

# Quaternion Multichannel SSA for Multicomponent Seismic Data

Breno Bahia and Mauricio D. Sacchi

Signal Analysis and Imaging Group (SAIG), Department of Physics, University of Alberta

## Summary

The quaternion signal model can naturally represent multi-component seismic data. In this paper, this model is employed to generalize the Singular Spectrum Analysis (SSA) filter to the case of vector measurements, named quaternion-SSA (QSSA). Adoption of the quaternion signal model requires the quaternionic version of signal processing techniques such as Fourier transforms and singular value decomposition. The rank-reduction operation in SSA is justified by means of quaternion vector autoregressive (VAR) models. Synthetic examples showing denoising of 3C seismic data are used to illustrate the method.

## Introduction

Multicomponent (MC) sensors have allowed the acquisition of vector fields. One approach to process vector data is to adopt the quaternion signal model (Le Bihan and Mars, 2004). The reasoning is that the quaternion domain offers a natural way of representing three- and four-dimensional signals, such as three- and four-component (3C and 4C) seismic data, while accounting and preserving mutual relations between its components (Le Bihan and Mars, 2004; Took and Mandic, 2011; Stanton and Sacchi, 2013).

The adoption of quaternions has been increasing among the signal processing community, which re-tools the field with new techniques based on quaternion algebra, such as Quaternion Fourier Transform (QFT) (Ell, 1992) and Quaternion Singular Value Decomposition (QSVD) (Le Bihan and Mars, 2004; Sangwine and Le Bihan, 2006). Such distinction happens because, in general, quaternion multiplication is not commutative, i.e.,  $pq \neq qp$  if  $p$  and  $q$  are quaternions. Owing to the non-commutativity of quaternions, for instance, is a variety of QFT definitions, as the positioning of the exponential kernel in the transform definition is now meaningful (see Ell et al. (2014) for more details). Quaternion-based techniques can be conveniently used thanks to the quaternion toolbox for MATLAB developed by Sangwine (2005).

SSA has already been applied to denoising and reconstruction of MC seismic data by Sacchi et al. (2017), where the authors employed long-vector notation to represent their data samples in the frequency-space ( $f-x$ ) domain. Its extension to the quaternion domain is given by Enshaeifar et al. (2016) in the analysis of multichannel electroencephalograms. The goal of this paper is to validate the QSSA as a tool for denoising multicomponent seismic data by exploiting the potential of the quaternion domain to represent vector fields.

## Theory

A quaternion ( $q$ ) can be defined as a four-dimensional hypercomplex number with cartesian form given by

$$q = a + bi + cj + dk, \quad (1)$$

where the tuple  $\{a, b, c, d\}$  are its components, and  $i^2 = j^2 = k^2 = ijk = -1$ . A quaternion is said to be pure if  $a = 0$ , and it reduces to a real number if its vectorial part is null, i.e.  $b = c = d = 0$ . Examples

of quaternion properties are its conjugate

$$q^* = a - bi - cj - dk, \quad (2)$$

and its involutions (self-inverse mappings)

$$q^i = -iqi = a + bi - cj - dk, \quad (3)$$

$$q^j = -jqj = a - bi + cj - dk, \quad (4)$$

$$q^k = -kqk = a - bi - cj + dk. \quad (5)$$

For the case of 2D-3C seismic data, where vector sensors records three orthogonal components, here denoted by  $(\mathbf{U}_x, \mathbf{U}_y, \mathbf{U}_z)$ , at  $nt$  discrete time samples and  $nx$  discrete space positions, it is possible to encode the three corresponding signals into the components of a matrix of pure quaternions  $\mathbf{Q}$  such as

$$\mathbf{Q}[n, x] = \mathbf{U}_x[n, x]i + \mathbf{U}_y[n, x]j + \mathbf{U}_z[n, x]k, \quad (6)$$

with  $n = [1, nt]$  and  $x = [1, nx]$ . As it is possible to notice,  $\mathbf{Q}$  represents the multicomponent seismic data as a whole. Of course, extension to 4C and data with higher dimensions is possible by means of arrays of full quaternions.

Application of QSSA to multicomponent seismic data follows the same underlying ideas as in Oropeza and Sacchi (2011) and Sacchi et al. (2017), where the process is applied to the signal at monochromatic frequency slices of the Fourier transformed  $\mathbf{Q}$ , represented by the vector  $\mathbf{q}_x \equiv \hat{\mathbf{Q}}[\omega, x]$ . Moreover, QSSA basically follows the same steps of decomposition and reconstruction as in the scalar case. Hence, it is possible to represent it with the following expression

$$\hat{\mathbf{q}}_x = \mathcal{A} \mathcal{R} \mathcal{H}[\mathbf{q}_x], \quad (7)$$

$\hat{\mathbf{q}}_x$  being the filtered version of  $\mathbf{q}_x$ , and  $\mathcal{H}$ ,  $\mathcal{R}$  and  $\mathcal{A}$  represent the operations of Hankelization, rank-reduction and anti-diagonal averaging, respectively (Oropeza and Sacchi, 2011).

The Hankelization operation consists in constructing a Hankel matrix  $\mathbf{T}$  of size  $L \times K$ , which columns are one-sample lagged versions of a length- $L$  window of the signal  $\mathbf{q}_x$ . In the SSA literature, this matrix is named trajectory matrix, and it is given by

$$\mathbf{T} = \begin{pmatrix} q_1 & q_2 & \dots & q_K \\ q_2 & q_3 & \dots & q_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ q_L & q_{L+1} & \dots & q_{nx} \end{pmatrix}, \quad (8)$$

with  $K = nx - L + 1$ . Enshaeifar et al. (2016) argues that second-order statistical information from the signal can be used to enhance QSSA by incorporating the three involutions (equations 3 to 5) in an augmented trajectory matrix  $\mathbf{T}^a$  of size  $4L \times K$  as

$$\mathbf{T}^a = [\mathbf{T}^T \quad \mathbf{T}^{iT} \quad \mathbf{T}^{jT} \quad \mathbf{T}^{kT}]^T. \quad (9)$$

As already discussed in the SSA theory applied to seismic data denoising and reconstruction, be it in the scalar or vector case, one might adopt autoregressive models to show that the trajectory matrix is of low-rank in the ideal case of complete noise-free seismic data. Even further, one can show that, for MC data consisting of  $p$  superposed linear (or quasi-linear) events in the time domain, the rank of the trajectory matrix is  $\text{rank}\{T\} = p$ . It is also known that the presence of noise in the data increases the rank of the trajectory matrix, which allows the formulation of denoising as a rank-reduction problem. It is therefore natural to turn to quaternion vector autoregressive (Q-VAR) models in order to justify the

generalization of the SSA filter to the quaternion case. The development of Q-VAR models is given by Ginzberg and Walden (2013) where a VAR process of order  $p$  can be written as

$$q_n = A_1 q_{n-1} + A_2 q_{n-2} + \dots + A_p q_{n-p}, \quad (10)$$

where  $A_i$  ( $i = [1, p]$ ) are the quaternion-AR coefficients, and it is assumed that the process has no innovation and it is zero-mean.

The rank-reduction operation in the quaternion domain is then justified by the Q-VAR model above. It is important to clarify that the rank-reduction operation is implemented via the quaternionic version of the singular value decomposition (QSVD). It is worth mentioning, for completeness, that the SVD of quaternion matrices is achieved via the classic SVD of complex or real bidiagonal matrices obtained by application of Householder transformations on the quaternion matrix (see Sangwine and Le Bihan (2006) for further details). The final step in the SSA filter, after rank-reduction, is to average the anti-diagonal entries of the resulting matrix to recover the filtered signal.

### Example

To illustrate the application of QSSA in multicomponent seismic data denoising, a 2D-3C synthetic data set is embedded in a quaternion array as in equation 6. Only a window of the data is used to approximate the condition of linear events in  $t - x$ . This window has 2 to 3 quasi-linear events, accounting for approximately 0.22 seconds of data in 51 traces, showing different polarities and also offset-dependent variations in amplitude. The number of time samples is  $nt = 111$ . The data is corrupted by noise such that the SNR is 2 for each component (Figure 1). The data is left-sided Fourier transformed using an eigenaxis  $\mu = \frac{1}{\sqrt{3}}(1, 1, 1)$ . The algorithm is applied in each frequency slice from 0 to 60 Hz.

One advantage of quaternion signal processing is the simultaneous processing of the data components. That can be illustrated by the spectrum of singular values obtained via QSVD of the trajectory matrix for both standard and augmented cases at a frequency of 20 Hz (Figure 2). It is possible to observe that two to three singular values differ from zero in (a). The augmented version has increased rank due to the addition of the involutions to the process. Since the events in the window are not perfectly linear, as it usually happens in real applications, the rank of the trajectory matrix is not equal to the number of events in the data. In this light, the rank is set to  $p = 4$  for all cases to account for small curvatures in the window. Figures 3 and 4 show the denoising results for standard and augmented QSSA, respectively.

The quality of the denoising process for each component can be quantified by the quality factor (in dB)

$$R_i = 10 \log_{10} \frac{\|\mathbf{U}_i^o\|_2^2}{\|\mathbf{U}_i - \mathbf{U}_i^o\|_2^2}, \quad (11)$$

where  $\mathbf{U}_i^o$  is the noise-free data, and  $\mathbf{U}_i$  is the denoised data, with  $i = \{x, y, z\}$  to represent each component. The following table compares the quality of each reconstruction.

Table 1: Denoising quality factors.

	QSSA	AQSSA
$R_x$ (dB)	9.1901	9.3532
$R_y$ (dB)	7.1951	7.5778
$R_z$ (dB)	8.1111	8.7950

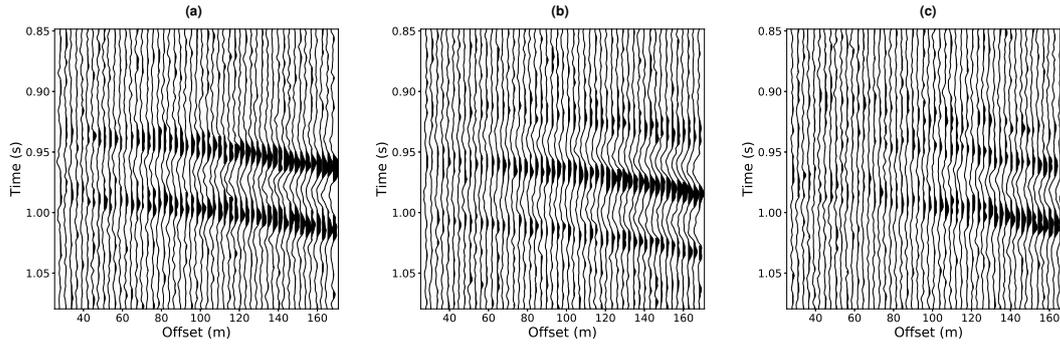


Figure 1: Window of 2D 3C synthetic seismic data contaminated with noise.

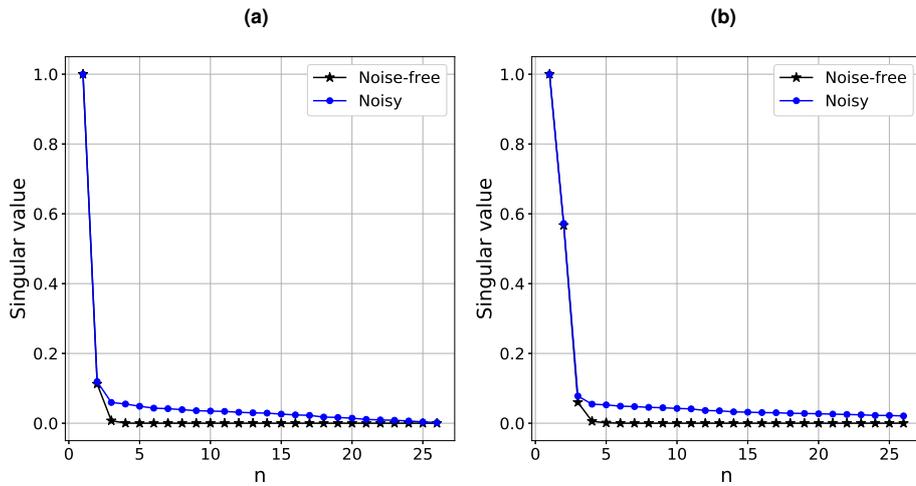


Figure 2: Singular values of trajectory matrix for frequency slice at  $20\text{ Hz}$  of the noise-free and noisy windowed data.

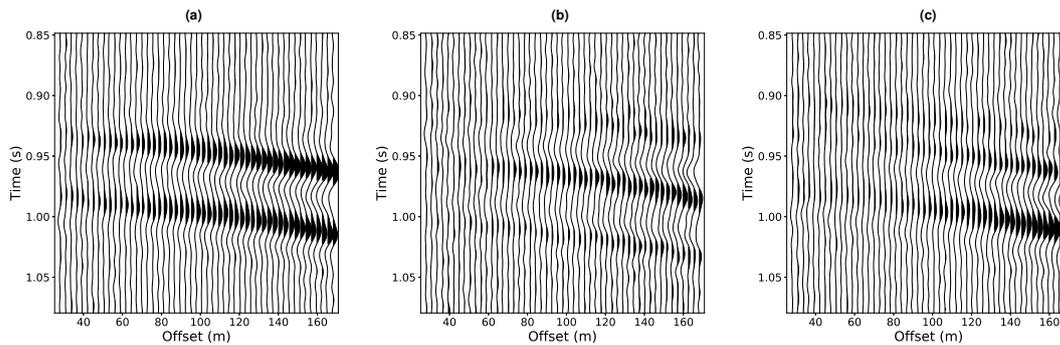


Figure 3: Result for QSSA denoising.

## Conclusion

This paper presented a generalization of the SSA filter to the vector field case via quaternion algebra. The multicomponent data is represented by an array of pure quaternions. Quaternion signal processing techniques, such as the QFT and QSVD, have to be employed to process such data. QSSA follows the same basic ideas of SSA in the scalar case, differing in the possibility of augmentation of the trajectory matrix to account for second-order statistical properties of the signal. As in the scalar case, noise increases the rank of the trajectory matrix, and denoising is posed as a rank-

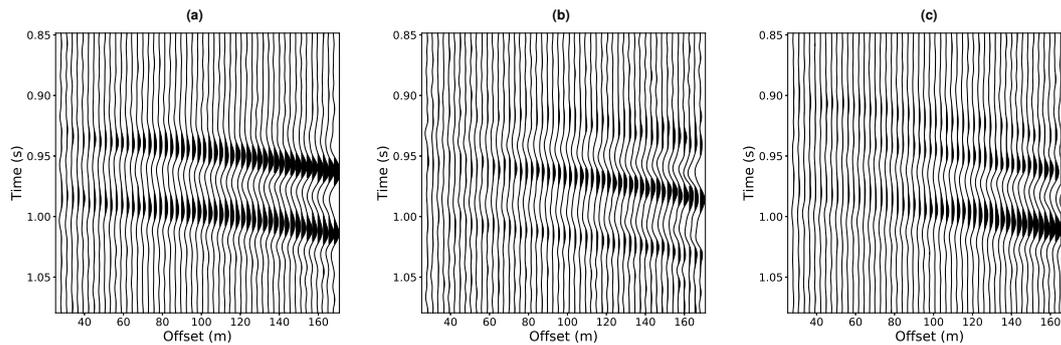


Figure 4: Result for augmented QSSA denoising.

reduction problem. It was possible to use the quaternion signal model to simultaneously attenuate random noise from all the components in a holistic fashion. The augmented version of the QSSA shows slightly better performance than its standard version, but such improvement is not significant.

## Acknowledgements

The authors are grateful to the sponsors of Signal Analysis and Imaging Group (SAIG) at the University of Alberta for their continued support. The developers of the QTFM are also acknowledged. Breno Bahia would like to thank the Canadian Association of Petroleum Producers for the financial support.

## References

- Ell, T. A., 1992, Hypercomplex spectral transformations: PhD Thesis.
- Ell, T. A., N. Le Bihan, and S. J. Sangwine, 2014, Quaternion Fourier transforms for signal and image processing: John Wiley & Sons.
- Enshaeifar, S., S. Kouchaki, C. C. Took, and S. Sanei, 2016, Quaternion singular spectrum analysis of electroencephalogram with application in sleep analysis: *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **24**, 57–67.
- Ginzberg, P., and A. T. Walden, 2013, Quaternion VAR modelling and estimation: *IEEE Transactions on Signal Processing*, **61**, 154–158.
- Le Bihan, N., and J. Mars, 2004, Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing: *Signal processing*, **84**, 1177–1199.
- Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: *Geophysics*, **76**, V25–V32.
- Sacchi, M. D., J. Cheng, and S. Janzen, 2017, Theory and application of vector singular spectrum analysis (SSA) for multicomponent seismic data reconstruction: *GeoConvention2017*.
- Sangwine, S., 2005, Quaternion toolbox for Matlab: <http://qtfm.sourceforge.net/>.
- Sangwine, S. J., and N. Le Bihan, 2006, Quaternion singular value decomposition based on bidiagonalization to a real or complex matrix using quaternion Householder transformations: *Applied mathematics and computation*, **182**, 727–738.
- Stanton, A., and M. D. Sacchi, 2013, Vector reconstruction of multicomponent seismic data: *Geophysics*, **78**, V131–V145.
- Took, C. C., and D. P. Mandic, 2011, Augmented second-order statistics of quaternion random signals: *Signal Processing*, **91**, 214–224.
- Trickett, S., 2008, F-xy Cadzow noise suppression, *in* SEG Technical Program Expanded Abstracts 2008: Society of Exploration Geophysicists, 2586–2590.