FASTER-THAN-NYQUIST SIGNALING FOR NEXT GENERATION COMMUNICATION ARCHITECTURES

Andrea Modenini[*]*, Fredrik Rusek*[†]*, and Giulio Colavolpe*[*]

[*]University of Parma, Department of Information Engineering, Viale delle Scienze 181A, Parma, Italy
[†]Department of Electrical and Information Technology, Lund University, Lund, Sweden

ABSTRACT

We discuss a few promising applications of the faster-than-Nyquist (FTN) signaling technique. Although proposed in the mid 70s, thanks to recent extensions this technique is taking on a new lease of life. In particular, we will discuss its applications to satellite systems for broadcasting transmissions, optical long-haul transmissions, and next-generation cellular systems, possibly equipped with a large scale antenna system (LSAS) at the base stations (BSs). Moreover, based on measurements with a 128 element antenna array, we analyze the spectral efficiency that can be achieved with simple receiver solutions in single carrier LSAS systems.

1. INTRODUCTION

Mazo's faster-than-Nyquist (FTN) signaling is a linear modulation technique that reduces the time spacing between two adjacent pulses (the symbol time) well below that ensuring the Nyquist condition, thus introducing controlled intersymbol interference (ISI) [1, 2]. If the receiver can cope with the ISI, the efficiency of the communication system is increased. In the original papers on FTN signaling [1–3], this optimal time spacing is obtained as the smallest value giving no reduction of the minimum Euclidean distance with respect to the Nyquist case. This ensures that, asymptotically, the ISI-free bit-error ratio (BER) performance is reached when optimal detectors are used. More recently, this concept has been extended to multicarrier transmissions in [3]. In this case, intentional intercarrier interference (ICI) is also introduced by reducing the frequency separation among carriers.

Some scepticism can be raised against this technique. From a practical point of view, FTN may require an optimal detector whose complexity, however, easily becomes unmanageable. No hints are provided in the original papers on how to perform the optimization in the more practical scenario where a reduced-complexity receiver is employed. From a theoretical point of view, although this technique has been proposed to increase the spectral efficiency of a communication system, the BER is used as figure of merit in place of the spectral efficiency itself.

Before discussing ways to solve these problems, we need to introduce a few definitions. Let us consider a multicarrier transmission and let F be the frequency separation between two adjacent carriers. We will denote by $\boldsymbol{x}^{(\ell)} = \{x_k^{(\ell)}\}$ the input symbols, drawn from an $\mathcal{M}$-ary constellation $\mathcal{X}$, transmitted over the ℓth carrier using a linear modulation with symbol time T. Without loss of generality, we assume that all carriers use the same shaping pulse $p(t)$. In general, the values of T and F are such that both ISI and ICI are observed at the receiver. At the receiver side, a discrete-time set of sufficient statistics is extracted using a bank of matched filters and we denote by $\boldsymbol{y}^{(\ell)} = \{y_k^{(\ell)}\}$ the samples at the output of the matched filter for the ℓth carrier.

Depending on the allowed complexity at the receiver, different strategies can be adopted for detection. For example, the receiver can neglect both ICI and ISI and adopt a symbol-by-symbol detector. In other words, instead of the optimal receiver for the actual channel, we could adopt the optimal receiver for a simplified auxiliary channel for which the combined effect of ISI and ICI is modeled as a zero-mean Gaussian process independent of the additive thermal noise. Once the simplified receiver has been selected, suboptimal for the channel at hand but optimal for the considered auxiliary channel, we can compute, by using the technique in [4], a lower bound on the information rate for that channel. The information rate, also called constrained capacity, is the mutual information when the input symbols are constrained to belong to our finite constellation $\mathcal{X}$. According to mismatched detection [5], this lower bound is *achievable* by that particular suboptimal detector. The achievable spectral efficiency (ASE) is defined as the ratio between the achievable lower bound on the information rate and the product FT

$$\text{ASE} = \frac{I(\boldsymbol{x}^{(\ell)}; \boldsymbol{y}^{(\ell)})}{FT}$$

where F is a measure of the bandwidth of the given subcarrier.

The most recent extension of the FTN principle is thus *time-frequency packing* in [6], where it is proposed to optimize F and T in order to maximize the ASE. The idea is very simple: by reducing T and F the achievable information rate $I(\boldsymbol{x}^{(\ell)}; \boldsymbol{y}^{(\ell)})$ will certainly degrade due to the increased interference. However, the spectral efficiency, i.e., $I(\boldsymbol{x}^{(\ell)}; \boldsymbol{y}^{(\ell)})/FT$, can improve. Hence, the main quantity of interest is not the BER performance (in the uncoded case or for a fixed code). We may accept a degradation of the information rate provided the spectral efficiency is increased. In other words, if we keep the same code, an improvement can be obtained by using a code with lower rate. Improving the spectral efficiency without increasing the constellation size is convenient since low-order constellations are more robust to impairments such as phase noise and nonlinearities, whose effects are already increased by the higher transmitted power needed to obtain higher spectral efficiency values.

In [6], the main concepts are elucidated with reference to a symbol-by-symbol detector and the additive white Gaussian noise (AWGN) channel, working on the samples at the

matched filters output. More sophisticated receiver architectures are considered in [7] still with reference to the AWGN channel. In particular, a significant gain, in terms of spectral efficiency can be be obtained through a proper *channel shortening* (CS) linear front end plus a maximum a posteriori (MAP) symbol detector which copes with only a portion of the original ISI intentionally introduced at the transmitter [8]. Further gains can be obtained by using algorithms, which detect more than one carrier at a time.

In this paper, we discuss other interesting scenarios where this technique gives a significant improvement. In particular, in Section 2 we will summarize the results in [9] related to the application of FTN to improve the spectral efficiency of 2nd-generation satellite digital video broadcasting (DVB-S2) systems, whereas in Section 3 we will consider the application to long-haul optical systems [10, 11]. The application to LSAS built upon single carrier technology will be considered in Section 4, while Section 5 will consider the performance of time-frequency packing for 3GPP test channel conditions.

2. SATELLITE COMMUNICATIONS

FTN can be used to improve the spectral efficiency of present satellite systems for broadcasting applications and, in particular, of DVB-S2 systems. In this case, at the satellite transponder nonlinear distortions originate from the presence of a high-power amplifier (HPA). Input and output multiplexing (IMUX and OMUX) filters placed before and after the HPA have to be also considered.

For this application, investigated in [9], different carriers undergo independent amplification by different transponders on board the satellite. Hence, each transponder works with a single carrier occupying its entire bandwidth. This case is particularly relevant for digital broadcasting services since it allows a more efficient use of on-board resources (in particular the HPA can work closer to saturation). Since the frequency separation between two adjacent transponders cannot be modified, only time packing can be applied.

A heuristic version of time-frequency packing can be obtained by optimizing the bandwidth W of the shaping pulse $p(t)$. In fact, in this case ICI and ISI is increased due to adjacent users occupying different transponders and to the IMUX and OMUX filters. Whereas on the AWGN channel this optimization is implicit in TF packing, in the sense that we can obtain the same ICI by fixing F and increasing W or by fixing W and decreasing F, this is no more true for this scenario since the IMUX and OMUX bandwidths are kept fixed. Hence, an increased value of W also increases the ISI. The bandwidth optimization can thus be employed as an alternative to frequency packing in this case where the transponder frequency plan cannot be modified and, hence, frequency packing is not an option.

Although ICI from signals occupying adjacent transponders is present, mainly due to the nonlinear spectral regrowth, only single-user detection is considered at the receiver. In [9] two different approaches to detection for nonlinear channels are considered, namely the use of a detector taking into account the nonlinear effects and a more traditional scheme based on predistortion and memoryless detection. In the case of predistortion, the dynamic data predistortion technique described in [12] has been considered, whereas in the case of advanced detection the receiver in [13] employing the CS technique is used.

With a detector which takes into account a memory of only one symbol, and thus with a very limited complexity increase, it is possible to obtain a gain more than 30% in terms of spectral efficiency with respect to the conventional use of the current standard. A gain up to 40% can be reached by changing the transponders' frequency plan.

3. LONG-HAUL OPTICAL SYSTEMS

Another scenario where FTN can provide a significant gain is represented by long-haul coherent optical systems with polarization multiplexing. This case has been investigated both in the absence [10] or presence [11] of nonlinear impairments and bit rates of the order of 1 Tb/s. In this case, with the current hardware limitations, more than one carrier is required. Time-frequency packing outperforms other solutions based on higher-order modulations proposed in the literature, showing that when nonlinear effects are present, the spectral efficiency cannot be trivially increased by increasing the modulation order. This result is confirmed also for ultra-long-haul links, up to 10,000 km, whereas better spectral efficiency values can be achieved using modulations with very high cardinality, such as a 256-ary quadrature amplitude modulation (QAM), but only for short range links where fiber nonlinearities have a weak effect, or resorting to unfeasible compensation techniques such as ideal full-complexity digital back propagation. Simulation results for a polarization-multiplexed time-frequency-packed quaternary phase shift keying (QPSK) system which, on a realistic optical link, reaches a spectral efficiency of 7.5 b/s/Hz (a loss of less than 1 b/s/Hz from the information-theoretic results) have been reported.

4. PERFORMANCE OF FTN IN MEASURED LARGE ANTENNA ARRAY CHANNELS

We next study the performance of single carrier FTN signaling with LSAS. The context for our discussion is measurements of actual channels conducted at Lund University. For simplicity, we study the transmission between a BS and a single user equipped with a single antenna. Our objective is twofolded: (1) Study to what degree the LSAS can simplify the ISI detection at the user side, and (2) Study to what extent FTN can boost the capacity. Compared with previous sections of the paper, we shall be more elaborate as these results have not been presented before.

The received continuous time signal at the user is described by the following complex baseband equation

$$y(t) = \gamma \sum_{m=1}^{M} \int h_m(t-\tau) s_m(\tau) \mathrm{d}\tau + n(t), \qquad (1)$$

where M is the number of antennas at the BS, γ is a constant that normalizes the transmit power, $h_m(t)$ is the impulse response between the mth BS antenna and the user, $s_m(t)$ is the transmitted signal from the mth antenna, and $n(t)$ is white complex Gaussian noise with spectral density N_0. We assume a fully reciprocal channel. Via training signals sent from the

user the BS can learn the impulse responses $\{h_m(t)\}$, and perfect channel estimation is assumed for simplicity. The BS uses time-reversal precoding [14] so that $s_m(t)$ equals

$$s_m(t) = \int h_m^*(-\tau)x(t-\tau)\mathrm{d}\tau, \tag{2}$$

where $x(t)$ is the linearly modulated signal

$$x(t) = \sum_k x_k p(t-kT). \tag{3}$$

In (3), $\{x_k\}$ are the information symbols, $p(t)$ denotes the modulation pulse shape, and T is the symbol period. With an allocated bandwidth $[-W, W]$, the optimal choice of modulation pulse is a sinc pulse $p(t) = \sin(\pi 2Wt)/(\pi 2Wt)$ accompanied with the choice $T = 1/2W$. In the absence of multi-path, an orthogonal system results and the symbols $\{x_k\}$ can be detected independently. However, a sinc is impractical and typically the smoother root-raised cosine (RRC) pulse with roll-off factor β is used. If the RRC pulse occupies the same bandwidth $[-W, W]$ we get that, in absence of multi-path, $T = (1+\beta)/2W$ yields an ISI free system. Thus, in comparison with a sinc pulse, the RRC suffers a spectral efficiency loss of an amount $1/(1+\beta)$. FTN signaling aims at reducing the loss by setting $T < (1+\beta)/2W$ and accepting ISI.

Inserting (2) into (1) and a taking the Fourier transform of $y(t)$ give

$$\begin{aligned} Y(f) &= \gamma \sum_{m=1}^{M} |H_m(f)|^2 X(f) + N(f) \\ &= \bar{H}(f)X(f) + N(f). \end{aligned} \tag{4}$$

Under the assumption that the antenna elements experience independent small scale fading, the channel transfer function $\bar{H}(f) = \gamma \sum_{m=1}^{M} |H_m(f)|^2$ seen by the user can be expected to be almost flat as the number of antennas grow large. Thus, usage of a large antenna array at the BS can mitigate the multi-path of the channel, rendering nearly a pure AWGN channel at the user. In the time-domain, (4) corresponds to

$$y(t) = \int \bar{h}(\tau)x(t-\tau)\mathrm{d}\tau + n(t) = \sum_k x_k g(t-kT) + n(t),$$

where $\bar{h}(t)$ is the inverse transform of $\bar{H}(f)$ and $g(t) = \int \bar{h}(\tau)p(t-\tau)\mathrm{d}\tau$ is the combination of the modulation pulse and the effective channel impulse response. For the user, there are two ways to implement the receiver: (i) The user estimates $g(t)$ and proceeds accordingly, or (ii) The user is unaware of $g(t)$, and proceeds as if $\bar{h}(t) = \delta(t)$. Clearly, there is a major overhead of (i) compared with (ii), but performance is superior. However, with large antenna arrays, $|\bar{H}(f)|$ may be nearly flat, which means that the difference between (i) and (ii) may be negligible. This greatly simplifies system design, and as we will demonstrate later, it is indeed the case for the measurements we have available. We set the constant γ such that $g(t)$ has unit energy, i.e., $\int |g(t)|^2 \mathrm{d}t = 1$. This is feasible even for receiver strategy (ii) as the BS has full control over $g(t)$ even if the user is assumed unaware of it. Further, the array gain of the large antenna is now harvested as a power saving at the BS side. We shall not investigate this power saving any further, so the relevant measure of SNR is $1/N_0$; the array gain then comes as an extra bonus on top, but is not included in any performance plots. With the optimal ML receiver based on strategy (i) and unit energy complex Gaussian inputs, the limiting rate (in nats/sec) of the transceiver system is

$$C = \int_0^{1/2T} \log\left(1 + \frac{2G_{\mathrm{fo}}(f)}{N_0}\right) \mathrm{d}f, \tag{5}$$

where $G_{\mathrm{fo}}(f)$ is the folded spectrum

$$G_{\mathrm{fo}}(f) = \sum_{k=-\infty}^{\infty} \left| G\left(f + \frac{k}{T}\right)\right|^2, \quad |f] \leq \frac{1}{2T}.$$

What is extremely important to note here is that the capacity in (5) is linearly growing with decreased signaling rate T as the support of the integral grows. However, since the RRC has no frequency support outside $|f| \geq W$, there is no point in decreasing T below $T = 1/2W$. Inserting this symbol time into (5) yields a limiting capacity of

$$C = \int_0^{W} \log\left(1 + \frac{2|G(f)|^2}{N_0}\right) \mathrm{d}f. \tag{6}$$

To reach (5) and (6) the ML receiver must be used. For a non-flat $\bar{H}(f)$ and/or FTN signaling, the ML receiver is of significant complexity since the memory of an ISI channel (2) can be large. As a low-complexity remedy, we consider the CS receiver from [8]. The CS receiver forces the memory of the ISI channel into only the K most recent input symbols by means of linear filtering. Whenever $K = \infty$, the ML receiver is obtained, while the case $K = 0$ corresponds to the MMSE receiver. The achievable rates of CS, analyzed in [8], are easy to compute for a given ISI channel. With strategy (ii) the user assumes that $g(t) = p(t)$. The CS receiver can still be applied, but there will be a slight mismatch since in reality we have $g(t) \neq p(t)$. With strategy (ii) and *without* any FTN, the only choice for the CS parameter K that makes any sense is $K = 0$, as the pulse $p(t)$ does not induce any ISI itself.

We next describe the measured LSAS channels. The measurements were conducted at an outdoor 15×30 meter courtyard, located in the center of the EE-building at Lund University, Sweden. The BS is a linear array comprising 128 antenna elements spaced half a wavelength apart and was positioned at the roof of the EE-building with no line-of-sight to the courtyard. The total bandwidth of the measurements is 20 MHz at 2.1 GHz. A single antenna user was placed at 10 random positions in the courtyard, and a single snapshot of the channel vector to the user was recorded at each position.

In Figure 1 we show capacity results (5) without any FTN, i.e., $T = (1+\beta)/2W$. The excess bandwidth is chosen as 20%, i.e., $\beta = 0.2$. This choice resembles the accumulated losses of a cyclic prefix and spectral guardbands in the LTE system. The solid set of curves corresponds to $M = 128$ antennas at the BS, while the dashed set of curves is a random selection of 4 antennas and then averaging over the random antenna selections. Within each set, the curve marked with a circle is the achievable capacity of receiver strategy (i) in combination with an ML receiver, the curve marked

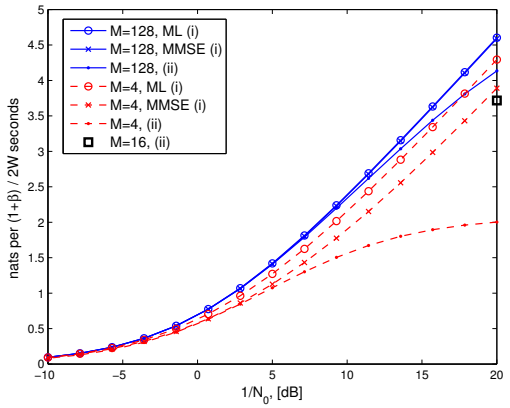

Fig. 1. Capacity results for non-FTN systems in measured LSAS channels with a RRC pulse having $\beta = 0.2$.

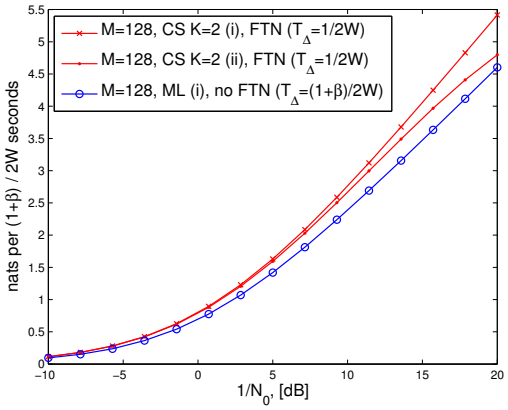

Fig. 2. Capacity results for FTN systems in measured LSAS channels with a RRC pulse having $\beta = 0.2$.

with x-es is for strategy (i) but with MMSE detection, and the curve marked with dots is the simple receiver strategy (ii). The heavy square at the rightmost part of the plot is the final point for $M = 16$ antennas with receiver strategy (ii). As we can see, with $M = 4$, there is a major loss of capacity if the receiver does not pay any attention to the effect of the channel's frequency selectivity. However, with $M = 128$, the frequency selectivity of the channel is far less pronounced so strategy (ii) is almost as good as strategy (i). Further, with (i), there is no loss of using an MMSE equalizer instead of an ML, while a 2 dB loss is observed for $M = 4$. From the given set of measurements we draw the conclusion that, with a large antenna array, a single carrier system works well. Indeed, the frequency selectivity of the channel averages out, and this enables the very simple receiver solution (ii).

In Figure 2 we focus on the $M = 128$ case but use FTN signaling to boost capacity. We set $T = 1/2W$ so that (5) collapses down into (6). The result for a CS receiver, $K = 2$ and strategy (i) is given in the top curve in the figure. By increasing the memory of the receiver K, we observed only very minor further gains, which implies that the CS receiver provides essentially the full ML performance, but at a much lower cost. By comparing this curve with that of the no-FTN case, we see a clear gain. Asymptotically in the SNR, the ratio of the FTN capacity to the non-FTN capacity will tend to $1+\beta = 1.2$. The middle curve marked with dots shows the result for the simpler strategy (ii). However, with FTN, this receiver must be based on a trellis due to the ISI induced by the FTN system itself, but it ignores the frequency selectivity of the channel. As we can see, even this receiver decisively outperforms the non-FTN system, although it will eventually saturate at high SNRs.

5. TIME AND FREQUENCY PACKING ON 3GPP CHANNELS

This section studies the application of time-frequency packing on doubly selective channels, typically encountered in mobile communications. The assumed channel model is time variant and has an impulse response of the form

$$h(t;\tau) = \sum_{i=0}^{L} c_i(t)\delta(\tau - \tau_i) \qquad (7)$$

where L is the number of channel taps and $\{\tau_i\}$ are the tap delays. The taps $c_i(t)$ have a Jakes power spectral density, which is completely described by the Doppler shift parameter f_D. The transmitted signal is modulated as filter bank multicarrier (FBMC, [15]) which has the general form

$$x(t) = \sum_{k,\ell} x_k^{(\ell)} p(t - k\delta_t T) e^{j2\pi\ell\delta_f \frac{t}{T}} \qquad (8)$$

where δ_t and δ_f are the time and frequency packing factors. Without frequency packing, FBMC has a frequency spacing of $1/T$, and orthogonality cannot be satisfied in both domains, even in absence of packing (i.e. $\delta_t = \delta_f = 1$) unless $p(t) = \sin(\pi t/T)/(\pi t/T)$. However since orthogonality is always corrupted by the channel, FBMC uses smoother pulses which introduce ISI and ICI. Thus, FBMC, similarly to time-frequency packing, introduce interference with the aim of increasing the spectral efficiency. Thus, using $\delta_t < 1$ and $\delta_f < 1$ is just a further degree of freedom in the system design.

Via over-sampling, the samples of the received signal can be gathered into a vector $\boldsymbol{y}$ and the channel is described through the matrix equation

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{P}\boldsymbol{x} + \boldsymbol{n}\,,$$

where $\boldsymbol{x}$ is a vector of transmitted symbols over all carriers and time indices, the matrix $\boldsymbol{C}$ gathers the samples of the fading coefficients $c_i(t)$, the matrix $\boldsymbol{P}$ models the shaping pulse, and the vector $\boldsymbol{n}$ is AWGN noise with variance N_0. We adopted a minimum mean square error (MMSE) equalizer possibly with interference cancellation (IC). Equalization is performed separately on each received FBMC symbol through the matrix

$$\boldsymbol{G}_{\mathrm{MMSE}} = \boldsymbol{\Lambda}^H(\boldsymbol{\Lambda}\boldsymbol{\Lambda}^H + N_0\boldsymbol{I} + \boldsymbol{\Lambda}_{\mathrm{I}}\boldsymbol{\Lambda}_{\mathrm{I}}^H)^{-1}\,,$$

where $\boldsymbol{\Lambda}$ gathers the rows of $\boldsymbol{CP}$ which refer to the FBMC symbol of interest and $\boldsymbol{\Lambda}_{\mathrm{I}}$ gathers the rows of interfering FBMC symbols. IC operations are as follow: Two layers of symbols $x_k^{(\ell)}$ are present, one layer contains all symbols with $k + \ell$ even, and the other all symbols with an odd value. At the first stage of the receiver process, the receiver decodes and

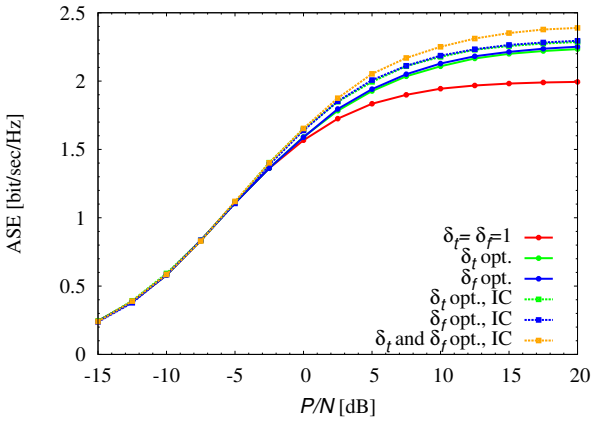

Fig. 3. ASE for 4QAM on ETU channel with f_D=300 Hz.

perfectly cancels the first layer. In the next step, the second layer is treated, but now the interference is far less as the first layer has been removed. This yields two capacities, one for the first layer, and one larger value for the second layer. We report the average of the two.

Figure 3 shows ASE curves for FBMC as a function of the ratio between signal power P and noise power N. The considered channel is (7) with tap delays and mean values described by 3GPP for extended typical urban (ETU) model and $f_D = 300$ Hz. The transmitted signal has a bandwidth of approximately 4.5 MHz, and is composed by 301 carriers with a reference frequency spacing $1/T = 15$ kHz. We use QPSK, and a RRC pulse with roll off $\beta = 0.1$ and bandwidth $W = 1.1/T = 16.5$ kHz. The figure shows that time and frequency packing gains w.r.t. $\delta_t = \delta_f = 1$ for any $P/N > -2.5$ dB, and further gains are achievable with IC.

However it must be noticed that the gains are close to the saturation of ASE curves, and they are not sufficient to outperform constellations with higher cardinality. Further gains could be obtained with more complex receivers but, on the other hand, we believe that is difficult to find substantial gains, since FBMC is already a sort of time-frequency packing. More research on this topic is required.

6. CONCLUSIONS

In this paper we have discussed possible application areas of FTN. We first surveyed satellite and optical communications, where FTN is already extensively discussed and is most promising. Then we went on towards areas where FTN is in infancy and we combined FTN with large scale antenna systems. By doing so, we have demonstrated that single carrier FTN systems with 128 BS antennas can use very simple receiver solutions, and reach high spectral efficiencies. We also considered FTN in doubly selective radio channels, and we observed spectral efficiencies, again for low-complexity receivers. Unfortunately, the reported gains are close to the saturation point of the capacity curves for finite constellations. This opens up for future work in the area.

7. ACKNOWLEDGEMENT

The authors would like to thank Fredrik Tufvesson for providing the channel measurements.

REFERENCES

[1] J. E. Mazo, "Faster-than-Nyquist signaling," *Bell System Tech. J.*, vol. 54, pp. 1450–1462, Oct. 1975.

[2] A. Liveris and C. N. Georghiades, "Exploiting faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 47, pp. 1502–1511, Sept. 2003.

[3] F. Rusek and J. B. Anderson, "The two dimensional Mazo limit," in *Proc. IEEE Intern. Symp. on Inform. Theory*, Adelaide, Australia, Nov. 2005, pp. 970–974.

[4] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006.

[5] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.

[6] A. Barbieri, D. Fertonani, and G. Colavolpe, "Time-frequency packing for linear modulations: spectral efficiency and practical detection schemes," *IEEE Trans. Commun.*, vol. 57, pp. 2951–2959, Oct. 2009.

[7] A. Modenini, G. Colavolpe, and N. Alagha, "How to significantly improve the spectral efficiency of linear modulations through time-frequency packing and advanced processing," in *Proc. IEEE Intern. Conf. Commun.*, Ottawa, Canada, June 2012, pp. 3299–3304.

[8] F. Rusek and A. Prlja, "Optimal channel shortening for MIMO and ISI channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 810–818, Feb. 2012.

[9] A. Piemontese, A. Modenini, G. Colavolpe, and N. Alagha, "Improving the spectral efficiency of nonlinear satellite systems through time-frequency packing and advanced processing," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3404–3412, Aug. 2013.

[10] G. Colavolpe, T. Foggi, A. Modenini, and A. Piemontese, "Faster-than-Nyquist and beyond: how to improve spectral efficiency by accepting interference," *Opt. Express*, vol. 19, no. 27, pp. 26600–26609, Dec 2011.

[11] G. Colavolpe and T. Foggi, "Next-generation long-haul optical links: higher spectral efficiency through time-frequency packing," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2013, vol. 1.

[12] G. Karam and H. Sari, "A data predistortion technique with memory for QAM radio systems," *IEEE Trans. Commun.*, vol. 39, no. 2, pp. 336–344, Feb. 1991.

[13] G. Colavolpe, A. Modenini, and F. Rusek, "Channel shortening for nonlinear satellite channels," *IEEE Commun. Letters*, vol. 16, pp. 1929–1932, Dec. 2012.

[14] F. Rusek, D. Persson, Buon Kiong Lau, E.G. Larsson, T.L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Mag.*, vol. 30, no. 1, pp. 40–60, 2013.

[15] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Processing Mag.*, vol. 28, no. 3, pp. 92–112, May 2011.