ABSTRACT This paper elucidates the relative benefits of bit loading and power loading for discrete multitone (DMT) modulation over wireline links. There is considerable variability in the use of bit and power loading in recent research on DMT for wireline applications. Here, we first compare different combinations of bit and power loading schemes in simulations. We propose a method for dynamically adjusting DMT bit and power profiles based on the bit error distribution across sub-channels, as determined by the receiver. We then present experimental results showing DMT operation at 75 Gb/s with a spectral efficiency of 2.5 b/sample over a channel with a smooth magnitude response exhibiting 26 dB attenuation at half the sampling rate and a loss of 33 dB at 1/4 the data rate (i.e., the Nyquist rate of a 4-level pulse-amplitude modulated link at the same data rate). The proposed dynamic bit and power allocation method reduces the bit error rate from $3 \times 10^{-3}$ to $2 \times 10^{-5}$, measured with 23,265,420 bits.

INDEX TERMS Bit error rate (BER)–aware, bit loading, discrete multitone (DMT), orthogonal frequency-division multiplexing (OFDM), power loading, transceiver, wireline communication.

I. INTRODUCTION

The ever-increasing demand for higher data rates in communication systems necessitates novel approaches that enable high data transfer speeds over low-cost wireline channels. After exhausting approaches to improve channel loss, a more recent approach is to encode more bits through multilevel signaling. However, multilevel signaling comes at a significant signal-to-noise ratio (SNR) penalty that has limited its adoption to 4-level pulse-amplitude modulation (4-PAM), where 2 bits are encoded per unit interval (UI). Discrete multitone (DMT) can enable fine-grained shaping of power spectral density for 200+ Gbps links. Historically, DMT has been used in telecommunication applications such as digital subscriber line (DSL)—with Asymmetric DSL (ADSL) being a sub-category of this application—to allow for efficient use of available frequency spectrum [1]. Traditional pulse-amplitude modulation (PAM) has a fixed number of bits per sample (hence a fixed spectral efficiency). On the contrary, in DMT spectral efficiency varies across the channel, often depending on available SNR. Therefore, the system designer can allocate as much information (bits per sample) as channel conditions and specified performance metrics, such as the target bit error rate (BER), allow. In fact, as we will show in this paper, BER in each sub-channel can be used alongside available SNR to set or fine-tune the bit allocation. In addition, DMT has several advantages. First, equalization is much simpler compared to amplitude modulation [1], [3]: difficult channels and high data rates necessitate very accurate channel equalization, which is done most
efficiently in the frequency domain. With the transition to the frequency domain made already (for equalization), there is not much overhead in switching to DMT modulation. Second, DMT can handle a channel discontinuity more effectively than amplitude modulation by avoiding any resulting notches in the frequency response through properly shaping the power spectrum of the modulation [4], [5], which is not possible with PAM. Finally, DMT has a high tolerance to sampling phase offset and jitter that can be exploited in timing recovery [2], [6].

Despite the advantages discussed above, DMT is not widely used in 60+ Gb/s wireline signaling. The transceiver reported in [7] operated at 56 Gb/s over a channel with 28 dB loss at a frequency of 1/4 the bit rate. That work used the same fixed quadrature-amplitude modulation (QAM) in each DMT sub-channel (except for the highest-frequency one) regardless of its SNR or BER. Another work [8] reported orthogonal frequency-division multiplexing (OFDM)—which is the same as DMT in this context—operation at a line rate of 212 Gb/s (with a payload of 184 Gb/s when error correction overhead is disregarded) but a channel loss of 9 dB at half the sampling rate. In the present article, we demonstrate via simulations and measurements that a significant data rate increase with DMT is achievable over channels with higher losses through a BER-informed bit loading and power loading approach.

The paper starts with a brief summary of recent research efforts and background on DMT in Section II, followed by a system overview in Section III. Section IV details the proposed BER-informed bit and power profile adjustment (BPPA) procedure. Section V presents simulation results comparing several cases with and without bit and/or power loading (including one with dynamic adjustments to the corresponding profiles). Section VI reports on an experimental setup and the associated results that incorporate the ideas discussed in the preceding sections. Finally, Section VII summarizes the findings of this paper and compares its results to those of other, similar works.

II. BACKGROUND

DMT modulation has been explored for wireline links (both electrical and optical) above 56 Gb/s in [2], [5], [6], [7], [8], [9], [10], [11], and [12]. The DMT-based receiver in [7] reported a BER of $2 \times 10^{-4}$ over a channel with a loss of 28 dB at 1/4 the data rate (the Nyquist frequency of a 4-PAM signal at the same data rate) and a loss of 20 dB at half the sampling rate ($f_s/2$). In [5], we demonstrated the effectiveness of DMT on a channel with a notch. However, the recent trend in most backplane and chip-to-chip channels is that they are quite smooth but can be lossy at high frequencies [13], [14], [15], [16]. Previously, we presented simulation results on a DMT link operating over a smooth channel in [6]. Here, our goal is to extend the loss compensation capability compared to previously reported DMT solutions in [7] and [6] by taking advantage of a dynamic bit and power loading approach that we will discuss further in the article. We now provide more details on the prior art in various aspects of the present work.

A. BIT AND/OR POWER LOADING

Even though a uniform bit (or power) allocation can also be considered a type of bit (or power) loading, in this paper we use the terms bit (and power) loading to refer strictly to the case where the corresponding profile is adjusted according to some criterion, such as sub-channel conditions. For clarity, different cases of inclusion or exclusion of bit or power loading are shown in Fig. 1. In short, bit loading involves assigning different orders to the constellations whereas power loading involves changing the sizes of the constellations.

Existing efforts in DMT for wireline signaling can be classified into three categories: those with only bit loading, those with only power loading, and those with both bit and power loading. The DMT implementations or models in [7],
[8], [10], and [12] used uniform bit allocation across their sub-channels (although [7] did bit-load in a very coarse way by reducing the constellation size of only 1 of 15 sub-channels), but they used power loading to improve the performance in the sub-channels facing a higher loss. By contrast, [5], [17] utilized bit loading but forwent power loading. Several other DMT-based studies [3], [18] performed both bit and power loading. Considering this variability among whether bit loading and/or power loading are applied in either electrical or optical links, a goal of this article is to take a closer look at the potential BER improvements resulting from allocating bits or power non-uniformly.

B. ADAPTIVE BIT LOADING AND/OR POWER LOADING

In the existing literature, bit and power loading is, for the most part, non-adaptive and applied at startup, without dynamically adapting to changing channel conditions [5], [7], [19]. While this approach may work for channels that do not change much (or change slowly) over time—and this is often the case in wireline communications—the analog front-end (AFE) response may see a significant variation with temperature. While there are relatively simple ways to combat AFE process, voltage and temperature (PVT) variations, such as changing transmitter pre-emphasis or receiver equalization, these methods result in sub-optimal use of the available channel response and SNR (for example, they lead to decreased transmit signal power, or to excessive noise enhancement at the receiver). Therefore, dynamically adjusting bit and power profiles according to the performance of the system can be beneficial. Moreover, even disregarding such changes in SNR or operating conditions, allowing dynamic adjustments eliminates the constraint of ensuring that the initial bit and power profiles are optimal, which would require accurate knowledge of the impairments in the link. On the other hand, as we will discuss below, such adjustments require the use of a backchannel, which entails additional hardware.

A feedback path from receiver to transmitter that could potentially enable dynamic adjustments was shown in [12] although its use for such adjustments was not mentioned in that work. Also, [20] discussed a “bit swapping” approach that employed a steady-state mean squared error (MSE) criterion to adjust the bit profile. While [20] did describe the corresponding procedure (which is based on computing the difference between the input and output of the receiver-side hard decision slicer), it did not provide information on its effects on performance. Moreover, [21] considered the one-bit-at-a-time updates in such an approach to be very slow and proposed “express bit swapping”, namely updating bit and power profiles of multiple sub-channels at once. Other works [22], [23] discussed approaches with feedback paths enabling power allocation updates involving only a limited set of predefined profiles.

Given that bit and power loading can be considered coarse and fine steps in distributing channel capacity among different sub-channels, they are intertwined and depend on one another. Therefore, we propose and investigate a dynamic bit and power allocation approach that iteratively adjusts the corresponding profiles until satisfactory performance is achieved.

C. BER-BASED ADAPTATION

To the extent that bit and power loading were applied in prior art, they were based on SNR criteria [7], [19]. In addition to making the transceiver more robust to factors such as channel aging and temperature variations, dynamic adjustment of bit and power profiles makes it possible to take into account all factors that affect SNR, such as noise and jitter. Note, however, that maximizing SNR, which aims to minimize pre–forward error correction (FEC) BER, does not always precisely minimize post-FEC BER [24], prompting us to consider BER as a bit and power profile adjustment criterion.

There have been examples of serializer/deserializer (SerDes) designs that aim for minimizing link BER as the objective function of their adaptation mechanism [25]. Similarly, our approach of dynamic BPPA utilizes error distribution profiles—i.e., knowledge of how susceptible to errors each sub-channel is—to improve BER performance.

The proposed approach is similar to express bit swapping [21] in its use of the errors detected at the receiver to adjust the bit and power profiles. However, [21] targets lower-data rate applications such as DSL, uses different update criteria and procedures than those in the present work, and primarily reports on the speed improvement resulting from its proposed approach, as opposed to on detection accuracy improvement and error reduction. Also, in contrast to [23], which considers a situation where communication is suddenly disturbed by narrow-band interference (NBI) and emphasizes the need to maintain connection between transmitter and receiver, our work is motivated by a desire to continually improve performance on an existing connection or to determine a suitable bit and power allocation that more accurately captures channel conditions.

III. SYSTEM OVERVIEW

Fig. 2 shows a DMT transceiver [1], [2], [5], [6], including Gray-coded QAM mapping, an inverse discrete Fourier transform (IDFT) block, cyclic prefix (CP) insertion (of length v samples) and a digital-to-analog converter (DAC). At the receiver, the signal is amplified, digitized by an analog-to-digital converter (ADC), processed through a discrete Fourier transform (DFT), and equalized in the frequency

1Note that the word “adaptive” used in the title of [3] refers to bit and power loading profiles designed to adapt to channel conditions and applied at startup, and it does not refer to dynamically adapting to changing channel conditions.
FIGURE 2. Overview of the DMT system under consideration. Modified from [5]. (I)FFT refers to (inverse) fast Fourier transform, an implementation of (I)DFT—(inverse) discrete Fourier transform—allowing fast computations. (a) Block diagram. The green blocks and elements are potentially time varying. For example, as the channel ages, the BPPA block adjusts the values of $b_k$ and $p_k$, necessitating changes to $\alpha_1$ and, by extension, to $\alpha_2$.
(b) Detailed view of the input mapping and the subsequent multiplications, highlighted with a dashed box in panel (a). The numbers (e.g., sub-channel indices $k$, the values of $\alpha_1$, $b_k$ and $p_k$) are arbitrary and for illustration purposes only, and $i$ indicates the imaginary unit.

Our goal is to introduce dynamic bit and power loading with minimal added complexity in the digital signal processing (DSP). A backchannel in the link, from the receiver to the transmitter, helps adaptively adjust the bit and power allocation in the digital domain based on receiver-side error statistics. For example, a bidirectional frequency division approach as in [26] could be used. Moreover, because the receiver needs knowledge of the bit and power loading profiles employed at the transmitter (and must be informed of changes to these profiles), these profiles can form a small part of the transmit sequence. Communication protocols allowing this coordination between transmitter and receiver with minimal resources and latency are discussed in [21].

Fig. 2(b) shows an example of how bits are mapped to a QAM constellation. With a $2N$-point (I)DFT, there are $N − 1$ sub-channels in the DMT system. So, the total DMT symbol length is $2N + v$, and the total number of allocated bits per DMT symbol is $B = \sum_{k=1}^{N-1} b_k$, where $b_k$ is the QAM constellation order of sub-channel $k$. These values can be grouped into a vector, $\mathbf{b}$. No bit loading indicates a constant value for all $b_k$. The resulting spectral efficiency is $B/(2N + v)$ bits/sample. Also, in our work, we set a limit on the maximum constellation size (the values of $b_k$) [5].

IV. BIT AND POWER PROFILE ADJUSTMENT

The receiver tracks the error count in each sub-channel, $e$ (obtained using its FEC decoder) and feeds it back to the transmitter via a required backchannel. The BPPA block then updates the values of $b_k$ and $p_k$ based on error statistics assumed to be available from FEC decoding at the receiver, considering the prevalence of FEC in modern wireline links [24], [27]. Such decoders use information such as syndromes and parity bits to determine where errors have occurred as part of the error detection and correction process. Even though the results presented in Sections V and VI make use of known transmit sequences at the receiver to determine error distributions, a future circuit implementation can use information from a FEC decoder for this purpose, although further investigation may be necessary to study how this approach (and the latency associated with it) for the adaptation scheme proposed in this work might affect calibration loop stability.
is used to decide if the errors are sufficiently balanced among sub-channels that contain errors. Imbalance can be quantified using variance, so this is the property that Algorithm 1 employs to decide whether the error distribution is no longer dominated by a few sub-channels with many errors, and further bit profile updates should involve an increased value of $\tau_1$ (which would make such updates less frequent and would shift the focus onto the power loading part of the procedure).\textsuperscript{2} If this is the case, the criterion for what constitutes a balanced distribution becomes stricter, i.e., $\tau_2$ decreases.

The values of the parameters discussed thus far should be determined empirically as they depend on specifics of the system under consideration: for example, changing the number of sub-channels may require a change in $\tau_1$ and $\tau_2$ are denoted $c_1$ and $c_2$, respectively. Note that whereas $\tau_1$ helps decide which sub-channels should be labelled dominant, $\tau_2$ controls (by changing $\tau_1$ based on the variance of the set of normalized error counts) what constitutes a dominant sub-channel. Another parameter $l$ establishes the maximum power decrease in a sub-channel to avoid abrupt changes.

The adaptation approach improves the overall link BER by coarsely normalize the error rate across sub-channels, and $\alpha_1$—shown in Fig. 2(a)—after each iteration of Algorithm 1, such that the desired IBO at the transmitter DAC is maintained, mitigating PAPR-related concerns.

### Algorithm 1 Bit and Power Profile Adjustment

```plaintext
\texttt{Algorithm 1 Bit and Power Profile Adjustment}
\texttt{\begin{align*}
\hat{e} & \leftarrow e/\sum_{k=1}^{N-1} e_k \\
\text{if } \text{Var}[\hat{e}] < \tau_2 \text{ then} & \\
\tau_1 & \leftarrow \tau_1 + c_1 \\
\tau_2 & \leftarrow \tau_2/c_2 \\
\text{end if}
\end{align*}}
\texttt{Sort } \hat{e} \text{ in descending order}
\texttt{for } k = 1 \text{ to } N - 1 \text{ do}
\texttt{if } \hat{e}_k > \tau_1 \text{ and } \exists \text{ an error- & flag-free sub-channel, } q, \text{ with } b_q \text{ below the QAM order limit then}
\texttt{Flag sub-channel } k
\texttt{\quad } b_k \leftarrow b_k - 1 \\
\texttt{\quad } b_q \leftarrow b_q + 1 \text{ (Pick lowest } q \text{ among possible options)} \\
\texttt{\quad } p_k \leftarrow 1 \\
\texttt{\quad } p_q \leftarrow 1 \\
\texttt{else if } \hat{e}_k > 0 \text{ then}
\texttt{\quad } p_k \leftarrow p_k + \hat{e}_k \\
\texttt{\quad for sub-channels } o \in O \text{ with no error history do}
\texttt{\quad\quad } p_o \leftarrow \min(\sum_{k=1}^{N-1} \hat{e}_k/|O|, l) \\
\texttt{\quad end for}
\texttt{end if}
\texttt{end for}
```

it is possible for the peak-to-average power ratio (PAPR) of the DMT waveform to change as a result of updated power loading, PAPR and IBO are related in that by maintaining a certain IBO, we balance the PAPR-related clipping against the SNR-related advantages of a higher average power \cite{5,6}. Therefore, the BPPA block also re-evaluates $\alpha_1$—shown in Fig. 2(a)—after each iteration of Algorithm 1, that the bit loading profiles are used to coarsely normalize the error rate across sub-channels, and power loading is used for fine-tuning. When we add “extra” bits to error-free constellations and take them away from the sub-channels with the highest BERs, a significant change in sub-channel SNR results. Therefore, in addition to updating the bit loading profile, we need to reallocate sub-channel power to optimize the available SNR.

The power adjustment part of Algorithm 1 is essentially a greedy search in that we seek to only allow decreases in the objective function that we aim to minimize (i.e., the BER) by reallocating power from error-free sub-channels to error-prone ones and flagging previously error-free sub-channels as soon as errors start to occur in them (indicated by the “with no error history” condition on the set $O$, i.e., we also have a power-related flag, separate from the one that is set when swapping bits). Therefore, if the new power allocation is worse in terms of the resulting BER, the latest power decrease is reversed in the next iteration by increasing the power in the recently flagged sub-channel(s).

The bit adjustment part of the algorithm is guided by the heuristic that having many errors concentrated in a few

\textsuperscript{2}If all sub-channels have a small BER, some sub-channels can be flagged as dominant even if they have a very small error count that is nevertheless several times larger than that of other sub-channels. In situations like this, a larger $\tau_1$ helps handle such small errors with gradual power increases as opposed to sudden changes in constellation sizes.
sub-channels indicates that the constellation sizes may be unnecessarily large there. While swapping bits could occasionally lead to BER increases, the algorithm ensures that the bit profile changes causing such increases are reversed (more accurately, that additional changes are made) by identifying the new regions of error concentration (that is, sub-channels with newly increased constellation sizes) and reducing the bit allocation there. Therefore, as our simulations and measurements have shown, occasional increases in BER are followed by further decreases, resulting in an overall downward pattern.

Also, in Sections V and VI, the receiver has knowledge of any changes to the bit and power profiles effected by the BPPA block, therefore when transitioning to new profiles, errors related to symbol unmapping caused by outdated profiles do not occur. As mentioned in Section III, there are ways to ensure this type of coordination between transmitter and receiver. Examples include [21] and an acknowledge signal–based scheme employed in DSL and described in the ANSI T1.413 ADSL standard [21], [28].

V. SIMULATION RESULTS

Using the model in [5], we compare different scenarios to assess the effectiveness of BER-based bit and power loading:
1) No bit loading or power loading
2) No bit loading, adaptive power loading
3) Bit loading only at startup, no power loading
4) Bit loading only at startup, adaptive power loading
5) Adaptive bit and power loading using BPPA

For Scenario 2, we used a modified form of Algorithm 1 with the second \( \pm \) block (for detecting and handling sub-channels with a disproportionately large number of errors) disabled and only power adjustments applied, while uniformly assigning a QAM order to all sub-channels. Scenario 3 takes a similar approach to that in [5] and—as described in that paper—uses a modified form of the Chow et al. algorithm in [19].

Scenario 4 makes use of the same bit loading algorithm at startup but additionally employs a modified form of the
procedure described in Section IV with the second $\lfloor \frac{1}{2}$ block disabled and power adjustments applied. Scenario 5 is meant to help assess the efficacy of the approach proposed in this paper. Note that the changes to bit loading effected by the BPPA block in Scenario 5 are in addition to an initial step of bit loading at startup, using the same initial allocation algorithm as in Scenarios 3 and 4.

To better establish the reliability of the simulation results and of the comparison between different approaches, we used three channels with different characteristics. The channels are obtained from [29], [30]—Channels 1–3 are from [31], [32], [33], respectively—and their magnitude responses are shown in Fig. 3.

In the simulations: $N = 128$, $v = 20$, 7-bit ideal DAC and ADC were used (since this resolution is common in recent DSP-based transceivers, for example [13], [14], [15]; moreover, [5] has considered lower and higher resolutions as well) with fullscale limits of 0.5 V ($1V_{pp}$) and 0.4 V ($0.8V_{pp}$), respectively, $\sigma_{z} = 0.7 \text{ mV}_{\text{rms}}$ and $\sigma_{n} = 1 \text{ mV}_{\text{rms}}$, which represent additive white Gaussian noise (AWGN) and AFE noise, respectively, intended to model all sources of noise and jitter in the system. As shown in [5] and [6], the impact of jitter, especially for $\sigma_{j}$ values up to around 0.01 UI (corresponding to 125 fs_{rms} when $f_s = 80 \text{ GS/s}$) is minimal. The largest constellation permitted in the bit loading process is 256-QAM. Moreover, $\text{IBO} = 12 \text{ dB}$ at both the DAC and ADC. This choice is supported by our previous results in [5] and [6]. Even though $N$ and the constellation sizes may be different in those works, [34] shows a lack of dependence of suitable IBO values on these parameters. Furthermore, [35] reports optimal IBO values that are similar to those reported in [5] and [6]. As for the BPPA used in Scenarios 2, 4 and 5, $l = 0.1$, $r_1 = 0.02$, $r_2 = 10^{-4}$, $c_1 = 0.03$ and $c_2 = 10$.

Fig. 4(a), Fig. 4(b) and Fig. 4(c) show the simulation results over Channel 1, Channel 2 and Channel 3, respectively. The spectral efficiencies are chosen to accommodate Scenarios 1 over Channel 1, Channel 2 and Channel 3, respectively. The Solomon (RS)(544, 514), or KP4, code and a post-FEC target $2.76 \text{ b/sample}$ correspond to all sub-channels being 16-QAM, the spectral efficiencies of $1.84 \text{ b/sample}$ and $2.30 \text{ b/sample}$ and $2.76 \text{ b/sample}$ correspond to all sub-channels being 16-QAM, 32-QAM and 64-QAM, respectively. This figure additionally indicates a FEC threshold, with the assumption of a Reed-Solomon (RS)(544, 514), or KP4, code and a post-FEC target BER of $10^{-15}$ [16], [36].

In Fig. 4, even though different sampling rates are shown (allowing the system designer to select one option at the design phase), each individual data point corresponds to a separate run with a fixed sampling rate. That is, the DMT operation does not involve changes to the sampling rate during runtime.

The reported BER values in Fig. 4 (over all three channels) for Scenario 4 and Scenario 5 are those obtained within 30 iterations of bit or power adjustments. The BER values for Scenario 2, which also uses dynamic adaptation, are those obtained within 60 iterations. We found that more iterations are required in Scenario 2 than Scenario 5 because the lack of bit loading results in many errors in the higher sub-channels, requiring more power increments for improved performance. Therefore, in Scenario 2, to be able to improve BER faster, one should consider larger power increments than in Algorithm 1.

Based on the results in Fig. 4, the combined BER-based bit and power loading method proposed here affords a BER improvement 1–3 orders of magnitude larger compared with the improvement afforded by only power loading or only bit loading. Moreover, we see BERs that are correctable with typical FEC at spectral efficiencies exceeding the 2 b/sample provided by standard 4-PAM signaling. For example, using a 120 GS/s DAC and ADC with DMT and BPPA over Channel 2 with a spectral efficiency of 1.84 b/sample—indicated by the bottom rightmost green (star) data point in Fig. 4(b)—results in a BER of $2 \times 10^{-4}$ with operation at 221 Gb/s, whereas we expect a 4-PAM link at this data rate to be unfeasible with reasonable equalization because of the notch appearing at 48 GHz, i.e., before the Nyquist frequency of 55 GHz. As another example, as shown in Fig. 4(c), with a sampling rate of 72 GS/s, the BER over Channel 3 with a spectral efficiency of 2.76 b/sample (yielding a bit rate of $\sim 200 \text{ Gb/s}$)—indicated by the bottom middle blue (star) data point—is $1 \times 10^{-4}$ when adaptive BPPA is applied. Achieving the same data rate with reasonable equalization using 4-PAM and 6-PAM would be difficult because of the notch at 37.5 GHz—besides, 4-PAM would require a 100 GS/s DAC and ADC instead of 72 GS/s.

To get an estimate of the achievable performance using PAM, focusing on Channel 3 (as in the second example above), we used the MATLAB code discussed in [37] to determine the channel operating margin (COM) [38], [39] with 6-PAM at a data rate of 200 Gb/s. (We did not consider 4-PAM in this case because its Nyquist frequency would exceed the maximum frequency available from the S-parameters of the considered channel.) Even with extremely generous COM configuration parameters, such as:

- 24 decision-feedback equalizer (DFE) taps,
- Almost no rise or fall times (to maintain consistency with the assumptions in the DMT simulations),
- No bandwidth limitations (again, for consistency),
- No jitter,
- Ignoring package trace lengths and impedances, and
- A transmitter-side SNR of 40 dB (meant to include distortion, which is expected to occur in DMT as well because of clipping),

the COM value—with a threshold set at $2 \times 10^{-4}$ based on our FEC threshold assumption—is negative, indicating no eye opening.

In Fig. 4(b), Scenario 2 (only power loading) works better than Scenario 4 (which contains an extra bit loading step) when the spectral efficiency is 2.30 b/sample. We confirmed that in the latter scenario the SNR-based bit loading done at startup caused a large initial BER in Sub-channel 1,
requiring a substantial increase in $p_1$ and leaving less room for additional power allocation to other sub-channels. This highlights the advantage of basing bit loading on BER instead of SNR. Also, in Fig. 4, even though Scenario 2 yields competitive performance in 1–2 isolated cases (such as the example mentioned above), overall the results suggest that some form of bit loading is essential for minimizing BER.

While Scenario 4 approaches the performance of Scenario 5 with some spectral efficiencies and channels, the latter consistently outperforms the former (as well as other scenarios) across almost all tested channels and spectral efficiencies, at times by close to an order of magnitude—see, for example, the 1.84 b/sample and 2.30 b/sample data (the dotted green and solid crimson traces) in Fig 4(b). Furthermore, in Fig. 4 we assume that the channel characteristics remain unchanged over time. However, in reality, in addition to channel aging, there can be AFE response variations with temperature. In such scenarios, we expect an adaptive bit and power loading approach (i.e., Scenario 5) to make a larger improvement in performance compared to situations with bit loading only at startup (i.e., Scenarios 3 and 4).

**VI. EXPERIMENTAL RESULTS**

**A. SETUP**

To experimentally validate the performance benefits of BPPA, we prepared a laboratory demonstration platform, shown in Fig. 5 and similar to those in [3] and [18]. It comprises an arbitrary waveform generator (AWG) (Keysight M8194A [40]) used as the transmitter, and a real-time oscilloscope (Tektronix DPO77002SX [41]) used as the receiver. Input data generation, received data processing and BPPA were done in MATLAB®. A 24” printed circuit board (PCB) trace on a BERTScope inter-symbol interference (ISI) board (BSA12500ISI [42]) was used as the channel. The total measured channel loss in Fig. 6 includes additional adapter and cable losses that are also part of the setup. Since [42] provides the magnitude response for this experimental channel up to 20 GHz, and we were able to characterize this channel accurately only up to 30 GHz (as shown in Fig. 6), we could not have used it in the simulations for comparison with the other channels in Section V over a wider range of frequencies.

For these experiments, the DMT parameters were: $N = 64$, $v = 10$, and sampling rate $f_s = 30$ GS/s with a target spectral efficiency of 2.5 b/sample and IBO of 12 dB at the DAC. (No variable-gain amplifier (VGA) appears in front of the oscilloscope to control its IBO in this experimental setup.) The peak-to-peak amplitude of the transmit signal was 0.73 V_pp. The largest constellation permitted in the bit loading process was 128-QAM [5]. The BPPA parameters $l = 0.1$, $\tau_1 = 0.08$, $\tau_2 = 10^{-3}$, $c_1 = 0.1$ and $c_2 = 10$ were used for these experiments. Note the larger values of $\tau_1$ and $c_1$ compared to those in the simulations presented in Section V, in line with the earlier discussion that reducing $N$ may necessitate increasing the values of these parameters.

**B. RESULTS**

This section describes how the link performance evolves with BPPA iterations. Fig. 7 shows the final state—after multiple iterations of BPPA—of the equalized received data in a sample set of sub-channels. Fig. 8 shows the BER across the sub-channels evolve over time during BPPA operation. Specifically, initial, intermediate, and final bit and power profiles are shown in Fig. 8(a), Fig. 8(b) and Fig. 8(c), respectively. The error distribution in each case is also included, corresponding to the highlighted BER data points in Fig. 9.

As shown in Fig. 9, the BER improved by at least 2 orders of magnitude over less than 15 iterations of BPPA, which required approximately 8.3 million waveform samples, corresponding to 23,265,420 bits. Given that $f_s = 30$ GS/s, this translates to less than 0.28 ms.

Interestingly, the overall link introduced an artifact around Sub-channels 52 and 53, distorting the shapes of these QAM symbols. This narrowband impairment, which may be related...
to spurs caused by the AWG’s 4-way time-interleaved DAC, is not evident in the measured channel response in Fig. 6, yet the BPPA is able to identify it and adapt bit and power loading around it. Note that we did take it into account in the initial bit loading shown in Fig. 8(a) to save time in running this experiment. When we did not do so, the initial BER was higher (approximately \(8 \times 10^{-3}\)), and the adaptive bit adjustments successfully reduced the QAM orders in the affected sub-channels to arrive at a similar final bit profile as in Fig. 8(c), albeit after a larger number of iterations.

C. DISCUSSION

Looking at Fig. 9, one observes that the last 4–5 iterations of dynamic BPPA are spaced over longer time intervals than the first 8–9. This is because for a fixed number of transmitted symbols, smaller BER values translate to fewer errors. A very small number of errors lacks statistical significance and does not exhibit a reliable error distribution that can be used for power profile adjustments. Therefore, in such conditions, longer times (more DMT symbols) are necessary between consecutive BPPA iterations. However, such fine-tuning under low-BER conditions can be performed slowly in the background since it is below the correctable limit of typical FEC.

In summary, a spectral efficiency of 2.5 b/sample is demonstrated, which is almost double the 1.23 b/sample in [5]. A spectral efficiency of 2.5 b/sample was also reported in [7], but at a BER of \(2 \times 10^{-4}\) with a channel loss of 28 dB at 1/4 the data rate of that work, whereas BPPA afforded a BER of \(2 \times 10^{-5}\) at 33 dB loss at 1/4 the data rate of our system. A higher spectral efficiency, namely 2.94 b/sample, was reported in [8]—although this number is based on the line rate and includes FEC overhead; the payload is 2.56 times the data converter sampling rate, not 2.94 times—but at 9 dB of channel loss at \(f_s/2\), which is 17 dB lower than the channel loss at \(f_s/2\) of our experimental demonstration, and with a pre-FEC BER of \(5 \times 10^{-4}\), an order of magnitude higher than the BER after BPPA in our work.

Although our experimental results are obtained using commercially available hardware as opposed to the silicon implementations in [7] and [8], the specifications, such as speed and resolution, are comparable to those of the reported ADCs and DACs in today’s wireline transceivers. Both the DAC and the ADC have 8 bits of vertical resolution [40], [41]—with the former having an effective number of bits (ENOB) of 5.5 bits at 30 GHz—and the measured noise in the setup is about 3.1 mVrms. The DAC and ADC in [7] also have at least 8 bits of resolution, and so does the DAC in [8]. Finally, further improvement in our setup may be possible if...
FIGURE 8. Bit and power profiles and the resulting error distribution: (a) initial (before BPPA), (b) intermediate (after 5 iterations of BPPA), (c) final (after 12 iterations of BPPA). The insets show zoomed-in views. Note the insets’ different y-axis scales.
We have demonstrated a practical method for adaptive bit and power loading in DMT wireline links that minimizes BER. Simulation results show the proposed approach offers significant improvements (up to 3 orders of magnitude) in BER compared to cases with only startup bit loading or only adaptive power loading. We also experimentally validated this performance improvement using an AWG and a real-time oscilloscope as the DMT transmitter and receiver, respectively. Operating at a sampling rate of 30 GS/s over a channel with a loss of 26 dB at half this sampling rate, the system transmitted data with a spectral efficiency of 2.5 b/sample, reaching a data rate of 75 Gb/s. The channel loss at 1/4 this data rate is 33 dB. Even though the initially observed BER (with bit loading at startup and no power loading) was about $3 \times 10^{-3}$, the adaptive BPPA procedure described in this paper helped bring the BER down to $2 \times 10^{-5}$. The amount of time spent in data transmission to enable this BER reduction was less than 0.3 ms. The proposed approach considerably relaxes the constraint of designing flaw-free initial (and static) bit and power profiles.

The experimental results showed a BER improvement consistent with the simulation results. (An exact match was not expected because of the differing noise conditions and spectral efficiencies.) Even though it is difficult to make direct comparisons with the works in [5], [7], and [8], we note that the present results demonstrate a higher spectral efficiency than [5], the same spectral efficiency but a lower BER at a higher channel loss than [7], and a larger channel loss compensation than [8]. With its performance optimized in this way, DMT may be useful in wireline links at data rates beyond 100 Gb/s since its spectral efficiency, 2.5 b/sample in this work, is higher than that of 4-PAM, allowing it to use lower-sample-rate data converters, yet it is more robust than 6-PAM and 8-PAM, which have shown high sensitivity to practical transceiver impairments [43], [44]. Additionally, recent implementations of DMT for high-speed wireline links [7], [8] and a power estimate provided in [5] show competitive power consumption compared to PAM-based transceivers. In particular, with similar system parameters to those in the present article, [5] estimated energy efficiency values of 2.4 pJ/b and 2.03 pJ/b for the receiver ADC and DFT, respectively. Also, comparing an estimate of transmitter power in DMT to that of a 4-PAM system with an 8-tap feed-forward equalizer (FFE), it concluded that the equalizer part of the latter requires only about 15% less logic than the IDFT of the former, assumed to have a similar power consumption to the receiver DFT.

ACKNOWLEDGMENT
The authors would like to thank Huawei Technologies Canada for helpful discussions and the lending of a test channel and also would like to thank the CMC Microsystems for test equipment loan. In particular, comments from Dr. Hossein Shakiba have directly influenced the content of the paper.

REFERENCES
[1] J. A. C. Bingham, “Multicarrier modulation for data transmission: An idea whose time has come,” IEEE Commun. Mag., vol. 28, no. 5, pp. 5–14, May 1990.
[2] J. Cosson-Martin, H. Shakiba, and A. Sheikholeslami, “Timing recovery and adaptive equalization for discrete multi-tone signalling in wireline applications,” IEEE Open J. Circuits Syst., vol. 2, pp. 856–868, 2021.
[3] J. Zhang, J. Liu, Z. Lin, J. Liu, L. Shen, and S. Yu, “Nonlinearity-aware adaptive bit and power loading DMT transmission over low-crosstalk ring-core fiber with mode group multiplexing,” J. Lightw. Technol., vol. 38, no. 21, pp. 5875–5882, Nov. 1, 2020.
[4] K. Gharibdoust, A. Tajalli, and Y. Leblebici, “Hybrid NRZ/multi-tone serial data transceiver for multi-drop memory interfaces,” IEEE J. Solid-State Circuits, vol. 50, no. 12, pp. 3133–3144, Oct. 2015.
[5] B. Vatankhahghadim, N. Wary, J. Bailey, and A. C. Carusone, “A study of discrete multitone modulation for wireline links beyond 100 Gb/s,” IEEE Open J. Circuits Syst., vol. 2, pp. 78–90, 2021.
[6] B. Vatankhahghadim, N. Wary, and A. C. Carusone, “Discrete multitone signalling for wireline communication,” in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), Oct. 2020, pp. 1–5.
[7] G. Kim, L. Kull, D. Luu, M. Braendli, C. Menolli, P.-A. Francesc, H. Yueksel, C. Aprile, T. Morf, M. Kossel, A. Ceverero, I. Ožkaya, A. Burg, T. Toifl, and Y. Leblebici, “A 161 mW 56 Gbh/s ADC-based discrete multitone wireline receiver data-path in 14 nm FinFET,” in Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC), Feb. 2019, pp. 476–477.
[8] T. Dickson et al., “A 72 GSa/s, 8-bit DAC-based wireline transmitter in 4 nm FinFET CMOS for 200+Gb/s serial links,” in Proc. IEEE Symp. VLSI Technol. Circuits (VLSI Technol. Circuits), Jun. 2022, pp. 28–29.
[9] G. Kim, W. Kwon, T. Toifl, Y. Leblebici, and H.-M. Bae, “Design considerations and performance trade-offs for 56 Gb/s discrete multi-tone electrical link,” in Proc. IEEE 62nd Int. Midwest Symp. Circuits Syst. (MWSCAS), Aug. 2019, pp. 1147–1150.
[10] G. Kim, “200 Gb/s: Beyond PAM4,” in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, 2022.
[11] J. Salinas, J. Cosson-Martin, M. Laghaei, H. Shakiba, and A. Sheikholeslami, “Performance comparison of baseband signaling and discrete multi-tone for wireline communication,” IEEE Open J. Circuits Syst., vol. 2, pp. 65–77, 2021.
[12] T. Beukema, “Signalling, modulation, coding and signal processing architectures to support 400GE and beyond,” in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, Feb. 2021. [Online]. Available: http://research.ibm.com/publications/signalling-modulation-coding-and-signal-processing-architectures-to-support-400ge-and-beyond

[13] M.-A. LaCroix, E. Chong, W. Shen, E. Nir, F. A. Musa, H. Mei, M.-M. Mohsenpour, S. Zebdeh, V. Zamanlooly, C. Carvalho, X. Qin, D. Petrov, H. Wong, H. Ho, Y. Xu, S. N. Shahi, P. Krotnov, C. Feist, H. Huang, and D. Tonietto, “A 116 Gb/s DSP-based wideband transceiver in 7 nm CMOS achieving 6 g/B at 45 dB loss in PAM-4/Dual-PAM-4 and 52 dB in PAM-2,” in Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC), Feb. 2021, pp. 132–133.

[14] T. Ali, E. Chen, H. Park, R. Yousry, Y.-M. Ying, M. Abdullahi, M. Gandara, C.-C. Liu, P.-W. Weng, H.-S. Chen, M. Elbady, Q. Nehler, K.-H. Tsai, K. Tan, Y.-C. Huang, C.-H. Tsai, Y. Chang, and Y.-H. Tung, “A 460 mW 112 Gb/s DSP-based transceiver with 38 dB loss compensation for next-generation data centers in 7 nm FinFET technology,” in Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC), Feb. 2020, pp. 118–119.

[15] M. Pisati et al., “A sub-250 mW 51-to-56 Gb/s continuous-range PAM-4 42.5 Gb/s IL ADC/DAC-based transceiver in 7 nm FinFET,” in Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC), Feb. 2019, pp. 116–117.

[16] J. Bailey, H. Shakiba, E. Nir, G. Marderfeld, P. Krotnov, M.-A. LaCroix, D. Cassan, and D. Tonietto, “A 112-Gb/s PAM-4 low-power nine-tap sliding-block DFE in a 7-nm FinFET wideband receiver,” IEEE J. Solid-State Circuits, vol. 57, no. 1, pp. 32–43, Jan. 2022.

[17] D. J. Love and R. W. Heath, “OFDM power loading using limited feedback,” IEEE Trans. Veh. Technol., vol. 49, no. 5, pp. 1773–1780, Sep. 2005.

[18] S. Edininger and C. Bauer, “Concepts of dynamic adaptive DMT modulation,” in Proc. Int. Symp. Circuits, Control Signal Process. (ISCSCP), Mar. 2006, pp. 1–4.

[19] E.-H. Chen, J. Ren, B. Leibowitz, H.-C. Lee, Q. Lin, K. S. Oh, F. Lambrecht, V. Stojanovic, J. Zerbe, and C.-K. K. Yang, “A near-optimal BERScope Differential ISI Board: Precision ISI Test Accessory,” in Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC), Feb. 1995, pp. 737–735.

[20] P. S. Chow, “Bandwidth optimized digital transmission techniques for spectrally shaped channels with impulse noise,” Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, 1993.

[21] C.-C. Liu, P.-S. Weng, H.-S. Chen, M. Elbady, Q. Nehler, K.-H. Tsai, K. Tan, Y.-C. Huang, C.-H. Tsai, Y. Chang, and Y.-H. Tung, “A 460 mW 112 Gb/s DSP-based transceiver with 38 dB loss compensation for next-generation data centers in 7 nm FinFET technology,” in Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC), Feb. 2020, pp. 118–119.

[22] M. Yang, S. Shahramian, H. Song, P. Krotnov, and A. Carusone, “P-Pre-FEC and post-FEC BER as criteria for optimizing wireline transceivers,” in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), May 2021, pp. 1–5.

[23] E.-H. Chen, J. Ren, B. Leibowitz, H.-C. Lee, Q. Lin, K. S. Oh, F. Lambrecht, V. Stojanovic, J. Zerbe, and C.-K. K. Yang, “Near-optimal equalizer and timing adaptation for I/O links using a BER-based metric,” IEEE J. Solid-State Circuits, vol. 43, no. 9, pp. 2144–2156, Sep. 2008.

[24] M. Bichan, M. Hossain, and A. Carusone, “Frequency-division bidirectional communication over chip-to-chip channels,” IEEE Trans. Adv. Packag., vol. 32, no. 2, pp. 298–305, May 2009.

[25] R. Mellitz. (2019). Synthesized 2 m QSFP CR channels: End to end IL 28.5 dB and cable assembly and IL 19.75 dB. [Online]. Available: https://www.ieee802.org/3/cp/public/19_11/mellitz_3ck_03_1119.pdf

[26] R. Mellitz. (2022). COM 3.70 With 3.1 Configuration Updates, Fixes, and Exploratory Features Including Package/Die Loud Ladder. [Online]. Available: https://www.ieee802.org/3/cp/public/adhoc/mar23_22_mellitz_3ck_adhoc_01_032322.pdf

[27] Keysight Technologies. MB194A 120 GS/s Arbitrary Waveform Generator Data Sheet. Accessed: Nov. 13, 2022. [Online]. Available: https://www.keysight.com/us/en/assets/2018-06341/data-sheets/5992-3361.pdf

[28] Tektronix. Scalable Perform. Oscilloscopes. DPI70000SX Series Datasheet. Accessed: Nov. 13, 2022. [Online]. Available: https://download.tek.com/datasheet/DP70000SX-Datasheet-FINAL-13July2022-small.pdf

[29] Tektronix. BERTScope Differential ISI Board: Precision ISI Test Accessory. Accessed: Nov. 13, 2022. [Online]. Available: https://download.tek.com/datasheet/B612000ISI-Datasheet-65W600451.pdf

[30] R. Mellitz. (2019). BERTScope Differential ISI Board: Precision ISI Test Accessory. Accessed: Nov. 13, 2022. [Online]. Available: https://www.mathworks.com/help/signal-integrity/ug/channel-operating-margxin.html

[31] J. Ali, “DSP based transceivers for 200 Gb/s: Challenges and the path forward,” in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, 2022.

[32] Z. Li, L. Wang, and H.-C. Kao, “BER-based back-off optimization in OFDM systems under ideal pre-distorters,” IEEE Trans. Veh. Technol., vol. 67, no. 1, pp. 284–297, Jan. 2020.

[33] B. Vatankhahghadim et al.: Spectrally Efficient DMT Operation With BER-Informed Dynamic Bit and Power Loading
MASUM HOSSAIN (Member, IEEE) received the B.Sc. degree from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2002, the M.Sc. degree from Queen’s University, Kingston, ON, Canada, in 2005, and the Ph.D. degree from the University of Toronto, Toronto, ON, Canada, in 2010.

From 2008 to 2010, he was with the Analog and Mixed Signal Division, Gennum Corporation, Burlington, ON, Canada. From 2010 to 2012, he was with the Rambus Laboratory, Sunnyvale, CA, USA, as a Senior Member of Technical Staff. In 2013, he joined the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada.

Dr. Hossain was a recipient of the Best Student Paper Award at the 2008 IEEE Custom Integrated Circuits Conference and the Analog Devices Outstanding Student Designer Award, in 2010. In 2021, he received EPS Society Nominated Best Paper Award in IEEE TRANSACTIONS ON COMPONENTS, PACKAGING AND MANUFACTURING TECHNOLOGY.

ANTHONY CHAN CARUSONE (Fellow, IEEE) received the Ph.D. degree from the University of Toronto, in 2002. He has been a Professor with the Department of Electrical and Computer Engineering, University of Toronto. He has also been a Consultant to Industry in the areas of integrated circuit design and digital communication, since 1997. He is currently the Chief Technology Officer of the Alphawave IP Group, Toronto, Canada.

He has coauthored best student papers at the 2007, 2008, 2011, and 2022 Custom Integrated Circuits Conferences, the Best Invited Paper at the 2010 Custom Integrated Circuits Conference, the Best Paper at the 2005 Compound Semiconductor Integrated Circuits Symposium, the Best Young Scientist Paper at the 2014 European Solid-State Circuits Conference, and the Best Paper at DesignCon 2021. He coauthored the popular textbooks Analog Integrated Circuit Design (along with D. Johns and K. Martin) and Microelectronic Circuits (8th Edition) (along with A. Sedra, K. C. Smith and V. Gaudet). He was the Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS, in 2009, and an Associate Editor for the IEEE JOURNAL OF SOLID-STATE CIRCUITS (2010–2017). He was a Distinguished Lecturer for the IEEE Solid-State Circuits Society (2015–2017) and has served on the Technical Program Committee of several IEEE conferences, including the International Solid-State Circuits Conference (2016–2021). He is currently the Editor-in-Chief of the IEEE SOLID-STATE CIRCUITS LETTERS.