

Spatio-Temporal Focusing for Elimination of Multipath Effects in High Rate Acoustic Communications

Milica Stojanovic

Massachusetts Institute of Technology, MIT E38-376, Cambridge, MA 02139

Abstract. High rate underwater communications have traditionally relied on equalization methods to overcome the intersymbol interference (ISI) caused by multipath propagation. An alternative technique has emerged in the form of time-reversal, which comes at virtually no cost in computational complexity, but sacrifices the data rate and relies on large arrays to reduce ISI. In this paper, optimal multipath suppression using spatio-temporal processing is addressed analytically. A communication link between a single element and an array is considered in several scenarios: uplink and downlink transmission, with and without channel state information and varying implementation complexity. Transmit/receive techniques are designed which simultaneously maximize the data detection SNR and minimize the residual ISI, while maintaining maximal data rate in a given bandwidth and satisfying a constraint on transmitted energy. The performance of various techniques is compared on a shallow water channel operating in the 15 kHz band. Results demonstrate benefits of optimal focusing whose performance is not conditioned on the array size.

INTRODUCTION

High rate acoustic communications have traditionally relied on adaptive equalization methods to overcome the intersymbol interference (ISI) caused by multipath propagation. Excellent performance of these receivers comes at a price of high computational complexity [1]. In standard equalization, all of the signal processing is performed at the receiver, while the transmitter uses standard signaling waveforms. A different, and possibly better approach is to split the signal processing between the transmitter and receiver. Such an approach forms the basis of spatio-temporal focusing.

In its simplest form, focusing is achieved by transmitting a time-reversed (or equivalently, phase-conjugated in the frequency domain) replica of a probe signal received earlier from the source location. This technique has been used for medical imaging, therapy, and material testing, while recent research has hailed it as a method that can replace traditional equalization, and eliminate the associated computational burden (e.g., [2]-[5]). Several research groups have been involved in application of time-reversal (TR) arrays to undersea acoustic communications, addressing active phase-conjugation for two-way communication [2], as well as passive phase-conjugation for one-way communication from a point source to an array [3]. These groups have been engaged in experimental work, emphasizing low-complexity processing using TR only. In parallel, analytical work addressed the use of adaptive channel estimation and low complexity equalization in conjunction with TR [4]. The common goal of these efforts was to eliminate ISI by

use of TR. Ensuring ISI-free transmission in a system that has multiple transmit/receive elements is a major asset in a channel whose bandwidth is severely limited. In particular, it lays ground for the exploitation of bandwidth-efficiency improvement available from multi-input multi-output (MIMO) signal processing [6].

A statement commonly encountered in the relevant literature is that TR “undoes” the effects of multipath. However, as the experimental results have shown, suppression of multipath effects by TR is achieved at the expense of reduced data throughput and/or the need for a large array, a consequence of relying on TR to eliminate ISI. TR recombines multipath energy in a manner of matched filtering, whose function is to maximize the SNR, and not to eliminate ISI. In fact, matched filtering increases temporal dispersion of the signal, and in a communication system where a sequence of pulses is transmitted at high rate, it must be followed by a sequence estimator or an equalizer. Increasing the number of elements in a TR array only helps to reduce residual ISI, but it does not eliminate it. The use of equalization in conjunction with TR may become necessary, but the advantage of this approach to standard equalization is not apparent. The use of spatio-temporal focusing for complete suppression of multipath effects thus remains an open question.

In this paper, a solution is proposed to the following problem: If the channel responses between a single element and an array are known, determine the optimal transmit/receive technique that the two can use to simultaneously (1) eliminate ISI *and* (2) maximize SNR, while maintaining *maximal* data rate in a given bandwidth and satisfying a constraint on transmitted energy. Note that because it allows for transmitter as well as receiver optimization, the solution differs from standard equalization. Also, because it explicitly requires minimization of ISI, it differs from TR. The resulting system does not depend on the number of array elements to minimize the multipath distortion, but instead provides an answer for a variety of applications that cannot afford large arrays. Optimal configurations are intended as a basis for adaptive system implementation in which channel estimates will replace the unknown, time-varying responses.

SYSTEM OPTIMIZATION

System optimization is addressed for uplink and downlink communication (to/from array), as shown in Fig.1. Performance is assessed using SNR as the figure of merit, and compared to that of TR, standard linear equalization, and TR in conjunction with equalization.

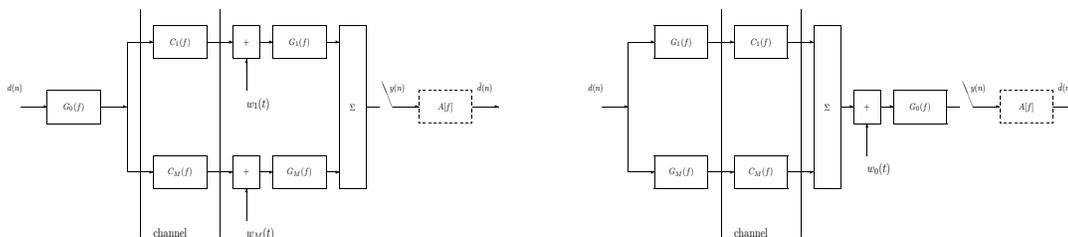


FIGURE 1. Uplink (left) and downlink (right) transmission. An equalizer may or may not be used.

Transmitter/receiver optimization for no ISI. The data sequence $d(n)$ is transmitted at symbol rate $1/T$. The problem is to find transmit/receive filters $G_0(f)$ and $G_1(f), \dots, G_M(f)$ such that the SNR at the receiver is maximized, subject to the constraint that there is no ISI in the decision variables $\hat{d}(n) = y(nT)$, and that finite transmitted energy per bit E is used. The channel responses $C_m(f), m = 1 \dots M$ and the power spectral density $S_w(f)$ of the uncorrelated noise processes $w_m(t), m = 0 \dots M$ are assumed to be known.

The composite channel response is denoted by $F(f) = G_0(f) \sum_{m=1}^M G_m(f) C_m(f)$. The received signal after filtering is then given by $y(t) = \sum_n d(n) f(t - nT) + z(t)$, where the noise $z(t)$ has power spectral density

$$S_z(f) = S_w(f) \sum_{m=1}^M |G_m^2(f)| \text{ for uplink, and } S_z(f) = S_w(f) |G_0^2(f)| \text{ for downlink.} \quad (1)$$

The requirement for no ISI is expressed as $F(f) = X(f)$, where $X(f)$ is Nyquist, i.e., bandlimited to $\pm 1/T$ and its waveform $x(t)$ satisfies $x(nT) = x_0 \delta_{n,0}$. $X(f)$ can be chosen as raised cosine, and without loss of generality we take that $X(f) = |X(f)|$. When there is no ISI, the sampled received signal is $y(nT) = d(n)x_0 + z(nT)$, and $SNR = \sigma_d^2 x_0^2 / \sigma_z^2$, where $\sigma_d^2 = E\{|d^2(n)|\}$ and $\sigma_z^2 = \int_{-\infty}^{+\infty} S_z(f) df$. The energy constraint is expressed as

$$E = \sigma_d^2 \int_{-\infty}^{+\infty} |G_0^2(f)| df \text{ for uplink, and } E = \sigma_d^2 \sum_{m=1}^M \int_{-\infty}^{+\infty} |G_m^2(f)| df \text{ for downlink.} \quad (2)$$

Let us first consider uplink transmission. Taking into account the no ISI requirement and the energy constraint, the SNR is expressed as

$$SNR = E x_0^2 \left[\int_{-\infty}^{+\infty} \frac{X^2(f)}{|\sum_m G_m(f) C_m(f)|^2} df \int_{-\infty}^{+\infty} S_w(f) \sum_{m=1}^M |G_m^2(f)| df \right]^{-1}. \quad (3)$$

This function is to be maximized with respect to the receive filters $G_m(f)$. To do so, we use a two-step procedure, each involving one Schwarz inequality. The first inequality states that

$$\left| \sum_{m=1}^M G_m(f) C_m(f) \right|^2 \leq \gamma(f) \sum_{m=1}^M |G_m^2(f)|; \quad \gamma(f) = \sum_{m=1}^M |C_m^2(f)| \quad (4)$$

where the equality holds for

$$G_m(f) = \alpha(f) C_m^*(f). \quad (5)$$

We note similarly with TR in that receiving filters should be proportional to the phase-conjugate of the channel transfer functions. However, there is room for additional improvement through optimization of the function $\alpha(f)$. Applying a second Schwarz inequality to the denominator of the SNR bound yields:

$$\int_{-\infty}^{+\infty} \frac{X^2(f)}{\gamma(f) \sum_m |G_m^2(f)|} df \int_{-\infty}^{+\infty} S_w(f) \sum_{m=1}^M |G_m^2(f)| df \geq \left[\int_{-\infty}^{+\infty} \frac{X(f)}{\sqrt{\gamma(f)}} \sqrt{S_w(f)} df \right]^2 \quad (6)$$

where the equality holds for

$$X(f)/\sqrt{\gamma(f)S_w(f)} = \beta \sum_{m=1}^M |G_m^2(f)| \quad (7)$$

and β is a constant. Combining the two conditions (5) and (7) we obtain the optimal value of $\alpha(f)$, which determines the the transmit filter $G_0(f) = X(f)/[\alpha(f)\gamma(f)]$, and the constant β then follows from the energy constraint (2). The desired solution is

$$G_0(f) = K(f)\sqrt{X(f)}\gamma^{-1/4}(f); G_m(f) = K^{-1}(f)\sqrt{X(f)}\gamma^{-3/4}(f)C_m^*(f), m = 1 \dots M$$

where $K(f) = \sqrt{E/\sigma_d^2} \left[\int_{-\infty}^{+\infty} S_w^{1/2}(f)X(f)\gamma^{-1/2}(f)df \right]^{-1/2} S_w^{1/4}(f)$. (8)

This selection of filters achieves maximal SNR,

$$SNR_2 = Ex_0^2 \left[\int_{-\infty}^{+\infty} \sqrt{S_w(f)} \frac{X(f)}{\sqrt{\gamma(f)}} df \right]^{-2} \quad (9)$$

where index ‘2’ indicates that both sides of the link adjust their filters in accordance with the channel. Optimization in the downlink case gives filters in identical form, except that the factor $K(f)$ is the reciprocal of that given in (8). The same maximal SNR is achieved.

We now turn to the situation in which one side of the link is constrained to have minimal complexity, such as when limited processing power is available at the single-element end. The no ISI condition still must hold, $F(f) = X(f)$, but the filter $G_0(f)$ may no longer be a function of the channel responses. A similar optimization procedure results in the following solution:

$$\text{uplink: } G_0(f) = K\sqrt{X(f)}; G_m(f) = K^{-1}\sqrt{X(f)}\gamma^{-1}(f)C_m^*(f), m = 1 \dots M$$

where $K = \sqrt{E/(\sigma_d^2 x_0)}$ (10)

$$\text{downlink: } G_0(f) = K^{-1}(f)\sqrt{X(f)}; G_m(f) = K(f)\sqrt{X(f)}\gamma^{-1}(f)C_m^*(f), m = 1 \dots M$$

where $K(f) = \sqrt{E/\sigma_d^2} \left[\int_{-\infty}^{+\infty} S_w(f)X(f)\gamma^{-1}(f)df \right]^{-1/2} S_w^{1/2}(f)$. (11)

This selection of filters achieves maximal SNR available with one-side adjustment,

$$SNR_1 = Ex_0 \left[\int_{-\infty}^{+\infty} S_w(f) \frac{X(f)}{\gamma(f)} df \right]^{-1}. \quad (12)$$

Comparing the SNR available with and without complexity constraint, we find that $SNR_1 \leq SNR_2$. The two SNRs are equal only when $\gamma(f)$ is proportional to $S_w(f)$. In what follows, we shall focus on the usual case of white noise, $S_w(f) = N_0$. Note that $K(f)$ then becomes a constant, and the same set of filters may be used for uplink and downlink transmission.

TR performance with residual ISI. When no care is taken to ensure focusing, samples of the received signal, $y(nT) = \sum_k f(kT)d(n-k) + z(nT)$, contain residual ISI. Assuming uncorrelated data symbols, the SNR is given by $SNR_0 = \sigma_d^2 |f^2(0)| / [\sigma_d^2 \sum_{k \neq 0} |f^2(kT)| + \sigma_z^2]$. We consider the following situation. On the uplink, the transmitter uses $G_0(f) = K_u \sqrt{X(f)}$, and the receiver uses $G_m(f) = G_0^*(f)C_m^*(f)$. This scenario is analogous to ideal (noiseless) passive phase-conjugation. On the downlink, the transmitter uses $G_m(f) = K_d \sqrt{X(f)}C_m^*(f)$. This scenario is analogous to active phase-conjugation. The receiver filter is $G_0(f) = \sqrt{X(f)}$. The constants K_u, K_d are determined from the energy constraint (2). The SNR in either case is obtained as

$$SNR_0 = \frac{E/N_0}{\rho E/N_0 + \frac{x_0}{\int_{-\infty}^{+\infty} X(f)\gamma(f)df}}; \rho = \frac{\sum_{k \neq 0} |f^2(kT)|}{|f^2(0)|} = \frac{\int_{-1/2T}^{+1/2T} |X\gamma[f]|^2 df T}{\int_{-\infty}^{+\infty} X(f)\gamma(f)df} - 1 \quad (13)$$

where $X\gamma[f]$ denotes the folded spectrum of $X(f)\gamma(f)$, $X\gamma[f] = \frac{1}{T} \sum_k X(f + \frac{k}{T})\gamma(f + \frac{k}{T})$. We note that as the noise vanishes, i.e. $E/N_0 \rightarrow +\infty$, unlike with optimal focusing where $SNR_{1,2} \rightarrow +\infty$, the performance of TR saturates, $SNR_0 \rightarrow 1/\rho$. The value of ρ depends on the channel through the function $\gamma(f)$, and on the system bandwidth through $X(f)$.

TR performance with equalization. The performance of TR saturates because of residual ISI. To overcome this limitation, an equalizer may be used. An optimal (MMSE) linear processor used on the downlink consists of the matched filter, $G_0(f) = [\sum_{m=1}^M G_m(f)C_m(f)]^*$, and a symbol-spaced equalizer with transfer function $A[f] = \sigma_d^2 F^*[f] / (\sigma_d^2 |F[f]|^2 + S_z[f])$, where $F[f]$ is the folded spectrum of the overall response $F(f) = |G_0^2(f)|$, and $S_z[f] = N_0 F[f]$ is the power spectral density of the discrete-time noise process $z(nT)$. The SNR at the equalizer output is

$$SNR = \left[\int_{-1/2T}^{+1/2T} \frac{df T}{1 + (\sigma_d^2/N_0)F[f]} \right]^{-1} - 1. \quad (14)$$

For the transmit filter selection as in active phase-conjugation, $G_m(f) = K_d \sqrt{X(f)}C_m^*(f)$, the SNR is

$$SNR_{3,tr} = \left[\int_{-1/2T}^{+1/2T} \frac{df T}{1 + \frac{E/N_0}{\int_{-\infty}^{+\infty} X(f)\gamma(f)df} X\gamma^2[f]} \right]^{-1} - 1 \quad (15)$$

where $X\gamma^2[f]$ is the folded spectrum of $X(f)\gamma^2(f)$.

Equalizer performance. A standard equalizer does not rely on TR at the transmitter, but instead uses pre-determined, channel-independent filters $G_m(f) = K_d \sqrt{X(f)}$. The SNR is computed from (14) as

$$SNR_{3,down} = \left[\int_{-1/2T}^{+1/2T} \frac{df T}{1 + \frac{E/N_0}{Mx_0} X\Sigma^2[f]} \right]^{-1} - 1 \quad (16)$$

where $X\Sigma^2[f]$ is the folded spectrum of $X(f)|\sum_m C_m(f)|^2$.

In the uplink scenario, the MMSE linear processor consists of a bank of matched filters, $G_m(f) = G_0^*(f)C_m^*(f)$ as in passive phase-conjugation, followed by the equalizer whose optimal transfer function now depends on $F(f) = |G_0^2(f)|\gamma(f)$. For the standard transmit filter selection, $G_0(f) = K_u\sqrt{X(f)}$, the SNR (14) becomes

$$SNR_{3,up} = \left[\int_{-1/2T}^{+1/2T} \frac{dfT}{1 + \frac{E/N_0}{x_0} X \gamma[f]} \right]^{-1} - 1. \quad (17)$$

Comparing uplink and downlink equalization, we have that $SNR_{3,up} \geq SNR_{3,down}$. The two are equal if the M channel transfer functions $C_m(f)$ are identical and constant within the signal bandwidth. It is not clear, however, how $SNR_{3,down}$ compares with $SNR_{3,tr}$, i.e. is there an advantage to using transmit TR in conjunction with equalization. This question gives rise to a broader one of optimal transmit filtering for equalization.

Equalizer performance with optimized transmit filter. If the requirement for no ISI is relaxed in the optimal system design, and an equalizer is used at the receiver, the question is what transmit/receive filtering should be used to maximize the SNR. Note that because this optimization criterion is less restrictive than that of optimal focusing (the no ISI constraint has been removed) improved performance may be expected.

Maximization of SNR (14) with respect to the transmit filter(s) is accomplished using the Lagrange method, yielding the following uplink/downlink solution when the system operates in minimal bandwidth $B = 1/T$ (see [7] for details):

$$SNR_4 = \left\{ 1 - \int_{B_L} dfT + \left[\int_{B_L} \frac{dfT}{\sqrt{\gamma(f)}} \right]^2 \left[\frac{E}{N_0} + \int_{B_L} \frac{dfT}{\gamma(f)} \right]^{-1} \right\}^{-1} - 1 \quad (18)$$

where $B_L = \{f : \gamma(f) \geq \gamma_L\}$, and γ_L is the smallest value of $\gamma(f)$ for which

$$\left[\frac{E}{N_0} + \int_{B_L} \frac{dfT}{\gamma(f)} \right] \left[\int_{B_L} \frac{dfT}{\sqrt{\gamma(f)}} \right]^{-1} \geq 1/\sqrt{\gamma_L}. \quad (19)$$

When $B_L = [-1/2T, 1/2T]$, it is easy to show that $SNR_4 = SNR_2 + SNR_2/SNR_1 - 1 \geq SNR_2$. Thus, this signaling scheme outperforms optimal focusing.

PERFORMANCE COMPARISON

To compare the performance of various techniques, a channel model based on the geometry of shallow water multipath is used. We look at repeated surface-bottom reflections and take into account P multipath arrivals, each characterized by a gain c_p , delay τ_p and angle of arrival θ_p . Nominal acoustic propagation loss that occurs for practical spreading at carrier frequency $f_c=15$ kHz is used to compute the gains. The channel transfer functions, observed at d -spaced elements $m = 1 \dots M$ are computed as

$$C_m(f) = \sum_{p=1}^P c_{m,p} e^{-j2\pi f \tau_p}, \text{ where } c_{m,p} = c_p e^{-j(m-1)\varphi_p} \text{ and } \varphi_p = 2\pi \frac{f_c}{c} d \sin \theta_p. \quad (20)$$

As an example, we use a channel of depth 75 m, range 3 km, and the system mounted near the bottom. The resulting multipath profile for $P = 3$ is shown in Fig.2. The channel function $\gamma(f)$ is shown for $M = 4$ and 32, together with the desired system response $X(f)$, chosen to provide maximal bit rate for ISI-free transmission in bandwidth $B = 1/T = 5$ kHz (10 kbps with 4-PSK, or 15 kbps with 8-PSK). The impulse response of the overall system obtained with TR is also shown, and is evidently far from ideal. As the number of array elements is increased, $\gamma(f)$ tends to flatten out, resulting in better, but not complete suppression of multipath through TR.

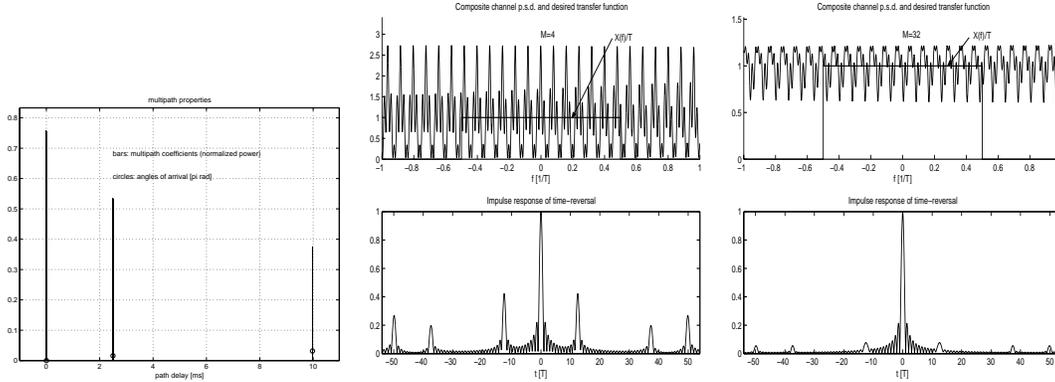


FIGURE 2. Multipath profile of the example channel (left). Composite channel power spectral density is $\gamma(f)$, and the impulse response of TR corresponding to $X(f)\gamma(f)$. Multipath coefficients are normalized such that $M \sum_p |c_p^2| = 1$, and half wavelength spacing between array elements is used.

Figure 3 summarizes performance results. Looking at $M = 4$ case, we first confirm that two-sided adjustment (9, solid ‘ \triangle ’) outperforms one-sided adjustment (12, dashed ‘ \triangle ’) in optimal focusing, but more interestingly, we observe that the difference in performance is small. This is an encouraging observation from the viewpoint of designing a practical system with restricted processing complexity. The performance of TR (13, dashed) is inferior to optimal focusing and to all other schemes at practical SNR, with loss becoming quite large even at a moderate E/N_0 of 10 dB - 15 dB. It saturates thereafter at a value $1/\rho$. Some of the loss is recovered by the use of equalizer in conjunction with TR (15, ‘+’); however, this system compares poorly with the standard equalizer (16, dashed ‘o’) as E/N_0 increases to more than a few dB. Finally, we confirm that equalization using optimized transmit filters (18, ‘*’) provides an upper bound on the performance of all other schemes. More importantly, we observe that this scheme offers negligible improvement over focusing, which allows a much easier implementation.

With $M = 32$, the performance of TR is improved; nonetheless, saturation is still notable. Equalization in conjunction with TR now outperforms standard equalization, while the performance of both focusing methods tends to the same optimal curve.

CONCLUSION

Because it ignores residual ISI, TR exhibits performance saturation, and strongly depends on the use of a large array. When this can be afforded (e.g. in a network whose base station uses a large array to spatially isolate multiple users) TR offers a solution for

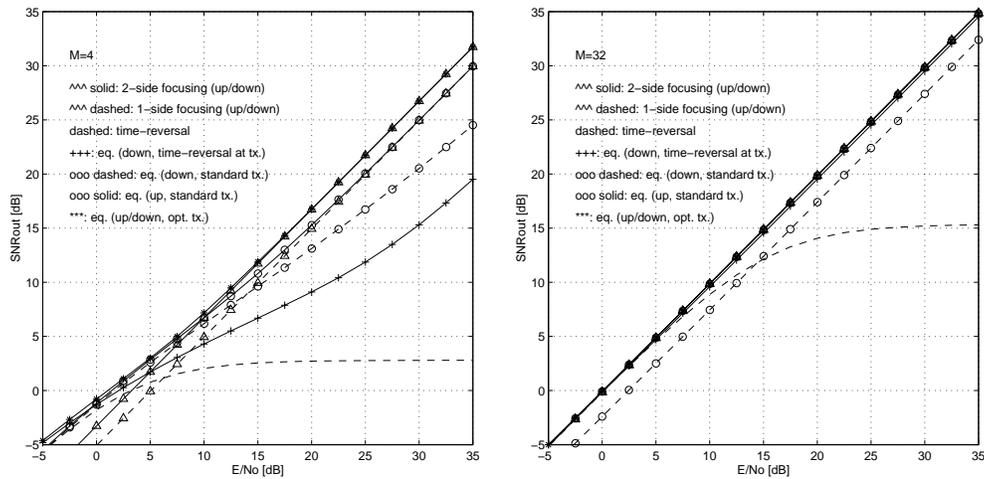


FIGURE 3. Performance of various techniques on the example channel.

minimal-complexity processing. With a smaller array, however, standard equalization outperforms TR, and the use of equalization in conjunction with TR does not guarantee performance improvement over standard equalization. Spatio-temporal focusing proposed in this paper guarantees maximal SNR and elimination of ISI for an arbitrary array size. It outperforms TR at the expense of additional filtering. If filter adjustment is constrained to the array side only, one-sided focusing offers an excellent trade-off between complexity and performance. It thus represents a solution for systems that cannot deploy large arrays and have limited processing power.

Future work will concentrate on experimental validation of spatio-temporal focusing aided by adaptive channel estimation. Two types of errors will guide the practical system performance: error due to noise and error due to time-variability of the channel. Analytical work will address system optimization with imperfect channel knowledge.

REFERENCES

1. M.Stojanovic, J.Catipovic and J.Proakis, "Reduced-complexity multichannel processing of underwater acoustic communication signals," *J. Acoust. Soc. Am.*, vol.98 (2), Pt.1, pp.961-972, Aug. 1995.
2. G.Edelmann, W.Hodgkiss, W.Kuperman and H.C.Song, "Underwater acoustic communication using time-reversal," in Proc. *IEEE Oceans'01 Conference*, Nov. 2001.
3. D.Rouseff et al., "Decision-directed passive phase-conjugation for underwater acoustic communication: experimental results," in Proc. *IEEE Oceans'02 Conference*, Oct. 2002.
4. J.Gomes and V.Barroso, "Time-reversed communication over Doppler-spread underwater channels," in Proc. *ICASSP'02 Conference*, pp.2849-2852(III), 2002.
5. T.C.Yang, "Temporal resolutions of time-reversal and passive phase-conjugation for underwater acoustic communications," *IEEE J. Oceanic Eng.*, vol.28, No.2, pp. 229-245, Apr. 2003.
6. D.Kilfoyle, J.Preisig and A.Baggeroer, "Spatial modulation over partially coherent multi-input / multi-output channels," *IEEE Trans. Sig. Proc.*, vol.51, No.3, pp.794-804, March 2003.
7. M.Stojanovic, "Retrofocusing techniques for high rate acoustic communications," MIT SG Internal Report, Jan. 2004. Available upon request (millitsa@mit.edu).