

# Single-carrier systems with MMSE linear equalizers: performance degradation due to channel and CFO estimation errors

Athanasios P. Liavas, *Member IEEE*, and Despoina Tspouridou

## 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 that

- 1) performance degradation is mainly caused by 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 training sequence.

We also highlight the fact that the placement of the training at the middle of the transmitted packet is a good practice.

## 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 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. Ciblat *et al.* considered a single-carrier system with an MMSE linear equalizer and computed the second-order statistics (power spectrum) of the TS that, under certain assumptions, minimizes the mean square estimation error [3].

In this work, we consider the same system but our aim is different. Performance degradation is caused by the fact that a *mismatched* MMSE linear equalizer is applied to channel output samples with *imperfectly canceled* CFO. Our aim is to uncover the relative importance of these error sources. We assume a *single-block* training and develop an asymptotic expression for the induced EMSE which, however, is difficult to interpret.<sup>1</sup>

Our main contribution lies in the fact that, assuming *small ideal MMSE*, we derive a simple and informative EMSE approximation, which reveals that

- 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.<sup>2</sup>

We also highlight the fact that the placement of the 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.  $\text{tr}(\cdot)$  denotes the trace operator,  $\text{Re}\{\cdot\}$  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  $\mathbf{P}_{\mathcal{R}(\mathbf{A})}^\perp$  denote, respectively, the orthogonal projector onto the column space of matrix  $\mathbf{A}$  and onto its orthogonal complement.

## II. CHANNEL AND CFO ESTIMATION

### A. The channel model

We consider a packet-based communication 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  $\phi$ . Then, the output at time

The authors are with the Department of Electronic and Computer Engineering, Technical University of Crete, GREECE. E-mail: (liavas, despoina)@telecom.tuc.gr. A. P. Liavas was partially supported by the FP-6 FET Open Project COOPCOM.

<sup>1</sup>We note that the same expression in terms of frequency domain quantities has been derived in [3].

<sup>2</sup>Thus, optimal TS design for CFO estimation is also highly relevant for *joint* channel and CFO estimation.

instant  $n$ , for  $n = 1, \dots, N + L$ , is

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

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

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

where

$$\mathbf{\Gamma}_{n:n-M}(\omega) \triangleq \text{diag}(e^{j\omega n}, \dots, e^{j\omega(n-M)}). \quad (3)$$

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

### B. Channel and CFO estimation

The  $N_{\text{tr}}$  consecutive symbols  $\mathbf{a}_{\text{tr}} \triangleq [a_{n_1} \cdots a_{n_2}]^T$ , with  $N_{\text{tr}} \triangleq n_2 - n_1 + 1$ , are used for training.<sup>3</sup> We collect the output samples that depend *only* on the training and construct

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

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

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

The joint ML CFO and channel estimates are [1]

$$\hat{\omega} = \underset{\tilde{\omega}}{\text{argmax}} \{ \mathbf{y}^H \mathbf{\Gamma}_{n_2:n_1+L}(\tilde{\omega}) \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{\Gamma}_{n_2:n_1+L}^H(\tilde{\omega}) \mathbf{y} \} \quad (6)$$

and

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

The estimation errors are  $\Delta\omega \triangleq \hat{\omega} - \omega$  and  $\Delta\mathbf{h} \triangleq \hat{\mathbf{h}} - \mathbf{h}$ . We assume that  $N_{\text{tr}}$  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

$$\mathbf{K} \triangleq \text{diag}(n_2, \dots, n_1 + L), \quad (8)$$

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

$$\sigma_{\Delta\omega}^2 \triangleq \mathcal{E}[(\Delta\omega)^2] = \frac{\sigma_w^2}{2 \text{tr}(\mathbf{h}^H \mathbf{A}^H \mathbf{K} \mathbf{P}_{\mathcal{R}(\mathbf{A})}^\perp \mathbf{K} \mathbf{A} \mathbf{h})} \quad (9)$$

$$\mathbf{\Psi} \triangleq \mathcal{E}[\Delta\mathbf{h} \Delta\mathbf{h}^H] = \sigma_w^2 (\mathbf{A}^H \mathbf{A})^{-1} + \sigma_{\Delta\omega}^2 (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{K} \mathbf{A} \mathbf{h} \mathbf{h}^H \mathbf{A}^H \mathbf{K} \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \quad (10)$$

$$\mathbf{\Psi}_t \triangleq \mathcal{E}[\Delta\mathbf{h} \Delta\mathbf{h}^T] = -\sigma_{\Delta\omega}^2 (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{K} \mathbf{A} \mathbf{h} \mathbf{h}^T \mathbf{A}^T \mathbf{K} \mathbf{A}^* (\mathbf{A}^H \mathbf{A})^{-T} \quad (11)$$

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

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

$$\mathbf{K} = n_2 \mathbf{I}_{N_{\text{tr}}-L} - \mathbf{D}_{N_{\text{tr}}-L-1}, \quad (13)$$

with  $\mathbf{D}_i \triangleq \text{diag}(0, 1, \dots, i)$ , then we can show that

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

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

<sup>3</sup>Training schemes with two or more blocks are beyond the scope of this work.

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 depends 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 term  $e^{j\omega\xi}$ , with  $\xi \triangleq n_1 + \frac{N_{\text{tr}}+L}{2}$ , i.e.,  $\xi$  is the middle position of  $\mathbf{y}$ , into channel  $\mathbf{h}$ , getting the ‘‘new channel’’  $\mathbf{h}' \triangleq e^{j\omega\xi}\mathbf{h}$ .<sup>4</sup> In this case, the channel output is expressed as

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

and (4) can be written as

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

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

$$\hat{\mathbf{h}}' = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{\Gamma}_{\frac{N_{\text{tr}}-L}{2}-1:-\frac{N_{\text{tr}}-L}{2}}^H(\hat{\omega}) \mathbf{y}. \quad (17)$$

We define

$$\mathbf{K}' \triangleq \text{diag}\left(\frac{N_{\text{tr}}-L}{2}-1, \dots, -\frac{N_{\text{tr}}-L}{2}\right) \quad (18)$$

and  $\Delta\mathbf{h}' \triangleq \hat{\mathbf{h}}' - \mathbf{h}'$ . The estimation error second-order statistics, denoted as  $\mathbf{\Psi}'$ ,  $\mathbf{\Psi}'_t$ ,  $\psi'$ , and  $\sigma_{\Delta\omega}^2$ , 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 perfectly cancel it before equalization. If we know the channel, then we can compute the order- $M$  delay- $d$  MMSE linear equalizer,  $\mathbf{f} \triangleq [f_0 \ \dots \ f_M]^T$ , as [4, Section 2.7.3]

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

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 MMSE symbol estimation error is

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

#### 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 their estimates as if they were the true values.

The mismatched MMSE equalizer is (see (19))

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

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

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

Vector  $\mathbf{s}_{n:n-M}$  which can be expressed as

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

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

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

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

<sup>4</sup>We shall say more on this topic later.

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

$$\mathbf{T}_1 \triangleq \text{tr}(\mathbf{R}_z^{-1} (\mathbf{R}^* \mathbf{\Psi}' \mathbf{R}^T + \mathbf{G} \mathbf{\Psi}'^* \mathbf{G}^H + \mathbf{G} \mathbf{\Psi}'^* \mathbf{R}^T + \mathbf{R}^* \mathbf{\Psi}' \mathbf{G}^H)) \quad (30)$$

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

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

#### IV. EMSE ANALYSIS

The EMSE at time instant  $n$  is defined as

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

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

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

where

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

$$\mathbf{r} \triangleq \mathbf{c} - \mathbf{e}_d \quad (28)$$

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  $n \in \mathcal{D} \triangleq \{d+1, \dots, n_1+d-1\} \cup \{n_2+d+1, \dots, N+d\}$ ,<sup>5</sup> can be expressed as*

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

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

$$\mathbf{D}'_{n:n-M} \triangleq \text{diag}((n-\xi), \dots, (n-M-\xi)) \quad (33)$$

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

**Proof:** The proof is provided in the Appendix.  $\square$

*Remark 1:* Term  $\mathbf{T}_1$  involves only the channel estimation error second-order statistics; it 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.

#### 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  $\|\mathbf{r}\|_2^2 \leq \text{MMSE}$ , which implies that  $\|\mathbf{r}\|_2 = \mathcal{O}(\sqrt{\text{MMSE}})$ . Thus, terms that involve matrix  $\mathbf{R}$ , which is constructed by vector  $\mathbf{r}$ , are “small” compared with terms that involve matrix  $\mathbf{G}$ .<sup>6</sup> Thus,  $\mathbf{T}_1$  and  $\mathbf{T}_3(n)$  of (30) and (32), respectively, can be approximated as

$$\mathbf{T}_1 \approx \text{tr}(\mathbf{R}_z^{-1} \mathbf{G} \mathbf{\Psi}'^* \mathbf{G}^H) \quad (34)$$

$$\mathbf{T}_3(n) \approx -2\sigma_{\Delta\omega}^2 \text{Re}\{\mathbf{h}'^T \mathbf{A}^T \mathbf{K}' \mathbf{A}^* (\mathbf{A}^H \mathbf{A})^{-T} \mathbf{G}^H \mathbf{R}_z^{-1} \mathbf{D}'_{n:n-M} \mathbf{H}' \mathbf{e}_d\}. \quad (35)$$

<sup>5</sup>We do not compute the EMSE for the training symbols  $a_n$ ,  $n = n_1, \dots, n_2$ .

<sup>6</sup>See the discussion before eq. (30) of [5].

### A. Time-average EMSE

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

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

If we write

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

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

$$\mathbf{T}_2(n) = \sigma_{\Delta\omega}^2 \left[ (n - \xi)^2 \text{Re}\{\mathbf{f}^H \mathbf{H}' \mathbf{e}_d\} - 2(n - \xi) \text{Re}\{\mathbf{f}^H \mathbf{D}_M \mathbf{H}' \mathbf{e}_d\} + \text{Re}\{\mathbf{f}^H \mathbf{D}_M^2 \mathbf{H}' \mathbf{e}_d\} \right] \quad (38)$$

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

If we define<sup>7</sup>

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

then

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

and

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

### 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<sup>8</sup> and terms  $\mathbf{T}_{22}$  and  $t_{31}$  vanish. In the sequel, we use this value of  $\xi$ ,<sup>9</sup> which implies that the training 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  $n_1 \approx \frac{N - N_{\text{tr}}}{2} + d - \frac{L}{2}$  (We prove it in the Appendix). Then, if we define

$$\mathcal{C} \triangleq \mathcal{C}_1 - \mathcal{C}_2^2 \quad (43)$$

we obtain

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

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

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 Appendix, we prove the following result.

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

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

That is, the EMSE is approximately proportional to the CFO estimation error variance with the factor of proportionality  $\mathcal{C}$  being *independent* of the training sequence. Thus, training sequences that are optimal for CFO estimation seem very good candidates for *joint* channel and CFO estimation.<sup>10</sup>

<sup>7</sup>Observe that  $\mathcal{C}_1 = O(N^2)$ , while  $\mathcal{C}_2 = O(N)$ .

<sup>8</sup>Observe that  $\mathbf{T}_{21} = O(N^2)$ , while the other component terms of  $\mathbf{T}_2$  and  $\mathbf{T}_3$  are much smaller.

<sup>9</sup>However, we do not claim that this value is optimal, in general.

<sup>10</sup>Optimal training sequence design for CFO estimation has been extensively studied; see, for example, [8]–[11]. This topic is beyond the scope of this paper.

**Table I**  
Channel Impulse Response  $\mathbf{h}$

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

*Remark 2:* In the Appendix, we essentially prove that  $\text{EMSE} \approx \mathbf{T}_2$ . 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 assumed 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, the values of  $\mathbf{T}_1$ ,  $\mathbf{T}_2$ , 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 training “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 this Section, we present simulation results for the channel of the Table I.<sup>11</sup> We set the equalizer order  $M = 12$ , the delay  $d = 5$ , the packet length  $N = 250$  and the TS length  $N_{\text{tr}} = 30$ . 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.

In Fig. 1, we plot the EMSE versus the time instances  $n$ , for SNR equal to 25 dB (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  $n_1 = N - N_{\text{tr}} + 1$ , i.e., the training block placed at the beginning and at the end of the packet, respectively. It is obvious that placing the TS close to the middle of the transmitted packet leads to significantly smaller maximum and time-average EMSE.

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  $\mathbf{T}_3$  in (30), (31) and (32). 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 EMSE, the theoretical EMSE and the simple EMSE approximation (46). We observe that the very simple and informative expression of (46) is indeed a very good EMSE approximation.

In Fig. 4 we plot the excess BER for the cases of  $n_1 = \frac{N - N_{\text{tr}}}{2} + d - \frac{L}{2}$ ,  $n_1 = 1$  and  $n_1 = N - N_{\text{tr}} + 1$ , i.e., the training block placed close to the middle, at the beginning and at the end of the transmitted packet, respectively. It is obvious that the placement of the training at the middle of the packet leads to significantly better BER performance.

In the sequel, we keep the parameters used for the previous simulation results but we take averages over different random channel realizations; we assume that the elements of  $\mathbf{h}$  are i.i.d, with  $h_i \sim \mathcal{CN}\left(0, \frac{1}{L+1}\right)$ , for  $i = 0, \dots, L$ .

In Fig. 5, 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  $\mathbf{T}_3$  in (30), (31) and (32), averaged over random channel realizations.

In Fig. 6 we plot the experimental EMSE, the theoretical EMSE and the simple EMSE approximation (46) averaged over random channels. We observe that the very simple and informative expression of (46) is indeed a very good EMSE approximation.

In Fig. 7 we plot the channel average excess BER for different training positions. We observe that again the placement of the TS close to the middle of the packet leads to better BER performance.

## VII. CONCLUSION

We considered the impact of the channel and CFO estimation errors on the performance of single-carrier systems with MMSE equalizers. We uncovered that, in many cases of high practical importance, the imperfectly canceled CFO is the main

<sup>11</sup>We observed analogous behavior in extensive simulation studies.

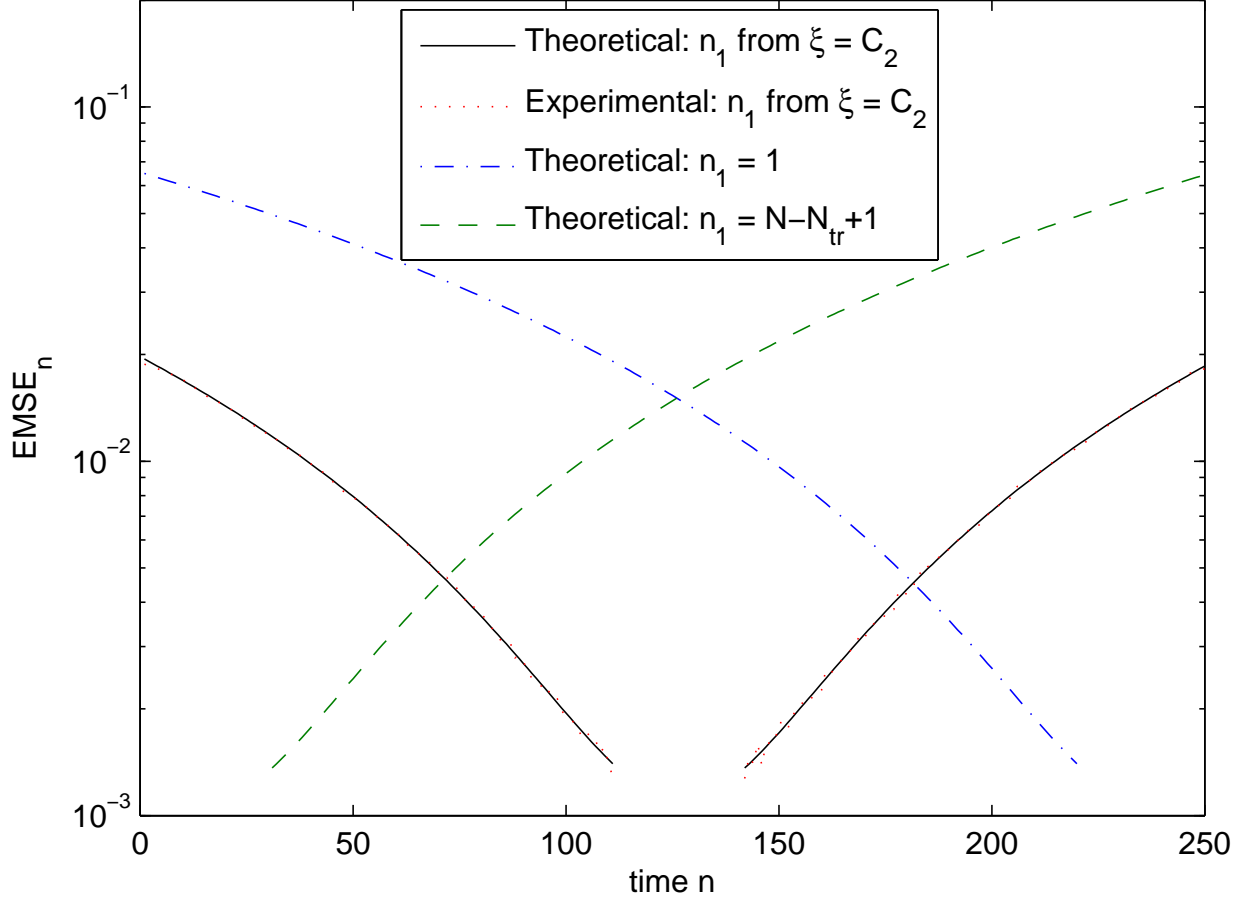


Fig. 1. Experimentally computed EMSE and EMSE theoretical approximation in (29) versus  $n$ , for different TS positions.

$$\begin{aligned}
 \text{MSE}_n(\hat{\mathbf{f}}, \hat{\omega}) &= \mathbf{f}^H \underbrace{\left( \mathbf{\Gamma}_{n:n-M}(-\Delta\omega) \mathbf{H}' \mathbf{H}'^H \mathbf{\Gamma}_{n:n-M}^H(-\Delta\omega) + \sigma_w^2 \mathbf{I}_{M+1} \right)}_{t_1} \mathbf{f} \\
 &+ \underbrace{\Delta \mathbf{f}^H \left( \mathbf{\Gamma}_{n:n-M}(-\Delta\omega) \mathbf{H}' \mathbf{H}'^H \mathbf{\Gamma}_{n:n-M}^H(-\Delta\omega) + \sigma_w^2 \mathbf{I}_{M+1} \right) \Delta \mathbf{f}}_{t_2} \\
 &+ 2 \operatorname{Re} \left\{ \mathbf{f}^H \underbrace{\left( \mathbf{\Gamma}_{n:n-M}(-\Delta\omega) \mathbf{H}' \mathbf{H}'^H \mathbf{\Gamma}_{n:n-M}^H(-\Delta\omega) + \sigma_w^2 \mathbf{I}_{M+1} \right) \Delta \mathbf{f}}_{t_3} \right\} \\
 &\underbrace{- 2 \operatorname{Re} \{ e^{j\Delta\omega\xi} \mathbf{f}^H \mathbf{\Gamma}_{n:n-M}(-\Delta\omega) \mathbf{H}' \mathbf{e}_d \}}_{t_4} \underbrace{- 2 \operatorname{Re} \{ e^{j\Delta\omega\xi} \Delta \mathbf{f}^H \mathbf{\Gamma}_{n:n-M}(-\Delta\omega) \mathbf{H}' \mathbf{e}_d \}}_{t_5} + 1.
 \end{aligned} \tag{47}$$

cause of the performance degradation. 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 training at the middle of the packet is a good practice. An interesting future topic is to consider multi-block training.

#### APPENDIX

##### A. Proof of Proposition 1:

If we use expression  $\Delta \mathbf{f} \triangleq \hat{\mathbf{f}} - \mathbf{f}$  in (25), we get (47) at the top of the next page. We define

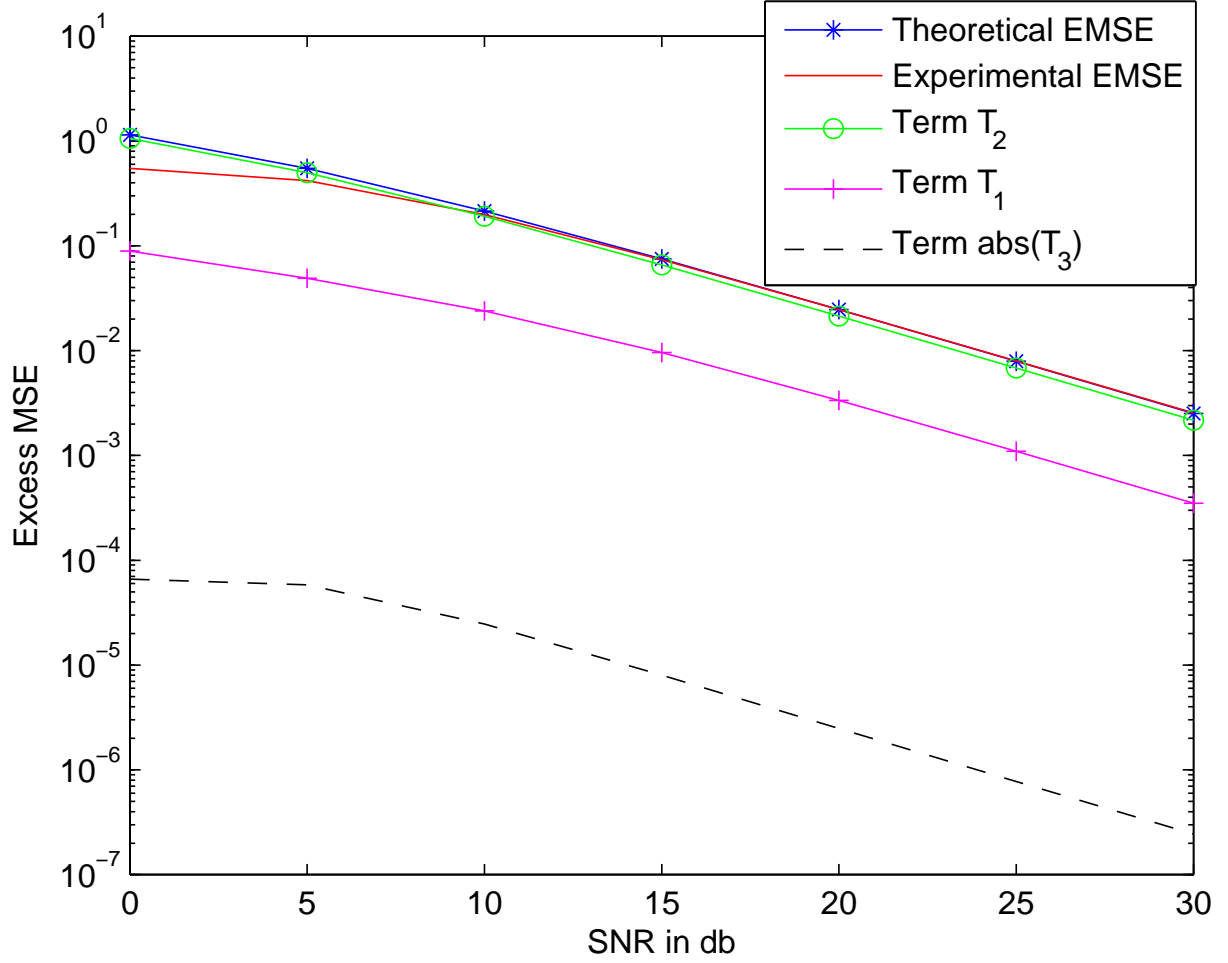


Fig. 2. Experimentally computed EMSE, EMSE theoretical approximation in (29) and the three terms  $\mathbf{T}_1$ ,  $\mathbf{T}_2$  and  $\mathbf{T}_3$ .

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

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

$$\mathbf{\Gamma}_{n:n-M}(-\Delta\omega) = \mathbf{I}_{M+1} - j\Delta\omega\mathbf{D}_{n:n-M} - \frac{1}{2}\Delta\omega^2\mathbf{D}_{n:n-M}^2 + \mathcal{O}_p\left(\frac{n^3\sigma_w^3}{R^{9/2}}\right) \quad (49)$$

$$\mathbf{\Gamma}'_{n:n-M}(-\Delta\omega) = \mathbf{I}_{M+1} - j\Delta\omega\mathbf{D}'_{n:n-M} - \frac{1}{2}(\Delta\omega)^2\mathbf{D}'_{n:n-M}{}^2 + \mathcal{O}_p\left(\frac{n^3\sigma_w^3}{R^{9/2}}\right). \quad (50)$$

We will write analytically the five terms defined in (47), using (49) and (50). We will also take the expected value of each term with respect to  $\Delta\mathbf{h}'$  and  $\Delta\omega$ .



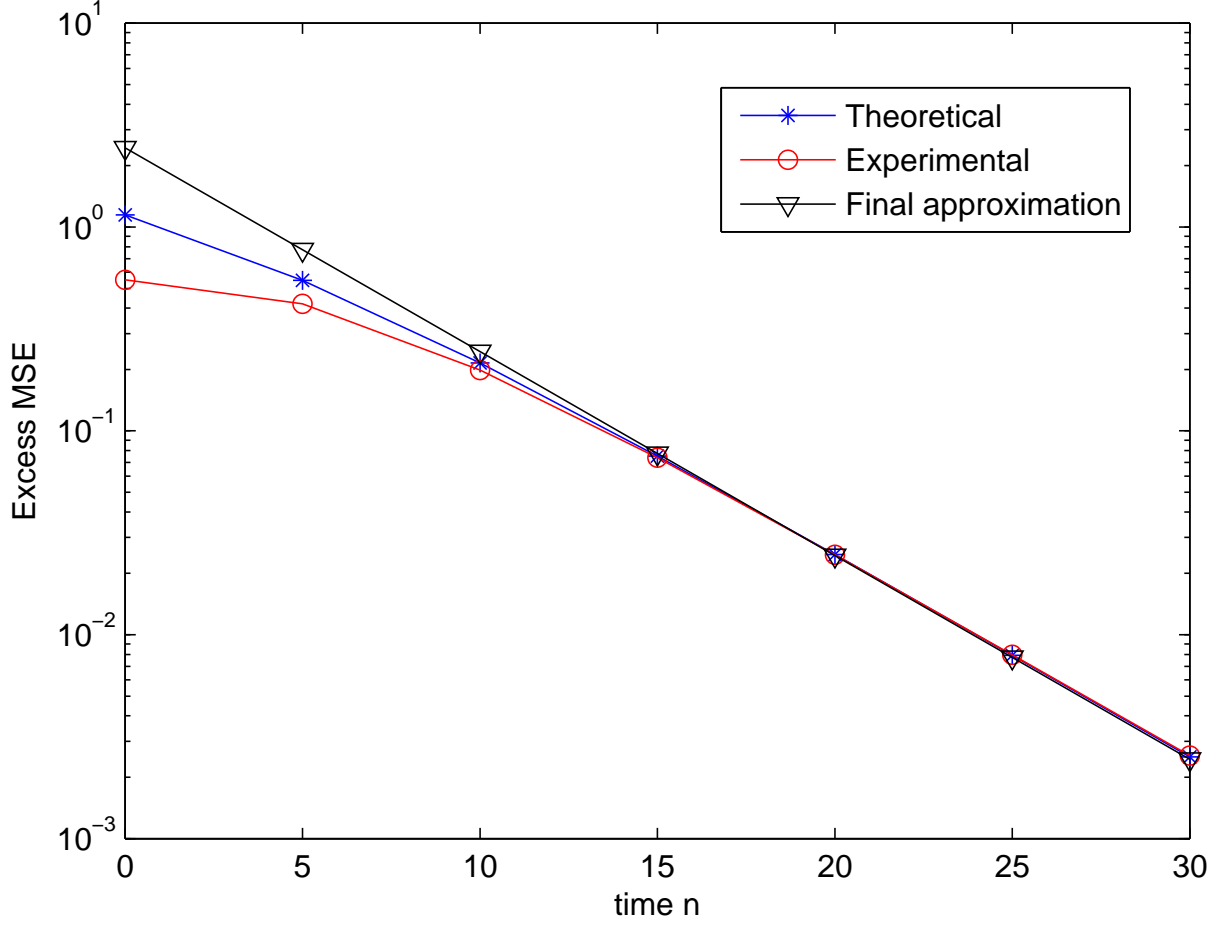


Fig. 3. Final EMSE theoretical expression in (46).

1) *Term  $t_1$* : Using (49) we obtain

$$\begin{aligned}
\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[t_1] &= \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ \mathbf{f}^H \left( \overset{(1)}{\mathbf{I}_{M+1}} - j\Delta\omega \overset{(2)}{\mathbf{D}_{n:n-M}} - \frac{1}{2}\Delta\omega^2 \overset{(3)}{\mathbf{D}_{n:n-M}^2} \right) \mathbf{H}'\mathbf{H}'^H \right. \\
&\quad \left. \times \left( \overset{(1)}{\mathbf{I}_{M+1}} + j\Delta\omega \overset{(2)}{\mathbf{D}_{n:n-M}} - \frac{1}{2}\Delta\omega^2 \overset{(3)}{\mathbf{D}_{n:n-M}^2} + \sigma_w^2 \mathbf{I}_{M+1} \right) \mathbf{f} \right] \\
&= \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ \mathbf{f}^H \left( \overset{(11)}{\mathbf{H}'\mathbf{H}'^H} + \sigma_w^2 \mathbf{I}_{M+1} \right) \mathbf{f} + j\Delta\omega \overset{(12)}{\mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}} \mathbf{f} \right. \\
&\quad - \frac{1}{2}\Delta\omega^2 \overset{(13)}{\mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}^2} \mathbf{f} - j\Delta\omega \overset{(21)}{\mathbf{f}^H \mathbf{D}_{n:n-M} \mathbf{H}'\mathbf{H}'^H} \mathbf{f} \\
&\quad + \Delta\omega^2 \overset{(22)}{\mathbf{f}^H \mathbf{D}_{n:n-M} \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}} \mathbf{f} + j\frac{1}{2}\Delta\omega^3 \overset{(23)}{\mathbf{f}^H \mathbf{D}_{n:n-M} \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}^2} \mathbf{f} \\
&\quad - \frac{1}{2}\Delta\omega^2 \overset{(31)}{\mathbf{f}^H \mathbf{D}_{n:n-M}^2 \mathbf{H}'\mathbf{H}'^H} \mathbf{f} - j\frac{1}{2}\Delta\omega^3 \overset{(32)}{\mathbf{f}^H \mathbf{D}_{n:n-M}^2 \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}} \mathbf{f} \\
&\quad \left. + \frac{1}{4}\Delta\omega^4 \overset{(33)}{\mathbf{f}^H \mathbf{D}_{n:n-M}^2 \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}^2} \mathbf{f} \right]. \tag{51}
\end{aligned}$$

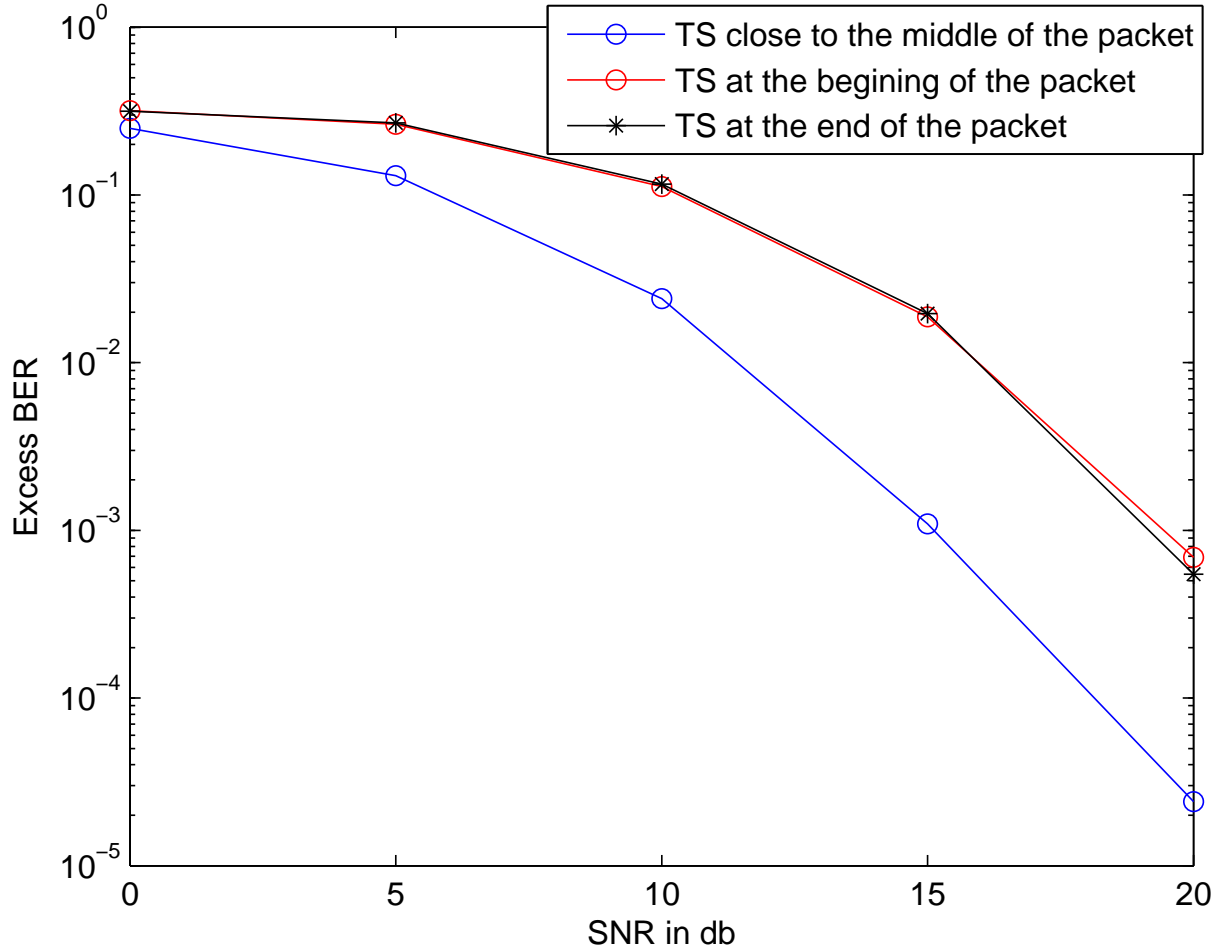


Fig. 4. Experimentally computed excess BERs for different TS positions.

Using that  $\mathcal{E}_{\Delta\omega}[\Delta\omega] = 0$ ,  $\mathcal{E}_{\Delta\omega}[\Delta\omega^3] = 0$  (the ML estimator is practically unbiased and Gaussian) and that the term proportional to  $\Delta\omega^4$  is practically equal to zero, we obtain

$$\begin{aligned} \mathcal{E}_{\Delta\omega}[\Delta\omega^2] &= \mathbf{f}^H \mathbf{R}_z \mathbf{f} + \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_{n:n-M} \mathbf{H}' \mathbf{H}'^H \mathbf{D}_{n:n-M} \mathbf{f} \\ &\quad - \frac{1}{2} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_{n:n-M}^2 \mathbf{f} - \frac{1}{2} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_{n:n-M}^2 \mathbf{H}' \mathbf{H}'^H \mathbf{f}. \end{aligned} \quad (52)$$

We write the last three terms of (52) using that  $\mathbf{D}_{n:n-M} = n\mathbf{I}_{M+1} - \mathbf{D}_M$  as

1)

$$\begin{aligned} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_{n:n-M} \mathbf{H}' \mathbf{H}'^H \mathbf{D}_{n:n-M} \mathbf{f} &= n^2 \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{f} + \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_M \mathbf{H}' \mathbf{H}'^H \mathbf{D}_M \mathbf{f} \\ &\quad - n \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_M \mathbf{H}' \mathbf{H}'^H \mathbf{f} - n \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_M \mathbf{f}. \end{aligned} \quad (53)$$

2)

$$\begin{aligned} -\frac{1}{2} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_{n:n-M}^2 \mathbf{f} &= -\frac{1}{2} n^2 \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{f} + n \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_M \mathbf{f} \\ &\quad - \frac{1}{2} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_M^2 \mathbf{f}. \end{aligned} \quad (54)$$

3)

$$\begin{aligned} -\frac{1}{2} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_{n:n-M}^2 \mathbf{H}' \mathbf{H}'^H \mathbf{f} &= -\frac{1}{2} n^2 \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{f} + n \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_M \mathbf{H}' \mathbf{H}'^H \mathbf{f} \\ &\quad - \frac{1}{2} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_M^2 \mathbf{H}' \mathbf{H}'^H \mathbf{f}. \end{aligned} \quad (55)$$

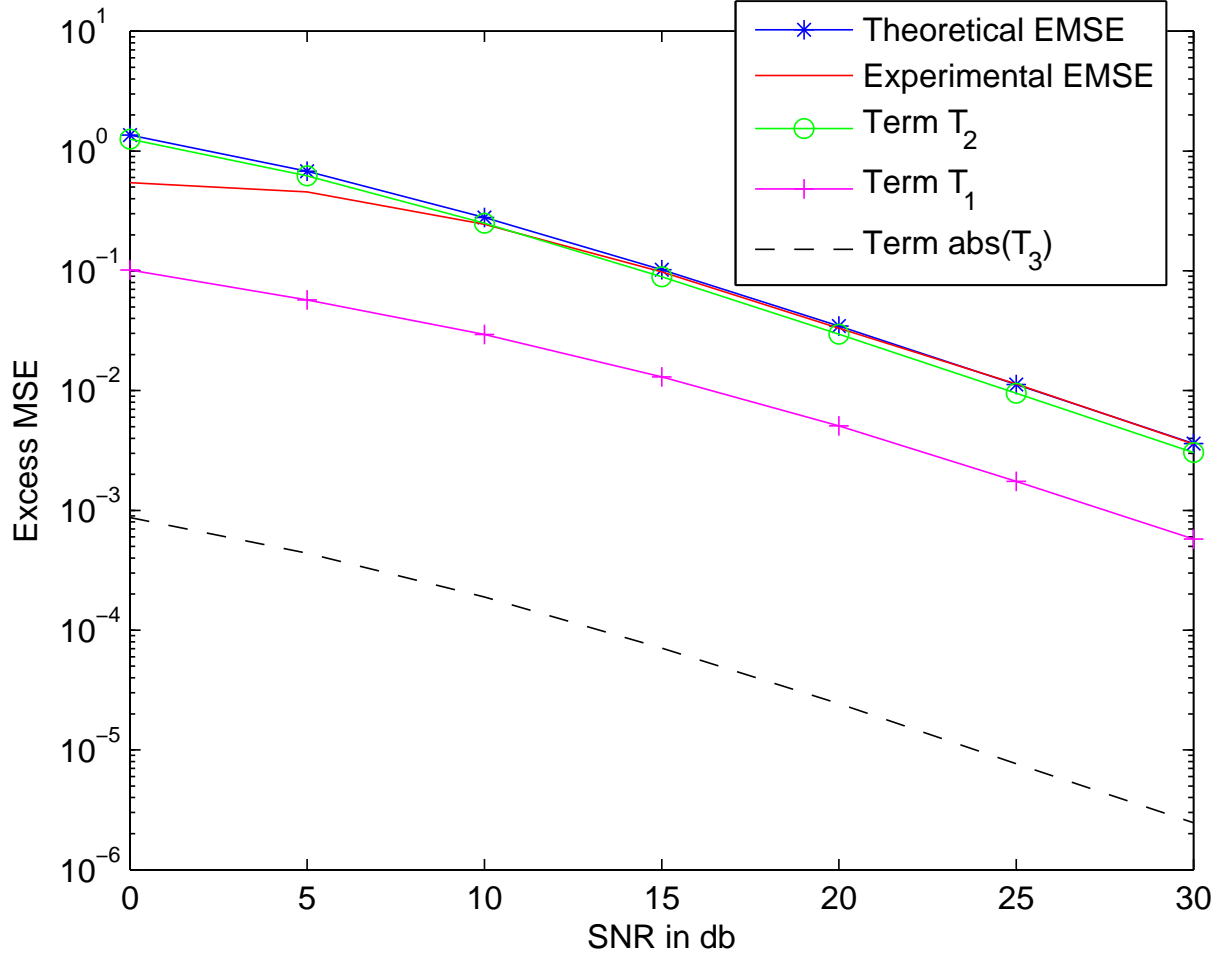


Fig. 5. Experimentally computed EMSE, EMSE theoretical approximation in (29) and the three terms  $\mathbf{T}_1$ ,  $\mathbf{T}_2$  and  $\mathbf{T}_3$  averaged over random channel realizations.

Using (53)-(55) in (52) we obtain

$$\begin{aligned}
 \mathcal{E}_{\Delta\mathbf{h}', \Delta\omega}[t_1] &= \mathbf{f}^H \mathbf{R}_z \mathbf{f} + \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_M \mathbf{H}' \mathbf{H}'^H \mathbf{D}_M \mathbf{f} \\
 &\quad - \frac{1}{2} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_M^2 \mathbf{f} - \frac{1}{2} \sigma_{\Delta\omega}^2 \mathbf{f}^H \mathbf{D}_M^2 \mathbf{H}' \mathbf{H}'^H \mathbf{f} \\
 &= \mathbf{f}^H \mathbf{R}_z \mathbf{f} + \mathcal{O}\left(\frac{M^2 \sigma_w^2}{R^3}\right).
 \end{aligned} \tag{56}$$

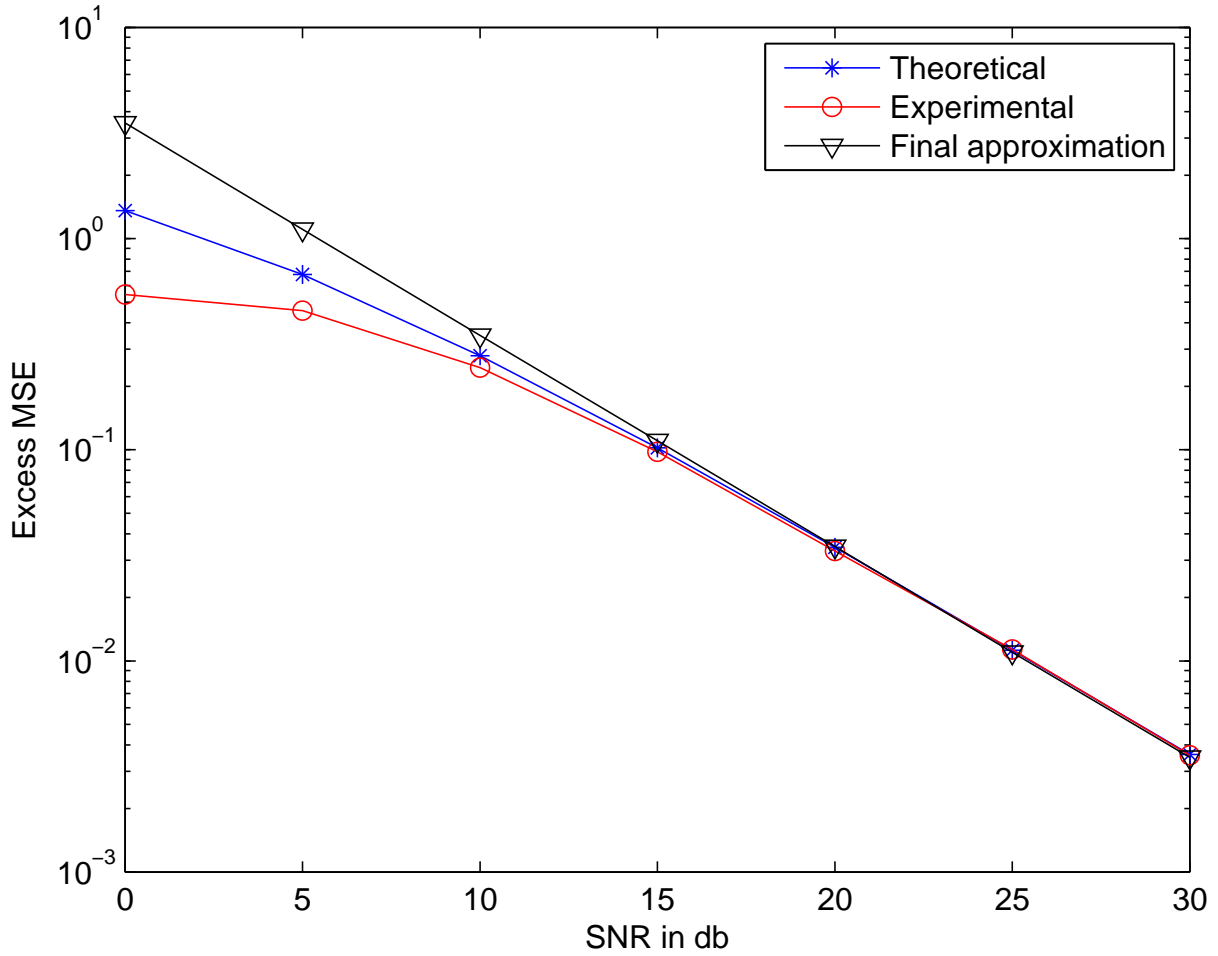


Fig. 6. Final EMSE theoretical expression in (46) averaged over random channels.

2) *Term  $t_2$* : Using (49) we obtain

$$\begin{aligned}
\mathcal{E}_{\Delta \mathbf{h}', \Delta \omega}[t_2] &= \mathcal{E}_{\Delta \mathbf{h}', \Delta \omega} \left[ \Delta \mathbf{f}^H \left( \overset{(1)}{\mathbf{I}_{M+1}} - j \Delta \omega \overset{(2)}{\mathbf{D}_{n:n-M}} - \frac{1}{2} \Delta \omega^2 \overset{(3)}{\mathbf{D}_{n:n-M}^2} \right) \mathbf{H}' \mathbf{H}'^H \right. \\
&\quad \left. \times \left( \overset{(1)}{\mathbf{I}_{M+1}} + j \Delta \omega \overset{(2)}{\mathbf{D}_{n:n-M}} - \frac{1}{2} \Delta \omega^2 \overset{(3)}{\mathbf{D}_{n:n-M}^2} + \sigma_w^2 \mathbf{I}_{M+1} \right) \Delta \mathbf{f} \right] \\
&= \mathcal{E}_{\Delta \mathbf{h}', \Delta \omega} \left[ \Delta \mathbf{f}^H \left( \overset{(11)}{\mathbf{H}' \mathbf{H}'^H} + \sigma_w^2 \mathbf{I}_{M+1} \right) \Delta \mathbf{f} + j \Delta \omega \Delta \mathbf{f}^H \overset{(12)}{\mathbf{H}' \mathbf{H}'^H} \overset{(13)}{\mathbf{D}_{n:n-M}} \Delta \mathbf{f} \right. \\
&\quad \left. - \frac{1}{2} \Delta \omega^2 \Delta \mathbf{f}^H \overset{(13)}{\mathbf{H}' \mathbf{H}'^H} \overset{(21)}{\mathbf{D}_{n:n-M}^2} \Delta \mathbf{f} - j \Delta \omega \Delta \mathbf{f}^H \overset{(21)}{\mathbf{D}_{n:n-M}} \overset{(22)}{\mathbf{H}' \mathbf{H}'^H} \Delta \mathbf{f} \right. \\
&\quad \left. + \Delta \omega^2 \Delta \mathbf{f}^H \overset{(22)}{\mathbf{D}_{n:n-M}} \overset{(23)}{\mathbf{H}' \mathbf{H}'^H} \overset{(23)}{\mathbf{D}_{n:n-M}} \Delta \mathbf{f} + j \frac{1}{2} \Delta \omega^3 \Delta \mathbf{f}^H \overset{(23)}{\mathbf{D}_{n:n-M}} \overset{(23)}{\mathbf{H}' \mathbf{H}'^H} \overset{(23)}{\mathbf{D}_{n:n-M}^2} \Delta \mathbf{f} \right. \\
&\quad \left. - \frac{1}{2} \Delta \omega^2 \Delta \mathbf{f}^H \overset{(31)}{\mathbf{D}_{n:n-M}^2} \overset{(31)}{\mathbf{H}' \mathbf{H}'^H} \Delta \mathbf{f} - j \frac{1}{2} \Delta \omega^3 \Delta \mathbf{f}^H \overset{(32)}{\mathbf{D}_{n:n-M}^2} \overset{(32)}{\mathbf{H}' \mathbf{H}'^H} \overset{(32)}{\mathbf{D}_{n:n-M}} \Delta \mathbf{f} \right. \\
&\quad \left. + \frac{1}{4} \Delta \omega^4 \Delta \mathbf{f}^H \overset{(33)}{\mathbf{D}_{n:n-M}^2} \overset{(33)}{\mathbf{H}' \mathbf{H}'^H} \overset{(33)}{\mathbf{D}_{n:n-M}^2} \Delta \mathbf{f} \right]. \tag{57}
\end{aligned}$$

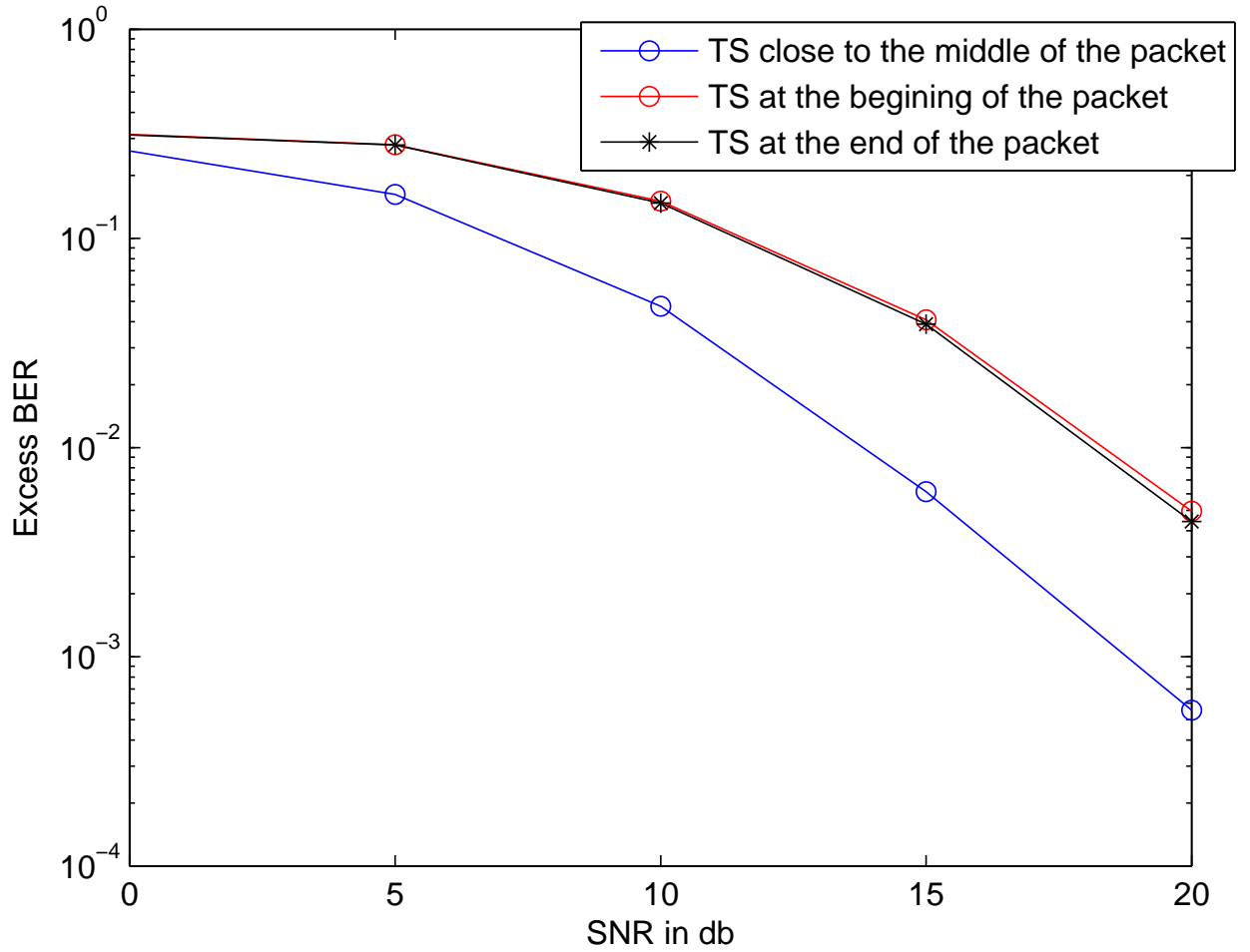


Fig. 7. Experimentally computed channel average excess BERs for different TS positions.

We keep second and third-order error terms (*we keep third-order terms just to see the lower-order term we neglect for  $t_2$* ). Then

$$\begin{aligned} \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[t_2] = \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ \Delta\mathbf{f}^H \mathbf{R}_z \Delta\mathbf{f} + j\Delta\omega \Delta\mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_{n:n-M} \Delta\mathbf{f} \right. \\ \left. - j\Delta\omega \Delta\mathbf{f}^H \mathbf{D}_{n:n-M} \mathbf{H}' \mathbf{H}'^H \Delta\mathbf{f} \right]. \end{aligned} \quad (58)$$

We write the last two terms of (58) using that  $\mathbf{D}_{n:n-M} = n\mathbf{I}_{M+1} - \mathbf{D}_M$  as

1)

$$\begin{aligned} \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ j\Delta\omega \Delta\mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_{n:n-M} \Delta\mathbf{f} \right] = \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ j n \Delta\omega \Delta\mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \Delta\mathbf{f} \right] \\ - \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ j\Delta\omega \Delta\mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \mathbf{D}_M \Delta\mathbf{f} \right]. \end{aligned} \quad (59)$$

2)

$$\begin{aligned} \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ -j\Delta\omega \Delta\mathbf{f}^H \mathbf{D}_{n:n-M} \mathbf{H}' \mathbf{H}'^H \Delta\mathbf{f} \right] = \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ -j n \Delta\omega \Delta\mathbf{f}^H \mathbf{H}' \mathbf{H}'^H \Delta\mathbf{f} \right] \\ + \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega} \left[ j\Delta\omega \Delta\mathbf{f}^H \mathbf{D}_M \mathbf{H}' \mathbf{H}'^H \Delta\mathbf{f} \right]. \end{aligned} \quad (60)$$

Using (59)-(60) in (58) we obtain

$$\begin{aligned}
\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[t_2] &= \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[\Delta\mathbf{f}^H \mathbf{R}_z \Delta\mathbf{f}] + \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[-j\Delta\omega\Delta\mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_M \Delta\mathbf{f}] \\
&\quad + \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[j\Delta\omega\Delta\mathbf{f}^H \mathbf{D}_M \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f}] \\
&= \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[\Delta\mathbf{f}^H \mathbf{R}_z \Delta\mathbf{f}] + 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[j\Delta\omega\Delta\mathbf{f}^H \mathbf{D}_M \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f}]\right\} \\
&= \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[\Delta\mathbf{f}^H \mathbf{R}_z \Delta\mathbf{f}] + \mathcal{O}\left(\frac{M\sigma_w^3}{R^{5/2}}\right).
\end{aligned} \tag{61}$$

3) *Term  $t_3$* : Using (49) we obtain

$$\begin{aligned}
\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[t_3] &= 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\mathbf{f}^H \left(\mathbf{I}_{M+1}^{(1)} - j\Delta\omega \mathbf{D}_{n:n-M}^{(2)} - \frac{1}{2}\Delta\omega^2 \mathbf{D}_{n:n-M}^{(3)}\right) \mathbf{H}'\mathbf{H}'^H \right. \right. \\
&\quad \left. \left. \times \left(\mathbf{I}_{M+1}^{(1)} + j\Delta\omega \mathbf{D}_{n:n-M}^{(2)} - \frac{1}{2}\Delta\omega^2 \mathbf{D}_{n:n-M}^{(3)}\right) + \sigma_w^2 \mathbf{I}_{M+1}\right) \Delta\mathbf{f}\right\} \\
&= 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\mathbf{f}^H \left(\mathbf{H}'\mathbf{H}'^H + \sigma_w^2 \mathbf{I}_{M+1}\right) \Delta\mathbf{f} + j\Delta\omega \mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}^{(12)} \Delta\mathbf{f} \right. \right. \\
&\quad \left. \left. - \frac{1}{2}\Delta\omega^2 \mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}^{(13)} \Delta\mathbf{f} - j\Delta\omega \mathbf{f}^H \mathbf{D}_{n:n-M}^{(21)} \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f} \right. \right. \\
&\quad \left. \left. + \Delta\omega^2 \mathbf{f}^H \mathbf{D}_{n:n-M}^{(22)} \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M} \Delta\mathbf{f} + j\frac{1}{2}\Delta\omega^3 \mathbf{f}^H \mathbf{D}_{n:n-M}^{(23)} \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}^2 \Delta\mathbf{f} \right. \right. \\
&\quad \left. \left. - \frac{1}{2}\Delta\omega^2 \mathbf{f}^H \mathbf{D}_{n:n-M}^{(31)} \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f} - j\frac{1}{2}\Delta\omega^3 \mathbf{f}^H \mathbf{D}_{n:n-M}^{(32)} \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M} \Delta\mathbf{f} \right. \right. \\
&\quad \left. \left. + \frac{1}{4}\Delta\omega^4 \mathbf{f}^H \mathbf{D}_{n:n-M}^{(33)} \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M}^2 \Delta\mathbf{f}\right\}.
\end{aligned} \tag{62}$$

We keep first and second-order error terms (we keep second-order terms just to see the lower-order term we neglect for  $t_3$ ). We will write the second-order terms of (62) in detail using that  $\mathbf{D}_{n:n-M} = n\mathbf{I}_{M+1} - \mathbf{D}_M$

1)

$$\begin{aligned}
2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega \mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_{n:n-M} \Delta\mathbf{f}\right]\right\} &= 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[jn\Delta\omega \mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f}\right]\right\} \\
&\quad - 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega \mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_M \Delta\mathbf{f}\right]\right\}.
\end{aligned} \tag{63}$$

2)

$$\begin{aligned}
2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[-j\Delta\omega \mathbf{f}^H \mathbf{D}_{n:n-M} \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f}\right]\right\} &= -2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[jn\Delta\omega \mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f}\right]\right\} \\
&\quad + 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega \mathbf{f}^H \mathbf{D}_M \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f}\right]\right\}.
\end{aligned} \tag{64}$$

Thus, using (63)-(64) in (62), and by ignoring terms of higher order, we obtain

$$\begin{aligned}
\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[t_3] &= 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\mathbf{f}^H \mathbf{R}_z \Delta\mathbf{f}\right]\right\} - 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega \mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_M \Delta\mathbf{f}\right]\right\} \\
&\quad + 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega \mathbf{f}^H \mathbf{D}_M \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f}\right]\right\} \\
&= 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\mathbf{e}_d^H \mathbf{H}'^H \Delta\mathbf{f}\right]\right\} - 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega \mathbf{f}^H \mathbf{H}'\mathbf{H}'^H \mathbf{D}_M \Delta\mathbf{f}\right]\right\} \\
&\quad + 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega \mathbf{f}^H \mathbf{D}_M \mathbf{H}'\mathbf{H}'^H \Delta\mathbf{f}\right]\right\} \\
&= 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\mathbf{e}_d^H \mathbf{H}'^H \Delta\mathbf{f}\right]\right\} + \mathcal{O}\left(\frac{M\sigma_w^2}{R^3}\right)
\end{aligned} \tag{65}$$

where we have used that  $\mathbf{f} = \mathbf{R}_z^{-1} \mathbf{H}' \mathbf{e}_d$ .

4) *Term  $t_4$* : Using (50) we obtain

$$\begin{aligned}
\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[t_4] &\approx -2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[e^{j\Delta\omega\xi}\mathbf{f}^H\mathbf{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{e}_d\right]\right\} \\
&= -2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\mathbf{f}^H\mathbf{\Gamma}'_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{e}_d\right]\right\} \\
&= -2 \operatorname{Re}\left\{\mathbf{f}^H\mathbf{H}'\mathbf{e}_d\right\} + 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega\mathbf{f}^H\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\right]\right\} \\
&\quad + \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\omega^2\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\right]\right\}.
\end{aligned} \tag{66}$$

Using that  $\mathcal{E}_{\Delta\omega}[\Delta\omega] = 0$ , we get

$$\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[t_4] \approx -2 \operatorname{Re}\left\{\mathbf{f}^H\mathbf{H}'\mathbf{e}_d\right\} + \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\omega^2\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\right]\right\}. \tag{67}$$

5) *Term  $t_5$* : Using (50) we obtain

$$\begin{aligned}
\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}[t_5] &\approx -2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[e^{j\Delta\omega\xi}\Delta\mathbf{f}^H\mathbf{\Gamma}_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{e}_d\right]\right\} \\
&= -2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\mathbf{f}^H\mathbf{\Gamma}'_{n:n-M}(-\Delta\omega)\mathbf{H}'\mathbf{e}_d\right]\right\} \\
&= -2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\mathbf{f}^H\mathbf{H}'\mathbf{e}_d\right]\right\} + 2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[j\Delta\omega\Delta\mathbf{f}^H\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\right]\right\} \\
&\quad + \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\omega^2\Delta\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\right]\right\} \\
&= -2 \operatorname{Re}\left\{\mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\mathbf{f}^H\mathbf{H}'\mathbf{e}_d\right]\right\} + \mathcal{O}\left(\frac{n^2\sigma_w^3}{R^{7/2}}\right).
\end{aligned} \tag{68}$$

From the definition of the EMSE and using (56), (61), (65), (67) and (68) we get

$$\begin{aligned}
\operatorname{EMSE}_n(\hat{\mathbf{f}}, \hat{\omega}) &= \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\mathbf{f}^H\mathbf{R}_z\Delta\mathbf{f} + \Delta\omega^2\operatorname{Re}\{\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\}\right. \\
&\quad \left.+ 2\operatorname{Re}\{j\Delta\omega\Delta\mathbf{f}^H\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\}\right] + \mathcal{O}\left(\frac{n^2\sigma_w^3}{R^{7/2}}\right) + \mathcal{O}\left(\frac{M^2\sigma_w^2}{R^3}\right).
\end{aligned} \tag{69}$$

In the sequel, we ignore the  $\mathcal{O}_p(\cdot)$  terms. Then

$$\begin{aligned}
\operatorname{EMSE}_n(\hat{\mathbf{f}}, \hat{\omega}) &= \mathcal{E}_{\Delta\mathbf{h}',\Delta\omega}\left[\Delta\mathbf{f}^H\mathbf{R}_z\Delta\mathbf{f} + \Delta\omega^2\operatorname{Re}\{\mathbf{f}^H\mathbf{D}'^2_{n:n-M}\mathbf{H}'\mathbf{e}_d\}\right. \\
&\quad \left.+ 2\operatorname{Re}\{j\Delta\omega\Delta\mathbf{f}^H\mathbf{D}'_{n:n-M}\mathbf{H}'\mathbf{e}_d\}\right].
\end{aligned} \tag{70}$$

The three terms of (70) are computed as follows

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

Proposition 1 is proved.  $\square$

### B. Components of $\mathbf{\Psi}'$

If the covariance matrix of the training sequence is  $\mathcal{R}$ , then [8, Appendix A]

$$\frac{1}{R}\mathbf{A}^H\mathbf{A} = \mathcal{R} + \mathcal{O}\left(\frac{1}{R}\right) \tag{71}$$

and

$$\frac{1}{R^2} \mathbf{A}^H \mathbf{D}_{R-1} \mathbf{A} = \mathcal{R} + \mathcal{O}\left(\frac{1}{R}\right). \quad (72)$$

If  $\mathcal{R}$  is invertible, then, using the first-order expansion

$$(\mathbf{A} + \Delta \mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \Delta \mathbf{A} \mathbf{A}^{-1} + \mathcal{O}(\|\Delta \mathbf{A}\|^2). \quad (73)$$

we obtain

$$(\mathbf{A}^H \mathbf{A})^{-1} = \frac{1}{R} \mathcal{R}^{-1} + \mathcal{O}\left(\frac{1}{R^2}\right) \mathcal{R}^{-2}. \quad (74)$$

Furthermore,

$$\begin{aligned} \mathbf{A}^H \mathbf{K}' \mathbf{A} &= \mathbf{A}^H \left( -\mathbf{D}_{R-1} + \left(\frac{R}{2} - 1\right) \mathbf{I}_R \right) \mathbf{A} \\ &= -\mathbf{A}^H \mathbf{D}_{R-1} \mathbf{A} + \left(\frac{R}{2} - 1\right) \mathbf{A}^H \mathbf{A} \\ &= -\frac{R^2}{2} \mathcal{R} - \mathcal{O}(R) + \left(\frac{R}{2} - 1\right) (R \mathcal{R} + \mathcal{O}(1)) \\ &= \mathcal{O}(R). \end{aligned} \quad (75)$$

Then

$$\begin{aligned} \mathbf{A}^H \mathbf{K}' \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} &= \mathcal{O}(R) \left( \frac{1}{R} \mathcal{R}^{-1} + \mathcal{O}\left(\frac{1}{R^2}\right) \mathcal{R}^{-2} \right) \\ &= \mathcal{O}(1) \mathcal{R}^{-1} + \mathcal{O}\left(\frac{1}{R}\right) \mathcal{R}^{-2} \\ &= \mathcal{O}(1). \end{aligned} \quad (76)$$

Using (76) and (90) it is easy to see that the second term of  $\Psi'$  is  $\mathcal{O}\left(\frac{\sigma_w^2}{R^3}\right)$ , while, using (74) we obtain that the first term of  $\Psi'$  is  $\mathcal{O}\left(\frac{\sigma_w^2}{R}\right)$ . Thus, if the covariance of the training sequence is invertible (and not very ill-conditioned), then the first term of  $\Psi'$  is much larger than the second.

### C. Proof of Proposition 2

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

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

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

$$\begin{aligned} \mathbf{H}'^H \left( \mathbf{H}' \mathbf{H}'^H + \sigma_w^2 \mathbf{I} \right)^{-1} \mathbf{H}' &\leq \mathbf{H}'^H \left( \mathbf{H}' \mathbf{H}'^H \right)^{-1} \mathbf{H}' \\ &= \mathbf{P}_{\mathbf{H}'^H} \leq \mathbf{I}_{M+L+1}. \end{aligned} \quad (77)$$

Using (77) and  $\text{tr}(\mathbf{A} \mathbf{B} \mathbf{A}^H) \leq \lambda_{\max}(\mathbf{B}) \text{tr}(\mathbf{A} \mathbf{A}^H)$  [7, p. 44], we obtain (recall the definition of  $\mathbf{R}_z$  in (19))

$$\begin{aligned} \mathbf{T}_1 &\approx \text{tr}(\Psi'^* \mathbf{F}^* \mathbf{H}'^H \mathbf{R}_z^{-1} \mathbf{H}' \mathbf{F}^T) \\ &\lesssim \text{tr}(\Psi'^* \mathbf{F}^* \mathbf{F}^T) \leq \lambda_{\max}(\Psi') \text{tr}(\mathbf{F}^* \mathbf{F}^T) \\ &= \lambda_{\max}(\Psi') \|\mathbf{F}\|_F^2 = \lambda_{\max}(\Psi') (L+1) \|\mathbf{f}\|_2^2. \end{aligned} \quad (78)$$

Using asymptotic arguments, we proved in the Appendix B that, if  $\mathbf{A}^H \mathbf{A}$  is invertible (and not very ill-conditioned), then the first term of  $\Psi'$  is much larger than the second. Thus,  $\lambda_{\max}(\Psi') \approx \sigma_w^2 / \lambda_{\min}(\mathbf{A}^H \mathbf{A})$  and

$$\mathbf{T}_1 \lesssim \frac{(L+1) \|\mathbf{f}\|_2^2 \sigma_w^2}{\lambda_{\min}(\mathbf{A}^H \mathbf{A})}. \quad (79)$$

2) *Term  $\mathbf{T}_2$* : Term  $\text{Re}\{\mathbf{f}^H \mathbf{H}' \mathbf{e}_d\}$  is the  $(d+1)$ -st coefficient of the combined (channel-equalizer) impulse response. Using the definition of  $\mathbf{f}$  in (19) and expression (77), it can be shown that  $\text{Re}\{\mathbf{f}^H \mathbf{H}' \mathbf{e}_d\}$  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  $\mathbf{T}_{23} = 2\text{Re}\{\mathbf{f}^H \mathbf{D}_M^2 \mathbf{H}' \mathbf{e}_d\} \leq 2M^2 k_2(\mathbf{H}')$ . 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

$$\mathbf{T}_2 \simeq \mathcal{C} \sigma_{\Delta\omega}^2. \quad (80)$$



3) *Term  $\mathbf{T}_3$* : In order to simplify term  $\mathbf{T}_3$  we first prove the following two lemmas.

**Lemma 1:** *Using the SVD of  $\mathbf{A}$ , it can be shown that*

$$\|\mathbf{A}^T \mathbf{K}' \mathbf{A}^* (\mathbf{A}^H \mathbf{A})^{-T}\|_2 \leq \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})} \|\mathbf{K}'\|_2 = \frac{R}{2} k_2(\mathbf{A}). \quad (81)$$

**Proof:** Using the singular value decomposition (SVD), matrix  $\mathbf{A}$  can be written as

$$\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H \quad (82)$$

where  $\mathbf{U}$  and  $\mathbf{V}$  are unitary matrices with dimensions  $(N_{\text{tr}} - L) \times (N_{\text{tr}} - L)$  and  $(L + 1) \times (L + 1)$  respectively, while  $\mathbf{\Sigma}$  is the  $(N_{\text{tr}} - L) \times (L + 1)$  matrix with the singular values of  $\mathbf{A}$  in its diagonal, and all the off the diagonal elements equal to zero. Using (82) we obtain

$$(\mathbf{A}^H \mathbf{A})^{-T} = \mathbf{V}^* \mathbf{\Sigma}_1 \mathbf{V}^T \quad (83)$$

where  $\mathbf{\Sigma}_1 = (\mathbf{\Sigma}^H \mathbf{\Sigma})^{-1} = \text{diag}(\sigma_1^{-2}(\mathbf{A}), \dots, \sigma_{L+1}^{-2}(\mathbf{A}))$ . Thus, in order to prove (81) we use the submultiplicative property of the matrix norms and expressions (82) and (83) to get

$$\begin{aligned} \|\mathbf{A}^T \mathbf{K}' \mathbf{A}^* (\mathbf{A}^H \mathbf{A})^{-T}\|_2 &= \|\mathbf{V}^* \mathbf{\Sigma}^T \mathbf{U}^T \mathbf{K}' \mathbf{U}^* \mathbf{\Sigma}^* \mathbf{\Sigma}_1 \mathbf{V}^T\|_2 \\ &\leq \|\mathbf{V}^*\|_2 \|\mathbf{\Sigma}^T\|_2 \|\mathbf{U}^T\|_2 \|\mathbf{K}'\|_2 \|\mathbf{U}^*\|_2 \|\mathbf{\Sigma}^*\|_2 \|\mathbf{\Sigma}_1\|_2 \|\mathbf{V}^T\|_2 \\ &= \sigma_{\max}(\mathbf{A}) \|\mathbf{K}'\|_2 \frac{1}{\sigma_{\min}(\mathbf{A})} = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})} \|\mathbf{K}'\|_2 = \frac{R}{2} k_2(\mathbf{A}) \end{aligned} \quad (84)$$

where for the last equality we used the definition of matrix  $\mathbf{K}'$  and the condition number with respect to the spectral norm.  $\square$

**Lemma 2:** *Using the SVD of  $\mathbf{H}'$ , it can be shown that*

$$\|\mathbf{G}^H \mathbf{R}_z^{-1} \mathbf{D}_M \mathbf{H}' \mathbf{e}_d\|_2 \leq \|\mathbf{F}\|_2 \frac{\sigma_{\max}(\mathbf{H}')}{\sigma_{\min}(\mathbf{H}')} M = M \|\mathbf{F}\|_2 k_2(\mathbf{H}'). \quad (85)$$

**Proof:** We first write the SVD of matrix  $\mathbf{H}'$  as

$$\mathbf{H}' = \mathbf{U}_1 \mathbf{\Lambda} \mathbf{V}_1^H \quad (86)$$

where  $\mathbf{U}_1$  and  $\mathbf{V}_1$  are unitary matrices with dimensions  $(M + 1) \times (M + 1)$  and  $(M + L + 1) \times (M + L + 1)$  respectively, while  $\mathbf{\Lambda}$  is the  $(M + 1) \times (M + L + 1)$  matrix with the singular values of  $\mathbf{H}'$  in its diagonal, and all the off the diagonal elements equal to zero. Using (86) and the definition of  $\mathbf{R}_z$  we obtain

$$\mathbf{R}_z^{-1} = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^H \quad (87)$$

where  $\mathbf{\Lambda}_1 = \text{diag}\left(\frac{1}{\sigma_1^2(\mathbf{H}') + \sigma_w^2}, \dots, \frac{1}{\sigma_{M+1}^2(\mathbf{H}') + \sigma_w^2}\right)$ .

Using (86), (87) and that  $\mathbf{G} = \mathbf{H}' \mathbf{F}^T$  we obtain

$$\begin{aligned} \|\mathbf{G}^H \mathbf{R}_z^{-1} \mathbf{D}_M \mathbf{H}' \mathbf{e}_d\|_2 &= \|\mathbf{F}^* \mathbf{V}_1 \mathbf{\Lambda}^H \mathbf{\Lambda}_1 \mathbf{U}_1^H \mathbf{D}_M \mathbf{U}_1 \mathbf{\Lambda} \mathbf{V}_1^H \mathbf{e}_d\|_2 \\ &\stackrel{(*)}{\leq} \|\mathbf{F}\|_2 \frac{1}{\sigma_{\min}(\mathbf{H}')} M \sigma_{\max}(\mathbf{H}') \\ &= \|\mathbf{F}\|_2 \frac{\sigma_{\max}(\mathbf{H}')}{\sigma_{\min}(\mathbf{H}')} M = M \|\mathbf{F}\|_2 k_2(\mathbf{H}') \end{aligned} \quad (88)$$

where at point (\*) we have used that  $\mathbf{\Lambda}^H \mathbf{\Lambda}_1 = \text{diag}\left(\frac{\sigma_1(\mathbf{H}')}{\sigma_1^2(\mathbf{H}') + \sigma_w^2}, \dots, \frac{\sigma_{M+1}(\mathbf{H}')}{\sigma_{M+1}^2(\mathbf{H}') + \sigma_w^2}\right) \leq \text{diag}\left(\frac{\sigma_1(\mathbf{H}')}{\sigma_1^2(\mathbf{H}')}, \dots, \frac{\sigma_{M+1}(\mathbf{H}')}{\sigma_{M+1}^2(\mathbf{H}')}\right) = \text{diag}\left(\frac{1}{\sigma_1(\mathbf{H}')}, \dots, \frac{1}{\sigma_{M+1}(\mathbf{H}')}\right)$ . Thus,  $\|\mathbf{\Lambda}^H \mathbf{\Lambda}_1\|_2 \leq \frac{1}{\sigma_{\min}(\mathbf{H}')} \square$

Thus, using (81) and (85)

$$\mathbf{T}_3 \leq (MR \|\mathbf{F}\|_2 k_2(\mathbf{A}) k_2(\mathbf{H}')) \sigma_{\Delta\omega}^2. \quad (89)$$

4) *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 (80) (recall that  $\mathcal{C} = O(N^2)$ ) and (89), we conclude that  $\mathbf{T}_2 \gg \mathbf{T}_3$ .

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

$$\sigma_{\Delta\omega}^2 \approx \frac{6\sigma_w^2}{R^2 \mathbf{h}^H \mathbf{A}^H \mathbf{A} \mathbf{h}}. \quad (90)$$

Thus

$$\mathbf{T}_2 \approx \frac{6\mathcal{C}\sigma_w^2}{R^2 \mathbf{h}^H \mathbf{A}^H \mathbf{A} \mathbf{h}}. \quad (91)$$

Using (79) and (91), we derive the following approximate bound

$$\begin{aligned} \frac{\mathbf{T}_1}{\mathbf{T}_2} &\lesssim \frac{(L+1)R^2 \|\mathbf{f}\|_2^2 \mathbf{h}^H \mathbf{A}^H \mathbf{A} \mathbf{h}}{6\mathcal{C} \lambda_{\min}(\mathbf{A}^H \mathbf{A})} \\ &\leq \frac{(L+1)R^2 \|\mathbf{f}\|_2^2 \lambda_{\max}(\mathbf{A}^H \mathbf{A}) \|\mathbf{h}\|_2^2}{6\mathcal{C} \lambda_{\min}(\mathbf{A}^H \mathbf{A})} \\ &= k_2(\mathbf{A}^H \mathbf{A})(L+1) \|\mathbf{f}\|_2^2 \|\mathbf{h}\|_2^2 \alpha \end{aligned} \quad (92)$$

where

$$\alpha \triangleq \frac{R^2}{6\mathcal{C}} = \mathcal{O}\left(\frac{R^2}{N^2}\right). \quad (93)$$

Thus, if  $\alpha$  is sufficiently small, i.e.,  $R$  is sufficiently small with respect to  $N$  (recall that  $R = N_{\text{tr}} - L$ ), 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 (80).

#### D. Value of $n_1$

We remind the definition of  $\xi$

$$\xi \triangleq n_1 + \frac{N_{\text{tr}} + L}{2}.$$

We found that

$$\xi = \mathcal{C}_2.$$

We remind

$$1 + 2 + \dots + n = \frac{n(n+1)}{2} \quad (94)$$

$$\begin{aligned} n_1 + \dots + n_2 &= (1 + \dots + n_2) - (1 + \dots + (n_1 - 1)) \\ &= \frac{n_2(n_2 + 1)}{2} - \frac{(n_1 - 1)n}{2} \\ &= \frac{n_2^2 - n_1^2 + n_2 + n_1}{2} \\ &= \frac{(n_2 + n_1)(n_2 - n_1 + 1)}{2} \\ &= \frac{(n_2 + n_1)N_{\text{tr}}}{2}. \end{aligned} \quad (95)$$

We compute  $\mathcal{C}_2$  as follows

$$\mathcal{C}_2 = \frac{1}{|\mathcal{D}|} \sum_{n \in \mathcal{D}} n = \frac{1}{N - N_{\text{tr}}} \mathcal{X} \quad (96)$$

where

$$\begin{aligned} \mathcal{X} &= \left( (d+1) + \dots + (d+n_1-1) \right. \\ &\quad \left. + (d+n_2+1) + \dots + (d+N) \right) \\ &= \left( (1 + \dots + (d+N)) - (1 + \dots + d) \right. \\ &\quad \left. - [(d+n_1) + \dots + (d+n_2)] \right) \\ &= \frac{(N+d)(N+d+1)}{2} - \frac{d(d+1)}{2} \\ &\quad - N_{\text{tr}}d - \frac{(n_2+n_1)N_{\text{tr}}}{2} \\ &= \frac{N^2 + 2dN + N - 2dN_{\text{tr}}}{2} - \frac{(n_2+n_1)N_{\text{tr}}}{2}. \end{aligned} \quad (97)$$

If we solve the equation

$$n_1 + \frac{N_{\text{tr}} + L}{2} = C_2 \quad (98)$$

we obtain

$$\begin{aligned} n_1 &= \frac{1}{2N} \left( N^2 + 2dN + N - 2dN_{\text{tr}} + N_{\text{tr}} - LN + LN_{\text{tr}} - NN_{\text{tr}} \right) \\ &\approx \frac{N - N_{\text{tr}}}{2} + d - \frac{L}{2}. \end{aligned} \quad (99)$$

#### REFERENCES

- [1] M. Morelli and U. Mengali, "Carrier frequency estimation for transmissions over selective channels," *IEEE Trans. on Comm.*, 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. on Comm.*, 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 Proc.*, vol. 56, no. 8, part 1, pp. 3424-3436, Aug. 2008.
- [4] A. Sayed. *Fundamentals of adaptive filtering*. 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 Proc.*, vol. 55, no. 7, part 1, pp. 3302-3311, July 2007.
- [6] R. A Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 1985.
- [7] H. Luetkepohl. *Handbook of matrices*. John Wiley, 1996.
- [8] O. Besson and P. Stoica, "Training sequence selection for frequency offset estimation in frequency-selective channels," *Digital Signal Processing*, vol. 13, no. 1, pp. 106-127, Jan. 2003.
- [9] Y. Ying, M. Ghogho, "Optimal pilot placement for frequency offset estimation and data detection in burst transmission systems," *IEEE Communications Letters*, vol.9, no. 6, 549-551, June 2005.
- [10] Y-D Kim, J. K. Lim, C. Suh, Y.H. Lee, "Designing training sequences for carrier frequency estimation in frequency-selective channels," *IEEE Trans. Vehicular Technology*, 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. on Comm.*, 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," Technical Report (available at <http://www.telecom.tuc.gr/liavas>)