1-Bit Quantization and Oversampling at the Receiver: Communication over Bandlimited Channels with Noise

Lukas Landau, *Member, IEEE,* Meik Dörpinghaus, *Member, IEEE,* and Gerhard P. Fettweis, *Fellow, IEEE*

***Abstract*—A bandlimited additive white Gaussian noise channel is considered where the output is 1-bit quantized and oversampled with respect to the Nyquist rate. We consider root raised cosine filters at the transmitter and receiver. In particular we focus on a roll-off factor equal to 1 and 0. Because of the oversampling the channel has infinite memory. An auxiliary channel law is proposed which describes the resulting received sequences based on a truncated waveform. The random distortion due to the residual sidelobes can be considered as an additional noise term in the auxiliary channel law. The auxiliary channel law is utilized for computing a lower bound on the achievable rate and in a further step for optimizing a Markov source model. Different signaling schemes have been considered, such as BPSK and ASK. Moreover, Nyquist signaling and *faster-than-Nyquist* signaling are considered. The resulting achievable rates are superior as compared to results from the literature on bandlimited channels with noise, 1-bit quantization and oversampling at the receiver.**

***Index Terms*—Quantization, 1-bit, oversampling, analog-to-digital converter, faster-than-Nyquist signaling, achievable rates.**

I. Introduction

COARSE quantization at the receiver is attractive in terms of energy efficiency. This is a key issue for multigigabit/s communication, e.g., wireless board-to-board communication [1] and for low-power applications in the context of the internet-of-things. The present work considers 1-bit quantization where the receiver has only sign information about the received signal. Oversampling w.r.t. Nyquist rate is applied to compensate for the loss in achievable rate brought by the quantization. The first time a gain from oversampling was reported in [2], where for a noiseless channel the achievable rate is 1.07 bits per Nyquist interval for two-fold oversampling. In [3] it was shown for the same noiseless channel, that by considering Zakai bandlimited processes significantly higher rates, namely,

$$I = \log_2(M_{\text{osr}} + 1) \text{ [bits per Nyquist interval]}, \qquad (1)$$

and higher are achievable, where M_{osr} denotes the oversampling factor w.r.t. the Nyquist rate. A benefit of oversampling on the capacity per unit cost has been pointed out in [4]. In [5] a marginal benefit of oversampling has been found for high signal-to-noise ratio (SNR). In addition, there are many promising studies on *non-strict-bandlimited* channels. In the present work, achievable rates comparable to (1) are calculated, for channels which are bandlimited and even noisy. We study the achievable rate of channels with root raised cosine (RRC) filters with fixed Nyquist bandwidth. We utilize an auxiliary channel law based on a truncation of infinite long waveforms to compute lower bounds on the achievable rate of the actual channel. With a truncation, the auxiliary channel refers to a channel with finite memory. The additional distortion due to the residual sidelobes can be described as an additional noise term in the auxiliary channel law. The capacity of channels with finite-state intersymbol interference can be approached by considering input sequences modeled by Markov sources [6]. Hence, a Markov source optimization strategy is proposed, following the principle in [7].

Sequences are denoted as $x^n = [x_1, \ldots, x_n]^T$ and sequences of vectors as $\boldsymbol{y}^n = [\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_n^T]^T$. A segment of a sequence is written as $x_{k-L}^k = [x_{k-L}, \ldots, x_k]^T$ and $\boldsymbol{y}_{k-L}^k = [\boldsymbol{y}_{k-L}^T, \ldots, \boldsymbol{y}_k^T]^T$, respectively. Convolution is denoted by $*$.

II. System Model

The input sequences are modeled as the output of a Markov source where the transmit symbols X_k depends on L_{src} previous symbols $P\left(X_k|X^{k-1}\right) = P\left(X_k|X_{k-L_{\text{src}}}^{k-1}\right) = P\left(S_k|S_{k-1}\right)$, with the source state $S_k = X_{k-L_{\text{src}}+1}^k$. In the following, we use the simplified notation $P_{i,j} = P\left(S_k = j|S_{k-1} = i\right)$. The corresponding stationary distribution is $\mu_i = P\left(S_k = i\right)$. The transmit and receive filter $h(t)$ and $g(t)$ are both RRC, given by

$$h(t) = \begin{cases} \frac{1}{\sqrt{T}}(1 - \beta + 4\frac{\beta}{\pi}), & t = 0 \\ \frac{\beta}{\sqrt{2T}}\left[(1 + \frac{2}{\pi})\sin(\frac{\pi}{4\beta}) + (1 - \frac{2}{\pi})\cos(\frac{\pi}{4\beta})\right], & t = \pm\frac{T}{4\beta} \\ \frac{1}{\sqrt{T}}\frac{\sin(\pi\frac{t}{T}(1-\beta)) + 4\beta\frac{t}{T}\cos(\pi\frac{t}{T}(1+\beta))}{\pi\frac{t}{T}(1-(4\beta\frac{t}{T})^2)}, & \text{else.} \end{cases}$$

where $T = (1+\beta)T_{\text{s}}$ with the Nyquist interval T_{s}. $\beta = 0$ yields the sinc pulse. The signaling rate is given by $\frac{M_{\text{Tx}}}{T_{\text{s}}}$, where the case of $M_{\text{Tx}} > 1$ implies faster-than-Nyquist signaling (FTN) [8]. With this, the transmit signal is denoted as

$$x(t) = \sum_{l=-\infty}^{\infty} \sqrt{\frac{E_{\text{s}}}{T_{\text{s}}}} x_l h\left(t - l\frac{T_{\text{s}}}{M_{\text{Tx}}}\right),$$

where E_{s} is the transmit energy per transmit pulse. The average transmit energy per Nyquist interval $\overline{E_{\text{s}}} = \lim_{n\to\infty} \frac{1}{n} \int_0^{\frac{nT_{\text{s}}}{M_{\text{Tx}}}} |x(t)|^2 \, dt$ depends on the input

This work has been supported in part by the German Research Foundation (DFG) within the SFB 912 “Highly Adaptive Energy-Efficient Computing”.

L. Landau was with TU Dresden and is now with CETUC/PUC-Rio, 22453-900 Rio de Janeiro, Brazil, (e-mail: lukas.landau@cetuc.puc-rio.br).

M. Dörpinghaus and G. Fettweis are with the Vodafone Chair Mobile Communications Systems and SFB 912 HAEC, Technische Universität Dresden, Dresden 01062, Germany (e-mail: {meik.doerpinghaus, gerhard.fettweis}@tu-dresden.de).

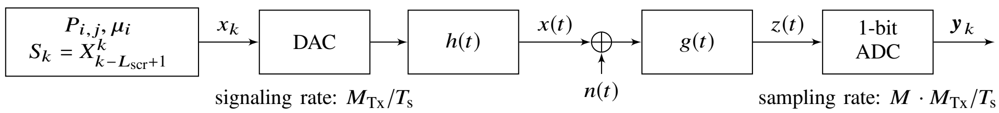

Fig. 1: System model, oversampling factor M and faster-than-Nyquist coefficient M_{Tx}

sequence design and is post computed. With $v(t) = h(t) * g(t)$ the output of the receive filter is given by

$$z(t) = \sum_{l=-\infty}^{\infty} \sqrt{E_{\mathrm{s}}} x_l v\left(t - l \frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right) + \eta(t),$$

with the filtered noise $\eta(t) = g(t) * n(t)$ based on the white Gaussian noise $n(t)$. The sampling rate at the receiver is $\frac{M_{\mathrm{osr}}}{T_{\mathrm{s}}}$ with oversampling factor $M_{\mathrm{osr}} = M \cdot M_{\mathrm{Tx}}$ w.r.t. the Nyquist rate. The received samples are given by

$$z\left(\left(k + \frac{m}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right) = \sum_{l=k-\infty}^{k+\infty} \sqrt{E_{\mathrm{s}}} x_l v\left(\left(k + \frac{m}{M} - l\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right) + \eta\left(\left(k + \frac{m}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right), \quad (2)$$

where $m \in \{0, \ldots, M-1\}$. In the following, we apply the more compact notation with indices $z_{k,m} = z\left(\left(k + \frac{m}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right)$. The received samples are represented in vectors $\boldsymbol{z}_k = [z_{k,0}, \ldots, z_{k,M-1}]^T$ of length M, where M is the oversampling factor w.r.t. a transmit symbol. Subsequently, the samples experience the quantization and its output is denoted by $\boldsymbol{y}_k = Q\{\boldsymbol{z}_k\}$. Its decision rule is $Q(z_{k,m} \geq 0) = 1$ and -1 otherwise. Thus the causal discrete time notation[1] is given by

$$\boldsymbol{y}_k = Q\{\boldsymbol{z}_k\} = Q\{\boldsymbol{V}\boldsymbol{U}x_{k-L_\infty}^{k} + \boldsymbol{G}\boldsymbol{n}_{k-L_\infty-1}^{k}\} \quad (3)$$
$$\boldsymbol{\eta}_k = \boldsymbol{G}\boldsymbol{n}_{k-L_\infty-1}^{k},$$

with the zero-inserting M-fold up-sampling matrix $\boldsymbol{U}$, filter matrices $\boldsymbol{V}$, $\boldsymbol{G}$ and the i.i.d. noise vector $\boldsymbol{n}_{k-L_\infty-1}^{k}$ with $M(L_\infty + 2)$ entries, each distributed with $\mathcal{N}(0, \sigma_n^2)$. $\boldsymbol{U}$ has dimension $(M(L_\infty+2)-1) \times (L_\infty+1)$. Its entries are given by

$$U_{i,j} = \begin{cases} 1 & \text{for } i = jM \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

L_∞ is chosen such that the error due to truncating the waveform causes a much smaller variance than the thermal noise. The filter matrices are structured by

$$\boldsymbol{V} = \begin{pmatrix} \left[\ \boldsymbol{v}^T\ \right] & 0 \cdots & 0 \\ 0 & \left[\ \boldsymbol{v}^T\ \right] & 0 \cdots 0 \\ & \ddots\ \ddots\ \ddots & \\ 0 \cdots & 0 & \left[\ \boldsymbol{v}^T\ \right] \end{pmatrix}, \quad \boldsymbol{G} = \frac{1}{\|\boldsymbol{g}\|_2} \begin{pmatrix} \left[\ \boldsymbol{g}^T\ \right] & 0 \cdots & 0 & 0 \\ 0 & \left[\ \boldsymbol{g}^T\ \right] & 0 \cdots 0 & 0 \\ & \ddots\ \ddots\ \ddots & & \\ 0 \cdots & 0 & \left[\ \boldsymbol{g}^T\ \right] & 0 \end{pmatrix},$$

with dimension $M \times (M(L_\infty+2)-1)$ and $M \times (M(L_\infty+2))$, respectively. The vectors $\boldsymbol{v}, \boldsymbol{g}$ of length $M(L_\infty+1)$ represent $v(t), g(t)$ with sampling rate $\frac{M \cdot M_{\mathrm{Tx}}}{T_{\mathrm{s}}}$ in the sense of

$$\boldsymbol{v} = \left[v\left(\frac{-(L_\infty+1)}{2}\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right), v\left(\left(\frac{-(L_\infty+1)}{2} + \frac{1}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right), \ldots, v\left(\left(\frac{(L_\infty+1)}{2} - \frac{1}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right)\right]^T. \quad (5)$$

[1]With respect to (2), in the matrix-vector notation in (3) the elements of the vectors $\boldsymbol{y}_k$ and $\boldsymbol{z}_k$ are shifted in time to simplify the notation.

III. The Auxiliary Channel Law

An auxiliary channel law $W(\cdot)$ w.r.t. the channel law $P(\cdot)$ can be utilized for computing a lower bound on the achievable rate [9], where the only constraint is that $W(\cdot) > 0$ whenever $P(\cdot) > 0$. In this section, we build a finite state channel law based on truncated waveforms with a length of $2c$ Nyquist intervals to provide an adequate statistical description of the channel (3). Rewriting a received sample yields, cf. (2)

$$\begin{aligned} z\left(\left(k + \frac{m}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right) = & \sum_{l=k-\zeta+1}^{k+\zeta} \sqrt{E_{\mathrm{s}}} x_l v\left(\left(k + \frac{m}{M} - l\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right) \\ & + \sum_{l=-\infty}^{k-\zeta} \sqrt{E_{\mathrm{s}}} x_l v\left(\left(k + \frac{m}{M} - l\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right) \\ & + \sum_{l=k+\zeta+1}^{\infty} \sqrt{E_{\mathrm{s}}} x_l v\left(\left(k + \frac{m}{M} - l\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right) \\ & + \eta\left(\left(k + \frac{m}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right), \end{aligned} \quad (6)$$

where $\zeta = cM_{\mathrm{Tx}}$. The variance of the 2nd term on the RHS of (6) is given by

$$\begin{aligned} \sigma_{\mathrm{lobe}}^2 &= \mathrm{Var}\left\{\sum_{l=-\infty}^{k-\zeta} \sqrt{E_{\mathrm{s}}} x_l v\left(\left(k + \frac{m}{M} - l\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right)\right\} \\ &= \mathrm{E}\left\{\left|\sum_{l=\zeta}^{\infty} \sqrt{E_{\mathrm{s}}} x_l v\left(\left(l + \frac{m}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right)\right|^2\right\}, \end{aligned} \quad (7)$$

for x_k being zero-mean. Explicitly for constructing an auxiliary channel law, we rely on the simplifying assumption that the x_k are independent[2]. With this, the sum is rewritten as

$$\sigma_{\mathrm{lobe}}^2 \approx \tilde{\sigma}_{\mathrm{lobe}}^2 = \sum_{l=\zeta}^{\infty} E_{\mathrm{s}} \mathrm{E}\left\{|x_l|^2\right\} \left|v\left(\left(l + \frac{m}{M}\right)\frac{T_{\mathrm{s}}}{M_{\mathrm{Tx}}}\right)\right|^2. \quad (8)$$

We consider three special cases where the variances of the 2nd and 3rd term on the RHS of (6) are equal or zero.

Case 1: $\beta = 0$, $M_{\mathrm{Tx}} = 1$, $M = 2$. For $m = 1$ the sum in (8) simplifies to

$$\begin{aligned} \tilde{\sigma}_{\mathrm{lobe}}^2|_{m=1} &= E_{\mathrm{s}} \mathrm{E}\left\{|x_l|^2\right\} \sum_{l=\zeta}^{\infty} \mathrm{sinc}^2\left(\pi\left(l + \frac{1}{2}\right)\right) \\ &= \frac{E_{\mathrm{s}} \mathrm{E}\left\{|x_l|^2\right\}}{\pi^2} \sum_{l=c}^{\infty} \left(l + \frac{1}{2}\right)^{-2}, \end{aligned} \quad (9)$$

where the convergence of the sum can be shown easily. For Nyquist signaling there is no intersymbol interference at the *perfect* sampling instances, which implies $\sigma_{\mathrm{lobe}}^2|_{m=0} = 0$.

[2]Note, this is not in contradiction to Markovian input sequences, as here the assumption on x_k is only made to calculate σ_{lobe}^2 of the auxiliary channel law.

Case 2: $\beta = 0$, $M_{\mathrm{Tx}} = 2$, $M = 1$ (FTN). With the considered FTN signaling scheme, the variance is independent of the time instance ($m = 0$ for every sample). It is given by

$$\tilde{\sigma}^2_{\mathrm{lobe}} = E_{\mathrm{s}} \mathrm{E}\left\{|x_l|^2\right\} \sum_{l=\zeta}^{\infty} \mathrm{sinc}^2\left(\frac{\pi}{2} l\right), \qquad (10)$$

which is equal to (9), as $\zeta = 2c$, and $\sin^2\left(\frac{\pi}{2} l\right) = 0$ for l even.

Case 3: $\beta = 1$, $c \geq 2$, M_{Tx}, M arbitrary. Here, we neglect residual sidelobes for the auxiliary channel law as $\tilde{\sigma}^2_{\mathrm{lobe}} \approx 0$.

The auxiliary channel law describes the outer sidelobes as additional Gaussian distributed distortion with the covariance matrix $\boldsymbol{R}_{\mathrm{lobes}}$. An effective noise covariance matrix is given by

$$\boldsymbol{R}_{\mathrm{eff}} = \mathrm{E}\left\{\boldsymbol{\eta}^k_{k-N}\left(\boldsymbol{\eta}^k_{k-N}\right)^T\right\} + \boldsymbol{R}_{\mathrm{lobes}}, \qquad (11)$$

where the description of the noise correlation is simplified by limiting it to $N + 1$ symbol durations. The auxiliary channel law $W(\cdot)$ provides an approximate statistical description of the channel output of the actual channel $P(\boldsymbol{y}_k|\boldsymbol{y}^{k-1}, x^n) \approx W(\boldsymbol{y}_k|\boldsymbol{y}^{k-1}, x^n)$. It takes into account N previous channel realizations and $L + N$ previous transmit symbols as described by

$$W(\boldsymbol{y}_k|\boldsymbol{y}^{k-1}, x^n) = W(\boldsymbol{y}_k|\boldsymbol{y}^{k-1}_{k-N}, x^k_{k-L-N}),$$

where $L \ll L_\infty$. Applying Bayes' rule yields

$$W(\boldsymbol{y}_k|\boldsymbol{y}^{k-1}_{k-N}, x^k_{k-L-N}) = \frac{W(\boldsymbol{y}^k_{k-N}|x^k_{k-L-N})}{W(\boldsymbol{y}^{k-1}_{k-N}|x^{k-1}_{k-L-N})}.$$

With (11) the auxiliary probability density function of the received signal is given by

$$w(z^k_{k-N}|x^k_{k-L-N}) = \left((2\pi)^{M(N+1)}\left|\boldsymbol{R}_{\mathrm{eff}}\right|\right)^{-\frac{1}{2}} \times \exp\left(-\frac{1}{2}(z^k_{k-N} - \boldsymbol{\mu}_x)^T \boldsymbol{R}^{-1}_{\mathrm{eff}}(z^k_{k-N} - \boldsymbol{\mu}_x)\right), \qquad (12)$$

where $\boldsymbol{\mu}_x = \boldsymbol{V}'(N)\boldsymbol{U}(N)x^k_{k-L-N}$ and $|\cdot|$ denotes the determinant. The entries of the upsampling matrix $\boldsymbol{U}(N)$ with dimension $((L + N + 2)M - 1) \times (L + N + 1)$ are given by (4). The filter matrix $\boldsymbol{V}'(N)$ is constructed similarly to $\boldsymbol{V}$, but it is based on the vector $\boldsymbol{v}'$, which belongs to the truncated waveform with $M(L + 1)$ coefficients. The vector $\boldsymbol{v}'$ is constructed according to (5) where L_∞ is replaced by L. The transition probabilities of the auxiliary channel are obtained by multivariate Gaussian integration over the quantization region $\mathbb{Y}^k_{k-N}$, which belong to $\boldsymbol{y}^k_{k-N}$ and are described by

$$W(\boldsymbol{y}^k_{k-N}|x^k_{k-L-N}) = \int_{z^k_{k-N} \in \mathbb{Y}^k_{k-N}} w(z^k_{k-N}|x^k_{k-L-N}) dz^k_{k-N}.$$

The case sensitive effective noise covariance matrix $\boldsymbol{R}_{\mathrm{eff}}$ is given for the three examples in the following.

Case 1: $\beta = 0$, $M_{\mathrm{Tx}} = 1$, $M = 2$, $N = 0$. When Nyquist signaling is considered only the *oversampling* samples are affected by the additional distortion

$$\boldsymbol{R}_{\mathrm{eff}} = \sigma^2_n \cdot \begin{bmatrix} 1 & 0.637 \\ 0.637 & 1 \end{bmatrix} + \frac{E_{\mathrm{s}} \mathrm{E}\left\{|x_l|^2\right\}}{\pi^2} \sum_{l=c}^{\infty} \left(l + \frac{1}{2}\right)^{-2} \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix},$$

where the 2$^{\mathrm{nd}}$ summand refers to (9) with factor two.

Case 2: $\beta = 0$, $M_{\mathrm{Tx}} = 2$ $M = 1$ $N = 1$ (FTN). For FTN signaling, the additional distortion affects each sample

$$\boldsymbol{R}_{\mathrm{eff}} = \sigma^2_n \cdot \begin{bmatrix} 1 & 0.637 \\ 0.637 & 1 \end{bmatrix} + \frac{E_{\mathrm{s}} \mathrm{E}\left\{|x_l|^2\right\}}{\pi^2} \sum_{l=c}^{\infty} \left(l + \frac{1}{2}\right)^{-2} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

where the 2$^{\mathrm{nd}}$ summand also refers to (9) with factor two.

Case 3: $\beta = 1$, $c \geq 2$, M_{Tx}, M arbitrary. Only thermal noise is considered with $\boldsymbol{R}_{\mathrm{eff}} = \mathrm{E}\left\{\boldsymbol{\eta}^k_{k-N}\left(\boldsymbol{\eta}^k_{k-N}\right)^T\right\}$.

IV. Markov Source Optimization

Using the auxiliary channel law $W(\cdot)$ the information rate of the channel in Fig. 1 is lower-bounded by

$$\lim_{n\to\infty} \frac{1}{n} I(X^n; \boldsymbol{Y}^n) \geq \frac{1}{n}(-\log_2 P(x^n) + \log_2 W(x^n|\boldsymbol{y}^n)), \qquad (13)$$

where $\boldsymbol{y}^n$ is a realization of a long sequence based on the true channel model (3). The auxiliary channel law corresponds to a channel with finite state memory, noise and an additional random distortion. For such finite state memory channels, Markovian input distributions are asymptotically capacity-achieving [6]. Thus we maximize the lower bound on the RHS of (13) by optimizing the transition probabilities of a Markov source model via the algorithm in [10], which follows the principle idea of the Blahut Arimoto algorithm in [7]. The validity of the lower bound can be shown by following the steps in [9] for the *reverse* formulation the mutual information instead. Rearranging (13) yields

$$\lim_{n\to\infty} \frac{1}{n} I(X^n; \boldsymbol{Y}^n) \geq \sum_{i,j} \mu_i P_{i,j} \left(\log_2\left(\frac{1}{P_{i,j}}\right) + \hat{T}_{i,j}\right), \qquad (14)$$

with coefficients

$$\hat{T}_{i,j} = \frac{\sum_{k\left|\substack{s_{k-1}=i \\ s_k=j}\right.} \log_2 W\left(s_k, s_{k-1}|\boldsymbol{y}^n\right)}{\sum_{k\left|\substack{s_{k-1}=i \\ s_k=j}\right.} 1} - \frac{\sum_{k-1|s_{k-1}=i} \log_2 W\left(s_{k-1}|\boldsymbol{y}^n\right)}{\sum_{k-1|s_{k-1}=i} 1},$$

where $W\left(s_k, s_{k-1}|\boldsymbol{y}^n\right)$ and $W\left(s_{k-1}|\boldsymbol{y}^n\right)$ are being computed with the BCJR algorithm [11], which relies on the auxiliary channel law $W(\cdot)$. The form of (14) allows for iterative optimization of the transition probabilities $P_{i,j}$ and with this μ_i. In this regard, the optimized $P_{i,j}$ are given by

$$P_{i,j} = \begin{cases} \frac{b_j}{b_i}\frac{A_{i,j}}{\lambda}, & \text{if the transition occurs in } x^n \\ 0, & \text{else,} \end{cases}$$

where $A_{i,j}$ are entries of the noisy adjacency matrix $\boldsymbol{A}$, computed with $A_{i,j} = 2^{\hat{T}_{i,j}}$. The coefficients b_j, b_i are the ith and jth entries of the eigenvector of $\boldsymbol{A}$, respectively, and λ is the largest real valued eigenvalue of $\boldsymbol{A}$. The optimized values $P_{i,j}$ are utilized for the generation of a realization of x^n, $\boldsymbol{y}^n$ and computation of $\hat{T}_{i,j}$ in the subsequent iteration. In general, to approach those rates in a practical manner, a coding scheme is required, where the information is mapped to codewords of fixed or variable length.

V. Numerical Results

The actual channel is simulated with waveforms (5) spanning over 1600 Nyquist intervals ($L_\infty = 1600M_{\text{Tx}} - 1$). With this, the modeling error is approximated with (8) as $2\tilde{\sigma}^2_{\text{lobe}}|_{\zeta=\frac{L_\infty+1}{2}}/(E_{\text{s}}\mathrm{E}\{|x_l|^2\}) \approx \frac{1}{4}\cdot 10^{-3}$ for $\beta = 0$ and even lower for $\beta = 1$. The lower bound in (13) and (14) is computed based on sequences of length $n = 10^5$. The optimization algorithm has been applied for 19 iterations with $L_{\text{src}} = L+N$. With the bandwidth $B = \frac{1}{T_{\text{s}}}$ the SNR is given by $\overline{E}_{\text{s}}/\frac{T_{\text{s}}}{M_{\text{Tx}}}(BN_0)^{-1}$, where $N_0 = \sigma_n^2$ applies due to the receive filter with unit energy. Also a runlength limited sequence (RLL) with an entropy rate of 0.6942 bit per symbol has been considered for FTN as suggested in [10]. Table I shows the signaling rates and input alphabets of x_k. The lower bound on the achievable rate in terms of bits per Nyquist interval is given in Fig. 2. Both, an increased input alphabet and FTN signaling can be exploited for sequence design. In most cases a higher maximum input entropy per Nyquist interval results in a higher achievable rate. Some configurations imply a relatively high number of samples close to the decision threshold, which reduces the degrees of freedom in sequence design. In this regard, for $\beta = 0$ the BPSK alphabet with FTN is superior to the 4-ASK alphabet, although both have the same maximum input entropy of 2 bits per Nyquist interval. The wide mainlobe for $\beta = 1$ hampers the design of zero-crossings, which however can be compensated by an input sequence with enough degrees of freedom. E.g., a BPSK alphabet with FTN provides a relative small number of distinguishable sequences with the RRC with $\beta = 1$. At high SNR a larger β can yield an increased achievable rate because of lower sidelobe distortion. Fig. 3 illustrates the maximum achievable rate together with results from literature. The proposed approach is superior to prior work [5] on communication over noisy and bandlimited channels with 1-bit quantization and oversampling at the receiver in terms of achievable rate. Our results are comparable with analytical results for a noiseless channel [3] and the achievable rate using a Nyquist rate *flash-ADC* with the same number of comparator operations per time.

TABLE I: Overview on Parameters for Computation

	c	M_{Tx}	M	L	N	$\frac{2\tilde{\sigma}^2_{\text{lobe}}}{E_{\text{s}}\mathrm{E}\{\lvert x_l\rvert^2\}}\Big\rvert_{\beta=0}$
⬠ optimized 8-ASK	2	1	2	3	0	$0 (m=0)$ $0.0994 (m=1)$
□ optimized 4-ASK	3	1	2	5	0	$0 (m=0)$ $0.0669 (m=1)$
△ optimized 3-ASK	2	2	1	6	0	0.0994
○ optimized BPSK	2	2	1	6	1	0.0994
◇ BPSK (RLL)	2	2	1	6	1	0.0994

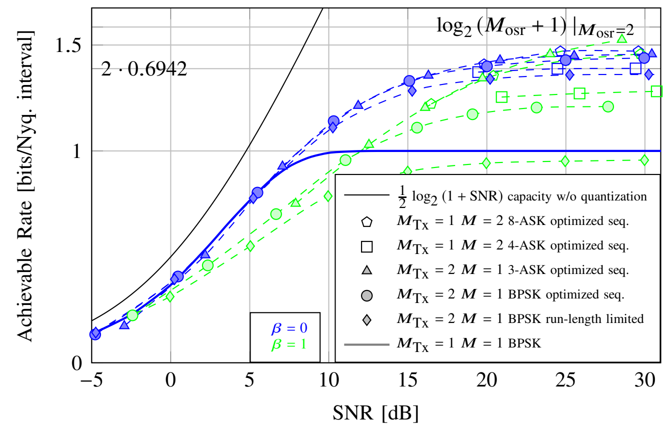

Fig. 2: Achievable rate versus SNR

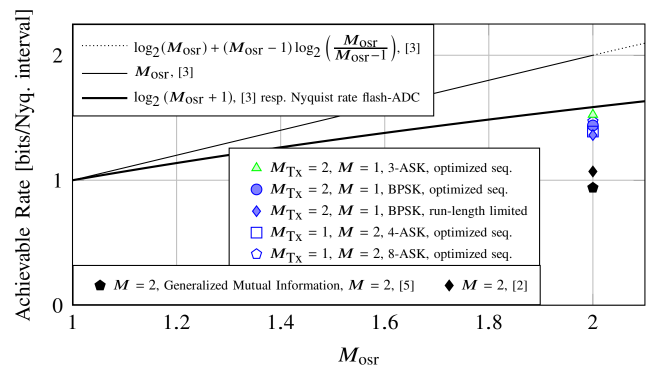

Fig. 3: Spectral efficiency versus oversampling factor

VI. Conclusions

We propose sequence based communication for noisy and bandlimited channels with 1-bit quantization and oversampling at the receiver. The considered RRC filters yield channels with infinite memory. An auxiliary channel law is introduced to compute a lower bound on the achievable rate of the actual channel. It relies on a truncation of the filters. The auxiliary channels have finite memory which is exploited for sequence design. Different symbol alphabets and signaling rates have been examined. The proposed approach is superior in terms of the achievable rate in comparison to the literature.

References

[1] G. P. Fettweis, N. ul Hassan, L. Landau, and E. Fischer, "Wireless interconnect for board and chip level," in *Proc. of the Conference on Design, Automation & Test in Europe*, 2013, (invited).

[2] E. N. Gilbert, "Increased information rate by oversampling," *IEEE Trans. Inf. Theory*, vol. 39, no. 6, pp. 1973–1976, 1993.

[3] S. Shamai, "Information rates by oversampling the sign of a bandlimited process," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1230–1236, 1994.

[4] T. Koch and A. Lapidoth, "Increased capacity per unit-cost by oversampling," in *Proc. of the IEEE Convention of Electrical and Electronics Engineers in Israel*, Eilat, Israel, Nov. 2010.

[5] W. Zhang, "A general framework for transmission with transceiver distortion and some applications," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 384–399, Feb. 2012.

[6] J. Chen and P. H. Siegel, "Markov processes asymptotically achieve the capacity of finite-state intersymbol interference channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1295–1303, 2008.

[7] A. Kavcic, "On the capacity of Markov sources over noisy channels," in *Proc. IEEE Glob. Comm. Conf. (GLOBECOM)*, vol. 5, San Antonio, Texas, USA, Nov. 2001, pp. 2997–3001.

[8] J. E. Mazo, "Faster-than-Nyquist signaling," *Bell System Technical Journal*, vol. 54, no. 1, pp. 1451–1462, 1975.

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

[10] L. Landau, M. Dörpinghaus, and G. P. Fettweis, "Communications employing 1-bit quantization and oversampling at the receiver: Faster-than-Nyquist signaling and sequence design," in *Proc. of the IEEE Int. Conf. on Ubiquitous Wireless Broadband*, Montreal, Canada, Oct. 2015.

[11] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.