Widely Linear Receivers for SMT Systems with TX/RX Frequency-Selective I/Q Imbalance

Aamir Ishaque and Gerd Ascheid
*Institute for Communication Technologies and Embedded Systems, RWTH Aachen University, Germany
Email: {ishaque,ascheid}@ice.rwth-aachen.de

***Abstract*—Staggered multitone (SMT) modulation scheme has received a considerable attention recently due to its higher spectral efficiency compared to the conventional OFDM systems and robustness to frequency offset and Doppler spread. However, channel equalization in SMT is complicated by the so called *intrinsic interference* as a result of non-flat channel response in each subchannel. In the presence of I/Q imbalance (IQI), this task becomes even more challenging as the *image band* and its intrinsic neighborhood contaminates the received signal. This paper deals with widely linear schemes for the effective channel equalization using time-domain and frequency-domain methodologies. Further iterative improvements are proposed for the later, combining enhanced equalization with per-subchannel-pair processing after intrinsic interference cancellation. The proposed techniques are simulated for both the uncoded and coded systems suffering from frequency-selective IQI and their effectiveness with multipath channels is demonstrated.**

Keywords: SMT, fading channels, equalizer, improper signals, I/Q imbalance, interference cancellation.

I. Introduction

In the past few years, there has been a growing interest in a potential alternative to OFDM-based systems that could achieve higher data rates over the wireless mobile channel while maintaining the benefits of the conventional OFDM i.e., efficiently combat the intersymbol/interchannel interference in dispersive propagation channels. Filter-bank multichannel modulation has come up as a promising set of solutions where optimized pulse shaping filters guarantee a sharper time-frequency localization. A constituent scheme in this class is the orthogonal frequency division multiplexing with offset QAM (OFDM/OQAM) [1], also known as staggered multitone (SMT), that increases the spectral efficiency by dropping the guard period and enhances system robustness to frequency offset and Doppler spread by enforcing time-frequency confinement of the pulse shapes. However, these advantages are realized at the cost of restricted orthogonality in the real fields. This gives rise to the intrinsic interference that is basically a projection of imaginary ambiguity function to real field due to the complex channel coefficients. By inherent extension to complex orthogonality, interference cancellation using decision-feedback could offer possibility for performance enhancement [2].

On the other hand, it has been reported that SMT is sensitive to the hardware imperfections, among them I/Q imbalance (IQI) being a key issue [3]. The problem of IQI has been analyzed in many scientific contributions for the conventional OFDM signals e.g., [4], [5]. The effect of IQI on SMT is two-fold. First, each subchannel suffers image interference from the negative subchannel similar to OFDM. Second, the intrinsic interference destroys the orthogonality; not only from the desired subchannel but also from the image subchannel neighborhood. A significant effort has been spent to compensate the effects of IQI in OFDM systems. Non-frequency selective (NFS) IQI compensation has been addressed in [6] using pilot-aided estimation technique. In [7], a joint channel and IQI compensation algorithm was proposed. The algorithm assumed combined acquisition of the propagation channel and IQI profile using appropriately designed pilot symbols and pairwise processing of the received observation. Whatsoever the transmission modulation is, [8] offers a generic blind approach benefiting from the IQI induced non-circularity [9]. However, it suffers from slow convergence and mitigates the RX-side IQI only.

This paper deals with the application of the widely linear (WL) equalization [10] concept to SMT systems with TX and RX IQI, exploiting the non-circularity property of the both; the input signal and received observation. In this regard, our contributions can be summarized as follows:

- We first develop the exact channel representation for SMT systems over fading channels and with a general IQI model incorporating TX and RX mismatches.
- Assuming that the effective channel is known, an FIR filter is derived using linear minimum mean square error (LMMSE) criterion and under the hypothesis that signals inherit reasonable improperness.
- The idea is then extended to per-subchannel-pair processing model and considering the approximate orthogonality offered by SMT, it mitigates the image channel leakage due to IQI.
- To achieve better performance, an intrinsic interference cancellation framework is proposed taking into account that the channel correlation dictates the energy dispersed from the neighboring subchannels.

Notation: $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote conjugation, transpose and conjugate transpose operations respectively; $\mathbf{I}$ and $\mathbf{0}$ represents the identity matrix and the null matrices respectively; $\mathbf{A}(m,n)$ provides the (m,n)-th element of the matrix $\mathbf{A}$; $\mathrm{Tr}\{\cdot\}$, $\mathcal{E}\{\cdot\}$, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ refer to trace, expectation, real and imaginary operators respectively; $\jmath^2 = -1$; $\otimes$ represents the convolution operation; $(\cdot)_{\mathcal{C}}$ and $(\cdot)_{\mathcal{L}}$ creates a circulant and lower-triangular matrix from the argument vector, respectively.

II. System Model

Consider the SMT scheme with K subchannels and that the complex baseband signal after pulse shaping over L_g adjacent

*Part of this work has been performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in METIS, although the views expressed are those of the authors and do not necessarily represent the project.

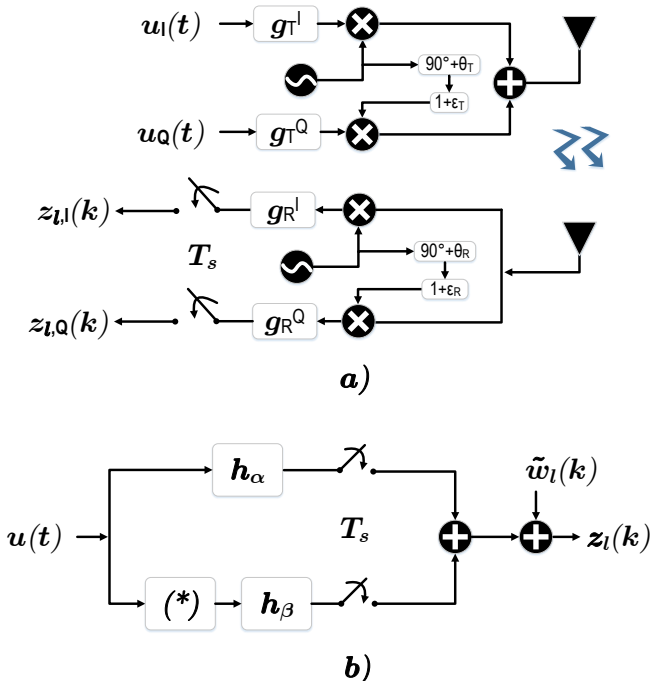

Fig. 1. a) Block diagram of TX/RX LO and LPF imbalance model, b) Equivalent baseband representation combining the propagation channel and TX/RX IQI into the effective impulse responses: $\boldsymbol{h}_\alpha(t)$ and $\boldsymbol{h}_\beta(t)$.

symbols at the transmitter can be expressed in discrete-time as

$$\boldsymbol{u}_l = \jmath^l \sum_{n=-L_g}^{L_g} \mathbf{F}_n \boldsymbol{s}_{n+l} \tag{1}$$

where the synthesis filterbank is collected into a $K/2 \times K$ precoding matrix $\mathbf{F}_n(k,m) = g(k - nK/2)\exp\left(\jmath 2\pi \frac{mk}{K}\right)\exp\left(\jmath\phi_{n,m}\right)$ with the subchannel phase rotation given by $\phi_{n,m} = (n+m)\,\pi/2 - nm\pi$ and $\boldsymbol{s}_n = [\boldsymbol{s}_n(0), \boldsymbol{s}_n(1), \cdots, \boldsymbol{s}_n(K-1)]^T$ is the real 2^{M_p}-ary pulse amplitude modulated (PAM) symbol vector at the n-th time instant. We assume $\boldsymbol{s}_n(m) \in \mathcal{X}_p$ to be independent and identically distributed (i.i.d) with zero mean and variance $\sigma_{\boldsymbol{s}}^2$. The TX-side IQI is represented by the direct channel $\boldsymbol{g}_{1T} = \alpha_T \boldsymbol{a}_T + \beta_T \boldsymbol{b}_T$ and the image channel $\boldsymbol{g}_{2T} = \alpha_T \boldsymbol{b}_T + \beta_T \boldsymbol{a}_T$ that generates a signal component with the conjugate of the ideal TX signal $\boldsymbol{u}_l$ as

$$\boldsymbol{x}_l = \boldsymbol{g}_{1T} \otimes \boldsymbol{u}_l + \boldsymbol{g}_{2T} \otimes \boldsymbol{u}_l^* \tag{2}$$

when $\alpha_T = \cos\theta_T + \jmath\varepsilon_T \sin\theta_T$ and $\beta_T = \varepsilon_T \cos\theta_T + \jmath \sin\theta_T$ represented the NFS IQI, and $\boldsymbol{a}_T = \left(\boldsymbol{g}_T^I + \boldsymbol{g}_T^Q\right)/2$ and $\boldsymbol{b}_T = \left(\boldsymbol{g}_T^I - \boldsymbol{g}_T^Q\right)/2$ model the frequency-selective (FS) IQI. As this waveform is transmitted over a multipath channel, it undergoes convolution with the sample-spaced channel impulse response $\boldsymbol{h} \in \mathbb{C}^{L_h \times 1}$ having statistically independent taps and correlation matrix $\mathbf{R}_{\boldsymbol{h}} = \mathcal{E}\left\{\boldsymbol{h}\boldsymbol{h}^H\right\}$. For TX signal (2), it is split into a lower-triangular self-fading part $\mathbf{H}_0$ and an intersymbol interference part $\mathbf{H}_1$. Assuming that the channel remains fixed over several consecutive blocks and its length L_h is shorter than the block length $M = K/2$, we have

$$\boldsymbol{v}_l = \mathbf{H}_0 \boldsymbol{x}_l + \mathbf{H}_1 \boldsymbol{x}_{l-1} + \boldsymbol{w}_l \tag{3}$$

with the additive proper Gaussian noise $\boldsymbol{w}_l \sim \mathcal{CN}(\mathbf{0}, \sigma_{\boldsymbol{w}}^2 \mathbf{I})$. After RX IQI mixing, a general I/Q mismatch match model is shown Fig. 1 and is defined using (2) and (3) as:

$$\boldsymbol{z}_l = \boldsymbol{g}_{1R} \otimes \boldsymbol{v}_l + \boldsymbol{g}_{2R} \otimes \boldsymbol{v}_l^* \tag{4}$$

$$= \boldsymbol{h}_\alpha \otimes \boldsymbol{u}_l + \boldsymbol{h}_\beta \otimes \boldsymbol{u}_l^* + \tilde{\boldsymbol{w}}_l \tag{5}$$

where $\boldsymbol{h}_\alpha \leftrightarrow \boldsymbol{h}_\beta \sim \mathbb{C}^{L\times 1}$, $\boldsymbol{g}_{1R} = \alpha_R \boldsymbol{a}_R + \beta_R^* \boldsymbol{b}_R$, $\boldsymbol{g}_{2R} = \alpha_R^* \boldsymbol{b}_R + \beta_R \boldsymbol{a}_R$ and $\tilde{\boldsymbol{w}}_l$ is the colored noise RV (for ideal RX i.e., $\boldsymbol{g}_{1R} = [1, 0, \ldots, 0]^T$, $\boldsymbol{g}_{2R} = [0, 0, \ldots, 0]^T$, and $\sigma_{\tilde{\boldsymbol{w}}}^2 \to \sigma_{\boldsymbol{w}}^2$). Finally, after demodulation, we get the received observation in a linear relation

$$\boldsymbol{y}_n = \sum_l \jmath^{-(l+n)} \mathbf{F}_{-l}^H \boldsymbol{z}_{n+l} \tag{6}$$

$$= \jmath^{-n} \sum_l \mathbf{G}_{n,l} \boldsymbol{s}_{n+l} + \bar{\boldsymbol{w}}_n. \tag{7}$$

By combining (1), (5) and (7), the effective frequency response can be derived as

$$\mathbf{G}_{n,l} = \underbrace{\mathbf{G}_{\alpha,l}^{(0)} + \jmath^{-1}\mathbf{G}_{\alpha,l}^{(1)}}_{=:\mathbf{G}_{\alpha,l}} + (-1)^n \underbrace{\left(\mathbf{G}_{\beta,l}^{(0)} + \jmath\mathbf{G}_{\beta,l}^{(1)}\right)}_{=:\mathbf{G}_{\beta,l}} \tag{8}$$

$$\mathbf{G}_{\alpha,l}^{(i)} = \sum_{\substack{l',l'' \\ l'+l''-i=l}} \mathbf{F}_{-l'}^H \mathbf{H}_\alpha^{(i)} \mathbf{F}_{l''} \tag{9}$$

$$\mathbf{G}_{\beta,l}^{(i)} = \sum_{\substack{l',l'' \\ l'+l''+i=l}} (-1)^{l'} \mathbf{F}_{-l'}^H \mathbf{H}_\beta^{(i)} \mathbf{F}_{l''}^* \tag{10}$$

where the $M \times M$ channel matrices are defined as $\mathbf{H}^{(0)} = (\boldsymbol{h})_{\mathcal{L}}$, $\mathbf{H}^{(1)} = \mathbf{H} - (\boldsymbol{h})_{\mathcal{L}}$ and $\mathbf{H} = (\boldsymbol{h})_{\mathcal{C}}$.

III. Widely Linear Receiver

In this section, we derive time-invariant FIR filters for the widely linear equalizers (WLE) that exploits the statistical dependence between $\boldsymbol{z}_l$ and $\boldsymbol{z}_l^*$ in an LMMSE framework. The problem of the time-domain (TD) WLE is to find non-causal[1] FIR filters $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ of length $2L'+1 \leq M$ in

$$\hat{\boldsymbol{u}}_l(k) = \boldsymbol{f}_0^H \boldsymbol{z}_{l,k} + \boldsymbol{f}_1^H \boldsymbol{z}_{l,k}^* \tag{11}$$

with $\boldsymbol{z}_{l,k} = [\boldsymbol{z}_l(k-L'+1), \cdots, \boldsymbol{z}_l(k), \cdots, \boldsymbol{z}_l(k+L'-1)]^T$ such that $\mathcal{E}\left\{|\boldsymbol{u}_l(k) - \hat{\boldsymbol{u}}_l(k)|^2\right\}$ is minimized. Assuming perfect CSI knowledge, the solution for $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ is given by [9]

$$\boldsymbol{f}_0 = \left(\mathbf{R}_{zz} - \mathbf{R}_{\overline{zz}}\mathbf{R}_{zz}^{-1^*}\mathbf{R}_{\overline{zz}}^*\right)^{-1}\left(\mathbf{R}_{zu} - \mathbf{R}_{\overline{zz}}\mathbf{R}_{zz}^{-1^*}\mathbf{R}_{\overline{zu}}^*\right) \tag{12}$$

$$\boldsymbol{f}_1 = \left(\mathbf{R}_{zz}^* - \mathbf{R}_{\overline{zz}}^*\mathbf{R}_{zz}^{-1}\mathbf{R}_{\overline{zz}}\right)^{-1}\left(\mathbf{R}_{\overline{zu}}^* - \mathbf{R}_{\overline{zz}}^*\mathbf{R}_{zz}^{-1}\mathbf{R}_{zu}\right). \tag{13}$$

The correlation matrices required for (12) and (13) are listed in Appendix IV. As the estimandum $\hat{\boldsymbol{u}}_l(k)$ is complex and $\boldsymbol{h}_\alpha \neq \boldsymbol{h}_\beta^*$, we have $\boldsymbol{f}_0 \neq \boldsymbol{f}_1^*$. Basically, (11) could benefit from the non-circularity of the TX signal $\boldsymbol{u}_l$ inheriting from the real input symbols $\boldsymbol{s}_l$ and the non-circularity introduced by IQI when $\mathbf{H}_\beta^{(i)} \neq \mathbf{0}$. One can note that in (31) and (33),

[1]Otherwise, we need to restrict to causal minimum phase channels $\boldsymbol{h}_\alpha$ and $\boldsymbol{h}_\beta$ so that the equalizer system is strictly causal. This is ensured when $\mathcal{E}\left\{|\boldsymbol{h}(0)|^2\right\} > \sum_{i=1}^{L_h-1}\mathcal{E}\left\{|\boldsymbol{h}(i)|^2\right\}$ that holds particularly for line-of-sight conditions [11] or exponentially decaying power delay profile (PDP) with small $r.m.s$ delay spread. Unfortunately, even if $\boldsymbol{h}_\alpha$ is minimum phase, $\boldsymbol{h}_\beta$ might still require non-causal filters.

there is no term with sole dependence on $\boldsymbol{h}_\alpha$. The reason is that the complimentary correlation (36) gives a (M-*periodic*) non-zero value only in conjunction with transmit precoding (e.g., subchannel weighting [12]) or non-orthogonal prototype filters. In other words, as the IQI gets insignificant, the complimentary auto-correlation of observation $\mathbf{R}_{\overline{zz}} \to \mathbf{0}$ and between observation and desired signal $\mathbf{R}_{\overline{zu}} \to \mathbf{0}$, resulting in $\boldsymbol{f}_1 \to \mathbf{0}$.

Nevertheless, taking (5) as an observation reference does not pose intrinsic contamination problems and depending on the IQI severity, this WLE structure could provide a mean square error (MSE) improvement over the strictly linear equalizer (SLE) as shown in (15),

$$\mathrm{MSE}_{WLE} = \mathcal{E}\left\{|\boldsymbol{u}_l(k) - \hat{\boldsymbol{u}}_l(k)|^2\right\} \tag{14}$$

$$= \mathrm{MSE}_{SLE} - \boldsymbol{r}^H \mathbf{R}^{-1} \boldsymbol{r} \tag{15}$$

using the facts that $\boldsymbol{r} = \mathbf{R}^*_{\overline{zu}} - \mathbf{R}^*_{\overline{zz}}\mathbf{R}^{-1}_{zz}\mathbf{R}_{zu}$, $\mathbf{R} = \mathbf{R}^*_{zz} - \mathbf{R}^*_{\overline{zz}}\mathbf{R}^{-1}_{zz}\mathbf{R}_{\overline{zz}}$, the inverse of the matrix $\mathbf{R}$ is a positive-definite matrix and when $\boldsymbol{h}_\beta \neq \mathbf{0}$, the Schur complement $\mathbf{R}^*_{\overline{zu}} - \mathbf{R}^*_{\overline{zz}}\mathbf{R}^{-1}_{zz}\mathbf{R}_{zu} > \mathbf{0}$ implying $\mathrm{MSE}_{WLE} < \mathrm{MSE}_{SLE}$.

IV. Frequency-Domain Equalization with Iterative Cancellation

As the WLE approach (11) needs computationally expensive operations, this section treats low complexity equalization operating over the demodulated signal (7). The proposed solution adapts the approach in [2]. The distinguishing features, however, are the pair-wise equalization (PWE) to suppress IQI and the use of the channel correlation knowledge for the optimal feedback design. The cases, when the interference cancellation has adverse impact on the performance of WLE, are identified.

A. WLE in Frequency-Domain (FD)

Lets us now consider an equivalent but concise channel model in which the *desired* $\boldsymbol{y}_n(m)$ and *image* $\boldsymbol{y}_n(K-m)$ signals are written together in a convenient way.[2] In this representation, each subchannel $\boldsymbol{y}_n(m)$ contains contribution from $\boldsymbol{s}_n(m)$ and $\boldsymbol{s}_n(K-m)$, and an interference leakage from other subchannels weighed by the ambiguity operator $\boldsymbol{\mathcal{I}}_{n'} = \sum_{l'+l''=n'} \mathbf{F}^H_{-l'}\mathbf{F}_{l''}$. The effects of the TX/RX caused IQI can be summarized in an equivalent channel matrix expressed as[3]

$$\mathbf{H}_m = \begin{pmatrix} \mathbf{G}_{\alpha,0}(m,m) & -\mathbf{G}_{\beta,0}(m,K-m) \\ -\mathbf{G}^*_{\beta,0}(K-m,m) & \mathbf{G}^*_{\alpha,0}(K-m,K-m) \end{pmatrix} \tag{16}$$

whose exemplary characteristics are depicted in Fig. 2. Note that although $\mathbf{H}_m$ is diagonally dominant, in LMMSE framework, the optimal strategy is to exploit this joint desired/image channel, rather than trying to cancel out the image interference as in [6], [8]. For the computation of WLE, one needs the joint channel frequency response statistics. The channel frequency response on different subchannels can be related by

[2]For presentation simplicity, we disregard the subchannels at 0 and $K/2$ in our further analysis as they rather suffer from self-image.

[3]Image channel $\mathbf{H}_m(i,j)$, $\forall i \neq j$ is almost surely weaker than the direct channel $\mathbf{H}_m(i,i)$ under realistic mismatch conditions allowing maximal information flow through $\mathbf{H}_m(i,i)$.

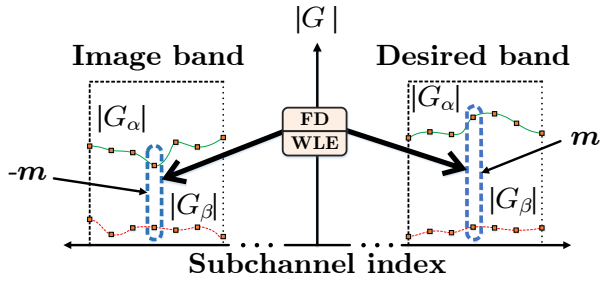

Fig. 2. FD-WLE estimates the desired and image signal simultaneously by exploiting the non-circularity of PAM symbols $\tilde{\boldsymbol{s}}_{n,m}$ and the cross-channel coefficients $\mathbf{H}_m(1,2)$ and $\mathbf{H}_m(2,1)$.

the subchannel correlation $\rho_{m'}$, namely

$$\mathbf{H}_{m+m'} = \rho_{m'}\mathbf{H}_m + \sqrt{1-|\rho_{m'}|^2}\boldsymbol{\Delta}_{m'} \tag{17}$$

$$= \tilde{\mathbf{H}}_m + \tilde{\boldsymbol{\Delta}}_{m'} \tag{18}$$

which is a combination of a fully correlated component $\tilde{\mathbf{H}}_m$ and an independent component $\tilde{\boldsymbol{\Delta}}_{m'}$ weighted by the correlation factor $\rho_{m'}$ and distributed according to $\mathcal{E}\left\{\mathbf{H}_{m+m'}\mathbf{H}^H_m\right\} = \rho_{m'}\mathbf{I}$ and $\mathbf{H}_m \leftrightarrow \boldsymbol{\Delta}_m \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$.[4,5]

Merging the subchannel pair $(m, K-m)$ into a matrix notation, we get

$$\tilde{\boldsymbol{y}}_{n,m} = \mathbf{H}_m\left(\tilde{\boldsymbol{s}}_{n,m} + \sum_{(n',m')\in\Omega} \rho_{m'}\boldsymbol{\mathcal{I}}_{n'}(0,m')\tilde{\boldsymbol{s}}_{n+n',m+m'}\right) + \underbrace{\sum_{(n',m')\in\Omega} \sqrt{1-|\rho_{m'}|^2}\boldsymbol{\mathcal{I}}_{n'}(0,m')\boldsymbol{\Delta}_{m'}\tilde{\boldsymbol{s}}_{n+n',m+m'}}_{=:\boldsymbol{\mathcal{K}}_{n,m}} + \bar{\boldsymbol{w}}_{n,m} \tag{19}$$

where $\tilde{\boldsymbol{y}}_{n,m} = \big(\boldsymbol{y}_n(m), \boldsymbol{y}_n(K-m)\big)^T$ and $\tilde{\boldsymbol{s}}_{n,m} = \big(\boldsymbol{s}_n(m), \boldsymbol{s}_n(K-m)\big)^T$. Even though $\bar{\boldsymbol{w}}_{n,m}$ is additive Gaussian noise that is colored by the receiver filter mismatch, we assume RX IQI is small so that $\sigma^2_{\bar{\boldsymbol{w}}} \approx \sigma^2_{\tilde{\boldsymbol{w}}}/2$. In the absence of intrinsic mitigation, the uncorrelated component constitutes an intrinsic contamination and $\boldsymbol{\mathcal{K}}_{n,m}$ must be treated as additive distortion component. In fact, for a frequency-flat channel, intrinsic contamination vanishes ($\sigma^2_{\boldsymbol{\mathcal{K}}} \to 0$) as $\rho_{m'} \to 1$.

In the case of improper signals, a suitable set of coefficients $\widetilde{\boldsymbol{W}}_m$ can be computed based on MSE minimization criterion using (19) such that

$$\widehat{\boldsymbol{s}}_{n,m} = 2\Re\left\{\boldsymbol{\mathcal{P}}_n\widetilde{\boldsymbol{W}}_m\boldsymbol{\mathcal{P}}_n\tilde{\boldsymbol{y}}_{n,m}\right\} \tag{20}$$

where we have defined

$$\boldsymbol{\mathcal{P}}_n = \begin{bmatrix} 1 & 0 \\ 0 & (-1)^n \end{bmatrix}. \tag{21}$$

Applying the sames principles of WLE as in Section III, we

[4]Correlation between the direct $\mathbf{H}_m(i,i)$ and image $\mathbf{H}_m(i,j), i \neq j$ channels, and the dependence on I/Q parameters are ignored for simplicity i.e., $\mathcal{E}\left\{|\mathbf{H}_m(i,i)|^2 + |\mathbf{H}_m(i,j)|^2\right\} \approx 1$. Note that due to the correlation symmetry, $\mathcal{E}\left\{\mathbf{H}_{m+m'}(i,i)\mathbf{H}^*_m(i,i)\right\} = \mathcal{E}\left\{\mathbf{H}_{m+m'}(i,j)\mathbf{H}^*_m(i,j)\right\} = \rho_{m'}$.

[5]For exponential PDP, it can be shown that $\rho_{m'} = 1/(1-j2\pi m'\tau_{rms})$ in a CP-OFDM waveform, where τ_{rms} is the normalized $r.m.s$ delay spread.

have

$$\widetilde{\boldsymbol{W}}_m = \left(\mathbf{R}_{YY} - \mathbf{R}_{\overline{YY}}\mathbf{R}_{YY}^{-1\,*}\mathbf{R}_{\overline{YY}}^{*}\right)^{-1} \times \left(\mathbf{R}_{YS} - \mathbf{R}_{\overline{YY}}\mathbf{R}_{YY}^{-1\,*}\mathbf{R}_{\overline{YS}}^{*}\right) \quad (22)$$

with $\mathbf{R}_{YS} = \mathcal{E}\left\{\tilde{\boldsymbol{y}}_{n,m}\tilde{\boldsymbol{s}}_{n,m}^H\right\} = \mathbf{R}_{\overline{YS}}$, $\mathbf{R}_{YY} = \mathcal{E}\left\{\tilde{\boldsymbol{y}}_{n,m}\tilde{\boldsymbol{y}}_{n,m}^H\right\}$ and $\mathbf{R}_{\overline{YY}} = \mathcal{E}\left\{\tilde{\boldsymbol{y}}_{n,m}\tilde{\boldsymbol{y}}_{n,m}^T\right\}$.

Proposition 1: The WL-PWE degenerates into SL-PWE provided that the channel is frequency-flat and the noise $\bar{\boldsymbol{w}}_{n,m}$ exhibits null complementary correlation i.e., $\mathbf{R}_{\overline{WW}} = \mathcal{E}\left\{\bar{\boldsymbol{w}}_{n,m}\bar{\boldsymbol{w}}_{n,m}^T\right\} = \mathbf{0}$.

Proof: Maximum performance gain $\delta_e^2 \propto \frac{\lambda_i}{\lambda_i^2 - \overline{\lambda}_i^2}$ over SL-PWE is achieved when the eigenvalues $\overline{\lambda}_i$ of $\mathbf{R}_{\overline{YY}} = \sigma_{\boldsymbol{s}}^2\left(1 + \sum_{(n',m')\in\Omega}\rho_{m'}^2\boldsymbol{\mathcal{I}}_{n'}^2(0,m')\right)\mathbf{H}_m\mathbf{H}_m^T + \mathbf{R}_{\overline{WW}}$ approach eigenvalues λ_i of $\mathbf{R}_{YY}$ [9]. For $\sigma_n^2 \to 0$ and when $\rho_i \ll 1, \forall\ i \neq 0$, $\sum_{(n',m')\in\Omega}\rho_{m'}^2\boldsymbol{\mathcal{I}}_{n'}^2(0,m') \approx 0$ and WLE promises significant MSE improvement. On the other hand, if $\rho_i \to 1, \forall\ i$, then $\sum_{(n',m')\in\Omega}\boldsymbol{\mathcal{I}}_{n'}^2(0,m') \to -1$, $\mathbf{R}_{\overline{YY}} \to \mathbf{0}$ and circularity can only be avoided by a non-zero $\mathbf{R}_{\overline{WW}}$. ■

Proposition 2: The upper bound on MSE performance of WL-PWE is

$$\sigma_{\boldsymbol{e}}^2 \leq \sigma_{\boldsymbol{s}}^2\left(1 - \frac{1}{2}\sum_{t=1}^{2}\lambda_t\left(\mathbf{B}\right)\lambda_t^{-1}\left(\mathbf{R}_{\boldsymbol{yy}}\right)\right) \quad (23)$$

given that $\mathbf{B} = \sigma_{\boldsymbol{s}}^2\left(\mathbf{H}_m^H\mathbf{H}_m + \mathbf{H}_m^T\mathbf{H}_m^*\right)$.

Proof: Let us reformulate our augmented signals as $\dot{\boldsymbol{y}}_{n,m} = \left(\tilde{\boldsymbol{y}}_{n,m}^T\ \ \tilde{\boldsymbol{y}}_{n,m}^H\right)^T$, $\dot{\boldsymbol{s}}_{n,m} = \left(\tilde{\boldsymbol{s}}_{n,m}^T\ \ \tilde{\boldsymbol{s}}_{n,m}^H\right)^T$, then the WL-PWE estimator for even n resembles a strictly-linear LMMSE filter, i.e., $\widehat{\dot{\boldsymbol{s}}}_{n,m} = \mathbf{W}_m\dot{\boldsymbol{y}}_{n,m} = \mathbf{R}_{\boldsymbol{sy}}\mathbf{R}_{\boldsymbol{yy}}^{-1}\dot{\boldsymbol{y}}_{n,m}$ and the error variance of an element in $\widehat{\dot{\boldsymbol{s}}}_{n,m}$ can be simply calculated as

$$\sigma_{\boldsymbol{e}}^2 = \frac{1}{4}\mathrm{Tr}\left\{\mathbf{R}_{\boldsymbol{ss}} - \mathbf{R}_{\boldsymbol{sy}}\mathbf{R}_{\boldsymbol{yy}}^{-1}\mathbf{R}_{\boldsymbol{sy}}^H\right\} \quad (24)$$

$$\overset{(a)}{\leq} \sigma_{\boldsymbol{s}}^2 - \frac{1}{4}\sum_{t=1}^{4}\lambda_t\left(\mathbf{R}_{\boldsymbol{sy}}^H\mathbf{R}_{\boldsymbol{sy}}\right)\lambda_t\left(\mathbf{R}_{\boldsymbol{yy}}^{-1}\right) \quad (25)$$

$$\overset{(b)}{=} \sigma_{\boldsymbol{s}}^2\left(1 - \frac{1}{2}\sum_{t=1}^{2}\lambda_t\left(\mathbf{B}\right)\lambda_t^{-1}\left(\mathbf{R}_{\boldsymbol{yy}}\right)\right) \quad (26)$$

where in step (a) we used cyclic trace property and matrix eigenvalue inequality [13] and step (b) results from the fact that $\mathbf{R}_{\boldsymbol{sy}}$ has only 2 non-zero eigenvalues. ■

Using Weyl's theorem and assuming the ordering $\lambda_1 \geq \lambda_2$, it is obvious from (23) that $\lambda_i\left(\mathbf{H}_m^H\mathbf{H}_m\right) + \lambda_2\left(\mathbf{H}_m^T\mathbf{H}_m^*\right) \leq \lambda_i\left(\mathbf{B}\right)/\sigma_{\boldsymbol{s}}^2 \leq \lambda_i\left(\mathbf{H}_m^H\mathbf{H}_m\right) + \lambda_1\left(\mathbf{H}_m^T\mathbf{H}_m^*\right)$ whereas the eigenvalues of $\mathbf{R}_{\boldsymbol{yy}}$ gets spread out. As error variance is decreasing function of the eigenvalues $\lambda_i\left(\mathbf{B}\right)$, we can deduce that WL-PWE outperforms SL-PWE. Interestingly, the presence of off-diagonal elements in $\mathbf{H}_m$ (due to IQI) causes the eigenvalues of $\mathbf{R}_{\boldsymbol{yy}}$ and $\mathbf{B}$ to spread out in such a manner that summation in (23) exceeds no IQI case (partly due to the fact that $\sum_{j=0}^{L-1}\mathcal{E}\left\{|\boldsymbol{h}_\alpha(j)|^2 + |\boldsymbol{h}_\beta(j)|^2\right\} \geq \sum_{i=0}^{L_h-1}\mathcal{E}\left\{|\boldsymbol{h}(i)|^2\right\}$) and hence improves the error performance.

B. Residual Interference Cancellation

In the signal model (19), $\boldsymbol{\mathcal{K}}_{n,m}$ accounts for the intrinsic contamination and depending on the channel correlation, it could limit the interference mitigation capability of (20). A residual interference canceler (RIC) reconstructs the interfering contribution $\widehat{\boldsymbol{\mathcal{K}}}_{n,m}$ using the decision mechanism on the basis of the past estimates $\widehat{\boldsymbol{s}}_n(m)$. Pre-processing the observation deprives the resultant signal of the intrinsic interference in

$$\bar{\boldsymbol{y}}_n(m) = \boldsymbol{y}_n(m) - \sum_{(n',m')\in\Omega'}\mathbf{G}_{\alpha,n'}(m,m+m')\widehat{\boldsymbol{s}}_{n+n'}(m+m') + (-1)^n\sum_{(n',m')\in\Omega'}\mathbf{G}_{\beta,n'}(m,m'-m)\widehat{\boldsymbol{s}}_{n+n'}(m'-m) \quad (27)$$

where $\Omega' \subseteq \Omega$ denotes RIC target space and hence $\Omega'' = \Omega \setminus \Omega'$ is the remaining contamination. It is noteworthy that RIC not only considers the direct link interference channel $\mathbf{G}_{\alpha,n'}$ but also the one originating from the image band $\mathbf{G}_{\beta,n'}$. As long as the channel state information (CSI) is accurate and the decisions $\widehat{\boldsymbol{s}}_n(m)$ are correct, RIC detector $\widehat{\bar{\boldsymbol{s}}}_{n,m} = \mathcal{D}\left\{\widetilde{\boldsymbol{W}}_m\bar{\boldsymbol{y}}_{n,m}\right\}$ outperforms the linear estimator (20).

1) Optimal Cancellation Order: For practical consideration, the performance of RIC will deteriorate if the CSI is erroneous (modeled here as $\hat{\mathbf{H}}_m = \mathbf{H}_m + \mathbf{E}_m$ with $\mathcal{E}\left\{\mathbf{E}_m\boldsymbol{\Delta}_m^H\right\} = \mathbf{0}$ and $\mathcal{E}\left\{\mathbf{E}_m\mathbf{E}_m^H\right\} = \sigma_{\boldsymbol{e_h}}^2\mathbf{I}$). In these cases, the subtraction of an intrinsic component $\widehat{\boldsymbol{\mathcal{K}}}_{n,m}$ will constitute another interfering source. Therefore, an optimization procedure for Ω' is introduced in the proposed design criterion to avoid performance deterioration with RIC. Let us define an MSE cost function after RIC and assuming perfect initial decisions as $\sigma_{\breve{\boldsymbol{e}}}^2 = \mathcal{E}\left\{\left|\boldsymbol{s}_n(m) - \widehat{\bar{\boldsymbol{s}}}_n(m)\right|^2\right\}$.

Remark 1: It is important to note that the eigenstructure of $\mathbf{R}_{YY}$ and $\mathbf{R}_{\overline{YY}}$ is a function of channel correlation $\rho_{m'}$ and RIC order $|\Omega'|$. As the RIC order $|\Omega'|$ increases, the eigenvalues of $\mathbf{R}_{YY}$ fall whereas those of $\mathbf{R}_{\overline{YY}}$ strengthen, meaning that a higher δ_e^2 is possible. If $\rho_{m'} \to 0$ for $m' \neq 0$, then a higher RIC order results in the reduction of intrinsic interference in $\mathbf{R}_{YY}$ composition and hence a lower $\sigma_{\breve{\boldsymbol{e}}}^2$.

Remark 2: The RIC comes at the cost of imperfect channel knowledge. When the CSI is significantly erroneous, the distortion due to the channel estimation error $\sigma_{\breve{\boldsymbol{h}}}^2 = \sigma_{\boldsymbol{e_h}}^2\sigma_{\boldsymbol{s}}^2\sum_{(n',m')\in\Omega'}|\boldsymbol{\mathcal{I}}_{n'}(0,m')|^2$ augmented as noise in $\mathbf{R}_{YY}$ supersedes the advantages of the interference cancellation. On the other hand, when $\sigma_{\boldsymbol{e_h}}^2 \ll 1 - |\rho_{m'}|^2$, one should target broader RIC, i.e., $|\Omega'| \gg 1$ to mitigate $\boldsymbol{\mathcal{K}}_{n,m}$.

2) Decision-Directed Cancellation in Coded Systems: In many practical wireless systems, bit interleaved coded modulation (shown in Fig. 3) offers a significant improvement over fading channels, in addition reducing decision errors in RIC. The data symbols $\boldsymbol{s}_n(m)$ are generated by concatenating a binary convolutional encoder producing the codeword $\boldsymbol{c}_i = \mathbf{G}\cdot\boldsymbol{b}_i$ with the PAM mapper $\mathcal{X}_p$ through a bitwise interleaver denoted by $\boldsymbol{\Pi}$. At the receiver, the tentative estimates $\widehat{\boldsymbol{s}}_n^{(0)}$ are obtained taking advantage of the approximate orthogonality of SMT (assuming $\boldsymbol{\mathcal{K}}_{n,m}$ is small). Afterwards, the demapper computes LLRs that are then rearranged with $\boldsymbol{\Pi}^{-1}$ and finally, the Viterbi decoder processes the received code bit steam $\hat{\boldsymbol{c}}_i$ to produce the information sequence $\hat{\boldsymbol{b}}$. The intrinsic contamination $\widehat{\boldsymbol{\mathcal{K}}}_{n,m}$ is reconstructed and subtracted from the observation group Ω'

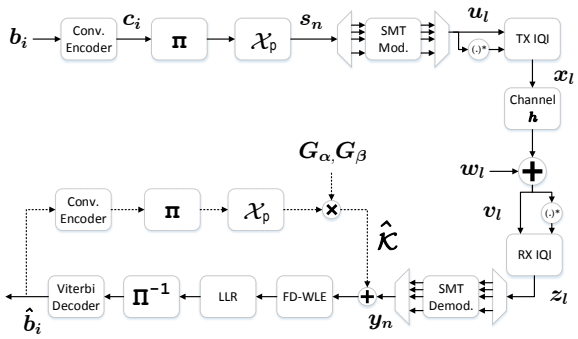

Fig. 3. Block diagram for interference cancellation in coded SMT systems operating over AWGN, TX/RX IQI and frequency-selective channel.

in the feedback loop. After executing RIC (27), the BICM decoding is reapplied to (20).

V. Numerical Analysis

A comprehensive analysis of the SMT WLE-based receiver in coherent Rayleigh-fading channel with TX/RX IQI is provided in terms of the coded and uncoded error probabilities. In the following simulations, the total number of subchannels is $K = 64$ and extended Gaussian pulses [1] are used with $\lambda = 2$ and $L_g = 4$. A 10-tap channel model is assumed with i.i.d complex Gaussian RVs and generated with exponential PDP. The variance of the channel taps is normalized by $\sum_{l=0}^{L_h-1} \mathcal{E}\left\{|\boldsymbol{h}(l)|^2\right\} = 1$. Relatively higher IQI is considered in RX w.r.t TX to mimic a downlink cellular scenario with NFS $\varepsilon_T = 0.04$, $\theta_T = 4°$, $\varepsilon_R = 0.08$, $\theta_R = 8°$ and FS $\boldsymbol{a}_T = (1, 0.091)^T$, $\boldsymbol{b}_T = (1, -0.05)^T$, $\boldsymbol{a}_R = (1, -0.2)^T$, $\boldsymbol{b}_R = (1, 0.3)^T$ parameters. The BICM structure has a rate $1/2$ convolutional code with generator polynomial $[171, 133]_8$ combined with a random interleaver $\boldsymbol{\Pi}$.

The first simulation addresses the analysis of the proposed WLE implementation from section III. According to Fig. 4, the widely linear method (TD-WLE) yields a good performance gain over SLE especially for higher constellations and effectively robust to the mismatch parameters. Note that efficient DFT-based computation of TD-WLE filters has a loss of e.g., 2-dB at BER of 3×10^{-3} and suffers from the error-flooring at high SNR. This is because convolution circularity is assumed in this method and the approximate frequency response $\mathcal{DFT}(\boldsymbol{h})$ deviates from the exact channel realization $\mathbf{G}_{n,l}$; an effect that strengthens with higher channel dispersion.

In Fig. 5, we plot the BERs of the WLE receivers in a frequency-domain mitigation approach for normalized r.m.s delay spreads of $1/3$ and 5. While in low-selectivity regime, PWE provides a substantial BER improvement, contamination due to the intrinsic interference $\boldsymbol{\mathcal{K}}_{n,m}$ prevails at high channel dispersion (when $\tau_{rms} \to \infty$ and $L_h \gg 1$) leading to poor performance and error floor. This behavior motivates us to look for means to suppress the intrinsic contamination via RIC scheme. Fig. 6 shows the relationship of symbol MSE performance with the channel time dispersion. Note that because most of the interference originates from the nearest subchannels/symbols, interference is almost entirely canceled out by the first few neighbors. The effect of RIC is more pronounced at higher values of τ_{rms}. The presence of IQI improves the symbol estimates as eigenvalues in (23) become more spread out. Also, wider RIC with imperfect channel is worse than narrower RIC with better CSI knowledge.

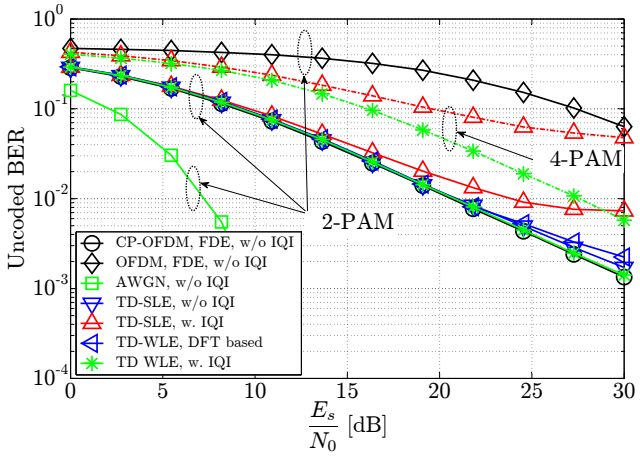

Fig. 4. Uncoded BER versus symbol SNR with 2- or 4-PAM (accordingly 4- or 16-QAM for CP-OFDM), $L' = 10$, and $\tau_{rms} = 1$.

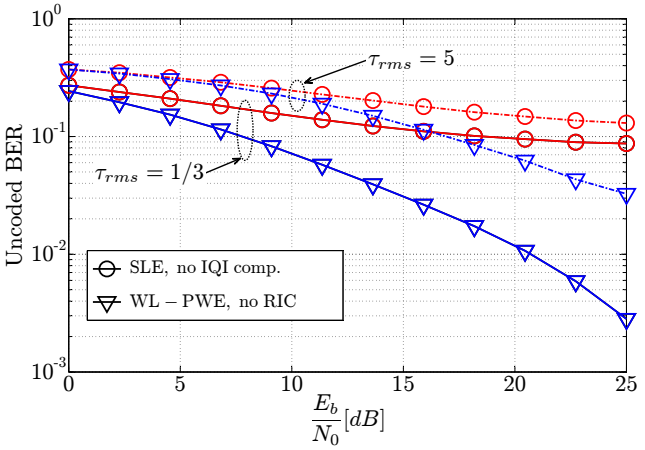

Fig. 5. BER versus bit energy per noise spectral density in an uncoded receiver with 4-PAM constellation and perfect CSI.

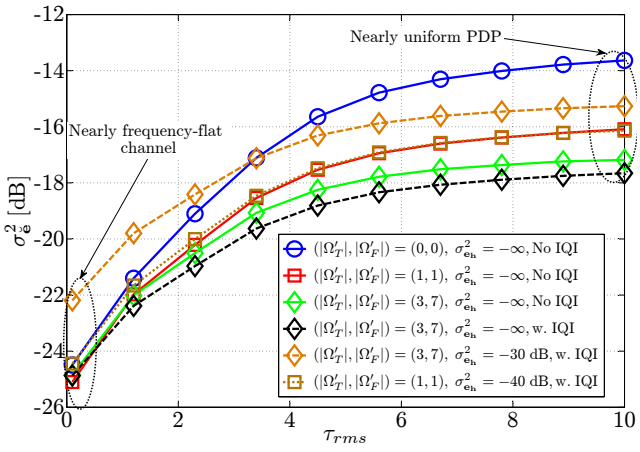

Fig. 6. Symbol MSE versus channel normalized $r.m.s$ delay spread for various RIC configurations and estimation accuracy levels at high SNR $= 30$ dB.

Finally, in Fig. 7 we show a comparison of the BER performance among the conventional (SLE) and proposed (PWE) equalizers with the convolution channel codes. The PWE with RIC gains around 3 dB at a BER of 10^{-4} in the case of 4-PAM signals. This is because higher constellations suffer more from intrinsic contamination. It is also obvious that exploiting RIC with decoding iterations helps improves WL-PWE performance. On the other hand, for small constellations, PWE without RIC is enough for the near-optimal performance.

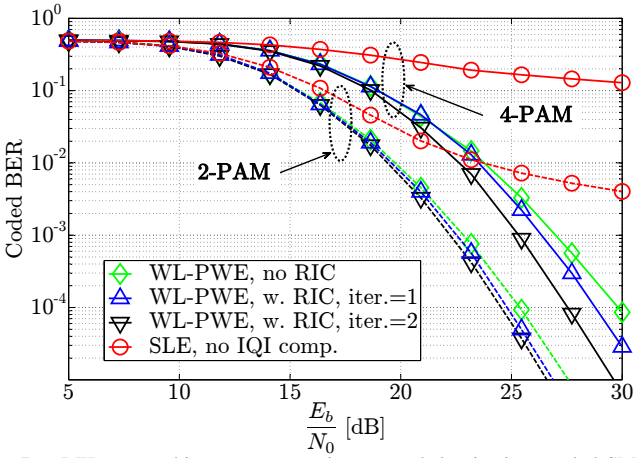

Fig. 7. BER versus bit energy per noise spectral density in a coded SMT system with $\sigma^2_{\boldsymbol{e_h}} = -40$ dB, $\tau_{rms} = 5$ and optimized RIC size.

VI. Conclusions

In this study, we considered the WLE implementation for SMT systems in the presence of IQI. First, we derived the WL solution that can directly process the modulated signal. Although the compensation for IQI distortion is implicit, it demands high computational burden for matrix inversion operations. Therefore, a novel iterative FD equalizer was proposed which consists of an initial detection stage assuming perfect orthogonality while combining the direct and image band processing. It was followed by a set of RIC filters to compensate for the intrinsic interference. Because RIC is influenced by channel selectivity and CSI imperfectness, we provide some guidelines on the selection of the RIC loop-filter length based on the MSE cost function. The use of RIC not only reduces the channel distortion but can also yield a higher non-circularity. The numerical experiments show that the proposed scheme outperforms the conventional methods and can effectively alleviate the loss of orthogonality caused by the channel selectivity and the IQI.

Appendix A
Correlation Matrices for TD-WLE

Maintaining the structure of (5), the correlation matrices can be written as

$$\tilde{\mathbf{R}}_{zz} = \mathcal{E}\left\{\boldsymbol{z}_l \boldsymbol{z}_l^H\right\} \tag{28}$$

$$= 2\sigma_{\boldsymbol{s}}^2 \Big(\mathbf{H}_\alpha^{(0)} {\mathbf{H}_\alpha^{(0)}}^H + \mathbf{H}_\alpha^{(1)} {\mathbf{H}_\alpha^{(1)}}^H + \mathbf{H}_\beta^{(0)} {\mathbf{H}_\beta^{(0)}}^H + \mathbf{H}_\beta^{(1)} {\mathbf{H}_\beta^{(1)}}^H \Big) + \mathbf{R}_{nn} \tag{29}$$

$$\tilde{\mathbf{R}}_{\overline{zz}} = \mathcal{E}\left\{\boldsymbol{z}_l \boldsymbol{z}_l^T\right\} \tag{30}$$

$$= 2\sigma_{\boldsymbol{s}}^2 \Big(\mathbf{H}_\alpha^{(0)} {\mathbf{H}_\beta^{(0)}}^T + \mathbf{H}_\alpha^{(1)} {\mathbf{H}_\beta^{(1)}}^T + \mathbf{H}_\beta^{(0)} {\mathbf{H}_\alpha^{(0)}}^T + \mathbf{H}_\beta^{(1)} {\mathbf{H}_\alpha^{(1)}}^T \Big) + \mathbf{R}_{\overline{nn}}. \tag{31}$$

It is easy to observe that $\tilde{\mathbf{R}}_\alpha = \mathbf{H}_\alpha^{(0)} {\mathbf{H}_\alpha^{(0)}}^H + \mathbf{H}_\alpha^{(1)} {\mathbf{H}_\alpha^{(1)}}^H$ is a Toeplitz matrix, except in top-right and bottom-left $L \times L$ corner blocks w.r.t the circulant matrix $\mathbf{R}_\alpha = \mathbf{H}_\alpha {\mathbf{H}_\alpha}^H$. For $\tilde{\boldsymbol{h}} = (0, \ldots, 0, \boldsymbol{h}(0), \boldsymbol{h}(1), \cdots, \boldsymbol{h}(L-1), 0, \cdots, 0)^T$, we have

$$\tilde{\mathbf{R}}_{zu} = \mathcal{E}\left\{\boldsymbol{z}_l \boldsymbol{u}_l^*(k)\right\} = 2\sigma_{\boldsymbol{s}}^2 \tilde{\boldsymbol{h}}_\alpha \tag{32}$$

$$\tilde{\mathbf{R}}_{\overline{zu}} = \mathcal{E}\left\{\boldsymbol{z}_l \boldsymbol{u}_l(k)\right\} = 2\sigma_{\boldsymbol{s}}^2 \tilde{\boldsymbol{h}}_\beta. \tag{33}$$

Now assuming a filter length of L' and substituting (29), (31), (32) and (33) in R.H.S of (12) and (13), one can obtain $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ from $\mathcal{F} = \{M/2 - L' + 1, \cdots, M/2, \cdots, M/2 + L' - 1\}$ elements of the product. It should be noted that expressions (29)-(33) are derived with the following characteristics of $\boldsymbol{u}_l$:

$$\mathcal{E}\left\{\boldsymbol{u}_l \boldsymbol{u}_{l+m}^H\right\} = \sigma_{\boldsymbol{s}}^2 \sum_n \mathbf{F}_n \mathbf{F}_{n-m}^H = \begin{cases} 2\sigma_{\boldsymbol{s}}^2 \mathbf{I}_M & m = 0 \\ \mathbf{0}_M & m \neq 0 \end{cases} \tag{34}$$

$$\mathcal{E}\left\{\boldsymbol{u}_l \boldsymbol{u}_{l+m}^T\right\} = \sigma_{\boldsymbol{s}}^2 j^{2l+m} \sum_n \mathbf{F}_n \mathbf{F}_{n-m}^T \tag{35}$$

or equivalently for TX signal $u(k)$ at k-th time instant, a (cyclostationary) complimentary correlation function results:

$$\mathcal{E}\left\{u(k)u(k+p)\right\} = \sigma_{\boldsymbol{s}}^2 \delta\left(2k + p + \eta M\right) \times \sum_n e^{jn\pi} g\left(k - nM\right) g\left(k + p - nM\right). \tag{36}$$

After being subject to the RX IQI, the noise $\tilde{\boldsymbol{w}}_l$ no longer remains white. Given that $\boldsymbol{w}_l$ has i.i.d elements and $\sum_l \mathbf{F}_l^H \mathbf{F}_l = \mathbf{I}$, $\sum_l \mathbf{F}_l^H \mathbf{F}_{l+n} = \mathbf{0}, n \neq 0$, the correlation and complementary correlation of the effective noise $\tilde{\boldsymbol{w}}_l$ are given as $\mathbf{R}_{nn} = \mathcal{E}\left\{\tilde{\boldsymbol{w}}_l \tilde{\boldsymbol{w}}_l^H\right\} = \sigma_{\boldsymbol{w}}^2 \left(\mathbf{A}_R \mathbf{A}_R^H + \mathbf{B}_R \mathbf{B}_R^H\right)$ and $\mathbf{R}_{\overline{nn}} = \mathcal{E}\left\{\tilde{\boldsymbol{w}}_l \tilde{\boldsymbol{w}}_l^T\right\} = \sigma_{\boldsymbol{w}}^2 \left(\mathbf{A}_R \mathbf{B}_R^T + \mathbf{B}_R \mathbf{A}_R^T\right)$ where $\mathbf{A}_R = (\boldsymbol{g}_{1R})_{\mathcal{L}}$ and $\mathbf{B}_R = (\boldsymbol{g}_{2R})_{\mathcal{L}}$.

References

[1] P. Siohan, C. Siclet, and N. Lacaille, "Analysis and design of OFDM/OQAM systems based on filterbank theory," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1170–1183, May 2002.

[2] Y. Cheng and M. Haardt, "Widely linear processing in MIMO FBMC/OQAM systems," in *Wireless Communication Systems (ISWCS 2013), Proceedings of the International Symposium on*, 2013, pp. 1–5.

[3] A. Ishaque and G. Ascheid, "I/Q imbalance and CFO in OFDM/OQAM systems: Interference analysis and compensation," in *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC 2013)*, 2013, pp. 381–386.

[4] Y. Yoshida, K. Hayashi, H. Sakai, and W. Bocquet, "Analysis and compensation of transmitter IQ imbalances in OFDMA and SC-FDMA systems," *IEEE Trans. Signal Process.*, vol. 57, no. 8, 2009.

[5] S. Krone and G. Fettweis, "Capacity analysis for OFDM systems with transceiver I/Q imbalance," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, 2008, pp. 1–6.

[6] Y.-H. Chung and S.-M. Phoong, "Joint estimation of I/Q imbalance, CFO and channel response for MIMO OFDM systems," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1485–1492, 2010.

[7] B. Narasimhan, S. Narayanan, H. Minn, and N. Al-Dhahir, "Reduced-complexity baseband compensation of joint Tx/Rx I/Q imbalance in mobile MIMO-OFDM," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1720–1728, 2010.

[8] L. Anttila, M. Valkama, and M. Renfors, "Circularity-based I/Q imbalance compensation in wideband direct-conversion receivers," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2099–2113, 2008.

[9] C. Cheong Took, S. Douglas, and D. Mandic, "On approximate diagonalization of correlation matrices in widely linear signal processing," *Signal Processing, IEEE Transactions on*, vol. 60, no. 3, March 2012.

[10] D. Darsena and et al., "Widely linear equalization and blind channel identification for interference-contaminated multicarrier systems," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 1163–1177, 2005.

[11] D. Cassioli and A. Mecozzi, "Minimum-phase impulse response channels," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3529–3532, 2009.

[12] H. Bolcskei, "Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 988–999, Jun. 2001.

[13] J. Liu, *Eigenvalue and Singular Value Inequalities of Schur Complements*, ser. Numerical Methods and Algorithms. Springer US, 2005.