MULTI-SCALE MULTI-LAG CHANNEL ESTIMATION VIA LINEARIZATION OF TRAINING SIGNAL SPECTRUM AND SPARSE APPROXIMATION

Sajjad Beygi[⋆] *Urbashi Mitra*[⋆⋆] *Mariane R. Petraglia*[†]

[⋆] Department of Electrical Engineering, University of Southern California, Los Angeles CA 90089

[†]Department of Electronics Engineering, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Emails:{beygihar, ubli}@usc.edu, and mariane@pads.ufrj.br

ABSTRACT

Multi-scale, multi-lag (MSML) models are adopted for time-varying (ultra)-wideband channels that are relevant for underwater acoustic, radar and ultrawideband radio applications. MSML channels are characterized by a limited number of paths, each parameterized by a delay, Doppler scale, and attenuation factor. Herein, a novel MSML channel estimator is proposed. First, in the Fourier domain, it is shown that there is an approximately linear relationship between the received signal and the Doppler scales that enables the recasting of channel estimation into a convex optimization problem. Second, the inherent sparsity of many MSML channels is exploited resulting in a further improvement in estimation performance of about 5 dB in low SNR relative to an unstructured estimation method. Finally, the resultant estimation strategy has very low implementation complexity.

Index Terms— Multi-scale multi-lag, underwater communication, sparsity, Doppler scale.

1. INTRODUCTION

Modern communication applications have driven the interest in time-varying, wideband communications. In particular, underwater acoustic (UWA) communications are of interest in surveillance, environmental monitoring and tsunami detection; the effects of the low speed of sound in water and high relative mobility must be combatted. Reliable UWA communication necessitates accurate channel estimation. These wideband communication channels, such as UWA and radar, can be well-represented by multi-scale multi-lag channel (MSML) models [1–3]. We observe that Doppler distortion in wideband channels reveals itself as a time scaling of the transmitted signal [3]. For terrestrial radio channels, there is strong interest in the use of OFDM signaling, due to the fact that it decomposes a static frequency selective channel into a number of flat channels, enabling low complexity detection at the receiver. In UWA communication, the time scaling of the signal causes different subcarriers to be shifted by different frequencies, resulting in significant inter-carrier interference [4, 7] which further underscores the need for accurate channel estimation in order to enable the equalization of the resulting interference.

While prior art has designed channel estimation algorithms for MSML channels by assuming a single, dominant scale [8, 10], this assumption can lead to performance degradation. Another approach is to explicitly consider the multiple scales. The methods proposed in [5, 6, 9] adapt classical subspace methods from array signal processing to estimate the channel parameters including the scales. While offering good performance, these methods suffer from high computational complexity as they require multiple singular value decompositions or matrix inverses. In particular, our proposed method has complexity that is squared with respect to the filter length whereas our prior method in [5] has complexity that is cubic in this parameter.

In this paper, we show that an approximately linear relationship between the received signal samples in the Fourier domain and Doppler scale parameters holds for small scale values. Furthermore, considering the sparse nature of underwater acoustic channels (doubly selective multipath channels), we enforce a sparsity constraint to estimate the channel gains and delays resulting in better accuracy. Our algorithm requires only multiplications and thresholding, and is thus of modest complexity. While we focus on OFDM signaling herein, our proposed method is easily adapted to other modulations. We observe that [10] also exploits the idea of linearity, but employs a piece-wise linear approximation via partial Fast Fourier Transform outputs. A single scale is assumed and estimated by a maximum-likelihood algorithm which searches over the candidate support of Doppler scale.

The remainder of the paper is organized as follows. Section 2 describes the OFDM signaling and MSML channel model. Section 3 develops the linear approximation for the MSML channel in the Fourier domain. In Section 4, we jointly exploit the linear relationship and the sparse structure of the MSML channel to design our channel estimation algorithm. Simulation results are presented in Section 5 and the paper is concluded in Section 6.

*This research has been funded in part by ONR N000144-09-1-0700, AFOSR FA9550-12-1-0215, DOT CA-26-7084-00, NSF CCF-1117896, NSF CNS-1213128, NSF CCF-1410009, NSF CPS-1446901, and CNP_q Brazil

2. SYSTEM MODEL

The OFDM signal in passband can be expressed as

$$\bar{x}(t) = \Re\left\{x(t)e^{j2\pi f_c t}\right\},$$

where $x(t) = \sum_{m=0}^{M-1} x_m e^{j2\pi f_m t} p(t)$, f_c is the carrier frequency, M is the number of sub-carriers, x_m is the data modulated onto the m-th subcarrier; and $f_m = m\Delta f$ is the m-th subcarrier frequency, where Δf is the sub-carrier spacing. A low-pass pulse shape $p(t)$ over the interval $t \in [0, T]$, where $T = \frac{1}{\Delta f}$ is the symbol duration, is employed to shape the transmitted signal. The passband time-varying channel impulse response for underwater acoustic communication is modeled as *multi-scale multi-lag* (MSML) channel,

$$\bar{h}(t,\tau) = \sum_{p=1}^{P} \bar{h}_p \delta(\tau - [\tau_p - a_p t]),$$

where P indicates the number of paths in the channel, and $\bar{h}_p$, τ_p, and a_p are channel gains, delays, and Doppler scale, respectively. All of these parameters are real-valued. The delay and scale values are assumed to lie in a finite range $\tau_p \in [0, \tau_{\max}]$ and $a_p \in [-a_{\max}, a_{\max}]$, where $\tau_{\max}$ denotes the delay spread of the channel and $a_{\min}$ is the maximum Doppler scale. The bandpass received signal after passing through a linear time-varying channel can be written as,

$$\bar{y}(t) = \int_{-\infty}^{+\infty} \bar{h}(t,\tau)\bar{x}(t-\tau)d\tau + \bar{z}(t),$$

where $\bar{z}(t)$ is the bandpass additive, Gaussian noise. Therefore, we have $\bar{y}(t) = \sum_{p=1}^{P} \bar{h}_p \bar{x}([1+a_p]t - \tau_p) + \bar{z}(t)$. If we consider that $\bar{y}(t) = \Re\left\{y(t)e^{j2\pi f_c t}\right\}$ and $\bar{z}(t) = \Re\left\{z(t)e^{j2\pi f_c t}\right\}$. Then, we can express the baseband system model as,

$$y(t) = \sum_{p=1}^{P} h_p x([1+a_p]t - \tau_p) + z(t). \quad (1)$$

where $h_p = \bar{h}_p e^{-j2\pi f_c \tau_p}$. Then, chip-matched filtering the received training signal (1) yields

$$y[n] = \sum_{p=1}^{P} h_p x_p[n - m_p] + z[n], \quad (2)$$

where $x_p[n] = x((1+a_p)T_s n)$ and $m_p = \frac{\tau_p}{(1+a_p)T_s}$. We assume that T_s is sufficiently small such that the delays m_p are approximately integer values.

3. LINEAR APPROXIMATION OF MSML CHANNEL

Herein, we design an algorithm to track the channel behavior at the receiver using the transmitted training signal in each transmission interval, as depicted in Fig. 1. As is clear from Fig. 1, we have two sets of unknown coefficients, *i.e.*, $\boldsymbol{\alpha}$ and $\mathbf{g}$. The vector $\mathbf{g}$ models the channel gains, the vector $\boldsymbol{\alpha}$ models the Doppler scale parameters, and the length of filter, L, is equal to the maximum discrete channel delay spread. Thus L is the length of the filter, whereas the number of paths in the channel is P and $L \gg P$ due to the sparse nature of the channel. To estimate variables $\boldsymbol{\alpha}$ and $\mathbf{g}$, we minimize the mean-squared error between the received signal $y[n]$ and its reconstruction, $\hat{y}[n]$, in Fig. 1, namely

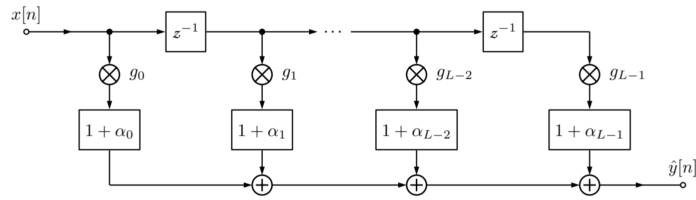

Fig. 1. Multi-scale multi-lag tracking filter schematic

$$\text{MSE} = \sum_{n=0}^{N-1} \frac{|y[n] - \hat{y}[n]|^2}{N} = \sum_{n=0}^{N-1} \frac{|Y[k] - \hat{Y}[k]|^2}{N^2}, \quad (3)$$

where $Y[k]$ and $\hat{Y}[k]$ denote the N-point discrete Fourier transform (DFT) of $y[n]$ and $\hat{y}[n]$, respectively. The equality in Eq. (3) holds due to Parseval's identity. Therefore, we minimize the MSE by minimizing the distance between the DFT of the received signal and our reconstructed signal. For the reconstructed signal we have,

$$\hat{y}[n] = \sum_{l=0}^{L-1} g_l x_l[n-l], \quad (4)$$

where $x_l[n]$ denotes the discretized scaled signal, *i.e.*, $x_l[n] = x(nT_s(1+\alpha_l))$. The N-point DFT of $\hat{y}[n]$ in (4) is computed as $\hat{Y}[k] = \sum_{l=0}^{L-1} g_l X_l[k] e^{-j\frac{2\pi}{N}kl}$, where $X_l[k]$ denotes the N-point DFT of $x_l[n]$. Thus, to compute $\hat{Y}[k]$, we need to compute $X_l[k]$. The $x_l[n]$ can be written as, $x_l[n] = \sum_{m=0}^{M-1} x_m e^{j2\pi f_m(1+\alpha_l)nT_s} p_l[n]$, where $p_l[n] = p((1+\alpha_l)nT_s)$. We know that the DFT of a signal is the sampled version of its discrete time Fourier transform (DTFT). Thus we first compute the DTFT of signal $x_l[n]$ and then perform sampling on the DTFT to compute the DFT of $x_l[n]$. The DTFT of $x_l[n]$ is as follows,

$$X_l(\omega) = \sum_{m=0}^{M-1} x_m P_l(\omega - 2\pi f_m T_s(1+\alpha_l)).$$

In addition, from the scaling property of the Fourier transform (if T_s is small enough to avoid aliasing), we have $P_l(\omega) = P\left(\frac{\omega}{1+\alpha_l}\right)$. Therefore,

$$X_l(\omega) = \sum_{m=0}^{M-1} x_m P\left(\frac{\omega}{1+\alpha_l} - 2\pi f_m T_s\right). \quad (5)$$

Note that $f_m T_s = (f_{\min} + m\Delta f)\frac{1}{M\Delta f} = \frac{f_{\min}}{M\Delta f} + \frac{m}{M}$. For simplicity, assume that $f_{\min} = m_0\Delta f$, where $m_0 \in \mathbb{N}$. Therefore, we have $f_m T_s = \frac{m_0+m}{M}$. Considering the fact that in underwater acoustic channel Doppler scales are small, *i.e.*, $\alpha_l \ll 1$, then we can approximate $\frac{1}{1+\alpha_l} \approx 1 - \alpha_l$, yielding

$$X_l(\omega) \simeq \sum_{m=0}^{M-1} x_m P\left((1-\alpha_l)\omega - 2\pi\frac{m_0+m}{M}\right).$$

Since the pulse shaping window is a low-pass signal in our design, to simplify the mathematical analysis, we consider the linear approximation to its Fourier transform. We seek the linear approximation of $P(\omega)$ inside the interval $[0, 2\pi]$ using M_l points as

$$P(\omega) \simeq P_i(\omega) \text{ for } \omega \in \mathcal{I}_i = \left[\frac{2\pi}{M_l} i, \frac{2\pi}{M_l}(i+1)\right), \quad (6)$$

and $P_i(\omega) = c_i\omega + b_i$ for $i = 0, 1, \ldots, M_l - 1$. Here c_i and b_i for $i = 1, \ldots, M_l$ are the complex numbers that minimize the following mean-square error

$$e_{M_l} = \sum_{i=0}^{M_l-1} \int_{\mathcal{I}_i} |P(\omega) - P_i(\omega)|^2 \, d\omega$$

To determine the c_i and b_i that minimize the value of I_M, we set the first order derivatives equal to zero and control the sign of the second derivatives. Hence, we obtain

$$\frac{\partial e_{M_l}}{\partial \mathrm{Re}(c_i)} + j\frac{\partial e_{M_l}}{\partial \mathrm{Im}(c_i)} = 2\int_{\mathcal{I}_i} [P_i(\omega) - P(\omega)]\omega d\omega = 0$$

$$\frac{\partial e_{M_l}}{\partial \mathrm{Re}(b_i)} + j\frac{\partial e_{M_l}}{\partial \mathrm{Im}(b_i)} = 2\int_{\mathcal{I}_i} [P_i(\omega) - P(\omega)] \, d\omega = 0$$

Let us define $\mu_{ik} := \int_{\mathcal{I}_i} \omega^k P(\omega) d\omega$. Then, using the equations above, we obtain c_i and b_i, for $i = 0, \ldots, M_l - 1$, as

$$c_i = \frac{3}{2}\left(\frac{M_l}{\pi}\right)^3 \left[\mu_{i1} - \frac{2\pi\mu_{i0}}{M_l}\left(i + \frac{1}{2}\right)\right], \quad (7)$$

$$b_i = \frac{M_l}{2\pi}\left[\mu_{i0} - \left(\frac{M_l}{2\pi}\right)^2 \left(i + \frac{1}{2}\right) c_i\right]. \quad (8)$$

Now, if we substitute the values of b_i and c_i from Eqs. (7) and (8) in Eq. (6), and then substitute the result in Eq. (5), we have

$$X_l(\omega) \approx \sum_{m=0}^{M-1} \left(\omega(1 + \alpha_l) - \frac{2\pi(m + m_0)}{M}\right) x_m c_i + x_m b_i,$$

such that $\omega \in 2\pi(1 + \alpha_l)\left[\frac{i}{M_l} + \frac{m+m_0}{M}, \frac{i+1}{M_l} + \frac{m+m_0}{M}\right)$. If we approximate the $X_l(\omega)$ by M_d-point DFT and $1 + \alpha_l \approx 1$, we have

$$X_l[k] \approx \sum_{m=0}^{M-1} 2\pi x_m c_i \left(\frac{k(1 - \alpha_l)}{M_d} - \frac{m + m_0}{M}\right) + x_m b_i,$$

such that $k \in M_d\left[\frac{i}{M_l} + \frac{m+m_0}{M}, \frac{i+1}{M_l} + \frac{m+m_0}{M}\right)$. Replacing i with $i - m\frac{M_l}{M} = i - mq$ in the above equation, we have

$$X_l[k] = v[k]\alpha_l + u[k], \quad (9)$$

where $v[k] = \frac{-2\pi}{M_d}\left(\sum_{m=0}^{M-1} x_m a_{i-mq}\right) k$ and

$$u[k] = \sum_{m=0}^{M-1} \left(b_{i-mq} + 2\pi\left(\frac{k}{M_d} - \frac{m + m_0}{M}\right) c_{i-mq}\right) x_m,$$

for $k \in M_d\left[\frac{i}{M_l} + \frac{m_0}{M}, \frac{i+1}{M_l} + \frac{m_0}{M}\right)$. Therefore, we find the linear relationship between the Doppler scales and Fourier transform of the scaled signal in each branch in Fig. 1. In the next section, we design an algorithm to compute the filter coefficients $\mathbf{g} = [g_0, g_1, \ldots, g_{L-1}]$ and Doppler scales $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_{L-1}]$ using the received data to minimize the MSE in Eq. (3).

4. COMPUTING UNKNOWN COEFFICIENTS

We know that in underwater acoustic communication channel, there are only a small number of dominant paths between the transmitter and the receiver; and the channel gains decrease exponentially by increasing the channel delay. Therefore, for proper sampling resolution, we expect that the number of active (non-zero) coefficients in $\mathbf{g}$ be a small number, *i.e.*, $P \ll L$, meaning that vector $\mathbf{g}$ is a sparse vector. Therefore to update the filter coefficients we propose the target optimization problem

$$\underset{\boldsymbol{\alpha},\mathbf{g}}{\operatorname{argmin}}\ J(\boldsymbol{\alpha}, \mathbf{g}) = \underset{\boldsymbol{\alpha},\mathbf{g}}{\operatorname{argmin}}\ L_{\boldsymbol{\alpha},\mathbf{g}}(Y, \hat{Y}) + \lambda\Omega(\mathbf{g}) \quad (10)$$

where $L_{\boldsymbol{\alpha},\mathbf{g}}(R, \hat{R})$ controls the closeness to the measurements, $\Omega(\mathbf{g})$ regularizes our *prior* knowledge about the channel structure, such as the sparsity of $\mathbf{g}$; λ is the Lagrange multiplier that trades off between sparsity and proximity to the measurements. We model the noise as additive and Gaussian, thus the proper choice for our loss function is

$$L_{\boldsymbol{\alpha},\mathbf{g}}(R, \hat{R}) = \sum_{k=0}^{N-1} \left(Y[k] - \hat{Y}[k]\right)^2;$$

to promote sparsity, we consider the convex l_1-norm as the regularizer (see *e.g.* [**?**]). Therefore, the objective function for the optimization problem in (10) can be written as

$$J(\boldsymbol{\alpha}, \mathbf{g}) := \sum_{k=0}^{N-1} \left(Y[k] - \hat{Y}[k]\right)^2 + \lambda\|\mathbf{g}\|_1. \quad (11)$$

Since $\hat{y}[n] = \sum_{l=0}^{L} g_l x_l[n - l]$, by taking the Fourier transform and using Eq. (9), we have

$$\hat{Y}[k] = \sum_{l=0}^{L} g_l \left(v_l[k]\alpha_l + u_l[k]\right) \quad (12)$$

where $v_l[k] = v[k]e^{-j\frac{2\pi}{N}kl}$ and $u_l[k] = u[k]e^{-j\frac{2\pi}{N}kl}$. If we substitute $\hat{Y}[k]$ from Eq. (12) into Eq. (11), we have

$$J(\boldsymbol{\alpha}, \mathbf{g}) = \sum_{k=0}^{N-1} \left(Y[k] - \sum_{l=0}^{L} g_l \left(v_l[k]\alpha_l + u_l[k]\right)\right)^2 + \lambda\|\mathbf{g}\|_1.$$

Let us define auxiliary variable $\alpha'_l = g_l\alpha_l$. Then, we can rewrite the above equation as

$$J(\boldsymbol{\alpha}, \mathbf{g}) = \|\mathbf{y} - (\mathbf{V}\boldsymbol{\alpha}' + \mathbf{U}\mathbf{g})\|_2^2 + \lambda\|\mathbf{g}\|_1 \quad (13)$$

where $\mathbf{y} = [Y[0], Y[1], \ldots, Y[N-1]]^T$, $\mathbf{V} \in \mathbb{C}^{N\times L}$ and $\mathbf{V}[i, j] = v_j[i]$, and $\mathbf{U} \in \mathbb{C}^{N\times L}$ and $\mathbf{U}[i, j] = u_j[i]$. As seen in Eq. (13), the objective function $J(\boldsymbol{\alpha}, \mathbf{g})$ is a convex function of design variables, $[\boldsymbol{\alpha}, \mathbf{g}]$. Thus there exists an optimal solution for the problem in Eq. (10). Therefore, first we derive the optimal solution for $\boldsymbol{\alpha}'$ given that $\mathbf{g}$ is known. Then we substitute this value in (13) and compute the optimal solution for $\mathbf{g}$. It is clear that the optimal solution of $\boldsymbol{\alpha}'$ is the least-squares solution of objective function in Eq. (13) and can be written as,

$$\boldsymbol{\alpha}' = \mathbf{V}^+ (\mathbf{y} - \mathbf{U}\mathbf{g}) \quad (14)$$

where $\mathbf{V}^+ = \left(\mathbf{V}^H\mathbf{V}+\right)^{-1}\mathbf{V}^H$ is the pseudo-inverse of matrix $\mathbf{V}$. Now if we define $\mathbf{V}_h = \mathbf{I} - \mathbf{V}\mathbf{V}^+$, then we can rewrite (13) as

$$J\left(\mathbf{g}\right) = \|\mathbf{V}_h\left(\mathbf{y} - \mathbf{U}\mathbf{g}\right)\|_2^2 + \lambda\|\mathbf{g}\|_1. \tag{15}$$

Since the objective function $J\left(\mathbf{g}\right)$ in (15) is a l_1 regularization problem, we can apply the alternating direction method of multipliers [?] to solve this optimization problem as follows. For the optimization problem in (15), ADMM consists of the following iterations,

$$\mathbf{w}^{n+1} = \left(\mathbf{U}^T\bar{\mathbf{V}}_h\mathbf{U} + \rho\mathbf{I}\right)^{-1}\left(\mathbf{U}^T\bar{\mathbf{V}}_h\mathbf{y} + \rho\mathbf{g}^n - \boldsymbol{\theta}^n\right),$$

$$\mathbf{g}^{n+1}[i] = \left((\mathbf{w}^{n+1} + \frac{1}{\rho}\boldsymbol{\theta}^n)[i] - \lambda\right)_+ \quad \text{for } i = 0, \ldots, L-1$$

$$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n + \rho\left(\mathbf{w}^{n+1} - \mathbf{g}^{n+1}\right),$$

where $\mathbf{w}$ is an auxiliary variable, $\bar{\mathbf{V}}_h = \mathbf{V}_h^T\mathbf{V}_h$, the constant $\rho > 0$ is the augmented Lagrangian parameter, $\boldsymbol{\theta}$ is the dual variable, index n indicates the iteration number, and $(x)_+ = \max(0, x)$. Note that in the $\mathbf{w}$ update step, the matrix $\left(\mathbf{U}^T\bar{\mathbf{V}}_h\mathbf{U} + \rho\mathbf{I}\right)^{-1}$ is known at the receiver and does not change in each iteration, thus it will be computed only once in the system design. Therefore, in each iteration, only a single matrix-vector multiplication is performed to update vector $\mathbf{w}$, which requires $L^2 + L$ flops. To update the vector $\mathbf{g}$ a simple scalar thresholding (comparison) is fulfilled, which requires L flops, and finally the dual variables are updated by a simple summation operation in L flops. After computing the channel gains, $\mathbf{g}$, using the ADMM algorithm, we substitute the evaluated $\mathbf{g}$ into Eq. (14) to compute the Doppler scale parameters $\boldsymbol{\alpha}'$ using a matrix-vector multiplication operation, which requires LN flops. Therefore, in total $\left[(L^2 + 3L)I_t + LN\right]$ flops are required to compile the overall algorithm, where I_t is the number of ADMM algorithm iterations, L is the length of tracking filter, and N is the number of measurements.

5. NUMERICAL RESULTS

We simulate the performance of our proposed method in terms of normalized mean square error (NMSE), *i.e.*, NMSE $= \mathbb{E}\left\{\|\boldsymbol{\nu} - \hat{\boldsymbol{\nu}}\|_2^2/\|\boldsymbol{\nu}\|_2^2\right\}$, where $\boldsymbol{\nu}$ is the true value and $\hat{\boldsymbol{\nu}}$ is the estimated value. We consider OFDM signaling for an UWA communication system with a minimum subcarrier frequency $f_c = 10$ kHz and subcarrier spacing $\Delta f = 10$ Hz. The sampling frequency is considered as $f_s = \frac{1}{T_s} = 39$ kHz. Furthermore, $L = 50 \times P$ is considered where $P \leq 10$, $N = 2M - 1$, and $M_d = M_l = M$ in all simulations. Furthermore, the OFDM pulse shape, $p(t)$ is a rectangular pulse with $T = \frac{1}{\Delta f}$. In Fig. 2, we examine the accuracy of our proposed approximation in Section 3. Thus, we consider that all the parameters (h_p, a_p, τ_p) for $p = 1, \ldots, P$, are known. Then, we compute the approximation error $\|\mathbf{y} - \hat{\mathbf{y}}\|_2$. It is clear from Eq. (4) to (9) that two key parameters in the approximation error are the number of subcarriers, M, and value of Doppler scale parameters, $a_{\max}$. Results in Fig. 2 indicate that by increasing the total number of subcarriers and the maximum Doppler scale, the approximation accuracy decreases. However, for a total number of subcarriers less than $M = 128$, our approximation results in an NMSE less

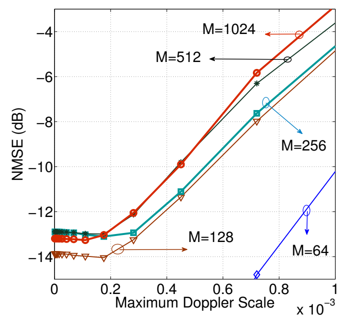

Fig. 2. Approximation error

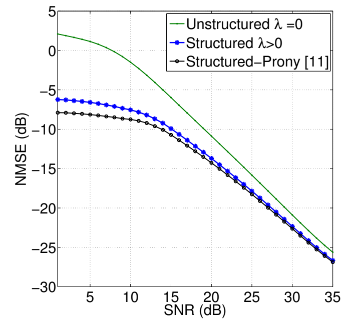

Fig. 3. MSML channel estimation

than 0.1 for $a_{\max} \leq 10^{-3}$. Furthermore, for $M \leq 1024$ and $a_{\max} \leq 5 \times 10^{-4}$, the NMSE is less than 0.1. Thus for systems designed within these ranges, our proposed method offers strong performance

In Fig. 3, we compare the performance our method proposed herein with our previously designed Structured-Prony method [5]. We have that $M = 512$, $a_{\max} = 10^{-4}$, and $I_t = 3$. We see that the Structured-Prony method offers a 1 dB improvement over the current method for SNRs less than 10. However, this performance gain comes at the cost of complexity. The Structured-Prony method [6], which is relatively low complexity algorithm among subspace methods (*e.g.* [9]), requires almost $\left(LN^2 + L^3 + 2\right)I_t + (PM)^3 + PM\log(PM)$ flops versus that of the current method ($\left[(L^2 + 3L)I_t + LN\right]$). In Fig. 3, we also examine the impact of enforcing sparsity on the quality of channel estimation. For the case where include the sparse structure of channel we have $\lambda > 0$; we can ignore such sparsity by setting $\lambda = 0$ in the optimization problem in Eq. (13). We see in Fig. 3 that there is about a 5 dB improvement in NMSE for SNR values less than 15 dB by considering the channel sparse structure.

6. CONCLUSIONS

In this paper, we have introduced a new, very low complexity channel estimation strategy for MSML channels. In particular, we show that one can well-approximate the received signal in the Fourier domain under the assumption of a training signal. In particular, the nonlinear effects of the multiple Doppler scales, can be linearized due to the assumption of small absolute scales. We additionally exploit the inherent sparsity of underwater acoustic channels to enhance the quality of channel estimation. The proposed method offers performance only slight worse than that of a previously proposed scheme but with complexity that is squared with the filter length versus cubed with the filter length. While we have focused on OFDM signaling, the proposed method can be extended to a broader class of signals, *e.g.*, Gabor frames.

7. REFERENCES

[1] Y. Jiang and A. Papandreou-Suppappola, "Discrete time-scale characterization of wideband time-varying systems," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1364–1375, 2006.

[2] A.R. Margetts, P. Schniter, and A. Swami, "Joint scale-lag diversity in wideband mobile direct sequence spread spectrum systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4308–4319, 2007.

[3] S.T. Rickard, R.V. Balan, H.V. Poor, and S. Verdu, "Canonical time-frequency, time-scale, and frequency-scale representations of time-varying channels," *J. Commun. Inf. Syst.*, vol. 5, no. 5, pp. 1–30, 2005.

[4] S. Beygi and U. Mitra, "Optimal bayesian resampling for OFDM signaling over multi-scale multi-lag channels," *IEEE Signal Process. Letters,* vol. 20, no. 11, pp. 1118–1121, 2013.

[5] S. Beygi and U. Mitra, "Multi-scale multi-lag channel estimation using low rank structure of received signal," *IEEE Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 4-9 May, 2014.

[6] S. Beygi, U. Mitra, "Multi-scale Multi-lag Channel Estimation Using Low Rank Approximation for OFDM," submitted to *IEEE Trans. Signal Process*.

[7] S. Yerramalli and U. Mitra, "Optimal resampling of OFDM signals for multiscale-multilag underwater acoustic channels," *IEEE J. Ocean. Eng.*, vol. 36, no. 1, pp. 126–138, 2011.

[8] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts," *IEEE J. Ocean. Eng.*, vol. 33, no. 2, pp. 1638–1649, Apr. 2008.

[9] C.R. Berger, S. Zhou, J.C. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1708–1721, 2010.

[10] S. Yerramalli, M. Stojanovic and U. Mitra, "Partial FFT demodulation: a detection method for Doppler distorted OFDM systems," *IEEE Trans. Signal Process.*, vol. 60, No. 11, , pp. 5906–5918, Nov. 2012.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," Foundations and Trends in Machine Learning, vol 3, no. 1, pp. 1–122, 2011.