Multi-tone Signaling for High-speed Backplane Electrical Links

Amir Amirkhany[1], Vladimir Stojanović[1,2], and Mark A. Horowitz[1]

[1]Department of Electrical Engineering, Stanford University, CA 94305, USA

[2]Rambus, Inc., Los Altos, CA 94022, USA

***Abstract*—A Multi-tone architecture is proposed for high-speed backplane serial links. To limit complexity, the links use analog multi-tone rather than the more modern DMT. The tradeoffs involved in the design of such a system are examined and the performance of a serial link based on this approach is compared to a baseband architecture in terms of data rate and complexity using a convex optimization framework. Slightly less than 2x improvement in data rate at reasonable complexity is shown to be achievable with the proposed architecture.**

Keywords-high-speed; electrical; links; multi-tone; convex optimization

I. INTRODUCTION

The required bit rates for backplane links, links that connect different boards in a system, continue to increase. Current systems use links running at 1-3.25Gb/s, and components that provide 6-12Gb/s have been demonstrated [1]. Although the transmission media in backplane links are copper traces that regularly do not exceed 10's of centimeters in length, signal reflections originating from impedance mismatches at the via stubs and signal power loss due to skin effect and dielectric loss create a challenging environment for signal transmission at 10's of gigabits per second.

As shown in Fig. 1, while from a digital communication perspective links channels have relatively well behaved spectra, implementation constraints make the design of these links challenging. In particular backplane links need to operate at Gsymbol/s rates and be both power and area efficient. As a result, production links only use baseband (BB) signaling with linear transmit precoding[1], while the next generation links have recently started using either Decision Feedback Equalization (DFE) or multi-level PAM constellations [1-3] to improve performance. However, as BB signaling extends to the regions of the channel with high signal attenuation, the marginal increase in rate with increased complexity gets smaller and smaller. In fact, our recent results show that there is a large gap between the fundamental limits of signaling on PCB traces and the rates achievable using equalization-based BB communication techniques [4]. Besides this fundamental limit, as signaling rates approach the fundamental limits of the available fabrication technology, BB processing becomes less efficient in terms of area and power.

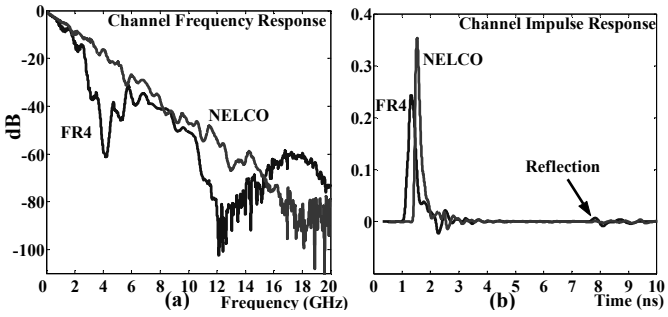

Fig. 1: Frequency Responses (a) and impulse responses (b) of typical link channels. The notches in the frequency response of the FR4 channel and the reflection taps towards the tail of its impulse response are the effects of via stubs.

This paper explores the feasibility of multi-tone (MT) signaling to reduce the existing performance gap. Since both analog to digital converters (ADC) and signal processors that run at Gsample/Sec rates are expensive, we particularly focus on a simple form of MT, which we call analog multi-tone (AMT), a technique from the 1960s [5]. However, as opposed to [5], in order to be relatively immune to channel distortions, we keep the overlap between the neighboring sub-channels to a minimum and try to cancel the remaining inter-channel interference (ICI). In particular, we demonstrate that an AMT system with (matrix) precoding at the transmitter outperforms a BB link with transmit precoding by about a factor of two for a modest additional complexity. Better ICI cancellation techniques might lead to even better performance while maintaining a reasonable cost.

AMT was almost abandoned after the development of high-performance DSP's and introduction of discrete multi-tone. In

Research supported by the MARCO Interconnect Focus Center and Rambus, Inc.

[1] It is the ease of building high-speed digital to analog converters compared to analog to digital converters that causes the links to use transmit precoding.

Section II, we will review and compare the general class of discrete multi-tone techniques with the AMT approach based on link-specific design constraints and point out the important implementation issues involved in the design of both systems.

An AMT system essentially creates a number of "baseband" channels that are frequency division multiplexed on the same wire. In Section III we will extend our convex optimization framework for baseband channels [4] to AMT systems with matrix transmit precoding. Since many links must run at very low BER, we use accurate noise models, and not their Gaussian approximations, in our analysis.

There are many design parameters in the implementation of an AMT system that significantly affect the performance, and the analysis of such a system is not possible without good understanding of those effects. Besides, the large number of parameters involved allows for many tradeoffs that the designer can use to advantage. Section IV enumerates the design parameters and points out the possible tradeoffs in some detail. Using the optimization framework from Section III and the insight from Section IV, the performance and complexity of an AMT system with matrix precoding is compared with a BB system with linear precoding in Section V and final concluding remarks are presented in Section VI.

Since the purpose of this paper is to demonstrate the feasibility of the MT approach and perform a fair comparison with BB rather than maximizing the throughput, we will mostly favor simplicity in our analysis and in the choice of system configurations. We will, however, occasionally comment on some of the opportunities to further boost performance.

II. Multi-tone approaches for high-speed links

A. Discrete Multi-tone

Discrete Multi-tone (DMT) [6] and other DFT-based MT techniques, like Filtered MT (FMT) [7], perform the necessary frequency up-conversion and down-conversion for MT operation in digital domain using inverse Discrete Fourier Transform (IDFT) at the transmitter (Tx) and DFT at the receiver (Rx). For this reason, all such techniques share the requirement for high resolution, high bandwidth (BW) ADC's at the Rx front-end. Unfortunately, building high resolution ADC's at several gigahertz frequency is a major challenge, and if the ADC does not have enough sampling resolution, quantization noise becomes a serious limiting factor.

To estimate the required ADC bandwidth and resolution, we use the capacity results from Stojanovic [4]. He shows that the channels in Fig. 1 have about 10GHz of usable bandwidth for signal transmission. Therefore, a regular discrete implementation would require a 20Gsamples/Sec ADC at the receiver. Thermal noise resulting from a 50Ω termination resistor at the receiver is about $1nV/\sqrt{Hz}$ and assuming about 13dB noise figure for the receiver sampler, a typical number for a BB link, gives a noise sigma of 0.45mV. Further assuming that the output transmitter swing is 1V, the ADC should at least have 10-bit resolution in order to have non-dominant quantization noise at the receiver. A few additional bits might be required as well to push quantization noise well below thermal noise level. A state of the art ADC only supports 8-bit quantization at 20Gsamples/Sec [8] and is far too complex for a serial link.

DFT-based MT techniques preserve the orthogonality of the sub-channels by adding a cyclic prefix to the time-domain DFT block, like DMT, or by avoiding spectral overlap between the sub-channels, like the so called 'non-critically sampled' FMT. For link systems, due to the long impulse response caused by the reflections, Fig. 1(b), the overhead for adding a cyclic prefix would be huge unless channel shortening filters are used. However, an FMT approach similar to [9], with relatively small number of sub-channels and per sub-channel precoding, shares many advantages of the AMT, as described in the following section, and could have been feasible if the ADC requirements did not exist.

B. Analog Multi-tone

A different approach to MT is to create several parallel BB streams of data and up-convert and down-convert them to the sub-channel frequencies using mixers. Regular BB equalization techniques can then be applied to each sub-channel to compensate for the remaining dispersion. We will refer to this approach as Analog MT (AMT).

Fig. 2 shows a general block diagram of such a system. Depending on the architecture, additional mixers or bandpass filters (BPF) might exist in the actual implementation, but the lowpass filters are architecture-independent and are referred to as channel-select filters. The signal processing blocks in the transmitter and the receiver are also architecture dependent and might not exist.

One clear advantage of AMT is that signal-processing rate in each sub-channel is a fraction of the overall signaling rate.

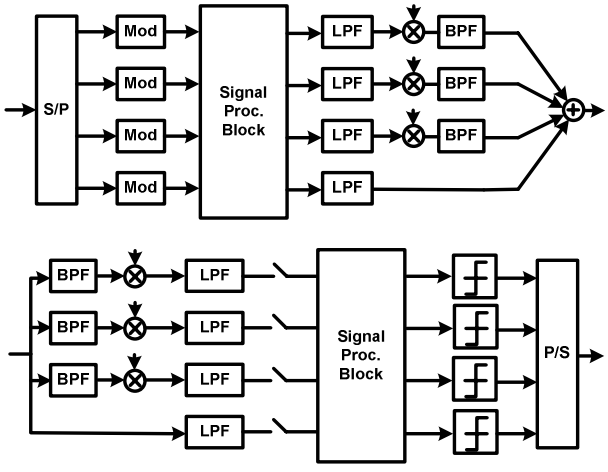

Fig. 2: General block diagram of an AMT transmitter (top) and receiver (bottom). Some blocks might not exist in a specific architecture.

Therefore, more sophisticated algorithms can be implemented compared to BB. The second advantage is that all existing experience in BB link design is readily applicable to each sub-channel; an appealing property for the industry. AMT also fits very well to the characteristics of link channels. Due to the smooth roll-off of link channels, it is expected that not many sub-channels are necessary to achieve the desired low equalization loss. Consequently, the overall complexity overhead due to the Radio Frequency (RF) components in each sub-channel could be tolerable. Nonetheless, some challenging implementation constraints exist as well.

One main issue in implementing an AMT system is to deal with inter-channel interference (ICI). Even though it is theoretically possible to eliminate ICI over an ideal channel if channel-select filters follow certain orthogonality conditions, due to channel distortions and bad tolerance of analog on-chip filters the actual ICI level might become intolerable if significant overlap is allowed between the sub-channels [5]. On the other hand, building several sharp on-chip analog filters that would allow non-overlapping sub-channels without loss of spectral efficiency is not easy as well, due to the bad quality of on-chip inductors and the large size of on-chip capacitors. A compromise is, therefore, to allow sufficient guard-band between the sub-channels so that the overlap is small and try to cancel the remaining ICI with signal processing. One good ICI cancellation method which is a natural extension of transmit precoding in BB links is matrix transmit precoding.

Besides ICI cancellation, implementation of good quality low noise linear gain elements and mixers is also important. The following section establishes an appropriate framework that helps us quantify the requirements on the analog front-end as well as the channel-select filters.

III. CONVEX OPTIMIZATION FRAMEWORK

We recently derived a convex optimization framework for minimizing the Bit Error Ratio (BER) in BB links that took into account accurate models for link-specific noise sources [4]. Since an AMT system is essentially a set of interacting BB links, we can easily extend that framework to AMT by writing the problem in vector form. However, we first need to model the system.

A. Signal Transmission Model

The baseband equivalent of the MT system with transmit matrix precoding is shown in Fig. 3. The model consists of a matrix precoder with finite length diagonal and cross precoders and a matrix channel. The matrix channel models the N sub-channels in the system and the interference between them. We will refer to the sub-channels as diagonal channels and to the interference channels as cross channels in the following discussion. At the receiver, data is sampled at sub-channel symbol rate and a decision is made immediately.

The formulation we will use in this paper is based on the NxN case of a system presented in [10]. Maximum ISI spread for all diagonal and cross channels, $\underline{\mathbf{p}}_{lj}$, is assumed to be ν, and

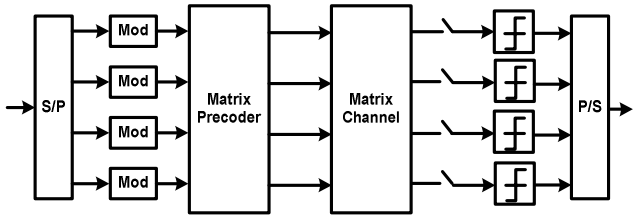

Fig. 3: Baseband equivalent of an AMT system with matrix precoding.

the length of all diagonal and cross precoders, $\underline{\mathbf{w}}_{lj}$, is L. The response of the system from input k to output l at sample time n can be expressed as:

$$\hat{x}_{lk}(n) = \underline{\mathbf{x}}_k^{\mathrm{T}}(n)\left(\mathbf{P}_{l1}\underline{\mathbf{w}}_{1k} + \ldots + \mathbf{P}_{lj}\underline{\mathbf{w}}_{jk} + \ldots + \mathbf{P}_{lN}\underline{\mathbf{w}}_{Nk}\right) \qquad (2)$$

where $\underline{\mathbf{x}}_k(n) = \left[x_k(n) \;\; \ldots \;\; x_k(n-L-\nu+1)\right]^{\mathrm{T}}_{(\mathbf{L}+\nu)\times\mathbf{1}}$ is the data vector at input k, $\underline{\mathbf{w}}_{jk}(n) = \left[w_{jk}(0) \;\; \ldots \;\; w_{jk}(L-1)\right]^{\mathrm{T}}_{\mathbf{L}\times\mathbf{1}}$ is the precoder from data input k to transmitter output j, and the channel convolution matrix $\mathbf{P}_{lj}$ is defined as:

$$\mathbf{P}_{lj} = \begin{bmatrix} p_{lj}(0) & \ldots & p_{lj}(\nu) & 0 & 0 & 0 \\ 0 & p_{lj}(0) & \ldots & p_{lj}(\nu) & 0 & 0 \\ 0 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & 0 & 0 & p_{lj}(0) & \ldots & p_{lj}(\nu) \end{bmatrix}^{\mathrm{T}}_{(\mathbf{L}+\nu)\times\mathbf{L}} \qquad (3)$$

The system is fully described by:

$$\hat{\underline{\mathbf{x}}}(n)_{\mathbf{N}\times\mathbf{1}} = \mathbf{X}(n)_{\mathbf{N}\times\mathbf{N}^2(\mathbf{L}+\nu)}\,\boldsymbol{\Psi}_{\mathbf{N}^2(\mathbf{L}+\nu)\times\mathbf{N}^2\mathbf{L}}\,\underline{\mathbf{W}}_{\mathbf{N}^2\mathbf{L}\times\mathbf{1}} + \underline{\mathbf{N}}(n)_{\mathbf{N}\times\mathbf{1}} \qquad (4)$$

where the channel matrix $\boldsymbol{\Psi}$ is defined as:

$$\boldsymbol{\Psi} = \begin{bmatrix} \mathbf{P} & & \mathbf{0} \\ & \ldots & \\ \mathbf{0} & & \mathbf{P} \end{bmatrix}_{\mathbf{N}^2(\mathbf{L}+\nu)\times\mathbf{N}^2\mathbf{L}} \qquad \mathbf{P} = \left[\mathbf{P}_1^{\mathrm{T}} \ldots \mathbf{P}_N^{\mathrm{T}}\right]^{\mathrm{T}}_{\mathbf{N}(\mathbf{L}+\nu)\times\mathbf{N}\mathbf{L}} \qquad (5)$$

with $\mathbf{P}_k = \left[\mathbf{P}_{k1}\cdots\mathbf{P}_{kN}\right], \forall k = 1\ldots N$, and the precoder filter $\underline{\mathbf{W}}$ is defined as:

$$\underline{\mathbf{W}} = \left[\underline{\mathbf{w}}_1^{\mathrm{T}} \;\; \ldots \;\; \underline{\mathbf{w}}_N^{\mathrm{T}}\right]^{\mathrm{T}}_{\mathbf{N}^2\mathbf{L}\times\mathbf{1}}, \; \underline{\mathbf{w}}_k = \left[\underline{\mathbf{w}}_{1k}^{\mathrm{T}} \;\; \ldots \;\; \underline{\mathbf{w}}_{Nk}^{\mathrm{T}}\right]^{\mathrm{T}}_{\mathbf{N}\mathbf{L}\times\mathbf{1}} \qquad (6)$$

and the input data matrix $\mathbf{X}(n)$ is defined as:

$$\mathbf{X}(n) = \left[\mathbf{X}_1(n) \;\; \ldots \;\; \mathbf{X}_N(n)\right]_{\mathbf{N}\times\mathbf{N}^2(\mathbf{L}+\nu)} \qquad (7)$$

$$\mathbf{X}_k(n) = \begin{bmatrix} \underline{\mathbf{x}}_k^{\mathrm{T}}(n) & & \underline{\mathbf{0}} \\ & \ldots & \\ \underline{\mathbf{0}} & & \underline{\mathbf{x}}_k^{\mathrm{T}}(n) \end{bmatrix}_{\mathbf{N}\times\mathbf{N}(\mathbf{L}+\nu)}$$

The vector $\underline{\mathbf{N}}(n)$ represents several noise terms including thermal noise, sampling jitter and carrier phase noise.

The spectral density of thermal noise originating from a 50Ω termination resistor is $(1nV)^2/Hz$ and the noise figure of all modules in the receiver data path including the sampler, gain elements, mixers and filters add to this value.

Carrier phase noise is the result of the random variations in the phase of the up/down converting oscillators at the Tx/Rx. Phase noise manifests itself as a voltage noise at the slicer input through the mixing of the real and imaginary parts of the QAM signal. The variance of the phase noise-induced voltage noise on the real/imaginary part of the signal, $\bar{\sigma}_C^2$, can be

described as:

$$\bar{\sigma}_C^2 = \left(\sigma^2_{carrier\_Tx} + \sigma^2_{carrier\_Rx}\right)\bar{E}_y \quad (8)$$

where $\bar{E}_y$ is received signal energy per dimension[2], and $\sigma^2_{carrier\_Tx/Rx}$ is the variance of phase variations at the Tx/Rx relative to the oscillator period.

Sampling jitter models the random variations in the phase of the signaling clock at the Tx and the sampling clock at the Rx. Ignoring the second order effects from the cross channels and cross precoders, we can directly use jitter expressions from Stojanovic [4] for the diagonal channels:

$$\bar{\sigma}^2_{J\_Tx/Rx,i} = \underline{\mathbf{w}}_{\mathbf{ii}}^{\mathbf{T}} \mathbf{S}_{\mathbf{i}}^{\mathbf{Tx/Rx}} \underline{\mathbf{w}}_{\mathbf{ii}},$$

$$\mathbf{S}_{\mathbf{i}}^{\mathbf{Tx}} = \bar{E}_{x,i} \sum_{j=-sbS}^{sbE} \sum_{k=-sbS}^{sbE} \mathbf{I}_{\mathbf{j-k}} [h_{ii,j-1}\ h_{ii,j}] \begin{bmatrix} R_{\varepsilon^{TX},i}(j-k) & -R_{\varepsilon^{TX},i}(j-k-1) \\ -R_{\varepsilon^{TX},i}(j-k+1) & R_{\varepsilon^{TX},i}(j-k) \end{bmatrix} \begin{bmatrix} h_{ii,k-1} \\ h_{ii,k} \end{bmatrix}$$

$$\mathbf{S}_{i}^{Rx} = \bar{E}_{x,i} R_{\varepsilon^{RX},i}(0) \sum_{j=-sbS}^{sbE} \sum_{k=-sbS}^{sbE} \mathbf{I}_{\mathbf{j-k}} [h_{ii,j-1}\ h_{ii,j}] \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} h_{ii,k-1} \\ h_{ii,k} \end{bmatrix} \quad (9)$$

where $\bar{\sigma}^2_{J\_Tx/Rx,i}$ is the jitter-induced voltage noise variance at the i^{th} sub-channel slicer input, $R_{\varepsilon^{Tx/Rx}}(m) = E(\varepsilon_k^{Tx/Rx}\varepsilon_{k+m}^{Tx/Rx})$ is the m^{th} sample of jitter correlation function, $\mathbf{I_n}$ is the identity matrix shifted right by n places, and $h_{ii,j}$ is the j^{th} sample of the impulse response of the i^{th} diagonal channel. $\bar{E}_{x,i}$ is the i^{th} sub-channel Tx signal energy per dimension, and *sbS* and *sbE* are start and end indices of the corresponding impulse response sequences.

B. Optimization Framework

Extending the convex optimization framework of [4] to a MT system with matrix transmit precoding as described in the previous section, we can obtain the following combinatorial optimization problem on the integer bit loading of the N sub-channels, $\bar{\underline{\mathbf{b}}} = [\bar{b}_1\ \bar{b}_2 \ldots \bar{b}_N]$, which reduces to a convex feasibility problem on $\mathbf{W}(\bar{\underline{\mathbf{b}}})$ for a given $\bar{\underline{\mathbf{b}}}$:

$$\text{Maximize DataRate} = \sum_{k=1}^{N} s_k \bar{b}_k$$

$$\text{Subject to } Q^{-1}\left(\frac{BER_{\text{target}}}{2s_k(1-2^{-\bar{b}_k})}\right)\bar{\sigma}_k - \sqrt{\frac{3}{2^{2\bar{b}_k}-1}}\left(\underline{\mathbf{1}}_{\Delta_k}^{\mathrm{T}}\boldsymbol{\Psi}\mathbf{W} - \frac{(2^{\bar{b}_k}-1)}{2}\left\|\underline{\mathbf{1}}_{PD}^{\mathrm{T}}\boldsymbol{\Psi}\mathbf{W}\right\|_1\right) + offset_k \le 0,\ k = 1\ldots N$$

$$\sum_{k=1}^{N}\sqrt{3s_k\frac{2^{\bar{b}_k}-1}{2^{\bar{b}_k}+1}}\left\|\underline{\mathbf{w}}_{\mathbf{kj}}\right\|_1 \le V_{peak},\ k = 1\ldots N \quad (10)$$

In the above inequalities $\underline{\mathbf{1}}_{\Delta_k}$ is a vector of length $N^2(L+\nu)$ that has all zero entries except at its $(N+1)(k-1)(L+\nu)+\Delta_k+1$ position where it is unity, $\underline{\mathbf{1}}_{PD}$ is a vector of the same size that has unity entries at the positions corresponding to the taps that are to be considered for peak distortion, and

$$\bar{\sigma}_k^2 = \mathbf{W}^{\mathbf{T}}\boldsymbol{\Psi}^{\mathbf{T}}\left(\mathbf{D}_k - \underline{\mathbf{1}}_{\Delta_k}\underline{\mathbf{1}}_{\Delta_k}^{\mathrm{T}}\right)\boldsymbol{\Psi}\mathbf{W} + \sigma^2_{Carrier,k}\mathbf{W}^{\mathbf{T}}\boldsymbol{\Psi}^{\mathbf{T}}\mathbf{D}_k\boldsymbol{\Psi}\mathbf{W} + \mathbf{W}_{\mathbf{kk}}^{\mathbf{T}}\mathbf{S}_{\mathbf{Jitter_k}}\mathbf{W}_{\mathbf{kk}} + \bar{\sigma}^2_{Thermal,k}$$

$$\sigma^2_{Carrier,k} = \sigma^2_{Tx\_Carrier,k} + \sigma^2_{Rx\_Carrier,k}, \quad \mathbf{S}_{\mathbf{Jitter_k}} = \mathbf{S}_{\mathbf{Tx\_Jitter_k}} + \mathbf{S}_{\mathbf{Rx\_Jitter_k}}$$

$$\mathbf{D}_{\mathbf{k}} = \frac{1}{s_k}E(\mathbf{X}^{\mathrm{T}}\mathbf{X}) = \frac{1}{s_k}diag\left(s_1\underline{\mathbf{1}}_{\Delta_1}^{\mathrm{T}}\ \ s_2\underline{\mathbf{1}}_{\Delta_2}^{\mathrm{T}}\ \ \ldots\ \ s_N\underline{\mathbf{1}}_{\Delta_N}^{\mathrm{T}}\right) \quad (11)$$

Δ_k is the main tap delay of the k^{th} sub-channel and s_k is 1 for the BB sub-channel and 2 for the PB sub-channels. Finally, offset_k is the total slicer sensitivity for the k^{th} sub-channel. We have also set $\bar{E}(x_k(n)) = 1$ (k=1,2,…,N) for all sub-channels, absorbing all the gain coefficients into the precoder.

The first set of inequalities in the feasibility problem are bit error ratio constraints for the sub-channels and the second set are peak voltage constraints at the Tx. Peak voltage constraint originates from the headroom requirement of the transmitter circuits, such that they maintain linearity of operation without introducing non-linear distortion. In the AMT architecture, the last points in the Tx data path where transistors are used are potentially the mixers, and therefore, it is essential that the peak voltages at the input to the mixers remain within the allowed range. However, the peak voltage expansion of the fixed analog LPF's preceding the mixers in the Tx data path is fixed and known through simulation. Therefore, the peak voltage constraint can alternatively be applied at the input to those analog LPF's.

Although the optimization problem is combinatorial, it is possible to obtain a tight upper bound for the $\bar{b}$'s by upper bounding the individual sub-channels in isolation when ICI does not exist. In a practical situation, due to complexity constraints, number of sub-channels is small and it is desired that ICI does not have a significant effect on the performance of the sub-channels. Consequently, starting from the upper bound, the maximum achievable bit loading can be obtained in reasonably small number of iterations.

Again following Stojanovic's procedure in [4], we can readily add per sub-channel DFE to the optimization framework; however, since the goal of this paper is to demonstrate the fundamental advantages of AMT to BB, we will only focus on linear transmit precoding in the subsequent sections.

IV. System Parameters and Design Tradeoffs

An AMT system has many design parameters that have to be determined before an optimization is performed. These parameters include the bandwidth of the sub-channels, channel-select filter type, order and bandwidth, and the guard bandwidth between the sub-channels. Due to the combinatorial nature of the optimization framework, a simple sweep of all these parameters is virtually impossible and provides little insight to the tradeoffs involved. An alternative is to try to quantify the effect of all the parameters in isolation and in a progressive manner. This is the procedure we will take on in this section.

Signal to Noise Ratio (SNR) is generally very high in link

[2] Variables represented with a bar are all per dimension, i.e. per real/imaginary part of the signal.

channels, at least in the low frequency range. Since in a MT implementation a large fraction of the overall data rate comes from the high-SNR portions of the channel, it is essential that the channel-select filters do not change the quality of the sub-channels by inducing ISI. Unfortunately regular analog filters are not Nyquist filters in general and can potentially affect the performance.

Analog filters follow a general tradeoff between the sharpness of their stopband fall-off and the linearity of their passband phase. The sharper the fall-off, the more nonlinear the phase, and consequently, the more ISI the filter can induce. There is also a general tradeoff between the bandwidth of a filter and the amount of ISI it can induce. Roughly speaking, the larger the BW, the shorter the time domain response of the filter, and the less ISI it creates when convolved with the channel. Simple Signal to Interference Ratio (SIR) analysis reveals that channel-select filters with moderate phase properties, like Butterworth (maximally flat) filters, induce less ISI compared to linear phase filters, like Bessel, or very sharp filters, like Chebyshev, when compared at same stopband roll-off. Consequently, we will focus on analog Butterworth filters to illustrate the tradeoffs.

Fig. 4(a) illustrates the tradeoff between filter characteristics and filter induced ISI. In order to have a quantitative measure for ISI, this figure plots the minimum number of required diagonal equalizer taps to achieve certain performance level versus filter excess bandwidth, where excess bandwidth is measured with respect to the Nyquist bandwidth. To generate the curves optimization is performed on a single passband channel and target performance is set to 3 bits per dimension[3] by setting the noise level. Clearly, as the signal attenuation increases in higher frequency sub-channels and thermal noise becomes more dominant, more levels of ISI can be tolerated and therefore, the requirements we have obtained for 3 bits/dimension performance level serve as an upper bound for high frequency sub-channels.

There is also a tradeoff between guard bandwidth and number of required cross equalizer taps to achieve a target performance. Fig. 4(b) illustrates this tradeoff. In order to generate the curves, optimization is performed on three adjacent sub-channels. The sub-channels each support 3 bits per dimension if they operate in isolation. Minimum number of required cross equalizer taps is thus obtained so that same performance is achieved when they operate together. As could be observed, even when a 4th order filter is used, the matrix precoder with only a few number of taps is able to cancel ICI to the desired level if moderate guard bandwidth is allowed between the sub-channels.

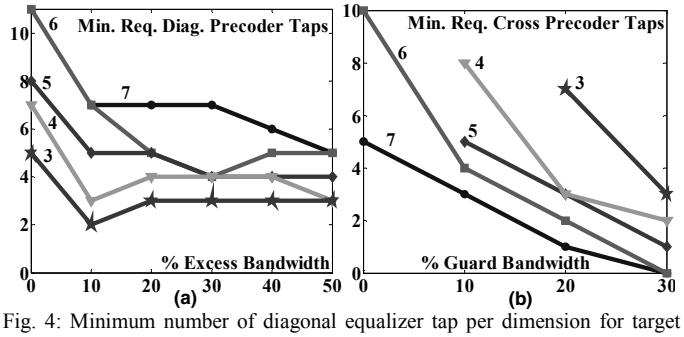

Fig. 4: Minimum number of diagonal equalizer tap per dimension for target performances of 3bits/dimension (a) and Minimum number of cross equalizer taps per dimension per neighbor for target performances of 3bits/dimension, (b). Butterworth filters are used both at the Tx and the Rx and ideal channel is assumed for both figures.

It is clear from Fig. 4 that although sharper filters need less cross precoder taps, they create more ISI, and hence, need more diagonal precoder taps, and a tradeoff exists. Using the data presented in Fig. 4, it is possible to plot the minimum overall required precoder taps (cross plus diagonal) versus the overall BW penalty (excess plus guard). Results indicate that for overall precoder complexity of 5 to 15 taps, 4th order filter induces the least BW penalty (between 40% to 20%). An overall complexity of 10 taps corresponds to 5% excess BW and 20% guard BW, and this is the configuration we choose for our analysis.

The appropriate bandwidth for the sub-channels very much depends on the channel characteristics and implementation constraints. For link channels similar to Fig. 1, ISI cancellation loss over 0.75GHz wide sub-channels is negligible and signal processing rate is low enough that allows using sophisticated algorithms if desired. Consequently reducing sub-channel bandwidth beyond 0.75GHz does not boost the performance while it increases the complexity overhead due to the RF front-end. On the other hand, increasing sub-channel BW beyond 1.25GHz leads to poor performance due to the deep notches in the channels. A notch in the channel can affect two neighboring sub-channels and when the sub-channels are wide, the resulting loss in the overall data rate could be significant. Therefore, for the analysis in this paper, we will set sub-channel BW to 1GHz independent of the channel being analyzed.

V. Performance and Complexity Comparison with BB

We are now in a position to perform a performance evaluation and complexity analysis on an AMT system. Since link performance is very implementation dependent, we will have to make some reasonable assumptions.

- For comparison, we will assume that the noise figure (NF) from the circuitry in each sub-channel in the MT system, and the NF in the BB system are the same and equal to 13dB.
- We will assume dedicated ring oscillators with 5° rms jitter for up/down conversion in each sub-channel. In practice better performance is achievable at higher cost if tuned oscillators are used, or if the low frequency clocks are derived from a single high-frequency clock.
- The clock used for sampling in all sub-channels is derived from the highest frequency oscillator in the system, and consequently, the sampling clock jitter is a percentage of the high-frequency clock period.
- Since inserting gain elements in the data path can

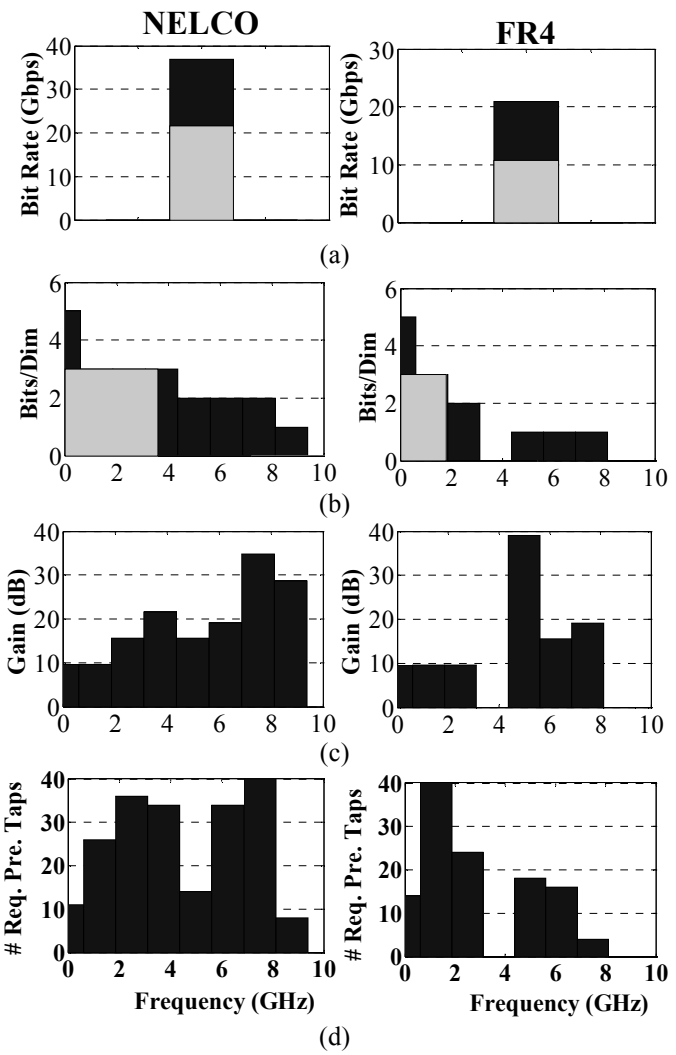

Fig. 5: (a) Total bit rate for MT and BB systems over NELCO (left) and FR4 (right) channels. (b) Bit loading over same channels. (c) Minimum required voltage gain per sub-channel. (d) Minimum number of required precoder taps (real and imaginary, diagonal and cross) per sub-channel.

proportionally reduce the effect of slicer sensitivity, we will assume negligible sensitivity for the slicers, but instead, will find the minimum required gain necessary for that assumption to hold if the actual sensitivity of the slicers was 10mV.

Fig. 5(b) shows sub-channel bit loading for the two channels in Fig. 1 when thermal noise, sampling jitter and carrier phase noise exist all together. Plotted in same figures are the bit loadings for BB systems operating at their optimum signaling-rates. The overall bit rates are shown in Fig. 5(a). Slightly less than 2x increase in data rate is achieved with MT over BB implementation.

As expected, over both channels carrier phase noise has mitigated the gain of MT in low frequency sub-channels. In other words, as long as both MT and BB systems are affected by signal dependent noise of the same order, there is no advantage in switching to MT[4]. However, a notch in the channel, FR4, prevents further increase in BB signaling rate without significant loss in performance while it has little effect on MT. Even if the channel does not have any notches, NELCO, as BB signaling rate increases, channel attenuation becomes so high that thermal noise becomes dominant, and therefore, MT prevails by suffering less from equalization loss.

Fig. 5(c) shows the minimum required voltage gain in each sub-channel so that our negligible slicer sensitivity assumption holds. For most sub-channels less than 20dB gain over a 1GHz bandwidth is sufficient to achieve the target performance. Part of this gain has to occur before the mixer so that the target NF of 13dB is realized, but the rest can occur at BB or any other convenient place in the data path. Part of the gain can also be absorbed in slicer sensitivity if ADC's with sensitivities less than 10mV can be designed at 1GHz data rate[5]. The required voltage gains for BB implementations are 12dB and 31dB over about 3.6GHz and 7.2GHz bandwidth for FR4 and NELCO respectively.

Fig. 5(d) shows the approximate minimum number of required precoder taps per sub-channel to achieve the performance shown in Fig. 5(b). Both real and imaginary dimensions for PB channels are counted. Overall, the two MT systems over FR4 and NELCO require about 100 and 200 precoder taps operating at 1GHz respectively. The sub-channels in the NELCO channels operate at higher rates compared to their counterparts in the FR4 channel, and therefore, require more ICI cancellation, and consequently, more cross precoder taps[6]. The BB systems over same channels require 32 and 77 taps operating at 3.6GHz and 7.2GHz respectively. Since power approximately scales with frequency, MT has a slight advantage over BB in terms of precoder complexity.

Finally, the MT system requires 15 ADC's operating at 1GHz while BB requires a single ADC operating at 3.6 GHz (7.2GHz) over FR4 (NELCO) channel.

In order to better understand the effect of thermal noise and carrier phase noise on the performance of a MT system, we now keep the same configuration as used for Fig. 5 and change thermal noise level and carrier phase noise independently. Fig. 6(a) shows the bit loadings for a MT system when NF is changed by ±6dB. Carrier phase noise has not changed. Fig. 6(b) shows same data for the case when carrier phase noise is changed by a factor of 2 (corresponding to ±6dB voltage noise) while NF is kept the same as in Fig. 5.

[4] To be exact, since interference contributes to both carrier phase noise and sampling jitter, MT would be affected less by these impairments, but the difference is not significant as long as interference is sufficiently canceled.

[5] In other words, required gain of g for one sub-channels means that the ratio of slicer sensitivity and signal gain in that sub-channel should be smaller than 10mV/g.

[6] Excessive number of precoder taps or required gain for some sub-channels indicates that they (or their neighboring sub-channels) are operating very close to the target BER, and therefore, require significant interference cancellation or gain. In general this can be fixed by assigning one less bit to those sub-channels (or their neighbors).

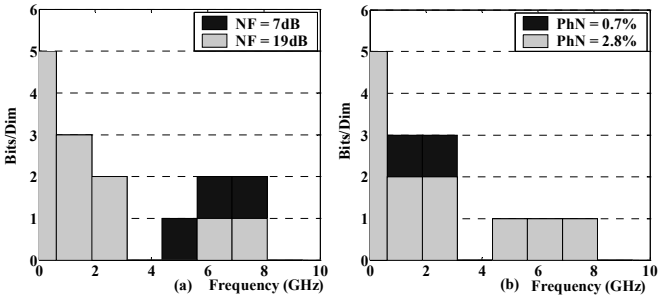

Fig. 6: Bit loadings for MT over FR4 channel when noise figure is changed by ±6dB, phase noise = 5° (a) and when carrier phase noise is changed by a factor of 2, noise figure = 13dB (b).

Comparison of the bit loadings in Fig. 6 with those in Fig. 5 reveals that the performance we obtained for MT in Fig. 5 is not very dependent on our assumptions about phase noise and noise figure. Furthermore, the bit loadings in Fig. 6 indicate that the change in the level of thermal noise only affects the high frequency sub-channels of the system while the change in carrier phase noise only affects the performance in the low frequency sub-channels. The reason is that the power of the voltage noise induced by carrier phase noise is proportional to received signal power while thermal noise power is signal independent. Therefore, in the low-frequency sub-channels where signal attenuation is low, thermal noise is much smaller than the noise induced by carrier phase noise. However, since for these sub-channels ICI power also depends on the strong neighboring sub-channels, reducing carrier phase noise improves the performance only as long as ICI is not dominant. Further analysis indicates that the performance of the 2nd sub-channel in Fig. 6 could have been improved by additional one bits/dimension if ICI was not present. As signal attenuation increases in high frequency sub-channels, both phase noise induced voltage noise and ICI scale with signal strength and thermal noise becomes dominant. A practical implication of this observation is that implementation constraints on clock quality are relaxed in the high frequency sub-channels where building high-quality circuits is difficult. However, RF front-end for same sub-channels should have good noise properties.

VI. CONCLUSION

Analog Multi-tone has the potential to significantly improve the performance of high-speed serial links beyond the limits achieved by baseband processing. The superiority of AMT originates from its fundamental capability to deal with ISI, as well as its inherent parallel structure that enables building better circuit components at lower frequencies. In fact, it is expected that better ICI cancellation techniques, when combined with better circuit components, lead to a lot better performance.

Due to the existence of RF components, implementation of an AMT system is a challenging task, however, our analysis indicates that the requirements are modest and far from being impossible. Therefore, the next main step would be to prove that RF circuits with the characteristics specified in this paper could be built in CMOS technology.

ACKNOWLEDGMENT

A. Amirkhany and V. Stojanović thank B. Nezamfar and I. Stojanović for full-hearted help on the paper and technical discussions.

REFERENCES

[1] J. Zerbe, C. Werner, V. Stojanović, F. Chen, J. Wei, G. Tsang, D. Kim, W. Stonecypher, A. Ho, T. Thrush, R. Kollipara , M. Horowitz, K. Donnelly, "Equalization and Clock Recovery for a 2.5 - 10Gb/s 2-PAM/4-PAM Backplane Transceiver Cell," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 12, pp. 2121-2130, Dec. 2003.

[2] S. Wu, S. Ramaswamy, B. Bhakta, P. Landman, R. Payne, V. Gupta, B. Parthasarathy, S. Deshpande, W. Lee, "Design of 6.25-Gbps Backplane SerDes with Top-Down Design Methodology," *DesignCon* 2004.

[3] J.T. Stonick *et al*, "An adaptive pam-4 5-Gb/s backplane transceiver in 0.25-µm CMOS," *IEEE J. Solid-State Circuits*, vol. 38, no. 3, pp. 436-443, March 2003.

[4] V. Stojanović, A. Amirkhany and M.A. Horowitz, "Optimal linear precoding with theoretical and practical data rates in high-speed serial-Link backplane communication," *IEEE International Conference on Communications,* June 2004.

[5] B. Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Transactions on Communications,* vol. 15, issue 6, pp 805-11, Dec. 1967.

[6] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency–division multiplexing using the Discrete Fourier Transform," *IEEE Transactions on Communications,* vol. 19, issue 5, pp 628-34, Oct. 1971.

[7] B. Hirosaki, "An orthogonally multiplexed QAM system using the Discrete Fourier Transform," *IEEE Transactions on Communications,* vol. 29, issue 6, pp 982-89, July 1981.

[8] K. Poulton *et al*, "A 20GS/s 8b ADC with a 1MB memory in 0.18µm CMOS," *IEEE International Solid-State Circuits Conference*, Feb. 2003, San Francisco.

[9] G. Cherubini, E. Eleftheriou, S. Oker and J. M. Cioffi, "Filter bank modulation techniques for very high speed digital subscriber lines," *IEEE Communications Magazine,* Vol. 38, issue 5, pp. 98-104, May 2000.

[10] P. A. Nelson, F. O-Bustamante and H. Hamada, "Inverse filter design and equalization zones in multichannel sound reproduction," *IEEE Transactions on Speech and Audio Processing*, vol. 3, no. 3, pp.185-92, May 1995.