

Spline-Based Adaptive Cancellation of Even-Order Intermodulation Distortions in LTE-A/5G RF Transceivers

Thomas Paireder¹, Graduate Student Member, IEEE, Christian Motz², Graduate Student Member, IEEE, and Mario Huemer³, Senior Member, IEEE

Abstract—Radio frequency transceivers operating in in-band full-duplex or frequency-division duplex mode experience strong transmitter leakage. Combined with receiver nonlinearities, this causes intermodulation products in the baseband, possibly with higher power than the desired receive signal. In order to restore the receiver signal-to-noise ratio in such scenarios, we propose two novel digital self-interference cancellation approaches based on spline interpolation. Both employ a Wiener structure, thereby matching the baseband model of the intermodulation effect. Unlike most state-of-the-art spline-based adaptive learning schemes, the proposed concept allows for complex-valued in- and output signals. The optimization of the model parameters is based on the stochastic gradient descent concept, where the convergence is supported by an appropriate step-size normalization. Additionally, we provide a gain control scheme and enable pipelining in order to facilitate a hardware implementation. An optional input transform improves the performance consistency for correlated sequences. In a realistic interference scenario, the proposed algorithms clearly outperform a state-of-the-art least mean squares variant with comparable complexity, which is specifically tailored to second-order intermodulation distortions. The high flexibility of the spline interpolation allows the spline-based Wiener models to match the performance of the kernel recursive least squares algorithm at less than 0.6 % of the arithmetic operations.

Index Terms—Adaptive learning, intermodulation distortion, LTE, self-interference cancellation, spline interpolation, RF transceivers.

I. INTRODUCTION

POWER efficiency is a key aspect for radio frequency (RF) transceivers in mobile communications equipment in order to ensure sufficient battery life at high data rates. Combined with other metrics, such as cost or area usage, this might lead to a decrease in receiver linearity. Given the low powers of the wanted receive (Rx) signal at the input, these design trade-offs

usually do not cause relevant distortions. However, in in-band full-duplex (IBFD) and frequency-division duplex (FDD) operation, a strong transmit (Tx) leakage in the receiver is unavoidable, leading to a severe deterioration of the Rx signal. When using the predominant direct-conversion receiver architecture, especially even-order intermodulation distortions (IMDs) are a major issue, since they fall directly into the Rx baseband (BB) independent of the Tx carrier frequency [1].

In the Long-Term Evolution Advanced (LTE-A) and 5G New Radio (NR) standards, an important transmission mode is FDD. The resulting separation of the Tx and Rx carriers enables the usage of band-selection filters, in communication transceivers these are typically duplexers. These components provide a limited suppression of the transmitter-to-receiver leakage of about 50–55dB [2]. A higher isolation is not feasible due to disadvantages such as higher cost or increased insertion losses. With Tx powers of up to 27 dBm at the output of the power amplifier (PA), the spectrally shaped leakage still has considerable power at the duplexer Rx port, causing IMD. Besides design changes in the analog front-end (AFE), several BB mitigation techniques exist to attenuate the IMD interference [3]. In this work, we focus on fully digital self-interference cancellation for FDD transceivers, where the interference is replicated based on the known Tx data. While increasing the computational burden in the digital BB compared to mixed-signal solutions, digital self-interference mitigation (DSIM) approaches do not require any changes to the AFE and scale well with smaller technology nodes. Most published approaches target second-order intermodulation products, both, with [4]–[7] and without [8], [9] a frequency-selective leakage path. However, also higher-order products are likely to occur but are rarely covered in literature [10], [11]. One category of suitable algorithms are general learning schemes, such as kernel adaptive filters (KAFs), support-vector machines (SVMs) or artificial neural networks (ANNs) with constant or adaptive activation functions [12]–[17]. Other concepts are adaptive truncated Volterra series and functional link adaptive filters [12], [18], which both utilize limited model knowledge. However, these models result in an unreasonably high number of parameters to be estimated, slowing down adaptation and increasing complexity. The generation of higher-order IMD is best represented by Wiener models, but the required nonlinear estimation is often assumed to be unsuitable for real-time

Manuscript received November 9, 2020; revised February 22, 2021 and April 6, 2021; accepted April 27, 2021. Date of publication April 30, 2021; date of current version July 8, 2021. This work was supported in part by the Austrian Federal Ministry for Digital and Economic Affairs, in part by the National Foundation for Research, Technology and Development, and in part by the Christian Doppler Research Association. The review of this article was coordinated by Dr. Yu Ted Su. (*Corresponding author: Christian Motz.*)

The authors are with the Christian Doppler Laboratory for Digitally Assisted RF Transceivers for Future Mobile Communications, Institute of Signal Processing, Johannes Kepler University Linz, Linz 4040, Austria (e-mail: thomas.paireder@jku.at; christian.motz@jku.at; mario.huemer@jku.at).

Digital Object Identifier 10.1109/TVT.2021.3076843

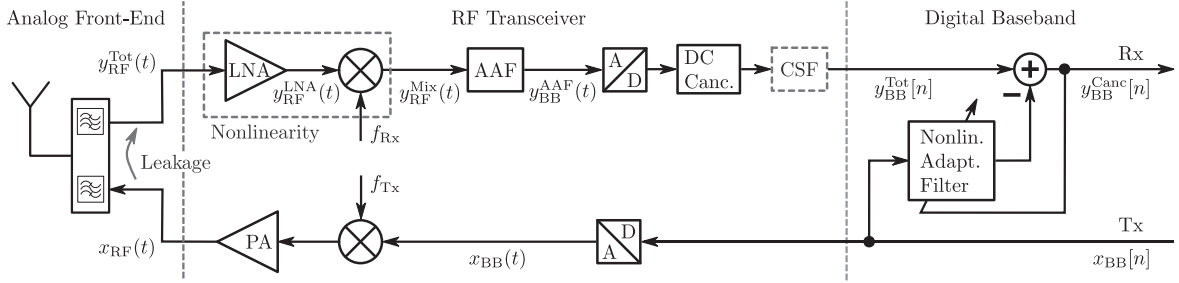


Fig. 1. Block diagram of FDD RF transceiver with digital cancellation of IMDx interference.

implementation on devices with limited computational resources. With the development of spline adaptive filters (SAFs) [19]–[21], low-complex adaptive Wiener models for a wide range of nonlinearities are available. However, most of them target real-valued functions and, thus, are not applicable to the complex-valued IMD cancellation problem. Two exceptions featuring spline interpolation with complex control points are presented in [22], [23]. Though, both show specialized concepts without normalization, which are not suited for the IMD problem.

In this work, we introduce two novel Wiener SAFs, comprised of a complex-valued linear system, a fixed internal nonlinearity and a real- or complex-valued spline output. The latter is achieved by combining two real-valued spline functions, where one models the real and one the imaginary part of the output. Both proposed SAF variants allow for a filtered output, which, for example, can cover a delayed update as it is caused by a pipelined hardware implementation. For the first time, internal clipping is avoided by means of an optimization constraint, that controls the norm of the linear filter weights. Additionally, an alternative limiter for the weight norm with lower complexity is proposed. The stochastic gradient descent (SGD) update used for learning is improved by means of an appropriate novel step-size normalization. Optionally, the learning is augmented with the well-known transform-domain (TD) concept to improve the performance consistency of the algorithms.

This paper is organized as follows: In Section II, we provide an in-depth analysis of the IMD effect, leading to a BB model that is essential for all further considerations. For completeness, in Section III, we give a brief overview of the spline interpolation concept. In Section IV, we derive normalized SGD update equations for the proposed SAFs, discuss possible optimizations and assess the computational complexity. Section V quantifies the IMD cancellation performance of our concepts on a real-world scenario.

II. SELF-INTERFERENCE DUE TO INTERMODULATION DISTORTIONS

Fig. 1 schematically depicts one Tx and one Rx path of FDD transceiver, operating simultaneously on a common antenna. As a consequence of non-ideal Tx-Rx isolation in the analog front-end, the Tx signal leaks into the receiver, where it causes nonlinear distortions that overlay the wanted Rx BB signal.

Note that in case of carrier aggregation, this problem potentially persists for all combinations of Tx and Rx chains.

A partial modeling of this effect is shown in [6], [24], which we use as a foundation to derive the interference components that occur in the Rx BB. Compared to the available literature, we include all important cross-terms of Tx leakage, wanted Rx and noise. This thorough analysis is essential in order to correctly interpret measurement results on the IMD effect. We start with the known digital Tx BB sequence $x_{BB}[n]$, which first passes the digital-to-analog converter (DAC). Next, the analog signal $x_{BB}(t)$ is up-converted to the carrier f_{Tx} . Since we focus on receiver nonlinearities, for our model we assume the up-conversion mixer and the PA to be ideal with a total gain of A_{PA} . This simplification is backed by simulations in Section V, where we include the saturation behavior of the PA. The cancellation performance of the proposed Wiener SAF is not impacted by the PA nonlinearity. After up-conversion and amplification, the resulting Tx RF signal is

$$x_{RF}(t) = A_{PA} \Re \{ x_{BB}(t) e^{j2\pi f_{Tx}t} \}. \quad (1)$$

The leakage path, comprising RF switches, diplexers and the duplexer, is modeled as

$$h_{RF}^{TxL}(t) = 2 \Re \{ h_{BB}^{TxL}(t) e^{j2\pi f_{Tx}t} \} \quad (2)$$

with the equivalent BB impulse response $h_{BB}^{TxL}(t)$. Depending on the context, $*$ denotes the time-continuous or time-discrete convolution. The factor 2 is introduced in the RF domain to compensate for the scaling effect of the convolution later on. With these definitions, the RF leakage signal can be written as:

$$\begin{aligned} y_{RF}^{TxL} &= x_{RF}(t) * h_{RF}^{TxL}(t) \\ &= \frac{A_{PA}}{2} (x_{BB}(t) e^{j2\pi f_{Tx}t} + x_{BB}(t)^* e^{-j2\pi f_{Tx}t}) \\ &\quad * (h_{BB}^{TxL}(t) e^{j2\pi f_{Tx}t} + h_{BB}^{TxL}(t)^* e^{-j2\pi f_{Tx}t}) \\ &= \frac{A_{PA}}{2} (x_{BB}(t) * h_{BB}^{TxL}(t)) e^{j2\pi f_{Tx}t} \\ &\quad + \frac{A_{PA}}{2} (x_{BB}(t)^* * h_{BB}^{TxL}(t)^*) e^{-j2\pi f_{Tx}t} \\ &= \Re \{ y_{BB}^{TxL}(t) e^{j2\pi f_{Tx}t} \}. \end{aligned} \quad (3)$$

$y_{BB}^{TxL}(t) = x_{BB}(t) * \tilde{h}_{BB}^{TxL}(t)$ is the BB equivalent leakage signal. To shorten the notation, we include the Tx gain in the impulse response of the leakage path $\tilde{h}_{BB}^{TxL}(t) = A_{PA} h_{BB}^{TxL}(t)$. Due to the

leakage, the total signal at the receiver input is

$$y_{\text{RF}}^{\text{Tot}}(t) = y_{\text{RF}}^{\text{Rx}}(t) + \eta_{\text{RF}}(t) + y_{\text{RF}}^{\text{TxL}}(t). \quad (4)$$

$y_{\text{RF}}^{\text{Rx}}(t)$ is the desired receive signal at the carrier f_{Rx}

$$y_{\text{RF}}^{\text{Rx}}(t) = \Re \{ y_{\text{BB}}^{\text{Rx}}(t) e^{j2\pi f_{\text{Rx}} t} \} \quad (5)$$

and $\eta_{\text{RF}}(t)$ is additive thermal noise from the antenna. Both components passed the duplexer and are therefore limited to the bandwidth of the selected Long-Term Evolution (LTE)/NR band. However, any noise components that are added after the duplexer might have substantially higher bandwidth. In order to limit the complexity of the model, we neglect any wideband noise in the following and define

$$\eta_{\text{RF}}(t) = \Re \{ \eta_{\text{BB}}(t) e^{j2\pi f_{\text{Rx}} t} \}. \quad (6)$$

Any components up to the quadrature mixer (I/Q mixer) potentially exhibit nonlinear behavior, which is modeled to be concentrated in the low-noise amplifier (LNA) and the mixer.

Based on measurements of the IMD products generated by an integrated complementary metal–oxide–semiconductor (CMOS) receiver we assume a polynomial nonlinearity of degree 3 for modeling the LNA [25], [26]:

$$y_{\text{RF}}^{\text{LNA}}(t) = \alpha_1 y_{\text{RF}}^{\text{Tot}}(t) + \alpha_2 (y_{\text{RF}}^{\text{Tot}}(t))^2 + \alpha_3 (y_{\text{RF}}^{\text{Tot}}(t))^3. \quad (7)$$

The coefficients α_i are real-valued quantities. Inserting $y_{\text{RF}}^{\text{Tot}}(t)$ into the LNA model yields

$$\begin{aligned} y_{\text{RF}}^{\text{LNA}}(t) &= \alpha_1 (y_{\text{RF}}^{\text{Rx}}(t) + \eta_{\text{RF}}(t) + y_{\text{RF}}^{\text{TxL}}(t)) \\ &+ \alpha_2 \left((y_{\text{RF}}^{\text{Rx}}(t))^2 + (\eta_{\text{RF}}(t))^2 + (y_{\text{RF}}^{\text{TxL}}(t))^2 \right. \\ &\quad + 2 y_{\text{RF}}^{\text{Rx}}(t) \eta_{\text{RF}}(t) + 2 y_{\text{RF}}^{\text{Rx}}(t) y_{\text{RF}}^{\text{TxL}}(t) \\ &\quad \left. + 2 \eta_{\text{RF}}(t) y_{\text{RF}}^{\text{TxL}}(t) \right) \\ &+ \alpha_3 \left((y_{\text{RF}}^{\text{Rx}}(t))^3 + (\eta_{\text{RF}}(t))^3 + (y_{\text{RF}}^{\text{TxL}}(t))^3 \right. \\ &\quad + 3 (y_{\text{RF}}^{\text{Rx}}(t))^2 \eta_{\text{RF}}(t) + 3 (y_{\text{RF}}^{\text{Rx}}(t))^2 y_{\text{RF}}^{\text{TxL}}(t) \\ &\quad + 3 y_{\text{RF}}^{\text{Rx}}(t) (\eta_{\text{RF}}(t))^2 + 3 y_{\text{RF}}^{\text{Rx}}(t) (y_{\text{RF}}^{\text{TxL}}(t))^2 \\ &\quad + 3 (\eta_{\text{RF}}(t))^2 y_{\text{RF}}^{\text{TxL}}(t) + 3 \eta_{\text{RF}}(t) (y_{\text{RF}}^{\text{TxL}}(t))^2 \\ &\quad \left. + 6 y_{\text{RF}}^{\text{Rx}}(t) \eta_{\text{RF}}(t) y_{\text{RF}}^{\text{TxL}}(t) \right). \quad (8) \end{aligned}$$

Due to the large number of terms, we refrain from inserting the equivalent BB definitions of $y_{\text{RF}}^{\text{Rx}}(t)$, $\eta_{\text{RF}}(t)$ and $y_{\text{RF}}^{\text{TxL}}(t)$.

Instead, in Table I, we provide an overview of the spectral contents of all terms. Due to space constraints, we dropped the time indices of the signals. While the LNA model does not include an explicit direct current (DC) component, signal components around DC occur due to intermodulation.

Following the LNA, the I/Q mixer performs a direct down-conversion to the complex BB using the local oscillator (LO) frequency f_{Rx} . The model of the mixer covers a DC feed-through, the desired down-conversion and an RF-to-LO terminal coupling, quantified by the coefficients β_0 , β_1 and β_2 , respectively.

TABLE I

CARRIER FREQUENCIES OF SIGNAL COMPONENTS AFTER LNA NONLINEARITY

Summands	Contained Carriers
$y_{\text{RF}}^{\text{Rx}}, \eta_{\text{RF}}$	$\pm f_{\text{Rx}}$
$y_{\text{RF}}^{\text{TxL}}$	$\pm f_{\text{Tx}}$
$(y_{\text{RF}}^{\text{Rx}})^2, (\eta_{\text{RF}})^2, y_{\text{RF}}^{\text{Rx}} \eta_{\text{RF}}$	$0, \pm 2 f_{\text{Rx}}$
$(y_{\text{RF}}^{\text{TxL}})^2$	$0, \pm 2 f_{\text{Tx}}$
$y_{\text{RF}}^{\text{Rx}} y_{\text{RF}}^{\text{TxL}}, \eta_{\text{RF}} y_{\text{RF}}^{\text{TxL}}$	$\pm f_{\text{Rx}} \pm f_{\text{Tx}}$
$(y_{\text{RF}}^{\text{Rx}})^3, (\eta_{\text{RF}})^3, (y_{\text{RF}}^{\text{Rx}})^2 \eta_{\text{RF}}, y_{\text{RF}}^{\text{Rx}} (\eta_{\text{RF}})^2$	$\pm f_{\text{Rx}}, \pm 3 f_{\text{Rx}}$
$(y_{\text{RF}}^{\text{TxL}})^3$	$\pm f_{\text{Tx}}, \pm 3 f_{\text{Tx}}$
$(y_{\text{RF}}^{\text{Rx}})^2 y_{\text{RF}}^{\text{TxL}}, (\eta_{\text{RF}})^2 y_{\text{RF}}^{\text{TxL}}, y_{\text{RF}}^{\text{Rx}} \eta_{\text{RF}} y_{\text{RF}}^{\text{TxL}}$	$\pm f_{\text{Tx}}, \pm 2 f_{\text{Rx}} \pm f_{\text{Tx}}$
$y_{\text{RF}}^{\text{Rx}} (y_{\text{RF}}^{\text{TxL}})^2, \eta_{\text{RF}} (y_{\text{RF}}^{\text{TxL}})^2$	$\pm f_{\text{Rx}}, \pm f_{\text{Rx}} \pm 2 f_{\text{Tx}}$

Hence, we have the following model at the mixer output:

$$y_{\text{RF}}^{\text{Mix}}(t) = \beta_0 y_{\text{RF}}^{\text{LNA}}(t) + \beta_1 y_{\text{RF}}^{\text{LNA}}(t) e^{-j2\pi f_{\text{Rx}} t} + \beta_2 (y_{\text{RF}}^{\text{LNA}}(t))^2. \quad (9)$$

β_0 and β_2 are possibly complex values, since the mixer is implemented using two independent branches. β_1 is assumed to be real-valued, which reflects a balanced down-conversion of the I and Q components of the wanted signal. This could be ensured by design or calibration. Besides the down-conversion of the wanted BB signal and several distortion terms at BB, the mixer also outputs spectral components at multiples of f_{Tx} and f_{Rx} . We assume that all components at additive combinations of these frequencies are suppressed by the anti-aliasing filter (AAF). A high Tx power and thus, strong leakage usually coincides with a low Rx power [27]. Therefore, we do not list all individual, usually weak intermodulation products of $y_{\text{BB}}^{\text{Rx}}(t)$ and $\eta_{\text{BB}}(t)$ but consider them as additive noise $\eta_{\text{BB}}^{\text{IMD}}(t)$.

Components at subtractive combinations of the carriers, i. e. located at the offsets $f_{\Delta} = \pm(a f_{\text{Tx}} - b f_{\text{Rx}})$, may fall within the AAF bandwidth. In case of $|f_{\Delta}| \leq 6 \text{BW}_{\text{Rx}}$ they potentially overlap with the main IMD components centered at DC. Here, a, b are positive integers and BW_{Rx} is the Rx bandwidth. The factor 6 is related to the combined degree of the nonlinearities of the LNA and the mixer. We assume that the bandwidth of $y_{\text{BB}}^{\text{TxL}}(t)$ is less than BW_{Rx} . One example is given by $f_{\text{Rx}} \approx 2 f_{\text{Tx}}$, which can occur in case of carrier aggregation (CA). Unlike the IMD components centered at DC, the intermodulation products around f_{Δ} are of even and odd order. The most significant components are of the form $c_k (y_{\text{BB}}^{\text{TxL}}(t))^p (y_{\text{BB}}^{\text{TxL}}(t))^q e^{j2\pi f_{\Delta} t}$ with $p + q \leq 6$. However, this interference class is not covered in this work. As a result, the subtractive combinations of the carriers are not considered in the estimation process but add to the total noise floor.

At the output of the AAF, our assumptions lead to

$$\begin{aligned} y_{\text{BB}}^{\text{AAF}} &= \frac{1}{2} \alpha_1 \beta_1 y_{\text{BB}}^{\text{Rx}} + \frac{1}{2} \alpha_1 \beta_1 \eta_{\text{BB}} + \eta_{\text{BB}}^{\text{IMD}} \\ &+ \left(\frac{1}{2} \alpha_1^2 \beta_2 + \frac{1}{2} \alpha_2 \beta_0 \right) |y_{\text{BB}}^{\text{TxL}}|^2 \\ &+ \left(\frac{3}{4} \alpha_1 \alpha_3 \beta_2 + \frac{3}{8} \alpha_2^2 \beta_2 \right) |y_{\text{BB}}^{\text{TxL}}|^4 \end{aligned}$$

$$\begin{aligned}
& + \frac{5}{16} \alpha_3^2 \beta_2 |y_{\text{BB}}^{\text{TxL}}|^6 \\
& + \frac{3}{4} \alpha_3 \beta_1 (y_{\text{BB}}^{\text{Rx}} + \eta_{\text{BB}}) |y_{\text{BB}}^{\text{TxL}}|^2 \\
& + \left(3\alpha_1 \alpha_3 \beta_2 + \frac{3}{2} \alpha_2^2 \beta_2 \right) \left(|y_{\text{BB}}^{\text{Rx}}|^2 + |\eta_{\text{BB}}|^2 \right) |y_{\text{BB}}^{\text{TxL}}|^2 \\
& + \frac{45}{16} \alpha_3^2 \beta_2 \left(|y_{\text{BB}}^{\text{Rx}}|^4 + |\eta_{\text{BB}}|^4 \right) |y_{\text{BB}}^{\text{TxL}}|^2 \\
& + \frac{45}{16} \alpha_3^2 \beta_2 \left(|y_{\text{BB}}^{\text{Rx}}|^2 + |\eta_{\text{BB}}|^2 \right) |y_{\text{BB}}^{\text{TxL}}|^4, \quad (10)
\end{aligned}$$

where we again dropped the time indices. The first term is the down-converted Rx signal and the second and third terms are noise caused by various sources. The pure intermodulation products of the leakage signal are represented by the terms 4 to 6. These components are targeted by the cancellation approaches presented later in this work. The remaining terms of $y_{\text{BB}}^{\text{AAF}}(t)$ are intermodulation products of the leakage with the Rx signal or noise. At very high transmit powers, these products could deteriorate the receiver signal-to-interference-plus-noise ratio (SINR). However, due to their dependence on both, $y_{\text{BB}}^{\text{Rx}}(t)$ and $y_{\text{BB}}^{\text{TxL}}(t)$, these terms are difficult to cancel with low to medium hardware complexity. Consequently, we consider them as noise and summarize the signal at the AAF output as

$$y_{\text{BB}}^{\text{AAF}}(t) = A_{\text{lin}} y_{\text{BB}}^{\text{Rx}}(t) + \eta_{\text{BB}}^{\text{AAF}}(t) + \sum_{k=1}^3 \gamma_k |y_{\text{BB}}^{\text{TxL}}(t)|^{2k} \quad (11)$$

with the linear gain $A_{\text{lin}} = \frac{1}{2} \alpha_1 \beta_1$, the combined noise $\eta_{\text{BB}}^{\text{AAF}}(t)$ and the combined coefficients γ_k .

Direct-conversion receivers usually suffer from spurious DC components, which, for example, could saturate the analog-to-digital converter (ADC). This issue is solved by employing a DC cancellation stage. Independent of its actual position in the receive chain, we model this stage as a notch filter $h_{\text{DC}}[n]$ in the digital domain directly following the (ideal) ADC. Additionally, the digital BB signal is commonly limited to the channel bandwidth by the channel-select filter (CSF). Since this filter might cause issues for the digital IMD cancellation, we propose to place the CSF after the DSIM point. The final digital BB model for all following considerations is

$$\begin{aligned}
y_{\text{BB}}^{\text{Tot}}[n] &= \underbrace{A_{\text{lin}} y_{\text{BB}}^{\text{Rx}}[n] * h_{\text{DC}}[n]}_{\tilde{y}_{\text{BB}}^{\text{Rx}}[n]} + \tilde{\eta}_{\text{BB}}[n] \\
&+ \underbrace{\sum_{k=1}^3 \gamma_k |x_{\text{BB}}[n] * \tilde{h}_{\text{BB}}^{\text{TxL}}[n]|^{2k} * h_{\text{DC}}[n]}_{y_{\text{BB}}^{\text{IMD}}[n]}. \quad (12)
\end{aligned}$$

with the filtered Rx BB signal $\tilde{y}_{\text{BB}}^{\text{Rx}}[n]$ and the BB interference $y_{\text{BB}}^{\text{IMD}}[n]$. Note that $\tilde{\eta}_{\text{BB}}[n]$ includes the quantization noise of the ADC $\eta_{\text{BB}}^{\text{ADC}}[n]$:

$$\tilde{\eta}_{\text{BB}}[n] = (\eta_{\text{BB}}^{\text{AAF}}[n] + \eta_{\text{BB}}^{\text{ADC}}[n]) * h_{\text{DC}}[n]. \quad (13)$$

Based on $x_{\text{BB}}[n]$ and $y_{\text{BB}}^{\text{Tot}}[n]$, the DSIM algorithm shall estimate the interference and subtract it from the Rx BB signal. When

denoting the replicated interference with $\hat{y}_{\text{BB}}^{\text{IMD}}[n]$, the enhanced Rx signal is given by $y_{\text{BB}}^{\text{Canc}}[n] = y_{\text{BB}}^{\text{Tot}}[n] - \hat{y}_{\text{BB}}^{\text{IMD}}[n]$.

III. BASICS OF SPLINE INTERPOLATION

In this work, we introduce two adaptive algorithms for simultaneous digital cancellation of multiple even-order IMD products, which both rely on spline interpolation to replicate the nonlinear function present in $y_{\text{BB}}^{\text{IMD}}[n]$. As a basis, we summarize the most important properties and definitions of the well-known spline interpolation method, which are then utilized in the derivations in Section IV.

A. B-Splines

In many applications, like numerical simulations or computer graphics, it is desired to approximate general nonlinear functions by simpler expressions in order to enable efficient evaluation and easier analysis. Moreover, discrete data series frequently have to be interpolated to obtain intermediate values or to enable analytic manipulation. A straight-forward approximation method for both types of applications is the use of a single polynomial over the whole domain of the target function. While Weierstrass' theorem states that this is generally possible to any desired accuracy, without further precautions oscillations occur (Runge's phenomenon) [28], [29]. This effect can be limited by solely using polynomials of low degree. The natural consequence is to employ piecewise polynomial functions, where the approximation accuracy is defined by the number of sections used.

The boundaries of the sections are defined by $M_{\text{sp}} = N_{\text{sp}} + Q_{\text{sp}}$ knots $[x_0, x_1, \dots, x_{M_{\text{sp}}-1}]$, which are sorted in a monotonically increasing order. N_{sp} is the number of points to interpolate and Q_{sp} is the order of the spline curve $S(x)$. The curve is composed of polynomial sections of degree $Q_{\text{sp}} - 1$. Depending on the continuity across the knots, the interpolation properties and the support of the base functions, different classes of spline curves are distinguished. We first focus on B-splines [29], [30], which provide $C^{Q_{\text{sp}}-2}$ smoothness and minimal support, but, in general¹, the curve does not exactly interpolate (i. e. pass through) its control points q_i [31]. $S(x)$ is a linear combination of the base functions $B_{Q_{\text{sp}},i}(x)$ weighted by N_{sp} control points q_i :

$$S(x) = \sum_{i=0}^{N_{\text{sp}}-1} q_i B_{Q_{\text{sp}},i}(x), \quad x_{Q_{\text{sp}}-1} \leq x < x_{N_{\text{sp}}}. \quad (14)$$

The domain of the open spline $S(x)$ is limited compared to the knot vector, because the first $Q_{\text{sp}} - 1$ and the last $Q_{\text{sp}} - 1$ intervals do not have full support. In order to limit the computational effort, we only cover uniform splines, where the lengths of all segments are identical, i. e. $\Delta x = x_{m+1} - x_m \forall m$. In this case, the control points q_i are located at

$$\bar{x}_i = x_i + \frac{Q_{\text{sp}}}{2} \Delta x, \quad i = 0, \dots, N_{\text{sp}} - 1. \quad (15)$$

¹In case of $Q_{\text{sp}} = 1$ (step function) and $Q_{\text{sp}} = 2$ (linear interpolation), the B-spline curve exactly passes through its control points.

The $B_{Q_{\text{sp}},i}(x)$ can be obtained by the Cox-de Boor recursion [29]. For illustrative purposes, we provide the explicit forms for $Q_{\text{sp}} = \{1, 2, 3\}$:

$$B_{1,i}(x) = \begin{cases} 1 & x_i \leq x < x_{i+1} \\ 0 & \text{else} \end{cases} \quad (16)$$

$$B_{2,i}(x) = \begin{cases} \frac{1}{\Delta x} (x - x_i) & x_i \leq x < x_{i+1} \\ \frac{1}{\Delta x} (x_{i+2} - x) & x_{i+1} \leq x < x_{i+2} \\ 0 & \text{else} \end{cases} \quad (17)$$

$$B_{3,i}(x) = \begin{cases} \frac{1}{2(\Delta x)^2} (x - x_i)^2 & x_i \leq x < x_{i+1} \\ \frac{1}{2(\Delta x)^2} ((x - x_i)(x_{i+2} - x) + (x_{i+3} - x)(x - x_{i+1})) & x_{i+1} \leq x < x_{i+2} \\ \frac{1}{2(\Delta x)^2} (x_{i+3} - x)^2 & x_{i+2} \leq x < x_{i+3} \\ 0 & \text{else} \end{cases} \quad (18)$$

Since the $B_{Q_{\text{sp}},i}(x)$ are non-zero only in the interval $[x_i, x_{i+Q_{\text{sp}}})$, a single control point q_i impacts a limited section of the curve. This important property enables an independent adjustment of the q_i by an iterative algorithm. Additionally, the base functions are normalized to form a partition of unity, i. e.

$$\sum_{i=0}^{Q_{\text{sp}}-1} B_{Q_{\text{sp}},m-i}(x) = 1, \quad x_m \leq x < x_{m+1}. \quad (19)$$

In order to simplify the evaluation of a uniform spline function, we replace the global input x by the normalized input $\nu \in [0, 1)$ relative to the lower limit of the interval ι :

$$x = \nu \Delta x + \iota \Delta x + x_0 \quad (20)$$

$$\nu = \frac{x}{\Delta x} - \left\lfloor \frac{x}{\Delta x} \right\rfloor, \quad \iota = \left\lfloor \frac{x - x_0}{\Delta x} \right\rfloor. \quad (21)$$

This allows to rewrite $S(x)$ into a matrix-vector product, as exemplarily shown for $Q_{\text{sp}} = 3$:

$$\begin{aligned} S(\nu, \iota) &= \sum_{i=0}^2 q_{\iota-i} B_{3,\iota-i}(\nu \Delta x + \iota \Delta x + x_0) \\ &= \frac{q_{\iota-2}}{2(\Delta x)^2} (\Delta x)^2 (1 - \nu)^2 \\ &\quad + \frac{q_{\iota-1}}{2(\Delta x)^2} (\Delta x)^2 ((\nu + 1)(1 - \nu) + (2 - \nu)\nu) \\ &\quad + \frac{q_{\iota}}{2(\Delta x)^2} (\Delta x)^2 \nu^2 \\ &= \begin{bmatrix} \nu^2 & \nu & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ -1 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} q_{\iota-2} \\ q_{\iota-1} \\ q_{\iota} \end{bmatrix} \\ &= \mathbf{v}_3^{\text{T}} \mathbf{B}_3^{\text{sp}} \mathbf{q}_{3,\iota}. \end{aligned} \quad (22)$$

For arbitrary spline order, the input vector $\mathbf{v}_{Q_{\text{sp}}}$ and the control point vector $\mathbf{q}_{Q_{\text{sp}},\iota}$ are

$$\mathbf{v}_{Q_{\text{sp}}} = \left[\nu^{Q_{\text{sp}}-1}, \nu^{Q_{\text{sp}}-2}, \dots, 1 \right]^{\text{T}} \quad (23)$$

$$\mathbf{q}_{Q_{\text{sp}},\iota} = \left[q_{\iota-Q_{\text{sp}}+1}, q_{\iota-Q_{\text{sp}}+2}, \dots, q_{\iota} \right]^{\text{T}}. \quad (24)$$

While the formulation (22) is useful for derivations, in an implementation it is advantageous to calculate (22) using Horner's method, which reduces the required number of multiplications. The B-spline basis matrices $\mathbf{B}_{Q_{\text{sp}}}^{\text{B}}$ for Q_{sp} up to 4, i. e. cubic interpolation, are

$$\mathbf{B}_1^{\text{B}} := 1, \quad \mathbf{B}_2^{\text{B}} := \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}, \quad (25)$$

$$\mathbf{B}_3^{\text{B}} := \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ -1 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}, \quad \mathbf{B}_4^{\text{B}} := \begin{bmatrix} -\frac{1}{6} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{6} \\ \frac{1}{2} & -1 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 \end{bmatrix}. \quad (26)$$

The normalization of the basis functions, shown in (19), is directly resembled in the matrices $\mathbf{B}_{Q_{\text{sp}}}$ since we have

$$\mathbf{v}_{Q_{\text{sp}}}^{\text{T}} \mathbf{B}_{Q_{\text{sp}}}^{\text{sp}} \mathbf{1}_{Q_{\text{sp}} \times 1} = \mathbf{v}_{Q_{\text{sp}}}^{\text{T}} \mathbf{e}_{Q_{\text{sp}}} = 1, \quad (27)$$

where $\mathbf{q}_{Q_{\text{sp}},\iota} = \mathbf{1}_{Q_{\text{sp}} \times 1}$ ensures equal weighting of the base functions, $\mathbf{1}_{K \times 1}$ is the all-ones vector and \mathbf{e}_k is the k -th unit vector.

B. Catmull-Rom Splines

As a trade-off between approximation accuracy and the risk of overfitting, cubic splines are preferred in many applications. Among all cubic splines, B-splines offer the highest degree of smoothness, but they do not exactly interpolate its control points. Conversely, this means that obtaining the control points of a well approximating spline curve based on a given set of function points necessitates to solve a system of equations. In adaptive learning, however, a simple way to initialize the control points to a given function is beneficial. Catmull-Rom (CR) splines feature the desired interpolation property at the cost of lower smoothness, which is only C^1 compared to C^2 for cubic B-splines [31]. The basis matrix for CR-splines is given by

$$\mathbf{B}^{\text{CR}} := \begin{bmatrix} -\frac{1}{2} & \frac{3}{2} & -\frac{3}{2} & \frac{1}{2} \\ 1 & -\frac{5}{2} & 2 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (28)$$

Note that CR-splines are limited to order 4.

IV. SPLINE-BASED CANCELLATION OF EVEN-ORDER IMD SELF-INTERFERENCE

An adaptive algorithm used for simultaneous digital cancellation of multiple even-order IMD products has to estimate the leakage path $\tilde{h}_{\text{BB}}^{\text{TxL}}[n]$ and the coefficients of the receiver nonlinearity γ_k . We aim to solve this problem by means of a linear adaptive filter followed by a nonlinear adaptive function based on splines. Hence, in the following we present two novel complex Wiener SAF algorithms, which are tailored to the IMD effect.

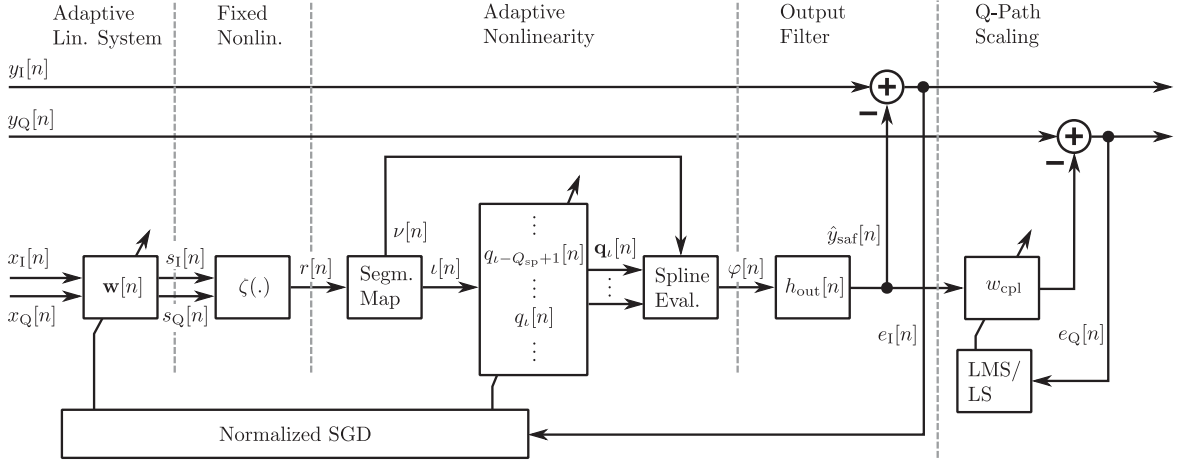


Fig. 2. Block diagram of adaptive nonlinear DSIM structure with real output and subsequent single-tap scaling for Q-path.

Besides the basic algorithm, we cover implementation-related aspects, such as pipelining, and use a constraint optimization technique to avoid internal clipping, an inherent issue of SAFs.

A. Wiener-SAF With Complex Input, Internal Fixed Nonlinearity and Fixed Output Filter

1) *Basic Algorithm:* We propose a Wiener model as shown in the block diagram in Fig. 2, which employs a fixed nonlinear function $\zeta: \mathbb{C} \rightarrow \mathbb{R}$ that transforms the complex filter output $s[n]$ into a real quantity $r[n]$. A suitable function for the IMD problem is $\zeta(s) = |s|^2$. Alternatively, $\zeta(s) = |s|$ could be used, which features an advantageous amplitude distribution. The signal $r[n]$ is the input of the spline function, which is capable of approximating a wide range of nonlinearities. Obviously, this includes the IMD products of the form $| \cdot |^{2k}$ and any linear combination of these terms. The latter case matches the interference $y_{\text{BB}}^{\text{IMD}}[n]$. In the following, we refer to the whole structure as complex-input Wiener SAF (CI-WSAF). Its output $\hat{y}_{\text{saf}}[n] \in \mathbb{R}$ is given by

$$\begin{aligned} \hat{y}_{\text{saf}}[n] &= \underbrace{\left(\mathbf{v}_{Q_{\text{sp}}}[n]^T \mathbf{B}_{Q_{\text{sp}}}^{\text{sp}} \mathbf{q}_{Q_{\text{sp}}, \iota}[n-1] \right)}_{\varphi[n]} * h_{\text{out}}[n] \\ &= \sum_{k=0}^{Q_{\text{out}}-1} \varphi[n-k] h_{\text{out}}[k] \end{aligned} \quad (29)$$

using the definitions

$$r[n] = \zeta(s[n]), \quad s[n] = \mathbf{w}[n-1]^T \mathbf{x}[n] \quad (30)$$

$$\nu[n] = \frac{r[n]}{\Delta r} - \lfloor \frac{r[n]}{\Delta r} \rfloor, \quad \iota[n] = \lfloor \frac{r[n] - r_0}{\Delta r} \rfloor. \quad (31)$$

$\hat{y}_{\text{saf}}[n]$ is an approximation of the (noisy) desired sequence $y[n]$, which is $y_{\text{BB},1}^{\text{Tot}}[n]$ in case of IMD cancellation. The corresponding input $x[n]$ of the algorithm is the known Tx BB stream $x_{\text{BB}}[n]$. The real impulse response $h_{\text{out}}[n]$ allows to apply a known filter to the output of the spline function. Depending on the setup, this filter could be a CSF or a pure delay. The latter is important to model pipelining stages in the output computation, which help to achieve the desired operating frequency. The length of

the adaptive filter $\mathbf{w}[n]$ is Q_{lin} , the length of the output filter $h_{\text{out}}[n]$ is Q_{out} and the spline order is Q_{sp} . We use uniform knots with a spacing of Δr and a lower limit of r_0 . When calculating $\iota[n]$, it is important to realize that its value must be in the range $[Q_{\text{sp}} - 1, N_{\text{sp}} - 1]$, where N_{sp} is the number of control points. For any values outside of this range, the correct spline output cannot be calculated due to an index underrun or overflow in $\mathbf{q}_{Q_{\text{sp}}, \iota}$. The vectors $\mathbf{v}_{Q_{\text{sp}}}[n]$, $\mathbf{q}_{Q_{\text{sp}}, \iota}[m]$, $\mathbf{w}[n]$ and $\mathbf{x}[n]$ are given by

$$\mathbf{v}_{Q_{\text{sp}}}[n] = \left[(\nu[n])^{Q_{\text{sp}}-1}, (\nu[n])^{Q_{\text{sp}}-2}, \dots, 1 \right]^T \quad (32)$$

$$\mathbf{q}_{Q_{\text{sp}}, \iota}[m] = \left[q_{\iota[m]-Q_{\text{sp}}+1}[m], q_{\iota[m]-Q_{\text{sp}}+2}[m], \dots, q_{\iota[m]}[m] \right]^T \quad (33)$$

$$\mathbf{w}[n] = \left[w_0[n], w_1[n], \dots, w_{Q_{\text{lin}}-1}[n] \right]^T \quad (34)$$

$$\mathbf{x}[n] = \left[x[n], x[n-1], \dots, x[n-Q_{\text{lin}}+1] \right]^T. \quad (35)$$

Note the difference in the time indices between the segment index $\iota[m]$ and the control points $q_{\iota}[n]$. The $q_{\iota}[n]$ and thus also $\hat{y}[n]$ are real-valued. For the following derivations, we also define the vector $\mathbf{q}[n] = [q_0[n], q_1[n], \dots, q_{N_{\text{sp}}-1}[n]]^T$ comprising all control points, and use the indexing scheme $q_{\iota}[n] = [\mathbf{q}[n]]_{\iota}$.

We propose to estimate the parameters \mathbf{w} and \mathbf{q} by means of SGD method. The corresponding cost function is the instantaneous squared estimation error, augmented by an additional penalty using the p -norm of the weights:

$$J[n] = |e[n]|^2 + \epsilon \underbrace{\left(\|\mathbf{w}[n-1]\|_p^p - \rho_w \right)^2}_{c_w[n]} \quad (36)$$

$$e[n] = y[n] - \hat{y}_{\text{saf}}[n]. \quad (37)$$

ρ_w is the target value for $\|\mathbf{w}[n]\|_p^p$ and ϵ is a weighting factor. The norm constrained SGD concept is known from sparse estimation, where, for instance, an approximation of the ℓ_0 norm is employed [32]. In the SAF application, the goal is to avoid any gain ambiguity between the weights and the control points on the one hand and to ensure a limited number range at the input of

the spline function on the other hand. This avoids saturation of the signal $r[n]$, which could lead to instability of the algorithm. From finite impulse response (FIR) filter implementations in fixed-point arithmetic it is known that the ℓ_1 norm of the weights poses a conservative bound on the output range for a given input number format [33]. A computationally less expensive approximation would be the ℓ_2 norm, where it holds that $\|\mathbf{w}\|_2 \leq \|\mathbf{w}\|_1$. Thus, the ℓ_2 norm is less restrictive and does not guarantee that overflows are eliminated completely. In order to minimize the cost function, we require the gradients of $J[n]$ with respect to the parameters.

Since the differentiation is not directly possible due to the time-delays caused by the convolution with $h_{\text{out}}[n]$, we introduce approximations, where the targeted parameter is assumed to be time-independent. Consequently, the approximated gradient with respect to \mathbf{q} is

$$\begin{aligned} \tilde{\mathbf{g}}_{\mathbf{q}}[n]^\top &= \frac{\partial J[n]}{\partial e[n]} \frac{\partial e[n]|_{\mathbf{q}[\cdot]=\mathbf{q}}}{\partial \mathbf{q}} \\ &= 2 e[n] \tilde{\mathbf{g}}_{e,\mathbf{q}}[n]^\top. \end{aligned} \quad (38)$$

Throughout this work, any derivative of the form $\partial f(\mathbf{x})/\partial \mathbf{x}$ shall be a row vector when \mathbf{x} is a column vector [34]. This allows for a straight-forward application of the chain rule. When evaluating the partial derivative $\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]$ we yield

$$\tilde{\mathbf{g}}_{e,\mathbf{q}}[n] = - \sum_{k=0}^{Q_{\text{out}}-1} h_{\text{out}}[k] \begin{bmatrix} \mathbf{0}_{(\iota[n-k]-Q_{\text{sp}}+1) \times 1} \\ (\mathbf{B}_{Q_{\text{sp}}}^{\text{sp}})^\top \mathbf{v}_{Q_{\text{sp}}}[n-k] \\ \mathbf{0}_{(N_{\text{sp}}-\iota[n-k]-1) \times 1} \end{bmatrix}, \quad (39)$$

where the all-zero vectors $\mathbf{0}_{K \times 1}$ are used to appropriately place the derivative of the spline output depending on the interval index at the time step $n-k$. Since the evaluation of (39) might be computationally expensive depending on Q_{out} , a possible approximation is to neglect small $h_{\text{out}}[k]$ or to replace the filter with its gain and group delay. The approximated gradient with respect to \mathbf{w} is obtained in a similar manner

$$\begin{aligned} \tilde{\mathbf{g}}_{\mathbf{w}}[n]^\top &= \frac{\partial J[n]}{\partial e[n]} \frac{\partial e[n]|_{\mathbf{w}[\cdot]=\mathbf{w}}}{\partial \mathbf{w}^*} + \frac{\partial J[n]}{\partial c_{\mathbf{w}}[n]} \frac{c_{\mathbf{w}}[n]}{\partial \mathbf{w}[n-1]} \\ &= 2 e[n] \tilde{\mathbf{g}}_{e,\mathbf{w}}[n]^\top + \epsilon \tilde{\mathbf{g}}_{c,\mathbf{w}}[n]^\top \end{aligned} \quad (40)$$

using the $\mathbb{C}\mathbb{R}$ (or Wirtinger) calculus [35], which conveniently guarantees the correct direction of the gradient for both, the real and the imaginary part of \mathbf{w} . The substitution of $\mathbf{w}[\cdot]$ shall affect only the spline input $\nu[\cdot]$, but not the spline segment indices $\iota[\cdot]$. We assume that $\zeta(s, s^*)$ and $c_{\mathbf{w}}(\mathbf{w}, \mathbf{w}^*)$ fulfill Brandwood's analyticity condition [36], thereby simplifying the $\mathbb{C}\mathbb{R}$ derivatives. The term $\tilde{\mathbf{g}}_{e,\mathbf{w}}[n]$ evaluates to

$$\begin{aligned} \tilde{\mathbf{g}}_{e,\mathbf{w}}[n] &= \sum_{k=0}^{Q_{\text{out}}-1} \frac{\partial e[n]}{\partial \mathbf{v}_{Q_{\text{sp}}}[n-k]} \frac{\partial \mathbf{v}_{Q_{\text{sp}}}[n-k]}{\partial \nu[n-k]} \\ &\quad \frac{\partial \nu[n-k]}{\partial r[n-k]} \frac{\partial r[n-k]}{\partial s[n-k]^*} \frac{\partial s[n-k]^*|_{\mathbf{w}[\cdot]=\mathbf{w}}}{\partial \mathbf{w}^{\text{H}}} \end{aligned}$$

$$\begin{aligned} &= -\frac{1}{\Delta r} \sum_{k=0}^{Q_{\text{out}}-1} h_{\text{out}}[k] \mathbf{v}'_{Q_{\text{sp}}}[n-k]^\top \mathbf{B}_{Q_{\text{sp}}}^{\text{sp}} \mathbf{q}_{Q_{\text{sp}},\iota[n-k]}[n-1-k] \\ &\quad \zeta'(s[n-k])^* \mathbf{x}[n-k]^* \end{aligned} \quad (41)$$

with $\zeta'(s) = \partial \zeta(s)/\partial s$ and

$$\begin{aligned} \mathbf{v}'_{Q_{\text{sp}}}[n] &= \\ &= \left[(Q_{\text{sp}}-1) (\nu[n])^{Q_{\text{sp}}-2}, (Q_{\text{sp}}-2) (\nu[n])^{Q_{\text{sp}}-3}, \dots, 0 \right]^\top. \end{aligned} \quad (42)$$

When using $\zeta(s) = |s|^2$, the $\mathbb{C}\mathbb{R}$ derivative of the fixed nonlinearity is $\zeta'(s) = s^*$. Within one spline segment, the index $\iota[\cdot]$ is constant with respect to \mathbf{w} due to the floor function. Again, small values of $h_{\text{out}}[k]$ could be neglected to reduce the complexity of evaluating $\tilde{\mathbf{g}}_{e,\mathbf{w}}[n]$. The derivative of the norm constraint $\tilde{\mathbf{g}}_{c,\mathbf{w}}[n]$ depends on the chosen norm p :

$$\tilde{\mathbf{g}}_{c,\mathbf{w}}[n] = \begin{cases} (\|\mathbf{w}[n-1]\|_1 - \rho_w) & p = 1 \\ \left[\frac{w_k[n-1]}{|w_k[n-1]|} \right]_{k=0,\dots,Q_{\text{in}}-1} & \\ 2 \left(\|\mathbf{w}[n-1]\|_2^2 - \rho_w \right) \mathbf{w}[n-1] & p = 2 \end{cases} \quad (43)$$

with $w_k = [\mathbf{w}]_k$. In an implementation it might be beneficial to use the equivalence $w_k/|w_k| = e^{j \arg(w_k)}$. Combining the above results, the update equations for iterative optimization of the parameters are given by

$$\mathbf{q}[n] = \mathbf{q}[n-1] - 2 \tau \mu[n] e[n] \tilde{\mathbf{g}}_{e,\mathbf{q}}[n] \quad (44)$$

$$\mathbf{w}[n] = \mathbf{w}[n-1] - \mu[n] (2 e[n] \tilde{\mathbf{g}}_{e,\mathbf{w}}[n] + \epsilon \tilde{\mathbf{g}}_{c,\mathbf{w}}[n]) \quad (45)$$

using the time-dependent non-negative step-size $\mu[n] \in \mathbb{R}_{\geq 0}$ and the non-negative coupling factor $\tau \in \mathbb{R}_{\geq 0}$.

2) *Step-Size Normalization*: The standard SGD approach is difficult to tune, since the effective learning rate depends on the dynamics of the involved signals. Thus, commonly a normalized variant is used. One method to derive the normalization is to consider a partial Taylor series expansion of the error signal [20], [37], which is truncated after the linear term

$$\begin{aligned} e[n+1] &\approx e[n] + \frac{\partial e[n]|_{\mathbf{q}[\cdot]=\mathbf{q}}}{\partial \mathbf{q}} \Delta \mathbf{q}[n] \\ &\quad + \frac{\partial e[n]|_{\mathbf{w}[\cdot]=\mathbf{w}}}{\partial \mathbf{w}} \Delta \mathbf{w}[n] + \frac{\partial e[n]|_{\mathbf{w}[\cdot]=\mathbf{w}}}{\partial \mathbf{w}^*} \Delta \mathbf{w}[n]^* \\ &\approx e[n] - 2 \tau \mu[n] e[n] \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2 \\ &\quad - 4 \mu[n] e[n] \|\tilde{\mathbf{g}}_{e,\mathbf{w}}[n]\|_2^2. \end{aligned} \quad (46)$$

$\Delta \mathbf{q}[n]$ and $\Delta \mathbf{w}[n]$ are the parameter changes from time step $n-1$ to n . Assuming ϵ to be small, we neglect the weight norm constraint in the normalization. Based on the Taylor expansion, it has been shown that the adaptation performance of SGD algorithms is improved if the step-size is chosen to fulfill the following condition:

$$\begin{aligned}
|e[n]| &\geq |e[n+1]| \\
|e[n]| &\geq |e[n]| \left| 1 - 2\tau\mu[n] \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2 - 4\mu[n] \|\tilde{\mathbf{g}}_{e,\mathbf{w}}[n]\|_2^2 \right| \\
1 &\geq \left| 1 - 2\mu[n] \left(\tau \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2 + 2 \|\tilde{\mathbf{g}}_{e,\mathbf{w}}[n]\|_2^2 \right) \right|. \quad (47)
\end{aligned}$$

Solving the inequality for $\mu[n]$ yields

$$0 \leq \mu[n] \leq \frac{1}{2 \|\tilde{\mathbf{g}}_{e,\mathbf{w}}[n]\|_2^2 + \tau \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2} \quad (48)$$

or alternatively

$$\mu[n] = \frac{\mu}{2 \|\tilde{\mathbf{g}}_{e,\mathbf{w}}[n]\|_2^2 + \tau \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2 + \xi}, \quad 0 \leq \mu \leq 1, \quad (49)$$

where the regularization value ξ places an upper bound on $\mu[n]$. The maximum adaptation rate is obtained for $\mu = 0.5$, higher values up to 1 will slow down the optimization again. While the normalization concludes the derivation of the basic CI-WSAF, we proceed by discussing several useful modifications and extensions.

3) *Complex Output*: In the IMD problem, a complex-valued interference signal $y_{\text{BB}}^{\text{IMD}}[n] = y_{\text{BB},\text{I}}^{\text{IMD}}[n] + jy_{\text{BB},\text{Q}}^{\text{IMD}}[n]$ has to be replicated, but the CI-WSAF only provides a real-valued output. When we assume nonlinearity coefficients of the form $\gamma_k = \gamma_{k,\text{I}}(1 + j\delta_{\text{Q}})$ with $\gamma_{k,\text{I}}, \delta_{\text{Q}} \in \mathbb{R}$, then the imaginary part of $y_{\text{BB}}^{\text{IMD}}[n]$ is just a scaled copy of the real part. The coupling δ_{Q} might be estimated by using a single-tap normalized least mean squares (N-LMS) or a weighted least squares (LS) algorithm, following a cascaded scheme as illustrated in Fig. 2. The interference replica used for cancellation is

$$\hat{y}_{\text{BB}}^{\text{IMD}}[n] = \hat{y}_{\text{saf}}[n] + jw_{\text{cpl}}[n] \hat{y}_{\text{saf}}[n] \quad (50)$$

with the weight calculations

$$\begin{aligned}
&w_{\text{cpl}}[n] \\
&= w_{\text{cpl}}[n-1] + \frac{\mu_{\text{cpl}}}{(\hat{y}_{\text{saf}}[n])^2 + \xi} e_{\text{Q}}[n] \hat{y}_{\text{saf}}[n], \quad 0 \leq \mu_{\text{cpl}} \leq 1 \quad (51)
\end{aligned}$$

for the 1-tap normalized LMS (N-LMS) using $e_{\text{Q}}[n] = y_{\text{Q}}[n] - w_{\text{cpl}}[n] \hat{y}_{\text{saf}}[n]$ or

$$w_{\text{cpl}}[n] = \frac{r_{y\hat{y}}[n]}{r_{\hat{y}\hat{y}}[n]} \quad (52)$$

$$r_{y\hat{y}}[n] = r_{y\hat{y}}[n-1] + \lambda_{\text{cpl}} y_{\text{Q}}[n] \hat{y}_{\text{saf}}[n], \quad r_{y\hat{y}}[-1] = 0 \quad (53)$$

$$r_{\hat{y}\hat{y}}[n] = r_{\hat{y}\hat{y}}[n-1] + \lambda_{\text{cpl}} (\hat{y}_{\text{saf}}[n])^2, \quad r_{\hat{y}\hat{y}}[-1] = 0 \quad (54)$$

for the exponentially weighted single-tap LS algorithm.

4) *Transform-Domain Concept*: Furthermore, it is well-known that SGD algorithms suffer from slow convergence in case of correlated input signals [38]. This property is particularly critical if the reference signal is LTE/NR sequence with narrow allocation, a common case for uplink. In [39], it has been shown that the TD concept provides an appropriate mitigation if a pre-computed power normalization is available. For the CI-WSAF, this approach is applied by replacing the input vector $\mathbf{x}[n]$ with

the transformed vector $\mathbf{v}[n]$:

$$\mathbf{v}[n] = \mathbf{P}^{-\frac{1}{2}} \mathcal{D}_{\text{Q}_{\text{lin}}} \mathbf{x}[n], \quad [\mathbf{P}]_{k,l} = \delta_{k,l} \left[\mathcal{D}_{\text{Q}_{\text{lin}}} \mathbf{C}_{\text{xx}} \mathcal{D}_{\text{Q}_{\text{lin}}}^{\text{T}} \right]_{k,l}. \quad (55)$$

$\mathcal{D}_{\text{Q}_{\text{lin}}} \in \mathbb{R}^{Q_{\text{lin}} \times Q_{\text{lin}}}$ is the discrete cosine transform (DCT) matrix and $\delta_{k,l}$ is the Kronecker delta. \mathbf{C}_{xx} is the auto-covariance matrix of the input vector $\mathbf{x}[n]$, which is used to compute the power normalization $\mathbf{P}^{-\frac{1}{2}}$. When calculating \mathbf{C}_{xx} , it is important to take any resampling of $x[n]$ into account. Besides operating on a different input, the CI-WSAF algorithm in the TD variant stays unaltered.

5) *Weight Norm Limiting*: Last, in the iterative optimization of $\mathbf{w}[n]$, the norm correction $\epsilon \tilde{\mathbf{g}}_{e,\mathbf{w}}[n]$ has to be applied in every iteration. In order to relax the requirements on a hardware implementation, we propose an alternative heuristic rescaling of the weights

$$\mathbf{w}_{\text{lim}}[n] = \begin{cases} \mathbf{w}[n] & \|\mathbf{w}[n-1]\|_p^p < \rho_w \\ \frac{\mathbf{w}[n]}{2} & \text{else} \end{cases}, \quad (56)$$

where the decision can be computed in parallel to the update step and the potential rescaling is a simple shift operation. Obviously, this scaling has to be compensated by the spline control points \mathbf{q} in order to maintain the output signal level. According to simulations, the iterative adaptation of \mathbf{q} is capable of this task if the rescaling is applied seldomly (cf. Fig. 7).

B. Extension to Complex Control Points

The CI-WSAF presented above operates on complex input data and uses complex filter weights, but in its basic form has a real-valued output. Still, a complex interference can be canceled by means of an additional single-tap adaptive filter, under the assumption that the interference in the Q-path is a scaled version of the I-path. When considering the IMD BB model (10), this constraint is not fulfilled in general. Therefore, we extend the algorithm by including separate spline functions for the I- and the Q-path.

1) *Basic Algorithm*: The structure of this algorithm is illustrated in Fig. 3. The output of the so-called complex-input-output Wiener SAF (CIO-WSAF) is given by

$$\begin{aligned}
\hat{y}_{\text{saf}}[n] &= \left(\mathbf{v}_{\text{Q}_{\text{sp}}}[n]^{\text{T}} \mathbf{B}_{\text{Q}_{\text{sp}}}^{\text{sp}} \right. \\
&\quad \left. (\mathbf{q}_{\text{I},\text{Q}_{\text{sp},\ell}}[n-1] + j\mathbf{q}_{\text{Q},\text{Q}_{\text{sp},\ell}}[n-1]) \right) * h_{\text{out}}[n] \\
&= \underbrace{\left(\mathbf{v}_{\text{Q}_{\text{sp}}}[n]^{\text{T}} \mathbf{B}_{\text{Q}_{\text{sp}}}^{\text{sp}} \mathbf{q}_{\text{Q}_{\text{sp},\ell}}[n-1] \right)}_{\varphi[n]} * h_{\text{out}}[n], \quad (57)
\end{aligned}$$

where $\mathbf{q}[n] \in \mathbb{C}^{N_{\text{sp}}}$ and $\hat{y}_{\text{saf}}[n] \in \mathbb{C}$. Despite this change, all signal definitions and the SGD cost function, (30)–(37), remain valid. In the IMD problem, the output of the CIO-WSAF can be directly used for interference cancellation, i. e. $\hat{y}_{\text{BB}}^{\text{IMD}}[n] = \hat{y}_{\text{saf}}[n]$. The approximate gradient of the cost function with respect to the control points now also requires the Wirtinger calculus. Applying the corresponding chain rule leads to:

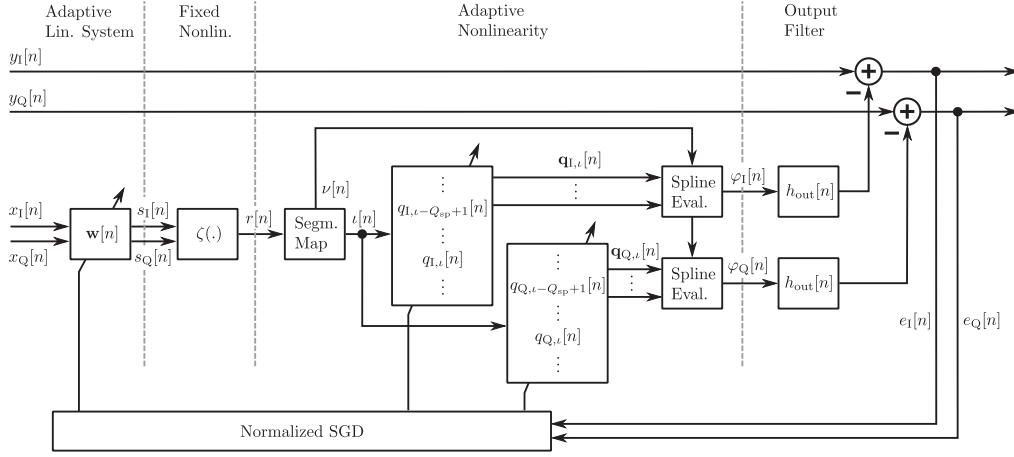


Fig. 3. Block diagram of adaptive nonlinear DSIM structure with complex output.

$$\begin{aligned}
 \tilde{\mathbf{g}}_{\mathbf{q}}[n] &= \frac{\partial J[n]}{\partial e[n]} \frac{\partial e[n]}{\partial \mathbf{q}^H} \Big|_{\mathbf{q}[\cdot]=\mathbf{q}} + \frac{\partial J[n]}{\partial e[n]^*} \frac{\partial e[n]^*}{\partial \mathbf{q}^H} \Big|_{\mathbf{q}[\cdot]=\mathbf{q}^*} \\
 &= -e[n] \sum_{k=0}^{Q_{\text{out}}-1} h_{\text{out}}[k] \begin{bmatrix} \mathbf{0}_{(\iota[n-k]-Q_{\text{sp}}+1) \times 1} \\ (\mathbf{B}_{Q_{\text{sp}}}^{\text{SP}})^T \mathbf{v}_{Q_{\text{sp}}}[n-k] \\ \mathbf{0}_{(N_{\text{sp}}-\iota[n-k]-1) \times 1} \end{bmatrix} \\
 &= e[n] \tilde{\mathbf{g}}_{e, \mathbf{q}}[n]. \tag{58}
 \end{aligned}$$

Due to the real spline input, this result can be interpreted as an independent optimization of the nonlinearities in the I- and the Q-path. The same structure of the chain rule also occurs in case of the approximate gradient with respect to the filter weights:

$$\begin{aligned}
 \tilde{\mathbf{g}}_{\mathbf{w}}[n]^T &= \frac{\partial J[n]}{\partial e[n]} \frac{\partial e[n]}{\partial \mathbf{w}^*} \Big|_{\mathbf{w}[\cdot]=\mathbf{w}} + \frac{\partial J[n]}{\partial e[n]^*} \frac{\partial e[n]^*}{\partial \mathbf{w}^*} \Big|_{\mathbf{w}[\cdot]=\mathbf{w}^*} \\
 &\quad + \frac{\partial J[n]}{\partial c_{\mathbf{w}}[n]} \frac{c_{\mathbf{w}}[n]}{\partial \mathbf{w}[n-1]^*} \\
 &= \tilde{\mathbf{g}}_{J, e, \mathbf{w}}[n]^T + \epsilon \tilde{\mathbf{g}}_{c, \mathbf{w}}[n]^T. \tag{59}
 \end{aligned}$$

The notation $\tilde{\mathbf{g}}_{J, e, \mathbf{w}}[n]$ indicates that no separation between the error $e[n]$ and the partial derivative of $e[n]$ is possible anymore, which is a major difference to the real case. The gradient $\tilde{\mathbf{g}}_{J, e, \mathbf{w}}[n]$ is given by

$$\begin{aligned}
 \tilde{\mathbf{g}}_{J, e, \mathbf{w}}[n] &= \frac{\partial J[n]}{\partial e[n]} \sum_{k=0}^{Q_{\text{out}}-1} \frac{\partial e[n]}{\partial \mathbf{v}_{Q_{\text{sp}}}[n-k]} \frac{\partial \mathbf{v}_{Q_{\text{sp}}}[n-k]}{\partial \nu[n-k]} \\
 &\quad \frac{\partial \nu[n-k]}{\partial r[n-k]} \frac{\partial r[n-k]}{\partial s[n-k]^*} \frac{\partial s[n-k]^*}{\partial \mathbf{w}^H} \Big|_{\mathbf{w}[\cdot]=\mathbf{w}} \\
 &\quad + \frac{\partial J[n]}{\partial e[n]^*} \sum_{k=0}^{Q_{\text{out}}-1} \frac{\partial e[n]^*}{\partial \mathbf{v}_{Q_{\text{sp}}}[n-k]} \frac{\partial \mathbf{v}_{Q_{\text{sp}}}[n-k]}{\partial \nu[n-k]} \\
 &\quad \frac{\partial \nu[n-k]}{\partial r[n-k]} \frac{\partial r[n-k]}{\partial s[n-k]^*} \frac{\partial s[n-k]^*}{\partial \mathbf{w}^H} \Big|_{\mathbf{w}[\cdot]=\mathbf{w}} \\
 &= -\frac{2}{\Delta r} \sum_{k=0}^{Q_{\text{out}}-1} h_{\text{out}}[k] \mathbf{v}'_{Q_{\text{sp}}}[n-k]^T \mathbf{B}_{Q_{\text{sp}}}^{\text{SP}} \\
 &\quad \Re\{e[n]^* \mathbf{q}_{Q_{\text{sp}}, \iota[n-k]}[n-1-k]\} \zeta'(s[n-k])^* \mathbf{x}[n-k]^*. \tag{60}
 \end{aligned}$$

This form is obtained by factoring out all common terms and using the identity $a + a^* = 2 \Re\{a\}$. The gradient of the norm constraint (43), as well as the limiter (56), remain unaffected by the complex control points and can be directly reused. Combining all results, the update equations of the CIO-WSAF are

$$\mathbf{q}[n] = \mathbf{q}[n-1] - \tau \mu[n] e[n] \tilde{\mathbf{g}}_{e, \mathbf{q}}[n] \tag{61}$$

$$\mathbf{w}[n] = \mathbf{w}[n-1] - \mu[n] (\tilde{\mathbf{g}}_{J, e, \mathbf{w}}[n] + \epsilon \tilde{\mathbf{g}}_{c, \mathbf{w}}[n]). \tag{62}$$

2) *Step-Size Normalization*: Similar to the real case, the adaptation performance of the algorithm is improved by calculating a partial linear Taylor approximation of the error signal and choosing the step-size such that the error decreases with each iteration. In order to simplify the corresponding inequality for $\mu[n]$, we replace the convolution with $h_{\text{out}}[n]$ by the group delay k_g and the passband gain h_g of the output filter.

$$\begin{aligned}
 e[n+1] &\approx e[n] + \frac{\partial e[n]}{\partial \mathbf{q}} \Big|_{\mathbf{q}[\cdot]=\mathbf{q}} \Delta \mathbf{q}[n] \\
 &\quad + \frac{\partial e[n]}{\partial \mathbf{w}} \Big|_{\mathbf{w}[\cdot]=\mathbf{w}; h_g} \Delta \mathbf{w}[n] + \frac{\partial e[n]}{\partial \mathbf{w}^*} \Big|_{\mathbf{w}^*[\cdot]=\mathbf{w}^*; h_g} \Delta \mathbf{w}^*[n]^* \\
 &\approx e[n] - \tau \mu[n] e[n] \|\tilde{\mathbf{g}}_{e, \mathbf{q}}[n]\|_2^2 \\
 &\quad - \mu[n] \frac{4}{(\Delta r)^2} h_g^2 \mathbf{v}'_{Q_{\text{sp}}}[n-k_g]^T \mathbf{B}_{Q_{\text{sp}}}^{\text{SP}} \mathbf{q}_{Q_{\text{sp}}, \iota[n-k_g]}[n-1-k_g] \\
 &\quad \mathbf{v}'_{Q_{\text{sp}}}[n-k_g]^T \mathbf{B}_{Q_{\text{sp}}}^{\text{SP}} \Re\{e[n]^* \mathbf{q}_{Q_{\text{sp}}, \iota[n-k_g]}[n-1-k_g]\} \\
 &\quad |\zeta'(s[n-k_g])|^2 \|\mathbf{x}[n-k_g]\|_2^2 \tag{63}
 \end{aligned}$$

In the classical derivation of the step-size bound, it is assumed that $|e[n+1]| \leq |e[n]|$. Since no closed-form solution of this inequality for $\mu[n]$ independent of $e[n]$ is possible, we instead place constraints in the real and imaginary part of $e[n]$ separately:

$$\begin{aligned}
 |e_I[n]| &\geq |e_I[n+1]| \\
 &\geq \left| e_I[n] - \tau \mu[n] e_I[n] \|\tilde{\mathbf{g}}_{e, \mathbf{q}}[n]\|_2^2 \right. \\
 &\quad \left. - \mu[n] e_I[n] \frac{4}{(\Delta r)^2} h_g^2 |\zeta'(s[n-k_g])|^2 \|\mathbf{x}[n-k_g]\|_2^2 \right|
 \end{aligned}$$

$$\left(\mathbf{v}'_{Q_{sp}}[n-k_g]^T \mathbf{B}_{Q_{sp}}^{sp} \mathbf{q}_{l,Q_{sp},\ell}[n-k_g][n-1-k_g] \right)^2 \Big| \geq |e_I[n]| |1 - \mu[n] b_{\mu,I}[n]| \quad (64)$$

$$|e_Q[n]| \geq |e_Q[n+1]|$$

$$\begin{aligned} &\geq \left| e_Q[n] - \tau \mu[n] e_Q[n] \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2 \right. \\ &\quad \left. - \mu[n] e_Q[n] \frac{4}{(\Delta r)^2} h_g^2 |\zeta'(s[n-k_g])|^2 \|\mathbf{x}[n-k_g]\|_2^2 \right. \\ &\quad \left. \left(\mathbf{v}'_{Q_{sp}}[n-k_g]^T \mathbf{B}_{Q_{sp}}^{sp} \mathbf{q}_{l,Q_{sp},\ell}[n-k_g][n-1-k_g] \right)^2 \right| \\ &\geq |e_Q[n]| |1 - \mu[n] b_{\mu,Q}[n]|. \quad (65) \end{aligned}$$

Here, we used the definitions

$$\begin{aligned} b_{\mu,I}[n] &= \tau \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2 + \frac{4}{(\Delta r)^2} h_g^2 |\zeta'(s[n-k_g])|^2 \|\mathbf{x}[n-k_g]\|_2^2 \\ &\quad \left(\mathbf{v}'_{Q_{sp}}[n-k_g]^T \mathbf{B}_{Q_{sp}}^{sp} \mathbf{q}_{l,Q_{sp},\ell}[n-k_g][n-1-k_g] \right)^2 \quad (66) \end{aligned}$$

$$\begin{aligned} b_{\mu,Q}[n] &= \tau \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2 + \frac{4}{(\Delta r)^2} h_g^2 |\zeta'(s[n-k_g])|^2 \|\mathbf{x}[n-k_g]\|_2^2 \\ &\quad \left(\mathbf{v}'_{Q_{sp}}[n-k_g]^T \mathbf{B}_{Q_{sp}}^{sp} \mathbf{q}_{l,Q_{sp},\ell}[n-k_g][n-1-k_g] \right)^2. \quad (67) \end{aligned}$$

In the inequality for $e_I[n]$, we neglect $e_Q[n]$ and vice versa. The solutions of both inequalities are of the form $0 \leq \mu[n] \leq 2/b_{\mu,IQ}[n]$. Since we did not include any coupling between $e_I[n]$ and $e_Q[n]$, the final upper bound for $\mu[n]$ is chosen conservatively as $\mu_{\max}[n] = 2/(b_{\mu,I}[n] + b_{\mu,Q}[n] + \xi)$. This value is guaranteed to be smaller than or equal to $2/b_{\mu,IQ}[n]$ since the $b_{\mu,IQ}[n]$ are non-negative. Following this approach, we obtain the final step-size normalization by choosing

$$\begin{aligned} \mu[n] &= \mu \mu_{\max}[n] = \mu \frac{2}{b_{\mu,I}[n] + b_{\mu,Q}[n] + \xi} \\ &= \mu \left(2 \frac{h_g^2}{(\Delta r)^2} |\zeta'(s[n-k_g])|^2 \|\mathbf{x}[n-k_g]\|_2^2 \right. \\ &\quad \left. \left| \mathbf{v}'_{Q_{sp}}[n-k_g]^T \mathbf{B}_{Q_{sp}}^{sp} \mathbf{q}_{l,Q_{sp},\ell}[n-k_g][n-1-k_g] \right|^2 \right. \\ &\quad \left. + \tau \|\tilde{\mathbf{g}}_{e,\mathbf{q}}[n]\|_2^2 + \xi \right)^{-1} \quad (68) \end{aligned}$$

with the constant step-size $0 \leq \mu \leq 1$. In order to improve the adaptation rate and the performance consistency of the CIO-WSAF, the TD concept (55) can be applied, too.

C. Computational Complexity

An important aspect in DSIM applications is the computational complexity of the estimation algorithm, since it directly determines the power consumption and real-time capability.

Therefore, we provide the general number of operations required by the CI-WSAF and the CIO-WSAF, depending on the filter length and the spline parameters. Additionally, we compare the complexity to two state-of-the-art algorithms for a specific configuration used in Section V.

1) *CI-WSAF*: In Table II, we break down the output and update equations of the CI-WSAF and its extension into real-valued operations per input sample. Besides additions/subtractions and multiplications we separately list division and square root operations. Due to their complexity, the latter two are usually approximated. We do not detail possible division and square root implementations in this work. For example, in [40], a low-complex approximation of the reciprocal is shown based on look-up tables. This method would imply an additional multiplication for every division, where the numerator is not 1. All operations that only involve constants are assumed to be precomputed. In Table II, we used the fixed nonlinearity $\zeta(s) = |s|^2$ as a baseline, the additional operations required for $\zeta(s) = |s|$ are listed separately. Observing the spline basis matrices in (25)–(27) reveals that the complexity of the product $\mathbf{v}'^T \mathbf{B}^{sp} \mathbf{q}$ heavily depends on the spline type. Besides the omission of products with zero, all powers of two would result in simple shift operations. However, in the following we assume a general \mathbf{B}^{sp} to ensure generality of the results. We also consider a general output filter of length Q_{out} . Consequently, the values in Table II represent an upper bound at the base sampling rate. If the Tx BB allocation necessitates an oversampling factor larger than 1, the effective complexity per Rx sample at the Rx BB rate increases accordingly. In case of the norm limiter, we neglect the scaling of the filter weights if the norm exceeds the target value ρ_w , since this step occurs very seldomly. The DCT (type-II) used by the TD extension is optimized for delay-line inputs, thereby reducing its complexity significantly [41]. This variant is also known as sliding cosine transform (SCT).

2) *CIO-WSAF*: The baseline complexity of the CIO-WSAF for $\zeta(s) = |s|^2$ is given in Table III. There, the additional operations for $\zeta(s) = |s|$ and the step-size normalization are provided, too. The extensions for controlling the norm of the weight vector and for decorrelating the input data are unaffected by the complex control points. Thus, their complexity can be found in Table II. Since the CIO-WSAF provides a complex-valued output, the single-tap scaler can be omitted. Overall, the costs of the CI-WSAF and CIO-WSAF are comparable, especially for short output filters.

3) *Comparison to State-of-The-Art Concepts*: The first chosen state-of-the-art algorithm is the IM2LMS [6], least mean squares (LMS) variant, which is limited to the cancellation of second-order IMD. It requires the Q-path interference signal to be derived from the I-path by scaling. Thus, it is combined with the single-tap LS scaler (52). The second method we use for comparison is the kernel recursive least squares (KRLS) algorithm, a very general adaptive learning concept [42]. We choose a real-valued Gaussian kernel with a complex-valued input [43], which allows the KRLS to model a wide range of complex-valued nonlinearities without the single-tap scaler. In order to limit its complexity, the KRLS needs an additional

TABLE II
 ARITHMETIC OPERATIONS PER SAMPLE OF CI-WSAF AND EXTENSIONS

Module		Add./Sub.	Mult.	Div.	Sqrt.
CI-WSAF, $\zeta(s) = s ^2$		$2 Q_{\text{lin}} (Q_{\text{out}} + 3) + 2 Q_{\text{sp}}^2$ $+ Q_{\text{sp}} (Q_{\text{out}} - 2) + Q_{\text{out}} + N_{\text{sp}} - 1$	$2 Q_{\text{lin}} (Q_{\text{out}} + 5) + 2 Q_{\text{sp}}^2$ $+ Q_{\text{sp}} (Q_{\text{out}} + 1) + Q_{\text{out}} + N_{\text{sp}} + 4$	0	0
Additional operations for $ s $		0	1	2	1
Normalization		$2 Q_{\text{lin}} + N_{\text{sp}}$	$2 Q_{\text{lin}} + N_{\text{sp}} + 2$	1	0
Constr. SGD	ℓ_1 norm	$4 Q_{\text{lin}}$	$6 Q_{\text{lin}} - 1$	$2 Q_{\text{lin}}$	Q_{lin}
	ℓ_2 norm	$4 Q_{\text{lin}}$	$6 Q_{\text{lin}} - 1$	0	0
Norm limiter	ℓ_1 norm	$2 Q_{\text{lin}}$	$2 Q_{\text{lin}}$	0	Q_{lin}
	ℓ_2 norm	$2 Q_{\text{lin}}$	$2 Q_{\text{lin}}$	0	0
Transf. input		$4 Q_{\text{lin}} + 10$	$6 Q_{\text{lin}} - 4$	0	0
1-tap scaler	N-LMS	3	4	1	0
	Weighted LS	3	4	1	0

 TABLE III
 ARITHMETIC OPERATIONS PER SAMPLE OF CIO-WSAF (FOR EXTENSIONS SEE TABLE II)

Module		Add./Sub.	Mult.	Div.	Sqrt.
CIO-WSAF, $\zeta(s) = s ^2$		$2 Q_{\text{lin}} (Q_{\text{out}} + 3) + 2 Q_{\text{sp}}^2$ $+ Q_{\text{sp}} Q_{\text{out}} + 2 Q_{\text{out}} + 2 N_{\text{sp}} - 2$	$2 Q_{\text{lin}} (Q_{\text{out}} + 5) + 2 Q_{\text{sp}}^2$ $+ Q_{\text{sp}} (Q_{\text{out}} + 3) + 2 Q_{\text{out}} + 2 N_{\text{sp}} + 6$	0	0
Additional operations for $ s $		0	1	1	1
Normalization		$2 Q_{\text{lin}} + 2 N_{\text{sp}}$	$2 Q_{\text{lin}} + 2 N_{\text{sp}} + 2$	1	0

 TABLE IV
 EXEMPLARY COMPLEXITY OF DSIM ALGORITHMS FOR IMD CANCELLATION

Algorithm		Add./Sub.	Mult.	Div.	Sqrt.	Exp.
IM2LMS + 1-tap scaler		164	201	2	0	0
CI-WSAF + 1-tap scaler		290	353	2	16	0
CIO-WSAF		299	363	1	16	0
KRLS, $ \mathcal{D} = 125$	Prediction	8125	4375	0	0	125
	Pred. + Train.	$7.1 \cdot 10^4$	$6.7 \cdot 10^4$	125	0	125
KRLS, $ \mathcal{D} = 700$	Prediction	$4.6 \cdot 10^4$	$2.5 \cdot 10^4$	0	0	700
	Pred. + Train.	$2.0 \cdot 10^6$	$2.0 \cdot 10^6$	700	0	700

sparsification method, in our case the approximate linear dependency (ALD) criterion. It maintains a growing dictionary \mathcal{D} of relevant input vectors, combined with a set of complex weights. In stationary scenarios, the dictionary size $|\mathcal{D}|$ can be considered to be settled at some point. In the IMD cancellation scenarios in Section V, the typical dictionary size is 125, with maximum values of up to 700 for high leakage powers.

Table IV explicitly gives the number of operations per sample for all discussed algorithms applied to IMD interference cancellation. The length of the input vector Q_{lin} is chosen to be 16 and Q_{sp} is 3. The spline-based methods use a normalized step-size, the weight norm limiter with $p = 1$ and the TD extension. In addition, the output filter is a simple delay, which allows for significant computational savings. In the equations in Table II this can be reflected by replacing N_{sp} with Q_{sp} . Hence, both spline models are considered to be implementable for real-time cancellation. The higher complexity of the CI(O)-WSAF compared to the IM2LMS is clearly outweighed by its much higher flexibility. The numbers for the KRLS assume a settled dictionary of given size, where only the weights are adjusted.

The Gaussian kernel requires the evaluation of $\exp(x)$, which would be approximated in practice. Still, the KRLS requires immense computing power for online cancellation at typical LTE/NR sampling rates. Even the sole prediction of new samples without altering the weights is too expensive for mobile devices. In the adaptation phase, the CIO-WSAF features less than 0.6% of the additions and multiplications of the KRLS.

V. PERFORMANCE EVALUATION

We conclude our investigations on spline-based DSIM algorithms with performance simulations on a realistic IMD interference scenario. Besides the steady-state cancellation with and without output delays, we also show the effect of the norm constraint on the filter weights.

A. Setup

The interference model used in the following is based on measurements on an integrated CMOS transceiver using the setup outlined in Fig. 4. The Tx signal is generated using vector signal generator (VSG) and optionally amplified by a discrete PA. A discrete duplexing filter provides the Tx-Rx isolation in the AFE. The signal after the duplexer is captured by vector signal analyzer (VSA) in order to obtain the actual input signal of the Rx chain. Simultaneously, the signal is passed to an external LNA with a gain of 15 dB. Unlike the internal LNA of the subsequent RF transceiver, the external LNA is assumed to be sufficiently linear to not contribute to the IMD. On the one hand, this setup enables a characterization of multiple duplexers for different LTE bands. On the other hand, it allows to extract the IMD components from the Rx BB sequence. The IMD interference power was measured for various leakage power

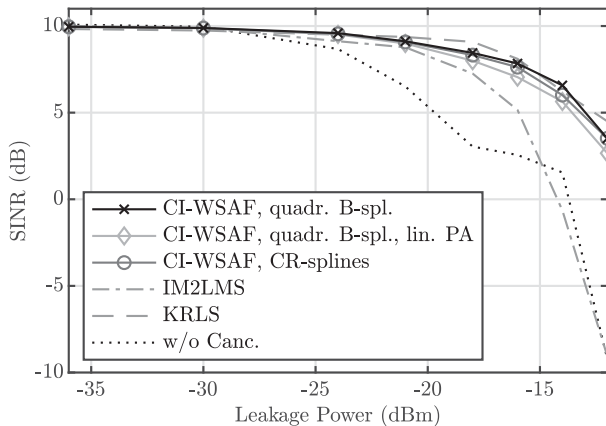


Fig. 6. Steady-state SINR improvement for various IMD cancellation schemes in case of scaled nonlinearity in Q-path.

performance is computed. This evaluation approach is vital especially for LMS-type algorithms, since they typically suffer from a high performance spread. In the simulation, we neglect any intermodulation products between the leakage signal, the wanted Rx signal and noise due to the wide bandwidth and low power of the wanted and noise components. Even if this assumption was not fulfilled, our simplification would only affect the optimum SNR, but not the relative performance between several DSIM algorithms. Due to the narrow allocation, no oversampling is required for the analyzed IMD products, i. e. up to the sixth order.

B. Scaled Nonlinearity in Q-Path

In the first test case, we assume a scalar coupling between the nonlinearities in the I- and Q-path. While this assumption is not fulfilled by the targeted receiver, it enables a comparison between the CI-WSAF and the comparably low-complex IM2LMS. In addition, we include the computationally intensive, but very general, KRLS algorithm. Since the ratio between the measured coefficients $\gamma_{k,I}$ and $\gamma_{k,Q}$ changes substantially over the leakage power range, we alter these values for this simulation. We keep the interference power unchanged and assume a coupling of $\delta_Q = -1$, leading to the modified coefficients

$$\tilde{\gamma}_k = (1 - j) \frac{\text{sgn}(\gamma_{k,I}) |\gamma_k|}{\sqrt{2}}. \quad (72)$$

The value of δ_Q does not influence the performance, since it is reliably estimated by the single-tap scaler.

The cancellation performance of the algorithms is compared by means of the SINR

$$\text{SINR} := \frac{\text{E} \left[|y_{\text{BB}}^{\text{Rx}}[n]|^2 \right]}{\text{E} \left[|y_{\text{BB}}^{\text{IMD}}[n] - \hat{y}_{\text{BB}}^{\text{IMD}}[n] + \tilde{\gamma}_{\text{BB}}[n]|^2 \right]}, \quad (73)$$

which is averaged over two LTE slots, while excluding the first symbol. Thus, the initial convergence does not affect the measured performance.

Fig. 6 depicts the SINR for two variants of the CI-WSAF and two algorithms for comparison. In addition, we show the

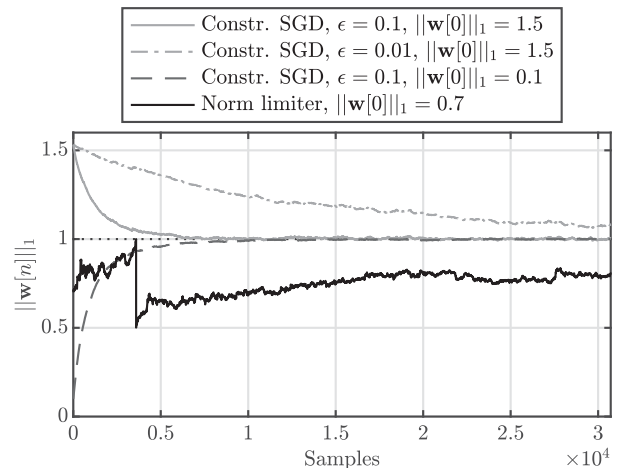


Fig. 7. Adaptation of ℓ_1 norm of filter weights due to constrained SGD optimization and limiter.

impact of the PA model. Each value is the ensemble average over the results for six different fitted duplexer impulse responses, where for each duplexer 50 runs with randomly generated signals were performed. The linear filter length of all algorithms is set to 16. All relevant parameters are optimized for the individual leakage power levels to guarantee the highest possible performance for a fair comparison. In case of the IM2LMS, previous knowledge about the sign of the real part of the IMD2 component is required in order to ensure correct operation. In practice, this information could be obtained by correlating $|x_{\text{BB}}[n]|^2$ with $y_{\text{BB}}^{\text{Tot}}[n]$. Since the output of the IM2LMS is real-valued, the Q-path interference is estimated using (50), similar to the CI-WSAF. Because of its superior performance, we employ the weighted LS solution with $\lambda_{\text{cpl}} = 0.9998$ for this purpose. The DC cancellation in the receiver is replicated by applying the notch filter $H(z) = (1 - z^{-1}) / (1 - 0.998z^{-1})$ to the IM2LMS output. The step-size of the IM2LMS is selected in the range $(0, 0.056]$, where extremely small values are used for low interference powers. The regularization parameter is chosen within $[0.01, 1]$. Despite the necessity of prior knowledge, the IM2LMS shows the lowest SINR among the compared algorithms. The CI-WSAF is used in the TD variant (55) with the norm limiting (56). In order to improve the adaptation rate of the spline-based algorithm, the weights of the linear section are initialized to random constants, which are unaltered for all runs. The fixed nonlinearity is $\zeta(s) = |s|^2$ to avoid a square root in the feedback loop.

Table V summarizes the chosen parameter ranges for the CI-WSAF, where most values tend to become larger with increasing leakage power. $\mathcal{N}_{Q_{\text{in}}}$ denotes a multivariate normal distribution. The parameter λ_{cpl} of the single-tap scaler was again 0.9998. A negative value for the first knot r_0 enables a domain of the spline function down to 0, as required by the fixed nonlinearity $|s|^2$. We have $h_{\text{out}}[n] = 1$, since we do not employ any pipelining and the cancellation is performed before the CSF. In Fig. 6, the CI-WSAF shows excellent performance, clearly outperforming the IM2LMS by up to 12.5 dB. For all leakage powers levels, it restores the SINR to values above the 0 dB threshold. For

TABLE V
PARAMETER VALUES OF CI-WSAF FOR IMD CANCELLATION

Parameter	Value(s)	Parameter	Value(s)
Q_{lin}	16	τ	[1, 660]
$\mathbf{w}[0]$	$\sim \mathcal{N}_{Q_{\text{lin}}}(\mathbf{0}, 0.01 \mathbf{I})$	N_{sp}	20
p	1	Q_{sp}	3, 4
ρ_w	3	r_0	-0.1, -0.15
μ	[0.0015, 0.015]	Δr	0.05
ξ	[0.0023, 0.45]		

the same number of control points, there is no benefit of using cubic CR-splines instead of quadratic B-splines, likely due to the smoothness of the IMD nonlinearity. Compared to a linear PA model, the clipping of the signals peaks by the Rapp PA model slightly improves the performance of the WSAF algorithms. Without clipping, signal peaks occur very seldomly, leading to slow adaptation of the corresponding spline control points and, in further consequence, higher estimation errors. Another minor SINR improvement is achievable by using an LS-based learning algorithm, in our case the KRLS. For this simulation, we used a complexified Gaussian kernel with a standard deviation in the range [5, 100]. The dictionary size was limited by means of the ALD approach with a threshold in the range $[10^{-5}, 5 \cdot 10^{-5}]$. Due to its fast convergence and high complexity, the KRLS weights were adapted for the first three LTE symbols, or about 6600 samples, only. Its maximum SINR advantage over the CI-WSAF is about 1 dB. A practical aspect to improve the convergence of all algorithms is to disable the adaptation for all samples close to the symbol boundaries within $x_{\text{BB}}[n]$. This measure avoids high error values, which would be caused by the bandwidth increase at symbol transitions [39].

Another interesting aspect of the CI-WSAF is the constraint on the filter weights, which helps to avoid internal clipping. In Fig. 7, the evolution of the ℓ_1 norm of $\mathbf{w}[n]$ is compared for a leakage power of -18 dBm and different parameters of the constraint. Using the constrained SGD method with a weighting of $\epsilon = 0.1$, $\|\mathbf{w}[n]\|_1$ converges close to the desired value of 1 within three symbols. A weighting of $\epsilon = 0.01$ already slows down the adaptation substantially, but still the norm tends towards 1, thereby reducing the clipping. Although this method shows the expected results, the second optimization goal potentially impacts the steady-state interference cancellation by slowing down the overall adaptation. In contrast, the norm limiting is also able to reduce clipping while not impacting the steady-state performance at all. However, the weight rescaling in the initial phase leads to a performance disadvantage in the first 5000 samples.

C. Independent Nonlinearity in I- and Q-Path

In the second major test case, we allow for independent coefficients $\gamma_{k,I}$ and $\gamma_{k,Q}$, a requirement indicated by measurements on the test receiver. The IM2LMS and the CI-WSAF are not suitable for this scenario, thus, we resort to a comparison between the CIO-WSAF and the KRLS. The first metric we

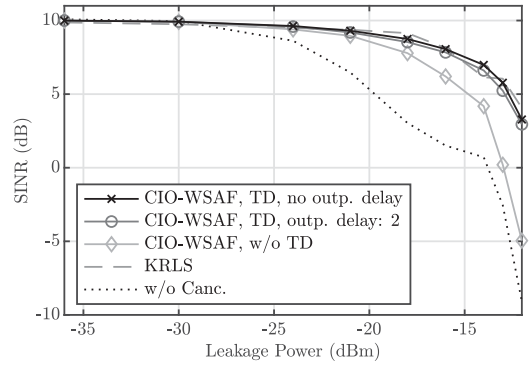


Fig. 8. Steady-state SINR improvement for general IMD problem with independent nonlinearities in I- and Q-path.

analyze is again the SINR, depicted in Fig. 8. The overall SINR degradation without any countermeasures is similar to the previous section, but now the underlying number of parameters is higher. Interestingly, this causes the previous small performance advantage of the KRLS to vanish completely. Again, its kernel standard deviation was chosen in the range [4.8, 141], whereas the ALD threshold was chosen within $[10^{-5}, 5 \cdot 10^{-5}]$. The KRLS weights and the dictionary were adapted for the first three LTE symbols. Unlike the KRLS, the CIO-WSAF with transformed input exhibits a slight SINR improvement compared to the CI-WSAF. The simulation shows that the approximated step-size normalization of the CIO-WSAF behaves just as the standard version used in the CI-WSAF. Compared to Table V, the step-size $\mu \in [4.4 \cdot 10^{-4}, 0.025]$, the coupling factor $\tau \in [1.3, 1200]$ and the regularization $\xi \in [0.017, 0.16]$ required readjustments. The other parameters could be reused. Since no output filter is present, we have $h_g = 1$ and $k_g = 0$. We only consider quadratic interpolation, thus, $Q_{\text{sp}} = 3$. In Fig. 8, we included two other variants of the CIO-WSAF. One incorporates a delay of two samples in the output signal, possibly caused by pipelining stages in the design. In the algorithm, this configuration requires $k_g = 2$. The performance cost of this measure is minor, supporting a real-time implementation of our approach. The second WSAF variant we included does not use a transformed input vector for the linear section and has no output delay (i. e. $k_g = 0$). Due to the properties of the LTE UL signal, which acts as a reference, the omission of the approximate decorrelation leads to SINR drop of up to 8.2 dB.

In addition to the integral performance quantified by the SINR, the power spectral density (PSD) of the Rx signal components in Fig. 9 gives a descriptive visualization of the cancellation process. As a baseline, we show the noise floor $\tilde{\eta}_{\text{BB}}[n]$ and the ideal Rx signal $\tilde{y}_{\text{BB}}^{\text{RX}}[n]$, which, combined with the IMD interference, form the total Rx signal $y_{\text{BB}}^{\text{Tot}}[n]$. In this example, we chose a leakage power of -14 dBm. For interference cancellation, we apply the CIO-WSAF with transformed input and no pipelining, which generates the interference replica $\hat{y}_{\text{BB}}^{\text{IMD}}[n]$. After DSIM, the remaining interference $e_{\text{BB}}^{\text{IMD}}[n] = y_{\text{BB}}^{\text{IMD}}[n] - \hat{y}_{\text{BB}}^{\text{IMD}}[n]$ is below the Rx level, which acts as additional noise from the perspective of the CIO-WSAF. The signal after DSIM, $y_{\text{BB}}^{\text{Canc}}[n]$, closely resembles the wanted Rx

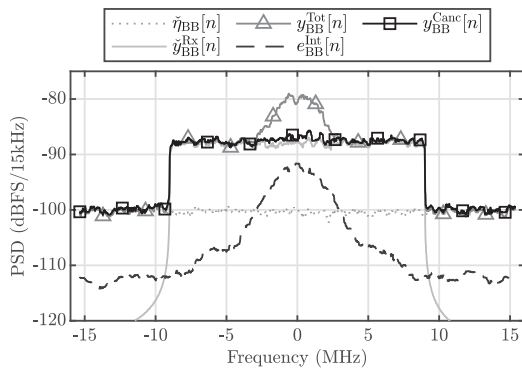


Fig. 9. PSD of important Rx signal components before and after IMD cancellation using CIO-WSAF.

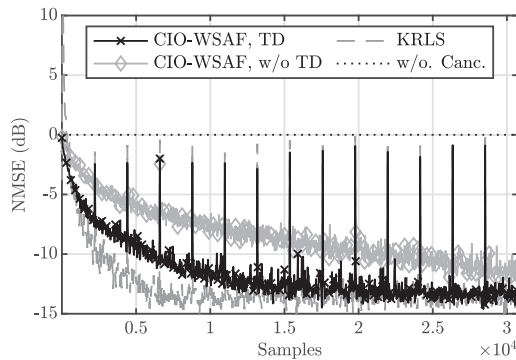


Fig. 10. Adaptation behavior of IMD cancellation schemes in case of independent nonlinearities in I- and Q-path.

signal. This indicates a successful cancellation, which agrees with the SINR of 7 dB shown in Fig. 8.

Another important metric for DSIM applications is the adaptation time of the estimation algorithm. Therefore, in Fig. 10 we compare the normalized mean square error (NMSE)

$$\text{NMSE}[n] := \frac{\text{E} \left[|y_{BB}^{\text{IMD}}[n] - \hat{y}_{BB}^{\text{IMD}}[n]|^2 \right]}{\text{E} \left[|y_{BB}^{\text{IMD}}[n]|^2 \right]} \quad (74)$$

for two SAF variants and the KRLS at a leakage power level of -18 dBm. As expected for RLS-type algorithms, the KRLS features very fast adaptation within the training sequence of 6600 samples. In contrast, the CIO-WSAF without an input transform requires more than $3 \cdot 10^4$ samples to reach steady-state. The TD variant manages to reduce the adaptation time to about $2 \cdot 10^4$ samples, which is a remarkable improvement for an SGD algorithm.

VI. CONCLUSION

We presented two novel adaptive learning schemes based on spline interpolation that allow for low-complex digital cancellation of transceiver self-interference caused by higher-order IMD. Based on a comprehensive modeling of the receiver nonlinearities, we derived the BB interference model, allowing to extract a Wiener structure suitable for online estimation. We proposed several extensions to the spline-based algorithm and

precisely assessed the computational complexity of all modules. The interference cancellation performance was evaluated in practically relevant IMD scenarios, which were based on measured receiver parameters. Compared to a general nonlinear estimation approach, our algorithms showed similar accuracy at a fraction of the computational costs.

REFERENCES

- [1] B. Razavi, “Design considerations for direct-conversion receivers,” *IEEE Trans. Circuits Systems II: Analog Digital Signal Processing*, vol. 44, no. 6, pp. 428–435, Jun. 1997.
- [2] “Saw duplexer, LTE / WCDMA band 3, series: B8625,” TDK, data sheet. [Online]. Available: <http://static6.arrow.com/arpdfconversion/d02a7e8b179789d6426087c1ffa6fa33cd180f0f/b8625.pdf>
- [3] S. Sadjina, C. Motz, T. Paireder, M. Huemer, and H. Pretl, “A survey of self-interference in LTE-Advanced and 5G new radio wireless transceivers,” *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 3, pp. 1118–1131, Feb. 2020.
- [4] A. Frotzschner and G. Fettweis, “Least squares estimation for the digital compensation of Tx leakage in zero-IF receivers,” in *Proc. IEEE Glob. Telecommun. Conf. (GLOBECOM)*, Nov. 2009, pp. 1–6.
- [5] A. Kiayani, L. Anttila, M. Kosunen, K. Stadius, J. Ryynänen, and M. Valkama, “Modeling and joint mitigation of TX and RX nonlinearity-induced receiver desensitization,” *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 7, pp. 2427–2442, Jul. 2017.
- [6] A. Gebhard, C. Motz, R. S. Kanumalli, H. Pretl, and M. Huemer, “Non-linear least-mean-squares type algorithm for second-order interference cancellation in LTE-A RF transceivers,” in *Proc. 51st Asilomar Conf. Signals, Syst., and Computers*, 2017, pp. 802–807.
- [7] A. Gebhard *et al.*, “A robust nonlinear RLS type adaptive filter for second-order-intermodulation distortion cancellation in FDD LTE and 5G direct conversion transceivers,” *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 5, pp. 1946–1961, May 2019.
- [8] A. Frotzschner and G. Fettweis, “A stochastic gradient LMS algorithm for digital compensation of Tx leakage in zero-IF-receivers,” in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, May 2008, pp. 1067–1071.
- [9] C. Lederer and M. Huemer, “LMS based digital cancellation of second-order TX intermodulation products in homodyne receivers,” in *Proc. IEEE Radio Wirel. Symp.*, Jan. 2011, pp. 207–210.
- [10] E. A. Keehr and A. Hajimiri, “Successive regeneration and adaptive cancellation of higher order intermodulation products in RF receivers,” *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 5, pp. 1379–1396, May 2011.
- [11] T. Paireder, C. Motz, S. Sadjina, and M. Huemer, “A robust mixed-signal cancellation approach for even-order intermodulation distortions in LTE-A/5G-transceivers,” *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 68, no. 3, pp. 923–927, Mar. 2021.
- [12] D. Comminiello and J. Principe, Eds., *Adaptive Learning Methods for Nonlinear System Modeling*. Elsevier Science, 2018.
- [13] C. Auer, K. Kostoglou, T. Paireder, O. Ploder, and M. Huemer, “Support vector machines for self-interference cancellation in mobile communication transceivers,” in *Proc. IEEE 91st Veh. Technol. Conf. (VTC2020-Spring)*, 2020, pp. 1–6.
- [14] C. Auer, T. Paireder, O. Lang, and M. Huemer, “Kernel recursive least squares based cancellation for receiver-induced self-interference,” in *Proc. 54th Asilomar Conf. Signals, Systems, Comput.*, 2020, *accepted for publication*.
- [15] L. Vecchi, F. Piazza, and A. Uncini, “Learning and approximation capabilities of adaptive spline activation function neural networks,” *Neural Netw.*, vol. 11, no. 2, pp. 259–270, 1998.
- [16] A. Uncini and F. Piazza, “Blind signal processing by complex domain adaptive spline neural networks,” *IEEE Trans. Neural Netw.*, vol. 14, no. 2, pp. 399–412, Mar. 2003.
- [17] P. Bohra, J. Campos, H. Gupta, S. Aziznejad, and M. Unser, “Learning activation functions in deep (spline) neural networks,” *IEEE Open J. Signal Process.*, vol. 1, pp. 295–309, 2020.
- [18] D. Comminiello, M. Scarpiniti, L. A. Azpicueta-Ruiz, J. Arenas-García, and A. Uncini, “Functional link adaptive filters for nonlinear acoustic echo cancellation,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 7, pp. 1502–1512, Jul. 2013.
- [19] M. Scarpiniti, D. Comminiello, R. Parisi, and A. Uncini, “Nonlinear spline adaptive filtering,” *Signal Process.*, vol. 93, no. 4, p. 772–783, Apr. 2013.

- [20] M. Scarpiniti, D. Comminiello, R. Parisi, and A. Uncini, "Hammerstein uniform cubic spline adaptive filters: Learning and convergence properties," *Signal Process.*, vol. 100, p. 112–123, Jul. 2014.
- [21] M. Scarpiniti, D. Comminiello, R. Parisi, and A. Uncini, "Novel cascade spline architectures for the identification of nonlinear systems," *IEEE Trans. Circuits Syst. I: Regular Papers*, vol. 62, no. 7, pp. 1825–1835, Jun. 2015.
- [22] P. P. Campo, D. Korpi, L. Anttila, and M. Valkama, "Nonlinear digital cancellation in full-duplex devices using spline-based hammerstein model," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–7.
- [23] P. Pascual Campo, L. Anttila, D. Korpi, and M. Valkama, "Cascaded spline-based models for complex nonlinear systems: Methods and applications," *IEEE Trans. Signal Process.*, vol. 69, pp. 370–384, Dec. 2020.
- [24] A. Gebhard, "Self-interference cancellation and rejection in FDD RF-transceivers," Ph.D. dissertation, Johannes Kepler University Linz, 2019.
- [25] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Self-interference cancellation with nonlinear distortion suppression for full-duplex systems," in *Proc. Asilomar Conf. Signals, Systems and Computers*, 2013, pp. 1199–1203.
- [26] C. Mollén, U. Gustavsson, T. Eriksson, and E. G. Larsson, "Impact of spatial filtering on distortion from low-noise amplifiers in massive MIMO base stations," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6050–6067, Dec. 2018.
- [27] C. Motz, T. Paireder, and M. Huemer, "Modulated spur interference cancellation for LTE-A/5G transceivers: A system level analysis," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC Spring)*, May 2020.
- [28] J. F. Epperson, "On the runge example," *The Amer. Math. Monthly*, vol. 94, no. 4, pp. 329–341, 1987.
- [29] C. de Boor, *A Practical Guide to Splines*. New York, NY, USA: Springer, 2001.
- [30] L. Schumaker, *Spline Functions: Basic Theory*, 3rd ed., Ser. Cambridge Mathematical Library. Cambridge, U.K.: Cambridge Univ. Press, 2007.
- [31] E. Catmull and R. Rom, "A class of local interpolating splines," in *Comput. Aided Geometric Des.*, R. E. Barnhill and R. F. Riesenfeld, Eds. Academic Press, 1974, pp. 317–326.
- [32] Y. Gu, J. Jin, and S. Mei, " l_0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 774–777, Sep. 2009.
- [33] D. C. v. Grünigen, *Digitale Signalverarbeitung: Mit Einer Einführung in Die Kontinuierlichen Signale Und Systeme*, 5th ed. Munich, Germany: Hanser, 2014 (in German).
- [34] P. Dhrymes, *Mathematics for Econometrics*, 4th ed. New York, NY, USA: Springer, 2013.
- [35] R. Remmert and R. Burckel, *Theory of Complex Functions, Ser. Graduate Texts in Mathematics*. New York, NY, USA: Springer, 1991.
- [36] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proc. F - Commun., Radar Signal Process.*, vol. 130, no. 1, pp. 11–16, Feb. 1983.
- [37] A. I. Hanna and D. P. Mandic, "A fully adaptive normalized nonlinear gradient descent algorithm for complex-valued nonlinear adaptive filters," *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2540–2549, Oct. 2003.
- [38] P. S. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 2nd ed. New York, NY, USA: Springer, 2002.
- [39] C. Motz, T. Paireder, and M. Huemer, "Improving digital interference cancellation in LTE-A/5G-transceivers by statistical modeling," in *Proc. 54th Asilomar Conf. Signals, Syst., Comput.*, 2020.
- [40] S. F. Obermann and M. J. Flynn, "Division algorithms and implementations," *IEEE Trans. Comput.*, vol. 46, no. 8, pp. 833–854, Aug. 1997.
- [41] V. Kober, "Fast algorithms for the computation of sliding discrete sinusoidal transforms," *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1704–1710, Jun. 2004.
- [42] Y. Engel, S. Mannor, and R. Meir, "The kernel recursive least-squares algorithm," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2275–2285, Jul. 2004.
- [43] R. Boloix-Tortosa, J. J. Murillo-Fuentes, and S. A. Tsaftaris, "The generalized complex kernel least-mean-square algorithm," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5213–5222, Oct. 2019.
- [44] S. Glock, J. Rascher, B. Sogel, T. Ussmueller, J. Mueller, and R. Weigel, "A memoryless semi-physical power amplifier behavioral model based on the correlation between AM-AM and AM-PM distortions," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 6, pp. 1826–1835, Jun. 2015.
- [45] C. Rapp, "Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting signal," *ESA Special Publication*, vol. 332, pp. 179–184, 1991.
- [46] 3GPP, "Evolved universal terrestrial radio access (E-UTRA); physical channels and modulation," *3rd Generation Partnership Project (3GPP)*, Technical Specification (TS) 36. 211, 04 2017, version 14.2.0.
- [47] T. Paireder, C. Motz, O. Lang, and M. Huemer, "Time-delay estimation for self-interference cancellation in LTE-A/5G transceivers," in *Proc. Austrochip Workshop Microelectron.* (Austrochip), 2019, pp. 21–28.



Thomas Paireder (Graduate Student Member, IEEE) received the bachelor's degree (Hons.) in information electronics and the master's degree (Hons.) in electronics and information technology from Johannes Kepler University Linz (JKU), Linz, Austria, in 2016 and 2018, respectively. His master's thesis was focused on signal processing in cooperation with the Research Center for Nondestructive Testing GmbH. In the thesis, he implemented a real-time processing system for laser-ultrasonic signals. He is currently working toward the Ph.D. degree with the Institute of Signal Processing, JKU. His topic is receiver interference cancellation by means of adaptive signal processing methods.



Christian Motz (Graduate Student Member, IEEE) was born in Vöcklabruck, Austria, in 1990. He received the bachelor's degree (Hons.) in hardware software design and the master's degree (Hons.) in embedded systems design from the School of Informatics, Communications and Media, University of Applied Sciences Upper Austria, Hagenberg, Austria, in 2013 and 2015, respectively. His master's thesis was focused on pattern matching in cooperation with the Software Competence Center Hagenberg GmbH, Hagenberg, Austria. He is currently working toward the Ph.D. degree with the Institute of Signal Processing, Johannes Kepler University Linz (JKU), Linz, Austria, in cooperation with DMCE GmbH (Intel Linz), Linz, Austria.

Since 2017, he has been a Member of the Institute of Signal Processing, JKU, and the CD Laboratory for Digitally Assisted RF Transceivers for Future Mobile Communications, where he is involved in receiver interference cancellation by means of adaptive signal processing methods.



Mario Huemer (Senior Member, IEEE) received the Dipl.-Ing. and Dr. Techn. degrees from the Johannes Kepler University (JKU) Linz, Austria, in 1996 and 1999, respectively. After holding positions in industry and academia, he was an Associate Professor with the University of Erlangen-Nuremberg, Erlangen, Germany, from 2004 to 2007, and a Full Professor with Klagenfurt University, Klagenfurt, Austria, from 2007 to 2013. Since September 2013, he has been heading the Institute of Signal Processing, JKU, as a Full Professor, and since 2017 he has been the Co-

Head of the Christian Doppler Laboratory for Digitally Assisted RF Transceivers for Future Mobile Communications. He has authored or coauthored more than 270 scientific papers in his research fields, which include statistical and adaptive signal processing, signal processing architectures, mixed signal processing with applications in information and communications engineering, radio frequency transceivers for communications, and radar, sensor and biomedical signal processing. From 2009 to 2015, he was a Member of the Editorial Board of the *International Journal of Electronics and Communications* (AEU), and from 2017 to 2019, he was an Associate Editor for the *IEEE SIGNAL PROCESSING LETTERS*. He was the recipient of the dissertation awards of the German Society of Information Technology (ITG) and the Austrian Society of Information and Communications Technology (GIT), the Austrian Kardinal Innitzer Award in natural sciences, and the German ITG Award.