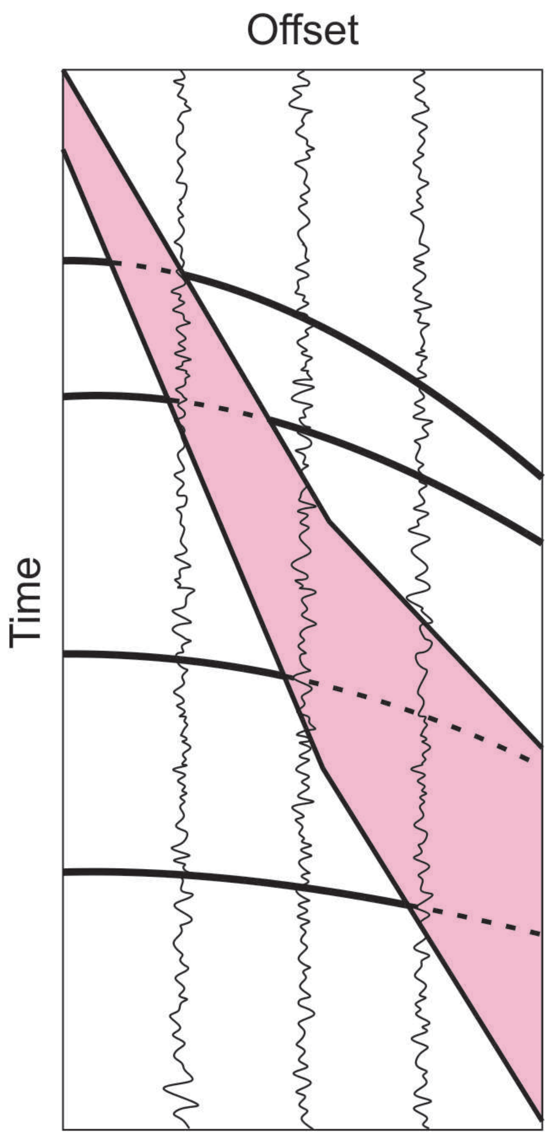
- f-k filtering
- Singular value decomposition (SVD)

Introduction



- f-k filtering
- Singular value decomposition (SVD)
- Optimum weighted stacking

Introduction



- f-k filtering
- Singular value decomposition (SVD)
- Optimum weighted stacking
- 2-stage optimum weighted stacking

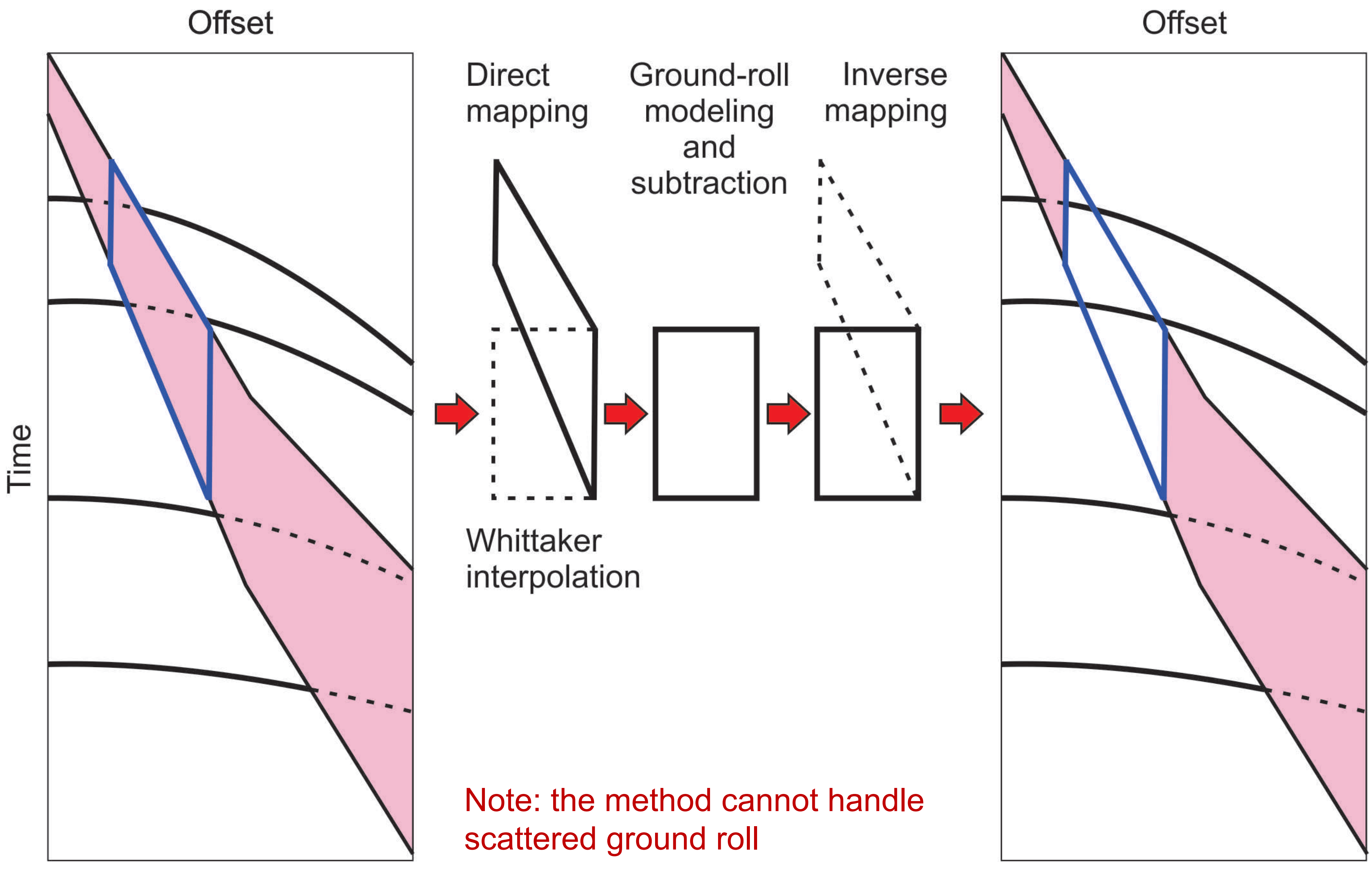
(Tyapkin Yu., Ursin B., Perroud H., Silinska O. and Tiapkina O., 2010. **Zero- and first-order approximations for least-squares estimation of seismic signal with coherent and random noise.** Journal of Geophysics and Engineering**,** **7**, 51–63)

Outline

- Description of methods
  - General scheme for ground roll subtraction
  - SVD-based filtering
  - Optimum weighted stacking
  - 2-stage optimum weighted stacking
- Practical examples
- Conclusions

Ground roll subtraction scheme in a sliding window

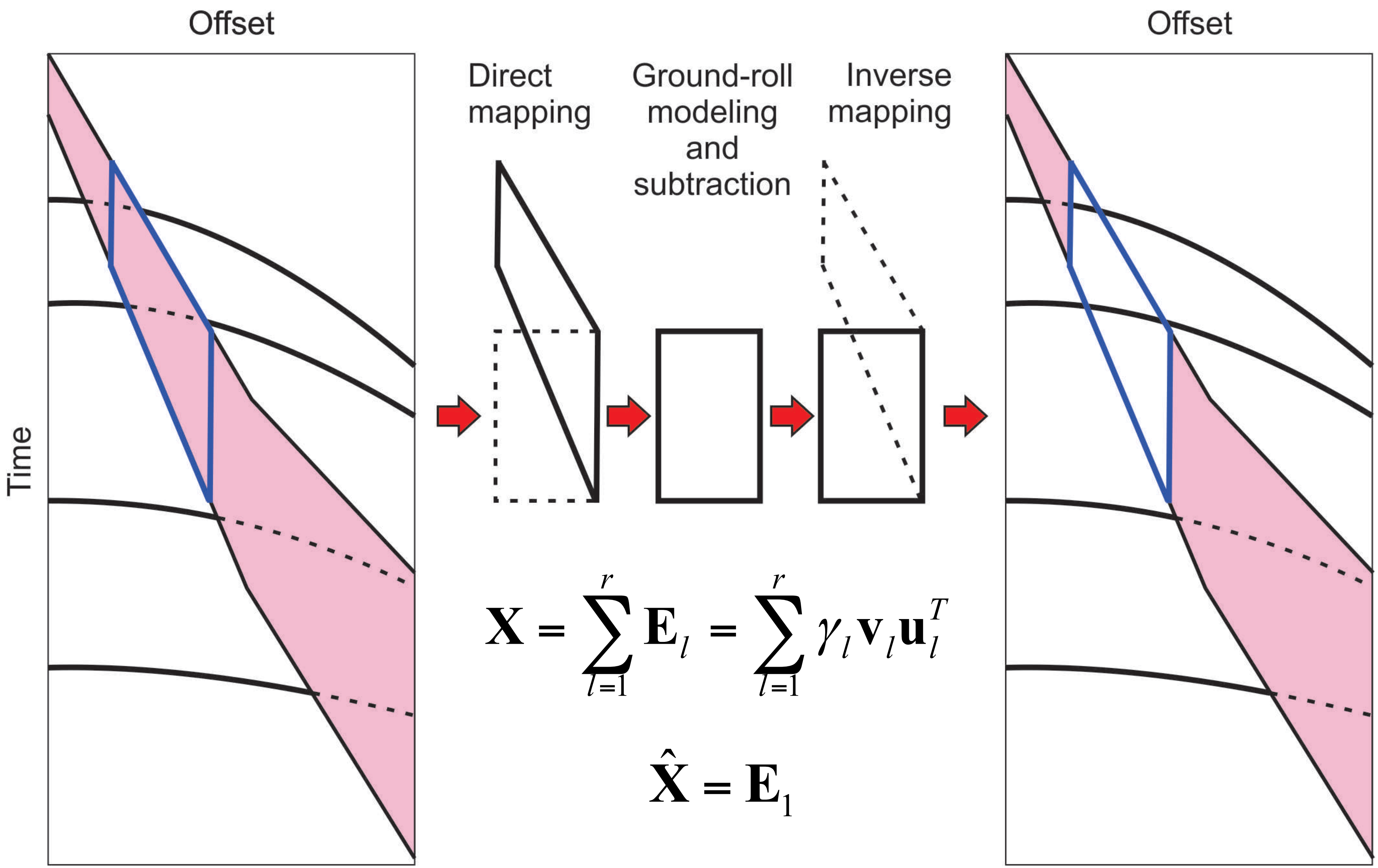
(Tyapkin et al., 2004, Liu, 1999)



Note: the method cannot handle scattered ground roll

SVD

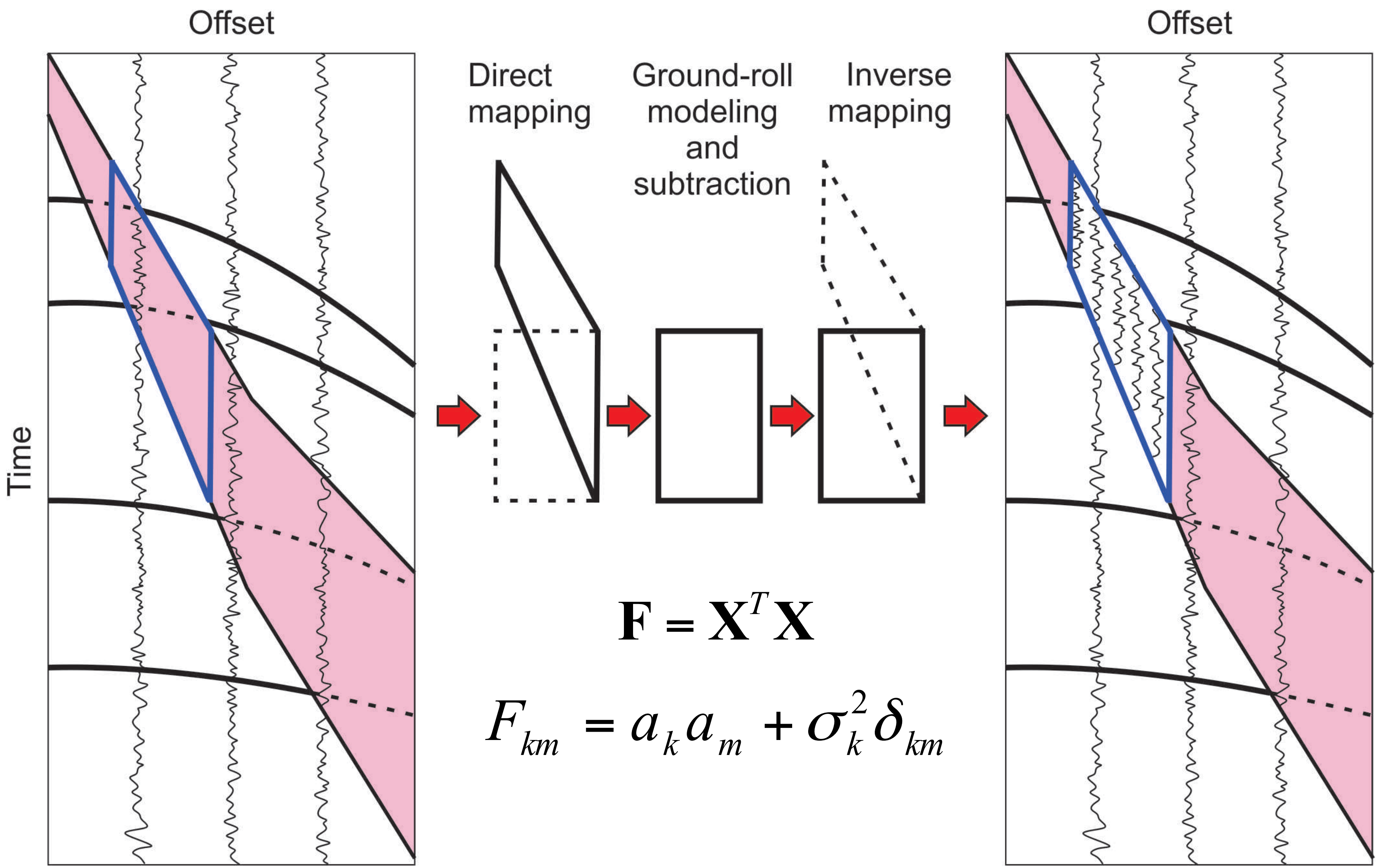
(Tyapkin et al., 2004, Liu, 1999)



$$\mathbf{X} = \sum_{l=1}^{r} \mathbf{E}_l = \sum_{l=1}^{r} \gamma_l \mathbf{v}_l \mathbf{u}_l^T$$

$$\hat{\mathbf{X}} = \mathbf{E}_1$$

SVD effect

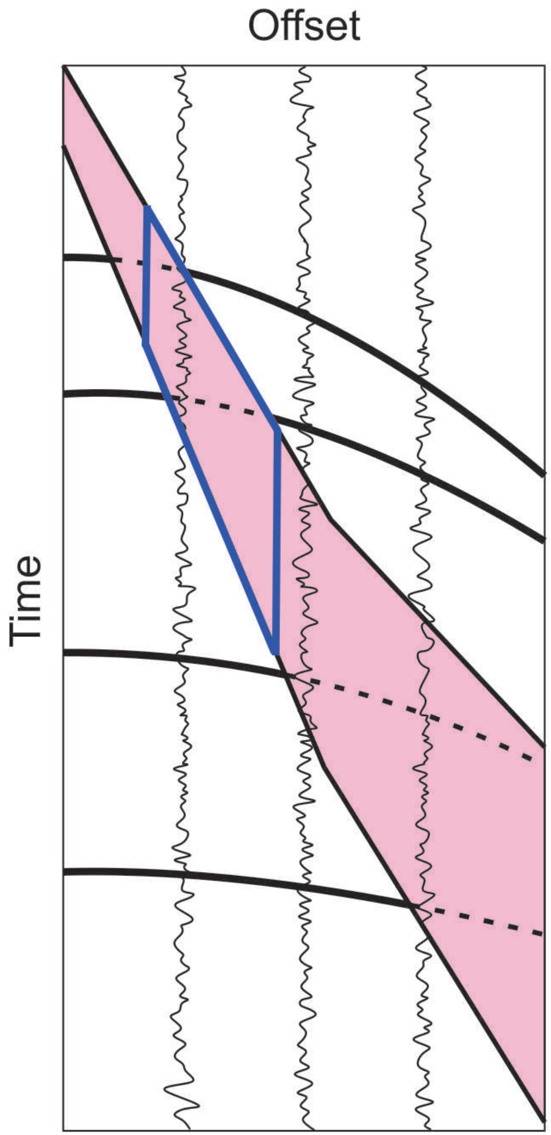


$$\mathbf{F} = \mathbf{X}^T\mathbf{X}$$

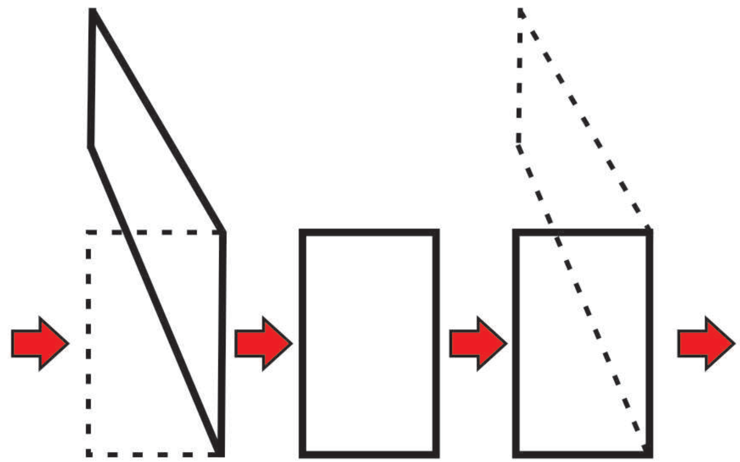
$$F_{km} = a_k a_m + \sigma_k^2 \delta_{km}$$

Optimum weighted stacking: first stage

(Tyapkin and Ursin, 2005)



$$\mathbf{X} = \mathbf{g}\mathbf{a}^T + \mathbf{N}$$



$$\hat{\mathbf{g}} = \mathbf{X}\mathbf{p}$$

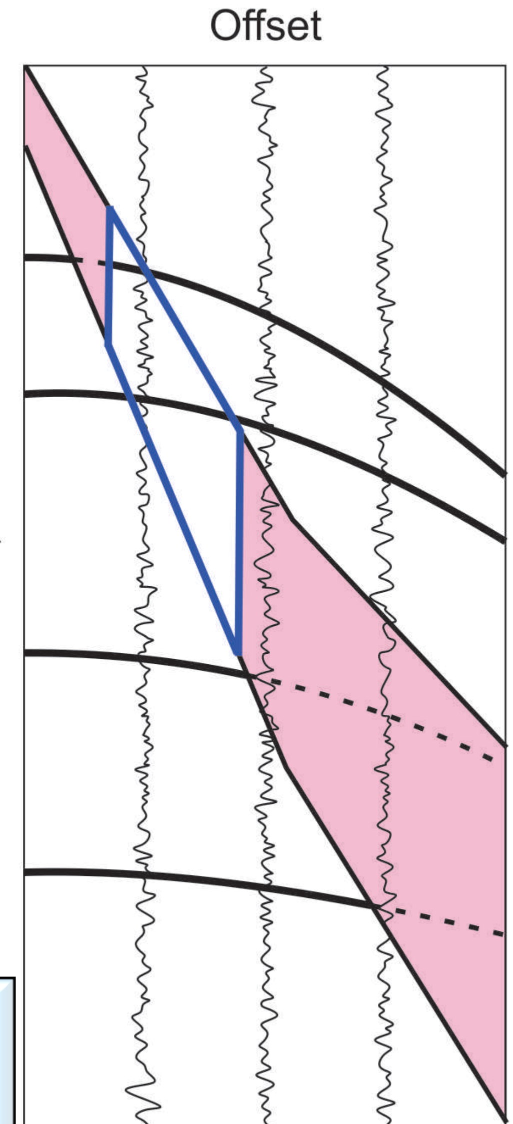
$$p_k = a_k \sigma_k^{-2} \Big/ \sum_{m=1}^{M} a_m^2 \sigma_m^{-2}$$

Gr.roll amplitudes:

$$\min_{a_k} \sum_{\substack{k,m=1 \\ k \neq m}}^{M} \left| F_{km} - a_k a_m \right|$$

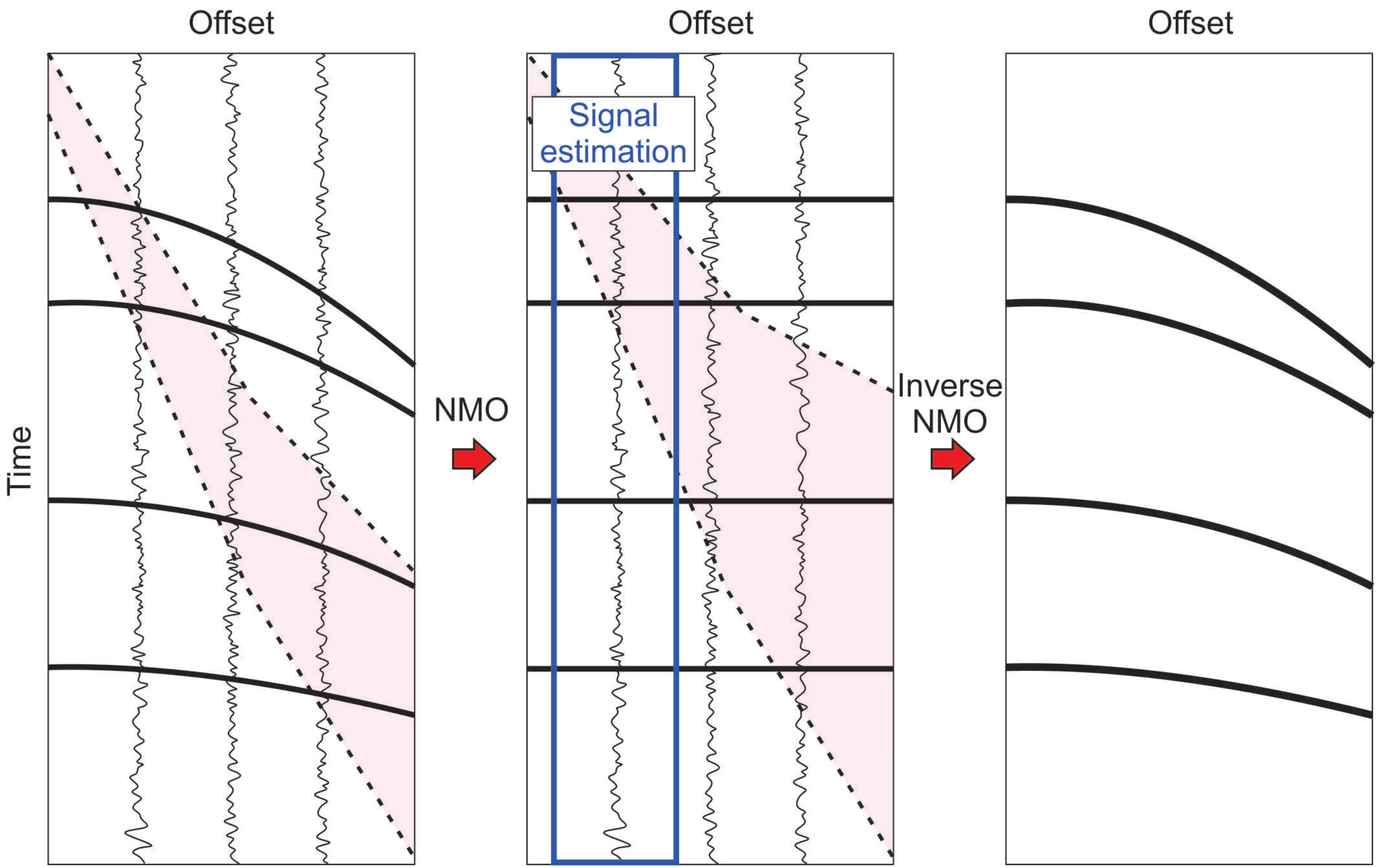
Noise variances:

$$\sigma_k^2 = F_{kk} - a_k^2$$



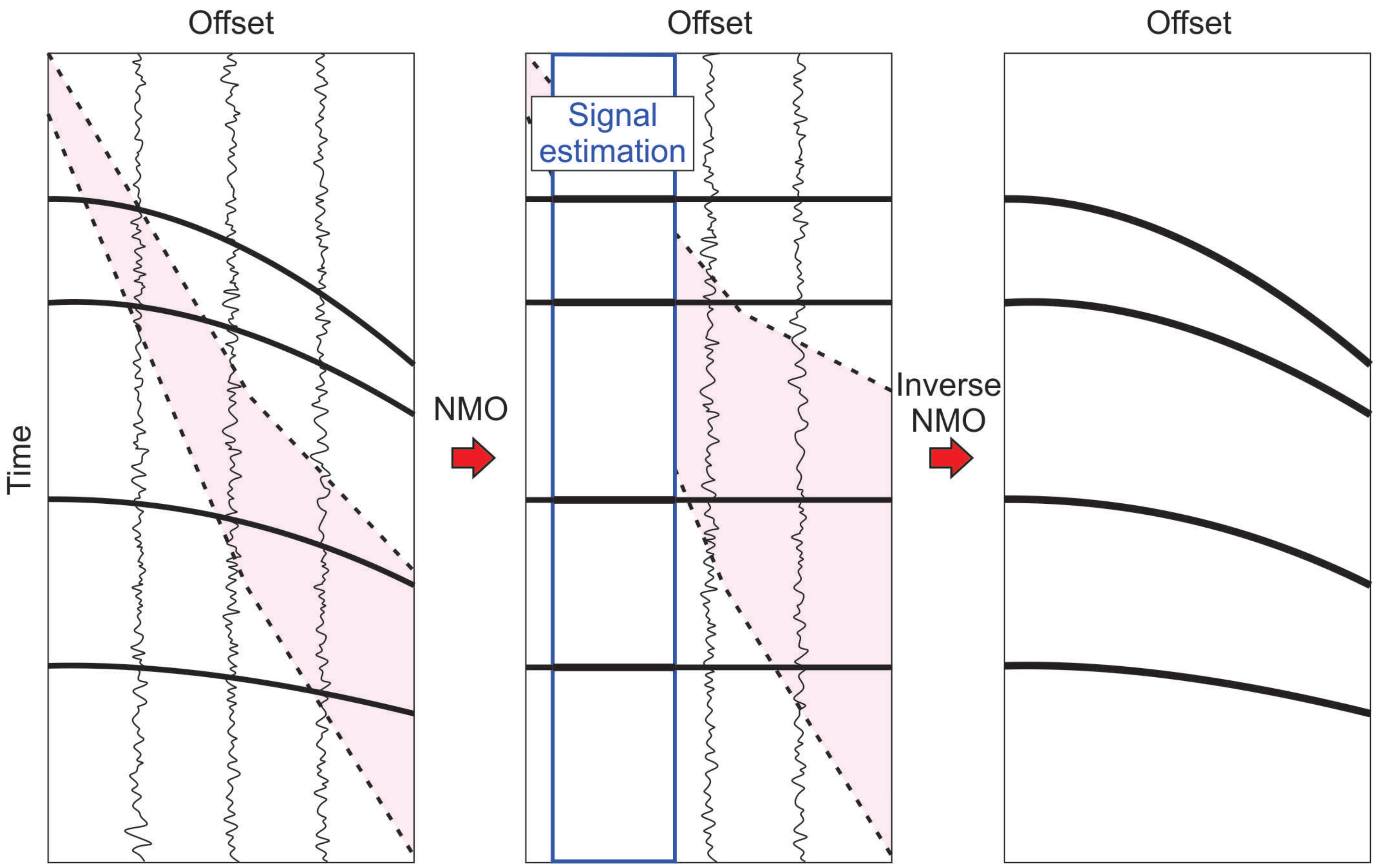
Optimum weighted stacking: second stage

(Tyapkin et al., 2010)



Optimum weighted stacking: second stage

(Tyapkin et al., 2010)



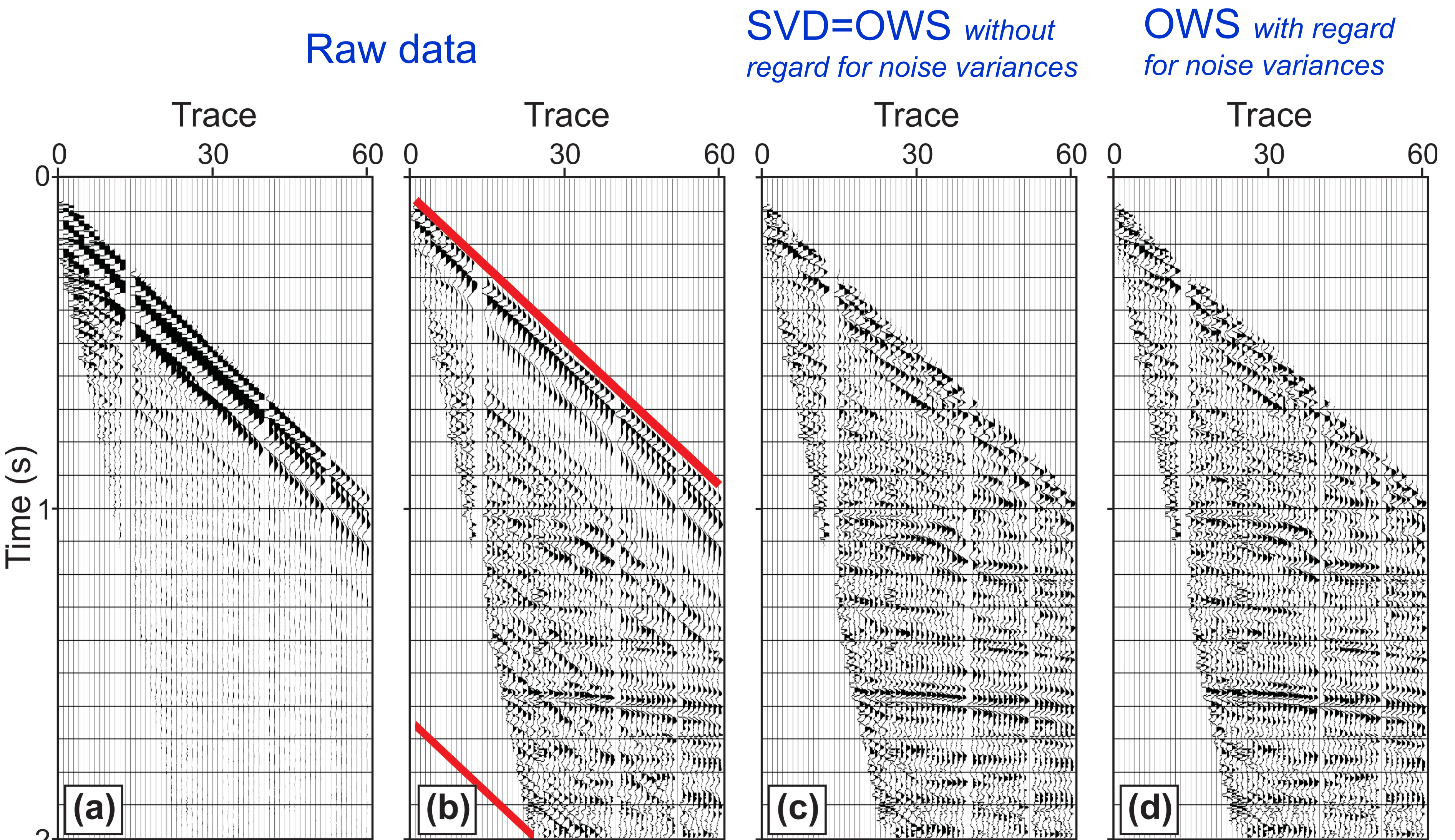
Practical examples

- Processing of 2 common-shot gathers from a 2D glacier survey, Spitsbergen, with

  - ✓ Almost non-dispersive ground roll
  - ✓ Dispersive ground roll

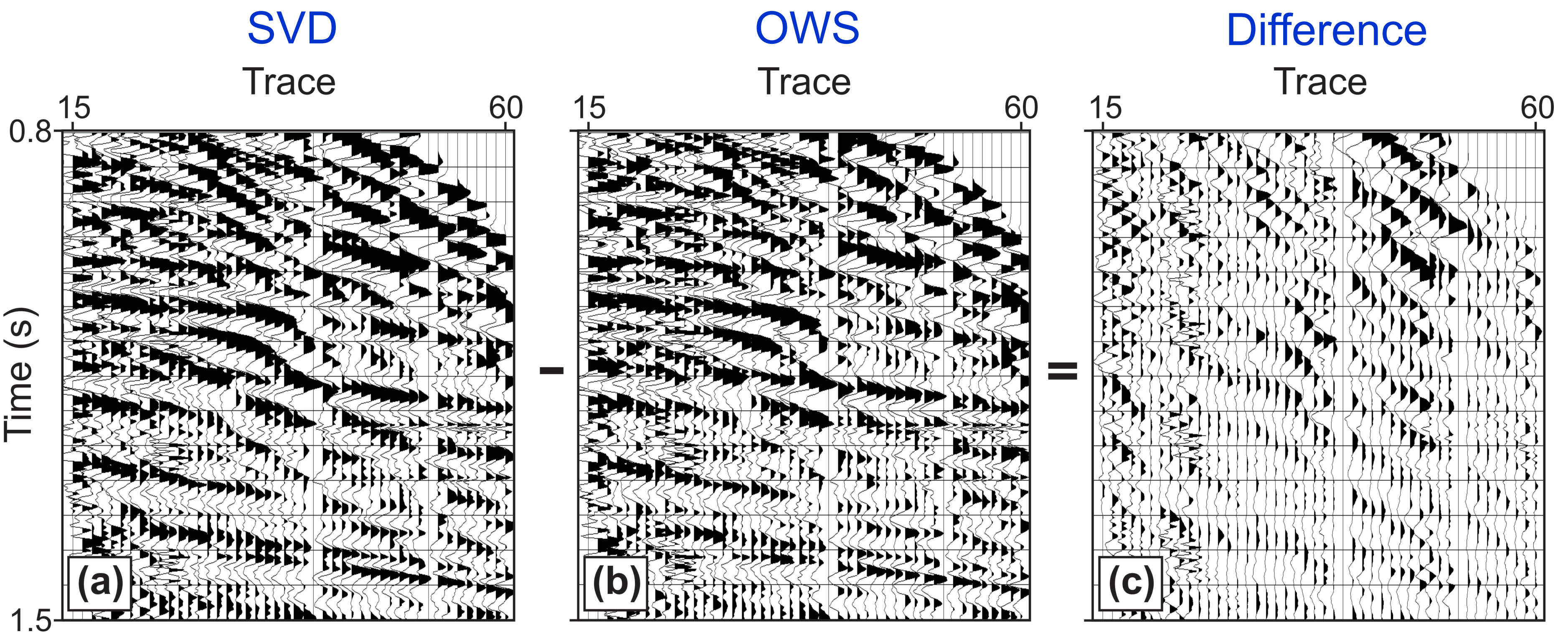
- Non-dispersive ground roll
- Dispersive ground roll

Raw record with non-dispersive ground roll and results of filtering



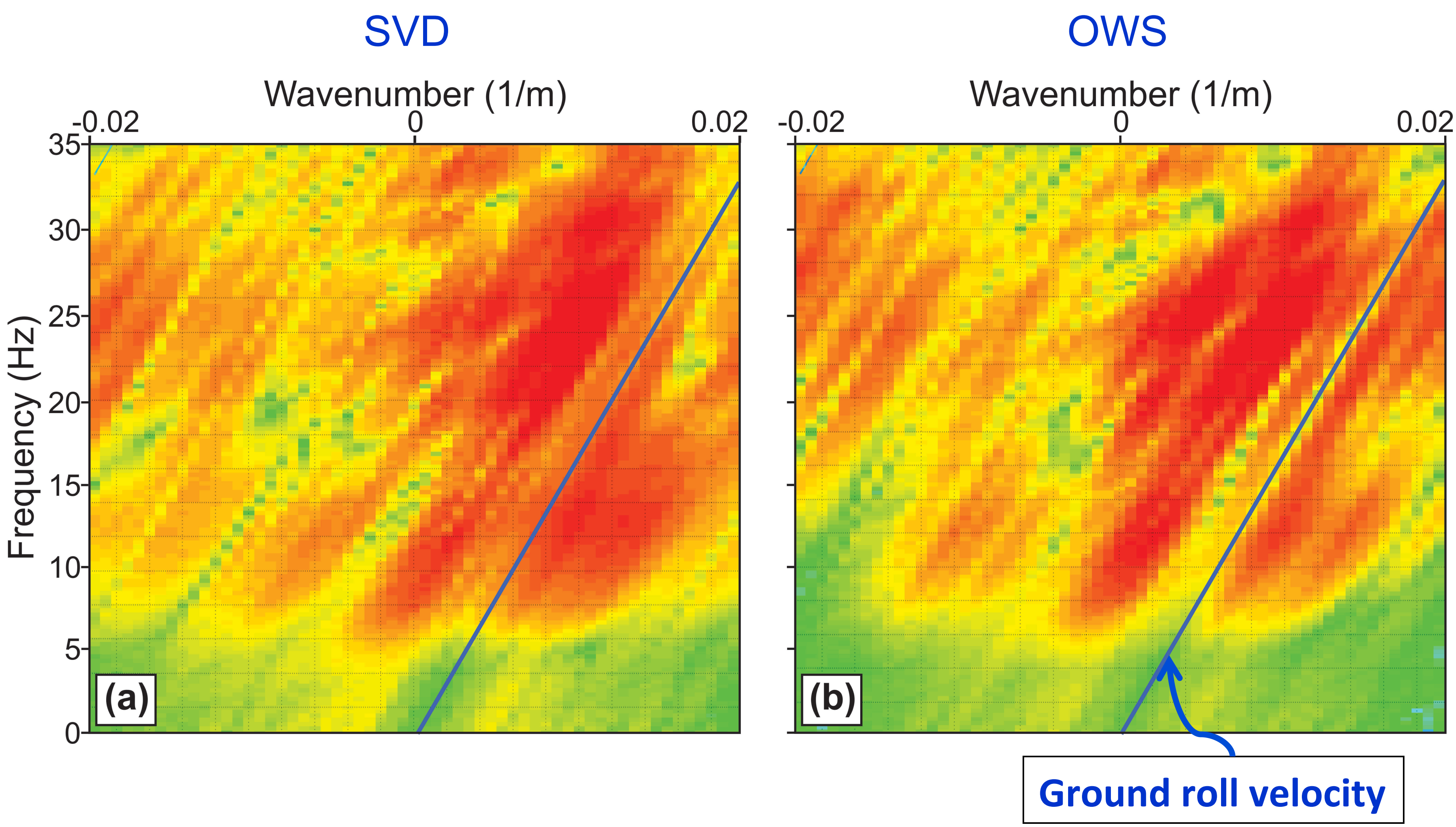
(a) Raw record without AGC; (b) Raw record with AGC and sector boundaries; (c) output of SVD-based filtering with AGC; (d) output of optimum weighted stacking with AGC

Difference between results of SVD and optimum weighted stacking



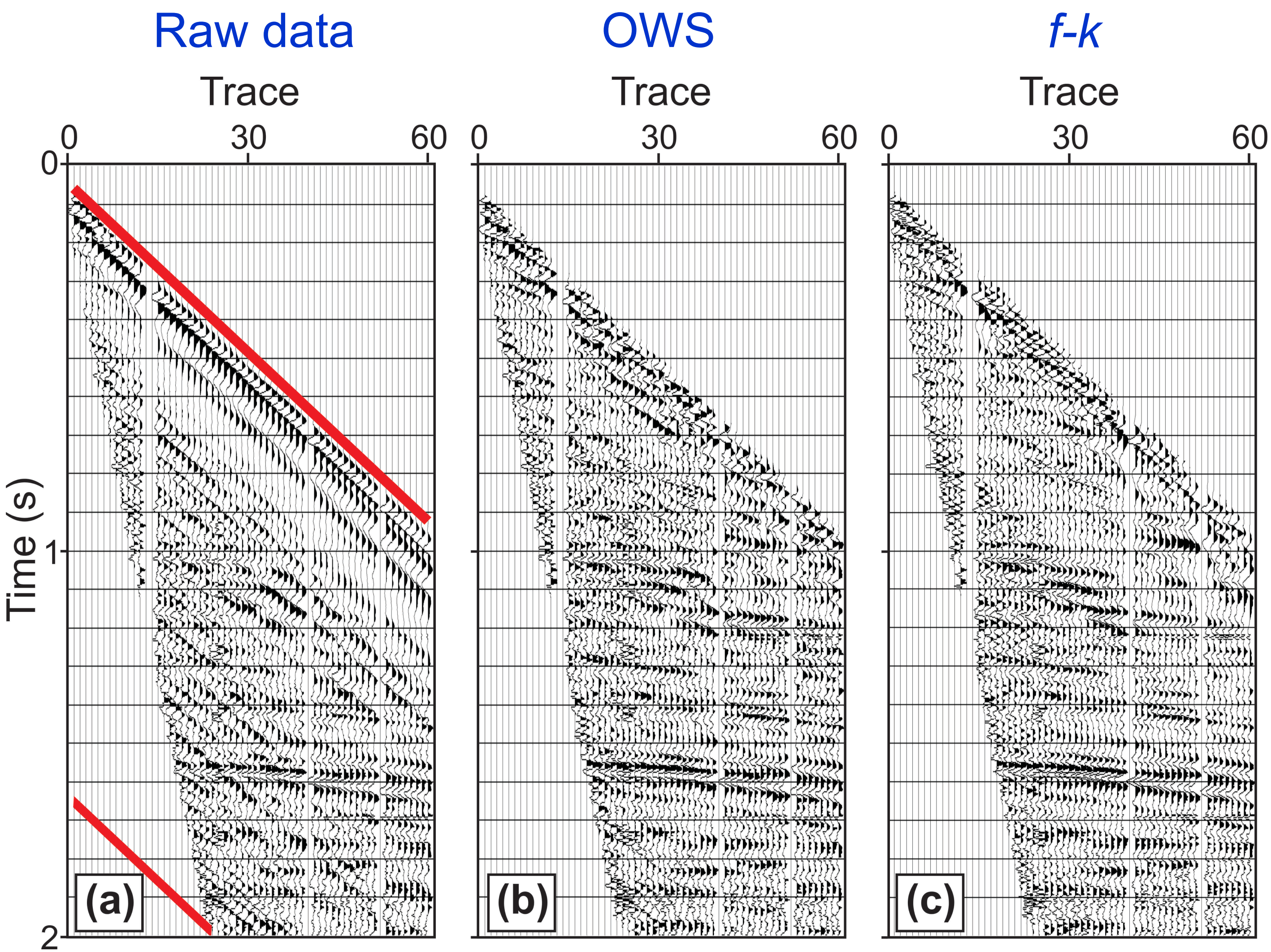
Panels of (a) output of SVD-based filtering, (b) output of optimum weighted stacking with regard for random noise variances, and (c) their difference (all without AGC)

2D spectra



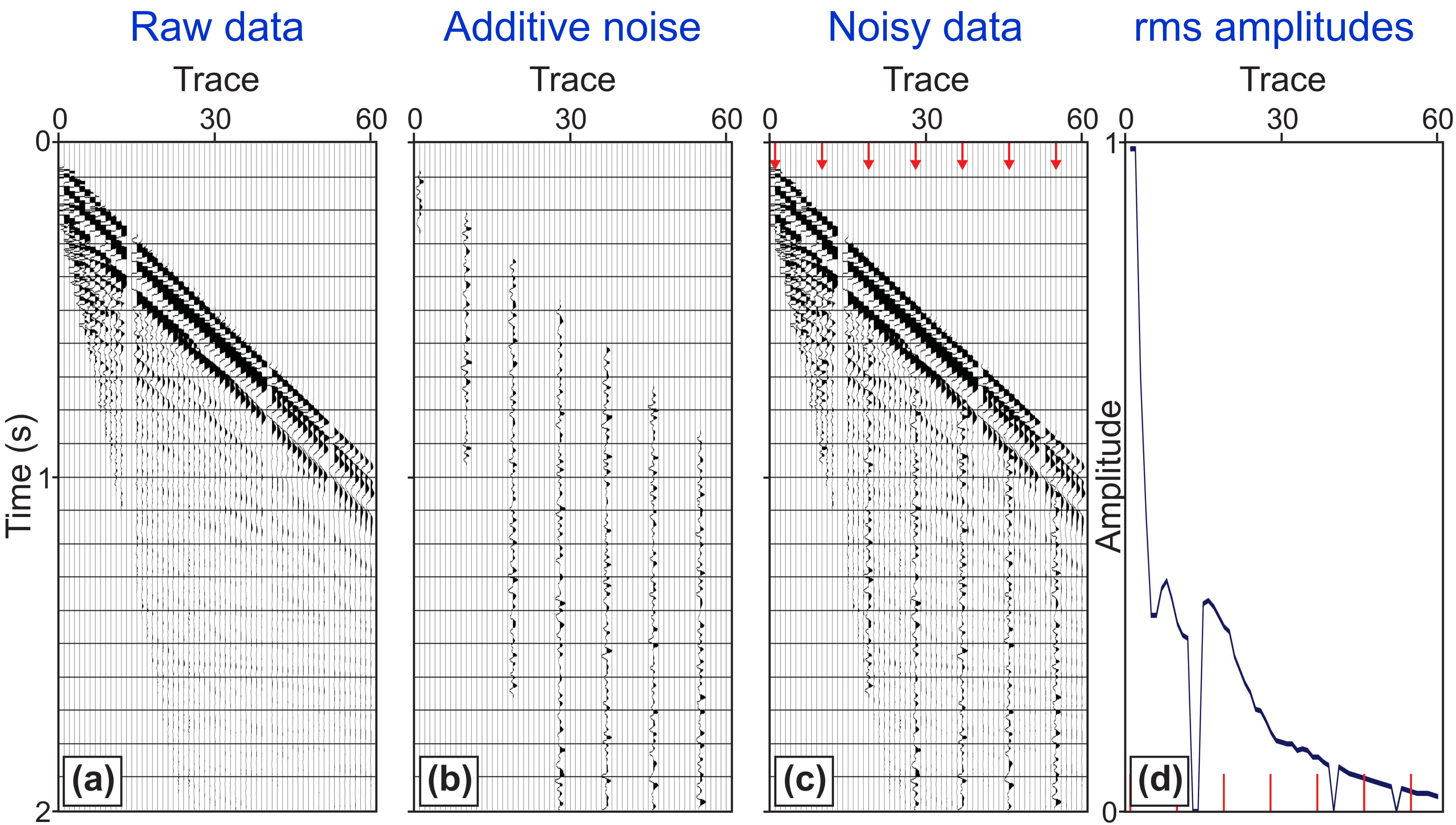
2D spectra of (a) output of SVD-based filtering and (b) output of optimum weighted stacking with regard for random noise variances

Optimum weighted stacking and *f-k* filtering



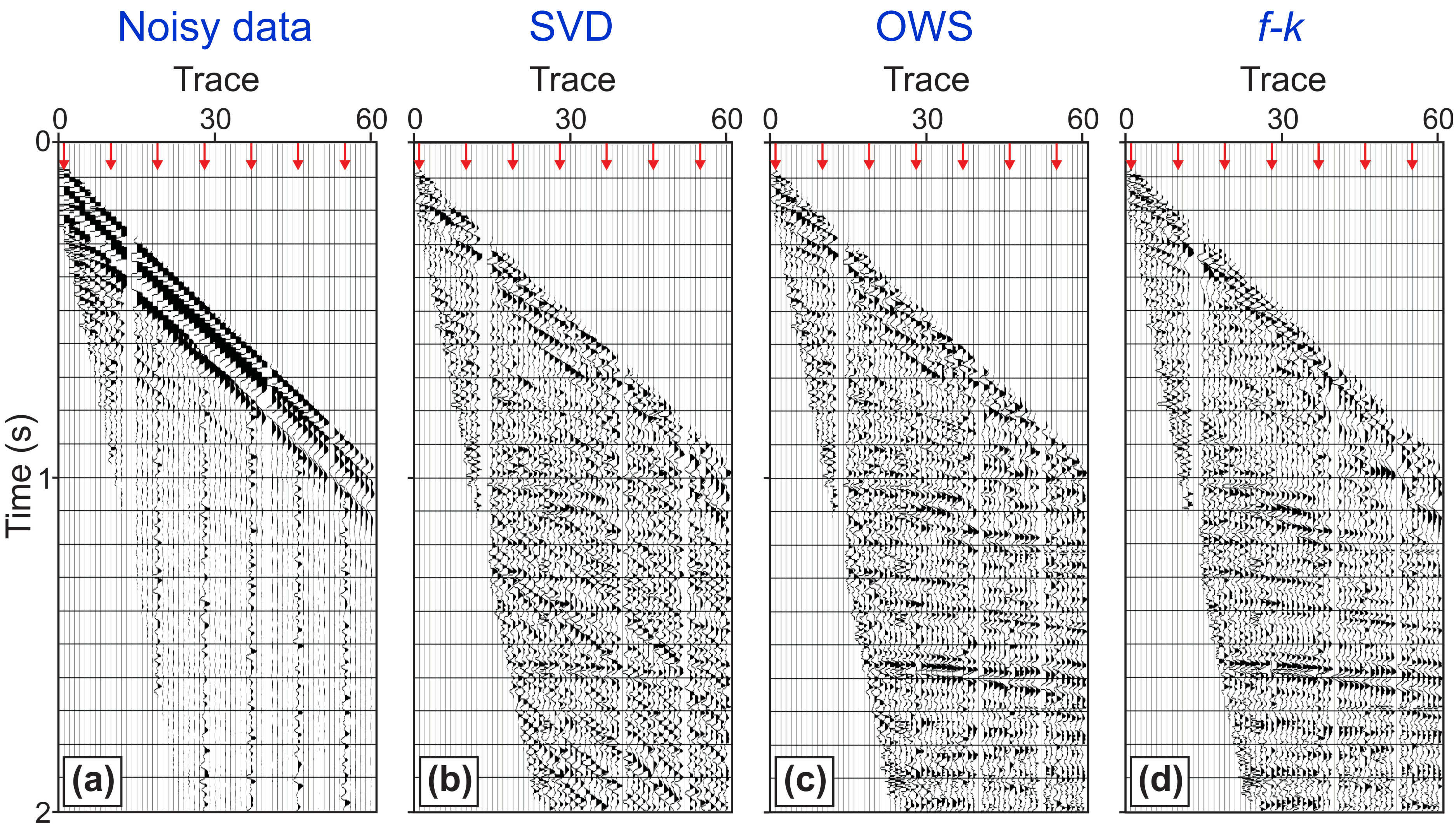
(a) Raw record; Outputs of (b) optimum weighted stacking, and of (c) *f-k* filtering (all with AGC)

Synthetic noise addition



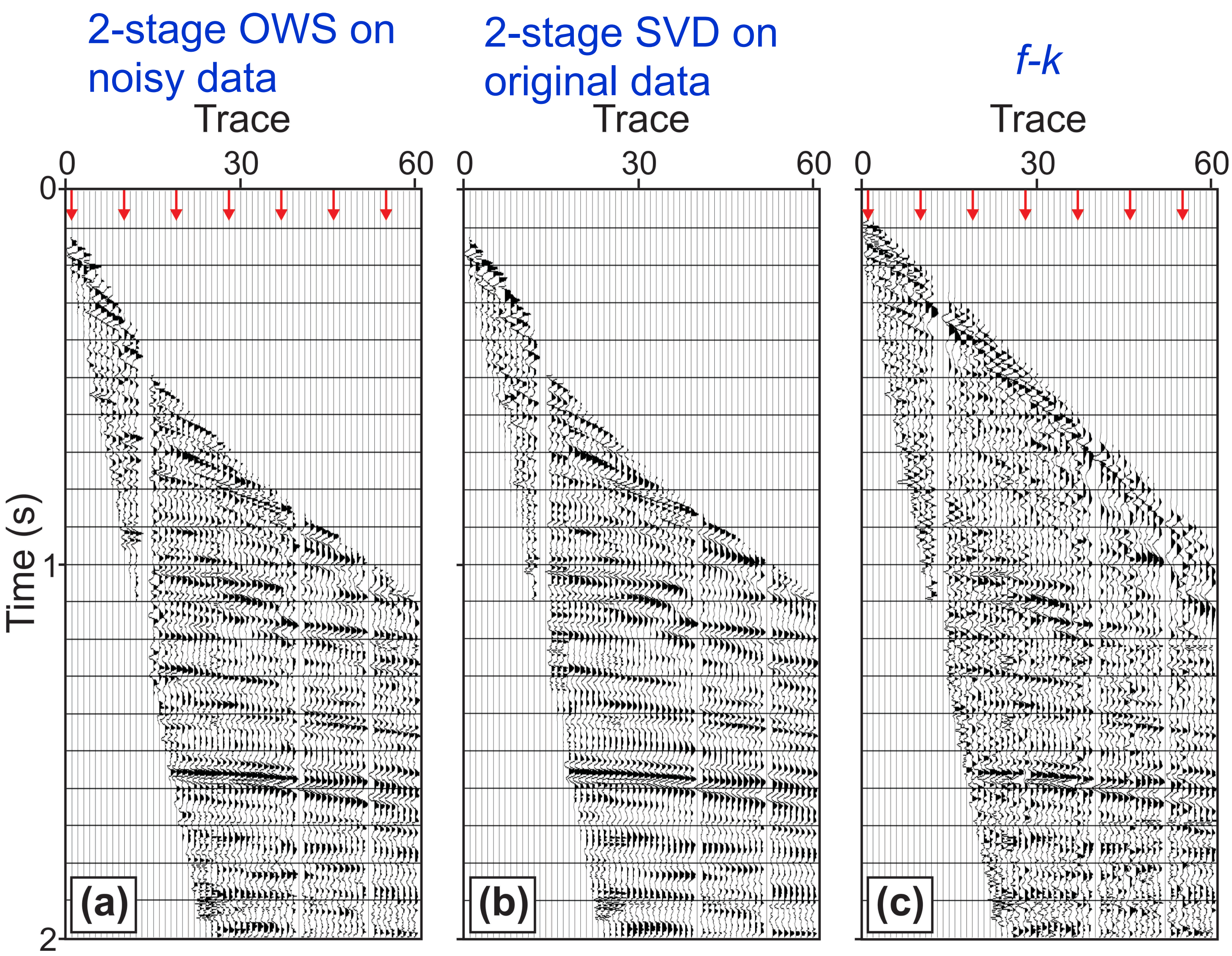
(a) Raw (‘pure’) record, (b) additive synthetic noise, and (c) noisy record (all without AGC).
(d) rms amplitudes of the raw record (blue curve) and of the additive noise (vertical red lines)

Filtering of noisy record



(a) Noisy record without AGC; Outputs of (b) SVD-based filtering, (c) optimum weighted stacking, and (d) *f-k* filtering (all with AGC)

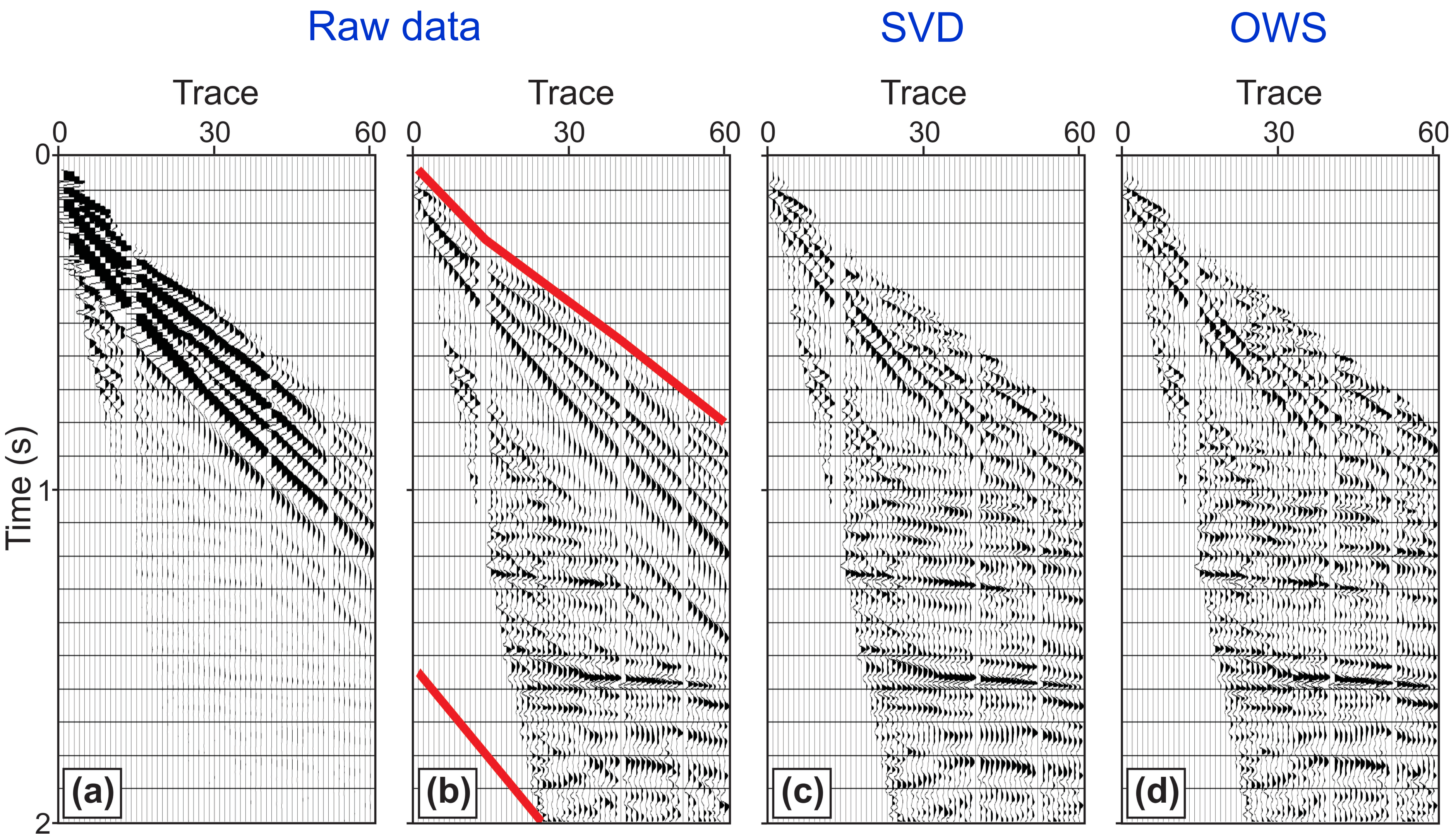
Optimum weighted stacking: second stage



Outputs of (a) the second step of optimum weighted stacking of the noisy record, (b) the second step of SVD-based filtering of the original record and (c) *f-k* filtering of the noisy record (all with AGC)

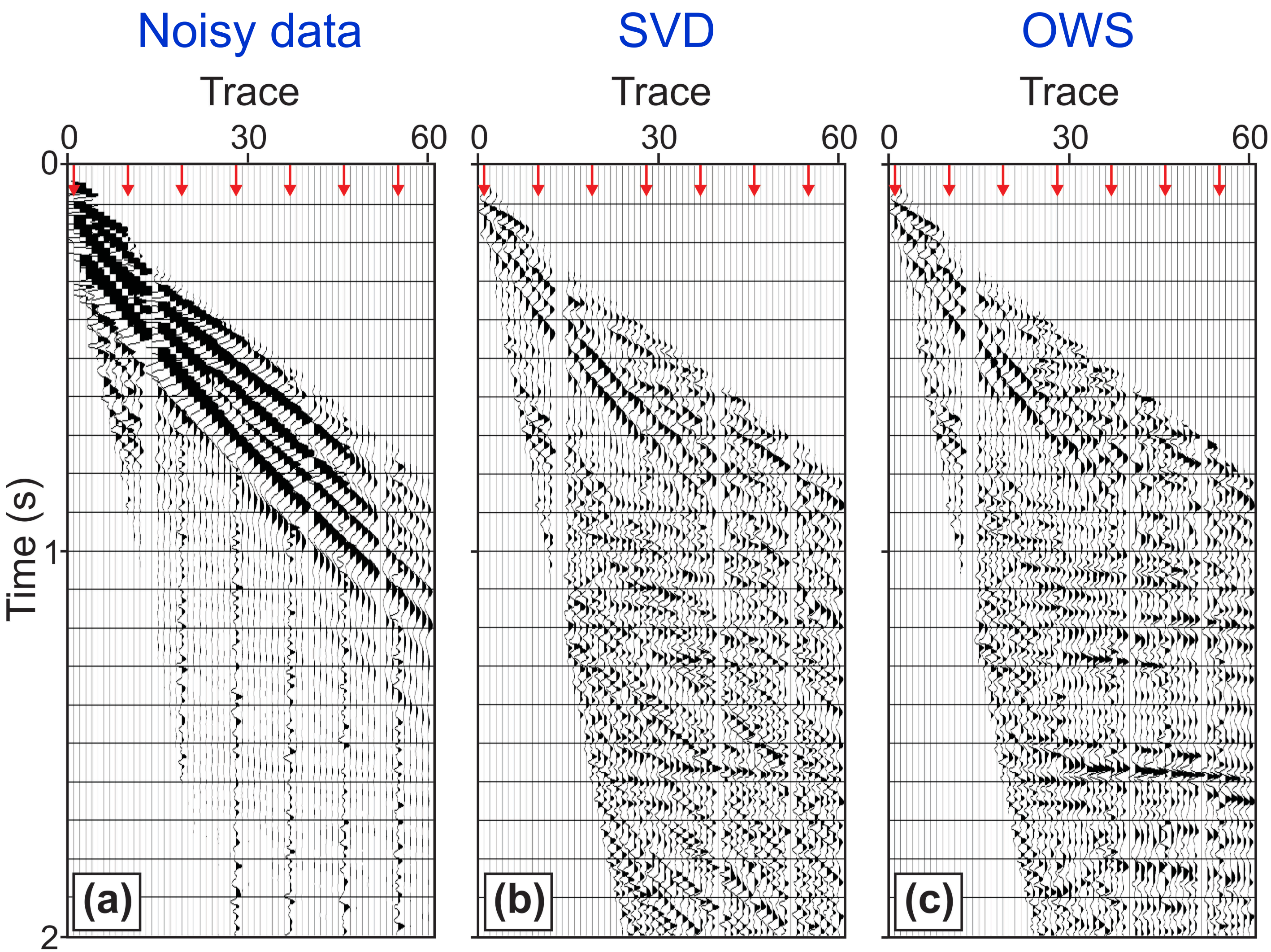
- Non-dispersive ground roll
- Dispersive ground roll

Raw record with dispersive ground roll and results of filtering



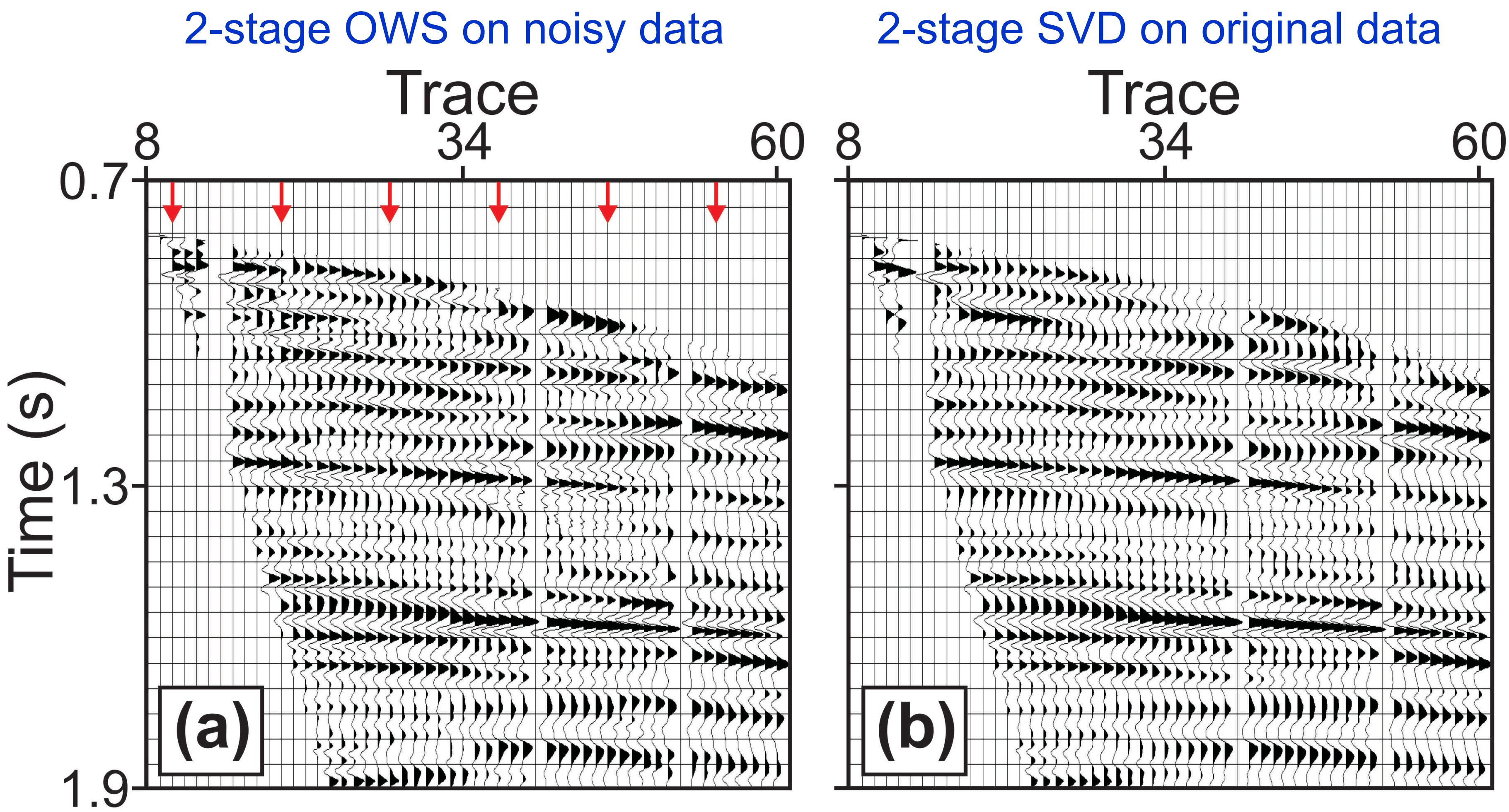
(a) Raw record without AGC; (b) Raw record with AGC and sector boundaries; (c) output of SVD-based filtering with AGC; (d) output of optimum weighted stacking with AGC

Filtering of noisy record



(a) Noisy record without AGC; Outputs of (b) SVD-based filtering and of (c) optimum weighted stacking (both with AGC)

Optimum weighted stacking: second stage



Outputs of (a) the second step of optimum weighted stacking of the noisy record and (b) the second step of SVD-based filtering of the original record (both with AGC)

Conclusions

- If the energy of additive random noise varies significantly between traces, SVD fails to subtract ground roll

- We have proposed optimum weighted stacking which accounts for variations in the noise energy across traces

- We have compared SVD-based filtering, optimum weighted stacking and f-k filtering on two common-shot gathers

- When the random noise is relatively small or stable from trace to trace, the three methods give comparable results, with SVD being slightly worse

- When the random noise varies significantly from trace to trace, two-stage optimum weighted stacking considerably outperforms SVD-based filtering and gives a better result than *f-k* filtering

Acknowledgements

- Tor Arne Johansen (University of Bergen)
- Statoil

Thank you

References

Liu, X., 1999, **Ground roll suppression using the Karhunen-Loeve transform**: Geophysics, **64**, 564-566.

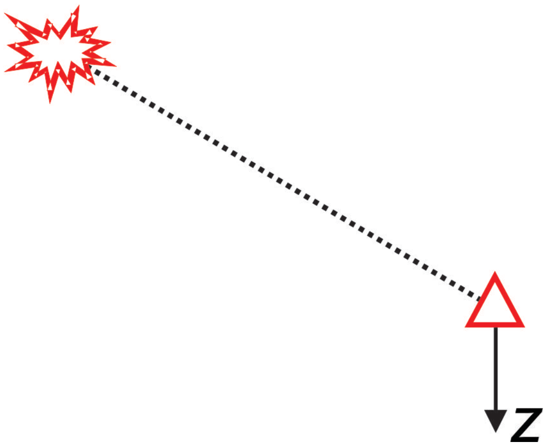
Tyapkin, Y.K., Marmalyevskyy, N.Y. and Gornyak, Z.V., 2004. **Suppression of source-generated noise using the singular value decomposition.** *66th EAGE Conference & Exhibition, Extended* Abstracts, D028.

Tyapkin, Y. and Ursin, B., 2005. **Optimum stacking of seismic record with irregular noise.** *Journal of Geophysics and Engineering*, **2,** 177–187.

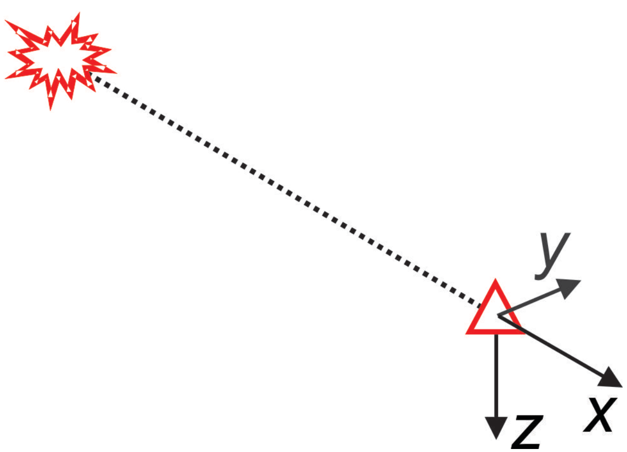
Tyapkin Yu., Ursin B., Perroud H., Silinska O. and Tiapkina O., 2010. **Zero- and first-order approximations for least-squares estimation of seismic signal with coherent and random noise.** Journal of Geophysics and Engineering**, 7**, 51–63.

Introduction

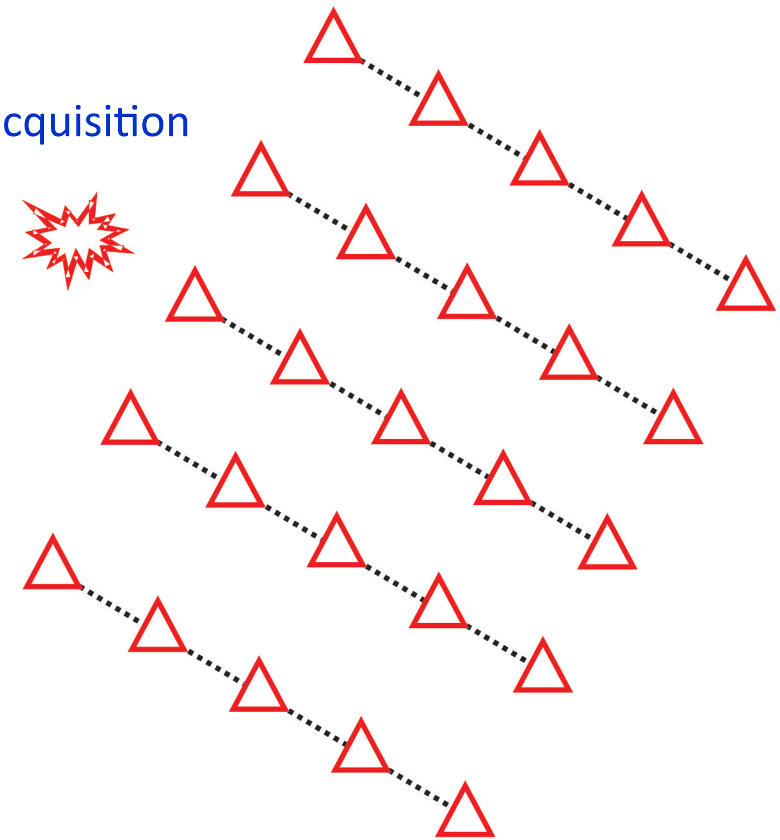
Single-component
2D data acquisition



Multi-component
2D data acquisition



3D data acquisition



Appendix

weighted stacking with a special linear pre-conditioning transform. This procedure is intended to compensate for the dispersive character of these surface waves and, therefore, to reduce the deviation of the actual data from the mathematical model adopted for both ground-roll subtraction techniques. Because these data are favorable for *f-k* filtering, we have also run it on one of the gathers for comparison. We have compared the performance of the three methods considering the two cases: when the energy of additive random noise is quite stable on different traces and when it significantly varies between the traces. In order to mimic such variations in the noise energy, some amount of synthetic noise has been added to some traces of both records. Our comparison indicates that when the random noise is small or stable from trace to trace, the three methods give comparable results, with the result of SVD-based filtering being slightly worse because it leaves some amount of residual coherent noise. When a record is corrupted by appreciable additive random noise whose energy varies significantly from trace to trace, the two-stage modification of optimum weighted stacking can considerably outperform SVD-based filtering and *f-k* filtering and should therefore be prescribed as a better choice than these conventional processes. Finally, there is no conceptual difficulty in extending the method developed in this paper for vector-sensor signal processing.

ACKNOWLEDGMENTS

We would like to express our gratitude to Tor Arne Johansen (University of Bergen) for providing the field data and allowing us to present these results. Special thanks are due to Pavlo Kuzmenko (SE Naukanaftogaz) for his help with seismic data processing. OT wishes to thank Statoil for funding her PhD. ML acknowledges financial support from NFR to the ROSE project at NTNU.

APPENDIX A

SVD-BASED FILTERING AND THE REASON OF ITS FAILURE WHEN THE ENERGY OF ADDITIVE RANDOM NOISE IS HIGHLY VARIABLE BETWEEN TRACES

Let the seismic data in a sliding window contaminated by coherent ground roll after alignment and compensation for the dispersive nature of these surface waves be given in matrix notation as

$$\mathbf{X} = \mathbf{G} + \mathbf{N}, \tag{A-1}$$

where $\mathbf{G} = \{g_{ik}\}$ and $\mathbf{N} = \{n_{ik}\}$ are, respectively, the ground-roll and additive noise of the record $\mathbf{X} = \{x_{ik}\}$, $i = 1, \ldots, L$, $k = 1, \ldots, M$; M is the number of traces and L is the number of samples per trace forming the window wherein the ground roll is modeled and then subtracted. The ground roll is assumed to have the same waveform, $\mathbf{g} = \{g_1, \ldots, g_L\}^T$, and arbitrary amplitudes, $\mathbf{a} = \{a_1, \ldots, a_M\}^T$, on different traces:

$$\mathbf{G} = \mathbf{g}\mathbf{a}^T \qquad \text{i.e.} \qquad g_{ik} = g_i a_k \tag{A-2}$$

where the superscript T denotes transpose. Note that in this model the presence of reflections is neglected.

Suppose that the additive noise is independent of the ground roll, stationary and Gaussian with a zero mean and the $ML \times ML$ positive definite covariance matrix $\mathbf{\Phi}$ having entries

$$\Phi_{ijkm} = E(n_{ik} n_{jm}), \tag{A-3}$$

where E is an expectation operator.

Taking into account the normal distribution of the additive noise, the problem of the maximum-likelihood (ML) estimation of the ground-roll waveform $\mathbf{g}$ can be formulated as the minimization of the following weighted quadratic form (Tyapkin and Ursin, 2005):

$$\sum_{i,j=1}^{L} \sum_{k,m=1}^{M} \Phi_{ijkm}^{-1} (x_{ik} - g_i a_k)(x_{jm} - g_j a_m). \quad \text{(A-4)}$$

To uniquely define the amplitude factors in (A-2) and (A-4), the energy of the ground-roll waveform should be normalized to unity:

$$\sum_{i=1}^{L} g_i^2 = 1. \quad \text{(A-5)}$$

Let the covariance matrix **Φ** have the form

$$\Phi_{ijkm} = \sigma^2 \delta_{ij} \delta_{km}, \quad \text{(A-6)}$$

which stands for an uncorrelated between traces white noise with a trace-independent variance σ^2, where δ_{ij} signifies the Kronecker delta function. Then, expression (A-4) gets simplified and the problem can be reduced to the conventional least-squares method that is equivalent to minimizing the square of the Frobenius norm (Horn and Johnson, 1986) of the data misfit (Tyapkin and Ursin, 2005):

$$\left\| \mathbf{X} - \mathbf{g}\mathbf{a}^T \right\|_F^2 = \sum_{i=1}^{L} \sum_{k=1}^{M} (x_{ik} - g_i a_k)^2. \quad \text{(A-7)}$$

To solve this, we use the SVD that permits any matrix **X** to be expressed (exactly) as a sum of specific matrices of unitary rank (Klema and Laub, 1980):

$$\mathbf{X} = \sum_{l=1}^{r} \gamma_l \mathbf{v}_l \mathbf{u}_l^T, \quad \text{(A-8)}$$

where $r = \min\{L, M\}$ and γ_l are, respectively, the rank and the lth singular value of **X**; $\mathbf{v}_l$ and $\mathbf{u}_l$ are the lth eigenvectors of the matrices $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$, respectively. The singular values γ_l are the positive square roots of λ_l, the eigenvalues of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$, while the left singular vectors $\mathbf{v}_l$ and the right singular vectors $\mathbf{u}_l$ are orthonormal ($\mathbf{v}_l^T\mathbf{v}_m = \delta_{lm}, \mathbf{u}_l^T\mathbf{u}_m = \delta_{lm}$). For convenience, the singular values are supposed to be arranged in non-ascending order: $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_r$.

The theorem of Eckart and Young (1936) asserts that the best (in a least-squares sense) approximation of the matrix **X** by another one of a lower rank, say $r_1 < r$, is attained when using the first r_1 terms in (A-8), with the others being omitted. Because in our case the approximating matrix $\mathbf{g}\mathbf{a}^T$ is of rank 1, the desired solution is the first term of the SVD of the matrix **X**,

$$\mathbf{g}\mathbf{a}^T = \gamma_1 \mathbf{v}_1 \mathbf{u}_1^T, \quad \text{(A-9)}$$

whence, in view of the condition of (A-5),

$$\mathbf{g} = \mathbf{v}_1, \quad \text{(A-10)}$$

$$\mathbf{a} = \gamma_1 \mathbf{u}_1. \quad \text{(A-11)}$$

Thus, in this case all the ground-roll parameters can be extracted directly from **X** via SVD.

Note it is the ground-roll estimate of (A-9) that is subtracted from the original record when performing SVD-based filtering.

It is instructive to demonstrate how the estimate of **g** from (A-10) relates to optimum weighted stacking considered in Appendix B. Because $\mathbf{u}_l$ are orthonormal, post-multiplication of (A-8) with $\mathbf{u}_1$ yields

$$\mathbf{X}\mathbf{u}_1 = \sum_{l=1}^{r} \gamma_l \mathbf{v}_l \mathbf{u}_l^T \mathbf{u}_1 = \gamma_1 \mathbf{v}_1 = \gamma_1 \mathbf{g}, \tag{A-12}$$

and it follows that

$$\mathbf{g} = \gamma_1^{-1} \mathbf{X}\mathbf{u}_1 = \gamma_1^{-2} \mathbf{X}\mathbf{a}. \tag{A-13}$$

Taking into account

$$\mathbf{a}^T\mathbf{a} = \gamma_1^2 \mathbf{u}_1^T \mathbf{u}_1 = \gamma_1^2, \tag{A-14}$$

equation (A-13) can be represented in the form

$$\mathbf{g} = \left(\mathbf{a}^T\mathbf{a}\right)^{-1} \mathbf{X}\mathbf{a}, \tag{A-15}$$

which is equivalent to optimum weighted stacking of equation (B-3) in Appendix B when the variance of additive random noise is trace independent (Tyapkin and Ursin, 2005).

It is thus seen from (A-10) and (A-15) that the sought-for solution can be derived either as the first principal component (Bruland, 1989) or as the output of optimum weighted stacking, considered in Appendix B. Both ways lead to the same result.

If the variance of additive random noise actually varies across the traces, which is a common occurrence, the entries of the matrix $\mathbf{F} = \mathbf{X}^T\mathbf{X}$, in view of equation (A-5), take the form

$$F_{km} = a_k a_m + \sigma_k^2 \delta_{km}, \tag{A-16}$$

where σ_k^2 is the noise variance on the kth trace.

Hence, the higher is the relative amount of the noise, the more the second term on the right-hand side of (A-16) affects the solution.

When the noise is large enough so that the presence of ground roll can be neglected, (A-16) becomes

$$F_{km} = \sigma_k^2 \delta_{km}. \tag{A-17}$$

Because **F** is now diagonal, all components of its first eigenvector $\mathbf{u}_1$ are equal to zero except for the one which is equal to unity and has the index equal to that of the trace with maximum noise variance. Thus, the SVD-based ground-roll estimator exclusively chooses such a trace, with the others being absolutely ignored. Consequently, this ground-roll estimate is entirely uncorrelated with the actual ground roll, which should be reproduced.

Let us also show how the unforeseen variations in the variance of additive random noise impact the energy of the output from the SVD-based ground-roll estimator, which is equal to the maximum eigenvalue λ_1 of the matrix **F**. As the matrices on the right-hand side of (A-16) are both Hermitian, it is convenient to make use of inequalities restricting the eigenvalues of a sum of such two matrices (Wilkinson, 1988). It yields (Tyapkin and Ursin, 2005)

$$\mathbf{a}^T\mathbf{a}+\sigma_{\max}^2 \geq \lambda_1 \geq \begin{cases} \sigma_{\max}^2 \\ \mathbf{a}^T\mathbf{a}+\sigma_{\min}^2 \end{cases}, \quad \text{(A-18)}$$

where $\sigma_{\max}^2$ and $\sigma_{\min}^2$ are the maximum and minimum variances, respectively. It follows from (A-18) that with a decreasing relative amount of ground-roll energy, the result of the SVD-based ground-roll estimation and therefore the result of subsequent SVD-based filtering become more and more dependent on the trace with maximum noise variance. In the limiting case of vanishing ground-roll, the energy of the ground-roll estimate is equal to $\sigma_{\max}^2$ and the procedure chooses this sole anomalous trace with the rest being completely ignored.

Moreover, when the variance of additive random noise on all traces is supposed to be constant, the optimum ground-roll estimate should maximize the sum of the squares of the inner products between this estimate and the individual traces (Bruland, 1989). The energy in the estimate cannot therefore be less than that in the trace of maximum energy. This also explains why the ground-roll estimate is now most influenced by the trace with highest noise energy.

APPENDIX B

ONE- AND TWO-STAGE MODIFICATIONS OF OPTIMUM WEIGHTED STACKING

Let the seismic data model adopted in Appendix A, which neglects the presence of reflections, be valid. It is, however, more realistic to suppose that the variance of additive random noise can in general vary from trace to trace in an arbitrary manner. Then, the covariance matrix of (A-3) in Appendix A takes the form

$$\Phi_{ijkm} = \sigma_k^2 \delta_{ij} \delta_{km}, \quad \text{(B-1)}$$

where σ_k^2 is the noise variance on the kth trace.

Substituting (B-1) into (A-4) yields

$$\sum_{i=1}^{L}\sum_{k=1}^{M} \sigma_k^{-2}\left(x_{ik} - g_i a_k\right)^2 . \quad \text{(B-2)}$$

This means that the problem of the ML estimation of the ground-roll waveform **g** reduces to the weighted least-squares method and is equivalent to minimizing the square of the Frobenius norm of the data misfit weighted with the reciprocals of the variances of additive random noise on different traces.

Differentiating (B-2) with respect to g_i and setting the result to zero yield the well-known formula for optimum weighted stacking (Tyapkin and Ursin, 2005):

$$g_i = \sum_{k=1}^{M} x_{ik} p_k , \quad \text{(B-3)}$$

where

$$p_k = a_k \sigma_k^{-2} \Big/ \sum_{m=1}^{M} a_m^2 \sigma_m^{-2} . \quad \text{(B-4)}$$

The effectiveness of optimum stacking is highly dependent on the accuracy of the necessary ground-roll and random-noise parameter determination, which is a key step in the overall methodology. For this purpose, we use the method presented by Tyapkin and Ursin (2005), which exploits the same data model and equation (A-16) from Appendix A. From this equation, it is seen that the off-diagonal entries of the matrix **F** depend only on the ground-roll amplitudes. As a result, the iteration scheme

$$a_k^{(i+1)} = \sum_{m \neq k} \frac{F_{km} a_m^{(i)}}{\left|F_{km} - a_k^{(i)} a_m^{(i)}\right|} \Bigg/ \sum_{m \neq k} \frac{a_m^{(i)2}}{\left|F_{km} - a_k^{(i)} a_m^{(i)}\right|}, \qquad \text{(B-5)}$$

where $a_k^{(i)}$ stands for the estimate of a_k at iteration step i, was derived for the determination of ground-roll amplitudes from the optimality criterion

$$\min_{a_k} \sum_{\substack{k,m=1 \\ k \neq m}}^{M} \left|F_{km} - a_k a_m\right|.$$

(B-6)

In this criterion, the cumulative least absolute deviation technique is applied in order to approximate the off-diagonal entries of the matrix **F** and minimize the data misfit. The explanation for choosing this criterion lies in this norm ensuring a more robust procedure than the conventional least-squares criterion when it handles certain types of errors, e.g. erratic data, and noise distribution, e.g. non-Gaussian (Claerbout and Muir, 1973). In our case, some deviations from the ideal model of the matrix **F**, which is described by equation (A-16) in Appendix A, have such a character. Many factors contribute to these deviations. Among them the main factor are reflections, which are neglected in the above-formulated record model, but distort the matrix **F** in a specific manner. Besides, residual statics and NMO, trace-to-trace variations in the ground-roll waveform, etc, are factors that cause both the actual record and the related matrix **F** to deviate from their assumed models.

Given the ground-roll amplitude estimates, the values of σ_k^2, needed for calculating the optimum weights p_k from equation (B-4), can be obtained from equation (A-16) in Appendix A as

$$\sigma_k^2 = F_{kk} - a_k^2.$$

(B-7)

If σ_k^2 is trace independent, equation (B-3) turns into (A-15). Because optimum weighted stacking accounts for trace-to-trace variations in the energy of additive random noise, it does not suffer from the above shortcoming of SVD-based filtering.

So far, for simplicity, the presence of reflections in the mathematical model of seismic data has been neglected. However, this defies reality, and now we introduce the more realistic model recently proposed by Tyapkin et al. (2010), which contains reflected waves.

Suppose that the kth trace in a sliding window that contains M traces may be written as

$$x_k(t) = b_k r\left(t - \tau_{(r)k}\right) + a_k g\left(t - \tau_{(g)k}\right) + n_k(t), \qquad k = 1, \ldots, M.$$

(B-8)

Here the signal (reflected waves) is described by the first term on the right-hand side of equation (B-8) and assumed to have an identical waveform $r(t)$ on each trace, with arbitrary trace-dependent amplitudes b_k and time delays $\tau_{(r)k}$. The amplitudes are permitted to have zero (a signal-free trace) and negative (e.g., due to the AVO effect) values. The second and third terms represent, respectively, the ground roll and additive random noise, with the above-described models. In addition, we introduce arbitrary trace-dependent time delays $\tau_{(g)k}$ into the ground roll and suppose that the signal and ground roll are stationary zero-mean Gaussian stochastic processes uncorrelated with random noise and with each other. Random noise is assumed to have trace-dependent variances σ_k^2. As well as the ground roll, the signal is supposed to have a waveform normalized to unity in order to uniquely define its amplitudes. It is worth noting that this mathematical model is a particular case of the more general one suggested in Tyapkin et al. (2010) and containing a superposition of several coherent noise wavetrains.

Given this multichannel data model, the ML estimate of the signal waveform $r(t)$ in the frequency domain reads (Tyapkin et al., 2010)

$$\hat{r} = c^{-1}\mathbf{f}^H\mathbf{D}^{-1}\left(\mathbf{I} - c_g^{-1}\mathbf{h}\mathbf{h}^H\mathbf{D}^{-1}\right)\mathbf{x}, \qquad \text{(B-9)}$$

where the scalar $\hat{r}$ is the sought-for ML estimate of the Fourier spectrum of $r(t)$; the column vector $\mathbf{x} = \{X_1^*, \ldots, X_M^*\}^H$ contains the Fourier spectra X_k of all the traces $x_k(t)$, $k = 1,\ldots,M$, with the superscripted asterisk and H standing for complex conjugation and complex (Hermitian) transpose, respectively; $\mathbf{f} = \{b_1\exp(i\omega\tau_{(r)1}), \ldots, b_M\exp(i\omega\tau_{(r)M})\}^H$, $\mathbf{h} = \{a_1\exp(i\omega\tau_{(g)1}), \ldots, a_M\exp(i\omega\tau_{(g)M})\}^H$, $\mathbf{D} = \mathrm{diag}\,\{\sigma_1^2, \ldots, \sigma_M^2\}$ with -1 signifying matrix inverse, $c = c_r - c_g^{-1}\left|c_{rg}\right|^2$, $c_r = \mathbf{f}^H\mathbf{D}^{-1}\mathbf{f} = \sum_{k=1}^{M} b_k^2\sigma_k^{-2}$, $c_g = \mathbf{h}^H\mathbf{D}^{-1}\mathbf{h} = \sum_{k=1}^{M} a_k^2\sigma_k^{-2}$, $c_{rg} = \mathbf{f}^H\mathbf{D}^{-1}\mathbf{h} = \sum_{k=1}^{M} b_k a_k \sigma_k^{-2}\exp\left[i\omega\left(\tau_{(r)k} - \tau_{(g)k}\right)\right]$; $\mathbf{I}$ is an identity matrix. For shot, hereafter the functional dependence on frequency is dropped.

The structure of (B-9) permits a straightforward interpretation. Following this equation, we should first estimate and then subtract the ground-roll wavetrain from the data $\mathbf{x}$. For this aim, the ground-roll waveform is estimated using optimum weighted stacking represented by the term $c_g^{-1}\mathbf{h}^H\mathbf{D}^{-1}\mathbf{x}$. This operation is performed with reference to the variances of random noise and the amplitudes and arrival times of this coherent noise. Then the result is multiplied by $\mathbf{h}$ in order to obtain the ultimate estimate of this coherent noise wavetrain, with its amplitudes and arrival times on all the traces. The entire procedure is represented by the second term in the parentheses on the right-hand side of (B-9).

Once the coherent noise wavetrain has been estimated and subtracted, the residual data undergo optimum weighted stacking described by the term $c^{-1}\mathbf{f}^H\mathbf{D}^{-1}$ in (B-9) and intended to obtain the final signal waveform estimate. This process is performed with regard to the variances of random noise and the amplitudes and arrival times of the signal. Multiplying the result by $\mathbf{f}$ produces the ultimate signal wavetrain estimate, with its amplitudes and arrival times on all the traces.

We call this technique two-stage optimum weighted stacking. Its first stage is equivalent to subtracting ground roll when, as considered above, reflected waves are supposed to be absent.

REFERENCES

Bekara, M., and M. van der Baan, 2007, Local singular value decomposition for signal enhancement of seismic data: Geophysics, **72**, V59–V65.

Bruland, L., 1989, A comparison of the Karhunen–Loeve stack with the conventional stack: Seismo-Series no. 42, Seismological Observatory, University of Bergen, Norway.

Chiu, S. K., and J. E. Howell, 2008, Attenuation of coherent noise using localized-adaptive eigenimage filter: 78th Annual International Meeting, SEG, Expanded Abstracts, 2541-2545.

Claerbout, J. F., and F. Muir, 1973, Robust modeling with erratic data: Geophysics, **38**, 826–844.