ABSTRACT
A novel background calibration technique for Time-Interleaved Analog-to-Digital Converters (TI-ADCs) is presented in this paper. This technique is applicable to equalized digital communication receivers. As shown in the literature, in a digital receiver it is possible to treat the TI-ADC errors as part of the communication channel and take advantage of the adaptive equalizer to compensate them. Therefore calibration becomes an integral part of channel equalization. No special purpose analog or digital calibration blocks or algorithms are required. However, there is a large class of receivers where the equalization technique cannot be directly applied because other signal processing blocks are located between the TI-ADC and the equalizer. The technique presented here generalizes earlier works to this class of receivers. The error backpropagation algorithm, traditionally used in machine learning, is applied to the error computed at the receiver slicer and used to adapt an auxiliary equalizer adjacent to the TI-ADC, called the Compensation Equalizer (CE). Simulations using a dual polarization optical coherent receiver model demonstrate accurate and robust mismatch compensation across different application scenarios. Several Quadrature Amplitude Modulation (QAM) schemes are tested in simulations and experimentally. Measurements on an emulation platform which includes an 8 bit, 4 GS/s TI-ADC prototype chip fabricated in 130nm CMOS technology, show an almost ideal mitigation of the impact of the mismatches on the receiver performance when 64-QAM and 256-QAM schemes are tested.

INDEX TERMS
Background calibration, error backpropagation, optical coherent receiver, TI-ADC, TI-ADC mismatch calibration.

I. INTRODUCTION
This paper proposes a novel background calibration technique for Time-Interleaved Analog-to-Digital Converters (TI-ADCs) used in equalized digital communication receivers. It generalizes a previously proposed technique [1], [2], [3]. Current and emerging digital receivers for ultra high-speed communication systems [4], [5], [6], [7], [8], [9], [10] require large bandwidth, high sampling rate ADCs. Please see [11] and references therein for a recent review of interleaving techniques for high speed data converters.

The TI-ADC has been the technique predominantly used to meet the demanding sampling rate and bandwidth requirements of high-speed transceivers [12], [13]. The performance of TI-ADCs is affected by mismatches among the interleaves [14], [15]. Mismatches of sampling time, gain, bandwidth, as well as DC offset, are the most common impairments. Many calibration techniques have been proposed in the literature. Please see [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41] and references therein for a thorough review and discussion.

Calibration techniques for general purpose TI-ADCs in general require dedicated calibration blocks and algorithms.
On the other hand, several authors [1], [2], [3] have shown that in the special case of an equalized digital receiver, it is possible to treat the TI-ADC errors as integral part of the communication channel and take advantage of the already existing adaptive equalizer to compensate them. Therefore calibration becomes an integral part of channel equalization. No special purpose analog or digital calibration blocks or algorithms are required. Equalizer-based compensation can compensate static as well as frequency-dependent errors such as bandwidth limitations and bandwidth or frequency response mismatches among the interleave of the ADC. Because the equalizer is adaptive, it also compensates time-dependent effects such as those caused by temperature and voltage variations, or by aging. Therefore, equalization becomes the TI-ADC compensation technique of choice in digital communication receivers.

Tsai et al. [1] provide a thorough description of the equalization technique and its advantages. However, there is a large class of receivers where this technique cannot be directly applied because other signal processing blocks are located between the TI-ADC and the equalizer. A block diagram of a typical receiver for high-speed digital communications is shown in Fig. 1.

An effective compensation of the TI-ADCs errors has been achieved in the referenced works [1], [2], [3] because the main receiver equalizer, or Feedforward Equalizer (FFE) is immediately located after the TI-ADC (in other words, the Signal Pre-Processing block of Fig. 1 is not present). Hence, the FFE can access and directly compensate the impairments of the different interleave. Also, the slicer error carries information about the impairments of the individual interleave and therefore the FFE adaptation algorithm can drive its coefficients to a solution that jointly compensates the channel and the TI-ADC impairments. Unfortunately, the application of Tsai’s technique to most types of receivers (e.g., for coherent optical communications) has been limited by the presence of signal pre-processing blocks (e.g., Timing Recovery (TR), Carrier Recovery (CR), or Bulk Chromatic Dispersion Equalizer (BCD)). These blocks cause signal components associated with different interleave of the TI-ADC to be combined in a way that makes the use of the FFE unsuitable to compensate them.

The main contribution of this work is a new background technique that overcomes the aforementioned limitations, and is especially well suited for complex digital receivers. A summary description of this technique was given in [42]. The basic idea consists in the use of an auxiliary, low complexity adaptive equalizer, called the Compensation Equalizer (CE), to compensate the mismatches of the TI-ADC. Although the CE solves the compensation problem, there is still a problem with the adaptation of the CE, because slicer error components associated with different interleave are also combined by the signal pre-processing blocks. Thus, the slicer error is not directly applicable to adapt the CE. To solve the adaptation problem, in this work we propose to adapt the CE using a post processed version of the error at the slicer of the receiver. The post processing is based on the backpropagation algorithm [43], widely used in machine learning applications [44]. Its main characteristic is that, in a multi-stage processing chain where several cascaded blocks have adaptive parameters, it is able to determine the contribution to the error generated by each one of these blocks and their associated parameters for all the stages. Backpropagation is used in combination with the Stochastic Gradient Algorithm (SGD) to adjust the coefficients of the CE in order to minimize the slicer Mean Squared Error (MSE). The use of the CE in combination with the backpropagation algorithm results in robust, fast converging background calibration. As we shall show, this proposal is not limited to the compensation of individual TI-ADCs (which is the case for most calibration techniques), but it extends itself to the entire receiver Analog Front End (AFE), enabling the compensation of impairments such as time skew, quadrature, and amplitude errors between the in-phase and the quadrature components of the signal in a receiver based on Phase Modulation (PM) or Quadrature Amplitude Modulation (QAM).

Because ultrafast adaptation is usually not needed, the backpropagation algorithm can be implemented in a highly subsampled hardware block which does not require parallel processing. Therefore, the implementation complexity of the proposed technique is low, as will be discussed in detail. Although the technique presented here is general and can be used in digital receivers for different applications, the primary example in this paper is a receiver for coherent optical communications. State of the art coherent optical receivers operate at symbol rates around 96 Giga-Baud (Gb/s) and require ADC sampling rates close to 150 GS/s and bandwidths of about 50 GHz. In the near future symbol rates will increase to 128–150 Gbd or higher, requiring bandwidths in the range of 65–75 GHz and sampling rates in the 200–250 GS/s range. High-order QAM schemes (e.g., 64-QAM, 256-QAM and higher) will be deployed to increase spectral efficiency [45]. High-order modulation schemes increase the resolution and overall performance requirements on the ADC. The benefits of the proposed technique are experimentally verified using 64-QAM and 256-QAM schemes.

The rest of this paper is organized as follows. Section II lists the requirements of calibration techniques suitable for digital receivers. These requirements set this application apart from more generic applications. Section II also compares the technique proposed here with other state-of-the-art techniques in the light of said requirements. Section III presents a discrete time model of the TI-ADC system in a Dual-Polarization (DP) optical coherent receiver. The error backpropagation based adaptive CE is introduced in Section IV.
Simulation results are discussed in Section V, while the experimental evaluation is presented in Section VI. The hardware complexity of the proposed scheme is discussed in Section VII and conclusions are drawn in Section VIII.

II. COMPENSATION REQUIREMENTS AND COMPARISON WITH THE STATE OF THE ART

The main requirements a calibration technique applicable to digital communication receivers should meet are the following:

- **Criterion 1: Background and continuously adaptive operation.** TI-ADC calibration should be transparent to the receiver operation, even in the face of variations in the TI-ADC parameters resulting from temperature, voltage, or other environmental changes. In most communications systems it is not possible to bring the link down to recalibrate the ADC. Foreground calibration during the system startup would be acceptable only if recalibration is guaranteed not to be required during normal operation, a condition difficult to meet in high speed systems using large constellations such as 64-QAM and 256-QAM due to their sensitivity to variations of AFE parameters.

- **Criterion 2: Joint compensation of the impairments of all TI-ADCs in the AFE.** Besides timing skew, gain, and offset mismatches, frequency-dependent effects such as bandwidth and frequency response mismatches should be addressed by the calibration technique. In addition to the impairments of individual TI-ADCs, in QAM receivers it is necessary to compensate errors affecting the in-phase and the quadrature components of the received signal, such as time skew (I/Q skew), quadrature, and gain errors. Although in principle it is possible to compensate these impairments with algorithms independent of the TI-ADC calibration, joint compensation is highly desirable because of its lower complexity and because it enables the global optimization of the receiver Signal-to-Noise Ratio (SNR).

- **Criterion 3: Global optimization of the receiver SNR.** Algorithms used in communication receivers, such as equalization, TR, CR, etc., usually seek to maximize the SNR (or, equivalently, minimize the MSE) at the decision point (the slicer). This criterion, under very general conditions, leads to the minimization of the Bit Error Rate (BER). A TI-ADC calibration algorithm that seeks to optimize the same criterion leads to optimal receiver performance, a condition that cannot be guaranteed with calibration algorithms based on heuristic criteria.

- **Criterion 4: No need for additional constraints on the statistical properties of the signal beyond those required by a receiver with ideal data converters.** Convergence of receiver algorithms usually can be guaranteed under very general conditions. It is undesirable to impose additional conditions on the signal only to ensure the correct operation of the TI-ADC calibration algorithms.

The technique proposed in this paper meets all the above requirements. In the following we compare it to some of the state of the art calibration techniques presented in the technical literature, with focus on the communications receiver applications. The following is a broad categorization of the most important techniques described in the literature and their comparison with the one proposed in this paper:

- **Group 1 (G1):** Techniques based on the autocorrelation of the quantized signal [21], [22], [23], [24], [25]. These techniques are well suited to estimating the sampling time errors, but do not provide information on frequency response mismatches or mismatches affecting different TI-ADCs in the AFE of a QAM receiver. This group of techniques satisfies Criterion 1 but not Criteria 2, 3, and 4.

- **Group 2 (G2):** Techniques based on statistical properties of the quantized signal [29], [30], [31], [32]. Mismatches of a single TI-ADC are estimated by calculations derived from sub-ADC outputs (e.g., using histograms). Similar to Group 1, Group 2 satisfies Criterion 1, but not Criteria 2, 3, and 4.

- **Group 3 (G3):** Techniques based on a reference channel [24], [25], [26], [27], [28], [29], [30]. Auxiliary hardware (sometimes operating at the sampling rate of the TI-ADC [29], [30]) is required to provide a reference or to enable the estimation. This group may satisfy Criteria 1 and 4 but not Criteria 2 and 3.

- **Group 4 (G4):** Techniques based on dither injection [33], [34], [35], [36]. Dither injection techniques are based on the addition of a known signal to the sample that is being quantized in order to estimate the calibration parameters of an individual TI-ADC. This group may meet Criteria 1 and 4 but not Criteria 2 and 3.

- **Group 5 (G5):** Techniques based on Machine Learning (ML) algorithms [37], [38], [39]. These techniques share with the one proposed in this paper the idea of using ML algorithms to estimate all the mismatches in a TI-ADC, although in substantially different ways. However, in the referenced works the calibration is essentially done in foreground for individual TI-ADCs, which does not meet Criteria 1 and 3.

- **Group 6 (G6):** Techniques based on matrix inversion [18], [19], [20], [21] combined with Taylor approximations or correlation-based techniques to estimate sampling time errors of a TI-ADC. This group may meet Criterion 1 but not Criteria 2, 3, and 4.

Table 1 summarizes the comparison of the above techniques with the one proposed in this paper on the basis of the Criteria 1 through 4 listed above.

It is important to notice that communications applications of TI-ADCs enjoy an important advantage over more general applications. This advantage is the availability of the global optimality criterion referred to as Criterion 3 above, in other words, the maximization of the SNR at the slicer. In general, applications other than digital communications receivers lack...
a global criterion such as the slicer SNR, whose optimization can be exploited to compensate the impairments of the TI-ADC, or more generally, of the AFE. Therefore, TI-ADCs for general purpose applications, unlike those used in digital communications receivers, often depend on heuristic methods to estimate the errors. In this sense, digital communications receivers enjoy a unique advantage over other applications, which should not be neglected in favor of techniques based on heuristic methods.

### III. SYSTEM MODEL OF HIGH-SPEED DIGITAL RECEIVERS BASED ON TI-ADC

Communication channels of interest in this work include, among others: (i) wireline, (ii) wireless, and (iii) optical. The primary example of application of the backpropagation-based compensation technique presented in the next sections is a DP coherent optical receiver [4], [5], [6]. However, it can be used in any high-speed digital receiver with minor modifications.

#### A. COHERENT OPTICAL CHANNEL

Dual-polarization coherent optical transceivers exploit the two polarizations of light in propagation over single-mode fibers (SMF) to transmit two independent signals. Chromatic Dispersion (CD) and Polarization-Mode Dispersion (PMD) are the most important linear effects experienced in a SMF link. The optical channel with CD and PMD can be modeled as a $2 \times 2$ Multiple-Input Multiple-Output (MIMO) complex-valued channel [46]. Thus, the input and output of the MIMO channel correspond to the transmitted and received signals in each polarization.

Let $P(\omega)$ be the Fourier transform of the total pulse, $p(t)$, which includes the transmit and receive filters. In the presence of CD and PMD, the frequency domain channel transfer matrix can be expressed as

$$
H(\omega) = H_{\text{CD}}(\omega, L)P(\omega)J(\omega),
$$

where $\omega$ is the angular frequency, $L$ is the fiber length, $H_{\text{CD}}(\omega, L)$ models CD, and $J(\omega)$ is the well-known Jones matrix defined by

$$
J(\omega) = \begin{bmatrix}
U(\omega) & W(\omega)
\end{bmatrix},
$$

where $^*$ denotes complex conjugate. Matrix $J(\omega)$ is unitary (i.e., $\det(J(\omega)) = |U(\omega)|^2 + |W(\omega)|^2 = 1$, $\forall \omega$) and models the effects of the PMD. Chromatic dispersion is modeled as:

$$
H_{\text{CD}}(\omega, L) \approx \exp \left( \frac{1}{2} \beta_2 L \omega^2 \right),
$$

where $\beta_2$ is related to the dispersion parameter $D = \frac{2\pi c}{\lambda^2} \beta_2$, with $c$ and $\lambda$ being the speed of light and the wavelength, respectively.

We highlight that a coherent receiver can compensate for PMD and CD impairments without noise enhancement or signal to noise ratio penalty [46].

#### B. DUAL-POLARIZATION COHERENT OPTICAL SYSTEM

A block diagram of the Optical Front End (OFE) and the AFE for a DP coherent receiver is shown in Fig. 2. The optical input signal is decomposed by the OFE into four signals, the in-phase and quadrature ($I$/$Q$) components of the horizontal and vertical ($H$/$V$) polarizations. Photodetectors are used to convert the optical signals to photocurrents which are amplified by Trans-Impedance Amplifiers (TIAs). Then, the AFE acquires the electrical signals and translates them to the digital domain.

Let $\hat{d}_k^{(H)} = d_k^{(1)} + jd_k^{(2)}$ and $\hat{d}_k^{(V)} = d_k^{(3)} + jd_k^{(4)}$ be the $k$-th QAM symbols to be transmitted in the horizontal ($H$) and vertical ($V$) polarizations, respectively. Notice that $d_k^{(i)}$ is the $i$-th QAM symbol sequence (e.g., $d_k^{(i)} \in \{\pm 1, \pm 3\}$ for 16-QAM). Complex symbols $\hat{d}_k^{(H)}$ and $\hat{d}_k^{(V)}$ are assumed to be independent and identically distributed such

$$
E \left\{ \hat{d}_k^{(H)} \hat{d}_m^{(V)\ast} \right\} = \delta_{m-k} \delta_k
$$

where $\delta_k$ being the discrete time impulse function. We also define $s^{(H)}(t)$ and $s^{(V)}(t)$ as the complex signals at the receiver input for polarizations $H$ and $V$, respectively. Then, the noise-free complex electrical signals provided by the optical demodulator can be expressed as [46]

$$
\begin{align*}
\hat{s}^{(H)}(t) &= s^{(1)}(t) + j s^{(2)}(t) \\
&= e^{j \phi^{(H)}(t)} \sum_k \hat{d}_k^{(H)} \hat{h}_{1,1}(t-kT) + \hat{d}_k^{(V)} \hat{h}_{1,2}(t-kT), \\
\hat{s}^{(V)}(t) &= s^{(3)}(t) + j s^{(4)}(t) \\
&= e^{j \phi^{(V)}(t)} \sum_k \hat{d}_k^{(H)} \hat{h}_{2,1}(t-kT) + \hat{d}_k^{(V)} \hat{h}_{2,2}(t-kT),
\end{align*}
$$

where $D$ is expressed in ps/(nm*km), representing the differential delay, or time spreading (in ps), for a source with a spectral width of 1 nm traveling on 1 km of fiber. It depends on the fiber type, and in the absence of equalization it would limit the error-free bit rate or the transmission distance.

---

**TABLE 1. Summary and comparison with other state-of-the-art techniques.**

| Criterion     | G1 | G2 | G3 | G4 | G5 | G6 | This Work |
|---------------|----|----|----|----|----|----|-----------|
| 1             | YES| YES| YES| YES| NO | YES| YES       |
| 2             | NO | NO | NO | NO | NO | NO | YES       |
| 3             | NO | NO | NO | NO | NO | NO | YES       |
| 4             | NO | NO | YES| YES| NO | NO | YES       |

Fig. 2. Optical/analog front-end for a TI-ADC-based coherent optical receiver. The optical signal is split into four electrical lanes that are converted by a TI-ADC. PBS: Polarization Beam Splitter; LO: Local Oscillator; 90° Hyb: 90° hybrid coupler.
where $1/T$ is the symbol rate, $\phi^{(H)}(t)$ and $\phi^{(V)}(t)$ are the carrier phase errors on each polarization, $s^{(1)}(t)$ and $s^{(2)}(t)$ ($s^{(3)}(t)$ and $s^{(4)}(t)$) are the in-phase and quadrature components of $\tilde{s}^{(H)}(t)$ ($\tilde{s}^{(V)}(t)$), respectively, while

$$
\tilde{h}_{11}(t) = F^{-1}[H_{CD}(\omega, L)P(\omega)U(\omega)],
$$

$$
\tilde{h}_{12}(t) = F^{-1}[H_{CD}(\omega, L)P(\omega)V(\omega)],
$$

$$
\tilde{h}_{21}(t) = F^{-1}[-H_{CD}(\omega, L)P(\omega)\tilde{V}^*(\omega)],
$$

$$
\tilde{h}_{22}(t) = F^{-1}[H_{CD}(\omega, L)P(\omega)\tilde{U}^*(\omega)],
$$

where $F^{-1}[,]$ denotes the inverse Fourier transform.

The four electrical signals $s^{(1)}(t)$, $s^{(2)}(t)$, $s^{(3)}(t)$, and $s^{(4)}(t)$ are sampled by four different TI-ADC’s. Digital receivers with a certain degree of oversampling (e.g., $T_s = \frac{T}{2}$ where $T_s$ is the sampling symbol period) are used to compensate the dispersion experienced in optical links [46]. Next, we formulate the model of the optical channel, including a TI-ADC system affected by mismatches, used in the remainder of this paper.

### C. AFE AND TI-ADC DISCRETE-TIME MODEL

A discrete-time model for the AFE and the TI-ADC system of Fig. 2 with their impairments is introduced in this section. A simplified representation of the analog path for one component is shown in Fig. 3. The responses of the electrical interconnections between the optical demodulator and the TIA, the TIA response itself, and any other components in the signal path up to a TI-ADC system are represented with a filter with impulse response $c^{(i)}(t)$. Time delay or skew between components $I$ and $Q$ of a given polarization is caused by mismatches between $c^{(I)}(t)$ and $c^{(Q)}(t)$ or $c^{(3)}(t)$ and $c^{(4)}(t)$ for polarization $H$ ($V$), and degrades the receiver performance. As we shall show, the proposed background calibration algorithm is able to compensate not only the mismatches of the TI-ADC, but also to the IQ skew, the quadrature and amplitude errors, and other impairments among the signal paths.

Blocks $f_{m}^{(i)}(t)$ with $i = 1, 2, 3, 4$ and $m = 0, \ldots, M-1$, model the independent responses of the $M$ Track and Hold (T&H) circuits in an $M$-channel TI-ADC system. Each one of the $M$ interleaved channels is sampled every $M/f_s = M T_s$ seconds with a proper sampling phase. The digitized high-frequency samples can be written as (see Appendix)

$$
y^{(i)}[n] = \sum_{l=0}^{L-1} \hat{h}^{(i)}[l]s^{(i)}[n - l] + \hat{e}^{(i)}[n] + q^{(i)}[n]
$$

where $\hat{h}^{(i)}[l]$ is the impulse response of a time-varying filter, which is an $M$-periodic sequence such $\hat{h}^{(i)}[l] = \hat{h}^{(i)}[l+M]$ defined by (36), and $q^{(i)}[n]$ is the quantization noise.

### D. COMPENSATION OF AFE MISMATCH AND TI-ADC IMPAIRMENTS

Errors and mismatches of the TI-ADC can be compensated by using digital finite impulse response (FIR) filters applied to each interleaved branch. In the case of a communication receiver, the digitized signal could be applied to a time-varying equalizer immediately following the TI-ADC (see [1] for more details). The practical implementation of this periodically time-varying equalizer is briefly addressed in Section IV-A, and in more detail in [3].

Similarly to what was done in previous works [1], [2], [3], [26], in the backpropagation-based architecture introduced in this paper we propose to adaptively compensate the TI-ADC mismatch, after the mitigation of the offset, using a filter with an $M$-periodic time-varying impulse response:

$$
x^{(i)}[n] = \sum_{l=0}^{L-1} \hat{g}^{(i)}[l]w^{(i)}[n - l], 
$$

where $\hat{g}^{(i)}[l]$ is the $M$-periodic time-varying impulse response of the compensation filter (i.e., $\hat{g}^{(i)}[l] = \hat{g}^{(i)}[l+M]$), $L_5$ is the number of taps of the compensation filters, and $w^{(i)}[n]$ is the DC offset-free signal given by

$$
w^{(i)}[n] = y^{(i)}[n] - \hat{e}^{(i)}[n],
$$

with $\hat{e}^{(i)}[n]$ being the $M$-periodic offset sequence estimation. The combination of the offset compensation blocks and the compensation filters $\hat{g}^{(i)}[l]$ constitutes the CE (see Fig. 4).

The adaptation algorithm of the CE as proposed in [1] or [2] cannot be implemented in coherent optical communication receivers. This is because of the presence of several signal pre-processing blocks placed between the CE and the slicers, such as the BCD or the MIMO FFE that compensates

\[ \text{FIGURE 3. Analog front-end model for the signal component $s^{(i)}(t)$ with $i = 1, 2, 3, 4$ in a TI-ADC-based dual-channel coherent optical receiver.} \]
PMD [4]. Thus, since the slicer errors are not available at the outputs of the CE, a proper strategy has to be defined to adapt the CE response. In the next section, we apply the backpropagation technique to adapt the CE coefficients.

IV. ERROR-BACKPROPAGATION-BASED COMPENSATION OF AFE AND TI-ADC IMPAIRMENTS IN DIGITAL RECEIVERS

Figure 4 depicts a block diagram of the AFE+TI-ADC in a DP optical coherent receiver with the adaptive compensation equalizer, including four instances of the real filter as defined by (12).²

The Digital Signal Processing (DSP) block of Fig. 4, performs the main receiver functions, operating with samples every $T_s$ seconds. In summary, some of the most important DSP algorithms used in optical coherent receivers are the BCD, the MIMO FFE, TR from the received symbols, the Fine CR (FCR) to compensate the carrier phase and frequency offset.³ Readers interested in more details on optical coherent receivers can see [4], [5], [47] and references therein.

A. PARALLEL IMPLEMENTATION OF THE COMPENSATION EQUALIZER

Before explaining the adaptation of the CE coefficients $g_m^{(i)}(l)$, it is important to highlight that no additional complexity is incurred in the CE by introducing $M$ independent time-interleaved responses when they are implemented in a parallel architecture. Let $g_m^{(i)}(l)$ with $i = 1, \cdots, 4$ be the filter impulse response $g_m^{(i)}(l)$ in one period defined as

$$ g_m^{(i)}(l) = g_m^{(i)}[l] = \sum_{n=0}^{M-1} m = 0, \cdots, M - 1, \quad (14) $$

where $l = 0, \cdots, L_g - 1$ and $n_0$ is an arbitrary time index multiple of $M$. Thus, notice that the CE in coherent receivers (see Fig. 4) comprises 4 sets of real valued FIR $g_m^{(i)}(l)$ with $i = 1, 2, 3, 4$, $m = 0, \cdots, M - 1$, and $l = 0, \cdots, L_g - 1$.

²The structure of the CE shown in Fig. 4 can be extended to include the compensation of the quadrature error of the optical demodulator. This will be addressed in a future work.

³Although the receiver DSP for wireline and wireless may include other algorithms, the technique presented here can be applied to them with minor modifications.

In high speed optical communication applications, the use of parallel implementations is mandatory. Typically, a parallelism factor on the order of 128 or higher is adopted. Furthermore, given the number of interleaves of the TI-ADC $M$, the parallelism factor $P$ can be selected to be a multiple of $M$, i.e., $P = q \times M$ with $q$ an integer. In this way, the different time multiplexed taps are located in fixed positions of the parallel implementation, and we do not incur significant additional complexity when compared to a filter with just one set of coefficients (see [2] for more details). The complexity of the resulting filter is similar to that of the I/Q-skew compensation filter already present in current coherent receivers [4]. Therefore, the typical skew correction filter can be replaced by the CE without adding significant penalties in area or power since the CE is also able to correct time skew.

B. ALL DIGITAL COMPENSATION ARCHITECTURE

The filter coefficients of the impulse response in (14) are adapted using the slicer error at the output of the receiver DSP block. Let $u_k^{(i)}$ be the equalized signal at the input of the slicer (see Fig. 4). The latter is a quantization device that makes the symbol decisions $\tilde{a}_k^{(i)}$ (e.g., $\tilde{a}_k^{(i)} \in \{±1, ±3\}$ for 16-QAM). Let $e_k^{(i)}$ be the slicer error defined as

$$ e_k^{(i)} = u_k^{(i)} - \tilde{a}_k^{(i)}, \quad i = 1, \cdots, 4. \quad (15) $$

As usual (e.g., see [48]), in the analysis we assume that there are no decision errors,⁴ and thus we use $\tilde{a}_k^{(i)}$ in place of $a_k^{(i)}$, i.e.,

$$ e_k^{(i)} = u_k^{(i)} - \tilde{a}_k^{(i)}, \quad i = 1, \cdots, 4. \quad (16) $$

Since the slicer operates at $1/T$ sampling rate, a subsampling of $T/T_s$ is needed after the receiver DSP block. Then, the total squared error at the slicer at time instant $k$ is defined as

$$ E_k = \sum_{i=1}^{4} |e_k^{(i)}|^2. \quad (17) $$

Let $E[\{E_k\}]$ be the MSE at the slicer with $E[.]$ denoting the expectation operator. In this work we iteratively adapt the real coefficients of the CE defined by (14) by using the Least Mean Squares (LMS) algorithm, in order to minimize the MSE at the slicer:

$$ g_m^{(i)}[p+1] = g_m^{(i)}[p] - \beta \nabla g_m^{(i)} E[\{E_k\}], \quad (18) $$

where $i = 1, \cdots, 4; m = 0, \cdots, M - 1; p$ denotes the number of iteration, $g_m^{(i)}$ is the $L_g$-dimensional coefficient vector at the $p$-th iteration given by

$$ g_m^{(i)} = [g_m^{(i)}[0], g_m^{(i)}[1], \cdots, g_m^{(i)}[L_g - 1]]^T, \quad (19) $$

⁴This assumption is justified by experience, which confirms that as long as the error rate is below ten percent or so, there is no appreciable effect on the equalizer operation [48]. This condition is verified in coherent optical communications, where the error rates before forward error correction are typically smaller than $4 \times 10^{-5}$.
gradient of the squared error (17) can be expressed as
\[ \nabla_{\hat{w}_{o}} E\{\epsilon_k^2\} \]

where \( \alpha \) is a certain constant, \( \mathbf{w}[n] \) is a vector with \( L_e \) input samples of the CE, i.e.,
\[ \mathbf{w}^{(i)}[n] = \left[ w^{(i)}[n], w^{(i)}[n-1], \ldots, w^{(i)}[n-L_e+1] \right]^T, \]

while \( \hat{\epsilon}^{(i)}[n] \) is the backpropagated error given by
\[ \hat{\epsilon}^{(i)}[n] = \sum_{j=1}^{4} \sum_{l=0}^{L_e-1} \Gamma_{1+2l}[j] e^{(j)}[n+l], \]

with \( e^{(j)}[n] \) being the oversampled slicer error generated from the slicer error at the baud-rate \( e_k^o \) in (16) as
\[ e^{(j)}[n] = \begin{cases} e_k^o \sqrt{2} & \text{if } n = 0, \pm 2, \pm 4, \ldots, \\ 0 & \text{otherwise} \end{cases} \]

Then, we can derive an all-digital compensation scheme using an adaptive CE with coefficients updated as
\[ \mathbf{g}_{m,p+1}^{(i)} = \mathbf{g}_{m,p}^{(i)} - \mu \nabla_{\hat{w}_{o}} E\{\epsilon_k^2\}, \]

where \( \mu = \alpha \beta \) is the adaptation step-size. Moreover, it is possible to estimate the DC offsets in the input samples, using the backpropagated error defined in (24), as follows
\[ \hat{o}_{m,p+1}^{(i)} = \hat{o}_{m,p}^{(i)} - \mu_o \hat{\epsilon}^{(i)}[n+m], \quad m = 0, \ldots, M-1, \]

where \( \hat{o}_{m,p}^{(i)} \) is the DC offset sequence estimation in one period (see (13)) at the \( p \)-th iteration, and \( \mu_o \) is the step-size of the DC offset estimator.

In order to avoid possible instability due to competition between the CE and any adaptive DSP blocks in \( \Gamma_n^{[j]}[l] \) (e.g., the MIMO FFE), an adaptation constraint must be included. This can be achieved by limiting one of the \( 4M \) sets of the CE coefficients to only be a time delay line. For example, \( g_0^{(i)}[l] = \delta_l l_d \) where \( l = 0, \ldots, L_e-1 \) and \( l_d = \frac{L_e+1}{2} \) (\( L_e \) is assumed odd).

Since channel impairments change slowly over time, the coefficient updates given by (26) and (27) do not need to operate at full rate, and subsampling can be applied. The latter allows implementation complexity to be significantly reduced. Additional complexity reduction is enabled by: 1) strobing the algorithms once they have converged, and/or 2) implementing them in firmware in an embedded processor, typically available in coherent optical transceivers. Practical aspects of the hardware implementation will be discussed in Section VII.

D. DISCUSSION

In this section we address the convergence properties of the proposed calibration algorithm, using the traditional LMS algorithm as a reference. In particular, it is well known that convergence of the latter is not affected by local minima...
of the error surface where the gradient algorithm could get trapped, because the error surface is quadratic. The proposed system is equivalent to a traditional LMS adaptive filter [50] with the exception that the gradient is not calculated directly with the error at the output of the adaptive equalizer but with an error back-propagated through the DSP filters. Thus, similarly to the traditional LMS, the algorithm proposed here cannot suffer from convergence issues due to local minima of the error surface. However, as a result of the backpropagation, the error (and the resulting gradient) could suffer from signal-dependent fading. This is a complex problem in the case of traditional neural networks due to their nonlinearities [51]. However, in the case proposed here the backpropagation is performed on linear blocks according to the filtering equation (24) where $\Gamma_n^{g,i} [l]$ is the combined filter response of the DSP which reduces the error fading problem to the cases where the error is frequency filtered. In particular, in coherent receivers the filters involved are low-pass filters whose bandwidth is equivalent to that of the received signal. Thus the gradient does not present fading within the spectrum of interest but may present it at high frequencies. The latter is due to the fact that the receiver works with a sampling frequency higher than that of the signal of interest, leaving the high frequency components filtered by the FFE in open loop since they do not affect the error and the gradient. The above, if not taken into account, may lead to divergence of the $g_m^{(i)} [l]$ coefficients of the filters controlled by the LMS because not all degrees of freedom are under the control of the adaptive loop. The way to avoid this is to use a regularization in the adaptation. The most commonly used regularization in digital communications is tap-leakage [50] in which the cost function is:

$$J_k = E \left\{ \sum_i e_k^{(i)} \right\}^2 + \zeta \sum_i \sum_l |g_m^{(i)} [l]|^2$$  \hspace{1cm} (28)

which prevents an overflow of the coefficients by providing a compromise (controlled by the real constant $\zeta$) between the minimization of the mean square error and the interest of the filter coefficients. This results in modifying the adaptation equation (18) according to:

$$g_{m,p+1}^{(i)} [l] = (1 - \mu \zeta) g_{m,p}^{(i)} [l] - \mu \nabla_{g_{m,p}} E \{ \mathcal{E}_k \}$$  \hspace{1cm} (29)

where except for the leakage factor $(1 - \mu \zeta)$ the adaptation is equivalent to that of the traditional LMS.

**E. ALGORITHM SUMMARY**

In the following, we summarize the proposed error-backpropagation based background calibration algorithm. For simplicity, we assume that the DSP block after the CE is described by a generic time-invariant $4 \times 4$ MIMO equalizer, $\Gamma_n^{g,i} [l] = \Gamma^{g,i} [l], \forall n$ with $i, j = 1, \cdots, 4; l = 0, \cdots, L_T - 1$:

- **Step 1.** Initialize the taps of the CE at the first iteration $p = 0$ (see (19)): $g_{m,p}^{(i)} = [0, \cdots, 1, \cdots, 0]^T$ with $i = 1, \cdots, 4; m = 0, \cdots, M - 1$.
- **Step 2.** Calculate the outputs of the CE, $x^{(i)} [n]$ with $i = 1, \cdots, 4$ (see (21)).
- **Step 3.** Calculate the downsampled outputs of the generic DSP block, $u_k^{(i)}$ with $i = 1, \cdots, 4$ (see (20)).
- **Step 4.** Evaluate the slicer outputs, $\bar{u}_k^{(i)} = Q(u_k^{(i)})$ with $i = 1, \cdots, 4$, where $Q(.)$ is the quantization device operation of the slicer.
- **Step 5.** Calculate the slicer errors, $e_k^{(i)}$ with $i = 1, \cdots, 4$ (see (15)).
- **Step 6.** Compute the oversampled slicer errors, $e_k^{(i)} [n]$ with $i = 1, \cdots, 4$ (see (25)).
- **Step 7.** Compute the backpropagated errors, $\bar{e}_k^{(i)}$ with $i = 1, \cdots, 4$ (see (24)).
- **Step 8.** Evaluate the instantaneous gradient, $\nabla e_k^{(i)} \mathcal{E}_k$ with $i = 1, \cdots, 4; m = 0, \cdots, M - 1$ (see (22)).
- **Step 9.** Update the taps of the CE (see (26)) and set $p = p + 1$.
- **Step 10.** Go to Step 2.

Since the technique operates in background, steps 2 through 10 run continuously to enable tracking of parameter variations caused by temperature, voltage, aging, etc. This happens even after the algorithm reaches convergence. However, since ultrafast compensation is usually not needed, steps 7 to 10 can be implemented in a highly subsampled hardware block. Consequently, the implementation complexity and power can be reduced.

**V. SIMULATION RESULTS**

In this section the proposed backpropagation based mismatch compensation technique is tested using computer simulations. The simulation setup is shown in Fig. 6. The simulated parameters are summarized in Table 2. TI-ADC mismatches are modeled as Uniformly Distributed Random Variables (UDRV). The electrical analog path responses (33) are modeled by first-order low-pass filters with 3dB-bandwidth defined by

$$B_m^{(i)} = B_0 + \Delta B_m^{(i)}, \hspace{0.5cm} i = 1, 2, 3, 4; \hspace{0.5cm} m = 0, \cdots, M - 1,$$  \hspace{1cm} (30)

where $B_0$ is the nominal BW and $\Delta B_m^{(i)}$ is the BW mismatch. Sampling phase errors and I/Q time skew are modeled by Lagrange interpolation filters. The I/Q time skew of each polarization is evenly distributed between all corresponding components (see Fig. 6). Errors of the TI-ADC are modeled as detailed in Section III-C. In particular, time skews among the interleaves are modeled using Lagrange interpolation filters (not to be confused with those used to model the I/Q skews). We consider a DP optical coherent system with a 64-QAM modulation scheme, and a symbol rate of $1/T = 96$ Gbd. Raised cosine filters with roll-off factor 0.1 for transmit pulse shaping are simulated (i.e., the nominal BW of the channel filters is $B_0 = 1.1 \times \frac{96 \text{ GHz}}{2} \approx 53$ GHz).
TABLE 2. Parameters used in simulations. (UDRV: Uniformly distributed random variable. VFS: Full-scale voltage.)

| Parameter                          | Value                                                                 |
|-----------------------------------|-----------------------------------------------------------------------|
| Modulation                        | 64-QAM                                                                |
| Symbol Rate ($f_{BB} = 1/T$)       | 96 GBd                                                                |
| Receiver Oversampling Factor ($T/T_s$) | 2                                                                   |
| Fiber Length                       | 100 km                                                                |
| Differential Group Delay (DGD)     | 10 ps                                                                 |
| Second Order Pol. Mode Disp. (SOPMD) | 1000 ps$^2$                                                           |
| Speed of Rotation of the Pol. at the Tx | 2 kHz                                                               |
| Speed of Rotation of the Pol. at the Rx | 10 kHz                                                             |
| TI-ADC Resolution                 | 8 bit                                                                 |
| TI-ADC Sampling Rate (all interleaves) | 192 GS/s                                                           |
| Number of Interleaves of TI-ADC ($M$) | 16                                                                  |
| Number of Taps of CE ($L_g$)       | 7                                                                     |
| Roll-off Factor                    | 0.10                                                                  |
| Nominal BW of Analog Paths ($B_a$) (see (30)) | 53 GHz                                                               |
| Gain Errors (see (10)) - UDRV      | $\Delta \epsilon(m) \in [-0.15]$                                    |
| Sampling Phase Errors - UDRV       | $\Delta \epsilon_i^{(m)} \in [-0.075]/T$                              |
| Bandwidth Mismatches (see (30)) - UDRV | $\tau_H, \tau_V \in [-0.075]/T$                      |
| I/Q Time Skew - UDRV               | $\alpha_i^{(m)} \in [-0.025]/VFS$                                   |
| DC Offsets - UDRV                  |                                                                       |

The Optical SNR (OSNR) is set to that required to achieve a BER of $\sim 1 \times 10^{-3}$ (see [52], [53] for the definition of OSNR). The oversampling factor in the DSP blocks is $T/T_s = 2$. The fiber length is 100 km with 10 ps of Differential Group Delay (DGD) and 1000 ps$^2$ of Second-Order PMD (SOPMD). Rotations of the State of Polarization (SOP) of 2 kHz and 10 kHz are included at the transmitter and receiver, respectively. Please see [54] for a comprehensive description of the aforementioned optical channel parameters. TI-ADCs with 8-bit resolution, 192 GS/s sampling rate, and $M = 16$ are simulated. The number of taps of the digital compensation filters is $L_g = 7$.

A. MONTECARLO SIMULATIONS OF THE ADAPTIVE CE

Each MonteCarlo test consists of 500 cases where the impairment parameters are obtained from a UDRV random number generator. Figs. 7 and 8 show the histograms of the BER for the receiver with and without the CE in the presence of sampling phase errors, gain errors, I/Q time skew, and BW mismatches. Only one effect is exercised in each case. Results for sampling phase and gain errors uniformly distributed in the interval $\delta_i^{(m)} \in [-0.075]/T$ and $\Delta \epsilon_i^{(m)} \in [-0.15]$ (see (10)), respectively, are depicted in Fig. 7, whereas Fig. 8 shows results for random BW mismatches (see (30)) and I/Q time skews uniformly distributed in the interval $\Delta \delta_i^{(m)} \in [-0.075]/B_0$ and $\tau_H, \tau_V \in [-0.075]/T$, respectively. For all the cases considered the proposed compensation technique is able to mitigate the impact on the performance of the receiver of all the impairments when they are exercised separately.

An improvement in BER of one order of magnitude can be achieved with the proposed technique. In particular, notice that the serious impact on the receiver performance of the I/Q time skew values of Table 2 is essentially eliminated by the proposed CE with $L_g = 7$ taps.

BER histograms for the receiver with and without the CE in the presence of the combined effects are shown in Fig. 9. Results of 4000 cases with random gain errors, sampling phase errors, I/Q time skews, BW mismatches, and DC offsets as defined in Table 2, are presented. Fig. 9 also depicts the performance of the CE with $L_g = 13$ taps. Without CE, a severe degradation on the receiver performance as a consequence of the combined effects of the TI-ADC mismatches is observed. However, note that the CE is able to compensate the impact of all combined impairments improving the BER in some cases by almost 100 times. Moreover, note that a slight performance improvement can be achieved increasing the number of taps $L_g$ from 7 to 13.

In multi-gigabit transceivers, the impairments of the AFE and TI-ADCs change very slowly over time, as mentioned in Section IV-C. Hence, decimation can be applied since the coefficient updates given by (26) and (27) do not need to be made at full rate. In ultra-high-speed transceiver implementations (e.g., for optical coherent communication), block processing and frequency domain equalization based on the Fast Fourier Transform (FFT) are widely used [4]. Therefore, we propose to update the CE performing block decimation over the error samples. The procedure is detailed as follows. Let $N$ be the block size in samples to be used for implementing the EBP. Define $D_B$ as the block decimation factor. In this way, the CE is updated using only one block of $N$ consecutive samples of the oversampled slicer error (25) every $D_B$ blocks, respectively.
i.e.,

\[ e^{(i)kND_B + n}, \quad n = 0, 1, \ldots, N - 1, \quad \forall k \]  

(31)

with \( k \) integer. By using this approach, Fig. 10 shows an example of the temporal evolution of the BER in the presence of combined impairments according to Table 2 for different values of \( D_B \) with \( N = 8192 \). A moving average filter of length 10 has been used to process the instantaneous BER. Gear shifting is used to reduce the steady-state MSE and speedup the convergence of the algorithm. We highlight that the impact on the resulting BER is negligible when block decimation is applied. In summary, block decimation drastically reduces implementation complexity.

VI. EXPERIMENTAL RESULTS

We demonstrate the benefits of our proposal using a digital communication platform especially designed to evaluate TI-ADC mismatch calibration techniques. The platform allows the capabilities of our proposal to be evaluated in communication links using several different modulation schemes.

A. RECONFIGURABLE EXPERIMENTAL PLATFORM

A block diagram of the experimental setup is shown in Fig. 11. A high-performance Field-Programmable Gate Array (FPGA) [55] is used to generate the symbols to be transmitted. The FPGA is also in charge of collecting the samples from the ADC and sending them to the receiver DSP, which is implemented on a host computer. Multiple Pseudo-Random Binary Sequences (PRBSs) with configurable length and seed are generated in the FPGA. The amplitude of the symbols and the Additive White Gaussian Noise (AWGN) can be set through the coefficients \( G_S \) and \( G_N \), respectively. Then, we are able to evaluate different SNR scenarios. The symbol with added noise is sent to a commercial, 16-bit Digital-to-Analog Converter (DAC) board [56] using an LVDS interface. The DAC synthesizes the samples at 1/\( T \) = 1 GS/s. This sampling rate is adopted due to limitations on the FPGA and DAC clocks. The communication channel is modeled as a low-pass filter with a \(-3 \text{ dB} \) cut-off frequency of 650 MHz [57]. Figure 12 shows the measured eye diagrams at the input and output of the channel with Binary Phase Shift Keying (BPSK) modulation. Notice that significant ISI is added by the channel. Although not explicitly shown, the impact of the ISI is even more significant for the higher order modulations used in the experiments, such as 8-PAM/64-QAM and 16-PAM/256-QAM. This ISI is an important part of the experiment since it enables the verification of the backpropagation technique, as discussed later in this section. On the receiver side, the signal is acquired by the TI-ADC described in [58], operating at a sampling rate of 2 GS/s (i.e., an oversampling ratio of \( T/T_s = 2 \) is used in the DSP blocks). The clocks for both DAC and ADC are generated from a single 10 MHz clock reference. More details of the experimental platform as well as the fabricated TI-ADC can be found in [58], [59].

B. MEASUREMENTS

As explained before, the available experimental setup has one TI-ADC. Therefore, a suitable signal for the coherent receiver has to be assembled by combining four independent measurements. This is done by collecting one set of samples for each
signal component. For a particular component of the complex signal, the platform is configured with a unique PRBS length. Furthermore, the configuration of the delay cells is also changed for each component. The post-processing routine on the computer incorporates a low complexity CE with 4 sets of 15 independent coefficients for each component.

We focus on the calibration of sampling phase and gain errors. Although the DC offset mismatch causes severe degradation of the performance of both TI-ADC and receiver, we do not consider its calibration here since it is already calibrated on-chip [59].

First, we illustrate the effectiveness of our proposal considering only the impairments of the TI-ADC for 64-QAM and 256-QAM in a noiseless channel. The resulting constellations are shown in Fig. 13. In the absence of compensation, the mismatch among the interleaves is large enough to enlarge the constellation points considerably. For a 256-QAM (see Fig. 13(b)) the degradation is such that the symbols in the received constellation are not distinguishable. In all cases, a drastic improvement can be observed when the proposed compensation technique is used.

The comparison of the BER curves for the receiver with and without the proposed technique is shown in Fig. 14. The performance of the receiver is severely affected when the TI-ADC mismatch is not mitigated. A sampling phase error of 4% has been set for Fig. 14(a), whereas 1% is set for Fig. 14(b). Setting a larger sampling phase error for QAM-256 would incur in issues related to the convergence of the receiver. A considerably high SNR penalty of 3 dB is measured for a 64-QAM modulation at a BER of $10^{-3}$. A similar penalty can be observed for a 256-QAM scheme, although the mismatch in this case is much smaller than in the previous case. After enabling the proposed technique, the performance of the receiver is restored to almost replicate the case without mismatch. This result indicates that our proposal is able to nearly eliminate the receiver penalty introduced by the mismatches of the TI-ADC.

The spectrum comparison for a 972 MHz sinusoidal input is shown in Fig. 15. Samples from one of the emulated channels are collected before and after running the technique on the communication setup of Fig. 11. Since the CE would not adapt properly with a sinusoidal signal, for this experiment the CE is frozen after being exercised with pseudo-random 64-QAM signals. Mismatches of $\pm 4\% T$ in the sampling phase and $\pm 5\%$ of gain with respect to the unity are applied. The input tone is identified with a $\nabla$, and spurs from mismatches in the TI-ADC are marked with a $\times$. Notice that the spurs caused by the mismatches among the interleaves seriously degrade both the Signal-to-Noise-plus-Distortion Ratio (SNDR) and Spurious-Free-Dynamic-Range (SFDR) to 19.4 dBFS and 21.9 dBFS, respectively. After applying the proposed technique, the performance of the TI-ADC is boosted to 39 dBFS and 46.6 dBFS, for SNDR and SFDR, respectively.

A comparison to other calibration techniques is reported in Table 3. The technique proposed in this paper is the only one that meets all four criteria established in Section II. As stated there, these criteria are important for digital communication receivers. In addition, our technique is the one that provides the largest improvement when the calibrated high frequency SNDR (HF SNDR) is compared with the uncalibrated one. The best way to compare performance of calibration techniques in the context of a digital communication receiver application is on the basis of the receiver BERs before and after calibration (see Figure 14). In the publications referenced in Table 3, this data is either not provided,
or provided for much smaller constellations, such as BPSK, QPSK, or QAM16. SNDR is not the best metric to use to compare performance of data converters in a communications application, but it is often the only metric provided. In terms of the calibrated SNDR, our result is the best except for Wei’s [23]. Wei achieves 44.44 db post-calibration HF SNDR, which is 5.4db better than ours. But Wei’s pre-calibration HF SNDR is also significantly higher than ours,\(^7\) which implies that Wei’s calibration did not have to work as hard. We believe that one of the key factors in Wei’s excellent performance is the extremely low jitter achieved. However, calibration only provides 0dB improvement in HF SNDR in his case, whereas in ours it provides 19.6db improvement.

**VII. HARDWARE COMPLEXITY ANALYSIS**

This section discusses some practical aspects of the implementation of the proposed compensation technique. We focus on the two main blocks of the all digital architecture: the compensation equalizer and the error backpropagation block.

### A. IMPLEMENTATION OF THE COMPENSATION EQUALIZER

As described in Section IV, the compensation equalizer in a DP optical coherent receiver comprises 4 real valued FIR filters \(g_{i}(l)\) with \(i = 1, 2, 3, 4, \) and \(l = 0, \ldots, L_q - 1\). From computer simulations of Section V it was observed that \(L_q = 7\) is enough to properly compensate the AFE and TI-ADC impairments. Therefore a time domain implementation is preferred for the CE.

Each of these filters has \(M\) independent impulse responses \(g_{i}(l)\) which are time multiplexed as

\[
g_{i}(l)[n] = g_{i}(l)[n/L_q] \quad (14)
\]

Note that time multiplexing of filters with independent responses does not translate to additional complexity when the filter is implemented with a parallel architecture. The use of parallel implementation is mandatory in high speed optical communication where parallelism factors on the order of 128 or higher are typical.

In these architectures, the parallelism factor \(P\) can be chosen to be a multiple of the ADC parallelism factor \(M\), i.e.,

\[
P = q \times M \quad \text{where} \quad q \text{ is an integer. Therefore the different}
\]

\[
\text{time multiplexed coefficients are used in fixed positions of the parallelism without incurring in significant additional complexity in relation to a filter with just one set of coefficients (see Fig. 16). We highlight that the resulting filter is equivalent in complexity to the I/Q-skew compensation filter already present in current coherent receivers [4]. Since the proposed scheme also corrects skew, the classical skew correction filter can be replaced by the proposed CE without incurring significant additional area or power.}

### B. IMPLEMENTATION OF THE ERROR BACKPROPAGATION BLOCK

A straightforward implementation of error backpropagation must include a processing stage for each DSP block located between the ADCs and the slicers. Typically these blocks comprise the BCD, FFE, TR interpolators, and the FCR. All these blocks can be mathematically modeled as a sub-case of the generic receiver DSP block used in Section IV-C and the Appendix. The EBP block is algorithmically equivalent to its corresponding DSP block with the only difference that the coefficients are transposed. Therefore, in the worst case, the EBP complexity would be similar to that of the receiver DSP block.\(^8\) Since doubling power and area consumption is not acceptable for commercial applications, important simplifications must be provided.

Considering that AFE and TI-ADC impairments change very slowly over time in multi-gigabit optical coherent transceivers, the coefficient updates given by (26) and (27) do not need to operate at full rate, and subsampling can be applied. The latter allows implementation complexity, and particularly power dissipation, to be drastically reduced. In Section V-A we evaluated the performance with block decimation where one block of \(N\) consecutive samples of the oversampled slicer error are used every \(DB\) blocks. Simulation results not included here have shown a good performance

\[
\text{Fig. 16. Example of a parallel implementation of the CE with } M = 4, \quad L_q = 3, \text{ and parallelism factor } P = 2M = 8.
\]

\[
P = q \times M \quad \text{where} \quad q \text{ is an integer. Therefore the different}
\]

\[
\text{time multiplexed coefficients are used in fixed positions of the parallelism without incurring in significant additional complexity in relation to a filter with just one set of coefficients (see Fig. 16). We highlight that the resulting filter is equivalent in complexity to the I/Q-skew compensation filter already present in current coherent receivers [4]. Since the proposed scheme also corrects skew, the classical skew correction filter can be replaced by the proposed CE without incurring significant additional area or power.}

\[
\text{Fig. 16. Example of a parallel implementation of the CE with } M = 4, \quad L_q = 3, \text{ and parallelism factor } P = 2M = 8.
\]

\[
P = q \times M \quad \text{where} \quad q \text{ is an integer. Therefore the different}
\]

\[
\text{time multiplexed coefficients are used in fixed positions of the parallelism without incurring in significant additional complexity in relation to a filter with just one set of coefficients (see Fig. 16). We highlight that the resulting filter is equivalent in complexity to the I/Q-skew compensation filter already present in current coherent receivers [4]. Since the proposed scheme also corrects skew, the classical skew correction filter can be replaced by the proposed CE without incurring significant additional area or power.}

\[
\text{Fig. 16. Example of a parallel implementation of the CE with } M = 4, \quad L_q = 3, \text{ and parallelism factor } P = 2M = 8.
\]
even with $N = 8192$ and $D_B = 256$. The block based decimation approach allows the EBP algorithm to be implemented in the frequency domain when necessary to reduce complexity (for example in the EBP of the BCD and FFE). This error decimation reduces the power dissipation of the EBP to only $1/D_B$ of the power of the corresponding DSP blocks, equivalent to less than 1% in the simulated example. However, the areas of the EBP blocks are still equivalent to the area of their corresponding DSP blocks. To reduce area, the EBP blocks could be implemented using a serial architecture\(^9\) or a lower parallelism factor. If a serial implementation is chosen, an area reduction proportional to the area, the EBP blocks could be implemented using a serial to the area of their corresponding DSP blocks. To reduce the EBP blocks, equivalent to less than 1% in the simulated example.

D/\times\text{FFT} / \text{IFFT pair). The latencies of the EBP blocks for the TR interpolators and FCR can be neglected. Therefore the CE adaptation speed is not reduced by a serial implementation of the EBP blocks if $2 \times (N_{BCD} + N_{FFE}) \times P = 2 \times N \times D_B$. Details of efficient architectures for implementing the error backpropagation block will be addressed in a future work.

VIII. CONCLUSION

A novel background calibration technique for TI-ADC mismatches based on the backpropagation algorithm has been introduced in this paper. The characteristics of the backpropagation algorithm are exploited in a digital communication receiver application, where the algorithm is used to generate a suitable error signal, which is processed to effectively mitigate the mismatches of a high-speed TI-ADC. The technique can compensate impairments of the entire AFE (e.g., I/Q time skew). Simulations performed in an application example with a DSP-based, DP optical coherent receiver have shown a fast, robust and almost ideal compensation/calibration of different TI-ADC mismatches. Sampling time, gain, offset, and bandwidth mismatches as well as I/Q time skew errors have been exercised both individually and combined. Measurements have been performed using an emulation platform based on an 8 bit, up to 4 GS/s TI-ADC test chip. We have shown that the degradation in the receiver performance is highly mitigated with this proposal for 64-QAM and 256-QAM schemes. We highlight that the proposed technique is able to compensate mismatches of several types simultaneously (i.e., sequential calibration of individual impairments of different types is not needed). Hardware complexity is minimized using decimation and serial processing in the backpropagation blocks. As the technique runs in background, the proposed technique is able to track parameter variations caused by temperature, voltage, aging, etc., without operational interruptions.

\(^9\)Typically, a serial implementation requires that hardware such as multipliers be reused with variable numerical values of coefficients, whereas in a parallel implementation hardware can be optimized for fixed coefficient values. This results in a somewhat higher power per operation in a serial implementation. Nevertheless, the drastic power reduction achieved through decimation greatly outweighs this effect.

![Figure 17](image1.png)

**FIGURE 17.** Modified model of the analog front-end and TI-ADC for signal $s^{(i)}(t)$ with $i = 1, 2, 3, 4$ in a DP coherent optical receiver.

![Figure 18](image2.png)

**FIGURE 18.** Equivalent discrete-time model of the analog front-end and TI-ADC system with impairments for the signal component given by (35) (i.e., without DC offsets and quantization noise) for signal $s^{(i)}(t)$ with $i = 1, 2, 3, 4$.

APPENDIX

**TI-ADC MODEL**

Next we review the model of the TI-ADC with impairments used in this paper (see Fig. 3). The effects of the sampling time errors $\delta_m^{(i)}$ and gain errors $\gamma_m^{(i)}$ can be modeled by analog interpolation filters with impulse responses $p_m^{(i)}(t)$ followed by ideal sampling [2], [3], as depicted in Fig. 17. The digitized high-frequency samples can be written as

$$y^{(i)}[n] = r^{(i)}[n] + \tilde{s}^{(i)}[n] + q^{(i)}[n],$$

(32)

where $r^{(i)}[n]$ is the signal component provided by the $M$-channel TI-ADC, and $q^{(i)}[n]$ is the quantization noise.

The total impulse response of the $m$-th interleaved channel is defined as

$$h_m^{(i)}(t) = \delta(t) \otimes f_m^{(i)}(t) \otimes p_m^{(i)}(t),$$

(33)

where $m = 0, \ldots, M - 1$ and $\otimes$ is the convolution operator. Let $H_m^{(i)}(\omega)$ and $S^{(i)}(\omega)$ be the Fourier Transforms (FTs) of $h_m^{(i)}(t)$ and $s^{(i)}(t)$, respectively. The spectral shaping commonly used in digital communication systems results in $|S^{(i)}(\omega)| \approx 0$ for $|\omega| \geq \pi/T_s$. Then, the analog filtering of Fig. 17 can be replaced (assuming $|H_m^{(i)}(\omega)| \approx 0$ for $|\omega| \geq \pi/T_s$) by a real discrete-time model, as depicted in Fig. 18, resulting

$$h_m^{(i)}(nT_s) = T_s h_m^{(i)}(nT_s), \quad m = 0, \ldots, M - 1.$$  

(34)

Therefore, it can be shown that the digitized high-frequency signal can be expressed as:

$$r^{(i)}[n] = \sum_{l} \tilde{h}_m^{(i)}[l] s^{(i)}[n - l],$$

(35)

where $s^{(i)}[n] = \tilde{s}^{(i)}(nT_s)$ and $\tilde{h}_m^{(i)}[l]$ is the impulse response of a time-varying filter, which is an $M$-periodic sequence such $\tilde{h}_m^{(i)}[l] = \tilde{h}_m^{(i)}[l + M]$, and defined by

$$\tilde{h}_m^{(i)}[l] = h_m^{(i)}[l], \quad n = 0, \ldots, M - 1, \quad \forall l,$$

(36)

with $h_m^{(i)}[l]$ given by (34).\(^10\) We highlight that the impact of both the AFE impairments and the $M$-channel TI-ADC

\(^10\)See [26] and references therein for more details about this formulation.
mismatches are included in (35). Finally, the digitized high-frequency sequence is obtained by replacing (35) in (32),

\[ y_0^* [n] = \sum_l h_0^* [l] y_0 [n - l] + \delta_0^* [n] + q_0^* [n]. \]  

(37)

REFERENCES

[1] T. H. Tsai, P. J. Hunten, and S. H. Lewis, “Correction of mismatches in a time-interleaved analog-to-digital converter in an adaptively equalized digital communication receiver,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 56, no. 2, pp. 307–319, Feb. 2009.

[2] G. C. Luna, D. E. Crivelli, M. R. Hueda, and O. E. Agazzi, “Compensation of track and hold frequency response mismatches in interleaved analog to digital converters for high-speed communications,” in Proc. IEEE Int. Solid-State Circuits Conf., Kos, Greece, May 2006, pp. 1631–1634.

[3] O. Agazzi et al., “A 90 nm CMOS DSP MLSI transceiver with integrated AFE for electronic dispersion compensation of multimode optical fibers at 10 Gb/s,” IEEE J. Solid-State Circuits, vol. 44, no. 12, pp. 3293–3297, Dec. 2008.

[4] D. Morero, M. A. Castrillón, A. Aguirre, M. R. Hueda, and O. E. Agazzi, “Design tradeoffs and challenges in practical coherent optical transceiver implementations,” J. Lightw. Technol., vol. 34, no. 1, pp. 121–136, Jan. 1, 2016.

[5] M. S. Faruk and S. J. Savory, “Digital signal processing for coherent transceivers employing multilevel formats,” J. Lightw. Technol., vol. 35, no. 5, pp. 1125–1141, Mar. 1, 2017.

[6] D. E. Crivelli, M. R. Hueda, H. S. Carrer, M. D. Barco, R. R. Lopez, P. Gianni, J. Finocchietto, N. Swenson, P. Voinis, and O. E. Agazzi, “Architecture of a single-chip 50 Gb/s DP-DOQSK/BPSK transceiver with electronic dispersion compensation for coherent optical channels,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 61, no. 4, pp. 1012–1025, Apr. 2014.

[7] S. Palermo, S. Hoyoos, A. Shafik, E. Z. Tabasy, S. Cai, S. Kiran, and K. Lee, “CMOS ADC-based receivers for high-speed electrical and optical links,” IEEE Commun. Mag., vol. 54, no. 10, pp. 168–175, Oct. 2016.

[8] Ethernet Alliance. (2020). The 2020 Ethernet Roadmap. [Online]. Available: https://ethernetalliance.org/technology/2020-roadmap/

[9] M. Brandomelli et al., “A 5-GS/s 150 mW 10 b SHA-less pipelined/SAR hybrid ADC for direct-sampling systems in 28 nm CMOS,” IEEE J. Solid-State Circuits, vol. 50, no. 12, pp. 2922–2934, Dec. 2015.

[10] Y. Xu, G. Wu, L. Belostotski, and J. W. Haslett, “5-bit 5-GS/s noninterleaved time-based ADC in 65-nm CMOS for radio-astronomy applications,” IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 24, no. 12, pp. 3513–3525, Dec. 2016.

[11] C. Schmidt, H. Yamazaki, G. Raybon, S. H. Wei, and E. Pincemin, S. J. B. Yoo, “Digitally assisted data converter design,” in Proc. IEEE Int. Solid-State Circuits Conf., Austin, TX, USA, Feb. 2014, pp. 194–195.

[12] F. Solis et al., “EBP-Based Background Calibration of TI-ADC et al.,” in IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 64, no. 6, pp. 1591–1595, Aug. 2017.

[13] H. Wei, P. Zhang, B. D. Sahoo, and B. Razavi, “An 8 bit 4 GS/s 120 mW CMOS ADC,” IEEE J. Solid-State Circuits, vol. 49, no. 8, pp. 1751–1761, Aug. 2014.

[14] M. El-Chammas and B. Murmann, “A 12-GS/s 81-mW 5-bit time-interleaved flash ADC with background timing skew calibration,” IEEE J. Solid-State Circuits, vol. 46, no. 4, pp. 838–847, Apr. 2011.

[15] S. K. Song, K. Ragab, X. Tang, and N. Sun, “A 10-b 800-MS/s time-interleaved SAR ADC with fast variance-based timing-skew calibration,” IEEE J. Solid-State Circuits, vol. 52, no. 10, pp. 2563–2575, Oct. 2017.

[16] S. Lee, A. P. Chandrakasan, and H. S. Lee, “A 1 GS/s 10 b 18.9-mW time-interleaved SAR ADC with background timing skew calibration,” IEEE J. Solid-State Circuits, vol. 49, no. 12, pp. 2846–2856, Dec. 2014.

[17] H. Mafi, M. Yargholi, and M. Yavari, “Digital blind background calibration of imperfections in time-interleaved ADCs,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 64, no. 6, pp. 1504–1514, Jun. 2017.

[18] C. Vogel, “The impact of combined channel mismatch effects in time-interleaved ADCs,” IEEE Trans. Instrum. Meas., vol. 54, no. 1, pp. 415–427, Feb. 2005.

[19] B. Murmann, “Digitally assisted data converter design,” in Proc. IEEE ESSCIRC, Bucharest, Romania, Sep. 2013, pp. 24–31.

[20] C. Vogel, “The impact of combined channel mismatch effects in time-interleaved ADCs,” IEEE Trans. Circuits Syst. I, Fundam. Theory Appl., vol. 48, no. 3, pp. 261–271, Mar. 2001.

[21] A. Adam, “A 12.2-b 10-GS/s interleaved pipelined ADC in 28-nm CMOS technology,” IEEE J. Solid-State Circuits, vol. 52, no. 12, pp. 3204–3218, Dec. 2017.

[22] A. M. S. Hish, and S. H. Lewis, “A time-interleaved SAR ADC with signal-independent background timing calibration,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 69, no. 2, pp. 620–633, Feb. 2022.

[23] F. Solis et al., “EBP-Based Background Calibration of TI-ADC et al.,” in IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 68, no. 4, pp. 1444–1457, Apr. 2021.
[39] Y. Zhao, S. Li, L. Li, and Z. Huang, “A novel time delay estimation and calibration method of TI-ADC based on a coherent optical communication system,” Photonics, vol. 8, no. 9, p. 398, Sep. 2021.

[40] B. T. Reyes, V. Gopinath, P. S. Mandolesi, and M. R. Hueda, “Joint sampling-time error and channel skew calibration of time-interleaved ADC in a multichannel fiber optic receivers,” in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), May 2012, pp. 2981–2984.

[41] B. T. Reyes, R. M. Sanchez, A. L. Pola, and M. R. Hueda, “Design and experimental evaluation of a time-interleaved ADC calibration algorithm for application in high-speed communication systems,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 64, no. 5, pp. 1019–1030, May 2017.

[42] F. Solis, A. F. Bocco, D. Morero, M. R. Hueda, and B. T. Reyes, “Back-ground calibration of time-interleaved ADC for optical coherent receivers using error backpropagation techniques,” in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), Seville, Spain, Oct. 2020, pp. 1–5.

[43] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” Nature, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[44] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[45] G. Raybon, “High symbol rate transmission systems for data rates from 400 Gb/s to 1 Tbs/,” in Proc. Opt. Fiber Commun. Conf. Los Angeles, CA, USA: OSA, 2015, pp. 1–25, Paper M3G.1.

[46] D. E. Crivelli, H. S. Carrer, and M. R. Hueda, “Adaptive digital equalization in the presence of chromatic dispersion, PMD, and phase noise in coherent fiber optic systems,” in Proc. IEEE Global Telecommun. Conf. (GLOBECOM), vol. 4, Dallas, TX, USA, Nov. 2004, pp. 2545–2551.

[47] C. R. S. Fludger, “Digital signal processing for coherent transceivers in next generation optical networks,” in Proc. Eur. Conf. Opt. Commun., Cannes, France, Sep. 2014, pp. 1–3.

[48] E. A. Lee and D. G. Messerschmitt, Digital Communication, 3rd ed. Dordrecht, The Netherlands: Springer, 2004.

[49] F. Solis, B. T. Reyes, D. A. Morero, and M. R. Hueda. “On the application of error backpropagation to the background calibration of time-interleaved ADC for digital communication systems,” Apr. 2020, arXiv:2008.02914.

[50] S. S. Haykin, Adaptive Filter Theory, 4th ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[51] S. Basodi, C. Ji, H. Zhang, and Y. Pan, “Gradient amplification: An efficient way to train deep neural networks,” Big Data Mining Anal., vol. 3, no. 3, pp. 196–207, Sep. 2020.

[52] W. Freude, R. Schmogrow, B. Nebendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, and J. Leuthold, “Quality metrics for optical signals: Eye diagram, Q-factor, OSNR, EVM and BER,” in Proc. Int. Conf. Transparent Opt. Netw. (ICTON), Jul. 2012, pp. 1–4.

[53] C. C. K. Chan, Optical Performance Monitoring: Advanced Techniques for Next-Generation Photonic Networks. New York, NY, USA: Academic, Feb. 2010.

[54] G. P. Agrawal, Fiber-Optic Communication Systems, 4th ed. New York, NY, USA: Wiley, Oct. 2010.

[55] Altera. (Mar. 2014). Stratix IV GX FPGA Development Kit User Guide. [Online]. Available: https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/ug/sivgx_fpga_dev_kit.pdf

[56] Texas Instruments. (2008). DACS681/812/82/EVM User’s Guide. [Online]. Available: https://www.ti.com/lit/ug/slaa286a/slaa286a.pdf

[57] Mini-Circuits. (2008). Coaxial Low Pass Filter VLF-190-4—Datasheet. [Online]. Available: https://www.minicircuits.com/pdfs/VLF-490+.pdf

[58] F. Solis, B. T. Reyes, D. A. Morero, and M. R. Hueda, “Design and experimental verification of a novel error-backpropagation-based background calibration for time interleaved ADC in digital communication receivers,” Apr. 2022, arXiv:2204.04806.

[59] F. Solis, A. F. Bocco, A. C. Galetto, L. Passetti, M. R. Hueda, and B. T. Reyes, “A 4 GS/s 8-bit time-interleaved SAR ADC with an energy-efficient architecture in 130 nm CMOS,” Int. J. Circuit Theory Appl., vol. 49, no. 10, pp. 3171–3185, Oct. 2021.

[60] R. L. Nguyen et al. “A highly reconfigurable 40–97 GS/s DAC and ADC with 40 GHz AFE bandwidth and sub-35 Uconv-step for 400 Gbs coherent optical applications in 7 nm FinFET,” in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, vol. 64, Feb. 2021, pp. 136–138.

[61] J. Cao, D. Cui, A. Nazemi, T. He, G. Li, B. Catli, M. Khanpouri, K. Hu, T. Ali, H. Zhang, H. Yu, B. Khew, S. Sheng, Y. Shim, B. Zhang, and A. Montaz, “A transmitter and receiver for 100 Gbs coherent networks with integrated 4×64 GS/s 8b ADCs and DACs in 20 nm CMOS,” in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2017, pp. 484–485.

[62] L. Kull, D. Luu, C. Menolfi, M. Brändli, P. A. Francese, T. Morf, M. Kossel, A. Cevrero, I. Ozkaya, and T. Toifi, “A 24–72-GS/s 8-b time-interleaved SAR ADC with 2.0–3.3-g/conversion and >30 dB SNDR at Nyquist in 14-nm CMOS FinFET,” IEEE J. Solid-State Circuits, vol. 53, no. 12, pp. 3508–3516, Dec. 2018.

FREDDY SOLIS (Graduate Student Member, IEEE) was born in Jujuy, Argentina, in 1990. He received the degree in electronic engineering from the Facultad Regional Córdoba (FRC), Universidad Tecnológica Nacional (UTN), Argentina, in 2017. He is currently pursuing the Ph.D. degree in mixed-signal integrated circuits design and calibration with high-speed digital communication systems applications. Since 2016, he has been a member of the Microelectronics and Analog Design Group, Fundación Fulgor. His research interests include high-speed mixed-signal circuits (such as ADCs and DACs), low-power analog design, and new technologies applied to electronic design and communications.

BENJAMÍN T. REYES was born in Río Tercero, Argentina, in 1982. He received the degree in electronic engineering from Universidad Tecnológica Nacional, Córdoba, Argentina, in 2008, and the Ph.D. degree in engineering from the Universidad Nacional del Sur, Bahía Blanca, Argentina, in 2015. He is currently a member of the Research Staff with Fundación Fulgor and a Senior Analog Designer at Marvell Argentina. His current research interests include high-speed analog-to-digital converters, high-speed mixed-signal circuits design, and digital-assisted analog design for digital communications applications.

DAMIÁN A. MORERO was born in Córdoba, Argentina. He received the degree (Hons.) in electronic engineering and the Ph.D. degree in engineering science from the National University of Córdoba (UNC), Córdoba, Argentina. From 2006 to 2009, he received the Ph.D. Fellowships from the Secretary of Science and Technology (SeCyT), Argentina. He is currently the Senior Director Research and Development at Marvell Argentina, where he has been engaged in the research and development of error correction coding schemes and digital signal processing algorithms for high-speed optical communications. His research interests include coding, information theory, signal processing, computer science, and VLSI design. In 2003 and 2005, he received the Academic Excellence Award from the Engineers Association of Córdoba Argentina and the UNC, respectively.

MARIO R. HUEDA was born in Jujuy, Argentina. He received the degree in electrical and electronic engineering and the Ph.D. degree from the Universidad Nacional de Córdoba (UNC), Argentina, in 1994 and 2002, respectively. Since 1997, he has been with the Digital Communications Research Laboratory, Department of Electronic Engineering, UNC. From 2005 to 2021, he was also with CONICET (National Scientific and Technological Research Council of Argentina). He is currently the VP Research and Development at Marvell Argentina. He has 50 patents issued or pending and has published more than 80 technical papers in journals and conferences. His research interests include digital communication, signal processing, and synchronization.