Sharp deconvolution in elimination of multiples

Gennady A. Ryzhikov and Marina S. Biryulina, University of Bergen, Department of Physics, Norway*

Summary

The problem of multiple suppression is of interest up to now despite of its long-standing. A new approach for attenuation of multiples in a "locally 1-D" environment is suggested, which is illustrated here with well-controlled processing of synthetic data. The method is based on a rigorous forward model and a non-Wiener filter design in the context of predictive deconvolution. Due to the linearity of the filter the method is of very low computational costs. The respective filter coefficients have quite definite physical sense: they are *primaries of a reflectivity series*. Therefore the method yields fairly good estimates of the reflection coefficients/acoustic impedance section **jointly** with elimination of multiples.

Introduction

To remove effects of unwanted signatures such as a long seismic wavelet or multiples by deconvolving them from the seismic trace, classical methods of deconvolution (Robinson and Treitel, 1980), (Webster and Levin, 1978) are implemented very successfully: *spiking* and *predictive* deconvolution respectively. Herewith we would like to show that development of mathematical aspects of solving ill-posed problems enables to contribute into the practice of seismic data processing.

Traditional multiple suppression is based on predictive deconvolution realized with the Wiener filter elaborated for stationary time series. We try to take into account that **a.** records of a pulse are highly non-stationary, **b.** information about reflections is contained not only in record cross-correlations but in autocorrelations as well: it is the starting point of the classical predictive deconvolution that the well-seen self-repetition of records is induced by distribution of reflectors.

Modified predictive deconvolution is designed with *Sharp deconvolution* (**SDec**):

- prewhitening is substituted with a *"precoloring"* which exploits an estimate of the data covariance matrix: *Self-adaptive regularization* (**SAR**)
- *an Entropy of Image Contrast* (**EnIC**) is involved: the value of the EnIC is minimal when the "image" of the earth's reflectivity is the most contrast, *sharp*, with the minimal number of reflectors.

The latter plays a role of an a priori information about filter coefficients while the SDec gives a powerful way to reduce noise-in-data effects.

Non-Wiener filter design

First of all we would like to give an example of synthetic experiment results. Based upon two media models shown with Fig. 1, `left column,` the input data (the upper "records" near the each model) are generated by convolving of impulse traces (normal incidence: Robinson code) with (zero-phase: Ricker) wavelet. The output of SDec is given with lower "records". It is well-seen that the matching of output with primaries of reflection series is perfect, with respect to not only kinematical but dynamical parameters.

Self-adaptive regularization (SAR)

The crucial part of deconvolution is the proper regularization, and to explain the relevant filter design and the necessity of the suggested regularization (SAR), let us recall a few items of regularization, in terms of linear algebra for simplicity:
given a system of linear equation with random $\mathbf{n}$ and with known covariance matrix $\mathbf{\Sigma} = E\,\mathbf{n}\mathbf{n}^{\dagger}$,

$$\mathbf{d} = \mathbf{W}\,\mathbf{r} + \mathbf{n} \tag{1}$$

find such a **vector** $\hat{\mathbf{r}}$ that minimize the following criterion:

$$\mathcal{J} = \mathbf{n}^{\dagger}\,\mathbf{\Sigma}^{-1}\mathbf{n} = (\,\mathbf{d} - \mathbf{W}\,\mathbf{r}\,)^{\dagger}\,\mathbf{\Sigma}^{-1}\,(\,\mathbf{d} - \mathbf{W}\,\mathbf{r}\,) \tag{2}$$

The normal equations $\nabla\mathcal{J} = \mathbf{0}$ have the form:

$$\mathbf{W}^{\dagger}\,\mathbf{\Sigma}^{-1}\mathbf{W}\,\mathbf{r} - \mathbf{W}^{\dagger}\,\mathbf{\Sigma}^{-1}\mathbf{d} = \mathbf{0} \tag{3}$$

which we rewrite as

$$\mathbf{\Phi}\,\mathbf{r} - \check{\mathbf{r}} = \mathbf{0} \tag{4}$$

with the Fisher's matrix/operator $\mathbf{\Phi} = \mathbf{W}^{\dagger}\,\mathbf{\Sigma}^{-1}\mathbf{W}$ and "back-projected" $\check{\mathbf{r}} = \mathbf{W}^{\dagger}\,\mathbf{\Sigma}^{-1}\mathbf{d}$.

It is the associated with $\mathbf{r}$ matrix $\mathbf{\Phi}$ that "creates" the troubles: there exist an infinite number of such a $\tilde{\mathbf{r}}$ that $\mathbf{\Phi}\tilde{\mathbf{r}} \approx \mathbf{0}$, i.e. $\mathbf{\Phi}(\mathbf{r}+\tilde{\mathbf{r}}) \approx \mathbf{\Phi}\mathbf{r}$, and therefore a computer will find one of $(\mathbf{r}+\tilde{\mathbf{r}})$ instead of $\mathbf{r}$ with the unit probability. Quantitatively: because $\mathbf{\Phi}$ is ill-conditioned, $\mathtt{cond}\,(\mathbf{\Phi}) = \lambda_{\max}/\lambda_{\min} \sim 10^5{-}10^{12}$ (example is given with Fig. 2), any high-frequency noise in $\check{\mathbf{r}}$ or in $\mathbf{\Phi}$ itself will be amplified with the factor $\mathtt{cond}\,(\mathbf{\Phi})$[1].

`Regularization` aims to suppress a strong instability (Tikhonov and Arsenin, 1977), (Franklin, 1970) while solving ill-posed problems, e.g., multiple artifacts while imaging.
The well-known changing of criterion $\mathcal{J}$ (2) with

$$\mathcal{J}_1 = \mathcal{J} + \Delta\mathcal{J} = \mathbf{n}^{\dagger}\,\mathbf{\Sigma}^{-1}\mathbf{n} + \alpha\,\mathbf{r}^{\dagger}\mathbf{H}\,\mathbf{r} \tag{5}$$

[1] Note here, that eq. 4 with data covariance (Toeplitz) matrix as $\mathbf{\Phi}$, $\mathbf{r}$ as Wiener filter coefficients and with cross-correlator $\check{\mathbf{r}}$ is the master equation of predictive deconvolution

Sharp Deconvolution

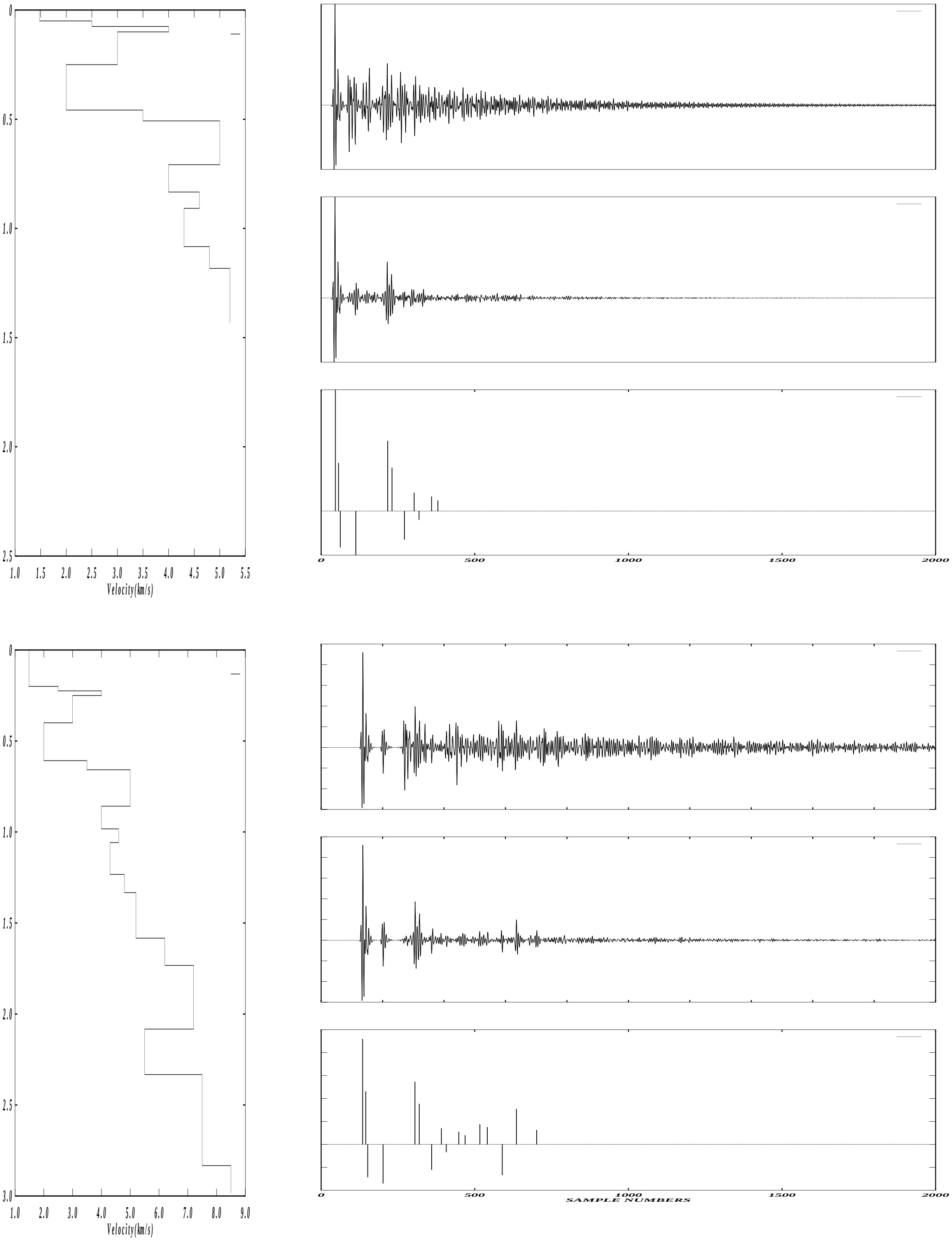

Fig. 1: Sharp deconvolution input-output. `Left column:` 2 velocity models `Top:` MODEL **"Shallow water".** `Below:` MODEL **"Long primaries' trace".** `Right column:` Traces `from top to bottom:` **input, output,** *primaries of impulse trace.*

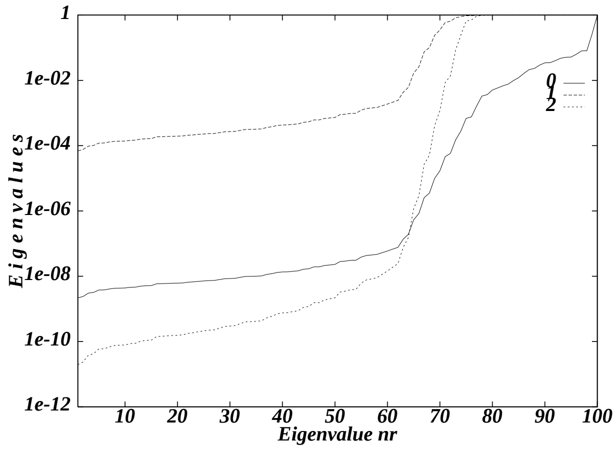

Fig. 2: Comparison of the filtering strength of conventional (**CR**) and self-adaptive (**SAR**) regularizations. Here **eigenvalue numbers** of the autocorrelation matrix $\mathbf{\Phi}$ for the Model "Shallow water" (Figure 1) are larger for larger eigenvalues. The latters are given by solid line (**0**). Eigenvalues of the filter operator $\mathbf{F}_\alpha$ for **CR** are represented with line **1**, and for **SAR** with line **2**.

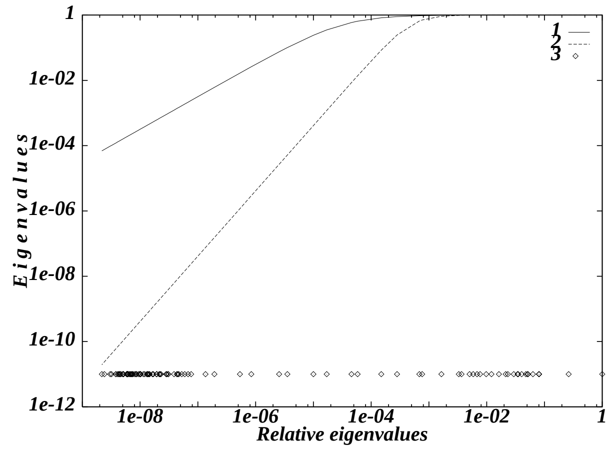

Fig. 3: $\alpha-$Filter Operator spectra for **CR** (line **1**) and for **SAR** (line **2**) in a scale of relative eigenvalues of the autocorrelation matrix $\mathbf{\Phi}$. The distribution of $\mathbf{\Phi}$-eigenvalues is marked along x-axis by **3**. It is well-seen how strongly the **SAR** eliminates all *non-informative* components ($\lambda/\lambda_{\max} < 10^{-4}$).

with a positive definite matrix $\mathbf{H} > 0$ does *regularize* the problem (2). The parameter of regularization α should be small enough: it defines the "penalty function" $\Delta\mathcal{J}$ in (5) and regulates a trade-off between a *bias* and a *noise*. Instead of (4) normal equations become now

$$(\mathbf{\Phi} + \alpha \mathbf{H})\,\mathbf{r} - \check{\mathbf{r}} = \mathbf{0} \qquad (6)$$

and due to $\mathbf{H} > 0$ the matrix $\mathbf{\Phi} + \alpha\mathbf{H}$ is better conditioned than $\mathbf{\Phi}$.

<u>`Conventional regularization`</u> (**CR**) selects the very explicit matrix $\mathbf{H}$ in the "penalty function" $\Delta\mathcal{J}$ (5): $\mathbf{H} = \mathrm{I}$ (*prewhitening* in terms of the Wiener filtering), so the criterion (5) looks like

$$\mathcal{J}_1 = \mathbf{n}^\dagger \mathbf{\Sigma}^{-1}\mathbf{n} + \alpha\,\mathbf{r}^\dagger \mathrm{I}\,\mathbf{r} \qquad (7)$$

which leads to the normal equations [2]

$$(\mathbf{\Phi} + \alpha\,\mathrm{I})\,\mathbf{r} - \check{\mathbf{r}} = \mathbf{0} \qquad (8)$$

<u>`Self-adaptive regularization`</u> (**SAR**), suggested here, is based on a **function on $\mathbf{\Phi}$**, that does not change the **$\mathbf{\Phi}$-eigenvectors.** It allows us to control easy the SAR-effect by α-parameter, which is dependent just on the noise level in $\check{\mathbf{r}}$ and/or desired accuracy of $\hat{\mathbf{r}}$ (6).

Formally [3] SAR deals with the regularizing matrix $\mathbf{H}$ in the form $\mathbf{H} = \alpha\,\mathbf{\Phi}^{-1} > 0$, so the normal equations look like

$$(\mathbf{\Phi} + \alpha\,\mathbf{\Phi}^{-1})\,\mathbf{r} - \check{\mathbf{r}} = \mathbf{0}, \qquad (9)$$

or

$$\mathbf{\Phi}^{-1}(\mathbf{\Phi}^2 + \alpha\,\mathrm{I})\,\mathbf{r} - \check{\mathbf{r}} = \mathbf{0} \qquad (10)$$

and the solution $\check{\mathbf{r}}$ is

$$\hat{\mathbf{r}} = (\mathbf{\Phi}^2 + \alpha\,\mathrm{I})^{-1}\,\mathbf{\Phi}\,\check{\mathbf{r}} \equiv \mathcal{R}_\alpha\,\check{\mathbf{r}} \qquad (11)$$

The remarkable feature of the $\mathcal{R}_\alpha$-family is that the latter yields much stronger regularization than the CR, have the power not less than the TR, [4] but in contrary to the TR its power is under control. In contrary to the CR, the relative SAR-solution **does not contain the null-space components**. Indeed, if matrix $\mathbf{\Phi}$ is degenerated and has an eigenvalue λ being null, then the SAR-inverse operator eliminates null-space components at all: $\lambda/(\lambda^2+\alpha) \overset{\lambda \to 0}{=} 0$, while for CR $1/(\lambda+\alpha) \overset{\lambda \to 0}{=} 1/\alpha$ (see, e.g. Fig. 2).

<u>`α−Filter Operator`</u> ($\mathbf{F}_\alpha$) for linear inverse problems seems to be a useful instrument for the comparative analysis of regularizations of different kind.
We define the $\alpha-$Filter Operator [5] by the equation

$$\hat{\mathbf{r}} = \mathcal{R}_\alpha\,\check{\mathbf{r}} \equiv \mathbf{F}_\alpha\,\mathbf{r}_0 \qquad (12)$$

where $\mathbf{r}_0 \equiv \mathbf{\Phi}^{-1}\check{\mathbf{r}}$ is an *unregularized*-, or LMS-solution of the Eiler equation 3, i.e.

$$\mathbf{F}_\alpha = \mathcal{R}_\alpha\,\mathbf{\Phi}^{-1}$$

and it defines the *filtering* features of any regularizing α-family $\{\mathcal{R}_\alpha\}$ with respect to the unregularized inversion of the Fisher operator $\mathbf{\Phi}$.

---

[2] The regularizer $\mathbf{H} = \mathrm{I} - \mathbf{D^2} > \mathbf{0}$ with *differential* operator $\mathbf{D^2}$ leads to Tikhonov regularization **TR** (Tikhonov and Arsenin, 1977). It implies the *ordering* of the $\mathbf{r}$-vector components, although it would be preferable if the regularization had an invariant form, e.g., with respect to permutation of $\mathbf{r}$-components.

[3] for simplicity we suppose here that $\mathbf{\Phi}$ has its inverse $\mathbf{\Phi}^{-1}$

[4] SAR can be written in terms of operators in Hilbert space: the $\mathcal{R}_\alpha$-family gives an α-approximation of the *unbounded* operator $\mathbf{\Phi}^{-1}$ by the family of the *compact* operators, and yields a strong convergence in a factor-space $(\cdot)/N(\mathbf{\Phi})$, if the null-space N of the operator $\mathbf{\Phi}$ is not empty.

[5] here again: just formally, in supposition that $\mathbf{\Phi}^{-1}$ exists

As long as the CR and SAR do not change the eigenvectors of $\mathbf{\Phi}$, we can write the $\mathbf{F}_\alpha$-eigenvalues via $\mathbf{\Phi}$-eigenvalues (Figs 2 and Fig. 3):

- for CR every eigenvalue of $\mathbf{\Phi}$ is mapped to the corresponding eigenvalue of $\mathbf{F}_\alpha$ by the following way (Fig. 2 and Fig. 3, the curve **1**) $\lambda \Rightarrow \lambda/(\lambda + \alpha)$
- for SAR (Fig. 2 and Fig. 3, the curve **2**) $\lambda \Rightarrow \lambda^2/(\lambda^2 + \alpha)$: the larger is the condition number of a matrix to be inverted the stronger is (self-adaptive) regularization (see fig. 3)

`Precoloring versus prewhitening` in predictive deconvolution: comparison of equations 8 and 9 gives the way to interpret the regularizer in SAR as an *a priori* information about filter coefficients $\mathbf{r}$, i.e. a *precoloring.*

Entropy of image contrast

The first aim of the Sharp Deconvolution is to *suppress* as much as possible all high-frequencies, associated with the noise, from the "image" $\hat{\mathbf{r}}$ of reflection series. The second part of the SDec is the concept of *revealing* the high-frequency components: to find such a function of "image" inside the confidential ellipsoid that has the most "contrast" shape (i.e. it should be represented with minimal number of non-zero parameters).

To maintain this problem we use the function that we introduced earlier for the 3-D nonlinear waveform inversion (Ryzhikov et al., 1995): *Entropy of image contrast* (**EniC**). The EnIC is a functional which quantifies contrast of a function $\mu(\mathbf{x})$ defined over a region Ω in a (multidimensional) space. The EnIC is defined in terms of the entropy [6]

$$E = -\int_\Omega p(\mathbf{x}) \, \ln p(\mathbf{x}) \, d\mathbf{x} \qquad (13)$$

of the pseudo-probability density function (*pseudo-pdf*)

$$p(\mathbf{x}) = (\nabla\mu(\mathbf{x}))^2 / \int_\Omega (\nabla\mu(\mathbf{x}))^2 d\mathbf{x} \qquad (14)$$

Large contrast of $\mu(\mathbf{x})$ - or high concentration of $|\nabla\mu(\mathbf{x})|^2$ - corresponds to low values of EnIC. EnIC is non-negative and assumes minima on suitably constrained sets of functions.

In the context of the predictive deconvolution gradients of image $\mu(\mathbf{x})$) are reduced to finite differences of the reflectivity "image" $\mathbf{r}(t)$) in time domain.

For the time being we apply EnIC by very simplified manner, using the lemma on `Gateaux derivative of entropy`, that we revealed in (Ryzhikov et al., 1995): the value of entropy $E(p_\kappa)$ is decreasing along the line $p(\mathbf{x}) + \kappa\,(p(\mathbf{x}) - p^*(\mathbf{x}))$, where p^* is a homogeneous distribution.

[6] the term *entropy* is a bit confusing: EnIC is not directly relevant to the statistical physics/information theory, we exploit *mathematical* features of the Boltzman/Shannon entropy

The EnIC allows us to *detect* principal parameters of the post-/pre-SAR images $\hat{\mathbf{r}}$. The final SDec-image is obtained then by solving of normal equations, reduced with respect to the extracted parameters.

Despite the problem of revealing the *distribution* from its band-limited mapping data, i.e. the image that contains even null-space components looks more spurious than ill-posed one, it can apparently be well posed.

Conclusions

Sharp deconvolution combines automatically predictive and spiking deconvolutions. When the image under reconstruction can be modelled by a few-parameter distribution, SDec has two advantages over conventional methods: in contrary to the trade-off enforced by the latters it yields more **robust & high-resolution** algorithms. The algorithms are very fast and thus feasible for industrial implementation.

Note also, that SAR can be applied for problems accosiated with computational instability, while SDec itself - for problems of imaging regardless of a space which contains an image.

Acknowledgements

The research was stimulated by a short-term project on elaboration of *Remul* (Lokshtanov, 1995), funded by Norsk Hydro (*NH*). Authors are grateful to D. Lokshtanov, and A. Sagehaug (*NH*) for discussions.

Our special thanks to Prof. Jakob J. Stamnes, (University of Bergen, Department of Physics): without his active participation the research would not have been carried out.

References

Franklin, J., 1970, Well-posed stochastic extension of ill-posed linear problems: J. Math. Anal. Appl, **31**, 682–716.

Robinson, E., and Treitel, S., 1980, Geophysical signal analysis: Prentice-Hall, Englewood Cliffs.

Ryzhikov, G., Biryulina, M., and Hanyga, A., 1995, 3D nonlinear inversion by entropy of image contrast optimization: Nonlinear Processes in Geophysics, **2**, 228–240.

Tikhonov, A., and Arsenin, V., 1977, Solution of ill-posed problems: Winston, Washington, DC.

Webster, G., and Levin, F., Eds., 1978, Deconvolution, Geophysics Reprint Series No. 1 Society of Exploration Geophysicists, Tulsa.