

# An Information Theoretic Take on Time Reversal for Non-Stationary Channels

Pawan Setlur and Natasha Devroye

**Abstract**—It has been shown that time-reversal (TR) techniques focus energy back to the dominant scatters, lead to super-resolution focusing, and gains in detection. Time reversal has so far mainly been studied when the channel remains invariant between the initial and time-reversed signal transmission times. In this letter, we relax this assumption and study the benefits of TR over time-varying channels. To do so, we compare a time-reversed and a non time-reversed system by comparing the mutual information between the channel impulse response and channel outputs given the transmitted signals. We present analytical results for a simple scalar problem which illustrates the impact of non-stationary channels on TR, and for general channels, numerically evaluate the difference in mutual informations, which demonstrate that, if the channels are non-stationary yet correlated, TR may still provide mutual information gains over non time-reversed systems.

**Index Terms**—Time reversal, Radar, Mutual Information, Stochastic time-varying channel, non-stationary channel

## I. INTRODUCTION

IN this letter, we analyze radar-based time reversal (TR) over time-varying channels using mutual information as comparison metric. In TR, a signal is initially radiated; the backscattered signal is then recorded, time-reversed, energy scaled and re-transmitted. TR may lead to super-resolution spatio-temporal focusing using multiple antennas, and detection gains for single and multiple antennas [1]–[6]. Most of these important contributions in TR were derived assuming the channel to be invariant from the initial signal transmission to the time-reversed re-transmission [1]–[6]. The question of whether time-reversal is beneficial in realistic time-varying channels remains. We make analytical progress by introducing and analyzing the *mutual information* in TR systems as compared to conventional systems for time-varying channels.

Our TR model is well suited to monostatic radar, where the receiver and transmitter are co-located but could be extended to other applications such as communications. In our model, a single antenna transceiver first probes the channel. It subsequently transmits the time-reversed signal it received from the initial probe. We will compare performance with a “conventional” model where the channel is probed twice with the same

signal (as opposed to recent waveform scheduling models [7]–[9]). The channels are assumed to be linear, stochastic, subject to additive Gaussian noise, and time-varying.

**Related work.** For invariant channels time-reversal in optics, ultrasound and acoustics, radar and communications, may be seen in for example [1]–[6] and references therein. Work on TR in time-varying channels is much more limited. It was acknowledged in [4, pg. 36-37] via experimental insights that TR focusing degrades in non-stationary time-varying environments such as the time-varying ocean surface and its volume. In communication applications (rather than our radar-focused application), it was experimentally shown that time varying channels affect the TR performance in [10], [11], and interestingly the conclusions drawn are similar to those drawn here – that TR may still be beneficial in channels which are correlated but not necessarily identical over time.

**Contributions and organization.** We first introduce the TR and conventional channel models in Section II. We then introduce the relevant mutual information quantity for these models in Section III, before analytically comparing the difference in mutual information for TR and conventional channels in Section IV, where we analytically work out an intuitive special scalar case which illustrates the effect of channel correlation between the two stages on the mutual information. Finally, in Section V we provide numerical evaluations. Our central contributions are 1) analyzing TR using information theoretic metrics for the first time, and 2) using this framework to quantitatively analyze the impact of time-varying channels on TR as compared to conventional systems.

## II. CHANNEL MODEL

We now outline the two transmission stages of the TR channel and the conventional channel models. Other “time-reversal” protocols consisting of more than two stages may be devised, but for simplicity, and to understand existing TR channel models, we limit ourselves to two stages.

During the first stage, let the baseband transmitted signal during the first scheduling instant or stage be given by the  $P$  baseband samples  $\mathbf{s} := [s(0), s(1), \dots, s(P-1)]^T$ , and let the matrix  $\mathbf{S} \in \mathbb{C}^{N \times M}$  be the Toeplitz convolution matrix comprised of the samples  $\mathbf{s}$  along with zeros padded appropriately – i.e. the  $i$ -th column consists of  $(i-1)$  zeros, the  $P$  samples of  $\mathbf{s}$ , and  $N-P-i+1$  zeros again. Note that  $N = M-1+P$ . The received signal at  $\mathbf{y}_1$  is then given by the  $N$  complex samples

$$\mathbf{y}_1 = \mathbf{S}\mathbf{a}_1 + \mathbf{v}_1, \quad (1)$$

Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

The authors are with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL, 60607 USA, e-mail: {setlur, devroye}@uic.edu. This work was partially supported by AFOSR under award FA9550-10-1-0239. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the AFOSR.

where  $\mathbf{v}_1 := [v_1(0), v_1(1), \dots, v_1(N-1)]^T$ , is the additive noise and  $\alpha_1 = [\alpha_{11}, \alpha_{12}, \dots, \alpha_{1M}]^T \in \mathbb{C}^M$  is the sampled channel impulse response.

During the second stage, the TR system transmitter then transmits the scaled, by  $\kappa = \frac{\|\mathbf{s}\|}{\|\bar{\mathbf{y}}_1\|}$ , time-reversed channel output  $\bar{\mathbf{y}}_1 := [y_1^*(N-1), y_1^*(N-2), \dots, y_1^*(0)]^T$ . The convolution matrix of the time-reversed output  $\mathbf{y}_1$  is denoted by  $\bar{\mathbf{Y}}_1 \in \mathbb{C}^{N \times M}$ . In the meantime the channel has changed and the new impulse response is given by  $\alpha_2 \in \mathbb{C}^M$  and the new noise by  $\mathbf{v}_2$ , which need not necessarily follow the same statistics as in the first stage. The “conventional” channel model, which we will use as comparison point in evaluating the utility of TR, transmits the same waveform  $\mathbf{s}$  during the second stage. Thus, the second stage outputs of the TR and conventional system, are given, respectively, by

$$\mathbf{y}_2^{\text{TR}} := \kappa \bar{\mathbf{Y}}_1 \alpha_2 + \mathbf{v}_2 \quad (\text{TR}) \quad (2)$$

$$\mathbf{y}_2^{\text{nTR}} := \mathbf{S} \alpha_2 + \mathbf{v}_2 \quad (\text{No TR}). \quad (3)$$

For analytical tractability, we assume that  $\alpha := [\alpha_1^T, \alpha_2^T]^T$  is jointly Gaussian with mean and covariance, respectively,

$$\mu := [\mu_1^T, \mu_2^T]^T, \quad \mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_{12} \\ \mathbf{C}_{12}^H & \mathbf{C}_2 \end{bmatrix}. \quad (4)$$

In (4),  $\mathbf{C}_i = \text{Cov}\{\alpha_i, \alpha_i\}$ ,  $i = 1, 2$  and  $\mathbf{C}_{12} = \text{Cov}\{\alpha_1, \alpha_2\}$ , where  $\text{Cov}\{\mathbf{x}, \mathbf{y}\}$  is the covariance (matrix) between vectors  $\mathbf{x}$  and  $\mathbf{y}$ . The concatenated noise vector,  $\mathbf{v} = [\mathbf{v}_1^T, \mathbf{v}_2^T]^T$  is independent of  $\alpha$  and is assumed to be zero-mean, jointly Gaussian with corresponding covariance matrix, given by,

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_{12} \\ \mathbf{V}_{12}^H & \mathbf{V}_2 \end{bmatrix}. \quad (5)$$

where  $\mathbf{V}_i = \text{Cov}\{\mathbf{v}_i, \mathbf{v}_i\}$ ,  $i = 1, 2$  and  $\mathbf{V}_{12} = \text{Cov}\{\mathbf{v}_1, \mathbf{v}_2\}$ .

### III. MUTUAL INFORMATION AS COMPARISON METRIC

In this work, we quantify the utility of TR channels by using the mutual information between the appropriate channel input/output quantities as our comparison metric. In radar channels, information theoretic metrics such as mutual information and conditional entropy have been used for a variety of purposes including waveform design [12], [13], waveform scheduling [8], [9], and sensor management [14], [15]. It is often motivated as a “surrogate” metric [7] because of its generic ability to quantify the information gained by certain measurements, which is not linked to a specific task such as detection or estimation (though links between mutual information and SNR or Fisher information may be made).

Over the two stages of transmission, the radar system wishes to learn about the channel  $\alpha$  from the received signals  $\mathbf{y} = [\mathbf{y}_1^T \ \mathbf{y}_2^{\text{TR}T}]^T$  (TR) or  $\mathbf{y} = [\mathbf{y}_1^T \ \mathbf{y}_2^{\text{nTR}T}]^T$  (conventional), given knowledge of the transmitted waveforms  $\mathbf{s}$ . The amount of information we can gain about  $\alpha$  from  $\mathbf{y} := [\mathbf{y}_1 \ \mathbf{y}_2]^T$  given knowledge of the transmitted waveform  $\mathbf{s}$  is given by the mutual information (MI) between  $\alpha$  and  $\mathbf{y}$ , where we note that  $\mathbf{s}$  is a known parameter, denoted by

$$I(\alpha; \mathbf{y}) := \mathbb{E}_{p(\alpha, \mathbf{y})} \left\{ \ln \left( \frac{p(\alpha, \mathbf{y})}{p(\alpha)p(\mathbf{y})} \right) \right\}, \text{ or equivalently}$$

$$I(\alpha; \mathbf{y}) := h(\alpha) - h(\alpha|\mathbf{y})$$

where  $h(\cdot)$  is defined as the conditional differential entropy, and  $p(\cdot|\cdot)$  is the conditional pdf, and  $\mathbb{E}$  denotes the expectation operator. We recall the definitions of differential entropy [16]:  $h(\alpha) := -\mathbb{E}_{p(\alpha)}\{\ln(p(\alpha))\}$  and  $h(\alpha|\mathbf{y}) = -\mathbb{E}_{p(\alpha, \mathbf{y})}\{\ln(p(\alpha|\mathbf{y}))\}$ . For our TR channel under Gaussian assumptions, we will be evaluating mutual information terms for which the following is useful [16]:

*Fact 1.* If  $\mathbf{z}$  is a multivariate Gaussian random vector with an arbitrary mean vector and a covariance matrix,  $\mathbf{C}_z$ , then the differential entropy (nats),  $h(\mathbf{z}) = \ln \det(\mathbf{C}_z) + \eta$ , where  $\eta$  is a constant and is related to the dimensions of  $\mathbf{z}$ .

*Remark 1. Time-Reversal: A channel with feedback or not?*

In a mono-static radar system which employs TR, the next waveform transmitted depends on the previously received data. Intuitively at least, TR appears analogous to an information theoretic channel where the encoder employs feedback (i.e. encoders at time  $i$  have access to previous channel outputs  $y_1, \dots, y_{i-1}$  and may let their subsequent channel inputs be functions of these outputs, i.e.  $x_i = f(m, y_1, \dots, y_{i-1})$  where  $m$  is the message). In our scenario, the time-reversed transmitted waveform appear to be just that – a specific function  $f(\cdot)$  of previously received outputs and hence may appear to be a feedback channel. For feedback channels, *directed information* (DI) between the inputs and outputs, rather than mutual information between them, is a more relevant metric (in terms of channel capacity) [17], [18]. In the context of the TR channel the relevant DI would be defined as  $I(\alpha \rightarrow \mathbf{y} | \mathbf{s}, \bar{\mathbf{y}}_1) := I(\alpha_1; \mathbf{y}_1) + I(\alpha; \mathbf{y}_2 | \mathbf{y}_1)$ . Interestingly, for our Gaussian channel model, the DI and MI may be shown to be equal (see [19]), which at first is somewhat surprising as often in channels where feedback is employed, DI is strictly less than MI. This may be explained by the fact that we are interested in the MI between the *channel impulse response*  $\alpha$  and the output  $\mathbf{y}$ , and the channel impulse response does NOT employ feedback. Hence, because of the problem’s application in learning about the channel rather than the inputs, what might appear to be a feedback channel does not result in different DI and MI.

### IV. COMPARING TR AND CONVENTIONAL CHANNELS

We compare the TR and conventional channels for time-varying channels using the difference between the either the DI or MI (as they are equal) as a metric. These are generally analytically intractable and we will use Monte Carlo simulations in Section V to evaluate the metric. However, to build intuition, we do present analytical results for a simple scalar problem which illustrates the impact of non-stationary channels on TR.

We consider the difference between the mutual information achieved by the TR and the conventional channel:

$$\Upsilon := I(\alpha_1, \alpha_2; \mathbf{y}_1, \mathbf{y}_2^{\text{TR}}) - I(\alpha_1, \alpha_2; \mathbf{y}_1, \mathbf{y}_2^{\text{nTR}}) \quad (6)$$

$$= I(\alpha; \mathbf{y}_2^{\text{TR}} | \mathbf{y}_1) - I(\alpha; \mathbf{y}_2^{\text{nTR}} | \mathbf{y}_1), \quad (7)$$

where (7) follows by the chain rule for mutual information [16], and the fact that  $\mathbf{y}_1$  is the same regardless of whether TR is used or not. If  $\Upsilon > 0$ , we may conclude that TR yields more information about the channel  $\alpha$  than the conventional channel, and vice-versa if  $\Upsilon < 0$ . We note that our derivation thus far has been for Gaussian, non-stationary environments, i.e. involving colored noise and channels  $\alpha$  which are correlated over the two time-instances.

The two terms in (7) may be evaluated as in (8) – (9) at the top of the next page. Although not explicitly shown, it is noted that  $\kappa$  is a function of  $\mathbf{y}_1$  and stays inside the expectation operation. From (2), note that the pdf of  $\mathbf{y}_2^{\text{TR}}$  given a  $\mathbf{y}_1$  is normally distributed with a mean given by  $\kappa \bar{\mathbf{Y}}_1 \mathbb{E}\{\alpha_2\}$  and covariance matrix  $\kappa^2 \bar{\mathbf{Y}}_1 \mathbf{C}_2 \bar{\mathbf{Y}}_1^H + \mathbf{V}_2$ , which yields (8). Unfortunately a closed form solution to (8) is not immediate, and hence Monte Carlo simulations are employed here. Similar Monte Carlo analysis was employed in [2], [3] but for comparing the detection performance of TR systems with its non TR counterparts. We now consider a simple scalar channel where we are able to evaluate the metric analytically.

#### A. Scalar special case

Consider the following scalar channel model:

$$y_1 = s\alpha_1 + v_1, \quad y_2^{\text{TR}} = \kappa y_1^* \alpha_2 + v_2, \quad y_2^{\text{NTR}} = s\alpha_2 + v_2$$

Assume for simplicity that the noise covariance matrix,  $\mathbf{V} = \sigma^2 \mathbf{I}$  and that the channel has covariance matrix,  $\mathbf{C} = \begin{bmatrix} \sigma_\alpha^2 & \sigma_\beta^2 \\ \sigma_\beta^2 & \sigma_\alpha^2 \end{bmatrix} \succ \mathbf{0}$ . For the scalar case  $\kappa = \sqrt{|s|^2/|y_1|^2}$ . Then,

$$\Upsilon = \mathbb{E}_{y_1} \{ \ln(\kappa^2 |y_1|^2 \sigma_\alpha^2 + \sigma^2) \} \quad (10)$$

$$- \ln\{(|s|^2 \sigma_\alpha^2 + \sigma^2) - (|s|^4 \sigma_\beta^4) / (|s|^2 \sigma_\alpha^2 + \sigma^2)\} \\ = -\ln(1 - \rho^2) \quad (11)$$

where  $\rho^2 = |s|^4 \sigma_\beta^4 / (|s|^2 \sigma_\alpha^2 + \sigma^2)^2$  is the square of the correlation coefficient between  $y_2^{\text{NTR}}$  and  $y_1$  for the conventional channel. Now from (10)  $\Upsilon$  is always positive for  $\rho^2 > 0$ . In other words, as long as the channel is correlated,  $\Upsilon > 0$ , which implies that TR is preferable to using the channel conventionally in terms of mutual information. Similar to  $\rho$ , it is envisioned that the matrix  $(\mathbf{S}\mathbf{C}_{12}\mathbf{S}^H + \mathbf{V}_{12})^H(\mathbf{S}\mathbf{C}_1\mathbf{S}^H + \mathbf{V}_1)^{-1}(\mathbf{S}\mathbf{C}_{12}\mathbf{S}^H + \mathbf{V}_{12})$  plays an analogous and dominant part in  $\Upsilon$  for vector channels.

#### V. SIMULATIONS

Two channels, termed  $\mathcal{A}, \mathcal{B}$  are considered in the simulations which differ in the channel covariances  $\mathbf{C}$ . The first is the simplest and is given by  $\mathbf{C} := \mathbf{C}_\mathcal{A} = \begin{bmatrix} \mathbf{I} & \rho_a \mathbf{I} \\ \rho_a \mathbf{I} & \mathbf{I} \end{bmatrix}$ . This implies that the  $M$  taps in  $\alpha_i = [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{iM}]^T, i = 1, 2$  are uncorrelated within each stage but are correlated across stages, depending on the values assumed by  $\rho_a$ . In channel  $\mathcal{B}$  the covariance matrix is  $\mathbf{C} := \mathbf{C}_\mathcal{B}$  that is Hermitian Toeplitz, and has a covariance function  $r_B(|m_1 - m_2|) = \exp(-\sigma_b^2 |m_1 - m_2|), (m_1, m_2) \in \{1, 2, \dots, 2M\}$ , correlating channel coefficients within and between stages, depending on  $\sigma_b^2$ . To purely analyze the effects of the channel, we assume white noise, i.e.  $\mathbf{V}_{12} = 0, \mathbf{V}_i = \mathbf{I}, i = 1, 2$ . Two waveforms for the transmitted  $s$  are analyzed, the first is a BPSK symbol

waveform comprising random  $\pm 1$ , the other is a radar chirp waveform. The analysis is carried out in baseband. We will assume  $M = 10$  for both  $\mathcal{A}, \mathcal{B}$ . The SNR is defined as  $|s|^2 M / \sigma^2$ . The number of Monte Carlo trials were set at 10,000 to evaluate the expectation operation.

In Fig.1, the value of  $\Upsilon$  versus  $\rho_a$  are shown for the BPSK and chirp waveforms for  $\mathcal{A}$  at SNR=0, 10, 20dBs. In Fig. 1(a), for SNR=10dB, and SNR=20dB, we see that the  $\Upsilon$  is positive for high correlation ( $\rho_a \in \pm(1, 0.5]$ ), indicating superior performance of the TR when compared to the conventional channel. For medium to low correlations, and not surprisingly, the opposite is true, i.e.  $\Upsilon$  becomes negative indicating that TR is not preferable when compared to using the channel conventionally. In particular, we see that for SNR=20dB, and for medium and low correlation, the metric assumes low values. Similar results are seen for the chirp waveform in Fig. 1(b). The break even points for Fig. 1(a) and Fig. 1(b), i.e.  $\Upsilon = 0$  are different for the same SNR, hence a waveform dependency is also noted. The processing for the chirp was performed in the baseband bandwidth [2] which contains 99% of the signal energy. For the implementation, spectral content outside the band was notched, and an inverse FFT (IFFT) was employed to return to the time domain. Such frequency domain processing is not required for the BPSK, as it is wideband.

In Fig.2, the results are shown for the BPSK waveform and the chirp for  $\mathcal{B}$ . The BPSK waveforms were different in Fig. 1(a) and Fig. 2(a). In Fig. 2 the parameter  $\sigma_b^2$  now controls the correlation. As  $\sigma_b^2$  increases, the channel coefficients start becoming uncorrelated. Identical conclusions to Fig. 1 may be drawn by observing Fig.2. For example and as before, for reasonable SNR and low correlation scenarios, the metric assumes low values, implying a harsh penalty for using the TR rather than using the channel conventionally.

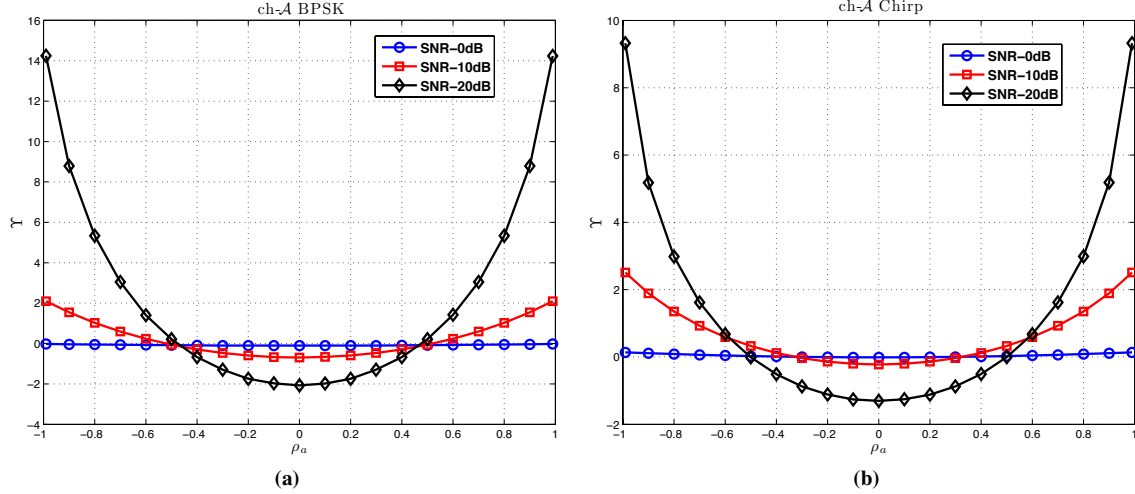
It is stressed that the metric  $\Upsilon$  evaluates the TR and conventional channel on the “average”. In other words, for low to medium correlation, we have seen instances for both  $\mathcal{A}, \mathcal{B}$ , where the difference between the DI between the TR and the conventional channels are actually positive, whereas on an average it is negative, i.e.  $\Upsilon < 0$ .

#### REFERENCES

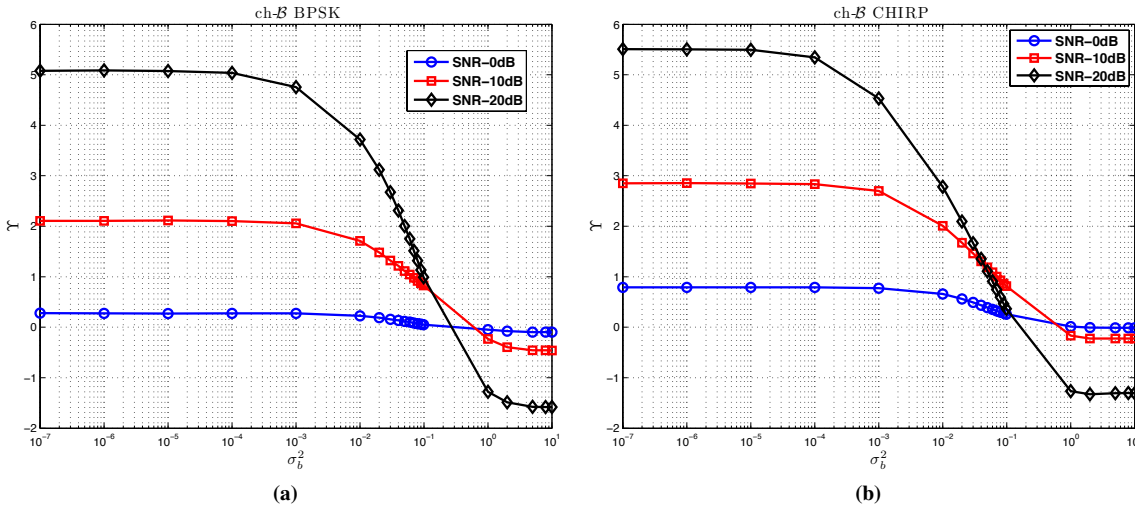
- [1] M. Fink, “Time reversal of ultrasonic fields I. basic principles,” *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 39, no. 5, pp. 555–566, Sep. 1992.
- [2] J. M. F. Moura and Y. Jin, “Detection by time reversal: Single antenna,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 1, pp. 187–201, Jan. 2007.
- [3] Y. Jin and J. M. F. Moura, “Time-reversal detection using antenna arrays,” *Trans. Sig. Proc.*, vol. 57, no. 4, pp. 1396–1414, Apr. 2009. [Online]. Available: <http://dx.doi.org/10.1109/TSP.2008.2010425>
- [4] W. A. Kuperman, W. S. Hodgkiss, H. C. Song, T. Akal, C. Ferla, and D. R. Jackson, “Phase conjugation in the ocean: Experimental demonstration of an acoustic time-reversal mirror,” *Journal of The Acoustical Society of America*, vol. 103, 1998.
- [5] B. I. Zeldovich, N. F. Pilipetskii, and V. V. Shkunov, *Principles of phase conjugation*. Springer-Verlag, 1985.
- [6] S. K. Lehman and A. J. Devaney, “Transmission mode time-reversal super-resolution imaging,” *J Acoust Soc Am*, vol. 113, no. 5, pp. 2742–53, 2003.
- [7] D. Cochran, S. Suvorova, S. Howard, and W. Moran, “Waveform libraries: Measures of effectiveness for radar scheduling,” *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 12–21, 2009.

$$I(\alpha; \mathbf{y}_2^{\text{TR}} | \mathbf{y}_1) = h(\mathbf{y}_2^{\text{TR}} | \mathbf{y}_1) - h(\mathbf{y}_2^{\text{TR}} | \mathbf{y}_1, \alpha_1, \alpha_2) = \mathbb{E}_{\mathbf{y}_1} \{ \ln \det(\kappa^2 \tilde{\mathbf{Y}}_1 \mathbf{C}_2 \tilde{\mathbf{Y}}_1^H + \mathbf{V}_2) \} - \ln \det(\mathbf{V}_2) \quad (8)$$

$$\begin{aligned} I(\alpha; \mathbf{y}_2^{\text{nTR}} | \mathbf{y}_1) &= h(\mathbf{y}_2^{\text{nTR}} | \mathbf{y}_1) - h(\mathbf{y}_2^{\text{nTR}} | \mathbf{y}_1, \alpha_1, \alpha_2) \\ &= \ln \det \left( \mathbf{S} \mathbf{C}_2 \mathbf{S}^H + \mathbf{V}_2 - (\mathbf{S} \mathbf{C}_{12} \mathbf{S}^H + \mathbf{V}_{12})^H (\mathbf{S} \mathbf{C}_1 \mathbf{S}^H + \mathbf{V}_1)^{-1} (\mathbf{S} \mathbf{C}_{12} \mathbf{S}^H + \mathbf{V}_{12}) \right) - \ln \det(\mathbf{V}_2) \end{aligned} \quad (9)$$



**Fig. 1:** For ch-A metric  $\Upsilon$  vs  $\rho_a$  for SNR={0,10,20}dBs (a) BPSK,  $N = 10$ , (b) Chirp waveform,  $N = 250$



**Fig. 2:** For ch-B metric  $\Upsilon$  vs  $\sigma_b^2$  for SNR={0,10,20}dBs (a) BPSK,  $N = 10$ , (b) Chirp waveform,  $N = 250$

- [8] W. Moran, S. Suvorova, and S. Howard, "Application of sensor scheduling concepts to radar," *Foundations and Applications of Sensor Management*, pp. 221–256, 2008.
- [9] P. Setlur and N. Devroye, "Adaptive waveform scheduling in radar: an information theoretic approach," pp. 836 103–836 103–11, 2012. [Online]. Available: + <http://dx.doi.org/10.1117/12.919213>
- [10] H. C. Song, "Time reversal communication in a time-varying sparse channel," *J Acoust Soc Am*, vol. 130, no. 4, p. EL161, 2011. [Online]. Available: <http://www.biomedsearch.com/nih/Time-reversal-communication-in-time/21974486.html>
- [11] I. Naqvi, P. Besnier, and G. Zein, "Robustness of a time-reversal ultra-wideband system in non-stationary channel environments," *IET Microwaves, Antennas and Propagation*, vol. 5, no. 4, pp. 468–475, 2011. [Online]. Available: <http://link.aip.org/link/?MAP/5/468/1>
- [12] M. Bell, "Information theory and radar waveform design," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1578–1597, 1993.
- [13] —, "Information theory and radar: mutual information and the design and analysis of radar waveforms and systems," Ph.D. dissertation, Caltech, 1988.
- [14] A. Hero, C. Kreucher, and D. Blatt, "Information theoretic approaches to sensor management," *Foundations and Applications of Sensor Management*, pp. 33–57, 2008.
- [15] J. Williams, "Information theoretic sensor management," Ph.D. dissertation, Massachusetts Institute of Technology, 2007.
- [16] T. Cover and J. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.
- [17] J. Massey, "Causality, feedback and directed information," in *Proc. IEEE Int. Symp. Inform. Theory and its Applications*, Honolulu, Nov. 1990.
- [18] G. Kramer, "Directed information for channels with feedback," Ph.D. dissertation, Swiss Federal Institute of Technology Zurich, 1998.
- [19] P. Setlur and N. Devroye, "On the mutual information of time reversal for non-stationary channels," in *ICASSP 2013*, In Preparation.