DOA Finding with Support Vector Regression Based Forward–Backward Linear Prediction

Jingjing Pan [1], Yide Wang [1,*], Cédric Le Bastard [1,2] and Tianzhen Wang [3]

[1] Institut d'Electronique et Télécommunications de Rennes, UMR CNRS 6164, Polytech Nantes, Rue Christian Pauc, BP 50609, 44306 Nantes CEDEX 3, France; jingjing.pan1@etu.univ-nantes.fr (J.P.); cedric.lebastard@cerema.fr (C.L.B.)

[2] Cerema, 49136 Les Ponts de Cé, France

[3] Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China; tzwang@shmtu.edu.cn

* Correspondence: yide.wang@univ-nantes.fr; Tel.: +33-2-4068-3230

Abstract: Direction-of-arrival (DOA) estimation has drawn considerable attention in array signal processing, particularly with coherent signals and a limited number of snapshots. Forward–backward linear prediction (FBLP) is able to directly deal with coherent signals. Support vector regression (SVR) is robust with small samples. This paper proposes the combination of the advantages of FBLP and SVR in the estimation of DOAs of coherent incoming signals with low snapshots. The performance of the proposed method is validated with numerical simulations in coherent scenarios, in terms of different angle separations, numbers of snapshots, and signal-to-noise ratios (SNRs). Simulation results show the effectiveness of the proposed method.

Keywords: direction-of-arrival (DOA); support vector regression (SVR); forward–backward linear prediction (FBLP); coherent signals; low snapshots

1. Introduction

Direction-of-arrival (DOA) estimation is of great importance in array signal processing [1,2]. A considerable number of techniques have been developed for the determination of the DOAs of incoming signals. Fourier-based methods belong to the traditional fast processing methods [1], but the resolution capability of the Fourier-based methods is limited by the physical size of the array (aperture). Subspace methods such as multiple signal classification (MUSIC) [3] and estimation of signal parameters via rational invariance technique (ESPRIT) [4] are popular because of their asymptotic infinite resolution and unbiased estimation performance. Apart from the successful applications in DOA finding, they are also very attractive in time-reversal imaging [5–7]. However, subspace methods require the eigendecomposition of the covariance matrix of the received signal to get the noise and signal subspaces, which adds to the computational burden. Besides, subspace methods cannot be applied directly in coherent scenarios [8]. Additional decorrelation techniques (e.g., spatial smoothing, SS) are usually necessary. SS was originally proposed for DOA estimation with coherent signals in [9], which was only a single-direction SS technique. Since then, various kinds of improvements of this technique have been developed. A forward–backward (FB) processing technique was introduced in [10] to improve the performance of the single-direction one. The work in [11] aimed at improving the computational efficiency of FB. In [12], the cross correlation in each sub-array was employed together with auto-correlation in SS processing. The authors in [13] focused on the correlated noise in coherent scenarios using FB. DOA finding of multi-group coherent signals was considered in [14]. Linear prediction (LP) is commonly used in time series problems, and can also be

used in array signal processing [15]. LP makes use of multiple overlapping observation sequences for the prediction of unknown ones. Similar to SS, it can directly deal with coherent signals [9,15]. It does not require eigendecomposition of the covariance matrix of the received signals. Besides, the family of LP methods has a higher resolution than the conventional Fourier-based approaches. However, LP fails to work when the number of snapshots is small [15,16].

The theory of LP is based on finding the weights which minimize the prediction error. Support vector regression (SVR) is a good sparse machine learning method capable of dealing with small samples [17]. Originally, the formulation of SVR was applied within the real domain, and later it was extended to complex data [18–20]. In the literature, SVR has many applications for signal processing problems, which can be classified into two approaches. One uses the training and testing pattern of SVR to approximate the mapping between the variables to be estimated and known features. The features are usually elements in the covariance matrix of the received signals. Details can be found in the estimation of DOA [21] and time delay [22]. The other approach combines the theory of SVR with classical signal processing methods. In [19], an SVR-based beamforming method was proposed to control the level of beam sidelobes. Similarly, the authors in [20] formulated SVR with Capon and then MUSIC in the estimation of the DOAs of incoming waves. However, [20] depended on an additional SS method or a recursive approach in coherent scenarios. The work in [20] also required the repetition of the SVR procedure along the DOA spectrum, which needs more computational time. Besides, there are combinations of SVR with linear signal processing methods such as auto-regressive model (AR) [18] and auto-regressive moving average (ARMA) [23] for frequency estimation and system identification problems. As described in [24], the principle of AR is closely related to the forward linear prediction (FLP). However, forward–backward linear prediction (FBLP) offers better performance than the one-directional prediction methods (FLP and backward linear prediction, BLP) [25]. Moreover, there is no explicit work about SVR-based LP models with coherent signals. Therefore, we propose to combine SVR with FBLP in the estimation of the DOAs of coherent incoming signals with a small number of snapshots.

The rest of this paper is organized as follows. Section 2 presents the signal model. Section 3 describes the framework of FBLP and the proposed SVR-based FBLP method. Section 4 gives some simulation results to show the effectiveness of the proposed method. Section 5 draws the conclusion.

Notations: $(.)^T$, $(.)^*$, and $(.)^H$ denote the transpose, conjugate, and transpose conjugate operations, respectively. $\mathbb{R}$ and $\mathbb{C}$ denote real and complex numbers, respectively. Vectors and matrices appear in boldface lowercase letters and boldface capital letters, respectively. $Re(\mathbf{z})$ and $Im(\mathbf{z})$ denote the real and imaginary parts of $\mathbf{z}$, respectively.

2. Signal Model

Consider a uniform linear array (ULA) with M isotropic antenna elements. There are K far-field narrow band incoming signals impinging on the antennas, which are corrupted by an additive Gaussian white noise (AGWN). If the number of snapshots is L, the output of antenna m can be expressed as [15]:

$$\mathbf{x}_m = [x_m(1),\ x_m(2),\ \ldots,\ x_m(L)]^T \tag{1}$$

with

$$x_m(t) = \sum_{k=1}^{K} s_k(t) e^{j(m-1)2\pi \frac{d}{\lambda} sin(\theta_k)} + n_m(t) \tag{2}$$

where $s_k(t)$ denotes the kth incoming signal received at the first antenna; θ_k is the corresponding DOA of the kth signal (with respect to the normal line of the array); $n_m(t)$ is the AGWN at the mth antenna; d denotes the distance between two adjacent sensors; and λ is the wavelength of the incoming signals.

3. Methodology

3.1. Forward–Backward Linear Prediction

LP methods estimate unknown values with a weighted sum of the known observations by minimizing the mean square prediction error. FBLP makes use of the observations from both forward and backward sequences in the estimation. When the order of the prediction filter is *P*, the prediction equation in matrix form can be expressed as follows [15,26]:

$$
\begin{bmatrix}
\mathbf{x}_P & \mathbf{x}_{P-1} & \cdots & \mathbf{x}_1 \\
\vdots & \vdots & & \vdots \\
\mathbf{x}_{M-1} & \mathbf{x}_{M-2} & \cdots & \mathbf{x}_{M-P} \\
\hdashline
\mathbf{x}_2^* & \mathbf{x}_3^* & \cdots & \mathbf{x}_{P+1}^* \\
\vdots & \vdots & & \vdots \\
\mathbf{x}_{M-P+1}^* & \mathbf{x}_{M-P+2}^* & \cdots & \mathbf{x}_M^*
\end{bmatrix}
\begin{bmatrix}
\omega_1 \\ \omega_2 \\ \vdots \\ \omega_P
\end{bmatrix}
=
\begin{bmatrix}
\mathbf{x}_{P+1} \\ \vdots \\ \mathbf{x}_M \\ \hdashline \mathbf{x}_1^* \\ \vdots \\ \mathbf{x}_{M-P}^*
\end{bmatrix}
\tag{3}
$$

We can observe that FLP and BLP can be expressed with the first and second halves (above and below the dotted line) of Equation (3), respectively. In a more compact way, Equation (3) can be rewritten as:

$$\mathbf{X}\omega = \mathbf{y} \tag{4}$$

where $\mathbf{y} \in \mathbb{C}^{N_T \times 1}$, $\mathbf{X} \in \mathbb{C}^{N_T \times P}$ and $\omega \in \mathbb{C}^{P \times 1}$, $N_T = 2(M - P)L$.

The weight coefficient vector ω can be estimated by the following equation:

$$\omega = \mathbf{R}^{-1}\mathbf{r} \tag{5}$$

where $\mathbf{R} = \mathbf{X}^H\mathbf{X}/L$ and $\mathbf{r} = \mathbf{X}^H\mathbf{y}/L$. The reverse of $\mathbf{R}$ requires the inequality: $N_T \geqslant P$ so that $\mathbf{R}$ is a non-singular matrix. Hence we have the constraint for the number of snapshots L as $L \geqslant P/[2(M - P)]$.

According to [26], the order of the prediction filter P should satisfy: $K \leqslant P \leqslant M - K/2$.

With the estimated weight vector ω, the power spectrum density (PSD) of FBLP can be expressed as:

$$P(\theta) = \frac{1}{\left| a^H(\theta) \begin{bmatrix} 1 \\ -\omega \end{bmatrix} \right|^2} \tag{6}$$

with $a(\theta) = [1, e^{j2\pi \frac{d}{\lambda} sin(\theta)}, \ldots, e^{j2\pi P \frac{d}{\lambda} sin(\theta)}]^T$, which is the steering vector.

3.2. Proposed Method: FBLP-SVR

The key issue in LP methods is to find the weight vector ω. However, in scenarios where the observation is insufficient (i.e., the number of snapshots is too small), LP methods might be unstable (even unsuccessful). Therefore, we propose to use the principle of SVR to deal with the situations of small samples. Inspired by [19], we propose a combined method of FBLP and SVR, which is called FBLP-SVR in the following.

According to [19], Equation (4) can be rewritten in terms of real and imaginary parts as:

$$\widetilde{\mathbf{y}} = \overline{\mathbf{X}}\widetilde{\omega} \tag{7}$$

where $\widetilde{\omega}^T = [Re(\omega^T)\ Im(\omega^T)]$, $\overline{\mathbf{X}} = \begin{bmatrix} Re(\mathbf{X}) & -Im(\mathbf{X}) \\ Im(\mathbf{X}) & Re(\mathbf{X}) \end{bmatrix}$ and $\widetilde{\mathbf{y}}^T = [Re(\mathbf{y}^T)\ Im(\mathbf{y}^T)]$.

It is worth noting that Equation (7) can be viewed as a typical form of SVR in real domain. The transformed vector $\widetilde{\boldsymbol{\omega}}$ belongs to $\mathbb{R}^{2P \times 1}$, while $\widetilde{\mathbf{y}} \in \mathbb{R}^{2N_T \times 1}$ and $\overline{\mathbf{X}} \in \mathbb{R}^{2N_T \times 2P}$. In regression problems, the objective of SVR is to build a hyper-plane to fit the data within a deviation less than a given value ϵ. We adopt the ϵ-intensive loss function here [17]. The optimization problem is to minimize the regression error both structurally and empirically:

$$\min_{\widetilde{\boldsymbol{\omega}}, \boldsymbol{\xi}, \widetilde{\boldsymbol{\xi}}} L_p(\widetilde{\boldsymbol{\omega}}, \boldsymbol{\xi}, \widetilde{\boldsymbol{\xi}}) = \min_{\widetilde{\boldsymbol{\omega}}, \boldsymbol{\xi}, \widetilde{\boldsymbol{\xi}}} [\frac{1}{2}\|\widetilde{\boldsymbol{\omega}}\|^2 + C \sum_{i=1}^{2N_T}(\xi_i + \widetilde{\xi}_i)] \tag{8}$$

$$s.t. \begin{cases} \widetilde{y}_i - \overline{\mathbf{x}}_i^T \widetilde{\boldsymbol{\omega}} \leqslant \epsilon + \xi_i, i = 1, \ldots, 2N_T \\ -\widetilde{y}_i + \overline{\mathbf{x}}_i^T \widetilde{\boldsymbol{\omega}} \leqslant \epsilon + \widetilde{\xi}_i, i = 1, \ldots, 2N_T \\ \xi_i, \widetilde{\xi}_i \geqslant 0, i = 1, \ldots, 2N_T \end{cases}$$

where $\overline{\mathbf{x}}_i$ and $\widetilde{y}_i$ are the ith column of $\overline{\mathbf{X}}^T$ and $\widetilde{\mathbf{y}}^T$, respectively; ξ_i and $\widetilde{\xi}_i$ are the corresponding slack variables to compensate for empirical errors; C controls the trade-off between the model complexity and empirical error.

The solution of Equation (8) is to construct a Lagrange function from the objective function (also called primal objective function) and the constraints by introducing a dual set of variables. The constructed Lagrange function has a saddle point that minimizes over the primal variables and maximizes over the dual variables [17]. Therefore, we deduce the following primal–dual objective function:

$$\begin{aligned} L_{pd} = &\frac{1}{2}\|\widetilde{\boldsymbol{\omega}}\|^2 + C \sum_{i=1}^{2N_T}(\xi_i + \widetilde{\xi}_i) - \sum_{i=1}^{2N_T}(\lambda_i \xi_i + \widetilde{\lambda}_i \widetilde{\xi}_i) \\ &+ \sum_{i=1}^{2N_T} \alpha_i(\widetilde{y}_i - \overline{\mathbf{x}}_i^T \widetilde{\boldsymbol{\omega}} - \epsilon - \xi_i) \\ &+ \sum_{i=1}^{2N_T} \widetilde{\alpha}_i(-\widetilde{y}_i + \overline{\mathbf{x}}_i^T \widetilde{\boldsymbol{\omega}} - \epsilon - \widetilde{\xi}_i) \end{aligned} \tag{9}$$

where α_i, $\widetilde{\alpha}_i$, λ_i, and $\widetilde{\lambda}_i$ are Lagrange multipliers.

All the partial derivatives of L_{pd} with respect to the primal variables ($\widetilde{\boldsymbol{\omega}}$, ξ_i, and $\widetilde{\xi}_i$) should be equal to zero at the saddle point:

$$\begin{cases} \frac{\partial L_{pd}}{\partial \widetilde{\boldsymbol{\omega}}} = \widetilde{\boldsymbol{\omega}} - \sum_{i=1}^{2N_T}(\alpha_i - \widetilde{\alpha}_i)\overline{\mathbf{x}}_i = 0 \\ \frac{\partial L_{pd}}{\partial \xi_i} = C - \alpha_i - \lambda_i = 0, \ i = 1, \ldots, 2N_T \\ \frac{\partial L_{pd}}{\partial \widetilde{\xi}_i} = C - \widetilde{\alpha}_i - \widetilde{\lambda}_i = 0, \ i = 1, \ldots, 2N_T \end{cases} \tag{10}$$

The weight vector $\widetilde{\boldsymbol{\omega}}$ is reformulated by means of the Lagrange multipliers: $\widetilde{\boldsymbol{\omega}} = \sum_{i=1}^{2N_T}(\alpha_i - \widetilde{\alpha}_i)\overline{\mathbf{x}}_i$. Substituting Equation (10) into Equation (9) yields the dual optimization problem:

$$\max_{\boldsymbol{\alpha}, \widetilde{\boldsymbol{\alpha}}} L_d(\boldsymbol{\alpha}, \widetilde{\boldsymbol{\alpha}}) = \max_{\boldsymbol{\alpha}, \widetilde{\boldsymbol{\alpha}}} [-\frac{1}{2}(\boldsymbol{\alpha} - \widetilde{\boldsymbol{\alpha}})^T(\overline{\mathbf{X}}\,\overline{\mathbf{X}}^T + \gamma\mathbf{I})(\boldsymbol{\alpha} - \widetilde{\boldsymbol{\alpha}}) - \epsilon\mathbf{1}^T(\boldsymbol{\alpha} + \widetilde{\boldsymbol{\alpha}}) + \widetilde{\mathbf{y}}^T(\boldsymbol{\alpha} - \widetilde{\boldsymbol{\alpha}})] \tag{11}$$

$$s.t. \ 0 \leqslant \alpha_i, \widetilde{\alpha}_i \leqslant C, \ i = 1, \ldots, 2N_T$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_{2N_T}]^T$ and $\widetilde{\boldsymbol{\alpha}} = [\widetilde{\alpha}_1, \widetilde{\alpha}_2, \ldots, \widetilde{\alpha}_{2N_T}]^T$; $\mathbf{1}$ is an all-one column vector with $2N_T$ elements.

In Equation (11), $\boldsymbol{\alpha}$ and $\widetilde{\boldsymbol{\alpha}}$ are the coefficients vectors maximizing the quadratic objective function, which can be found using any quadratic programming (QP) solvers [20]. A small identity term $\gamma\mathbf{I}$

is added in the dual objective function in case of ill-conditional matrix of $\overline{\mathbf{X}}\,\overline{\mathbf{X}}^T$ in Equation (11) [18]. The QP techniques are computationally demanding. Nevertheless, the number of samples in the considered situations being small, the increase of the computational burden is not significant, which will be shown qualitatively in the simulation part.

Finally, the weight vector $\widetilde{\omega}$ is obtained. We can rewrite it back into the complex domain as $\omega_i = \widetilde{\omega}_i + j\widetilde{\omega}_{i+P}$ for $i = 1, \ldots, P$ and then find the DOAs by searching the peak positions of the power spectrum density defined in Equation (6).

4. Simulation Results

In this section, the performance of the proposed method is evaluated with four simulations. We assume a ULA with ten isotropic sensors (i.e., $M = 10$). The distance between two adjacent sensors is half the wavelength of incoming signals, $d = \lambda/2$.

The research results in [15] show that the decorrelation ability of LP methods is at the expense of the real effective array aperture. In order to maintain an effective array aperture, we chose $P = 9$. Therefore, the number of snapshots L in the standard FBLP should be greater than five in order to make the covariance matrix invertible. The SVR parameters used in all the simulations are $\epsilon = 0.1$, $C = 0.1$, and $\gamma = 10^{-6}$, as in [20].

### 4.1. Performance with Power Spectrum Density

In the first simulation, we examine the PSD of the standard FBLP and the proposed FBLP-SVR not only with two sources, but also with three sources. In the two-source case, the signals come from $\theta_1 = 10°$, $\theta_2 = 20°$. When there are three sources, the signals come from $\theta_1 = -20°$, $\theta_2 = 0°$, $\theta_3 = 8°$. SNR = 10 dB. In both cases, two different numbers of snapshots are considered: $L = 100$ and $L = 5$. The spatial spectrum search is performed over $[-90°, 90°]$ with step size $0.01°$. The angles corresponding to the top two highest peak positions in the spectrum allow estimation of the DOAs of incoming signals. Figures 1 and 2 show the PSD of the methods with two and three coherent sources, respectively. The vertical dashed lines indicate the true values of DOAs.

As shown in the two-source case in Figure 1, FBLP-SVR and FBLP perform better with 100 snapshots than with 5 snapshots in terms of the positions of the peaks and stability of curves across the spectrum. When there are 100 snapshots, these two methods are able to detect the true DOAs. The PSD curves of FBLP and FBLP-SVR are similar to each other. However, in the scenario where there are only five snapshots, the FBLP-SVR method can accurately detect the signals while the standard FBLP fails. Similar results can be observed in a three-source case, as shown in Figure 2. The curves of FBLP-SVR have no false peaks, even when the number of snapshots is small.

(a)

(**b**)
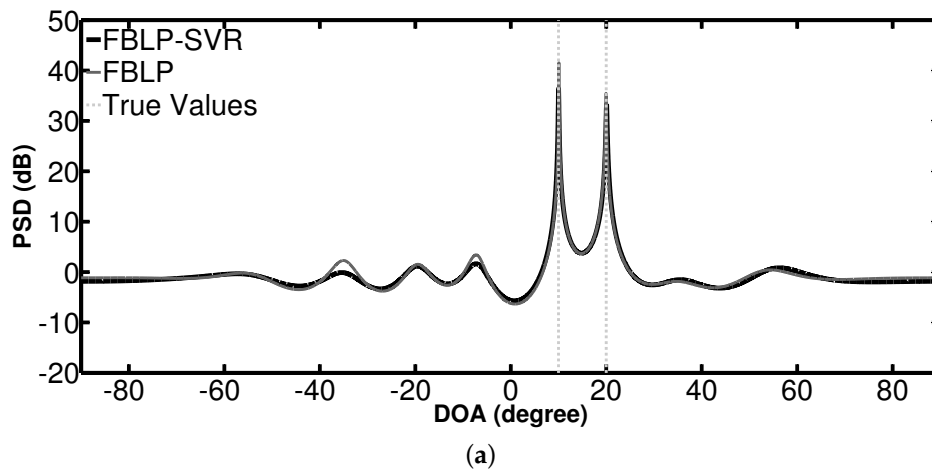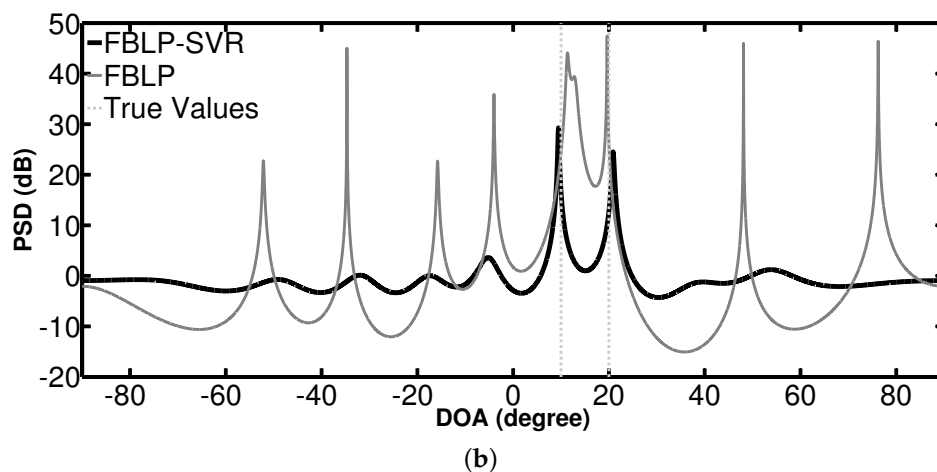
Figure 1. Power spectrum density (PSD) of combined forward–backward linear prediction and support vector regression (FBLP-SVR), and FBLP with 10 antennas and 2 coherent signals coming from $\theta_1 = 10°$, $\theta_2 = 20°$. (**a**) number of snapshots = 100; (**b**) number of snapshots = 5.


(**a**)

(**b**)
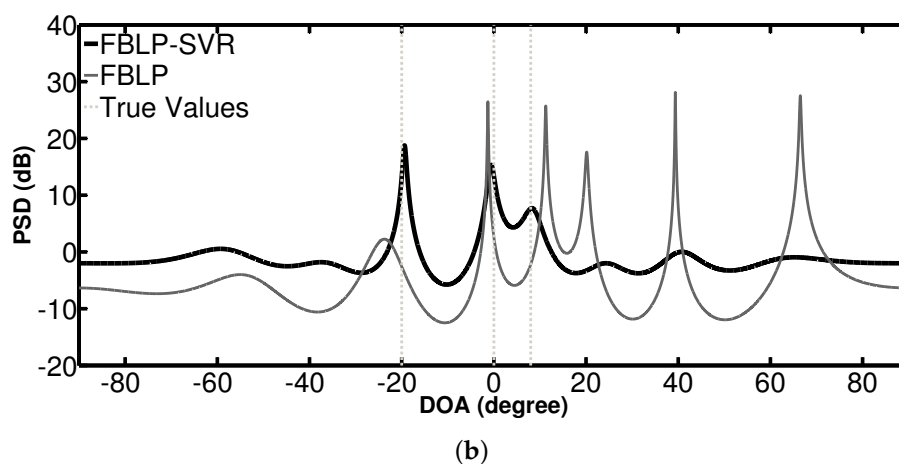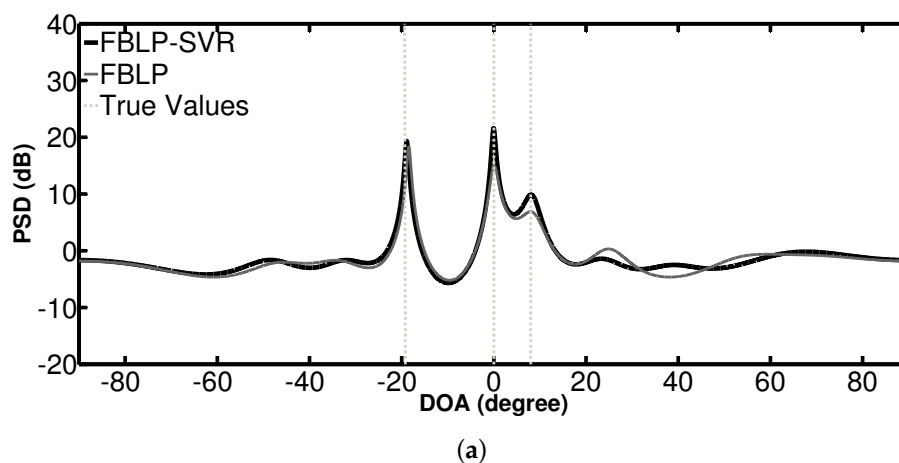
Figure 2. PSD of FBLP-SVR and FBLP with 10 antennas and 3 coherent signals coming from $\theta_1 = -20°$, $\theta_2 = 0°$, $\theta_3 = 8°$. (**a**) number of snapshots = 100; (**b**) number of snapshots = 5.

4.2. Performance versus Angle Separation

In the second simulation, the statistical performance of the proposed FBLP-SVR versus the angle separation between the incoming signals is assessed. For comparison, we also collect the results of the standard FBLP, FLP, and SVR-based FLP method (FLP-SVR). There are two sources in the performance analysis. One source is fixed at direction $\theta_1 = 0°$, while the other comes from $\theta_2 = \theta_1 + \Delta\theta$ with the same power. SNR = 15 dB and $\Delta\theta$ varies from $1°$ to $10°$. In the evaluation, root mean square error (RMSE) is applied based on 500 independent realizations. The definition of RMSE is given by [8]:

$$RMSE = \sqrt{\frac{1}{KJ} \sum_{j=1}^{J} \sum_{k=1}^{K} (\hat{\theta}_{kj} - \theta_k)^2} \tag{12}$$

where $\hat{\theta}_{kj}$ is the estimate of θ_k at the jth independent trial. J is the total number of trials. Figure 3 depicts the RMSE of DOA estimation versus angle separation for both 5 and 100 snapshots. As can be seen from Figure 3, all methods fail to detect the DOAs when the angle separation is small. When there are 100 snapshots and the angle separation is larger than $3°$, the results of SVR methods (FBLP-SVR, FLP-SVR) are similar to the standard LP methods (FBLP, FLP); the RMSE of FBLP-based methods (FBLP, FBLP-SVR) is smaller than that of FLP methods (FLP-SVR, FLP) for every angle separation. When there are 5 snapshots, the standard FBLP and FLP cannot work, regardless of the angle separation. The performance of FBLP-SVR and FLP-SVR with five snapshots is inferior to that with 100 snapshots. FBLP-SVR has the best accuracy and outperforms FBLP, FLP, and FLP-SVR at low angle separations.
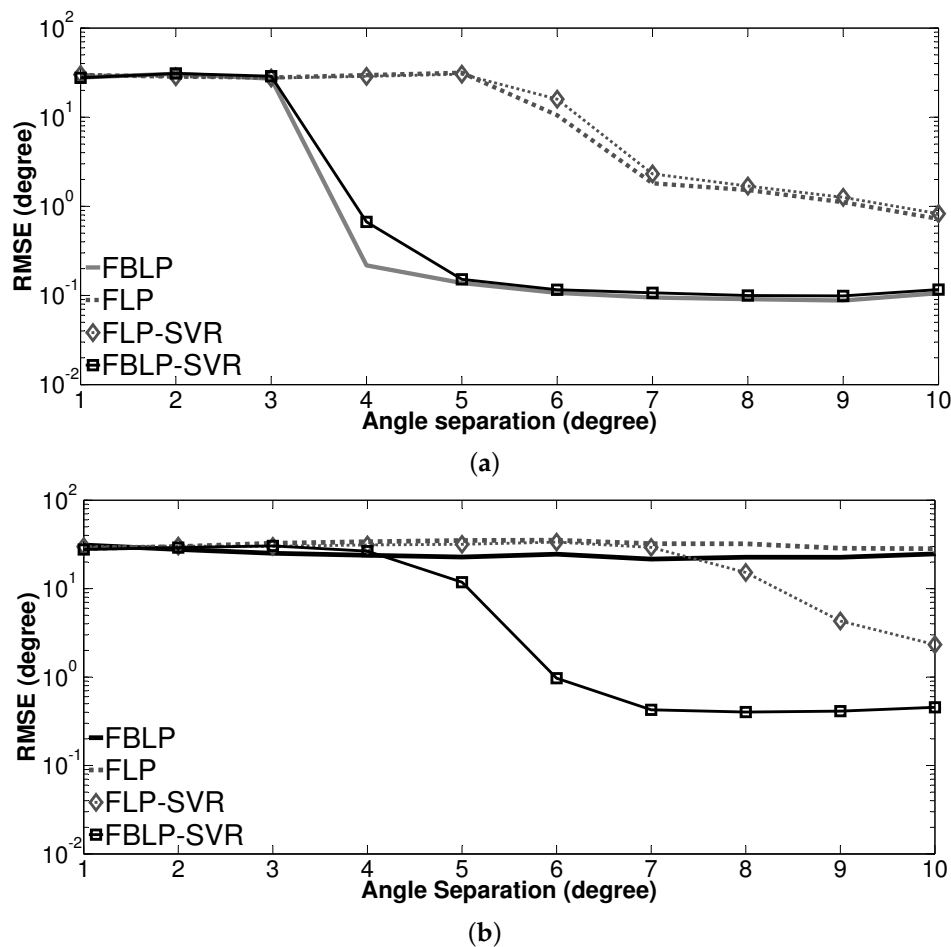
Figure 3. Root mean square error (RMSE) of direction-of-arrival (DOA) estimation as a function of angle separation. (**a**) number of snapshots = 100; (**b**) number of snapshots = 5.

4.3. Performance versus Number of Snapshots

In the third simulation, we test the performance of the proposed method as a function of the number of snapshots. The simulation conditions are similar to the second simulation, except the two coherent sources are from $\theta_1 = 0°$ and $\theta_2 = 6°$. The comparison is conducted between the standard FBLP and the proposed FBLP-SVR. The number of snapshots L varies from $[5, 10, 15, 20, \ldots, 50]$. At each number of snapshots, the statistical analysis is conducted with 500 independent trials. Figure 4 shows the RMSE of DOA estimation versus the number of snapshots. The RMSE of FBLP-SVR and FBLP decreases with the increase of the number of snapshots. When the number of snapshots is small, the proposed FBLP-SVR gives much better estimation result than the standard FBLP. As L gets larger, FBLP and FBLP-SVR achieve similar results.
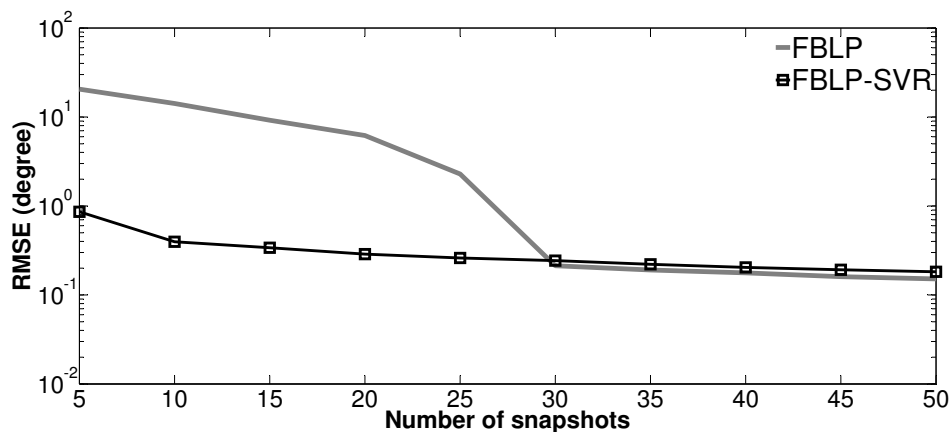

Figure 4. RMSE of DOA estimation versus number of snapshots.

In order to get an idea about the computational burden, the execution time is evaluated with $L = 5$ during 500 simulations. The average time for a single simulation using the proposed FBLP-SVR is 0.2728 s, while the corresponding time for the standard FBLP is 0.2595 s with a computer equipped with a processor unit (CPU) of 2.7 GHz and 16 GB of RAM. This indicates that the integration of SVR with FBLP can greatly improve the estimation performance with only a small additional computational burden, especially when the number of observation snapshots is limited.

4.4. Performance versus SNR

In the last simulation, we evaluate the performance of the proposed method with respect to SNR in coherent scenarios. The two sources come from directions $\theta_1 = 0°$ and $\theta_2 = 6°$, respectively. Only five snapshots are considered, and SNR varies from 0 to 30 dB. The RMSE is calculated with 500 independent trials for each SNR. The RMSE of angle estimation against SNR with 5 snapshots is plotted in Figure 5. It can be seen that FBLP-SVR has a more significant decrease of RMSE when SNR increases, compared with the standard FBLP.
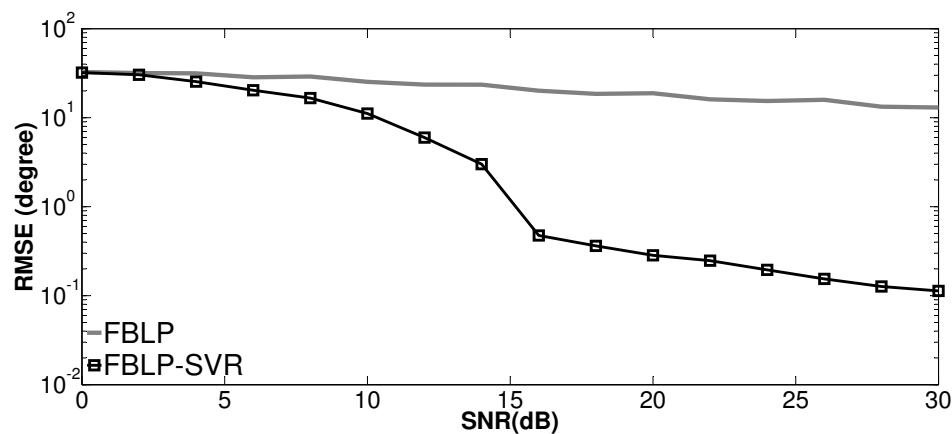
Figure 5. RMSE of DOA estimation versus signal-to-noise ratio (SNR) via 5 snapshots.

5. Conclusions

This paper proposes a combined version of SVR and FBLP in the estimation of the DOAs of coherent incoming signals, taking advantage of their inherent features. The proposed FBLP-SVR allows to directly deal with coherent signals, and remains applicable with a limited number of snapshots. Simulation results prove the stability and robustness of FBLP-SVR with coherent signals and low snapshots, in comparison with FLP, FLP-SVR, and FBLP.

Author Contributions: The general idea was proposed by Jingjing Pan and Yide Wang. Jingjing Pan designed and performed the simulations, analyzed the results and wrote the paper. Yide Wang and Cédric Le Bastard supervised the work, revised the manuscript and provided many valuable suggestions. All authors participated in amending the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DOA	Direction-of-Arrival
FBLP	Forward-Backward Linear Prediction
SVR	Support Vector Regression
SNR	Signal-to-noise
SS	Spatial smoothing
FB	Forward–backward
LP	Linear prediction
AR	Auto-regressive
ARMA	Auto-regressive moving average
FLP	Forward linear prediction
BLP	Backward linear prediction
MUSIC	Multiple Signal Classification
ESPRIT	Estimation of Signal Parameters via Rational Invariance Technique
ULA	Uniform Linear Array
AGWN	Additive Gaussian white noise
PSD	Power Spectrum Density
QP	Quadratic programming
RMSE	Root Mean Square Error

References

1. Krim, H.; Viberg, M. Two decades of array signal processing research: The parametric approach. *IEEE Signal Process. Mag.* **1996**, *13*, 67–94.
2. Tuncer, T.E.; Friedlander, B. Narrowband and wideband DOA estimation for uniform and nonuniform linear arrays. In *Classical and Modern Direction-of-Arrival Estimation*; Academic Press: Burlington, MA, USA, 2009; pp. 138–173.
3. Schmidt, R.O. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* **1986**, *34*, 276–280.
4. Roy, R.; Kailath, T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 984–995.
5. Marengo, E.A.; Gruber, F.K.; Simonetti, F. Time-reversal MUSIC imaging of extended targets. *IEEE Trans. Image Process.* **2007**, *16*, 1967–1984.
6. Ciuonzo, D.; Romano, G.; Solimene, R. Performance analysis of time-reversal MUSIC. *IEEE Trans. Signal Process.* **2015**, *63*, 2650–2662.
7. Ciuonzo, D.; Salvo Rossi, P. Noncolocated time-Reversal MUSIC: High-SNR distribution of null spectrum. *IEEE Signal Process. Lett.* **2017**, *24*, 397–401.
8. Wang, Y.; Trinkle, M.; Ng, B.W.H. DOA estimation under unknown mutual coupling and multipath with improved effective array aperture. *Sensors* **2015**, *15*, 30856–30869.
9. Shan, T.J.; Wax, M.; Kailath, T. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Trans. Acoust. Speech Signal Process.* **1985**, *33*, 806–811.
10. Pillai, S.U.; Kwon, B.H. Forward/backward spatial smoothing techniques for coherent signal identification. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 8–15.
11. Linebarger, D.A.; DeGroat, R.D.; Dowling, E.M. Efficient direction-finding methods employing forward/backward averaging. *IEEE Trans. Signal Process.* **1994**, *42*, 2136–2145.
12. Du, W.; Kirlin, R.L. Improved spatial smoothing techniques for DOA estimation of coherent signals. *IEEE Trans. Signal Process.* **1991**, *39*, 1208–1210.
13. Qi, C.; Wang, Y.; Zhang, Y.; Han, Y. Spatial difference smoothing for DOA estimation of coherent signals. *IEEE Signal Process. Lett.* **2005**, *12*, 800–802.
14. Wei, J.; Xu, X.; Luo, D.; Ye, Z. Sequential DOA estimation method for multi-group coherent signals. *Signal Process.* **2017**, *130*, 169–174.
15. Wang, Y.L.; Chen, H.; Peng, Y.; Wan, Q. Linear prediction. In *Spatial Spectrum Estimation Theory and Algorithm*; Tsinghua University: Beijing, China, 2004; pp. 83–109.
16. Xin, J.; Sane, A. Linear prediction approach to direction estimation of cyclostationary signals in multipath environment. *IEEE Trans. Signal Process.* **2001**, *49*, 710–720.
17. Vapnik, V.N. *Statistical Learning Theory*; Wiley: New York, NY, USA, 1998.
18. Martínez-Ramón, M.; Christodoulou, C. Support vector machines for antenna array processing and electromagnetics. In *Synthesis Lectures on Computational Electromagnetics*; Morgan and Claypool: San Rafael, CA, USA, 2005; Volume 1, pp. 1–120.
19. Gaudes, C.C.; Santamaria, I.; Via, J.; Gómez, E.M.; Paules, T.S. Robust array beamforming with sidelobe control using support vector machines. *IEEE Trans. Signal Process.* **2007**, *55*, 574–584.
20. El Gonnouni, A.; Martinez-Ramon, M.; Rojo-Álvarez, J.L.; Camps-Valls, G.; Figueiras-Vidal, A.R.; Christodoulou, C.G. A support vector machine MUSIC algorithm. *IEEE Trans. Antennas Propag.* **2012**, *60*, 4901–4910.
21. Pastorino, M.; Randazzo, A. A smart antenna system for direction of arrival estimation based on a support vector regression. *IEEE Trans. Antennas Propag.* **2005**, *53*, 2161–2168.
22. Le Bastard, C.; Wang, Y.; Baltazart, V.; Derobert, X. Time delay and permittivity estimation by ground-penetrating radar with support vector regression. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 873–877.
23. Rojo-Álvarez, J.L.; Martínez-Ramón, M.; de Prado-Cumplido, M.; Artés-Rodríguez, A.; Figueiras-Vidal, A.R. Support vector method for robust ARMA system identification. *IEEE Trans. Signal Process.* **2004**, *52*, 155–164.
24. Haykin, S.S. Linear prediction. In *Adaptive Filter Theory*; Pearson Education: New Delhi, India, 2008; pp. 136–148.

25. Chen, Y.H.; Chiang, C.T. Kalman-based spatial domain forward-backward linear predictor for DOA estimation. *IEEE Trans. Aerosp. Electron. Syst.* **1995**, *31*, 474–479.

26. Tufts, D.W.; Kumaresan, R. Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood. *Proc. IEEE* **1982**, *70*, 975–989.