[19] M. Jansson, P. Wirfalt, K. Werner, and B. Ottersten, "ML estimation of covariance matrices with Kronecker and persymmetric structure," in *Proc. IEEE DSP/SPE Workshop*, Marco Island, FL, Jan. 2009, pp. 298–301.

[20] J. Ward, "Space-time adaptive processing for airborne radar," Lincoln Lab., MIT, Cambridge, MA, Tech. Rep. 1015, Dec. 1994.

[21] K. J. Sohn, H. Li, and B. Himed, "Recursive parametric tests for multichannel adaptive signal detection," *IET Radar, Sonar, Navig.*, vol. 2, no. 1, pp. 63–70, Feb. 2008.

[22] Y. I. Abramovich, "Controlled method for adaptive optimization of filters using the criterion of maximum SNR," *Radio Eng. Electron. Phys.*, vol. 26, pp. 87–95, Mar. 1981.

[23] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, Jul. 1988.

# Single-Carrier Systems With MMSE Linear Equalizers: Performance Degradation due to Channel and CFO Estimation Errors

Athanasios P. Liavas and Despoina Tsipouridou

***Abstract*—We assess the impact of the channel and the carrier frequency offset (CFO) estimation errors on the performance of single-carrier systems with MMSE linear equalizers. Performance degradation is caused by the fact that a *mismatched* MMSE linear equalizer is applied to channel output samples with *imperfectly canceled* CFO. Assuming a single-block training, we develop an asymptotic expression for the excess mean square error (EMSE) induced by the channel and CFO estimation errors and derive a simple EMSE approximation which reveals the following: 1) performance degradation is mainly caused by the imperfectly canceled CFO and 2) the EMSE is approximately proportional to the CFO estimation error variance, with the factor of proportionality being independent of the training sequence. We also highlight the fact that the placement of the single-block training at the middle of the packet is a good practice.**

***Index Terms*—Joint channel and CFO estimation, linear MMSE equalization.**

## I. Introduction

A problem that frequently arises in packet-based wireless communication systems is the *joint* estimation of the frequency selective channel and the CFO [1], [2]. Optimal training sequence (TS) design for this problem has been considered in [2], where the optimized cost function was the *asymptotic* Cramér-Rao bound (CRB). However, in [2], the channel and CFO estimation errors were assigned equal weight which might be *suboptimal* since "*... presumably channel estimation errors will have a different impact, e.g., on bit-error rate, than frequency estimation errors*" [2].

It seems that the *unequal weighting* problem cannot be resolved unless we consider specific communication systems. An important structure to study is a single-carrier system with an MMSE linear equalizer. Performance degradation is caused by the fact that a *mismatched* MMSE linear equalizer is applied to channel output samples with *imperfectly canceled* CFO. This system has been considered in [3], where the authors derived an expression for the EMSE and designed optimal TSs.

We consider the same system as in [3] but our aim is different. More specifically, we assume a *single-block* training and, as done in [3], we develop an asymptotic expression for the induced EMSE which, however, is difficult to interpret. Our main contribution lies in the fact that, assuming *small ideal MMSE*, we derive a simple and informative EMSE approximation, which reveals the following:

1) the dominant error source is the imperfectly canceled CFO;
2) the EMSE is approximately proportional to the CFO estimation error variance, with the factor of proportionality being *independent* of the TS.[1]

We also highlight the fact that the placement of the single-block TS at the middle of the transmitted packet is a good practice.

*Notation:* Superscripts $^T$, $^H$, and $^*$ denote transpose, conjugate transpose, and elementwise conjugation, respectively. $\mathrm{tr}(\cdot)$ denotes the trace operator, $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex number, and $\mathbf{I}_M$ denotes the $M \times M$ identity matrix. $\sigma_{\max}(\cdot)$, $\sigma_{\min}(\cdot)$, $\|\cdot\|_2$, $\|\cdot\|_F$, and $k_2(\cdot)$ denote, respectively, the maximum singular value, the minimum singular value, the spectral norm, the Frobenius norm, and the condition number, with respect to the spectral norm, of the matrix argument. $\mathcal{E}[X]$ denotes the expected value of $X$. $\mathbf{P}_{\mathcal{R}(\mathbf{A})}$ and $\mathbf{P}^{\perp}_{\mathcal{R}(\mathbf{A})}$ denote, respectively, the orthogonal projector onto the column space of matrix $\mathbf{A}$ and its orthogonal complement.

## II. Channel and CFO Estimation

### A. The Channel Model

We consider a packet-based single-carrier system with input packet length $N$. We assume that the baseband-equivalent frequency-selective channel has impulse response $\mathbf{h} \triangleq [h_0 \cdots h_L]^T$, angular CFO $\omega$, and phase $\phi$. The output at time instant $n$, for $n = 1, \ldots, N + L$, is

$$r_n = \mathrm{e}^{j(\omega n+\phi)} \sum_{l=0}^{L} h_l a_{n-l} + w_n \tag{1}$$

where $\{a_n\}_{n=1}^{N}$ and $\{w_n\}_{n=1}^{N+L-1}$ denote the channel input and additive channel noise, respectively. The input symbols are i.i.d. unit variance circular. The noise samples are i.i.d. circular Gaussian with variance $\sigma_w^2$. In the sequel, we absorb term $e^{j\phi}$ into channel $\mathbf{h}$.

The channel output vector $\mathbf{r}_{n:n-M} \triangleq [r_n \ \cdots \ r_{n-M}]^T$ can be expressed as

$$\mathbf{r}_{n:n-M} = \mathbf{\Gamma}_{n:n-M}(\omega)\mathbf{H}\mathbf{a}_{n:n-L-M} + \mathbf{w}_{n:n-M} \tag{2}$$

where

$$\mathbf{\Gamma}_{n:n-M}(\omega) \triangleq \mathrm{diag}(\mathrm{e}^{j\omega n}, \ldots, \mathrm{e}^{j\omega(n-M)}) \tag{3}$$

and $\mathbf{H}$ is the $(M+1)\times(M+L+1)$ Toeplitz filtering matrix constructed by $\mathbf{h}$.

[1]Thus, optimal TS design for CFO estimation is also highly relevant for *joint* channel and CFO estimation.

Manuscript received June 20, 2011; revised October 31, 2011 and January 19, 2012; accepted February 13, 2012. Date of publication March 05, 2012; date of current version May 11, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yimin D. Zhang.

The authors are with the Department of Electronic and Computer Engineering, Technical University of Crete, Greece (e-mail: liavas@telecom.tuc.gr; despoina@telecom.tuc.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2012.2189858

*B. Channel and CFO Estimation*

The $N_{\rm tr}$ consecutive symbols $\mathbf{a}_{\rm tr} \triangleq [a_{n_1} \ \cdots \ a_{n_2}]^T$, with $N_{\rm tr} \triangleq n_2 - n_1 + 1$, are used for training.[2] We collect the output samples that depend *only* on the training and construct

$$\mathbf{y} \triangleq \mathbf{r}_{n_2:n_1+L} = \mathbf{\Gamma}_{n_2:n_1+L}(\omega)\mathbf{A}\mathbf{h} + \mathbf{w}_{n_2:n_1+L} \tag{4}$$

where $\mathbf{A}$ is the $(N_{\rm tr} - L) \times (L+1)$ Hankel matrix

$$\mathbf{A} \triangleq \begin{bmatrix} a_{n_2} & \cdots & a_{n_2-L} \\ \vdots & \ddots & \vdots \\ a_{n_1+L} & \cdots & a_{n_1} \end{bmatrix}. \tag{5}$$

The joint ML CFO and channel estimates are [1]

$$\hat{\omega} = \arg\max_{\tilde{\omega}} \{\mathbf{y}^H \mathbf{\Gamma}_{n_2:n_1+L}(\tilde{\omega})\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{\Gamma}^H_{n_2:n_1+L}(\tilde{\omega})\mathbf{y}\} \tag{6}$$

and

$$\hat{\mathbf{h}} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{\Gamma}^H_{n_2:n_1+L}(\hat{\omega})\mathbf{y}. \tag{7}$$

The estimation errors are $\Delta\omega \triangleq \hat{\omega} - \omega$ and $\Delta\mathbf{h} \triangleq \hat{\mathbf{h}} - \mathbf{h}$. We assume that $N_{\rm tr}$ is sufficiently large so that the above ML estimates are unbiased and efficient. Thus, the second-order statistics of $\Delta\omega$ and $\Delta\mathbf{h}$ are determined by the *finite sample* CRBs [2]. More specifically, if we define

$$\mathbf{K} \triangleq \mathrm{diag}(n_2, \ldots, n_1 + L), \tag{8}$$

then, working as in [2], we can show that

$$\sigma^2_{\Delta\omega} \triangleq \mathcal{E}\left[(\Delta\omega)^2\right] = \frac{\sigma_w^2}{2\,\mathrm{tr}\left(\mathbf{h}^H\mathbf{A}^H\mathbf{K}\mathbf{P}^{\perp}_{\mathcal{R}(\mathbf{A})}\mathbf{K}\mathbf{A}\mathbf{h}\right)} \tag{9}$$

$$\begin{aligned}\mathbf{\Psi} &\triangleq \mathcal{E}\left[\Delta\mathbf{h}\Delta\mathbf{h}^H\right] = \sigma_w^2(\mathbf{A}^H\mathbf{A})^{-1} \\ &\quad + \sigma^2_{\Delta\omega}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{K}\mathbf{A}\mathbf{h}\mathbf{h}^H\mathbf{A}^H\mathbf{K}\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1} \end{aligned} \tag{10}$$

$$\begin{aligned}\mathbf{\Psi}_t &\triangleq \mathcal{E}[\Delta\mathbf{h}\Delta\mathbf{h}^T] = -\sigma^2_{\Delta\omega}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{K}\mathbf{A}\mathbf{h}\mathbf{h}^T \\ &\quad \times \mathbf{A}^T\mathbf{K}\mathbf{A}^*(\mathbf{A}^H\mathbf{A})^{-T} \end{aligned} \tag{11}$$

$$\boldsymbol{\psi} \triangleq \mathcal{E}[\Delta\omega\Delta\mathbf{h}] = j\,\sigma^2_{\Delta\omega}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{K}\mathbf{A}\mathbf{h}. \tag{12}$$

Since $\mathbf{K}$ depends on the training positions, it seems that the quantities defined in (9)–(12) also depend on the training positions. However, if we express $\mathbf{K}$ as

$$\mathbf{K} = n_2\mathbf{I}_{N_{\rm tr}-L} - \mathbf{D}_{N_{\rm tr}-L-1} \tag{13}$$

with $\mathbf{D}_i \triangleq \mathrm{diag}(0, 1, \ldots, i)$, then it can be shown that

$$\sigma^2_{\Delta\omega} = \frac{1}{2}\,\sigma_w^2\left[\mathrm{tr}\left(\mathbf{h}^H\mathbf{A}^H\mathbf{D}_{N_{\rm tr}-L-1}\mathbf{P}^{\perp}_{\mathcal{R}(\mathbf{A})}\mathbf{D}_{N_{\rm tr}-L-1}\mathbf{A}\mathbf{h}\right)\right]^{-1}. \tag{14}$$

That is, $\sigma^2_{\Delta\omega}$ is *independent* of the training positions.

On the other hand, the accuracy of $\hat{\mathbf{h}}$ is determined by the CFO estimation error that exists in $\Gamma_{n_2:n_1+L}(\hat{\omega})$ and does depend on the training positions. The structure of $\Gamma_{n_2:n_1+L}(\hat{\omega})$ suggests that an accurate channel estimate might be obtained if we absorb into channel $\mathbf{h}$ term $e^{j\omega\xi}$, with $\xi \triangleq n_1 + \frac{N_{\rm tr}+L}{2}$, i.e., $\xi$ is the index of the output sample that lies at the middle position of $\mathbf{y}$, getting the "new channel" $\mathbf{h}' \triangleq \mathrm{e}^{j\omega\xi}\mathbf{h}$.[3] Then, the channel output is written as

$$r_n = \mathrm{e}^{j\omega(n-\xi)}\sum_{l=0}^{L} h'_l a_{n-l} + w_n \tag{15}$$

and (4) can be expressed as

$$\mathbf{y} = \mathbf{\Gamma}_{\frac{N_{\rm tr}-L}{2}-1:-\frac{N_{\rm tr}-L}{2}}(\omega)\mathbf{A}\mathbf{h}' + \mathbf{w}_{n_2:n_1+L}. \tag{16}$$

In the sequel, we assume that the true system model is given by (15). The ML estimate of $\omega$ is still given by (6), while

$$\hat{\mathbf{h}}' = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{\Gamma}^H_{\frac{N_{\rm tr}-L}{2}-1:-\frac{N_{\rm tr}-L}{2}}(\hat{\omega})\,\mathbf{y}. \tag{17}$$

We define

$$\mathbf{K}' \triangleq \mathrm{diag}\left(\frac{N_{\rm tr}-L}{2} - 1, \ldots, -\frac{N_{\rm tr}-L}{2}\right) \tag{18}$$

and $\Delta\mathbf{h}' \triangleq \hat{\mathbf{h}}' - \mathbf{h}'$. The estimation error second-order statistics, denoted by $\mathbf{\Psi}'$, $\mathbf{\Psi}'_t$, $\boldsymbol{\psi}'$, and $\sigma^2_{\Delta\omega}$, are given by (10)–(12) and (14), with $\mathbf{h}$ and $\mathbf{K}$ substituted by $\mathbf{h}'$ and $\mathbf{K}'$, respectively. Finally, we assume that the noise variance, $\sigma_w^2$, is known at the receiver, i.e., the noise variance estimation error is negligible compared with the channel and CFO estimation error.

## III. CFO Correction and MMSE Linear Equalization

*A. The Ideal Case*

If we know the CFO, then we can cancel it perfectly before equalization. If we know the channel, then we can compute the order-$M$ delay-$d$ MMSE linear equalizer, $\mathbf{f} \triangleq [f_0 \ \cdots \ f_M]^T$, as [4, Section 2.7.3]

$$\mathbf{f} = \left(\mathbf{H}'\mathbf{H}'^H + \sigma_w^2\mathbf{I}_{M+1}\right)^{-1}\mathbf{H}'\mathbf{e}_d = \mathbf{R}_z^{-1}\mathbf{H}'\mathbf{e}_d \tag{19}$$

where $\mathbf{e}_d$ is the $(M+L+1)\times 1$ vector with 1 at the $(d+1)$-st position and zeros elsewhere. It can be shown that the mean-square input symbol estimation error at the output of the MMSE linear equalizer is

$$\mathrm{MSE}(\mathbf{f}) = 1 - \mathbf{f}^H\mathbf{R}_z\mathbf{f}. \tag{20}$$

*B. Mismatched CFO Correction and MMSE Equalization*

If we do not know the true channel and CFO, then we can adopt the so-called mismatched approach, that is, estimate them and use the estimates as if they were the true quantities.

The mismatched MMSE equalizer is [see (19)]

$$\hat{\mathbf{f}} = \left(\hat{\mathbf{H}}'\hat{\mathbf{H}}'^H + \sigma_w^2\mathbf{I}_{M+1}\right)^{-1}\hat{\mathbf{H}}'\mathbf{e}_d \tag{21}$$

with mismatch $\Delta\mathbf{f} \triangleq \hat{\mathbf{f}} - \mathbf{f}$. After CFO correction, we obtain

$$s_n \triangleq \mathrm{e}^{-j\hat{\omega}(n-\xi)}r_n. \tag{22}$$

The vector $\mathbf{s}_{n:n-M}$ can be expressed as

$$\begin{aligned}\mathbf{s}_{n:n-M} &= \mathrm{e}^{j\Delta\omega\xi}\mathbf{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{a}_{n:n-L-M} \\ &\quad + \mathrm{e}^{j\hat{\omega}\xi}\mathbf{\Gamma}_{n:n-M}(-\hat{\omega})\mathbf{w}_{n:n-M}. \end{aligned} \tag{23}$$

[2]Training schemes with two or more blocks are beyond the scope of this work.

[3]We shall say more on this topic later.

The input symbol estimation error at the output of the mismatched equalizer at time instant $n$ is

$$\hat{e}_n \triangleq \hat{\mathbf{f}}^H \mathbf{s}_{n:n-M} - \mathbf{e}_d^H \mathbf{a}_{n:n-L-M} \tag{24}$$

and the time-dependent mean square error is given by (25) at the bottom of the page.

## IV. EMSE Analysis

The EMSE at time instant $n$ is defined as

$$\mathrm{EMSE}_n(\hat{\mathbf{f}}, \hat{\omega}) \triangleq \mathcal{E}_{\Delta \mathbf{h}', \Delta\omega}[\mathrm{MSE}_n(\hat{\mathbf{f}}, \hat{\omega})] - \mathrm{MSE}(\mathbf{f}). \tag{26}$$

Using slightly different notation, it has been proved in [5, eqs. (22) and (27)] that the mismatched equalizer $\hat{\mathbf{f}}$ can be expressed as

$$\hat{\mathbf{f}} = \mathbf{f} - \mathbf{R}_z^{-1}\left(\mathbf{R}^* \Delta\mathbf{h}' + \mathbf{G}\Delta\mathbf{h}'^*\right) + \mathcal{O}\left(\|\Delta\mathbf{h}'\|^2\right) \tag{27}$$

where

1) $\mathbf{R}$ is the $(M+1)\times(L+1)$ Hankel matrix constructed by vector

$$\mathbf{r} \triangleq \mathbf{c} - \mathbf{e}_d \tag{28}$$

where $\mathbf{c}$ is the combined (channel-equalizer) impulse response, i.e., $\mathbf{c} \triangleq \mathbf{H}'^T \mathbf{f}^*$;

2) $\mathbf{G} \triangleq \mathbf{H}'\mathbf{F}^T$, where $\mathbf{F}$ is the $(L+1)\times(L+M+1)$ Toeplitz filtering matrix constructed by $\mathbf{f}$.

The following proposition provides an asymptotic EMSE expression. We note that the same result, expressed in terms of frequency domain quantities, has been derived in [3].

*Proposition 1:* . The EMSE induced by the channel and CFO estimation errors at time instant $n$, for $n \in \mathcal{D} \triangleq \{d+1, \ldots, n_1+d-1\} \cup \{n_2+d+1, \ldots, N+d\}$,[4] can be expressed as

$$\mathrm{EMSE}_n(\hat{\mathbf{f}}, \hat{\omega}) \approx \mathbf{T}_1 + \mathbf{T}_2(n) + \mathbf{T}_3(n) \tag{29}$$

where $\mathbf{T}_1$, $\mathbf{T}_2(n)$, and $\mathbf{T}_3(n)$ are defined in (30)–(32) at the bottom of the page,

$$\mathbf{D}'_{n:n-M} \triangleq \mathrm{diag}((n-\xi), \ldots, (n-M-\xi)) \tag{33}$$

and $R \triangleq N_{\mathrm{tr}} - L$.

*Proof:* The proof is provided in the Appendix. □

*Remark 1:* Term $\mathbf{T}_1$ involves only the channel estimation error second-order statistics and is the EMSE that would result if the mismatched equalizer were applied to perfectly CFO-corrected channel output samples [5, eq. (28)]. Term $\mathbf{T}_2(n)$ involves only the CFO estimation error variance and is the EMSE that would result if the ideal MMSE equalizer were applied to imperfectly CFO-corrected samples. Term $\mathbf{T}_3(n)$ involves both the channel and the CFO estimation errors.

[4]We do not compute the EMSE for the TS $\{a_n\}_{n=n_1}^{n_2}$.

## V. "Small Ideal MMSE" Assumption

Expression (29) is complicated and difficult to interpret. In order to derive a simple and insightful EMSE approximation, we assume that the *ideal* MMSE is sufficiently small, i.e., the equalizer length is sufficiently large, the SNR is sufficiently high and the delay is chosen carefully. This assumption is of high practical importance because it refers to the cases where the MMSE linear equalizer is *effective*. Under this assumption, vector $\mathbf{r}$, defined in (28), becomes "small." More specifically, it has been proved in [5, eq. (29) ] that $\|\mathbf{r}\|_2^2 \le \mathrm{MMSE}$, which implies that $\|\mathbf{r}\|_2 = \mathcal{O}\left(\sqrt{\mathrm{MMSE}}\right)$. Thus, terms that involve matrix $\mathbf{R}$, which is constructed by vector $\mathbf{r}$, are "small" compared with terms that involve matrix $\mathbf{G}$.[5] Thus, $\mathbf{T}_1$ and $\mathbf{T}_3(n)$ of (30) and (32), respectively, can be approximated as

$$\mathbf{T}_1 \approx \mathrm{tr}\left(\mathbf{R}_z^{-1}\mathbf{G}\mathbf{\Psi}'^*\mathbf{G}^H\right) \tag{34}$$

$$\begin{aligned}\mathbf{T}_3(n) \approx\; & -2\sigma_{\Delta\omega}^2 \,\mathrm{Re}\Big\{\mathbf{h}'^T\mathbf{A}^T\mathbf{K}'\mathbf{A}^*(\mathbf{A}^H\mathbf{A})^{-T} \\ & \times \mathbf{G}^H\mathbf{R}_z^{-1}\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\Big\}.\end{aligned} \tag{35}$$

### A. Time-Average EMSE

In the sequel, as done in [3], we study the EMSE time-average across the time instances that correspond to the unknown data

$$\mathrm{EMSE}(\hat{\mathbf{f}}, \hat{\omega}) \triangleq \frac{1}{|\mathcal{D}|}\sum_{n\in\mathcal{D}} \mathrm{EMSE}_n(\hat{\mathbf{f}}, \hat{\omega}). \tag{36}$$

If we write

$$\mathbf{D}'_{n:n-M} = (n-\xi)\,\mathbf{I}_{M+1} - \mathbf{D}_M \tag{37}$$

then terms $\mathbf{T}_2(n)$ of (31) and $\mathbf{T}_3(n)$ of (35) become

$$\begin{aligned}\mathbf{T}_2(n) = \sigma_{\Delta\omega}^2\Big[&(n-\xi)^2\,\mathrm{Re}\left\{\mathbf{f}^H\mathbf{H}'\mathbf{e}_d\right\} - 2(n-\xi)\,\mathrm{Re}\left\{\mathbf{f}^H\mathbf{D}_M\mathbf{H}'\mathbf{e}_d\right\} \\ &+ \mathrm{Re}\left\{\mathbf{f}^H\mathbf{D}_M^2\mathbf{H}'\mathbf{e}_d\right\}\Big]\end{aligned} \tag{38}$$

[5]See the discussion before (30) of [5].

---

$$\begin{aligned}\mathrm{MSE}_n(\hat{\mathbf{f}}, \hat{\omega}) &\triangleq \mathcal{E}_{a,w}\left[|\,\hat{e}_n|^2\right] \\ &= \hat{\mathbf{f}}^H\left(\mathbf{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{H}'^H\mathbf{\Gamma}_{n:n-M}^H(-\Delta\omega) + \sigma_w^2\mathbf{I}_{M+1}\right)\hat{\mathbf{f}} - 2\mathrm{Re}\{\mathrm{e}^{j\Delta\omega\xi}\hat{\mathbf{f}}^H\mathbf{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{e}_d\} + 1.\end{aligned} \tag{25}$$

---

$$\mathbf{T}_1 \triangleq \mathrm{tr}\left(\mathbf{R}_z^{-1}\left(\mathbf{R}^*\mathbf{\Psi}'\mathbf{R}^T + \mathbf{G}\mathbf{\Psi}'^*\mathbf{G}^H + \mathbf{G}\mathbf{\Psi}_t'^*\mathbf{R}^T + \mathbf{R}^*\mathbf{\Psi}_t'\mathbf{G}^H\right)\right) \tag{30}$$

$$\mathbf{T}_2(n) \triangleq \sigma_{\Delta\omega}^2 \mathrm{Re}\{\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\} \tag{31}$$

$$\mathbf{T}_3(n) \triangleq 2\sigma_{\Delta\omega}^2 \mathrm{Re}\left\{\mathbf{h}'^H\mathbf{A}^H\mathbf{K}'\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{R}^T\mathbf{R}_z^{-1}\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d - \mathbf{h}'^T\mathbf{A}^T\mathbf{K}'\mathbf{A}^*(\mathbf{A}^H\mathbf{A})^{-T}\mathbf{G}^H\mathbf{R}_z^{-1}\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\right\} \tag{32}$$

TABLE I
CHANNEL IMPULSE RESPONSE $\mathbf{h}$

| |
|---|
| $h_0 = -0.1538 + 0.4229 * j$ |
| $h_1 = -0.5923 - 0.0134 * j$ |
| $h_2 = 0.0446 + 0.1164 * j$ |
| $h_3 = 0.1023 + 0.0621 * j$ |
| $h_4 = -0.4077 - 0.0664 * j$ |
| $h_5 = 0.4235 + 0.2581 * j$ |

$$\mathbf{T}_3(n) \approx -2\sigma_{\Delta\omega}^2 \mathrm{Re}\left\{\mathbf{h}'^T \mathbf{A}^T \mathbf{K}' \mathbf{A}^* (\mathbf{A}^H \mathbf{A})^{-T} \times \mathbf{G}^H \mathbf{R}_z^{-1} \left((n-\xi)\,\mathbf{I}_{M+1} - \mathbf{D}_M\right) \mathbf{H}' \mathbf{e}_d \right\}. \tag{39}$$

If we define[6]

$$\mathcal{C}_1 \triangleq \frac{1}{|\mathcal{D}|} \sum_{n\in\mathcal{D}} n^2, \quad \mathcal{C}_2 \triangleq \frac{1}{|\mathcal{D}|} \sum_{n\in\mathcal{D}} n \tag{40}$$

then

$$\mathbf{T}_2 \triangleq \frac{1}{|\mathcal{D}|}\, \mathbf{T}_2(n) = \sigma_{\Delta\omega}^2 \Bigg[ \underbrace{\left(\mathcal{C}_1 - 2\,\mathcal{C}_2\xi + \xi^2\right) \mathrm{Re}\{\mathbf{f}^H \mathbf{H}' \mathbf{e}_d\}}_{\mathbf{T}_{21}} \underbrace{-2(\mathcal{C}_2-\xi)\mathrm{Re}\{\mathbf{f}^H \mathbf{D}_M \mathbf{H}' \mathbf{e}_d\}}_{\mathbf{T}_{22}} \underbrace{+\mathrm{Re}\{\mathbf{f}^H \mathbf{D}_M^2 \mathbf{H}' \mathbf{e}_d\}}_{\mathbf{T}_{23}} \Bigg] \tag{41}$$

and

$$\mathbf{T}_3 \triangleq \frac{1}{|\mathcal{D}|} \mathbf{T}_3(n) \approx -2\sigma_{\Delta\omega}^2 \mathrm{Re}\left\{ \mathbf{h}'^T \mathbf{A}^T \mathbf{K}' \mathbf{A}^* (\mathbf{A}^H \mathbf{A})^{-T} \mathbf{G}^H \mathbf{R}_z^{-1} \times \left( \underbrace{(\mathcal{C}_2 - \xi)}_{t_{31}} \mathbf{I}_{M+1} - \mathbf{D}_M \right) \mathbf{H}' \mathbf{e}_d \right\}. \tag{42}$$

### B. A Simple Approximation

Both $\mathbf{T}_2$ and $\mathbf{T}_3$ depend on $\xi$. It turns out that there does *not* exist a *unique, channel independent,* $\xi$ that is optimal, i.e., always attains minimum EMSE. If we put $\xi = \mathcal{C}_2$, then term $\mathbf{T}_{21}$ is minimized[7] and terms $\mathbf{T}_{22}$ and $t_{31}$ vanish. In the sequel, we use this value of $\xi$,[8] which implies that the TS is placed "close to the middle" of the packet; indeed, using the definition of $\xi$ after (16) and the fact that $\xi = \mathcal{C}_2$, it can be shown that $n_1 \approx \frac{N-N_{\mathrm{tr}}}{2} + d - \frac{L}{2}$. Then, if we define

$$\mathcal{C} \triangleq \mathcal{C}_1 - \mathcal{C}_2^2 \tag{43}$$

we obtain

$$\mathbf{T}_2 = \sigma_{\Delta\omega}^2 \left[ \mathcal{C}\, \mathrm{Re}\{\mathbf{f}^H \mathbf{H}' \mathbf{e}_d\} + \mathrm{Re}\{\mathbf{f}^H \mathbf{D}_M^2 \mathbf{H}' \mathbf{e}_d\} \right] \tag{44}$$

[6]Observe that $\mathcal{C}_1 = O(N^2)$, while $\mathcal{C}_2 = O(N)$.

[7]Observe that $\mathbf{T}_{21} = O(N^2)$, while the other component terms of $\mathbf{T}_2$ and $\mathbf{T}_3$ are much smaller.

[8]However, we do not claim that this value is optimal, in general.

$$\mathbf{T}_3 \approx 2\sigma_{\Delta\omega}^2\, \mathrm{Re}\left\{ \mathbf{h}'^T \mathbf{A}^T \mathbf{K}' \mathbf{A}^* (\mathbf{A}^H \mathbf{A})^{-T}\ \mathbf{G}^H \mathbf{R}_z^{-1} \mathbf{D}_M \mathbf{H}' \mathbf{e}_d \right\}. \tag{45}$$

Thus, the EMSE time-average is approximately equal to the sum of the three terms in (34), (44), and (45), which is still complicated. In the Appendix, we prove the following result.

*Proposition 2:* If $\xi = \mathcal{C}_2$, $N_{\mathrm{tr}}$ is sufficiently small with respect to $N$, and matrices $\mathbf{A}$ and $\mathbf{H}'$ are not very ill-conditioned, then

$$\mathrm{EMSE}(\hat{\mathbf{f}}, \hat{\omega}) \simeq \mathcal{C}\sigma_{\Delta\omega}^2. \tag{46}$$

That is, the EMSE is approximately proportional to the CFO estimation error variance with the factor of proportionality $\mathcal{C}$ being *independent* of the TS. Thus, TSs that are optimal for CFO estimation seem very good candidates for *joint* channel and CFO estimation.[9]

*Remark 2:* In the Appendix, we essentially prove that $\mathrm{EMSE} \approx \mathbf{T}_2$. Recall that $\mathbf{T}_2$ is the EMSE that would result if a perfect equalizer were applied to imperfectly CFO-corrected output samples. Thus, (46) implies that, under the stated assumptions, the main cause of the performance degradation is the imperfectly canceled CFO.

*Remark 3:* In the proof, we assume that $k_2(\mathbf{H}')$ is not "very large." By construction, if just one of the elements of $\mathbf{h}$ is nonzero, then the rows of $\mathbf{H}$ are linearly independent and, thus, $\mathbf{H}$ has full rank. Thus, in general, $\mathbf{H}$ is not close to rank deficient matrices and its condition number is not "very large."

*Remark 4:* It turns out that, for *fixed* training positions, the EMSE remains the same *irrespective* of the value of $\xi$ in (15). Of course, $\mathbf{T}_1$, $\mathbf{T}_2$, and $\mathbf{T}_3$ depend on $\xi$. Considering $\mathbf{h}'$ instead of $\mathbf{h}$ leads to "accurate" channel estimates and, thus, to "small" $\mathbf{T}_1$. Setting $\xi = \mathcal{C}_2$, that is, putting the TS "at the middle" of the packet, has two effects. The first is that it makes $\mathbf{T}_2$ much larger than $\mathbf{T}_3$ leading to the simple expression (46). The second, and more important, is that it minimizes $\mathbf{T}_{21}$, which is the most significant EMSE term. Thus, it leads to good performance.

## VI. SIMULATION RESULTS

In our simulations, we set the equalizer order $M = 12$, the delay $d = 5$, the packet length $N = 250$ and the TS length $N_{\mathrm{tr}} = 30$. The data symbols are i.i.d. BPSK. The training symbols, which are also i.i.d. BPSK, have been placed close to the middle of the transmitted packet, i.e., $\xi = \mathcal{C}_2$. The binary sequence we use corresponds to the hexadecimal number $172D97E1$. When we use a single channel realization, we use the channel of Table I while when we average over the channels we use i.i.d. $h_i \sim \mathcal{CN}\left(0, \frac{1}{L+1}\right)$, $i = 0, \ldots, L$.

In Fig. 1, we plot the EMSE versus the time $n$, for SNR equal to 25 dB for the channel of Table I (as mentioned above, we do not compute the EMSE for the known training symbols). The experimentally computed EMSE and the EMSE theoretical approximation (29) practically coincide. We observe that the EMSE *increases as we move away from the training symbol positions*. We also plot the EMSE theoretical approximation (29) for $n_1 = 1$ and $n_1 = N - N_{\mathrm{tr}} + 1$, i.e., the training block placed at the beginning and at the end of the packet, respectively. Obviously, placing the TS close to the middle of the transmitted packet leads to significantly smaller maximum and time-average EMSE.

[9]Optimal TS design for CFO estimation has been extensively studied; see, for example, [8]–[11]. This topic is beyond the scope of this paper.

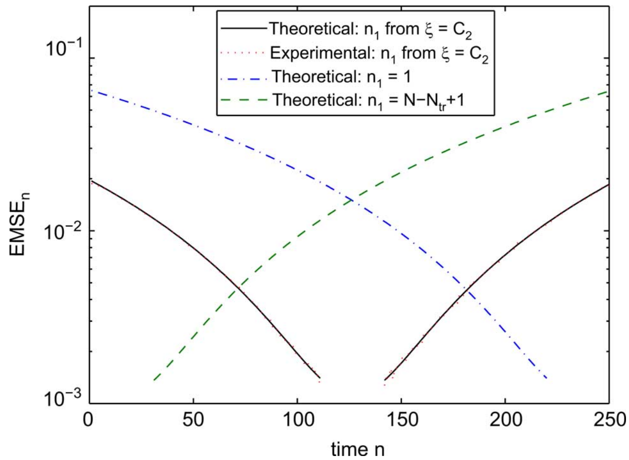


Fig. 1. Experimental EMSE and EMSE theoretical approximation (29) versus $n$, for different TS positions.

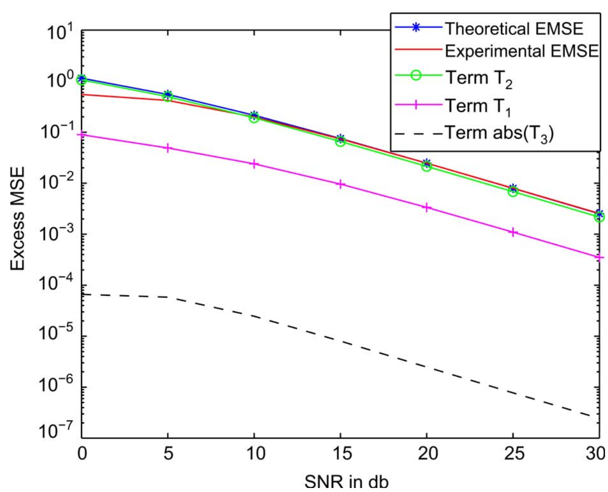


Fig. 2. Experimental EMSE, EMSE theoretical approximation (29) and terms $\mathbf{T}_1$, $\mathbf{T}_2$, and $\mathbf{T}_3$.

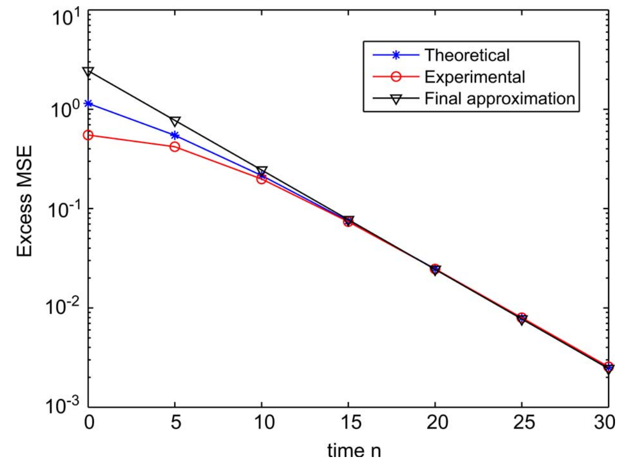


Fig. 3. Experimental, theoretical and approximate EMSE (46).

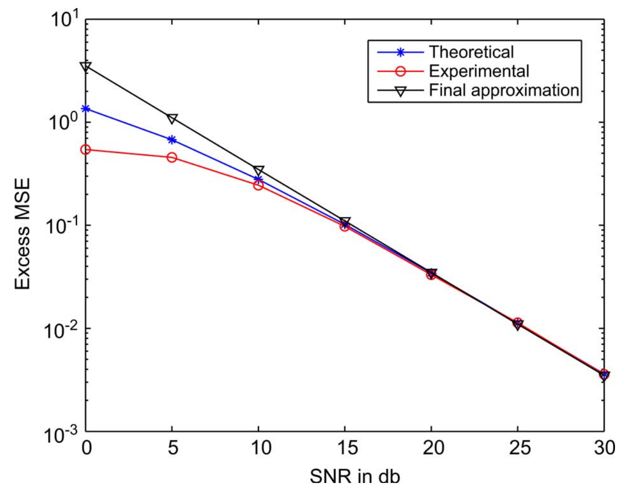


Fig. 4. Experimental, theoretical and approximate EMSE (46), averaged over random channels.

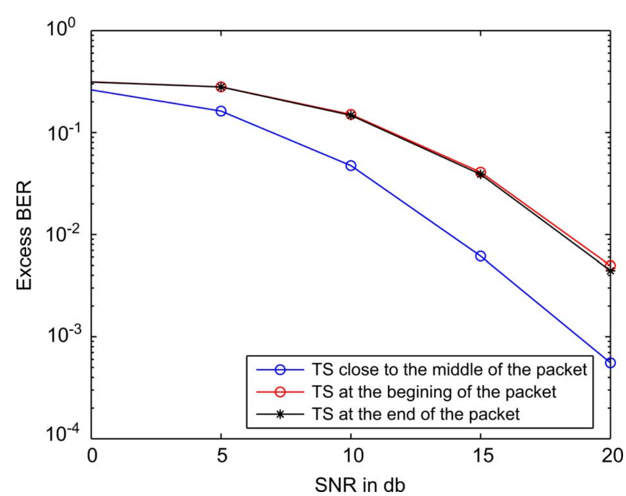


Fig. 5. Experimentally computed average excess BERs for different TS positions.

In Fig. 2, we plot the experimentally computed time-average EMSE, the time-average of the EMSE theoretical approximation in (29), and the time-averages of the three EMSE terms $\mathbf{T}_1$, $\mathbf{T}_2$, and $\mathbf{T}_3$ of (30), (31), and (32), respectively, for the channel of Table I. We observe that approximation (29) practically coincides with the true EMSE for SNR higher than 10 dB. We observe that $\mathbf{T}_2$ is very close to the EMSE, while terms $\mathbf{T}_1$ and $\mathbf{T}_3$ are much smaller.

In Fig. 3, we plot the experimental, the theoretical and the approximate EMSE (46) for the channel of Table I, while in Fig. 4, we plot the experimental average of the same quantities over $10^4$ channel realizations. In both cases, we observe that the very simple and informative expression of (46) is a very good EMSE approximation.

In Fig. 5, we plot the experimental average, over $10^4$ channel realizations, excess BER for different training positions. We observe that the placement of the TS close to the middle of the packet leads to smaller BER.

## VII. Conclusion

We considered the impact of the channel and CFO estimation errors on the performance of single-carrier frequency-selective systems with MMSE equalizers. We uncovered that, in many cases of high practical importance, the imperfectly canceled CFO is the main performance degradation cause. In these cases, the EMSE is approximately proportional to the CFO estimation error variance, with the factor of proportionality being independent of the TS. Thus, optimal TS design for CFO estimation is also highly relevant for *joint* CFO and channel estimation. We also highlighted the fact that placing the single-block TS at the middle of the packet is a good practice. An interesting future topic is the study of multi-block TSs.

## Appendix

*Proof of Proposition 1:* If we use expression $\Delta\mathbf{f} \triangleq \hat{\mathbf{f}} - \mathbf{f}$ in (25), we get

$$\begin{aligned}
&\mathrm{MSE}_n(\hat{\mathbf{f}}, \hat{\omega}) \\
&= \mathbf{f}^H \left(\boldsymbol{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{H}'^H\boldsymbol{\Gamma}^H_{n:n-M}(-\Delta\omega) + \sigma_w^2\mathbf{I}_{M+1}\right)\mathbf{f} \\
&\quad + \Delta\mathbf{f}^H\left(\boldsymbol{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{H}'^H\boldsymbol{\Gamma}^H_{n:n-M}(-\Delta\omega) + \sigma_w^2\mathbf{I}_{M+1}\right)\Delta\mathbf{f} \\
&\quad + 2\,\mathrm{Re}\Big\{\mathbf{f}^H\Big(\boldsymbol{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{H}'^H\boldsymbol{\Gamma}^H_{n:n-M}(-\Delta\omega) \\
&\qquad\qquad + \sigma_w^2\mathbf{I}_{M+1}\Big)\Delta\mathbf{f}\Big\} \\
&\quad - 2\,\mathrm{Re}\{\mathrm{e}^{j\Delta\omega\xi}\mathbf{f}^H\boldsymbol{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{e}_d\} \\
&\quad - 2\,\mathrm{Re}\{\mathrm{e}^{j\Delta\omega\xi}\Delta\mathbf{f}^H\boldsymbol{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{e}_d\} + 1.
\end{aligned} \tag{47}$$

We define

$$\boldsymbol{\Gamma}'_{n:n-M}(-\Delta\omega) \triangleq \mathrm{e}^{j\Delta\omega\xi}\boldsymbol{\Gamma}_{n:n-M}(-\Delta\omega). \tag{48}$$

Using the expression $\exp(x) = 1 + x + \frac{x^2}{2} + O(x^3)$, we obtain

$$\begin{aligned}
\boldsymbol{\Gamma}'_{n:n-M}(-\Delta\omega) = \mathbf{I}_{M+1} - j\Delta\omega\mathbf{D}'_{n:n-M} \\
-\frac{1}{2}(\Delta\omega)^2\mathbf{D}'^2_{n:n-M} + \mathcal{O}_p\left(\frac{n^3\sigma_w^3}{R^{\frac{9}{2}}}\right).
\end{aligned}$$

After some straightforward calculations, we obtain

$$\begin{aligned}
\mathrm{MSE}_n(\hat{\mathbf{f}}, \hat{\omega}) = {} & \mathrm{MSE}(\mathbf{f}, \omega) + \Delta\mathbf{f}^H\mathbf{R}_z\Delta\mathbf{f} \\
& + (\Delta\omega)^2\mathrm{Re}\{\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\} \\
& + 2\mathrm{Re}\{j\Delta\omega\Delta\mathbf{f}^H\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\} \\
& + g(\Delta\omega) + g_1\left((\Delta\omega)^3\right) \\
& + \mathcal{O}_p\left(\frac{n^2\sigma_w^3}{R^{\frac{7}{2}}}\right) + \mathcal{O}_p\left(\frac{M^2\sigma_w^2}{R^3}\right)
\end{aligned}$$

where $g(\Delta\omega)$ and $g_1\left((\Delta\omega)^3\right)$ denote terms that are linear functions of $\Delta\omega$ and $(\Delta\omega)^3$, respectively. In the sequel, we ignore the $\mathcal{O}_p(\cdot)$ terms. Then

$$\begin{aligned}
&\mathrm{EMSE}_n(\hat{\mathbf{f}}, \hat{\omega}) \\
&= \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\Big[\Delta\mathbf{f}^H\mathbf{R}_z\Delta\mathbf{f} + \Delta\omega^2\mathrm{Re}\{\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\} \\
&\qquad + 2\mathrm{Re}\{j\Delta\omega\Delta\mathbf{f}^H\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\}\Big].
\end{aligned} \tag{49}$$

The three terms of (49) are computed as follows:

$$\begin{aligned}
\mathbf{T}_1 &\triangleq \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\mathbf{f}^H\mathbf{R}_z\Delta\mathbf{f}\right] \\
&\overset{(27)}{=} \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\Big[\left(\Delta\mathbf{h}'^H\mathbf{R}^T + \Delta\mathbf{h}'^T\mathbf{G}^H\right) \\
&\qquad \times\mathbf{R}_z^{-1}\left(\mathbf{R}^*\Delta\mathbf{h}' + \mathbf{G}\Delta\mathbf{h}'^*\right)\Big] \\
&\overset{(10),(11)}{=} \mathrm{tr}\Big(\mathbf{R}_z^{-1}\Big(\mathbf{R}^*\boldsymbol{\Psi}'\mathbf{R}^T + \mathbf{G}\boldsymbol{\Psi}'^*\mathbf{G}^H \\
&\qquad + \mathbf{G}\boldsymbol{\Psi}_t'^*\mathbf{R}^T + \mathbf{R}^*\boldsymbol{\Psi}'_t\mathbf{G}^H\Big)\Big). \\
\mathbf{T}_2(n) &\triangleq \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\omega^2\mathrm{Re}\{\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\}\right] \\
&= \sigma^2_{\Delta\omega}\mathrm{Re}\{\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\}. \\
\mathbf{T}_3(n) &\triangleq \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[2\mathrm{Re}\{j\Delta\omega\Delta\mathbf{f}^H\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\}\right] \\
&\overset{(27)}{=} 2\mathrm{Re}\Big\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\Big[j\Delta\omega\left(\Delta\mathbf{h}'^H\mathbf{R}^T + \Delta\mathbf{h}'^T\mathbf{G}^H\right) \\
&\qquad \times\mathbf{R}_z^{-1}\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\Big]\Big\} \\
&= 2\sigma^2_{\Delta\omega}\,\mathrm{Re}\Big\{\mathbf{h}'^H\mathbf{A}^H\mathbf{K}'\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1} \\
&\qquad \times\mathbf{R}^T\mathbf{R}_z^{-1}\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d \\
&\qquad - \mathbf{h}'^T\mathbf{A}^T\mathbf{K}'\mathbf{A}^*(\mathbf{A}^H\mathbf{A})^{-T} \\
&\qquad \times\mathbf{G}^H\mathbf{R}_z^{-1}\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\Big\}.
\end{aligned}$$

This proves Proposition 1. □

*Proof of Proposition 2:*

*Term* $\mathbf{T}_1$*:* From (34) and the definition of $\mathbf{G}$ [below (28)], we obtain

$$\mathbf{T}_1 \approx \mathrm{tr}\left(\boldsymbol{\Psi}'^*\mathbf{G}^H\mathbf{R}_z^{-1}\mathbf{G}\right) = \mathrm{tr}\left(\boldsymbol{\Psi}'^*\mathbf{F}^*\mathbf{H}'^H\mathbf{R}_z^{-1}\mathbf{H}'\mathbf{F}^T\right).$$

Using the facts i) $\mathbf{A} \geq \mathbf{B}$ implies that $\mathbf{A}^{-1} \leq \mathbf{B}^{-1}$[6, p. 471] and ii) $\mathbf{P}_{\mathcal{R}(\mathbf{H}'^H)} \leq \mathbf{I}_{L+M+1}$, we obtain

$$\begin{aligned}
\mathbf{H}'^H\left(\mathbf{H}'\mathbf{H}'^H + \sigma_w^2\mathbf{I}\right)^{-1}\mathbf{H}' &\leq \mathbf{H}'^H\left(\mathbf{H}'\mathbf{H}'^H\right)^{-1}\mathbf{H}' \\
&= \mathbf{P}_{\mathbf{H}'^H} \leq \mathbf{I}_{M+L+1}.
\end{aligned} \tag{50}$$

Using (50) and $\mathrm{tr}\left(\mathbf{A}\mathbf{B}\mathbf{A}^H\right) \leq \lambda_{\max}(\mathbf{B})\,\mathrm{tr}\left(\mathbf{A}\mathbf{A}^H\right)$ [7, p. 44], we obtain [recall the definition of $\mathbf{R}_z$ in (19)]

$$\begin{aligned}
\mathbf{T}_1 &\approx \mathrm{tr}\left(\boldsymbol{\Psi}'^*\mathbf{F}^*\mathbf{H}'^H\mathbf{R}_z^{-1}\mathbf{H}'\mathbf{F}^T\right) \\
&\lessapprox \mathrm{tr}\left(\boldsymbol{\Psi}'^*\mathbf{F}^*\mathbf{F}^T\right) \leq \lambda_{\max}\left(\boldsymbol{\Psi}'\right)\mathrm{tr}\left(\mathbf{F}^*\mathbf{F}^T\right) \\
&= \lambda_{\max}\left(\boldsymbol{\Psi}'\right)\|\mathbf{F}\|_F^2 = \lambda_{\max}\left(\boldsymbol{\Psi}'\right)(L+1)\|\mathbf{f}\|_2^2.
\end{aligned} \tag{51}$$

Using asymptotic arguments, it can be shown [12] that the first term of $\boldsymbol{\Psi}'$ is much larger than the second. Thus, $\lambda_{\max}\left(\boldsymbol{\Psi}'\right) \approx \frac{\sigma_w^2}{\lambda_{\min}(\mathbf{A}^H\mathbf{A})}$ and

$$\mathbf{T}_1 \lessapprox \frac{(L+1)\|\mathbf{f}\|_2^2\,\sigma_w^2}{\lambda_{\min}(\mathbf{A}^H\mathbf{A})}. \tag{52}$$

*Term* $\mathbf{T}_2$*:* Term $\mathrm{Re}\{\mathbf{f}^H\mathbf{H}'\mathbf{e}_d\}$ is the $(d+1)$-st coefficient of the combined (channel-equalizer) impulse response. Using the definition of $\mathbf{f}$ in (19) and expression (50), it can be shown that $\mathrm{Re}\{\mathbf{f}^H\mathbf{H}'\mathbf{e}_d\}$ is always smaller than 1, and, under the small MMSE assumption, it is very close to 1. Thus, $\mathbf{T}_{21} \approx \mathcal{C}$. On the other hand, using the definition of $\mathbf{f}$ in (19), the submultiplicative property of the matrix norms, and the singular value decomposition (SVD) of $\mathbf{H}'$, it can be shown that $\mathbf{T}_{23} = 2\mathrm{Re}\{\mathbf{f}^H \mathbf{D}_M^2 \mathbf{H}' \mathbf{e}_d\} \leq 2M^2 k_2(\mathbf{H}')$. If $N$ is sufficiently large with respect to $M$ and $\mathbf{H}'$ is not very ill-conditioned, then $\mathbf{T}_{21} \gg \mathbf{T}_{23}$ and

$$\mathbf{T}_2 \simeq \mathcal{C}\sigma_{\Delta\omega}^2. \tag{53}$$

*Term* $\mathbf{T}_3$*:* Using the SVD of $\mathbf{A}$, it can be shown that

$$\|\mathbf{A}^T \mathbf{K}' \mathbf{A}^* (\mathbf{A}^H \mathbf{A})^{-T}\|_2 \leq \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})} \|\mathbf{K}'\|_2 = \frac{R}{2} k_2(\mathbf{A}).$$

In the same manner, we can show (recall that $\mathbf{G} = \mathbf{H}'\mathbf{F}^T$)

$$\|\mathbf{G}^H \mathbf{R}_z^{-1} \mathbf{D}_M \mathbf{H}' \mathbf{e}_d\|_2 \leq \|\mathbf{F}\|_2 \frac{\sigma_{\max}(\mathbf{H}')}{\sigma_{\min}(\mathbf{H}')} M = M \|\mathbf{F}\|_2 k_2(\mathbf{H}'). \tag{54}$$

Thus,

$$\mathbf{T}_3 \leq \left(MR \|\mathbf{F}\|_2 k_2(\mathbf{A}) k_2(\mathbf{H}')\right) \sigma_{\Delta\omega}^2. \tag{55}$$

*Comparison of* $\mathbf{T}_2$ *and* $\mathbf{T}_3$*:* If $N$ is sufficiently large and $\mathbf{A}$ and $\mathbf{H}'$ are not very ill-conditioned, then, from (53) (recall that $\mathcal{C} = O(N^2)$) and (55), we conclude that $\mathbf{T}_2 \gg \mathbf{T}_3$.

*Comparison of* $\mathbf{T}_1$ *and* $\mathbf{T}_2$*:* Using [8, eq. (10)], we can derive the following asymptotic expression:

$$\sigma_{\Delta\omega}^2 \approx \frac{6\,\sigma_w^2}{R^2\,\mathbf{h}^H \mathbf{A}^H \mathbf{A}\mathbf{h}}. \tag{56}$$

Thus

$$\mathbf{T}_2 \approx \frac{6\,\mathcal{C}\,\sigma_w^2}{R^2 \mathbf{h}^H \mathbf{A}^H \mathbf{A}\mathbf{h}}. \tag{57}$$

Using (52) and (57), we derive the following approximate bound:

$$\begin{aligned}\frac{\mathbf{T}_1}{\mathbf{T}_2} &\lessapprox \frac{(L+1)R^2\,\|\mathbf{f}\|_2^2\,\mathbf{h}^H \mathbf{A}^H \mathbf{A}\mathbf{h}}{6\,\mathcal{C}\,\lambda_{\min}(\mathbf{A}^H\mathbf{A})} \\ &\leq \frac{(L+1)R^2 \|\mathbf{f}\|_2^2 \lambda_{\max}(\mathbf{A}^H\mathbf{A}) \|\mathbf{h}\|_2^2}{6\,\mathcal{C}\,\lambda_{\min}(\mathbf{A}^H\mathbf{A})} \\ &= k_2(\mathbf{A}^H\mathbf{A})(L+1)\|\mathbf{f}\|_2^2\,\|\mathbf{h}\|_2^2\,\alpha\end{aligned} \tag{58}$$

where

$$\alpha \triangleq \frac{R^2}{6\,\mathcal{C}} = \mathcal{O}\left(\frac{R^2}{N^2}\right). \tag{59}$$

Thus, if $\alpha$ is sufficiently small, i.e., $R$ is sufficiently small with respect to $N$ (recall that $R = N_{\mathrm{tr}} - L$), and $\mathbf{A}$ is not very ill-conditioned, then term $\mathbf{T}_2 \gg \mathbf{T}_1$. Thus, $\mathbf{T}_2$ is much larger than $\mathbf{T}_1$ and $\mathbf{T}_3$. Proposition 2 is proved using (53).

## REFERENCES


[1] M. Morelli and U. Mengali, "Carrier frequency estimation for transmissions over selective channels," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1580–1589, Sep. 2000.

[2] P. Stoica and O. Besson, "Training sequence design for frequency offset and frequency-selective channel estimation," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1910–1917, Nov. 2003.

[3] P. Ciblat, P. Bianchi, and M. Ghogho, "Training sequence optimization for joint channel and frequency offset estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3424–3436, Aug. 2008.

[4] A. Sayed, *Fundamentals of Adaptive Filtering*. New York: IEEE-Wiley, 2003.

[5] A. P. Liavas and D. Tsipouridou, "On the performance of the mismatched MMSE and the LS linear equalizers," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3302–3311, Jul. 2007.

[6] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[7] H. Luetkepohl, *Handbook of Matrices*. New York: Wiley, 1996.

[8] O. Besson and P. Stoica, "Training sequence selection for frequency offset estimation in frequency-selective channels," *Digit. Signal Process.*, vol. 13, no. 1, pp. 106–127, Jan. 2003.

[9] Y. Ying and M. Ghogho, "Optimal pilot placement for frequency offset estimation and data detection in burst transmission systems," *IEEE Commun. Lett.*, vol. 9, no. 6, pp. 549–551, Jun. 2005.

[10] Y.-D. Kim, J. K. Lim, C. Suh, and Y. H. Lee, "Designing training sequences for carrier frequency estimation in frequency-selective channels," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 151–157, Jan. 2006.

[11] H. Minn, X. Fu, and V. Bhargava, "Optimal periodic training signal for frequency offset estimation in frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 54, pp. 1081–1096, Jun. 2006.

[12] A. P. Liavas and D. Tsipouridou, "Single-carrier systems with MMSE linear equalizers: Performance degradation due to channel and CFO estimation errors," Tech. Rep., Jun. 2011 [Online]. Available: http://www.telecom.tuc.gr/ liavas/