Defensive Compressive Time Delay Estimation Using Information Bottleneck

Yiming Li, Rundong Wang, Yulin Hu, *Senior Member, IEEE*, Junan Yang, and Xiaoxia Cai

***Abstract*—Time delay estimation (TDE) is of great importance in reconnaissance and passive localization. Though TDE could leverage compressive sensing (CS) kernel to enjoy the advantage of sub-Nyquist samples, the estimation accuracy and robustness are inclined to be reduced by jamming or noise. To address the issue, this letter proposes a scheme of defensive compressive time delay estimation (DCTDE) with a novel sensing kernel optimization method and a defense augmentation strategy for the kernel. Inspired by the information bottleneck (IB) theory, we deduce a new analytic principle to enhance the estimation capability of the kernel, which focuses on the fidelity to the non-compressive scheme and the relevant information provided for TDE task. The strategy adopts an adversarial idea for the anti-spoofing design of kernel. Via simulations, the proposed scheme shows promising estimation accuracy, robustness, and high adaptability in both Gaussian noise and jamming environments, and more competitiveness than the Nyquist scheme.**

***Index Terms*—Compressive sensing (CS), information bottleneck (IB), sensing kernel optimization, anti-spoofing, time delay estimation (TDE).**

I. Introduction

TIME delay estimation (TDE) has been a fundamental research topic in the signal processing field, especially for the design of real-world applications. Generally, TDE schemes refer to finding the time differences-of-arrival between signals received at several sensors. A common approach is to find the peak of the cross-correlation function of signal samples. Hence, the estimation precision relies on a high sampling rate, which results in a heavy computational burden. To relieve the burden, compressive sensing (CS) technology [1] has been widely employed, since it can significantly lower the sampling rate to acquire the signal by capturing the sparse nature of signals.

Differing in the standard procedure of CS, which includes signal acquisition and reconstruction, compressive TDE schemes usually seek to directly estimate delay from the compressive samples [2], [3]. In general, these schemes acquire signals by utilizing a family of predefined sensing kernels, such as 1) random kernel, 2) Fourier basis kernel, 3) circulant kernel. More specifically, it is observed in [4], [5], the random kernel, including Gaussian, Bernoulli, ternary, *Etc.*, owns a high level of kernel variability to exploit the sparse nature of signals at the cost of estimation accuracy and computational efficiency. In addition, the authors in [6] demonstrate a Fourier-kernel-based TDE scheme achieving a higher estimation accuracy and computation efficiency while offering less kernel variability, due to a strict assumption of the kernel. As to the circulant kernel, the scheme proposed in [7] can enjoy an accurate estimation by exploiting a partial commutation property of the kernel. Moreover, more computational efficiency and comparable kernel variability are shown in [8] compared with the random-kernel-based scheme.

On the other hand, existing predefined kernels in TDE schemes are usually generated randomly, which make the estimation performance be unpredictable and uncontrolled. One approach to solve this problem involves the use of mutual information [9] to improve the estimation accuracy in a Gaussian environment. For instance, the authors in [10] proposed a scheme that utilizes the prior information of delay distribution based on the Bayesian framework. Likewise, a likelihood-oriented scheme, proposed in [11], adopts the Fisher information matrix to optimize the sub-Nyquist filter in delay-Doppler estimation. Compared to Gaussian noise, spoofing attacks pose greater threats to the applications [12]. Without anti-spoofing design, traditional TDE schemes are vulnerable to such attacks [13]. Consequently, CS TDE schemes suffer more due to the lossy acquisition.

Hence, it is essential and challenging to improve the accuracy and robustness of compressive TDE schemes due to its severe vulnerability to spoofing behaviors. In this letter, different from the above discussed schemes without anti-spoofing design, we propose a defensive compressive TDE (DCTDE) scheme in adversarial prerequisite which mainly includes a novel kernel optimization method and a defense augmentation strategy. Inspired by the information bottleneck (IB) theory, an analytic principle is deduced in the kernel optimization method. The strategy aims to augment the kernel to discriminate the actual delay relating to target signal and attenuate the effect induced by the jamming effect. The main contributions are as follows: 1) A kernel optimization method is proposed to utilize signal features, with an analytic principle deduced, 2) A defense augmentation strategy designed in spoofing prerequisite, and 3) the investigation of the feasibility of the proposed scheme in a jamming environment.

II. Defensive Compressive Time Delay Estimation

A. Time Delay Model and Delay Retrieval Method

All the related notations are shown in Table I. In signal processing field, given time delay d, the discrete delay is $\tau \approx d/\Delta$,

Manuscript received May 29, 2021; revised July 26, 2021; accepted August 12, 2021. Date of publication August 18, 2021; date of current version October 13, 2021. This work was supported in part by Anhui Provincial Natural Science Foundation under Grants 1908085MF202 and 2008085QF283. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xiangui Kang. *(Corresponding author: Junan Yang.)*

Yiming Li, Junan Yang, and Xiaoxia Cai are with the College of Electronic Engineering, National University of Defense Technology, Hefei 230009, China (e-mail: yimingli_jamin@qq.com; yangjunan@ustc.edu; cxxygg@163.com).

Rundong Wang is with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore (e-mail: rundong001@e.ntu.edu.sg).

Yulin Hu is with the School of Electronic Information, Wuhan University, Wuhan 430072, China (e-mail: yulin.hu@whu.edu.cn).

Digital Object Identifier 10.1109/LSP.2021.3105928

TABLE I
LIST OF NOTATIONS

Notation	Description (number field denoted by $\mathbb{F}$)
	Integer
d	the time delay of continuous waveform ($\mathbb{R}$)
n	the length of discrete waveform ($\mathbb{Z}_+$)
τ	the discrete delay of discrete waveform ($\mathbb{Z}$)
$\tilde{\tau}$	the discrete delay solution of the proposed scheme ($\mathbb{Z}$)
k, k_i, k_j	an arbitrary, the ith and the jth delay estimate ($\mathbb{Z}$)
Δ	the reciprocal of received waveform bandwidth ($\mathbb{R}$)
	Vector
$\boldsymbol{x_1}$, $\boldsymbol{x_2}$	the received discrete waveform ($\mathbb{R}^n$ *or* $\mathbb{C}^n$)
$\boldsymbol{s}$, $\boldsymbol{s_\tau}$	original signal and the delayed signal ($\mathbb{R}^n$ *or* $\mathbb{C}^n$)
$\boldsymbol{a_1}$, $\boldsymbol{a_2}$	the amplitude vectors of the received waveform ($\mathbb{R}^n$)
$\boldsymbol{s}_{a_1}$, $\boldsymbol{s}_{a_2}$	the received attack signals ($\mathbb{R}^n$ *or* $\mathbb{C}^n$)
$\boldsymbol{y_1}$, $\boldsymbol{y_2}$	the measurements acquired by CS receivers ($\mathbb{R}^{pq}$ *or* $\mathbb{C}^{pq}$)
$\boldsymbol{p_\tau}$, $\boldsymbol{1}_n$	patch signal ($\mathbb{R}^n$ *or* $\mathbb{C}^n$), vector of all ones ($\mathbb{R}^n$)
	Matrix
$\boldsymbol{\Phi}$	the sensing kernel with circulant structure ($\mathbb{R}^{pq\times n}$)
$\boldsymbol{\Psi}_i$	the ith minimal element of circulant matrix ($\mathbb{R}^{p\times n}$)
$\boldsymbol{H}_s^t$, $\boldsymbol{D}_s^t$, $\overline{\boldsymbol{D}}_s^t$	the auxiliary matrices for vector operation ($\mathbb{R}^{s\times s}$)
$\boldsymbol{0}_t$, $\boldsymbol{I}_t$	square matrix of all zeros, eye matrix ($\mathbb{R}^{t\times t}$)

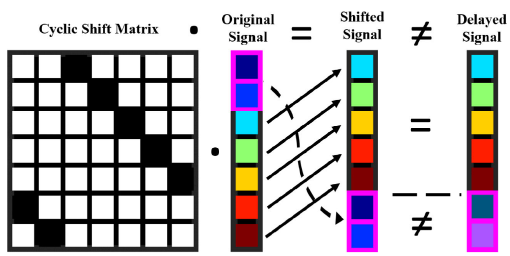

Fig. 1. The difference between cyclic shifted waveform and delayed one.

which is an integer. Then, the waveforms received by two receivers can be modeled as

$$\boldsymbol{x_1} = \boldsymbol{a_1} \circ \boldsymbol{s} + \boldsymbol{s}_{a_1}, \quad \boldsymbol{x_2} = \boldsymbol{a_2} \circ \boldsymbol{s_\tau} + \boldsymbol{s}_{a_2}, \tag{1}$$

where " $\circ$ " denotes hadamard product (entry-wise product), and $\boldsymbol{a_1}, \boldsymbol{a_2}$ indicate the amplitude vectors of received waveforms. To study the error of the model, the difference between shifted waveform and delayed one is introduced as patch signal

$$\boldsymbol{p_\tau} = \boldsymbol{s_\tau} - \boldsymbol{D_n^\tau s}, \tag{2}$$

where $\boldsymbol{D}_n^\tau$ is cyclic shift matrix defined in Eq. (18). As depicted in Fig. 1, the inevitable occurrence of the patch signal may cause the error of delay solution. To minimize it, we give assumptions as follows:

1) $\boldsymbol{s}$, $\boldsymbol{s}_{a_1}$ and $\boldsymbol{s}_{a_2}$ are zero-mean.
2) $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ are independent with $\boldsymbol{s}$ and $\boldsymbol{p_\tau}$.
3) $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ follow correlated Gaussian distribution.
4) $n \gg \tau$.

The measurements acquired by the CS receivers are

$$\boldsymbol{y_i} = \boldsymbol{\Phi x_i}, \quad (i = 1, 2) \tag{3}$$

where the kernel is a circulant matrix, which can be formulated as $\boldsymbol{\Phi} = [\boldsymbol{\Psi}_1{}^{\mathrm{T}}, \boldsymbol{\Psi}_2{}^{\mathrm{T}}, \ldots, \boldsymbol{\Psi}_q{}^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{\Psi}_i = \boldsymbol{\Psi}_1 \boldsymbol{D}_n^{(i-1)}$. Generally, kernel parameters should satisfy $pq \ll n$ for low-rate sampling. Then, the element $\boldsymbol{\Psi}_1$ is chosen as Gaussian type (vector or matrix) to generate a kernel. In addition, we summarize partial commutation property [7] of the circulant matrix as

$$\boldsymbol{H}_{pq}^k \boldsymbol{\Phi} \boldsymbol{D}_n^k = \overline{\boldsymbol{D}}_{pq}^k \boldsymbol{\Phi}, \tag{4}$$

where $\boldsymbol{H}_s^t$ and $\overline{\boldsymbol{D}}_s^t$ are defined in Eq. (18). Assumed with noise-free environment and fixed amplitude of received waveforms, the measurements can be expressed as

$$\boldsymbol{y_1} = \boldsymbol{\Phi s}, \;\; \boldsymbol{y_2} = \boldsymbol{\Phi D_n^\tau s}. \tag{5}$$

We multiply $\boldsymbol{D}_n^\tau \boldsymbol{s}$ after Eq. (4) and simplify the results as

$$\boldsymbol{H}_{pq}^k \boldsymbol{\Phi}\, \boldsymbol{D}_n^k \boldsymbol{D}_n^\tau \boldsymbol{s} = \overline{\boldsymbol{D}}_{pq}^k\, \boldsymbol{y_2}. \tag{6}$$

We note that, when $k = -\tau$, the operation matrix owns that $\boldsymbol{D}_n^k \boldsymbol{D}_n^\tau = \boldsymbol{I}$, which could be substituted in Eq. (6) and yields

$$\boldsymbol{H}_{pq}^{-\tau}\, \boldsymbol{y_1} = \overline{\boldsymbol{D}}_{pq}^{-\tau}\, \boldsymbol{y_2}. \tag{7}$$

Eq. (7) demonstrates the delay retrievablity of CS measurements. Then, we have CS correlation function

$$\tilde{\boldsymbol{C}}(k) = F_k \langle \boldsymbol{H}_{pq}^k\, \boldsymbol{y_1}, \overline{\boldsymbol{D}}_{pq}^k\, \boldsymbol{y_2} \rangle, \tag{8}$$

where factor F_k, later presented in Eq. (17), guarantees a fair yardstick for different delay estimates. Additionally, we find an identical relation relating to the inner product in Eq. (8) as

$$\langle \boldsymbol{H}_{pq}^k\, \boldsymbol{y_1}, \overline{\boldsymbol{D}}_{pq}^k\, \boldsymbol{y_2} \rangle = \langle \boldsymbol{y_1}, \overline{\boldsymbol{D}}_{pq}^k\, \boldsymbol{y_2} \rangle \tag{9}$$

Then, we substitute Eq. (9) into Eq. (8) and get delay solution

$$\tilde{\tau} = -\arg\max_k F_k \langle \boldsymbol{y_1}, \overline{\boldsymbol{D}}_{pq}^k \boldsymbol{y_2} \rangle. \tag{10}$$

Thus, time delay solution is $\tilde{\tau}\Delta$, and we demand $q > |\tau|_{\max}$ for a valid TDE solution. Given the same sampling rate, the resolution of CS scheme is much higher than the native one of Nyquist scheme. However, CS inherent tradeoff between SNR loss and accuracy gain is imposing the restrictions on a broader and more practical application.

B. Kernel Estimation Capability Optimization

The previous studies of improving kernel capability fall into two typical categories. It is shown in [6] that, under a mild condition, delay can be retrieved from a very sparse measurement by a strict kernel. Additionally, it is observed in [14] that, a less strict kernel may precisely fit the need of pattern detection problem. According to information theory, the former scheme aims to wipe off the extraneous information in waveform for a more compound measurement, while the objective of the latter is to increase task information for a more meaningful measurement. Inspired by information bottleneck (IB), we can optimize a kernel to let receivers acquire both compound and meaningful measurements. We recall the variational principle [15] in IB as

$$\mathcal{L} = I(\boldsymbol{x}, \tilde{\boldsymbol{x}}) - \beta\, [I(\boldsymbol{y}, \tilde{\boldsymbol{x}})], \tag{11}$$

where $I(:,:)$ denotes mutual information entropy, and β is a Lagrange multiplier.

The original thought in IB is that, through a limited codewords $\tilde{\boldsymbol{x}}$, the information of $\boldsymbol{x}$ is squeezed to predict the $\boldsymbol{y}$. In a passive scenario, the compressed correlation $\tilde{\boldsymbol{c}}$ is chosen as codewords $\tilde{\boldsymbol{x}}$ because it contains the information acquired by the two receivers. For a more compound measurement, we can measure its fidelity by comparing it with original correlation $\boldsymbol{c}$; for a more meaningful measurement, we should evaluate its relevant information relating to the TDE task (estimated parameter τ). By exploiting prior delay distribution, Eq. (11) is expressed as

$$\mathcal{L} = I(\boldsymbol{c}, \tilde{\boldsymbol{c}}) - \beta\, \mathrm{E}_\tau[\, I(\tau, \tilde{\boldsymbol{c}})\,], \tag{12}$$

which could be simplified as

$$\mathcal{L} = (1-\beta) h(\tilde{\boldsymbol{c}}) - \beta\, \mathrm{E}_\tau[\, h(\tilde{\boldsymbol{c}} | \tau)\,]. \tag{13}$$

By substituting Eq. (1) and Eq. (3) in Eq. (8), we get the compressed correlation free of attack signal, (15)–(18) are shown at

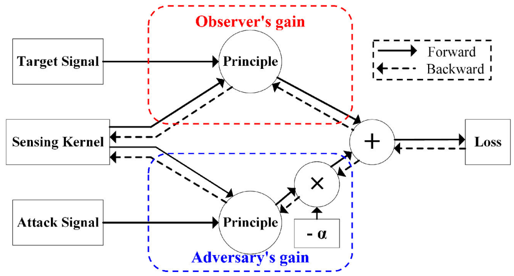

Fig. 2. Kernel defense augmentation strategy (calculation graph).

bottom of the page.

$$\tilde{\boldsymbol{c}}(k) = F_k \langle \boldsymbol{\Phi}(\boldsymbol{a_1} \circ \boldsymbol{s}), \boldsymbol{H}_{pq}^k \boldsymbol{\Phi}\, \boldsymbol{D}_n^k(\boldsymbol{a_2} \circ \boldsymbol{s_\tau}) \rangle. \tag{14}$$

As the central limit theorem [16] indicates, the asymptotic normality could be determined if samples of weak dependence are obtained in enough quantity. Since the CS correlation meets the above condition, we assume the correlation follows the Gaussian distribution $\tilde{\boldsymbol{c}} \sim \mathcal{N}(\boldsymbol{\mu_c}^{\mathrm{T}}, \boldsymbol{\Sigma}_{\tilde{\boldsymbol{c}}})$, where $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{c}}} = \mathrm{E}_\tau[\tilde{\boldsymbol{c}}\tilde{\boldsymbol{c}}^{\mathrm{T}}]$. Then, we get conditional covariance given by element-wise manner $(\boldsymbol{\Sigma}_{\tilde{\boldsymbol{c}}}|\tau(i,j))$ in Eq. (15). By substituting Eq. (15) with total probability covariance law: $\mathrm{Cov}[\boldsymbol{y_1}, \boldsymbol{y_2}] = \mathrm{E}[\mathrm{Cov}[\boldsymbol{y_1}, \boldsymbol{y_2}|\boldsymbol{x}]] + \mathrm{Cov}[\mathrm{E}[\boldsymbol{y_1}|\boldsymbol{x}], \mathrm{E}[\boldsymbol{y_2}|\boldsymbol{x}]]$, we get the total covariance $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{c}}}$ in Eq. (16). Finally, we get the analytic principle of kernel estimation capability optimization

$$\mathcal{L}(\boldsymbol{s}; \boldsymbol{\Phi}) = (1-\beta)\, |\boldsymbol{\Sigma}_{\tilde{\boldsymbol{c}}}| - \beta\, \mathrm{E}_\tau[\, |\boldsymbol{\Sigma}_{\tilde{\boldsymbol{c}}}|_\tau|\,]. \tag{19}$$

We note that the patch signal is eliminated through the derivation of the principle. That is to say, the native error introduced by TDE stage would not affect optimization stage. By gradient descent method, the solution is obtained as

$$\boldsymbol{\Phi} = \arg\min_{\boldsymbol{\Phi}} \mathcal{L}(\boldsymbol{s}; \boldsymbol{\Phi}) \tag{20}$$

C. Kernel Defense Augmentation Strategy

With the frequent occurence of jamming and spoofing behaviors, the robustness of compressive TDE schemes are seriously threatened due to its low-rate sampling. Therefore, the spoofing scenario is given as: the observer tries to conduct compressive TDE task, but deceived by the adversary who masks the target signal $\boldsymbol{s}$ with similar attack signal $\boldsymbol{s}_a$.

To counter the threat, we devise the strategy in light of adversarial learning, involving observer's gain and adversary's. Namely, observer's gain can be determined by how effectively the kernel is optimized, which could be evaluated by the principle $\mathcal{L}(\boldsymbol{s})$, and adversary's gain can be obtained by the extent to which observer's optimization is spoofed by the attack denoted by $\mathcal{L}(\boldsymbol{s}_a)$. In the scenario, the maximum of observer's gain is bounded by the adversary. Therefore, the mitigation of the adversary's gain could further raise the upper-bound of anti-spoofing capability by exploiting the attack signal. Then, we incorporate adversary's gain as a penalty term to modify the principle as

$$\mathcal{L}_{\mathrm{Def}}(\boldsymbol{s}, \boldsymbol{s}_a; \boldsymbol{\Phi}) = \mathcal{L}(\boldsymbol{s}) - \alpha \mathcal{L}(\boldsymbol{s}_a), \tag{21}$$

where α is a hyperparameter controlling the strength of anti-spoofing capability. Strategy diagram is presented in Fig. 2.

Moreover, the penalty would not only augment anti-spoofing capability, but also moderately weaken TDE capability of target signal due to the similarity between target and attack signals. Since enhancing the latter capability is obviously prioritized than the former, observer's gain should always outweigh adversary penalty as $\alpha \in [0, 1]$ for an eligible and meaningful optimization.

III. Simulations

In this section, we compared the proposed scheme with that of Gaussian random sensing, information-theoretic sensing (IT sensing) [10] and Nyquist GCC-PHAT [17]. We consider passive scenario which requires at least two receivers. However, the TDE stage of IT sensing is not suitable for the scenario. For this reason, we adopt the optimization stage of IT sensing and simulation parameter setups are provided in Table II. In the comparison, the metrics used in the simulation include correct retrieval probability and MSE. We note that one correct retrieval is counted as $\tilde{\tau} \in [d/\Delta - \frac{1}{2}, d/\Delta + \frac{1}{2}]$.

In Fig. 3(a), we compared the correct retrieval probability in Gaussian noise. The sampling rate of CS schemes is the 1/4 Nyquist rate of target signal. As shown in the plot, optimal sensing outperforms random sensing from -5 dB to 20 dB SNR. Hence, the effectiveness of the proposed scheme is confirmed. Another important finding is that, at the interval SNR $\in$ [0 dB,15

$$\boldsymbol{\Sigma}_{\tilde{\boldsymbol{c}}}|_\tau(i,j) = F_{k_i}F_{k_j}\{\mathrm{D}(\boldsymbol{a_1}\boldsymbol{a_2})[\mathcal{E}(\boldsymbol{s})^{\mathrm{T}}\boldsymbol{G}_1\mathcal{E}(\boldsymbol{s})] + \mathrm{E}[\boldsymbol{a_1}^2]\mathrm{E}[\boldsymbol{a_2}^2][\mathcal{E}(\boldsymbol{s})^{\mathrm{T}}\boldsymbol{G}_2\mathcal{E}(\boldsymbol{s})] + \mathrm{E}[\boldsymbol{a_1}\boldsymbol{a_2}]^2[\mathcal{E}(\boldsymbol{s})^{\mathrm{T}}\boldsymbol{G}_3\mathcal{E}(\boldsymbol{s})]\} \tag{15}$$

$$\begin{aligned}\boldsymbol{\Sigma}_{\tilde{\boldsymbol{c}}}(i,j) = F_{k_i}F_{k_j}\{&\boldsymbol{D}(\boldsymbol{a_1}\boldsymbol{a_2})[\mathcal{E}(\boldsymbol{s})^{\mathrm{T}}\mathrm{E}_\tau[\boldsymbol{G}_1(\tau)]\mathcal{E}(\boldsymbol{s})] + \mathrm{E}[\boldsymbol{a_1}^2]\mathrm{E}[\boldsymbol{a_2}^2][\mathcal{E}(\boldsymbol{s})^{\mathrm{T}}\mathrm{E}_\tau[\boldsymbol{G}_2(\tau)]\mathcal{E}(\boldsymbol{s})] \\ &+ \mathrm{E}[\boldsymbol{a_1}\boldsymbol{a_2}]^2[\mathcal{E}(\boldsymbol{s})^{\mathrm{T}}\mathrm{E}_\tau[\boldsymbol{G}_3(\tau) + \boldsymbol{G}_4(\tau)]\mathcal{E}(\boldsymbol{s})]\}\end{aligned} \tag{16}$$

$$\begin{array}{lll} \boldsymbol{G}_1(\tau) = (\boldsymbol{I} \circ \boldsymbol{B}_{-\tau}^{k_i}) \cdot (\boldsymbol{I} \circ \boldsymbol{B}_{-\tau}^{k_j}) & \boldsymbol{G}_2 = (\boldsymbol{B}^{k_i} - \boldsymbol{I} \circ \boldsymbol{B}^{k_i}) \circ (\boldsymbol{B}^{k_j} - \boldsymbol{I} \circ \boldsymbol{B}^{k_j}) & \boldsymbol{G}_3(\tau) = (\boldsymbol{B}_{-\tau}^{k_i} - \boldsymbol{I} \circ \boldsymbol{B}_{-\tau}^{k_i}) \\ & & \quad \circ((\boldsymbol{B}_{-\tau}^{k_j})^{\mathrm{T}} - \boldsymbol{I} \circ (\boldsymbol{B}_{-\tau}^{k_j})^{\mathrm{T}}) \\ \boldsymbol{G}_4(\tau) = ((\boldsymbol{1}_n)^{\mathrm{T}} \cdot (\boldsymbol{I} \circ \boldsymbol{B}_{-\tau}^{k_i}))^{\mathrm{T}} \cdot ((\boldsymbol{1}_n)^{\mathrm{T}} \cdot (\boldsymbol{I} \circ \boldsymbol{B}_{-\tau}^{k_j})) & \boldsymbol{B}^k = \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{H}_{pq}^k\boldsymbol{\Phi}\boldsymbol{D}_n^k & \boldsymbol{B}_{-\tau}^k = \boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{H}_{pq}^k\boldsymbol{\Phi}\boldsymbol{D}_n^k\boldsymbol{D}_n^{-\tau} \\ \mathcal{E}(\boldsymbol{x}) = \boldsymbol{x} \circ \boldsymbol{x},\, \boldsymbol{x} \in \mathbb{R}^n \;\mathit{or}\; \Sigma_{i=1}^n |\boldsymbol{x}_i|_2^2,\, \boldsymbol{x} \in \mathbb{C}^n & \mathrm{E}[\boldsymbol{a_i}^2] = \sigma_i^2 + \mu_i^2 \quad (i = 1, 2) & \mathrm{E}[\boldsymbol{a_1}\boldsymbol{a_2}] = \rho\sigma_1\sigma_2 + \mu_1\mu_2 \\ \mathrm{E}[\boldsymbol{a_1}^2\boldsymbol{a_2}^2] = (1+3\rho^2)\sigma_1^2\sigma_2^2 + (1-4\rho^2)\mu_1^2\sigma_2^2 & \mathrm{D}(\boldsymbol{a_1}\boldsymbol{a_2}) = (1+2\rho^2)\sigma_1^2\sigma_2^2 + (1-4\rho^2)\mu_1^2\sigma_2^2 & F_k = n/(n - |k|) \\ \quad + \mu_2^2\sigma_1^2 + \mu_1^2\mu_2^2 + 4\rho\mu_1\mu_2\sigma_1\sigma_2 & \quad + \sigma_1^2\mu_2^2 + 2\rho\mu_1\mu_2\sigma_1\sigma_2 & \end{array} \tag{17}$$

$$\boldsymbol{D}_s^t = \begin{cases} \begin{bmatrix} & \boldsymbol{I}_{s-t} \\ \boldsymbol{I}_t & \end{bmatrix}, & t \geq 0 \\ (\boldsymbol{D}_s^{-t})^{\mathrm{T}}, & t < 0 \end{cases} \quad \boldsymbol{H}_s^t = \begin{cases} \begin{bmatrix} \boldsymbol{I}_{s-t} & \\ & \boldsymbol{0}_t \end{bmatrix}, & t \geq 0 \\ \boldsymbol{M}\boldsymbol{H}_s^{-t}\boldsymbol{M}, & t < 0 \end{cases}, \boldsymbol{M} = \begin{bmatrix} & & 1 \\ & 1 & \\ \cdots & & \end{bmatrix} \overline{\boldsymbol{D}}_s^t = \begin{cases} \begin{bmatrix} & \boldsymbol{I}_{s-t} \\ \boldsymbol{0}_t & \end{bmatrix}, & t \geq 0 \\ (\overline{\boldsymbol{D}}_s^{-t})^{\mathrm{T}}, & t < 0 \end{cases} \tag{18}$$

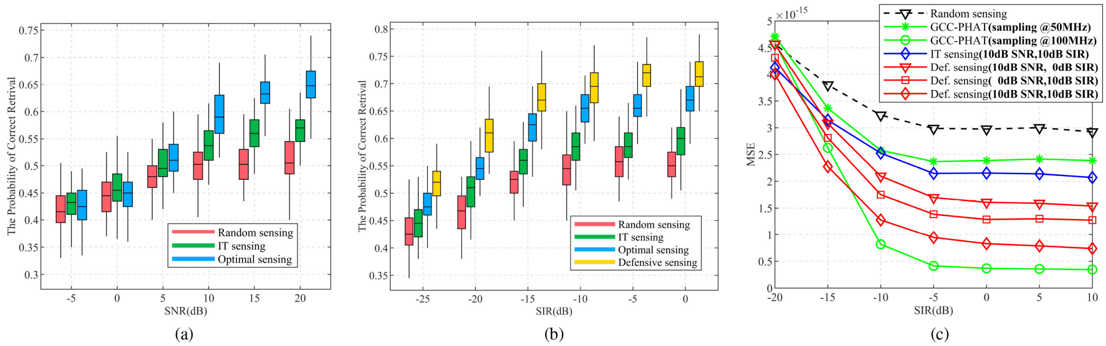

Fig. 3. TDE Performance comparison of different schemes. (a) The probability of correct retrieval in Gaussian noise environment; (b) The probability of correct retrieval in jamming environment at SNR = 10 dB; (c) MSE comparison in jamming environment at SNR = 10 dB.

TABLE II
SIMULATION AND SCHEME PARAMETER SETTING

Item	Learning rate	Scheme parameter	Kernel setting	Amplitudes distribution	*a prior* information	Resolution	Sampling rate	Signal character
Value	1e-6	$\alpha = 0.3$ $\beta = 0.01$	p=1, q=10, n=200 (a), (b) p=1, q=20, n=200 (c)	$\mathcal{N}(\mathbf{1}, 0.1\boldsymbol{I})$, $\rho = 0.9$	$d \sim \mathbf{U}[-1\mu s, 1\mu s]$	$0.1\mu s$ (All schemes)	5MHz (a), (b) 10MHz (c)	bpsk B=10MHz (Target) bpsk B=7MHz (Attack)

dB], the performance of optimal sensing (kernel estimation capability optimization) increases at a more sloping angle than other schemes. This promising behavior reveals that, when target signal is less noise-masked, the proposed scheme enable CS receivers to obtain more meaningful information for TDE task than others. Furthermore, it is exciting to notice that the proposed scheme significantly extends the asymptotic performance limit of CS scheme at high SNR region(e.g., above 15 dB). When CS rate is fixed, there is a inherent tradeoff between SNR loss and TDE performance which draws the boundary of application. We have pushed this inherent boundary forward, so the proposed scheme could offer a better solution of kernel.

We further study the scheme performance in a spoofing scenario, where the adversary spoofs the observer by using the same knowledge of $a\ prior$ delay probability distribution and signal character. It is shown in Fig. 3(b), defensive sensing (defense augmentation strategy) obtains more accurate results than others. To illustrate that, even at SIR $= -25$ dB, the proposed scheme can maintain 0.5 correct probability in most cases. Furthermore, most schemes reach their asymptotic performance limit at SIR $\in [-10$ dB,0 dB]. The results in this region shows that, a kernel can develop more anti-spoofing capability if defensive sensing scheme is used. In other words, even in a spoofing environment, the inherent tradeoff between SIR loss and TDE performance can be significantly improved. We also note that, even without the use of the defensive strategy, anti-spoofing capability can also be enhanced. These results demonstrate the robustness of the two proposed schemes in spoofing environment.

Generally, training samples play a vital role in the optimization scheme. In previous experiments, we supposed that the observer obtains noise-free samples and prior information before training. Unfortunately, neither of them can be easily obtained in practice. Hence, it is essential to study the effectiveness, sensitivity and performance of the schemes, if its training samples are contaminated in a spoofing scenario.

Firstly, the proposed scheme shows better MSE performance than random sensing. In fact, the sample condition SIR $= 0$ dB is the minimum guarantee for a valid training, otherwise the condition SIR < 0 dB leads to an opposite and false labelling (target or attack signal) of training samples. Therefore, the results demonstrate the pragmatic use of the proposed schemes.

Secondly, we conducted a sensitivity test to investigate the impact of different sample noise and jamming power on TDE performance. As the red line illustrating in the figure, at SIR $\in [-10$ dB,10d B], the proposed scheme with sample SIR deteriorating from 10 dB to 0 dB undertakes nearly 1.5 X amount of the loss induced by the same deteriorating of SNR. Namely, the optimization stage is more sensitive to jamming than Gaussian noise. Thus, anti-spoofing strategy is crucial to a scheme design at both optimization and TDE stages.

Thirdly, we compare the performance of IT sensing scheme and the defensive scheme. Obviously, trained with the same samples (e.g., 10 dB SNR and 10 dB SIR), the proposed one can unleash a stronger anti-spoofing capability of the CS TDE scheme than the compared one.

Finally, Nyquist anti-spoofing scheme GCC-PHAT [17] is also tested. If sampling at 100 MHz (10X CS rate), Nyquist scheme owns the same resolution with the compared CS schemes. As shown in the figure, Nyquist scheme performance is perfect at such sampling rate, while the sampling burden seriously weakens its practical competitiveness. Otherwise, Nyquist scheme undertakes significant resolution loss at 50 MHz (5X CS rate). Thus, considering both the burden and loss, DCTDE scheme has a wider application than the Nyquist one.

IV. CONCLUSION

This letter introduced DCTDE scheme with a new TDE stage and optimization stages. In the optimization stages, we deduced a new principle using IB theory for an enhanced TDE capability, and designed a novel strategy for augmented anti-spoofing capacity. A set of simulation results demonstrated and confirmed the effectiveness and robustness of the proposed schemes. Finally, we verified the high adaptability of the proposed schemes, when training samples are contaminated by jamming and Gaussian noise. From the above results, the scheme is more competitive than Nyquist one and compared one.

REFERENCES

[1] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] S. Salari, F. Chan, Y.-T. Chan, and W. Read, “TDOA estimation with compressive sensing measurements and hadamard matrix,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 6, pp. 3137–3142, Dec. 2018.

[3] A. Jain, P. Sarvepalli, S. Bhashyam, and A. P. Kannu, “Algorithms for change detection with sparse signals,” *IEEE Trans. Signal Process.*, vol. 68, pp. 1331–1345, 2020.

[4] A. G. Zebadua, P. Amblard, E. Moisan, and O. J. J. Michel, “Compressed and quantized correlation estimators,” *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 56–68, Jan. 2017.

[5] X. Ding and H.-T. Wu, “On the spectral property of kernel-based sensor fusion algorithms of high dimensional data,” *IEEE Trans. Inf. Theory*, vol. 67, no. 1, pp. 640–670, Jan. 2021.

[6] H. Ohlsson, Y. C. Eldar, A. Y. Yang, and S. S. Sastry, “Compressive shift retrieval,” *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4105–4113, Aug. 2014.

[7] D. Valsesia and E. Magli, “Compressive signal processing with circulant sensing matrices,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2014, pp. 1015–1019.

[8] L. Ding *et al.*, “Approximate kernel selection via matrix approximation,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4881–4891, Nov. 2020.

[9] B. Tang and J. Li, “Spectrally constrained MIMO radar waveform design based on mutual information,” *IEEE Trans. Signal Process.*, vol. 67, no. 3, pp. 821–834, Feb. 2019.

[10] Y. Gu and N. A. Goodman, “Information-theoretic compressive sensing kernel optimization and Bayesian Cramér-Rao bound for time delay estimation,” *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4525–4537, Sep. 2017.

[11] A. Lenz, M. S. Stein, and A. L. Swindlehurst, “Joint transmit and receive filter optimization for sub-Nyquist delay-doppler estimation,” *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2542–2556, May 2018.

[12] A. Luo, E. Li, Y. Liu, X. Kang, and Z. J. Wang, “A capsule network based approach for detection of audio spoofing attacks,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2021, pp. 6359–6363.

[13] R. T. Ioannides, T. Pany, and G. Gibbons, “Known vulnerabilities of global navigation satellite systems, status, and potential mitigation techniques,” *Proc. IEEE*, vol. 104, no. 6, pp. 1174–1194, Jun. 2016.

[14] S. Rousseau and D. Helbert, “Compressive color pattern detection using partial orthogonal circulant sensing matrix,” *IEEE Trans. Image Process.*, vol. 29, pp. 670–678, 2020.

[15] N. Tishby, F. C. Pereira, and W. Bialek, “The information bottleneck method,” 2000, *physics/0004057*.

[16] M. H. Neumann, “A central limit theorem for triangular arrays of weakly dependent random variables,”, 2010.

[17] C. Knapp and G. Carter, “The generalized correlation method for estimation of time delay,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-24, no. 4, pp. 320–327, Aug. 1976.