

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Iterative Equalization for OFDM Systems over Wideband MSML Channels

EURASIP Journal on Wireless Communications and Networking 2012,
2012:280 doi:10.1186/1687-1499-2012-280

Tao Xu (t.xu@tudelft.nl)
Zijian Tang (z.tang@tudelft.nl)
Rob Remis (r.f.remis@tudelft.nl)
Geert Leus (g.j.eus@tudelft.nl)

ISSN 1687-1499

Article type Research

Submission date 20 March 2012

Acceptance date 10 August 2012

Publication date 31 August 2012

Article URL <http://jwcn.eurasipjournals.com/content/2012/1/280>

This peer-reviewed article can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

For information about publishing your research in *EURASIP WCN* go to

<http://jwcn.eurasipjournals.com/authors/instructions/>

For information about other SpringerOpen publications go to

<http://www.springeropen.com>

Iterative equalization for OFDM systems over wideband Multi-Scale Multi-Lag channels

Tao Xu^{1*}

*Corresponding author

Email: t.xu@tudelft.nl

Zijian Tang^{1,2}

Email: z.tang@tudelft.nl

Rob Remis¹

Email: r.f.remis@tudelft.nl

Geert Leus¹

Email: G.J.T.Leus@tudelft.nl

¹Delft University of Technology, Mekelweg 4, Delft, 2628CD, The Netherlands

²TNO, Oude Waalsdorperweg 63, The Hague, 2597 AK, The Netherlands

Abstract

OFDM suffers from inter-carrier interference (ICI) when the channel is time varying. This article seeks to quantify the amount of interference resulting from wideband OFDM channels, which are assumed to follow the multi-scale multi-lag (MSML) model. The MSML channel model results in full channel matrices both in the frequency and time domains. However, banded approximations are possible, leading to a significant reduction in the equalization complexity. Measures for determining whether a time-domain or frequency-domain approach should be undertaken are provided based on the interference analysis, and we propose to use the conjugate gradient (CG) algorithm to equalize the channel iteratively. The suitability of a preconditioning technique, that often accompanies the CG method to accelerate the convergence, is also discussed. We show that in order for the diagonal preconditioner to function properly, optimal resampling is indispensable.

Introduction

With many desirable properties such as a high spectral efficiency and inherent resilience to the multipath dispersions of frequency-selective channels, the orthogonal frequency division multiplexing (OFDM) technology shows attractive features to wireless radio applications [1]. OFDM relies on the assumption that the channel stays constant within at least one OFDM symbol period. When Doppler effects due to temporal channel variation cannot be ignored, this assumption does not hold any more since the communication channel becomes time varying. The Doppler effects corrupt the orthogonality among OFDM subcarriers by inducing

non-negligible inter-carrier interference (ICI) [2], and can therefore severely deteriorate the system performance. For traditional terrestrial radio systems, compensation of ICI in channel equalization has widely been researched for *narrowband* systems. Due to the small relative signal bandwidth (actual bandwidth divided by the center frequency) of narrowband systems, the Doppler effects can be modeled primarily by frequency shifts [3, 4], in which case it is reasonable to assume that each OFDM subcarrier experiences a statistically identical frequency offset [2]. Consequently, the effective channel matrix of a narrowband OFDM system in the presence of Doppler can be approximated as banded. Efficient equalization schemes for such a banded channel matrix have been studied in, e.g., [5–7].

In a *wideband* system, where the relative signal bandwidth is large, the Doppler effects should be more appropriately modeled as scalings of the signal waveform [3, 4]. Wideband systems arise in, e.g., underwater acoustic (UWA) systems or wideband terrestrial radio frequency systems such as ultra wideband (UWB). Due to multipath, a wideband linear time-varying (LTV) channel can be more accurately described by a multi-scale multi-lag (MSML) model [3, 8]. Many signaling schemes have been studied for wideband systems. For instance, [9, 10] consider direct-sequence spread spectrum (DSSS). Recently, the use of OFDM for UWA or UWB has received considerable attention. To counteract the scaling effect due to Doppler, [11] proposes a multi-band OFDM system such that within each band, the narrowband assumption can still be valid. More commonly, many works assume a single-scale multi-lag (SSML) model for the wideband LTV channel. Based on the SSML assumption, after a resampling operation the channel can be approximated by a time-invariant channel but subject to a carrier frequency offset (CFO) [12, 13]. However, since the channel should be more accurately described by an MSML model, determining the optimal resampling rate is not trivial [14].

In this article, we consider OFDM transmission based on an MSML model. The resulting channel, which is a full matrix in the presence of Doppler, will be equalized by means of the conjugate gradient (CG) algorithm [15], whose performance is less sensitive to the condition of the channel matrix than, e.g., a least-squares approach. On the other hand, the convergence rate of CG is inversely proportional to the channel matrix condition number. This is especially of significance if a truncated CG is to be used in practice, which halts the algorithm after a limited number of iterations in order to reduce the overall complexity. Therefore, it is desired that the channel matrix is well-conditioned to ensure a fast convergence. To this end, preconditioning techniques can be invoked to enforce the eigenvalues of the channel matrix to cluster around one [16]. To achieve a balance between performance and complexity, we restrict the preconditioner to be a diagonal matrix, whose diagonal entries can be designed by following the steps given in [17]. We notice that a circulant preconditioner in the time domain was introduced in [18], which is equivalent to a diagonal preconditioner in the frequency domain. This preconditioner is introduced based on a basis expansion model (BEM), which is often used to approximate the channel's time-variation for a narrowband system. For a wideband system as considered in this article, it can be shown that this preconditioner in the frequency domain is equal to the inverse of the diagonal entries of the frequency-domain channel matrix.

What is not considered in [17, 18] is the resampling operation at the receiver, which is an indispensable and crucial step for wideband LTV channels. Different from the trivial

resampling scheme for SSML channel models, an optimum resampling method is proposed in [14] for MSML channels, which aims at minimizing the average error of approximating the MSML channel by an SSML model. This article studies the resampling from a preconditioning point of view. It is observed that if the major channel energy is located on the off-diagonals of the channel matrix, a diagonal preconditioner will deteriorate the channel matrix condition rather than improve it, thereby reducing the convergence rate of CG instead of increasing it as opposed to the claim of [17]. The energy distribution of the channel matrix is governed by the resampling. Different from [14], which only considers rescaling the received signal, and [19], which considers both rescaling and frequency synchronization, this article will show that for OFDM systems, all these three resampling parameters can have a significant impact on the system performance (i.e., rescaling, frequency synchronization and time synchronization). More specifically, we will extend the results of [19, 20] by jointly optimizing these three resampling parameters both in the frequency domain and the time domain.

Notation: Upper (lower) bold-face letters stand for matrices (vectors); superscripts T , H , and $*$ denote transpose, Hermitian transpose and conjugate transpose, respectively; we reserve j for the imaginary unit, $\lfloor k \rfloor$ and $\lceil k \rceil$ for integer rounding and ceiling of a number k , $\|\mathbf{x}\|_2$ for the two norm of the vector \mathbf{x} , $\|\mathbf{A}\|_{\text{Fro}}$ for the Frobenius norm of the matrix \mathbf{A} , $[\mathbf{A}]_{k,m}$ for the (k, m) th entry of the matrix \mathbf{A} ; $\text{diag}(\mathbf{x})$ for a diagonal matrix with \mathbf{x} on its main diagonal, and \odot for the Hadamard product of two matrices.

System model based on an MSML channel

Continuous data model

Suppose that the baseband transmit signal $s(t)$ consists of K subcarriers, and can be written as

$$s(t) = \frac{1}{\sqrt{KT}} \sum_{k=0}^{K-1} b_k e^{j2\pi f_k t} u(t), \quad -T_{\text{pre}} < t \leq KT + T_{\text{post}} \quad (1)$$

where the data symbol b_k is modulated on the k th subcarrier $f_k = k\Delta f$, for $k = 0, 1, \dots, K-1$, with Δf being the OFDM subcarrier spacing. With $T = 1/(K\Delta f)$, KT is the effective duration of an OFDM symbol. The cyclic prefix and postfix are given as T_{pre} and T_{post} , respectively. The cyclic prefix is assumed to be longer than the delay spread and the cyclic postfix is long enough to ensure signal completeness in case of scaling, which will be defined later on. The rectangular pulse $u(t)$ is defined to be 1 within $t \in [-T_{\text{pre}}, KT + T_{\text{post}}]$ and 0 otherwise. Prior to transmission, $s(t)$ is up-converted to passband, yielding $\bar{s}(t) = \Re\{s(t)e^{j2\pi f_c t}\}$, where f_c denotes the carrier frequency. With sufficient cyclic extensions, the interference from adjacent OFDM symbols can be neglected and hence we are allowed to consider an isolated OFDM symbol in this article without loss of generality. Although this article discusses the scenario when cyclic extensions are used, the analysis can be directly applied to zero padding OFDM (ZP-OFDM) with minor modifications.

The considered signal is transmitted over a wideband LTV channel, which is assumed to comprise multiple resolvable paths. The l th path can mathematically be characterized by the

following three parameters: \bar{h}_l , the path gain; v_l , the radial velocity which is uniquely determined by the incident angle of this path; and τ_l , the delay due to the propagation time. In compliance with the wideband assumption, the received signal resulting from the l th path is given by $\bar{h}_l \sqrt{\alpha_l} \bar{s}(\alpha_l(t - \tau_l))$, where $\alpha_l = \frac{c+v_l}{c-v_l} \approx 1 + \frac{2v_l}{c}$ is the scaling factor with c the speed of the communication medium (normally $c \gg v_l$) and $\sqrt{\alpha_l}$ is added as a normalization factor. Depending on the sign of v_l , the received signal waveform via this path can be either dilated (a negative v_l) or compressed (a positive v_l).

With a collection of $L + 1$ paths, the actual received signal $\bar{y}(t)$ is given by

$$\begin{aligned} \bar{y}(t) &= \sum_{l=0}^L \bar{h}_l \sqrt{\alpha_l} \bar{s}(\alpha_l(t - \tau_l)) + \bar{w}(t), \\ &= \sum_{l=0}^L \bar{h}_l \sqrt{\alpha_l} \Re\{s(\alpha_l(t - \tau_l)) e^{j2\pi f_c \alpha_l(t - \tau_l)}\} + \bar{w}(t), \end{aligned} \quad (2)$$

where $\bar{w}(t)$ stands for the passband noise. In the above, if there exist at least two paths l and l' , for which $\alpha_l \neq \alpha_{l'}$ and/or $\tau_l \neq \tau_{l'}$, the channel exhibits a multi-scale multi-lag (MSML) character. For a practical channel, it is realistic to assume that $\alpha_l \in [1, \alpha_{\max}]$ and $\tau_l \in [0, \tau_{\max}]^a$, where $\alpha_{\max} \geq 1$ and $\tau_{\max} \geq 0$ determines the scale spread and delay spread, respectively. Note that in many prior works [12, 13], the approximation $\alpha_l \approx \alpha_{l'}$ for any $l \neq l'$ is adopted for the sake of analytical ease, which gives rise to an SSML model.

The equivalent complex baseband received signal of $\bar{y}(t) = \Re\{y(t) e^{j2\pi f_c t}\}$ is given by

$$y(t) = e^{-j2\pi f_c t} \sum_{l=0}^L \bar{h}_l \sqrt{\alpha_l} s(\alpha_l(t - \tau_l)) e^{j2\pi \alpha_l f_c (t - \tau_l)} + w(t), \quad (3)$$

where $w(t)$ stands for the baseband noise. By substituting (1) in the above, we can rewrite $y(t)$ as

$$\begin{aligned} y(t) &= \sum_{l=0}^L \bar{h}_l e^{-j2\pi f_c \alpha_l \tau_l} \sqrt{\alpha_l} \left(\frac{1}{\sqrt{KT}} \sum_{k=0}^{K-1} b_k e^{j2\pi f_k \alpha_l (t - \tau_l)} u(\alpha_l(t - \tau_l)) \right) e^{j2\pi (\alpha_l - 1) f_c t} + w(t) \\ &= \frac{1}{\sqrt{KT}} \sum_{k=0}^{K-1} b_k h_k(t) e^{j2\pi f_k t} + w(t), \end{aligned} \quad (4)$$

where

$$h_k(t) = \sum_{l=0}^L \bar{h}_l \sqrt{\alpha_l} e^{-j2\pi (f_c + f_k) \alpha_l \tau_l} e^{j2\pi (\alpha_l - 1) (f_c + f_k) t} u(\alpha_l t - \alpha_l \tau_l), \quad (5)$$

which stands for the time-varying channel frequency response seen by the k th subcarrier. From the definition of $h_k(t)$, we notice that the k th subcarrier experiences a frequency offset of $(\alpha_l - 1)(f_c + f_k)$ over the l th path.

Remark 1 The cyclic prefix is assumed to be longer than the delay spread and the cyclic postfix has a duration long enough to ensure signal continuity in the observation window for $t \in [0, KT]$. Specifically, it is required that $u(\alpha_l t - \alpha_l \tau_l) = 1$ within this window for all paths. In other words, because $u(\alpha_l t - \alpha_l \tau_l)$ gives a time support on $t \in \left[\frac{-T_{\text{pre}} + \alpha_l \tau_l}{\alpha_l}, \frac{KT + T_{\text{post}} + \alpha_l \tau_l}{\alpha_l} \right]$, we should then always satisfy $\frac{-T_{\text{pre}} + \alpha_l \tau_l}{\alpha_l} \leq 0$ and also $\frac{KT + T_{\text{post}} + \alpha_l \tau_l}{\alpha_l} \geq KT$ for any $l \in \{0, 1, \dots, L\}$. It leads to

$$T_{\text{pre}} \geq \alpha_{\max} \tau_{\max} \quad (6)$$

$$T_{\text{post}} \geq (\alpha_{\max} - 1)KT. \quad (7)$$

When the above conditions are satisfied, we are allowed to drop the notation of the rectangular pulse $u(t)$ embedded in $h_k(t)$ in the sequel for the sake of notational ease.

Discrete data model

For MSML channels, discretizing the received signal and achieving time/frequency synchronization is not trivial [12, 14]. We illustrate such difficulty in Figure 1, where we assume the transmit signal propagates via three paths. Since the received symbol is the summation of these three paths, it invites the following questions:

1. Which point should we consider as the starting point of the OFDM symbol (time synchronization)?
2. What sampling rate should we adopt to discretize the received signal over MSML channels (rescaling)?
3. What frequency shift should we apply to remove the residual carrier frequency offset (frequency synchronization)?

Figure 1 Illustration of the synchronization and resampling problem; α_l stands for the scaling factor due to the l th path, and β for the rescaling factor adopted by the receiver during resampling

These problems can mathematically be described by determining β , ϕ and σ in the following expression

$$r^{(\beta, \phi, \sigma)}(t) = \sqrt{\frac{1}{\beta}} y\left(\frac{t}{\beta} - \sigma T\right) e^{j2\pi f_c \phi t / \beta}, \quad (8)$$

where β is a positive number within $[1, \alpha_{\max}]$ and βT represents the sampling rate at the receiver; σ is the time shift factor, which is used to represent time synchronization; and likewise, ϕ is the phase shift factor used for frequency synchronization. $\sqrt{\frac{1}{\beta}}$ is a normalization factor. Later on, we will show that a different choice of (β, ϕ, σ) can influence the energy

distribution of the channel matrix significantly. For the moment, we leave the values of these parameters open to allow for a general treatment of the problems. It is clear that when $(\beta, \phi, \sigma) = (1, 0, 0)$, there is no resampling operation carried out.

After resampling, the noiseless sample obtained at the n th time instance in the time domain is given by (see Appendix 2 for the detailed derivation)

$$\begin{aligned} r_n^{(\beta, \phi, \sigma)} &= r^{(\beta, \phi, \sigma)}(nT) \\ &= \sum_{l=0}^L h_l^{(\beta, \sigma)} e^{j2\pi\omega \frac{(\alpha_l - 1 + \phi)}{\beta} \frac{n}{K}} \times \left(\sum_{k=0}^{K-1} b_k e^{j2\pi \frac{\alpha_l}{\beta} \frac{nk}{K}} \right) e^{-j2\pi\alpha_l(\lambda_l + \sigma) \frac{k}{K}}, \end{aligned} \quad (9)$$

where we use

$$\omega = \frac{f_c}{\Delta f}$$

to denote the normalized carrier frequency and

$$\lambda_l = \frac{\tau_l}{T}$$

to denote the normalized delay of the l th path; and the discrete channel coefficient is given by

$$h_l^{(\beta, \sigma)} = \sqrt{\frac{\alpha_l}{\beta KT}} \bar{h}_l e^{-j2\pi f_c(\alpha_l \tau_l + (\alpha_l - 1)\sigma T)}.$$

In (9), the term $e^{j2\pi\omega \frac{(\alpha_l - 1 + \phi)}{\beta} \frac{n}{K}}$ corresponds to the residual CFO related with the l th path after resampling; the term $e^{-j2\pi f_c \alpha_l(\lambda_l + \sigma) \frac{k}{K}}$ corresponds to the phase changes due to the time shift along the l th path; and the summation $\sum_{k=0}^{K-1} b_k e^{j2\pi \frac{\alpha_l}{\beta} \frac{nk}{K}}$ is the adapted version of the transmitted OFDM signal due to the channel time variation in the l th path.

Let us now stack the received samples $r_n^{(\beta, \phi, \sigma)}$, for $n = 0, \dots, K-1$, into a vector $\mathbf{r}_T^{(\beta, \phi, \sigma)} = [r_0^{(\beta, \phi, \sigma)}, \dots, r_{K-1}^{(\beta, \phi, \sigma)}]^T$, and similarly let $\mathbf{b} = [b_0, \dots, b_{K-1}]^T$. In the noiseless case, it follows that

$$\mathbf{r}_T^{(\beta, \phi, \sigma)} = \sum_{l=0}^L h_l^{(\beta, \sigma)} \mathbf{D}_l^{(\beta, \phi)} \mathbf{F}_{\alpha_l/\beta}^H \mathbf{\Lambda}_l^{(\sigma)} \mathbf{b}, \quad (10)$$

where \mathbf{F}_α denotes a fractional normalized discrete Fourier transform (DFT) matrix, whose (m, k) th entry is defined as

$$[\mathbf{F}_\alpha^H]_{m,k} = \frac{1}{\sqrt{K}} e^{j2\pi\alpha \frac{mk}{K}}. \quad (11)$$

Obviously, \mathbf{F}_1 reduces to a regular normalized DFT matrix. In addition,

$$\mathbf{\Lambda}_l^{(\sigma)} = \text{diag}([1, e^{j2\pi\alpha_l(\lambda_l + \sigma) \frac{1}{K}}, \dots, e^{j2\pi\alpha_l(\lambda_l + \sigma) \frac{K-1}{K}}]^T), \quad (12)$$

and

$$\mathbf{D}_l^{(\beta, \phi)} = \text{diag}([1, e^{j2\pi\omega \frac{\alpha_l - 1 + \phi}{\beta} \frac{1}{K}}, \dots, e^{j2\pi\omega \frac{\alpha_l - 1 + \phi}{\beta} \frac{K-1}{K}}]^T), \quad (13)$$

where the superscript (β, ϕ) in $\mathbf{D}_l^{(\beta, \phi)}$ and (σ) in $\Lambda_l^{(\sigma)}$ reflects the dependence on the specific resampling parameters. This convention will hold throughout this article.

Interference analysis

Normally speaking, equalization of an OFDM channel is implemented in the frequency domain. To this end, the received signal $\mathbf{r}_T^{(\beta, \phi, \sigma)}$ is first transformed into the frequency domain by means of the DFT, which in the absence of noise yields

$$\mathbf{r}_F^{(\beta, \phi, \sigma)} = \mathbf{F}_1 \mathbf{r}_T^{(\beta, \phi, \sigma)} = \mathbf{H}_F^{(\beta, \phi, \sigma)} \mathbf{b}, \quad (14)$$

where $\mathbf{H}_F^{(\beta, \phi, \sigma)}$ stands for the frequency-domain (FD) channel matrix, which is defined as

$$\begin{aligned} \mathbf{H}_F^{(\beta, \phi, \sigma)} &= \sum_{l=0}^L h_l^{(\beta, \sigma)} \mathbf{F}_1 \mathbf{D}_{\alpha_l}^{(\beta, \phi)} \mathbf{F}_{\alpha_l/\beta}^H \Lambda_{\lambda_l}^{(\sigma)} \\ &= \sum_{l=0}^L h_l^{(\beta, \sigma)} \mathbf{H}_{F,l}^{(\beta, \phi)} \Lambda_{\lambda_l}^{(\sigma)}, \end{aligned} \quad (15)$$

with $\mathbf{H}_{F,l}^{(\beta, \phi)} = \mathbf{F}_1 \mathbf{D}_{\lambda_l}^{(\beta, \phi)} \mathbf{F}_{\alpha_l/\beta}^H$ being its l th component, whose (m, k) th entry is specified as

$$\begin{aligned} [\mathbf{H}_{F,l}^{(\beta, \phi)}]_{m,k} &= \frac{1}{K} \sum_{n=0}^{K-1} e^{-j2\pi \frac{mn}{K}} e^{j2\pi \omega \frac{\alpha_l - 1 + \phi}{\beta} \frac{n}{K}} e^{j2\pi \frac{\alpha_l}{\beta} \frac{nk}{K}} \\ &= \frac{1}{K} \sum_{n=0}^{K-1} e^{-jn \frac{2\pi}{K} ((m-k) - (\xi_{l,F1} k + \xi_{l,F2}))} \\ &= e^{-j \frac{(K-1)\pi}{K} ((m-k) - (\xi_{l,F1} k + \xi_{l,F2}))} \times \\ &\quad \frac{\text{sinc}((m-k) - (\xi_{l,F1} k + \xi_{l,F2}))}{\text{sinc}(\frac{1}{K} ((m-k) - (\xi_{l,F1} k + \xi_{l,F2})))}, \end{aligned} \quad (16)$$

where $\xi_{l,F1} = \frac{\alpha_l - \beta}{\beta}$ and $\xi_{l,F2} = \frac{\alpha_l - 1 + \phi}{\beta} \omega$ with $\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$.

It is obvious from (16) that in the absence of Doppler effects, i.e., $\alpha_l = 1$ for $l = 0, 1, \dots, L$, no rescaling and frequency synchronization is necessary, hence $\beta = 1$ and $\phi = 0$, which leads to a diagonal $\mathbf{H}_{F,l}^{(1,0)}$ with $[\mathbf{H}_{F,l}^{(\beta, \phi)}]_{m,k} = \delta_{m-k}$. In another special case where $\alpha_l \equiv \alpha$ for

$l = 0, 1, \dots, L$, we can also enforce a diagonal $\mathbf{H}_{F,l}^{(\beta, \phi)}$ by letting $\beta = \alpha$ and $\phi = 1 - \alpha$, a scenario considered in, e.g., [13]. For a realistic wideband LTV channel, however, the channel energy distribution in $\mathbf{H}_{F,l}^{(\beta, \phi)}$ is governed by a Dirichlet kernel, where the center of this Dirichlet kernel is offset by

$$\Delta_{F,l}^{(\beta, \phi)}(k) = \langle \xi_{l,F1} k + \xi_{l,F2} \rangle. \quad (17)$$

Clearly, such an offset is not only dependent on the Doppler spread α and the carrier

frequency f_c , but also on the subcarrier frequency $f_k = k\Delta f$. The dependence of the signal energy offset on the subcarrier index is unique to wideband channels, and is also referred to as nonuniform Doppler shifts in [13]. In contrast, the frequency offset for narrowband channels is statistically identical for all the subcarriers [2].

The Dirichlet kernel in (16) also suggests that the signal energy is mostly concentrated in subcarrier $k + \Delta_{F,l}^{(\beta,\phi)}(k)$ and its nearby subcarriers, and decays fast in subcarriers farther away. To appreciate how fast the signal energy decays, let us introduce $B_{F,l}^{(\beta,\phi)}(k)$ to quantify the number of subcarriers where most of the energy of b_k is located, which can thus be viewed as the bandwidth of $\mathbf{H}_{F,l}^{(\beta,\phi)}$ along its k th column. $B_{F,l}^{(\beta,\phi)}(k)$ is obtained as the smallest B for which

$$\begin{aligned} & \sum_{m=k+\Delta_{F,l}^{(\beta,\phi)}(k)-B}^{k+\Delta_{F,l}^{(\beta,\phi)}(k)+B} \left| [\mathbf{H}_{F,l}^{(\beta,\phi)}]_{m,k} \right|^2 > \gamma \sum_{m=0}^{K-1} \left| [\mathbf{H}_{F,l}^{(\beta,\phi)}]_{m,k} \right|^2 \\ \Leftrightarrow & \sum_{m=k+\Delta_{F,l}^{(\beta,\phi)}(k)-B}^{k+\Delta_{F,l}^{(\beta,\phi)}(k)+B} \left| \frac{\text{sinc}(\pi((m-k) - (\xi_{l,F1}k + \xi_{l,F2})))}{\text{sinc}(\frac{\pi}{K}((m-k) - (\xi_{l,F1}k + \xi_{l,F2})))} \right|^2 \\ & > \gamma \sum_{m=0}^{K-1} \left| \frac{\text{sinc}(\pi((m-k) - (\xi_{l,F1}k + \xi_{l,F2})))}{\text{sinc}(\frac{\pi}{K}((m-k) - (\xi_{l,F1}k + \xi_{l,F2})))} \right|^2, \end{aligned} \quad (18)$$

where γ is a positive threshold no larger than 1. In the left plot of Figure 2, the relationship between $\max_k B_{F,l}^{(\beta,\phi)}(k)$ and γ for the case $\beta = 1$ and $\alpha = 0$ (no resampling and frequency synchronization) is plotted. It is clear that most of the signal energy of b_k is captured within a limited bandwidth. For example, with a bandwidth $\max_k B_{F,l}^{(1,0)}(k) = 5$, roughly 98 % of the signal energy of b_k is captured. Notably, this bandwidth is almost independent of $\xi_{l,F1}$ and $\xi_{l,F2}$ as suggested by the left plot of Figure 2.

Figure 2 Bandwidth of $\mathbf{H}_{F,l}^{(\beta,\phi)}$ and $\mathbf{H}_{T,l}^{(\beta,\sigma)}$

Since each $\mathbf{H}_{F,l}^{(\beta,\phi)}$ is roughly banded, it is therefore reasonable to approximate $\mathbf{H}_F^{(\beta,\phi,\sigma)}$, which is a weighted sum of different $\mathbf{H}_{F,l}^{(\beta,\phi)}$ matrices, also as banded. As an example, we plot in Figure 3 the structure of $\mathbf{H}_{F,l}^{(\beta,\phi)}$, where we assume that there are in total two paths. Obviously, the approximate bandwidth of $\mathbf{H}_F^{(\beta,\phi,\sigma)}$ at the k th column, denoted as $B_F^{(\beta,\phi)}(k)$, is

$$\begin{aligned} B_F^{(\beta,\phi)}(k) &= \\ & \max_l \left(k + \Delta_{F,l}^{(\beta,\phi)}(k) + B_{F,l}^{(\beta,\phi)}(k) \right) - \min_l \left(k + \Delta_{F,l}^{(\beta,\phi)}(k) - B_{F,l}^{(\beta,\phi)}(k) \right) \\ & \approx \max_l \left(\Delta_{F,l}^{(\beta,\phi)}(k) \right) - \min_l \left(\Delta_{F,l}^{(\beta,\phi)}(k) \right) + 2 \max_l \left(B_{F,l}^{(\beta,\phi)}(k) \right), \end{aligned} \quad (19)$$

which is independent of σ . We refer the reader to Figure 3 for the physical meaning of the notations. It is important to underscore that since the bandwidth $B_F^{(\beta,\phi)}(k)$ is dependent on the

subcarrier index k , the boundaries of the band are not parallel to each other as in the narrowband case. A banded approximation of the channel matrix is crucial to many low-complexity equalizers, e.g., [5–7, 18]. The equalizer considered in this article will also adopt this approximation to reduce the complexity. More specifically, we first define a matrix $\mathbf{B}_F^{(\beta, \phi)}$, whose (m, k) th entry is equal to 1 if $\min_l \left(k + \Delta_{F,l}^{(\beta, \phi)}(k) - B_{F,l}^{(\beta, \phi)}(k) \right) \leq m \leq \max_l \left(k + \Delta_{F,l}^{(\beta, \phi)}(k) + B_{F,l}^{(\beta, \phi)}(k) \right)$, and 0 otherwise, and we then consider the matrix

$$\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)} = \mathbf{B}_F^{(\beta, \phi)} \odot \mathbf{H}_F^{(\beta, \phi, \sigma)} \quad (20)$$

as the banded approximation of $\mathbf{H}_F^{(\beta, \phi, \sigma)}$.

With the banded approximation, let us rewrite (14) as

$$\mathbf{r}_F^{(\beta, \phi, \sigma)} = \tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)} \mathbf{b} + \bar{\mathbf{v}}_F^{(\beta, \phi, \sigma)}, \quad (21)$$

where $\bar{\mathbf{v}}_F^{(\beta, \phi, \sigma)} = \left(\mathbf{H}_F^{(\beta, \phi, \sigma)} - \tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)} \right) \mathbf{b}$.

Figure 3 Illustration of the FD matrix $\mathbf{H}_F^{(\beta, \phi, \sigma)}$ for two paths

The above analysis can also be applied in the time domain in an analogous manner. See Appendix 3 for the details. Here we only want to highlight that, different from the energy distribution in the FD channel matrix which is influenced by the rescaling factor β and the phase-shift factor ϕ [c.f. $\xi_{l,F1}$ and $\xi_{l,F2}$ in (16)], the energy distribution in the TD channel matrix is affected by the rescaling factor β and the time-shift factor σ [c.f. $\xi_{l,T1}$ and $\xi_{l,T2}$ in (39)]. However, similarly as the FD channel matrix, we can also understand from the right subplot of Figure 2 that $\mathbf{H}_{T,l}^{(\beta, \sigma)}$ is roughly banded along the l th path in the time domain, and so is the overall time-domain channel matrix $\mathbf{H}_T^{(\beta, \phi, \sigma)}$.

Channel equalization scheme

Let us now focus on the channel frequency-domain equalization, which is depicted in Figure 4. In this figure, it is clear that, prior to the equalization, we propose an optimum resampling operation to achieve $(\beta, \phi, \sigma) = (\beta_{F,*}, \phi_{F,*}, \sigma_{F,*})$, which is different from [14, 19] as mentioned previously. Specifically, the resampling method proposed in [14] only considers the rescaling parameter β while [19] ignores the time-shift parameter σ . Afterwards, the banded matrix $\tilde{\mathbf{H}}_F^{(\beta_{F,*}, \phi_{F,*}, \sigma_{F,*})}$ is adopted to approximate $\mathbf{H}_F^{(\beta_{F,*}, \phi_{F,*}, \sigma_{F,*})}$ according to the approach mentioned in the last section. Our banded method induces a non-parallel bandwidth structure which is different from the banded approach used in narrowband OFDM systems [5–7, 21]. In order to speed up the convergence of the iterative equalization, we then design a diagonal preconditioner to improve the condition of this banded matrix. It is noteworthy here that our preconditioner design is adapted from [17, 18] to enhance its suitability for our MSML scenario. Finally, iterative equalization is proposed on the preconditioned channel matrix. Although we choose the CG method in this article, other

iterative methods can also be applied, such as the LSQR algorithm [22].

Figure 4 Depiction of our equalization scheme

Additionally, we would like to highlight that just as a single-carrier channel can be equalized in the frequency domain, it is also possible to equalize an OFDM channel in the time domain. Due to the similarity, we again refer the reader to Appendix 3 for a detailed mathematical derivation of the time-domain method. The question in which domain the wideband channel should be equalized, shall be addressed in the following section.

Iterative equalization

To better motivate the other components of our equalization scheme, we first introduce the channel equalization method itself. A zero-forcing equalizer in the frequency domain is considered, given by

$$\hat{\mathbf{b}} = \left(\bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)H} \bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)} \right)^{-1} \bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)H} \mathbf{r}_{\mathbf{F}}^{(\beta, \phi, \sigma)}, \quad (22)$$

where $\hat{\mathbf{b}}$ is the obtained estimate of \mathbf{b} . Because the original channel matrix $\mathbf{H}_{\mathbf{F}}^{(\beta, \phi, \sigma)}$ is a full matrix, its inversion inflicts a complexity of $\mathcal{O}(K^3)$ and is thus not desired for a practical system. To lower the complexity, $\mathbf{H}_{\mathbf{F}}^{(\beta, \phi, \sigma)}$ has been replaced by the banded approximation $\bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)}$ in (22).

Besides, the matrix inversion in (22) will be implemented iteratively using the CG algorithm. An advantage of using CG rather than inverting the matrix directly is that the resulting data estimates yielded by CG are always constrained in the Krylov subspace, making its performance less susceptible to the spectral distribution of $\bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)}$. In practice, a truncated CG, which halts the algorithm after a limited number of iterations, is desired to further reduce the complexity. It is well-known that the convergence of the CG algorithm can be accelerated by applying preconditioning on $\bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)}$ [16, 17, 23]. With $\mathbf{C}_{\mathbf{F}}$ denoting such a preconditioner, the I/O relationship given in (21) in the noiseless case can be rewritten as

$$\begin{aligned} \mathbf{r}_{\mathbf{F}}^{(\beta, \phi, \sigma)} &= \left(\bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)} \mathbf{C}_{\mathbf{F}} \right) \left(\mathbf{C}_{\mathbf{F}}^{-1} \mathbf{b} \right) \\ &= \bar{\mathbf{H}}_{\mathbf{FC}}^{(\beta, \phi, \sigma)} \mathbf{b}_{\mathbf{C}} \end{aligned} \quad (23)$$

from which an estimate of $\mathbf{b}_{\mathbf{C}} = \mathbf{C}_{\mathbf{F}}^{-1} \mathbf{b}$ is first obtained by applying CG on the preconditioned matrix $\bar{\mathbf{H}}_{\mathbf{FC}}^{(\beta, \phi, \sigma)} = \bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)} \mathbf{C}_{\mathbf{F}}$. Afterwards, $\hat{\mathbf{b}} = \mathbf{C}_{\mathbf{F}} \hat{\mathbf{b}}_{\mathbf{C}}$ is computed to obtain the final data estimates. For details about our CG equalization, see Appendix 4.

The optimal design of $\mathbf{C}_{\mathbf{F}}$ can be exhaustive [23]. Inspired by [17], we find our preconditioner by minimizing a cost function based on the Frobenius norm, which clusters most of the eigenvalues of $\bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)} \mathbf{C}_{\mathbf{F}}$ around 1 with the exception of a few outliers. Further, observing that the design of $\mathbf{C}_{\mathbf{F}}$ itself, as well as the operation of $\bar{\mathbf{H}}_{\mathbf{F}}^{(\beta, \phi, \sigma)} \mathbf{C}_{\mathbf{F}}$, inflicts an

additional complexity, a common approach is to impose a sparse structure on \mathbf{C}_F , e.g., diagonal [17] as

$$\mathbf{C}_F = \text{diag}\{[c_{F,0}, c_{F,1}, \dots, c_{F,K-1}]^T\}. \quad (24)$$

Diagonal preconditioning

In this section, we will show that the normal approach to design the diagonal preconditioner as described in [17] will not necessarily cluster eigenvalues around one. To realize this, let us consider the diagonal preconditioner $\mathbf{C}_{F,\star}$ that minimizes the cost function in the Frobenius norm [17] given by

$$\mathbf{C}_{F,\star} = \arg \min_{\mathbf{C}_F} \left\| \bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)} \mathbf{C}_F - \mathbf{I}_{K \times K} \right\|_{\text{Fro}}^2$$

which leads to

$$\begin{aligned} c_{F,k,\star} &= \arg \min_{c_{F,k}} \left\| \bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)} c_{F,k} \mathbf{e}_k - \mathbf{e}_k \right\|_2^2, \\ &= \frac{[\bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)}]_{k,k}^*}{\left\| \bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)} \mathbf{e}_k \right\|_2^2}, \end{aligned} \quad (25)$$

where \mathbf{e}_k is the k th column of the identity matrix.

One problem of the above diagonal preconditioner designed by (25) is that the eigenvalues may, in some situations, tend to cluster around zero instead of one, with the consequence that the condition number of the preconditioned channel matrix increases considerably. To understand this, assume there exists a $\epsilon_1 > 0$ such that

$$\left\| \bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)} \mathbf{e}_{kC_{F,k}} - \mathbf{e}_k \right\|_2^2 \leq \epsilon_1^2, \quad (26)$$

for $k = \{0, 1, \dots, K-1\}$. At the same time, assume there exists a $\epsilon_0 > 0$ such that

$$\left\| \bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)} \mathbf{e}_{kC_{F,k}} \right\|_2^2 \leq \epsilon_0^2 \quad (27)$$

for $k \in \{0, \dots, K-1\}$.

If we denote the k th eigenvalue of the preconditioned channel matrix $\bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)} \mathbf{C}_F$ as μ_k , (27) indicates that (for details see Appendix 5)

$$\sum_{k=0}^{K-1} |\mu_k|^2 \leq K\epsilon_0^2$$

which means that all μ_k 's lie inside a disk of radius $\sqrt{K}\epsilon_0$ centered around zero. Similarly,

from (26) we have

$$\sum_{k=0}^{K-1} |\mu_k - 1|^2 \leq K\epsilon_1^2$$

which implies that all μ_k 's at the same time lie inside a disk of $\sqrt{K}\epsilon_1$ centered around one. It is clear that if $\epsilon_0 < \epsilon_1$, then minimizing $\left\| \tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)} \mathbf{C}_F - \mathbf{I}_{K \times K} \right\|_{\text{Fro}}^2$ will at the same time minimize the Frobenius norm $\left\| \tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)} \mathbf{C}_F \right\|_{\text{Fro}}^2$ itself, making the eigenvalues more clustered around zero rather than one.

With $c_{F,k,\star}$ defined in (25), we can show that

$$\epsilon_1 = \max_k \frac{\sum_{m=0}^{K-1} |[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)}]_{m,k}|^2 - |[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)}]_{k,k}|^2}{\sum_{m=0}^{K-1} |[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)}]_{m,k}|^2}, \quad (28)$$

and

$$\epsilon_0 = \max_k \frac{|[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)}]_{k,k}|^2}{\sum_{m=0}^{K-1} |[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)}]_{m,k}|^2}. \quad (29)$$

Obviously, if $|[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)}]_{k,k}|^2 < \sum_{m=0}^{K-1} |[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)}]_{m,k}|^2$, for $k = 0, \dots, K-1$, then the optimal diagonal preconditioner will cluster the eigenvalues in a “wrong” area. This case arises when the sum of the off-diagonal power in each column is higher than the power on the diagonal. Such a situation could occur in multi-scale channels where significant channel power is located on off-diagonal entries as we argued in the previous section (see Figure 3 for instance). In the upper-left plot of Figure 5, the eigenvalues of such a matrix, with and without preconditioning, are displayed on a complex plane. It can be seen that diagonal preconditioning indeed clusters the eigenvalues around zero rather than one.

Figure 5 Left plots: eigenvalues with and without preconditioning; Right plots: convergence performance with and without preconditioning; FD matrix for top two plots corresponds to the original channel, FD matrix for bottom two plots is obtained after our optimum resampling; The MSML channel is set according to Table 1

To evaluate the impact of such a preconditioner on the convergence of CG, we compute the mean squared error (MSE) as

$$\text{MSE} = \frac{\|\mathbf{b} - \hat{\mathbf{b}}^{(i)}\|^2}{\|\hat{\mathbf{b}}\|^2}, \quad (30)$$

with $\hat{\mathbf{b}}^{(i)}$ being the result obtained at the i th iteration of our CG equalization as mentioned in Appendix 4. In the top-right plot of Figure 5, it is clear that the CG convergence with such a diagonal preconditioner is even worse than without any preconditioning. This illustrates that the diagonal preconditioning defined in (25) may not always yield a better performance than without preconditioning, as opposed to what is claimed in [17, 18]. Using a more complex structured preconditioner can avoid this, which is, however, not desired due to complexity and

Table 1 Channel I: a frequency-domain case

Channel I	path	scale α_l	delay λ_l	path gain \bar{h}_l
$(T = 0.2 \text{ ms}$	$l = 0$	1.0150	0.00	0 dB
$\omega = 256$	$l = 1$	1.0154	10.15	-3 dB
$K = 128)$	$l = 2$	1.0201	20.40	-5 dB
Resampl.	Orig.	$(\beta, \phi, \sigma) = (1, 0, 0)$		
Para.	Resampl.	$(\beta_{F,\star}, \phi_{F,\star}, \sigma_{F,\star}) = (1.0150, -0.0150, -15.00)$		
	Orig./no precondition.	4.26×10^5		
Cond. Num.	Orig./with precondition.	1.19×10^6		
for FD	Resampl./no precondition.	23.36		
	Resampl./with precondition.	7.17		
FD	Orig.	$\rho_F^{(1,0,0)} = 0.0021$		
Ratio	Resampl.	$\rho_F^{(\beta_{F,\star}, \phi_{F,\star}, \sigma_{F,\star})} = 0.9279$		

implementation considerations.

To alleviate this problem, we adapt the diagonal preconditioner in (24) and (25) as follows

$$c_{F,k,\star} = \begin{cases} \frac{[\bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)}]_{k,k}^*}{\|\bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)} \mathbf{e}_k\|_2^2}, & \text{if } \zeta_{\text{FC}}^{(\beta,\phi,\sigma)}(k) \geq 1 \\ 1, & \text{otherwise} \end{cases} \quad (31)$$

where

$$\zeta_{\text{FC}}^{(\beta,\phi,\sigma)}(k) = \frac{\left| [\bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)}]_{k,k} \right|^2}{\sum_{m=0, m \neq k}^{K-1} \left| [\bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)}]_{m,k} \right|^2}. \quad (32)$$

In Section ‘Optimal resampling’, we will show how to enhance (32) with a higher probability by means of optimal resampling.

Optimal resampling

From the previous subsections, we understand that the effectiveness of a diagonal preconditioner depends on the energy distribution of the channel matrix. It is desired that the channel matrix should have most of its energy concentrated on the main diagonal. The analysis in Section ‘Discrete data model’ learns that the resampling operation (β, ϕ, σ) plays an important role in governing the energy distribution of the channel matrix, and so far we have left (β, ϕ, σ) open for choice. Recall that resampling is a standard step taken in many wideband LTV communication systems to compensate for the Doppler effect. For example, optimizing β is considered in [14], while β and ϕ are jointly optimized in [21]. In this sense, the optimal resampling proposed in this article can be considered as a generalization of [14, 21].

Next, we shall discuss how to jointly optimize the resampling parameters (β, ϕ, σ) . Focusing

on the FD matrix $\mathbf{H}_F^{(\beta, \phi, \sigma)}$, we desire $\left| \left[\mathbf{H}_F^{(\beta, \phi, \sigma)} \right]_{k,k} \right|^2 > \sum_{m \neq k} \left| \left[\mathbf{H}_F^{(\beta, \phi, \sigma)} \right]_{m,k} \right|^2$ for all $k \in \{0, 1, \dots, K-1\}$. However, satisfying the above condition for each index k individually is expensive. As a relaxation, we practically seek $\sum_k \left| \left[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)} \right]_{k,k} \right|^2 > \sum_k \sum_{m \neq k} \left| \left[\tilde{\mathbf{H}}_F^{(\beta, \phi, \sigma)} \right]_{m,k} \right|^2$.

To this end, let us denote the diagonal energy ratio as

$$\rho_F^{(\beta, \phi, \sigma)} = \frac{\sum_{k=0}^{K-1} \left| \left[\mathbf{H}_F^{(\beta, \phi, \sigma)} \right]_{k,k} \right|^2}{\sum_{k=0}^{K-1} \sum_{m=0}^{K-1} \left| \left[\mathbf{H}_F^{(\beta, \phi, \sigma)} \right]_{m,k} \right|^2}, \quad (33)$$

and define our resampling operation by solving

$$(\beta_{F,\star}, \phi_{F,\star}, \sigma_{F,\star}) = \arg \max_{\beta, \phi, \sigma} \rho_F^{(\beta, \phi, \sigma)}, \quad (34)$$

which leads to the maximal ratio $\rho_F^{(\beta_{F,\star}, \phi_{F,\star}, \sigma_{F,\star})}$. One can also explain this resampling as minimizing the total amount of ICI in the frequency domain.

Since the energy governing mechanism is determined by the sinc function as indicated in (16), we can equivalently rewrite (34) by only maximizing the diagonal energy of $\mathbf{H}_F^{(\beta, \phi, \sigma)}$ as

$$\begin{aligned} (\beta_{F,\star}, \phi_{F,\star}, \sigma_{F,\star}) &= \arg \max_{\beta, \phi, \sigma} \sum_{k=0}^{K-1} \left| \left[\mathbf{H}_F^{(\beta, \phi, \sigma)} \right]_{k,k} \right|^2 \\ &= \arg \max_{\beta, \phi, \sigma} \sum_{k=0}^{K-1} \left| \sum_{l=0}^L h_l^{(\beta, \sigma)} e^{-j \frac{(K-1)\pi}{K} (\xi_{l,F1} k + \xi_{l,F2})} \times \right. \\ &\quad \left. \frac{\text{sinc}(\xi_{l,F1} k + \xi_{l,F2})}{\text{sinc}(\frac{1}{K} (\xi_{l,F1} k + \xi_{l,F2}))} \times e^{j 2\pi (\lambda_l + \sigma) \frac{k}{K}} \right|^2, \end{aligned} \quad (35)$$

where again $\xi_{l,F1} = \frac{\alpha_l - \beta}{\beta}$ and $\xi_{l,F2} = \frac{\alpha_l - 1 + \phi}{\beta} \omega$. It is noteworthy that all three parameters, β , ϕ and σ , play a role in (35), indicating that separately considering one or two parameters as in [14, 21] might lead to a local maximum.

To illustrate our resampling approach in the frequency domain, we consider the channel example specified in Table 1, where we also compare the properties of the resampled FD channel (i.e., the condition number and diagonal power ratio of the channel matrix) with the original MSML FD channel. A geometric interpretation may help to understand our resampling operation since β rotates the FD matrix through $\xi_{l,F1} = \frac{\alpha_l - \beta}{\beta}$, ϕ shifts the FD matrix through $\xi_{l,F2} = \frac{\alpha_l - 1 + \phi}{\beta} \omega$ in (16), and σ influences the phase of each element in (35). The joint effect of these actions maximizes the matrix diagonal energy. The yielded resampling $(\beta_{F,\star}, \phi_{F,\star}, \sigma_{F,\star}) = (1.015, -0.015, -15.00)$ corresponds to a maximal diagonal power ratio $\rho_F^{(\beta_{F,\star}, \phi_{F,\star}, \sigma_{F,\star})} = 0.9279$. We underscore that the condition number is already

significantly reduced, solely by the optimum resampling, from 4.26×10^5 to 23.36. In comparison, the resampling method proposed in [14] yields $(\beta, \phi, \sigma) = (1.016, 0, 0)$ and $\rho_F^{(1.016, 0, 0)} = 0.3623$. Its corresponding condition number is 432.78, which is larger than our condition number after resampling. This is not surprising since the criterion adopted in [14] focuses only on minimizing the aggregate errors between the multi-scale channel and its single-scale approximation, which is different from our criterion.

In the lower plots of Figure 5, we show the effectiveness of diagonal preconditioning applied to the resampled channel in Table 1. It is clear that, after our resampling procedure, the diagonal preconditioner clusters the eigenvalues of the preconditioned FD channel matrix closer to one than without preconditioning, which further reduces the condition number from 23.36 to 7.17. In contrast, without optimal resampling, the preconditioner “wrongly” pushes the eigenvalues closer to zero. In this case, the matrix condition number increases from 4.26×10^5 to 1.19×10^6 , and hence the CG equalizer performs even worse than without preconditioning as shown in the top two plots of Figure 5.

Similarly, we can show that optimal resampling can also improve the performance of the CG in the time domain, for which we just provide Table 2 and Figure 6 here due to space limitations. From them, we can make the same observations as from Table 1 and Figure 5 for the frequency domain case.

Table 2 Channel II: a time-domain case

Channel II	path	scale α_l	delay λ_l	path gain \bar{h}_l
$(T = 0.2 \text{ ms}$	$l = 0$	1.0161	1.00	0 dB
$\omega = 640$	$l = 1$	1.0180	0.80	-3 dB
$K = 128)$	$l = 2$	1.0244	3.00	-5 dB
Resampl.	Orig.	$(\beta, \phi, \sigma) = (1, 0, 0)$		
Para.	Resampl.	$(\beta_{T,*}, \phi_{T,*}, \sigma_{T,*}) = (1.0160, -0.0210, -1.00)$		
Cond. Num. for TD	Orig./no precondition.	2.54×10^4		
	Orig./with precondition.	7.37×10^4		
	Resampl./no precondition.	50.78		
	Resampl./with precondition.	15.03		
TD	Orig.	$\rho_F^{(1,0,0)} = 0.0021$		
Ratio	Resampl.	$\rho_F^{(\beta_{F,*}, \phi_{F,*}, \sigma_{F,*})} = 0.9168$		

Figure 6 Left plots: eigenvalues with and without preconditioning; Right plots: convergence performance with and without preconditioning; TD matrix for top two plots corresponds to the original channel, TD matrix for bottom two plots is obtained after our optimum resampling; The MSML channel is set according to Table 2.

Frequency-domain or time-domain equalization?

In the previous sections, we showed that the equalization of an OFDM channel can be implemented in either the frequency or the time domain. With the CG algorithm specified in Appendix 4, it is clear that the cost of equalization in the frequency domain will be

upper-bounded by $\mathcal{O}(B_F^{(\beta,\phi)} K)$ with $B_F^{(\beta,\phi)} = \max_k B_F^{(\beta,\phi)}(k)$ for each CG iteration. Likewise, the cost of equalization in the time domain will be upper-bounded by $\mathcal{O}(B_T^{(\beta,\sigma)} K)$ with $B_T^{(\beta,\sigma)} = \max_m B_T^{(\beta,\sigma)}(m)$. By assuming that the number of CG iterations is predetermined and identical in both domains, we can use the ratio $B_F^{(\beta,\phi)} / B_T^{(\beta,\sigma)}$ as a criterion to choose in which domain the equalization will be realized in order to minimize the complexity.

However, the evaluation of $B_F^{(\beta,\phi)} / B_T^{(\beta,\sigma)}$ is cumbersome and lacks the insight of the channel physics. For simplicity reasons, we equivalently consider the proportion given by

$$\epsilon = \frac{B_F^{(\beta,\phi)} - 2B_{\text{rul}}}{B_T^{(\beta,\sigma)} - 2B_{\text{rul}}} = \frac{\max_k \left(\max_l (\Delta_{F,l}^{(\beta,\phi)}(k)) - \min_l (\Delta_{F,l}^{(\beta,\phi)}(k)) \right)}{\max_m \left(\max_l (\Delta_{T,l}^{(\beta,\sigma)}(m)) - \min_l (\Delta_{T,l}^{(\beta,\sigma)}(m)) \right)}, \quad (36)$$

where we reasonably assume $B_{\text{rul}} = \max_{l,k} B_{F,l}(k) \approx \max_{l,m} B_{T,l}(m)$ [see Figure 2]. One may argue that the above evaluation is still cumbersome. However, if a realistic channel allows us to assume, for all $l \in \{1, 2, \dots, L\}$, that

$$\begin{aligned} |\alpha_l - \beta_{F,\star}| / \beta_{F,\star} &\ll 1/(K-1), \\ |\alpha_l - \beta_{T,\star}| / \beta_{T,\star} &\ll 1/(K-1), \end{aligned}$$

which indicates that the Doppler scale spread is well-limited, it follows that $\max_{l,k} (|\xi_{l,F1}|k) \ll 1$ and $\max_{l,m} (|\xi_{l,T1}|m) \ll 1$. In other words, $\Delta_{F,l}^{(\beta_{F,\star}, \phi_{F,\star})}(k) \approx \langle \xi_{l,F2} \rangle$ and $\Delta_{T,l}^{(\beta_{T,\star}, \sigma_{T,\star})}(m) \approx \langle \xi_{l,T2} \rangle$, both of which are independent of the symbol index. With these assumptions, ϵ can further be simplified as

$$\begin{aligned} \epsilon &\approx \frac{\max_l (\langle \xi_{l,F2} \rangle) - \min_l (\langle \xi_{l,F2} \rangle)}{\max_l (\langle \xi_{l,T2} \rangle) - \min_l (\langle \xi_{l,T2} \rangle)} \\ &= \frac{\left\langle (\max_l (\alpha_l) - 1 + \phi_{F,\star}) \frac{\omega}{\beta_{F,\star}} \right\rangle - \left\langle (\min_l (\alpha_l) - 1 + \phi_{F,\star}) \frac{\omega}{\beta_{F,\star}} \right\rangle}{\left\langle \max_l (\alpha_l (\lambda_l + \sigma_{T,\star})) \right\rangle - \left\langle \min_l (\alpha_l (\lambda_l + \sigma_{T,\star})) \right\rangle} \end{aligned}$$

which suggests that if the maximum difference between the Doppler shifts of each path (i.e., $\frac{\alpha_l - 1}{\beta} \omega$) is smaller than the maximum difference between the time shifts of each path (i.e., $\alpha_l \lambda_l$), then equalization should be realized in the frequency domain; otherwise, a time-domain approach will be preferred. A similar conclusion has been made for narrowband systems [24], though its extension to wideband systems is not straightforward as shown above.

To illustrate the above idea, we again use the channel examples specified in Tables 1 and 2, respectively. We use $B_{\text{rul}} = 5$ to roughly capture $\gamma = 98\%$ of the channel energy in both domains where γ is introduced in (18). In this way, we have $\epsilon \approx 0.10 < 1$ for the channel in Table 1, while for the channel in Table 2, we have $\epsilon \approx 2.00 > 1$.

For both channels, we compare the equalization performance in different domains. OFDM with $K = 128$ subcarriers using QPSK is transmitted and the receiver is assumed to have perfect channel knowledge. We examine the bit error rate (BER) results of our CG equalization with a fixed CG iteration number (e.g., $i_{F,\max} = i_{T,\max} = 100$). We use different bandwidths for the banded approximation $\tilde{\mathbf{H}}_F^{(\beta_{F,*}, \phi_{F,*}, \sigma_{F,*})}$ and $\tilde{\mathbf{H}}_T^{(\beta_{T,*}, \phi_{T,*}, \sigma_{T,*})}$ during the equalization and the values for $(\beta_{F,*}, \phi_{F,*}, \sigma_{F,*})$ and $(\beta_{T,*}, \phi_{T,*}, \sigma_{T,*})$ have also been given in Tables 1 and 2, respectively. After our optimal resampling in either domain, the CG equalization is carried out using the appropriate preconditioner design.

The left subplot of Figure 7 plots the BER performance as a function of signal-to-noise ratio (SNR) for Channel I. Note that $(\beta_{F,*}, \phi_{F,*}, \sigma_{F,*}) = (1.015, -0.015, -15)$ and $(\beta_{T,*}, \phi_{T,*}, \sigma_{T,*}) = (1.015, -0.016, 0.00)$ for this channel. It can be seen that the performance of the FD equalizer (FDE) based on $\tilde{\mathbf{H}}_F^{(\beta_{F,*}, \phi_{F,*}, \sigma_{F,*})}$ outperforms the TD equalizer (TDE) based on $\tilde{\mathbf{H}}_T^{(\beta_{T,*}, \phi_{T,*}, \sigma_{T,*})}$ using the same bandwidth $B_F^{(\beta_{F,*}, \phi_{F,*})} = B_T^{(\beta_{T,*}, \sigma_{T,*})}$. In other words, FDE is more attractive than TDE in this case.

Figure 7 BER versus SNR for the two channels given in Tables 1 and 2

The BER performance for Channel II is illustrated in the right subplot of Figure 7, where the optimal resampling parameters are $(\beta_{T,*}, \phi_{T,*}, \sigma_{T,*}) = (1.016, -0.021, -1)$ and $(\beta_{F,*}, \phi_{F,*}, \sigma_{F,*}) = (1.016, -0.016, -3)$. In this case, it is evident that the TD equalizer is more appealing.

These observations made for the channels in Tables 1 and 2 confirm our metric ϵ for determining which domain is more suitable for channel equalization. Additionally, we like to point out that, in either domain, with a larger bandwidth the BER performance of our CG equalization will be increased.

Numerical results

In this section, we randomly generate two different types of wideband channels as specified in Table 3: $\epsilon < 1$ (Case I) represents wideband LTV channels where the Doppler differences among the multipath are more pronounced than the delay differences; and $\epsilon > 1$ (Case II) is the case where the Doppler differences among the multipath are less pronounced than the delay differences. For all simulations, OFDM with $K = 128$ subcarriers is considered with QPSK. The wideband channels are assumed to have $L = 5$ paths, whose channel gains (i.e., \bar{h}_l 's) are modeled to be identically and independently distributed. The path delay (τ_l) is chosen as a random variable that has a uniform distribution within the range $[0, \tau_{\max}]$. Likewise, the path scale (α_l) is chosen as a random variable that obeys a uniform distribution within the range $[1, 1 + \alpha_{\text{sp}}]$ with α_{sp} the scale spread. For both cases, the receiver is assumed to have perfect channel knowledge and the cyclic extensions at the transmitter are $T_{\text{pre}} = 32T$ and $T_{\text{post}} = 10T$ which satisfy (6) and (7). In all simulations, a banded approximation of the channel matrix is adopted in both domains with the same bandwidth (e.g.,

$$B_F^{(\beta_F, \star, \phi_F, \star)} = B_T^{(\beta_T, \star, \sigma_T, \star)} = 11).$$

Table 3 Channel parameters.

Case 1: $\epsilon < 1$ $K = 128, \omega = 256$			Case 2: $\epsilon > 1$ $K = 128, \omega = 640$		
L	α_{sp}	τ_{\max}/T	L	α_{sp}	τ_{\max}/T
5	0.008	30.00	5	0.010	4.00

In Figure 8, the convergence of the CG equalization is plotted in terms of the bit error rate (BER) against the number of iterations at SNR = 30 dB for Case I. Since $\epsilon < 1$, frequency-domain equalization (FDE) is carried out. It is clear that the receiver, which simply adopts a diagonal preconditioner in (25) without resampling, performs worst. The performance is already considerably improved if optimal resampling is applied. Moreover the use of our preconditioner given by (31) boosts the performance even further.

Figure 8 BER versus number of iterations for Case I channels at SNR = 30 dB.

The proposed resampling and preconditioning method can also benefit from other Krylov-based algorithms. For instance, the LSQR algorithm exploiting a full channel matrix is studied in [18]. Note that [18] focuses on a narrowband LTV system where no resampling is required. Further, the preconditioner given in [18] is based on a truncated basis expansion model (BEM) which is usually used for the approximation of a narrowband time-varying channel. Because it is not clear whether such a truncated BEM is still suitable for a wideband LTV channel, in order to emulate a similar approach as in [18] for constructing the preconditioner, we utilize a (trivial) full-order critically-sampled complex exponential BEM (the CCE-BEM [25]) in the simulation. The preconditioner in [18] then boils down to the inverse of the diagonal of the frequency-domain channel matrix, which is obviously sub-optimal in the Frobenius norm sense. Consequently, it is no surprise that directly applying the equalizer of [18] to wideband LTV channels yields a bad performance as shown in Figure 8. In comparison, the LSQR algorithm benefiting from the optimal resampling and our preconditioner renders the fastest convergence rate and lowest BER amongst all the equalization schemes. Of course, such an improved BER performance is achieved by leveraging the full channel matrix at the cost of a higher complexity, compared to our proposed method using banded matrices.

Figure 9 exhibits the BER versus SNR for the CG-based equalization schemes, where a truncated CG is used which halts at the 5th iteration. It can be seen in the figure that the equalizer leveraging the full channel matrix gives the best BER performance but inflicts more complexity. When using a banded channel matrix approximation, the frequency-domain approach performs much better than the time-domain approach because we have $\epsilon < 1$ for this type of channel. Additionally, the equalization approach in [14] is carried out and its performance is also shown in Figure 9. As we discussed earlier, the resampling operation in [14] is solely focused on the rescaling parameter ignoring the impact of frequency and time synchronization, which is therefore sub-optimal. Besides, the equalizer in [14] approximates the channel matrix to be diagonal (i.e., using a bandwidth of one for the banded matrices), and thus its performance becomes inferior in the presence of higher scale differences among the

multipath as in the tested channel here.

Figure 9 BER versus SNR for Case I channels.

The performance of the equalizers for Case II is depicted in Figure 10, where the significance of optimal resampling and our adapted preconditioner is again illustrated just like in Figure 8. Similarly, we can see that the LSQR algorithm in [18] also works well for this type of channel if optimal resampling and preconditioning are included.

Figure 10 BER versus number of iterations for Case II Channels at SNR = 30 dB.

Different from Case I, the channels of Case II are subject to a larger delay spread than a Doppler spread (i.e., $\epsilon > 1$). In this case, a time-domain equalizer will be more effective than its frequency-domain counterpart as validated in Figure 11. The equalizer in [14] yields a much worse performance than ours since the Doppler scale spread differences in this case are even higher than for Case I.

Figure 11 BER versus SNR for Case II channels.

Conclusion

In this article, we have discussed iterative equalization of wideband channels using the conjugate gradient (CG) algorithm for OFDM systems. The channel follows a multi-scale multi-lag (MSML) model, and suffers therefore from interferences in both the frequency domain and time domain. To lower the equalization complexity, the channel matrices are approximated to be banded in both domains. A novel method of optimal resampling is proposed, which is indispensable for wideband communications. A diagonal preconditioning technique, that accompanies the CG method to accelerate the convergence, has also been adapted to enhance its suitability. Experimental results have shown that our equalization scheme allows for a superior performance to those schemes based on a single-scale resampling method, without any resampling operation, or using a traditional preconditioning procedure. In addition, we gave a simple criterion to determine whether to use a frequency-domain or time-domain equalizer, depending on the channel situation, to obtain the best BER performance with the same complexity. Such a criterion is also validated by experiments.

Appendix 1

Detailed derivation of the discrete data model

Here we give the derivation of (9), assuming no noise is present. We start from (8) given by

$$\begin{aligned} r_n^{(\beta, \phi, \sigma)} &= r^{(\beta, \phi, \sigma)}(nT) \\ &= \frac{1}{\sqrt{\beta KT}} \sum_{k=0}^{K-1} b_k h_k \left(\frac{nT}{\beta} - \sigma T \right) e^{j2\pi(\phi f_c + f_k) \frac{nT}{\beta}} e^{-j2\pi f_k \sigma T} \end{aligned}$$

where $h_k(t)$ is defined in (5) and the embedded $u(t)$ in $h_k(t)$ is considered to be one for the concerned observation window as clarified in Remark 1.

Now, we substitute $h_k(t)$ to obtain

$$\begin{aligned} r_n^{(\beta, \phi, \sigma)} &= \frac{1}{\sqrt{\beta KT}} \sum_{k=0}^{K-1} b_k \times \sum_{l=0}^L \bar{h}_l \sqrt{\alpha_l} e^{-j2\pi(f_c + f_k)\alpha_l \tau_l} e^{j2\pi(\alpha_l - 1)(f_c + f_k)(\frac{nT}{\beta} - \sigma T)} \\ &\quad \times e^{j2\pi(\phi f_c + f_k) \frac{nT}{\beta}} e^{-j2\pi f_k \sigma T} \\ &= \frac{1}{\sqrt{\beta KT}} \sum_{k=0}^{K-1} b_k \times \sum_{l=0}^L \bar{h}_l \sqrt{\alpha_l} \times \left(e^{-j2\pi f_c \alpha_l \tau_l} e^{j2\pi(\alpha_l - 1)f_c(\frac{nT}{\beta} - \sigma T)} e^{j2\pi \phi f_c \frac{nT}{\beta}} \right) \\ &\quad \times \left(e^{-j2\pi f_k \alpha_l \tau_l} e^{j2\pi(\alpha_l - 1)f_k(\frac{nT}{\beta} - \sigma T)} e^{j2\pi f_k \frac{nT}{\beta}} e^{-j2\pi f_k \sigma T} \right) \\ &= \sum_{l=0}^L \left(\sqrt{\frac{\alpha_l}{\beta KT}} \bar{h}_l e^{-j2\pi f_c(\alpha_l \tau_l + (\alpha_l - 1)\sigma T)} \right) e^{j2\pi f_c \frac{(\alpha_l - 1 + \phi)nT}{\beta}} \left(\sum_{k=0}^{K-1} b_k e^{j2\pi f_k \frac{\alpha_l nT}{\beta}} \right) e^{-j2\pi f_k \alpha_l(\tau_l + \sigma T)} \\ &= \sum_{l=0}^L h_l^{(\beta, \sigma)} e^{j2\pi f_c \frac{(\alpha_l - 1 + \phi)nT}{\beta}} \times \left(\sum_{k=0}^{K-1} b_k e^{j2\pi f_k \frac{\alpha_l nT}{\beta}} \right) e^{-j2\pi f_k \alpha_l(\tau_l + \sigma T)}, \end{aligned}$$

where the channel coefficient is given by

$$h_l^{(\beta, \sigma)} = \sqrt{\frac{\alpha_l}{\beta KT}} \bar{h}_l e^{-j2\pi f_c(\alpha_l \tau_l + (\alpha_l - 1)\sigma T)}$$

Now, if we denote

$$\omega = \frac{f_c}{\Delta f}$$

for the normalized carrier frequency and

$$\lambda_l = \frac{\tau_l}{T}$$

for the normalized delay of the l th path, we have

$$r_n^{(\beta, \phi, \sigma)} = \sum_{l=0}^L h_l^{(\beta, \sigma)} e^{j2\pi \omega \frac{(\alpha_l - 1 + \phi)}{\beta} \frac{n}{K}} \times \left(\sum_{k=0}^{K-1} b_k e^{j2\pi \frac{\alpha_l}{\beta} \frac{nk}{K}} \right) e^{-j2\pi \alpha_l (\lambda_l + \sigma) \frac{k}{K}}$$

which gives (9).

Appendix 2

System model in the time domain and time-domain equalization

To derive the time-domain model, let us rewrite (10) as

$$\mathbf{r}_T^{(\beta, \phi, \sigma)} = \mathbf{H}_T^{(\beta, \phi, \sigma)} \mathbf{s}, \quad (37)$$

where $\mathbf{s} = \mathbf{F}_1^{-1} \mathbf{b}$, and $\mathbf{H}_T^{(\beta, \phi, \sigma)}$ stands for the time-domain (TD) channel matrix

$$\mathbf{H}_T^{(\beta, \phi, \sigma)} = \sum_{l=0}^L h_l^{(\beta, \sigma)} \mathbf{D}_{\alpha_l}^{(\beta, \phi)} \mathbf{H}_{T,l}^{(\beta, \sigma)} \quad (38)$$

with $\mathbf{H}_{T,l}^{(\beta, \sigma)} = \mathbf{F}_{\alpha_l/\beta}^H \Lambda_{\lambda_l}^{(\sigma)} \mathbf{F}_1$ being its l th component. The (m, k) th entry of $\mathbf{H}_{T,l}^{(\beta, \sigma)}$ is given by

$$\begin{aligned} \left[\mathbf{H}_{T,l}^{(\beta, \sigma)} \right]_{m,k} &= \frac{1}{K} \sum_{n=0}^{K-1} e^{j2\pi \frac{\alpha_l}{\beta} \frac{mn}{K}} e^{j2\pi \alpha_l (\lambda_l + \sigma) \frac{n}{K}} e^{-j2\pi \frac{nk}{K}} \\ &= e^{-j \frac{(K-1)\pi}{K} ((k-m) - (\xi_{l,T1}m + \xi_{l,T2}))} \times \\ &\quad \frac{\text{sinc}(\pi ((k-m) - (\xi_{l,T1}m + \xi_{l,T2})))}{\text{sinc}(\frac{\pi}{K} ((k-m) - (\xi_{l,T1}m + \xi_{l,T2})))}, \end{aligned} \quad (39)$$

where $\xi_{l,T1} = \frac{\alpha_l - \beta}{\beta}$ and $\xi_{l,T2} = \alpha_l (\lambda_l + \sigma)$.

Observing the analogy between (16) and (39), a similar interference analysis can be made on \mathbf{H}_T . By defining

$$\Delta_{T,l}^{(\beta, \sigma)}(m) = \langle \xi_{l,T1}m + \xi_{l,T2} \rangle, \quad (40)$$

we can introduce the symbol $B_{T,l}^{(\beta,\sigma)}(m)$ defined as [c.f. (18)]

$$\begin{aligned}
B_{T,l}^{(\beta,\sigma)}(m) &= \min\{B\}, \\
\text{s.t.} \quad &\sum_{k=m+\Delta_{T,l}^{(\beta,\sigma)}(m)-B}^{m+\Delta_{T,l}^{(\beta,\phi)}(m)+B} \left| \frac{\text{sinc}(\pi((k-m) - (\xi_{l,T1}m + \xi_{l,T2})))}{\text{sinc}(\frac{\pi}{K}((k-m) - (\xi_{l,T1}m + \xi_{l,T2})))} \right|^2 \\
&> \gamma \sum_{k=0}^{K-1} \left| \frac{\text{sinc}(\pi((k-m) - (\xi_{l,T1}m + \xi_{l,T2})))}{\text{sinc}(\frac{\pi}{K}((k-m) - (\xi_{l,T1}m + \xi_{l,T2})))} \right|^2,
\end{aligned} \tag{41}$$

which determines the index set of the data symbols that contribute the most to the m th received signal $[\mathbf{r}_T^{(\beta,\phi,\sigma)}]_m$ via the l th path. Note that $B_{F,l}^{(\beta,\phi)}(k)$ in (18) depends on the resampling factor β and the frequency shift factor ϕ , whereas $B_{T,l}^{(\beta,\sigma)}(m)$ in (18) depends on the resampling factor β and the time shift factor σ .

Similarly as in the frequency domain, we obtain a banded approximation of $\mathbf{H}_T^{(\beta,\phi,\sigma)}$ by introducing

$$B_T^{(\beta,\sigma)}(m) \approx \max_l \left(\Delta_{T,l}^{(\beta,\sigma)}(m) \right) - \min_l \left(\Delta_{T,l}^{(\beta,\sigma)}(m) \right) + 2\max_l \left(B_{T,l}^{(\beta,\sigma)}(m) \right) \tag{42}$$

and a selection matrix $\mathbf{B}_T^{(\beta,\sigma)}$, whose (m, k) th entry is equal to 1 if $\min_l \left(m + \Delta_{T,l}^{(\beta,\sigma)}(m) - B_{T,l}^{(\beta,\sigma)}(m) \right) \leq k \leq \max_l \left(m + \Delta_{T,l}^{(\beta,\sigma)}(m) + B_{T,l}^{(\beta,\sigma)}(m) \right)$, and 0 otherwise. Then the banded approximation of $\mathbf{H}_T^{(\beta,\phi,\sigma)}$ is obtained by

$$\bar{\mathbf{H}}_T^{(\beta,\phi,\sigma)} = \mathbf{B}_T^{(\beta,\sigma)} \odot \mathbf{H}_T^{(\beta,\phi,\sigma)}. \tag{43}$$

We can then rewrite (37) as

$$\mathbf{r}_T^{(\beta,\phi,\sigma)} = \bar{\mathbf{H}}_T^{(\beta,\phi,\sigma)} \mathbf{s} + \bar{\mathbf{v}}_T^{(\beta,\phi,\sigma)}, \tag{44}$$

where $\bar{\mathbf{v}}_T^{(\beta,\phi,\sigma)} = \left(\mathbf{H}_T^{(\beta,\phi,\sigma)} - \bar{\mathbf{H}}_T^{(\beta,\phi,\sigma)} \right) \mathbf{s}$.

The time-domain equalization can be presented in an analogous manner as in the frequency domain. Similar to its FD counterpart in (23), we here rewrite the noiseless case for (44) as

$$\begin{aligned}
\mathbf{r}_{TC}^{(\beta,\phi,\sigma)} &= \mathbf{C}_T \mathbf{r}_T^{(\beta,\phi,\sigma)} = \left(\mathbf{C}_T \bar{\mathbf{H}}_T^{(\beta,\phi,\sigma)} \right) \mathbf{s} \\
&= \bar{\mathbf{H}}_{TC}^{(\beta,\phi,\sigma)} \mathbf{s} = \bar{\mathbf{H}}_{TC}^{(\beta,\phi,\sigma)} \mathbf{F}_1^H \mathbf{b}
\end{aligned} \tag{45}$$

where $\mathbf{s} = \mathbf{F}_1^H \mathbf{b}$, \mathbf{C}_T is the preconditioner applied in the time domain and $\bar{\mathbf{H}}_{TC}^{(\beta,\phi,\sigma)} = \mathbf{C}_T \bar{\mathbf{H}}_T^{(\beta,\phi,\sigma)}$. We first estimate \mathbf{s} by applying the CG algorithm on $\mathbf{r}_{TC}^{(\beta,\phi,\sigma)}$ to invert $\bar{\mathbf{H}}_{TC}^{(\beta,\phi,\sigma)}$ iteratively, and afterwards we obtain $\hat{\mathbf{b}} = \mathbf{F}_1^H \hat{\mathbf{s}}$.

We highlight that the adopted diagonal preconditioner

$\mathbf{C}_{T,\star} = \text{diag}\{[c_{T,0,\star}, c_{T,1,\star}, \dots, c_{T,K-1,\star}]^T\}$ is defined in a similar manner as in the frequency domain. Specifically, we use

$$c_{T,m,\star} = \begin{cases} \frac{[\mathbf{H}_T^{(\beta,\phi,\sigma)}]_{m,m}^*}{\|\mathbf{e}_m^T \mathbf{H}_T^{(\beta,\phi,\sigma)}\|_2^2}, & \text{if } \zeta_{\text{TC}}^{(\beta,\phi,\sigma)}(m) \geq 1 \\ 1, & \text{otherwise} \end{cases} \quad (46)$$

where

$$\zeta_{\text{TC}}^{(\beta,\phi,\sigma)}(m) = \frac{\left| [\mathbf{H}_T^{(\beta,\phi,\sigma)}]_{m,m} \right|^2}{\sum_{k=0, k \neq m}^{K-1} \left| [\mathbf{H}_T^{(\beta,\phi,\sigma)}]_{m,k} \right|^2}. \quad (47)$$

To enhance the suitability of the preconditioner, the optimal resampling operation is needed as given by

$$(\beta_{T,\star}, \phi_{T,\star}, \sigma_{T,\star}) = \arg \max_{\beta, \phi, \sigma} \sum_{m=0}^{K-1} \left| \sum_{l=0}^L h_l^{(\beta,\sigma)} e^{j2\pi \omega \frac{\alpha_l - 1 + \phi}{\beta} \frac{m}{K}} \times \right. \\ \left. e^{-j \frac{(K-1)\pi}{K} (\xi_{l,T1} m + \xi_{l,T2})} \times \frac{\text{sinc}(\xi_{l,T1} k + \xi_{l,T2})}{\text{sinc}(\frac{1}{K} (\xi_{l,T1} m + \xi_{l,T2}))} \right|^2. \quad (48)$$

Appendix 3

Equalization using the conjugate gradient algorithm

If we consider to solve the preconditioned system in (23) in a similar manner as (22), we have

$$\hat{\mathbf{b}}_C = \bar{\mathbf{M}}_{\text{FC}}^{(\beta,\phi,\sigma)^{-1}} \bar{\mathbf{H}}_{\text{FC}}^{(\beta,\phi,\sigma)^H} \mathbf{r}_F^{(\beta,\phi,\sigma)}$$

where $\bar{\mathbf{M}}_{\text{FC}}^{(\beta,\phi,\sigma)} = \bar{\mathbf{H}}_{\text{FC}}^{(\beta,\phi,\sigma)^H} \bar{\mathbf{H}}_{\text{FC}}^{(\beta,\phi,\sigma)}$, and $\hat{\mathbf{b}}_C$ is the estimate of $\mathbf{b}_C = \mathbf{C}_F^{-1} \mathbf{b}$.

Its implementation using CG is described in the frequency domain as follows

1. Define $\mathbf{d}_F = \bar{\mathbf{H}}_F^{(\beta,\phi,\sigma)^H} \mathbf{r}_F^{(\beta,\phi,\sigma)}$ and $i = 0$;

2. Perform the following iterations:

$$\begin{aligned}
& \text{Loop} \\
& \mathbf{g}^{(i)} = \mathbf{d}_F - \bar{\mathbf{M}}_{FC}^{(\beta, \phi, \sigma)} \hat{\mathbf{b}}_C^{(i)}, \\
& \mathbf{a}^{(i)} = \frac{\|\mathbf{g}^{(i)}\|_{\text{Fro}}^2}{\|\mathbf{g}^{(i-1)}\|_{\text{Fro}}^2} \mathbf{a}^{(i-1)} + \mathbf{g}^{(i)}, \\
& u^{(i)} = \frac{\xi^{(i)}}{\mathbf{a}^{(i)H} \bar{\mathbf{M}}_{FC}^{(\beta, \phi, \sigma)} \mathbf{a}^{(i)}}, \\
& \hat{\mathbf{b}}_C^{(i)} = \hat{\mathbf{b}}_C^{(i-1)} + u^{(i)} \mathbf{a}^{(i)} \\
& \text{End Loop;}
\end{aligned} \tag{49}$$

where $\mathbf{a}^{(0)} = \mathbf{g}^{(0)} = \mathbf{d}_F$, $u^{(0)} = \frac{\|\mathbf{d}_F\|^2}{\mathbf{d}_F^H \bar{\mathbf{M}}_{FC}^{(\beta, \phi, \sigma)} \mathbf{d}_F}$ and $\hat{\mathbf{b}}^{(0)} = u^{(0)} \mathbf{d}_F$;

3. Perform $\hat{\mathbf{b}}^{(i)} = \mathbf{C}_F \mathbf{b}_C^{(i)}$, which is the i th output of the equalization process, and the index i is incremental from 0 to i_{\max} where i_{\max} is the iteration number when the stopping criterion of the CG is satisfied.

Notably, the optimal stopping criterion for CG can be case dependent, e.g., as discussed in [23], and is not included in this article. When our CG iterations stop, we finally have $\hat{\mathbf{b}} = \hat{\mathbf{b}}^{(i_{\max})}$, which is the data estimate.

It is worthy to note that the computational complexity of each CG iteration above is determined by the complex multiplication (CM) of $\bar{\mathbf{M}}_{FC}^{(\beta, \phi, \sigma)}$ with a vector (e.g. $\hat{\mathbf{b}}^{(i)}$ or $\mathbf{a}^{(i)}$), e.g., as in (49). When $\mathbf{C}_F^{(\beta, \phi, \sigma)}$ is a diagonal preconditioner as considered in this article, the bandwidth of the preconditioned $\bar{\mathbf{H}}_{FC}^{(\beta, \phi, \sigma)}$ equals that of $\bar{\mathbf{H}}_C^{(\beta, \phi, \sigma)}$, and consequently $\bar{\mathbf{M}}_{FC}^{(\beta, \phi, \sigma)}$ is banded with a bandwidth $2B_F^{(\beta, \phi)}$ where $B_F^{(\beta, \phi)} = \max_k B_F^{(\beta, \phi)}(k)$ with $B_F^{(\beta, \phi)}(k)$ defined in (19). In this case, the computational complexity of each iteration is upper-bounded by $\mathcal{O}(B_F^{(\beta, \phi)} K)$ which is linear in the vector size K .

One can also repeat the above derivations using the TD notations for the TD CG equalization.

Appendix 4

Eigenvalue locations

We consider the diagonal matrix $\mathbf{C}_F = \text{diag}\{[c_{F,0}, c_{F,1}, \dots, c_{F,K-1}]^T\}$, and denote the eigenvalues of $\bar{\mathbf{H}}_{FC}^{(\beta, \phi, \sigma)} = \bar{\mathbf{H}}_F^{(\beta, \phi, \sigma)} \mathbf{C}_F$ as $\{\mu_1, \mu_2, \dots, \mu_{K-1}\}$.

Let $\mathbf{U}\mathbf{W}\mathbf{U}^H$ be a Schur decomposition of $\bar{\mathbf{H}}_{FC}^{(\beta, \phi, \sigma)}$ such that $\mathbf{U}\mathbf{U}^H = \mathbf{I}_{K \times K}$ and the diagonal

elements of \mathbf{W} equal $\{\mu_1, \mu_2, \dots, \mu_{K-1}\}$. Then

$$\begin{aligned} \sum_{k=0}^{K-1} |\mu_k|^2 &= \|\text{diag}\{\mathbf{W}\}\|_2^2 \leq \|\mathbf{W}\|_{\text{Fro}}^2 \\ &= \|\bar{\mathbf{H}}_{\text{FC}}^{(\beta, \phi, \sigma)}\|_{\text{Fro}}^2 = \|\bar{\mathbf{H}}_{\text{F}}^{(\beta, \phi, \sigma)} \mathbf{C}_{\text{F}}\|_{\text{Fro}}^2. \end{aligned}$$

Note that $\bar{\mathbf{H}}_{\text{F}}^{(\beta, \phi, \sigma)} \mathbf{e}_{k\text{CF},k} = \bar{\mathbf{H}}_{\text{F}}^{(\beta, \phi, \sigma)} \mathbf{C}_{\text{F}} \mathbf{e}_k$, where \mathbf{e}_k stands for an all-zero vector except for its k th entry which equals 1, as defined in (25) for $k = 0, 1, \dots, K$. We then recall (27), which holds for any $k \in \{0, 1, \dots, K-1\}$, and thus upper-bounds the above expressions as

$$\sum_{k=0}^{K-1} |\mu_k|^2 \leq K \|\bar{\mathbf{H}}_{\text{F}}^{(\beta, \phi, \sigma)} \mathbf{e}_{k\text{CF},k}\|_{\text{Fro}}^2 \leq K \epsilon_0^2.$$

Similarly, we can also prove that $\sum_{k=0}^{K-1} |\mu_k - 1|^2 \leq K \epsilon_1^2$ associated with (26).

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The first author wants to thank the National University of Defense Technology, China, and also the China Scholarship Council for the financial support. This work was supported in part by NWO-STW under the VICI program (project 10382). The work of Z. Tang is also supported in part by the European Defence Agency (EDA) project RACUN (Robust Acoustic Communication in Underwater Networks). In addition, we would like to thank Dr. Magnus Lundberg Nordenvaad from the Lulea University of Technology, Sweden, Prof. Urbashi Mitra from the University of Southern California, U.S., and Prof. Huihuang Chen from the Xiamen University, China, who participated in valuable discussions.

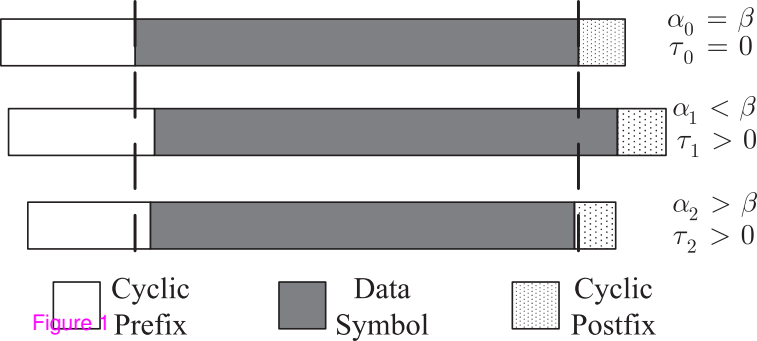
Endnote

^a As a matter of fact, the case where $\alpha_l < 1$ or $\tau_l < 0$ can be converted to the current situation by means of proper resampling and timing at the receiver. This justifies the assumption of a compressive and causal scenario without loss of generality.

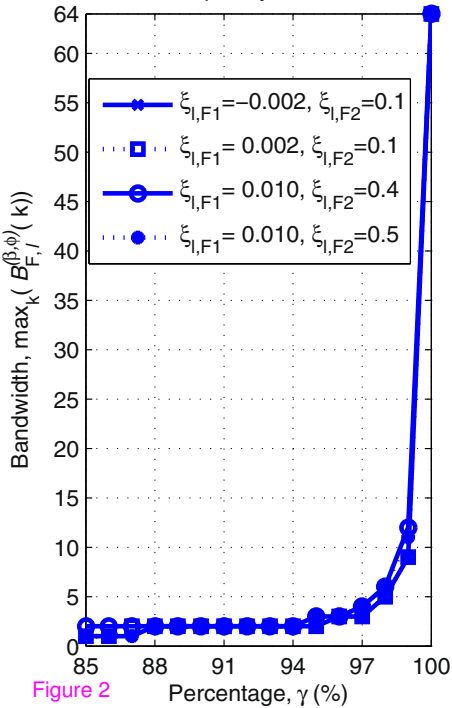
References

1. Z Wang, GB Giannakis, Wireless multicarrier communications: where Fourier meets Shannon. *IEEE Signal Process. Mag.* **17**(3), 29–48 (2000)
2. X Cai and GB Giannakis, Bounding performance and suppressing intercarrier interference in wireless mobile OFDM. *IEEE Trans. Commun.* **51**(12), 2047–2056 (2003).
3. LG Weiss, Wavelets and wideband correlation processing. *IEEE Signal Process. Mag.* **11**, 13–32 (1994)
4. DW Ricker, The Doppler sensitivity of large TW phase modulated waveforms. *IEEE Trans. Signal Process.* **40**(10), 2406–2413 (1992)
5. P Schniter, Low-complexity equalization of OFDM in doubly selective channels. *IEEE Trans. Signal Process.* **52**(4), 1002–1011 (2004)
6. L Rugini, P Banelli, G Leus, Simple equalization of time-varying channels for OFDM. *IEEE Commun. Lett.* **9**(7), 619–621 (2005)
7. K Fang, L Rugini, G Leus, Low-complexity block turbo equalization for OFDM systems in time-varying channels. *IEEE Trans. Signal Process.* **56**(11), 5555–5566 (2008)
8. Y Jiang, A Papandreou-Suppappola, Discrete time-scale characterization of wideband time-varying systems. *IEEE Trans. Signal Process.* **54**(4), 1364–1375 (2006)
9. EM Sozer, JG Proakis, R Stojanovic, JA Rice, A Benson, M Hatch, Direct sequence spread spectrum based modem for under water acoustic communication and channel measurements, in *Proc. Conf. IEEE/MTS OCEANS*, vol. 1 (Boston, MA, USA, 1999) pp. 228–233
10. AR Margetts, P Schniter, A Swami, Joint scale-lag diversity in wideband mobile direct sequence spread spectrum systems. *IEEE Trans. Wirel. Commun.* **6**(12), 4308–4319 (2007)
11. G Leus, P van Walree, Multiband OFDM for covert acoustic communications. *IEEE J. Sel. Areas Commun.* **26**(9), 1662–1673 (2008)
12. AB Salberg, A Swami, Doppler and frequency-offset synchronization in wideband OFDM. *IEEE Trans. Wirel. Commun.* **4**(6), 2870–2881 (2005)
13. B Li, S Zhou, M Stojanovic, L Freitag, P Willett, Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts. *IEEE J. Oceanic Eng.* **33**(2), 198–209 (2008)
14. S Yerramalli, U Mitra, Optimal resampling of OFDM signals for multiscale-multilag underwater acoustic channels. *IEEE J. Oceanic Eng.* **36**, 126–138 (2011)
15. GH Golub, CF van Loan, *Matrix Computations*, 3rd edn. (Johns Hopkins University Press, 1996)
16. M Benzi, Preconditioning techniques for large linear systems: a survey. *J. Comput. Phys.* **182**, 418–477 (2002)

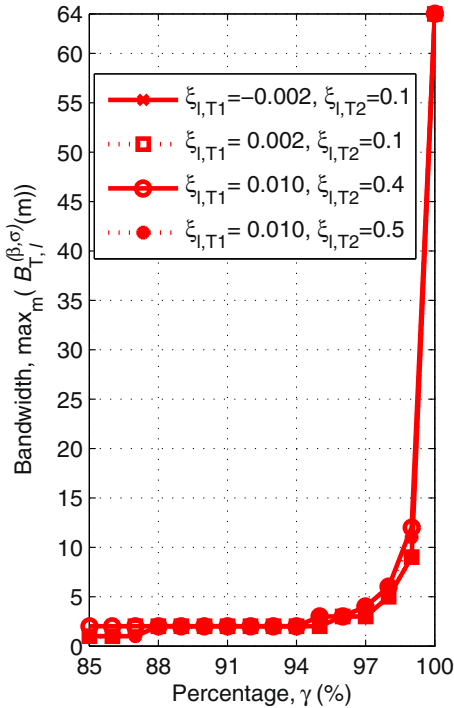
17. MJ Groote, T Huckle, Parallel preconditioning with sparse approximate inverses. *SIAM J. Sci. Comput.* **18**, 838–853 (1997)
18. T Hrycak, S Das, G Matz, HG Feichtinger, Low complexity equalization for doubly selective channels modeled by a basis expansion. *IEEE Trans. Signal Process.* **58**(11), 5706–5719 (2010)
19. Z Tang, R Remis, G Leus, T Xu, ML Nordenvaad, Equalization for multi-scale/multi-lag OFDM channels, in *Proc. Allerton Conf. Comm., Control, and Comput.*, vol. 1 (Monticello, IL, USA, 2011), pp. 654–661
20. T Xu, Z Tang, G Leus, U Mitra, Time- or frequency-domain equalization for wideband OFDM channels? in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 5 (Kyoto, Japan, 2012) pp. 3029–3032
21. Z Tang, R Remis, ML Nordenvaad, On preconditioned conjugate gradient method for time-varying OFDM channel equalization, in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 5 (Kyoto, Japan, 2012) pp. 2119–2122
22. CC Paige, MA Saunders, LSQR: an algorithm for sparse linear equations and sparse least square problems. *ACM Trans. Math. Softw.* **8**, 43–71 (1982)
23. M Hanke, *Conjugate Gradient Type Methods for ill-Posed Problems* (Longman Scientific & Technical, 1992)
24. D Falconer, SL Ariyavisitakul, A Benyamin-Seeyar, B Eidson, Frequency domain equalization for single-carrier broadband wireless systems. *IEEE Commun. Mag.* **40**(4), 58–66 (2002)
25. GB Giannakis, C Tepedelenlioglu, Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels. *Proc. IEEE* **86**(10), 1969–1986 (1998)

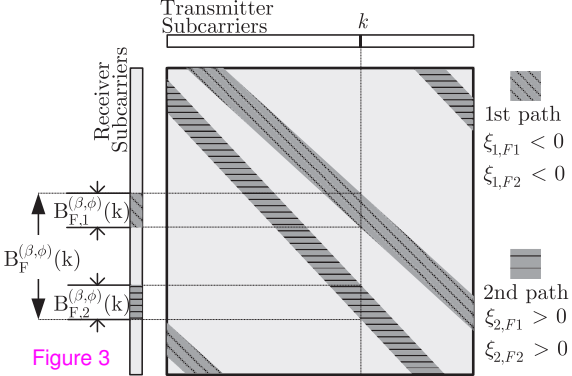


Frequency Domain



Time Domain





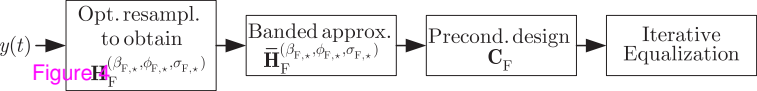
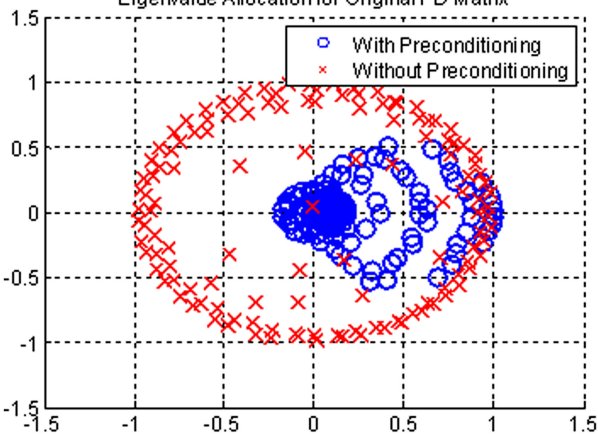
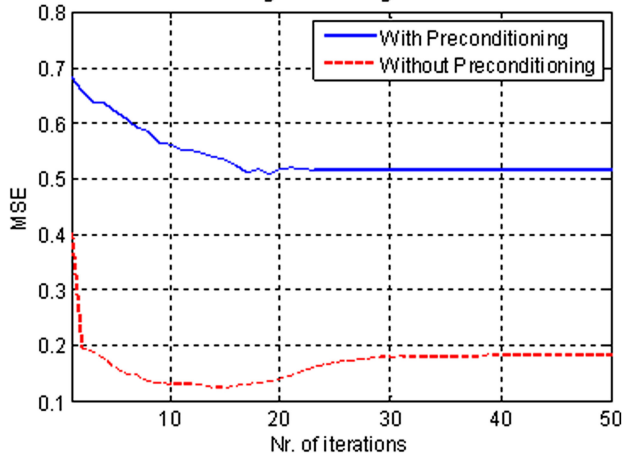


Figure 4

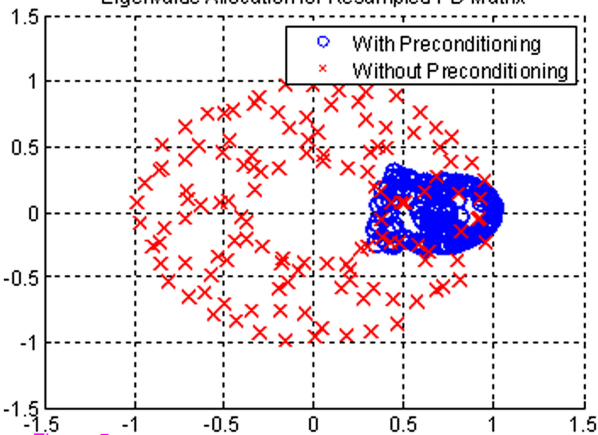
Eigenvale Allocation for Original FD Matrix



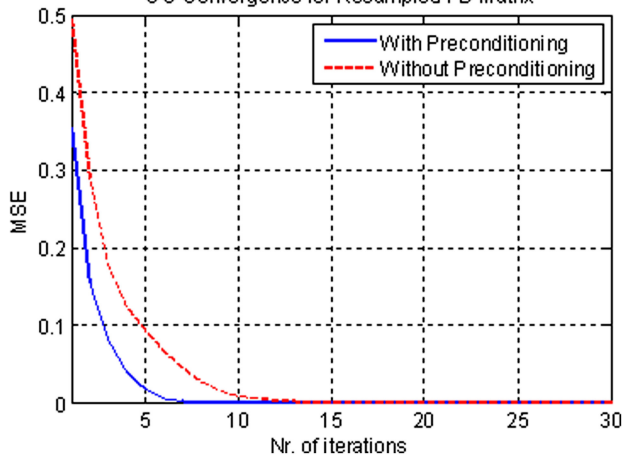
CG Convergence for Original FD Matrix



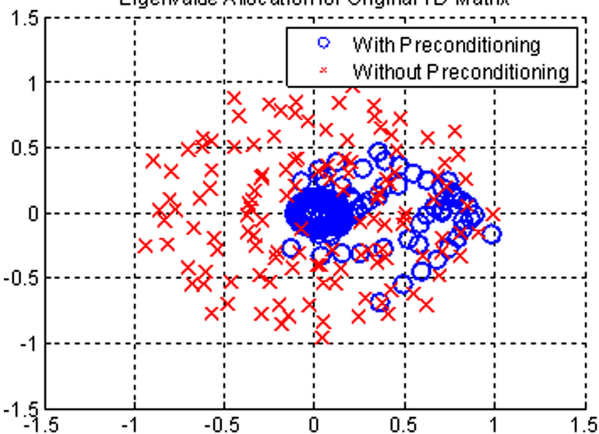
Eigenvale Allocation for Resampled FD Matrix



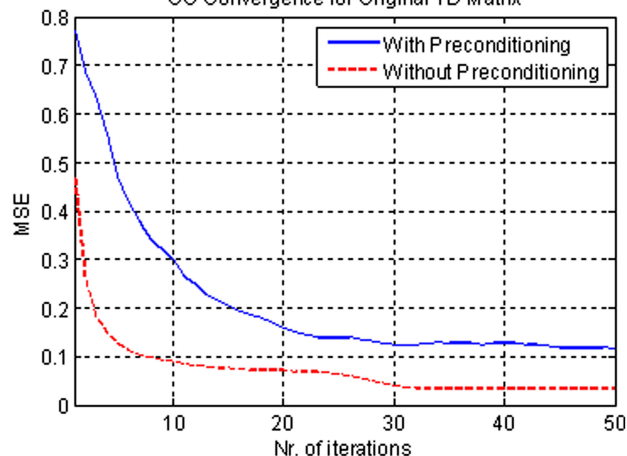
CG Convergence for Resampled FD Matrix



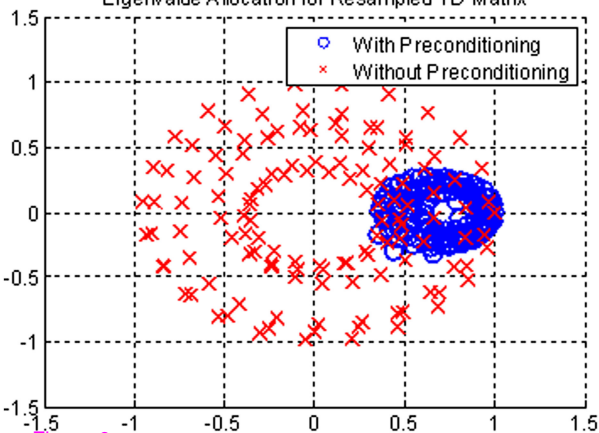
Eigenvalue Allocation for Original TD Matrix



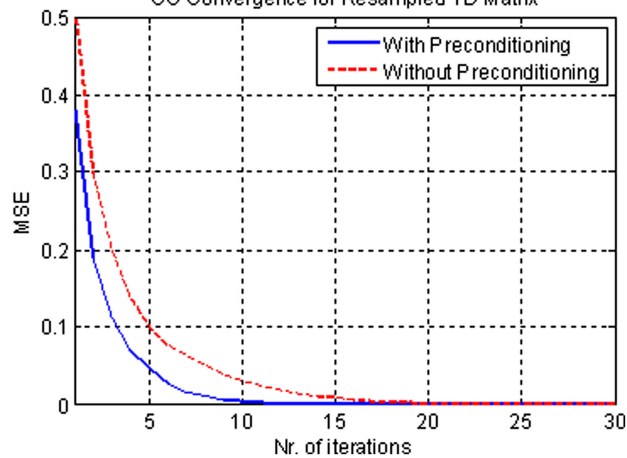
CG Convergence for Original TD Matrix



Eigenvalue Allocation for Resampled TD Matrix



CG Convergence for Resampled TD Matrix



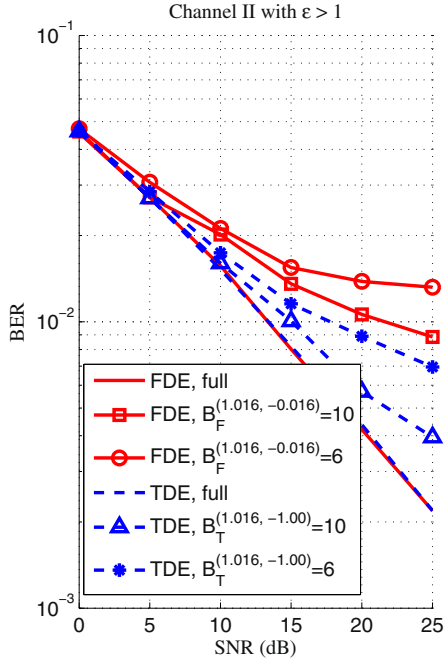
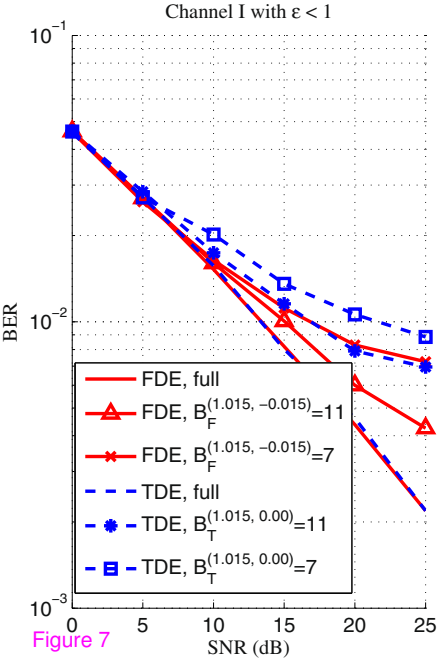
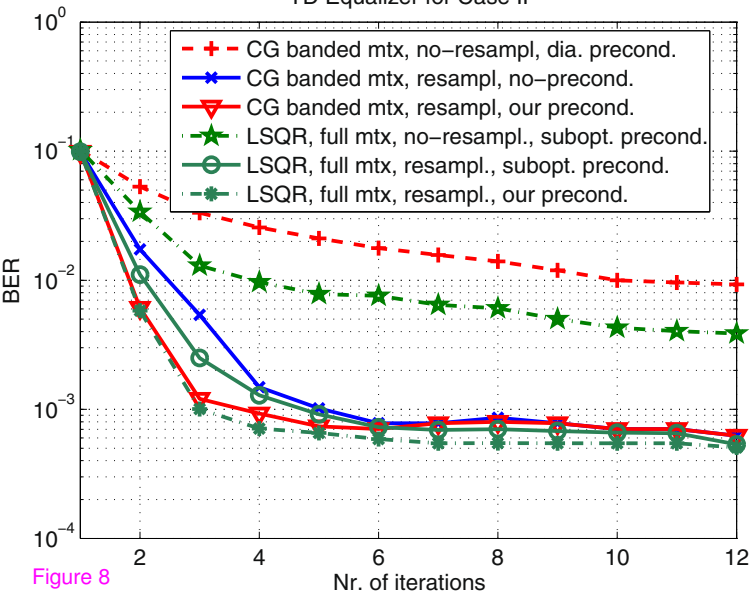


Figure 7

TD Equalizer for Case II



CG Equalizer for Case I

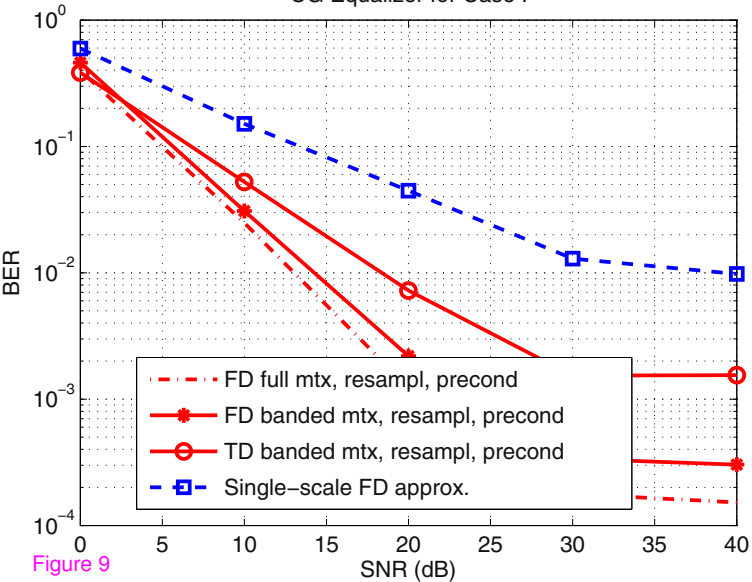
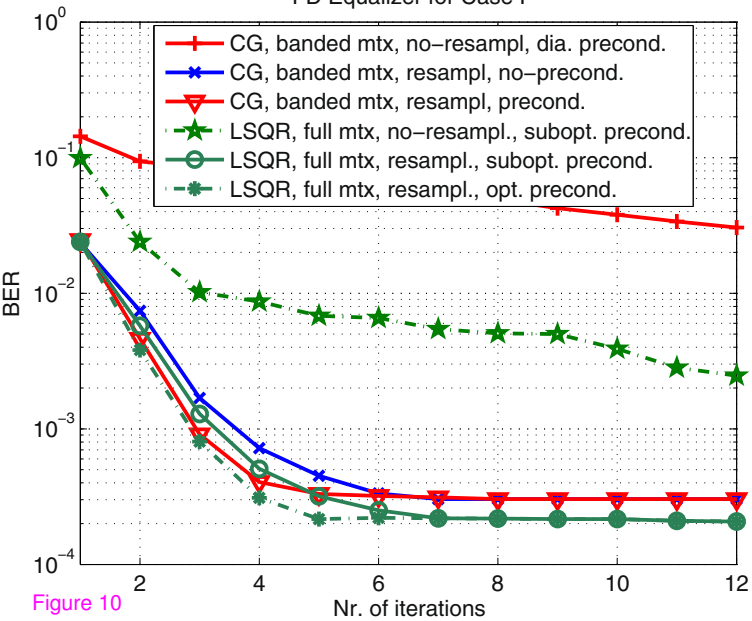


Figure 9

FD Equalizer for Case I



CG Equalizer for Case II

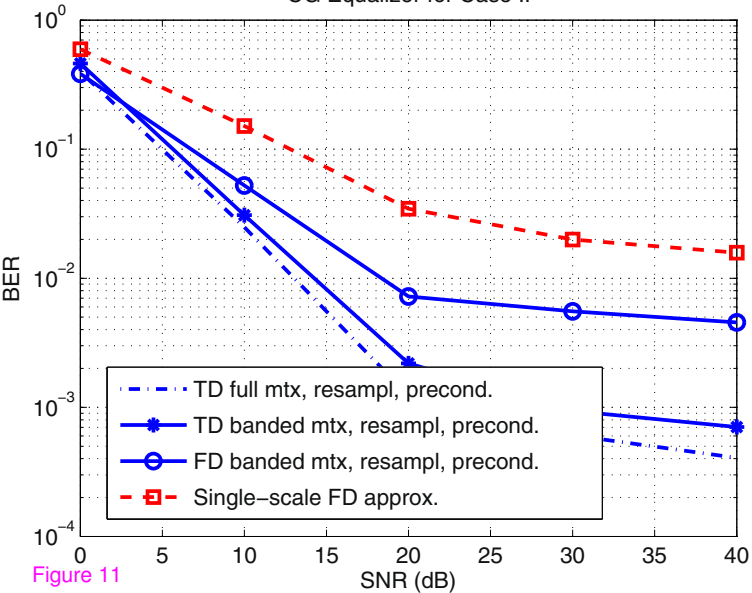


Figure 11

Additional files provided with this submission:

Additional file 1: AQC.pdf, 95K

<http://jwcn.eurasipjournals.com/imedia/6239092777653928/supp1.pdf>