DIAGONALLY LOADED NORMALISED SAMPLE MATRIX INVERSION (LNSMI) FOR OUTLIER-RESISTANT ADAPTIVE FILTERING

Y. I. Abramovich

ISRD, Defence Science & Technology Organisation (DSTO), PO Box 1500, Edinburgh SA 5111, Australia
Yuri.Abramovich@dsto.defence.gov.au

N. K. Spencer *

Adelaide Research & Innovation Pty Ltd (ARI), 200 Labs, ISRD, DSTO, PO Box 1500, Edinburgh SA 5111, Australia
Nick.Spencer@adelaide.edu.au

ABSTRACT

Instead of a "hard" decision on ignoring "outlier" training samples in constructing the covariance matrix estimate, we propose a "softer" method that reduces the impact of such abnormal data samples on adaptive filter performance. Specifically, we introduce a diagonally loaded covariance matrix estimate that is normalised by a generalised inner product (GIP), which is more robust against outliers. We demonstrate the efficiency of this technique on high-frequency (HF) over-the-horizon radar (OTHR) data.

Index Terms— Adaptive signal processing, array signal processing, HF radar, covariance matrices, robustness

1. INTRODUCTION

In adaptive radar applications, an estimate of the unknown interference covariance matrix is usually computed from a "training set" of observed (sample) data that contains only interference. This is called "supervised training", and it has long been known that the inadvertent presence of desired-signal (*ie*. target) returns in the training data can severely degrade adaptive filter performance [1, 2]. For this reason, the training data (also known as "secondary" sample data) is usually sourced from resolution cells close to, but not including, the "primary" resolution cell (that is being tested for the presence or absence of a target).

Firstly, this approach means that there are as many adaptive filters as there are resolution cells (in range and Doppler frequency, for example), which is far greater than the number of possible targets. Secondly, this technique by itself cannot guarantee the absence of a (strong) target in the training data, which severely curtails its ability to detect a (weak) target in the primary data. In fact, we need to assume that the vast majority of the available data contains only interference, with targets or target-like "outliers" appearing in a small minority. Moreover, the standard homogeneity condition (that requires all training data to be described by the same covariance matrix) significantly limits the total number of samples available. Also, the same covariance matrix does not necessarily mean that the samples ("snapshots") are indeed identically distributed, since they can have quite different powers and/or be well-represented by the class of spherically invariant random processes (SIRPs) [3].

The problem of outlier-resistant adaptive matched filtering has been addressed in [4], under the assumption that the interference-only training data is independent identically distributed (i.i.d.). That proposed solution was called a "re-iterated fast maximum-likelihood method", and eliminates from the training set any data that is deemed to be an outlier. The suggested method to identify outliers is a GIP test that any given snapshot $\boldsymbol{x}$ is described by a covariance matrix $\hat{R}$ [5]:

$$\boldsymbol{x}^H \hat{R}^{-1} \boldsymbol{x} \lessgtr c \tag{1}$$

where the n greatest GIP values across the training data are declared to be outliers [4].

The results presented in [4] demonstrate that this approach is quite efficient for simulated i.i.d. training data, but the practical application to HF OTHR, at least, is not completely satisfactory. One reason is that HF OTHR training data usually contains samples of varying power. For example, the undesirable "spread clutter" [6] that propagates via highly perturbed ionospheric layers mostly replicates the sea-clutter Doppler spectrum, and so has different power across the (training) Doppler cells. Similarly, unpremeditated co-channel interference may have a (Doppler) spectrum that is far from uniform. In such cases, the GIP test cannot be used by itself for outlier identification.

In this paper, we introduce a method of outlier-resistant adaptive beamforming for highly inhomogeneous power over the training data, and demonstrate its efficiency using real HF OTHR data.

2. PROBLEM FORMULATION

In radar applications with highly inhomogeneous clutter power over the resolution cells, the following model is typically used. For the observed N i.i.d. Gaussian M-variate snapshots that are specified by the covariance matrix R, *ie.* $\boldsymbol{x}_j \sim \mathcal{CN}_M(0,\, R)$, we consider the transformed snapshots

$$\boldsymbol{y}_j = c_j\, \boldsymbol{x}_j \qquad \text{for} \quad j = 1, \ldots, N \tag{2}$$

where the c_j are non-negative scalars. If these power-scaling factors $\boldsymbol{c} \equiv [c_1, \ldots, c_N]^T$ are random values and the probability density function (p.d.f.) $f_{\boldsymbol{c}}(\boldsymbol{c})$ is somehow known, then we are dealing with a SIRP [3].

In some cases where the characteristic p.d.f. $f_{\boldsymbol{c}}(\boldsymbol{c})$ is unknown, the scaling factors $\boldsymbol{c}$ are treated as unknown deterministic parameters rather than random variables. This leads to the "deterministic maximum likelihood" (DML) [7], as distinct from the "stochastic maximum likelihood" (SML) with its fixed number of unknown parameters that does not grow with the sample size N. While a consistent estimate of the scaling factors $\boldsymbol{c}$ does not yet exist, DML is widely used. Its likelihood function (LF) is defined as [8]

$$f[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N \,|\, R,\, \boldsymbol{c}] = \prod_{j=1}^{N} \frac{\exp[-\boldsymbol{y}_j^H R^{-1} \boldsymbol{y}_j / c_j^2]}{\pi^M c_j^{2M} \det R} \tag{3}$$

* Supported by the Centre of Expertise in Phased Array and Microwave Radar Systems (CEPAMiR) established in partnership between the Defence Science and Technology Organisation (DSTO) and the University of Adelaide, Australia. Adelaide Research & Innovation Pty Ltd (ARI) is the commercial development company of the University of Adelaide.

1-4244-0728-1/07/$20.00

which has a maximum with respect to $\boldsymbol{c}$ of [9]

$$\max_{\boldsymbol{c}} f[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N \,|\, R, \boldsymbol{c}] = \prod_{j=1}^{N} \frac{M^M \exp[-M]}{(\pi \boldsymbol{y}_j^H R^{-1} \boldsymbol{y}_j)^M \det R} \tag{4}$$

since

$$[\hat{c}_j^2]_{ML} = \frac{1}{M} \boldsymbol{y}_j^H R^{-1} \boldsymbol{y}_j \,, \tag{5}$$

and when the covariance matrix R is an arbitrary non-negative-definite Hermitian matrix, a direct solution of the equation

$$\frac{\partial}{\partial R} \log f[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N \,|\, R, \boldsymbol{c}] = 0 \tag{6}$$

leads to [9]

$$\hat{R}_{ML} = \frac{M}{N} \sum_{j=1}^{N} \frac{\boldsymbol{y}_j \boldsymbol{y}_j^H}{\boldsymbol{y}_j^H \hat{R}_{ML}^{-1} \boldsymbol{y}_j} \,. \tag{7}$$

Convergence of the iterative estimation

$$\hat{R}_{ML} = \lim_{k \to \infty} \hat{R}^{(k)} \,, \qquad \hat{R}^{(k)} \equiv \frac{M}{N} \sum_{j=1}^{N} \frac{\boldsymbol{y}_j \boldsymbol{y}_j^H}{\boldsymbol{y}_j^H \hat{R}^{(k)-1} \boldsymbol{y}_j} \tag{8}$$

has been proven in [9] for $N \geq M$. Another well-known covariance matrix estimate (CME) that is also invariant with respect to the scaling factors $\boldsymbol{c}$ is [10]

$$\hat{\mathcal{R}}_{ML} = \sum_{j=1}^{N} \frac{\boldsymbol{y}_j \boldsymbol{y}_j^H}{\boldsymbol{y}_j^H \boldsymbol{y}_j} \,, \tag{9}$$

but the benefit of the estimate $\hat{R}_{ML}$ (7) is that the adaptive matched filter (AMF) detector that is constructed from this CME has the important constant false-alarm rate (CFAR) property (with respect to the covariance matrix R) [11].

However, the CME $\hat{R}_{ML}$ has an additional important property, since the denominator in (7) is actually analogous to the GIP (1) that was used in [4, 5] as a nonhomogeneity detector.

For convenience, let us introduce the notation

$$Y_j(R) \equiv \frac{\boldsymbol{y}_j \boldsymbol{y}_j^H}{\boldsymbol{y}_j^H R^{-1} \boldsymbol{y}_j} \tag{10}$$

so that our CME (7) is just expressed as

$$\hat{R}_{ML} = \frac{M}{N} \sum_{j=1}^{N} Y_j(\hat{R}_{ML}) \,. \tag{11}$$

For interference-only samples $\boldsymbol{y}_j$, the normalisation in (7) just makes all contributions $Y_j(\hat{R}_{ML})$ statistically identically distributed, regardless of the power-scaling factors $\boldsymbol{c}$. On the contrary, for an outlier sample that contains a target, $\boldsymbol{y}_t$ say, the GIP in the denominator of $Y_t(\hat{R}_{ML})$ is no longer proportional to the scaling $\boldsymbol{c}$. Indeed, the GIP value of a sufficiently high (clairvoyant) signal-to-interference output ratio for the "corrupted" sample $\boldsymbol{y}_t$ will (with high probability) exceed the GIP value of the interference-only sample $\boldsymbol{y}_j$ [4, 5], *ie*.

$$\boldsymbol{y}_t^H \hat{R}_{ML}^{-1} \boldsymbol{y}_t \overset{p}{>} \boldsymbol{y}_j^H \hat{R}_{ML}^{-1} \boldsymbol{y}_j \tag{12}$$

and hence

$$Y_t(\hat{R}_{ML}) \overset{p}{<} Y_j(\hat{R}_{ML}) \,. \tag{13}$$

Based on this property, we propose computing the (diagonally) loaded normalised CME for adaptive filter design using the iterative scheme

$$\hat{R}_{LNSMI} = \lim_{k \to \infty} \hat{R}^{(k)} \tag{14}$$

with

$$\hat{R}^{(k+1)} \equiv \frac{1}{C(k)} \left[\beta(k) I_M + \sum_{j=1}^{N} Y_j(\hat{R}^{(k)}) \right] , \quad \hat{R}^{(0)} = \frac{I_M}{M} \tag{15}$$

where the normalisation factor $C(k)$ is chosen such that

$$\operatorname{tr} \hat{R}^{(k+1)} = 1 \,. \tag{16}$$

Here I_M is the M-variate identity matrix, and the loading factor $\beta(k)$ is chosen so that the training sample $\boldsymbol{y}_t$ that is "corrupted" by a strong target is "suppressed" to a level below the diagonal loading:

$$\beta(k) I_M \overset{p}{>} Y_t(\hat{R}^{(k)}) \,. \tag{17}$$

At the same time, weak targets that do not significantly contribute to the CME are only mildly suppressed. The specific loading factor $\beta(k)$ that satisfies this requirement must be chosen based on the (expected) interference rejection factor (interference-to-noise ratio). As in most diagonally loaded routines, the performance of this adaptive algorithm is expected to be quite robust with respect to the loading factor for sufficiently strong interference.

Due to a complicated (nonlinear) interaction between interference-only and interference-plus-target data samples in (14), an accurate theoretical analysis of the convergence and efficiency of the LNSMI estimate is not yet available. For this reason, we illustrate its efficiency by practical results.

3. PERFORMANCE ANALYSIS

We demonstrate the performance of the LNSMI estimate by considering an example of external-noise mitigation by adaptive beamforming in HF OTHR. This real data was collected at the output of the multichannel digital receiver of an OTHR facility and then range processed. As for most HF OTHRs that use a periodic continuous-wave linear frequency-modulated (CW LFM) waveform, this particular radar makes available a rather limited number of (operationally important) range cells. Each range cell may contain target(s) or target-like signal(s), for example generated by cooperative transponders, as well as backscattered sea-clutter and co-channel interference.

Fig. 1 shows the usual "range-Doppler maps" that are the result of OTHR data processing; these are plots of the amplitude of the output signal from some beamforming (azimuthal) direction for every operational range and Doppler resolution cell. In addition, Fig. 1 shows "range-cut" diagrams (Doppler spectra) that are just one-dimensional versions of the range-Doppler maps. This data was chosen to illustrate strong co-channel interference, which can be seen as overwhelming "vertical stripes" across all Doppler cells in the upper-left subfigure that shows the result of conventional beamforming (CBF). A weak artificial target (circled) has been injected into the real data in order to study the effect of various processing in a controlled fashion.

The subfigure below that one is the corresponding range-Doppler map for the output of the conventional adaptive beamformer, implemented by the (diagonally) loaded sample matrix inversion (LSMI) algorithm [12]:

$$\hat{R}_{LSMI} = \beta I_M + \sum_{j \in \Omega} \boldsymbol{x}_j \boldsymbol{x}_j^H \tag{18}$$

with the filter

$$\hat{\boldsymbol{w}}_{LSMI} = \frac{\hat{R}_{LSMI}^{-1}\, \boldsymbol{s}(\theta)}{\boldsymbol{s}^H(\theta)\, \hat{R}_{LSMI}^{-1}\, \boldsymbol{s}(\theta)} \tag{19}$$

where $\boldsymbol{s}(\theta)$ is the standard antenna array steering vector that forms a beam in the azimuthal direction θ, and Ω is the training region in range-Doppler space. Here we chose the training region to be Doppler cells 1–40 and range cells in blocks of ten:

$$t = \{1 + 10(n-1), \ldots, 10n\} \qquad \text{for} \quad n = 1, \ldots, 11 \tag{20}$$

so that, respectively, the training region is free of strong sea clutter, and has sufficient snapshots ($N = 400$) to process the $M = 372$ channels of data. We used a loading factor of $\beta = 10^{-3}$ with respect to the maximum eigenvalue of the sample matrix $\sum_j \boldsymbol{x}_j \boldsymbol{x}_j^H$. The same method was applied to the LMSMI algorithm.

The next lower subfigure shows the corresponding range-Doppler map that is the output of the loaded normalised sample matrix inversion (LNSMI) algorithm (14), trained over the same region, with the analogous filter to (19), $\hat{\boldsymbol{w}}_{LNSMI}$.

The lower-left subfigure compares the above three differently processed range cuts for the weak artificial target (range cell 32), with the target arrowed (Doppler cell 30).

We see that this weak target is submerged in the external noise, and is not detectable by CBF. On the contrary, both adaptive beamforming techniques have equally efficient external-noise mitigation and have enhanced the "subclutter visibility" (ratio of peak to noise floor) of the weak target by more than 10 dB, despite the fact that the target-containing cell is included in the training set for covariance matrix estimation. This example confirms the well-known fact that relatively weak targets are not significantly affected by adaptive processing when they are included in the training region.

Overall, the range-Doppler maps clearly show that both adaptive techniques have largely removed the co-channel interference.

In the same format, the right-hand subfigures show the processing results for the same data, but with the addition of a strong ("outlier") target injected at range cell 37 and Doppler cell 20, which is therefore located in the same training region as the weak target. In this new scenario, the LSMI ("outlier-sensitive") and LNSMI ("outlier-resistant") adaptive beamformers produce completely different results. As expected [1, 2], we see from the right-hand range cuts that the influence of the strong target has drastically degraded the efficiency of the LSMI processing, to the point where the weak target is not longer detectable, since the external noise is practically unsuppressed. Of course, the presence of this outlier in the training region can be easily detected by CBF or LSMI processing and thresholding, and then excluded from re-training, as discussed in [4] for example. However, the benefit of our LNSMI method is that it does not require a "hard" decision on outlier detection and censoring. A comparison of the range-cut subfigures clearly demonstrates that weak-target detectability for LNSMI here is not affected by the presence of the strong outlier.

Finally, it is important to note that only $k = 2$ iterations of the LNSMI formula (14) were required to obtain these results, while the element-wise convergence

$$\max_{ij} |\hat{R}_{ij}^{(k)} - \hat{R}_{ij}^{(k-1)}| < 10^{-2} \times \lambda_{\min}[\hat{R}^{(k)}] \tag{21}$$

occured at $k = 14$ iterations for the range cells of interest. A similar analysis conducted for the loading factors $\beta = 10^{-2}$ and 10^{-4} results in practically the same performance as in Fig. 1, which demonstrates the expected robustness of the LNSMI algorithm for this practical problem.

4. SUMMARY AND CONCLUSIONS

We have proposed a (diagonally) loaded normalised sample matrix inversion (LNSMI) method of implementing an outlier-resistant adaptive filter, suitable for weak target detection in both i.i.d. and SIRP interference. Using real HF OTHR data, we have demonstrated that this method can be very efficient at mitigating interference, and is virtually unaffected by the presence of strong outliers in the data sample training set. We have shown that only two iterations of the algorithm can be sufficient, and this imposes a small computational burden in practical implementation. Indeed, the LNSMI technique significantly simplifies adaptive processing, since the "primary" (tested) cell and the "guard" cells no longer need to be excluded from the training set.

5. REFERENCES

[1] Y. Abramovich, "Maximum convergence rates of adaptive tuning processes in interference compensation systems using an unclassified training sample," *Radio Engineering and Electronic Physics*, vol. 27 (8), pp. 74–78, 1982, English translation of *Radioteknika i Elektronika*.

[2] K. Gerlach, "The effects of signal contamination on two adaptive detectors," *IEEE Trans. Aero. Elect. Sys.*, vol. 30, no. 1, pp. 297–309, Jan. 1995.

[3] M. Rangaswamy, D. Weiner, and A. Oztürk, "Non-Gaussian random vector identification using spherically invariant random processes," *IEEE Trans. Aero. Elect. Sys.*, vol. 29, no. 1, pp. 111–124, Jan. 1993.

[4] K. Gerlach, "Outlier resistant adaptive matched filtering," *IEEE Trans. Aero. Elect. Sys.*, 38(3), pp. 885–901, July 2002.

[5] P. Chen, W. Melvin, and M. Wicks, "Screening among multivariate normal data," *Journal of Multivariate Analysis*, vol. 69, pp. 10–29, 1999.

[6] S. Kraut, K. Harmanci, and J. Krolik, "Space-time adaptive processing for over-the-horizon spread-Doppler clutter mitigation," in *Proc. SAM-2000*, Cambridge, MA, USA, 2000, pp. 245–249.

[7] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust. Sp. Sig. Proc.*, vol. 38 (10), pp. 1783–1795, 1990.

[8] E. Conte, A. De Maio, and G. Ricci, "Covariance matrix estimation for adaptive CFAR detection in compound-Gaussian clutter," *IEEE Trans. Aero. Elect. Sys.*, vol. 38, no. 2, pp. 415–426, Apr. 2002.

[9] ——, "Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection," *IEEE Trans. Sig. Proc.*, vol. 50, no. 8, pp. 1908–1915, Aug. 2002.

[10] E. Conte, M. Lops, and G. Ricci, "Adaptive radar detection in compound-Gaussian clutter," in *Proc. EUSIPCO-94*, Edinburgh, 1994, pp. 526–529.

[11] F. Robey, D. Fuhrmann, E. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aero. Elect. Sys.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[12] Y. Abramovich, "A controlled method for adaptive optimization of filters using the criterion of maximum SNR," *Radio Eng. Electron. Phys.*, vol. 26 (3), pp. 87–95, 1981.

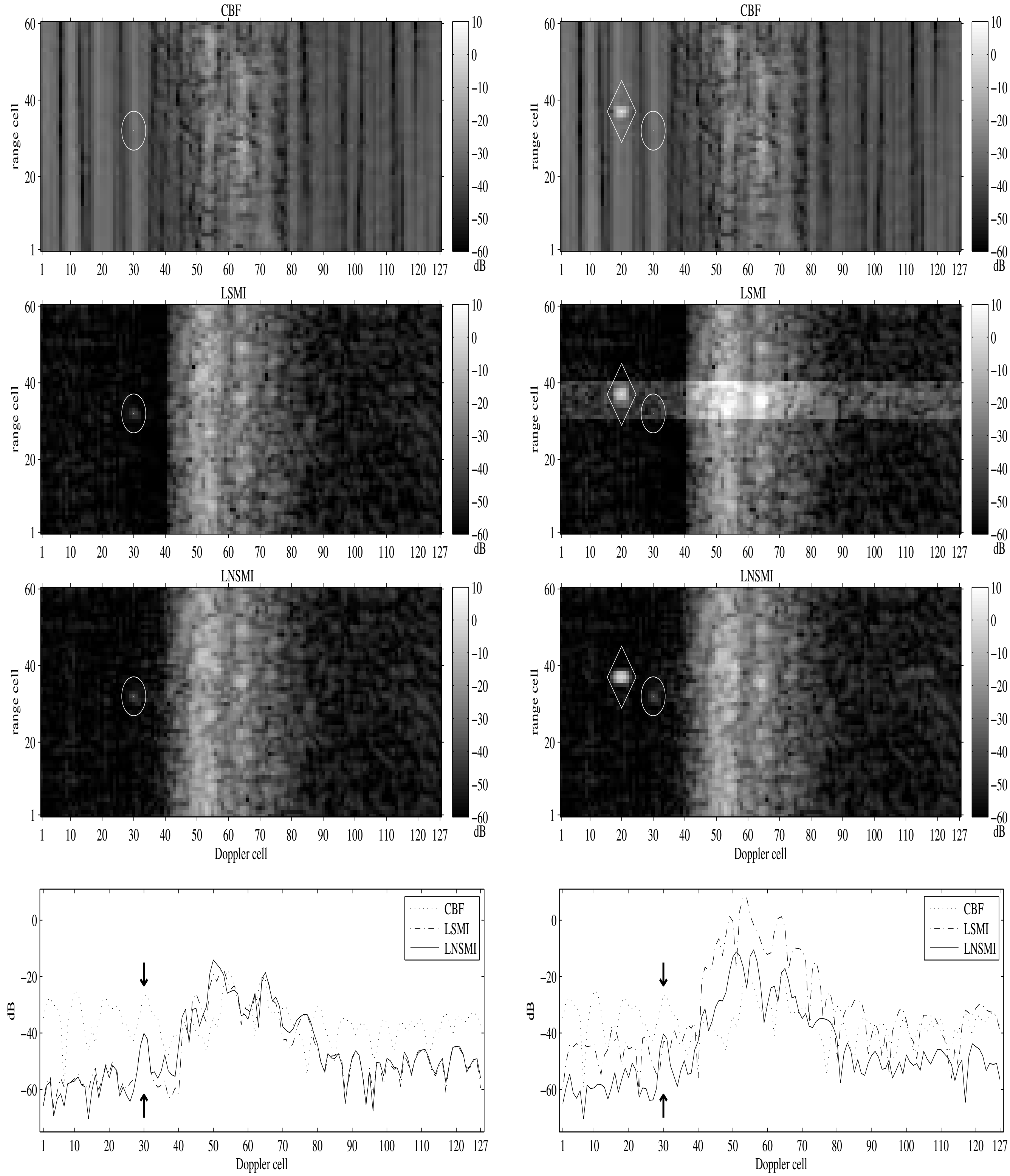

Fig. 1. Comparative results of various processing on real OTHR data; the right-hand side results are for the same data and processing as the left-hand side, but with a strong target injected nearby the weak target; the upper six subfigures are range-Doppler maps, the lower two subfigures are range cuts for the range bin of the weak target.