SHANNON MEETS NYQUIST: CAPACITY LIMITS OF SAMPLED ANALOG CHANNELS

Yuxin Chen[*], *Yonina C. Eldar*[*†], *Andrea J. Goldsmith*[*]

[*]Department of Electrical Engineering, Stanford University.
[†]Department of Electrical Engineering, Technion, Israel Institute of Technology.
Email: yxchen@stanford.edu, yonina@ee.technion.ac.il, andrea@ee.stanford.edu

ABSTRACT

We explore several fundamental questions at the intersection of sampling theory and information theory. In particular, we study how capacity is affected by a given sampling mechanism below the channel's Nyquist rate, and what sampling strategy should be employed to maximize capacity. Two classes of sampling mechanisms are investigated: uniform sampling with filtering and uniform sampling with a filter bank. Optimal filters that maximize capacity are identified for both cases. We also highlight connections between capacity and minimum mean squared error (MMSE) estimation from sampled data. Our results indicate that maximizing capacity of sampled analog channels is a joint optimization problem over both the transmission strategy and the sampling technique.

Index Terms— Sampling rate, channel capacity, sampled analog channels, sub-Nyquist sampling, Landau rate.

1. INTRODUCTION

Capacity of waveform channels has been extensively studied in the continuous domain [1]. Essentially all communication receivers today convert their analog received signals into a digital sequence through some form of sampling. From Nyquist's theorem, by sampling at a rate equal to twice the signal's bandwidth over positive frequencies, the information content of the signal is preserved. Capacity analysis of analog channels thus assumes super-Nyquist sampling (e.g. [2]) without accounting for hardware limitations that may preclude sampling at this rate.

The Nyquist sampling rate is the rate required for perfect reconstruction of bandlimited analog signals or, more generally, the class of signals lying in shift-invariant subspaces [3]. This holds for both uniform and nonuniform sampling. However, the Nyquist rate may be excessive for signals that possess further structure. Examples of such structured signals include multiband signals, whose spectral content resides continuously within several subbands over a wide spectrum. For this class of signals, the sampling rate requirement for perfect recovery is the sum of the subband bandwidths (including both positive and negative frequencies), termed the *Landau rate* [4].

Uniform sampling with a filter bank is an effective sampling mechanism for multiband signals [5]. The basic paradigm of this technique, which includes recurrent non-uniform sampling as a special case, is to apply a bank of prefilters to the analog signal, each followed by a uniform sampler. Most sampling theoretic work aims at finding optimal sampling and reconstruction mechanisms that achieve (a) perfect reconstruction of a class of analog signals from *noiseless* samples, or (b) minimum reconstruction error from *noisy* samples based on certain statistical measures (e.g. minimum mean squared error (MMSE)). Here we consider optimal sampling techniques based on an information theoretic metric, namely, the channel capacity that can be achieved through noisy samples of the channel output. Guo *et. al.* [6] explored the connection between mutual information and MMSE for Gaussian channels, but they did not account for undersampling of analog signals.

In this work, we investigate how to optimize sampling to maximize capacity. This turns out to be a joint optimization problem over both the channel input and the sampling technique. We derive optimal structures for both uniform sampling with a single filter and uniform sampling with a filter bank. As we show, the optimal sampling structures in both cases also minimize the MSE between the original and the reconstructed signals, provided that the input is a wide-sense stationary stochastic signal. This illuminates a connection between channel capacity and MMSE under multiple forms of sub-Nyquist rate sampling.

The remainder of the paper is organized as follows. We derive the optimal filter for uniform sampling with a filter in Section 2. The optimal filters under sampling with filter banks is derived in Section 3. Numerical results illustrating the impact of these different techniques on capacity are also provided.

2. UNIFORM SAMPLING WITH FILTERING

Ideal uniform sampling is one of the simplest sampling mechanisms, which is performed by sampling the analog signal uniformly at a rate $f_s = T_s^{-1}$. The resulting samples are given by

$$y_s[n] \stackrel{\Delta}{=} y_c(nT_s). \tag{1}$$

In order to avoid aliasing and suppress out-of-band noise, a prefilter is often added prior to the ideal uniform sampler [3]. This prefilter may also model linear distortion in practical sampling devices. We therefore model our sampling process to include a general analog prefilter, as illustrated in Fig. 1(a).

Suppose that the transmit signal $x(t)$ is time constrained to the interval $(0, T]$. The channel is modeled as a linear time-invariant (LTI) filter with impulse response $h(t)$ and frequency response $H(f)$. Before sampling we prefilter the received signal with an LTI filter that has impulse response $s(t)$ and frequency response $S(f)$. Suppose that $h(t)$ and $s(t)$ are both bounded and continuous. The filtered output is observed over $(0, T]$ and can be written as

$$y_c(t) = \begin{cases} s(t) * \left(h(t) * x(t) + \eta(t)\right), & \text{if } t \in (0, T]; \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

This work was supported in part by the NSF Center for Science of Information, the Interconnect Focus Center of the Semiconductor Research Corporation, and BSF Transformative Science Grant 2010505.

We then sample $y_c(t)$ using an ideal uniform sampler, which leads to the sampled sequence $\{y_s[n]\}_{(0,T]}$, where we explicitly indicate the interval $(0,T)$ over which the samples are taken.

We consider the problem of finding the capacity $C(f_s)$ of the sampled channel at rate f_s:

$$C(f_s) = \lim_{T\to\infty} \frac{1}{T} \max_{p(x)} I\left(x[0,T]; \{y_s[n]\}_{(0,T]}\right).$$

Here, $x[0,T]$ denotes all transmit signals over time $[0,T]$, and the maximum is taken over all input distributions $p(x)$ subject to a power constraint $\frac{1}{T}\mathbb{E}(\int_0^T |x(\tau)|^2 \,\mathrm{d}\tau) \leq P$. We will use the terminology *Nyquist-rate channel capacity* for the analog channel capacity, which is commensurate with sampling at or above the Nyquist rate of the received signal after optimized pre-filtering.

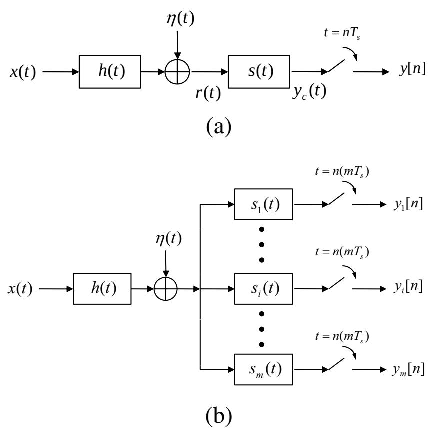

Fig. 1. (a) uniform sampling with a single filter; (b) uniform sampling with a filter bank.

The sampled capacity is given in the following theorem. Due to aliasing, the sampled capacity turns out to be a folded version of the non-sampled capacity.

Theorem 1 *Consider the setup in Fig. 1(a) for a given filter $s(t)$. Assume that $\eta(t)$ has power spectral density $\mathcal{N}(f)$. The capacity $C(f_s)$ of the sampled channel with a power constraint P is*

$$C(f_s) = \int_{f\in\mathcal{F}(W)} \frac{\log\left(W\frac{\sum_{l=-\infty}^{+\infty}|H(f-lf_s)S(f-lf_s)|^2}{\sum_{l=-\infty}^{+\infty}\mathcal{N}(f-lf_s)|S(f-lf_s)|^2}\right)}{2} df, \quad (3)$$

where $\mathcal{F}(W)$ and W are chosen based on a water-filling power allocation strategy.

The capacity expression of Theorem 1 is somewhat similar to the non-sampled waveform channel capacity derived in [1, Theorem 8.5.1]. The difference is that in (3) the integral is now over the sampling bandwidth, and the corresponding SNR at each $f \in \left[-\frac{f_s}{2}, \frac{f_s}{2}\right]$ is replaced by a folded version of the original SNR. This reflects the aliasing effect due to reduced-rate sampling.

Different prefilters lead to different channel capacities. The filter maximizing the capacity is given by the following theorem.

Theorem 2 *Consider the setup in Fig. 1(a). The capacity in (3) is maximized by the filter with frequency response*

$$S(f-kf_s) = \begin{cases} 1, & \text{if } \frac{|H(f-kf_s)|^2}{\mathcal{N}(f-kf_s)} = \sup_{l\in\mathbb{Z}} \frac{|H(f-lf_s)|^2}{\mathcal{N}(f-lf_s)}; \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

for all $f \in [-f_s/2, f_s/2]$ and all $k \in \mathbb{Z}$.

It can be verified that the SNR at each $f \in \left[-\frac{f_s}{2}, \frac{f_s}{2}\right]$ is a convex combination of $\left\{\frac{|H(f-lf_s)|^2}{\mathcal{N}(f-lf_s)}, l\in\mathbb{Z}\right\}$, where the weights are determined by the prefilters. That said, the prefilter assigns different weights to those frequencies that are equivalent modulo f_s. The optimal prefilter puts all its mass in those frequency components with the highest SNR.

In general, the frequency response of the optimal prefilter is discontinuous, which may be hard to realize in practice. However, for certain classes of channel models, the prefilter has a smooth frequency response. One example is a *monotone channel*, whose channel response $|H(f)|$ is non-increasing in f. Theorem 2 implies that the optimizing prefilter for a monotone channel is a *low-pass* filter with cutoff frequency f_s.

Since we have full control over both $X(f)$ and $S(f)$, the problem of finding optimal prefilters can be treated as a *joint* optimization over all input and filter responses. This joint optimization problem yields a solution different from the matched filter (i.e. the one that sets $S(f) = H^*(f)$, which is optimal when $X(f)$ is fixed). The optimal solution is to perform selection combining by setting some $S(f-lf_s)$ to one, as well as noise suppression by setting all other $S(f-lf_s)$ to zero. The comparison of this optimal filter and the matched filter and their corresponding capacity is illustrated in Fig. 2.

After passing through an optimal prefilter, all frequency components that do not possess the highest SNR are removed. Unless there exist multiple branches that possess the highest SNR, the optimal uniform sampler would indeed *suppress aliasing* as well as noise.

The prefilter (4) also turns out to minimize the MSE between the original and the reconstructed signals. In particular, suppose that the input is a wide-sense stationary signal with a given power spectral density (PSD) $\mathcal{S}_X(f)$, which satisfies the power constraint $\int_{-\infty}^{\infty}\mathcal{S}_X(f) \leq P$. Sampling theory allows us to find the MMSE between the input signal and the reconstructed signals over all reconstruction schemes for a given input. The joint optimization problem of interest can be posed as follows: how to choose the input PSD and the prefilter in order to minimize the MSE? We have shown that this problem can be reduced to ([7])

$$\underset{\{S(f-lf_s),l\in\mathbb{Z}\},\mathcal{S}_X}{\text{maximize}} \quad \frac{\sum_{l\in\mathbb{Z}}|H_l|^2\mathcal{S}_{X,l}|S_l|^2}{\sum_{l\in\mathbb{Z}}\left\{\left(|H_l|^2\mathcal{S}_{X,l}+\mathcal{N}_l\right)|S_l|^2\right\}} \quad (5)$$

subject to a power constraint $\int_{-\infty}^{\infty}\mathcal{S}_X(f) \leq P$, where $H_l := H(f-lf_s)$, $S_l := S(f-lf_s)$, $\mathcal{N}_l = \mathcal{N}(f-lf_s)$ and $\mathcal{S}_{X,l} := \mathcal{S}_X(f-lf_s)$. The objective function at frequency $f\in\left[-\frac{f_s}{2}, \frac{f_s}{2}\right]$ is a convex combination of $\left\{\frac{|H_l|^2\mathcal{S}_{X,l}}{|H_l|^2\mathcal{S}_{X,l}+\mathcal{N}_l} : l\in\mathbb{Z}\right\}$. Hence, for a given input PSD, the optimizing prefilter satisfies

$$S(f-kf_s) = \begin{cases} 1, & \text{if } \frac{|H_k|^2\mathcal{S}_{X,k}}{\mathcal{N}_k} = \max_{l\in\mathbb{Z}} \frac{|H_l|^2\mathcal{S}_{X,l}}{\mathcal{N}_l}; \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Allocating all input power to the frequency components with the highest $\frac{|H_k|^2}{\mathcal{N}_k}$ yields the optimal value among all input PSDs, which results in an optimal filter as given by (4).

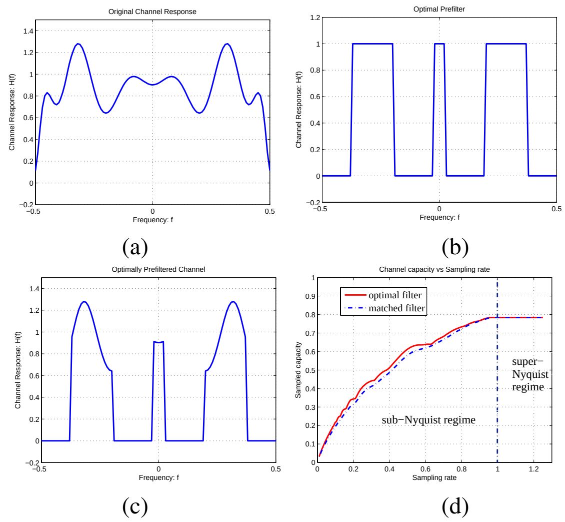

Fig. 2. The sampled capacity for a given channel with channel bandwidth 1. (a) frequency response of the original channel; (b) optimal prefilter associated with this channel with sampling rate 0.4; (c) optimally prefiltered channel response with sampling rate 0.4; (d) capacity vs sampling rate for the optimal prefilter and for the matched filter. The optimal prefilter has support size f_s in the frequency domain, hence its output allows alias-free sampling. In the sub-Nyquist regime, this aliasing-suppressing filter leads to higher capacity than the matched filter.

We note that channel capacity results include maximizing mutual information for a given system, while sampling theory considers optimal prefiltering and reconstruction schemes for a given class of input signals. The perspectives from both theories coincide via the filter design: the filter that maximizes mutual information for a given input distribution also minimizes the MSE for that input distribution.

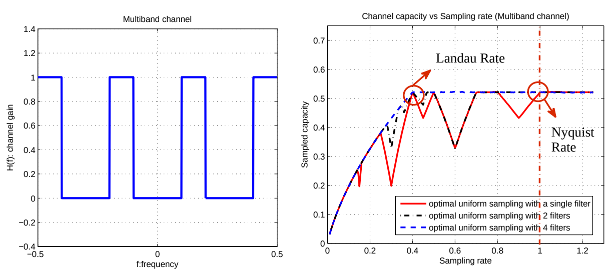

Fig. 3. Channel capacity vs sampling rate for (1) uniform sampling with a single filter; (2) uniform sampling with 2 filters; (3) uniform sampling with 4 filters. The channel bandwidth is assumed to be $\left[-\frac{1}{2}, \frac{1}{2}\right]$, the power constraint $P = 10$, and the noise power $\sigma_\eta^2 = 1$. The frequency response is given as $H(f) = 1$ for $|f| \in \left[\frac{1}{10}, \frac{1}{5}\right] \bigcup \left[\frac{2}{5}, \frac{1}{2}\right]$ and $H(f) = 0$ otherwise.

When the channel gain is not monotonic in f, the channel capacity $C(f_s)$ may not monotonically increase with the sampling rate f_s under uniform sampling with filtering. As an example, consider the multiband channel as illustrated in Fig. 3. If the channel is sampled at a rate $f_s = \frac{3}{5} f_{\rm NYQ}$ (above the Landau rate), aliasing occurs and leads to only one active interval (and hence one degree of freedom). However, if $f_s = \frac{2}{5} f_{\rm NYQ}$ (the Landau rate), it can be easily verified that two shifted subbands remain non-overlapping, resulting in two degrees of freedom. The tradeoff curve between capacity and sampling rate with optimal prefilters is plotted in Fig. 3. The underlying reason for the non-monotonic behavior of the capacity is that uniform sampling largely constrains our ability to exploit channel and signal structures.

3. UNIFORM SAMPLING WITH A FILTER BANK

As shown in the previous section, oversampling above the Landau rate with a known multiband channel may not achieve Nyquist-rate capacity if it does not favor the channel structure. Specifically, when operating upon signals with specific structure, uniform sampling with a filter may suppress information by collapsing degrees of freedom in the channel. This indicates that optimal sampling to maximize capacity may require more general non-uniform sampling methods [5] [8].

In order to incorporate non-uniform sampling, we replace the sampler block in Fig. 1(a) by a filter bank followed by uniform sampling at rate f_s/m for each branch, as illustrated in Fig. 1(b). Denote by $s_i(t)$ and $S_{i,c}(f)$ the impulse response and frequency response of the ith filter, respectively. The filtered output in the ith branch is

$$y_{c,i}(t) = \begin{cases} s_i(t) * (h(t)x(t) + \eta(t)) & \text{if } t \in (0, T]; \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

The filtered signals are passed through m ideal samplers to yield

$$y_i[n] \triangleq y_{c,i}(nmT_s) \quad \text{and} \quad \mathbf{y}_s[n] \triangleq [y_1[n], \cdots, y_m[n]]^T. \quad (8)$$

The capacity can then be expressed as

$$C(f_s) = \lim_{T\to\infty} \frac{1}{T} \max_{p(x):\frac{1}{T}\mathbb{E}\left(\int_0^T |x(\tau)|^2 \mathrm{d}\tau\right) \le P} I\left(x[0,T]; \{\mathbf{y}_s[n]\}\right),$$

and is given by the following theorem.

Theorem 3 *Consider the setup in Fig. 1(b). The sampled capacity $C(f_s)$ is*

$$C(f_s) = \max_{Q} \int_{-\frac{f_s}{2m}}^{\frac{f_s}{2m}} \frac{1}{2} \log\det\left(\mathbf{I} + \mathbf{F}_s^{\ddagger}\mathbf{F}_h\mathbf{Q}\mathbf{F}_h^{*}\mathbf{F}_s^{\ddagger *}\right) \mathrm{d}f$$

where $\mathbf{F}_s^{\ddagger} \triangleq (\mathbf{F}_s\mathbf{F}_s^*)^{-\frac{1}{2}}\mathbf{F}_s$ and Q is the power allocation matrix subject to a power constraint P. In addition, $\mathbf{F}_s$ is an infinite matrix of m rows and infinitely many columns and $\mathbf{F}_h$ is a diagonal infinite matrix such that

$$\begin{cases} (\mathbf{F}_s(f))_{i,l} & = S_i\left(f - \frac{lf_s}{m}\right) \\ (\mathbf{F}_h(f))_{l,l} & = H\left(f - \frac{lf_s}{m}\right) \end{cases} \quad (\forall 1 \le i \le k, \forall l \in \mathbb{Z}) \quad (9)$$

The optimal $\{Q(f)\}$ corresponds to the water-filling power allocation strategy based on the singular values of the equivalent channel matrix $\mathbf{F}_s^{\ddagger}\mathbf{F}_h$, where $\mathbf{F}_h$ is associated with the original channel and $\mathbf{F}_s^{\ddagger}$ is due to prefiltering and noise whitening.

In general, $\log\det\left(\mathbf{I}_m + \mathbf{F}_s^{\ddagger}\mathbf{F}_h\mathbf{Q}\mathbf{F}_h^{*}\mathbf{F}_s^{\ddagger *}\right)$ is not perfectly determined by $\mathbf{F}_s^{\ddagger}(f)$ and $\mathbf{F}_h(f)$ at a single frequency f, since the optimal power allocation strategy relies on the power constraint P/σ_η^2 as well as $\mathbf{F}_s$ and $\mathbf{F}_h$ across all f. In other words, $\log\det\left(\mathbf{I}_m + \mathbf{F}_s^{\ddagger}\mathbf{F}_h\mathbf{Q}\mathbf{F}_h^{*}\mathbf{F}_s^{\ddagger *}\right)$ is a function of all singular values of $\mathbf{F}_s^{\ddagger}\mathbf{F}_h$ and the universal water-level associated with optimal power allocation. Given two sets of singular values, we cannot determine which set is preferable without accounting for the water-level, unless one set is element-wise larger than the other. That said, if there exists a prefilter that maximizes all singular values simultaneously, then this prefilter will be universally optimal regardless of the water-level. Encouragingly, such optimal schemes exist, as we characterize in the following theorem.

Theorem 4 *The maximum capacity using uniform sampling with a filter bank can be achieved through a bank of filters whose frequency response in the k^{th} branch is given by*

$$S_k\left(f-\frac{lf_s}{m}\right)=\begin{cases}1, & \text{if } \left|H\left(f-\frac{lf_s}{m}\right)\right|^2=\left(\mathbf{F}_h(f)\mathbf{F}_h^*(f)\right)_k, \\ 0, & \text{otherwise},\end{cases} \tag{10}$$

for all $l\in\mathbb{Z}$, $1\leq k\leq m$ and $f\in\left[-\frac{f_s}{2m},\frac{f_s}{2m}\right]$.

Denote by $\left(\mathbf{F}_h\mathbf{F}_h^*\right)_k$ the k^{th} largest entry of $\mathbf{F}_h\mathbf{F}_h^*$, or equivalently, the k^{th} largest element in $\left\{\left|H(f-\frac{lf_s}{m})\right|^2:l\in\mathbb{Z}\right\}$. Then Theorem 4 is derived by observing that the k^{th} largest singular value σ_k of $\mathbf{F}_s^{\ddagger}\mathbf{F}_h$ obeys

$$\sigma_k\leq\sqrt{\left(\mathbf{F}_h\mathbf{F}_h^*\right)_k},\qquad\forall\, 1\leq k\leq m. \tag{11}$$

More importantly, all these upper bounds can be simultaneously achieved by the prefilter (10), which is hence universally optimal regardless of the water level.

Since $\mathbf{F}_s^{\ddagger}$ has orthonormal rows, it acts as an orthogonal projection and outputs the m-dimensional subspace closest to the channel space spanned by $\mathbf{F}_h$. By observing that the rows of the diagonal matrix $\mathbf{F}_h$ are orthogonal to each other, the subspace with the strongest signal strength corresponds to the m rows of $\mathbf{F}_h$ that contain the highest channel gain out of the entire correlated frequency set $\left\{f-\frac{lf_s}{m}\mid l\in\mathbb{Z}\right\}$.

In a *monotone* channel, the filter bank will sequentially crop out the m best frequency bands, each of bandwidth f_s/m. Concatenating all of them results in a low-pass filter with cut-off frequency $f_s/2$, which is equivalent to the optimal prefilter under uniform sampling with a single filter. In other words, using filter banks harvests no gain in capacity compared with a single filter. In this case, the sampled capacity with the optimal filter bank increases monotonically with the sampling rate up to the Nyquist-rate capacity. For more general channels, however, the capacity is not a monotone function of f_s, which again results from the use of uniform sampling. As illustrated in Fig. 3, when we apply a bank of two filters prior to uniform sampling, the capacity curve is still non-monotonic but outperforms sampling with a single filter.

Another consequence of our results is that when the number of prefilters is appropriately chosen, the Nyquist-rate channel capacity can be achieved by sampling at or above the Landau rate. In order to show this, we introduce the following notion of channel permutation. We call $\tilde{H}(f)$ a *permutation* of a channel response $H(f)$ at rate f_s if for any f, $\left\{\tilde{H}(f-lf_s):l\in\mathbb{Z}\right\}=\left\{H(f-lf_s):l\in\mathbb{Z}\right\}$. In other words, $[\cdots,\tilde{H}(f-f_s),\tilde{H}(f),\tilde{H}(f+f_s),\cdots]$ is a permutation of $[\cdots,H(f-f_s),H(f),H(f+f_s),\cdots]$ for every f. The following proposition, proved in [7], characterizes a sufficient condition that allows the Nyquist-rate channel capacity to be achieved at any sampling rate at or above the Landau rate.

Proposition 1 *If there exists a permutation $\tilde{H}(f)$ of $H(f)$ at rate $\frac{f_s}{m}$ such that the support of $\tilde{H}(f)$ is $\left[-\frac{f_L}{2},\frac{f_L}{2}\right]$, then optimal uniform sampling with a bank of m filters achieves Nyquist-rate capacity when $f_s\geq f_L$.*

Examples of channels satisfying Proposition 1 include any multiband channel with N subbands among which K subbands have non-zero channel gain. For any $f_s\geq f_L=\frac{K}{N}f_{\text{NYQ}}$, we are always able to permutate the channel at rate f_s/K to generate a band-limited channel of bandwidth f_L. Hence, oversampling with K filters achieves the Nyquist-rate channel capacity. This is illustrated in Fig. 3 where, for a multiband channel with $K=4$ and $N=10$, sampling with 4 prefilters outperforms that with a single prefilter and with two prefilters and achieves the Nyquist-rate capacity whenever $f_s\geq\frac{2}{5}f_{\text{NYQ}}$.

The joint optimization of input processes and filter banks can also be interpreted from a sampling theoretic perspective. We capture this in the following proposition, also proved in [7], where we again restrict ourselves to the class of zero-mean wide-sense stationary input processes and linear reconstruction for simplicity.

Proposition 2 *Suppose the input signal $x(t)$ is wide-sense stationary with power spectral density $S_X(f)$. For a given system, let $\hat{x}(t)$ denote the optimal linear estimate of $x(t)$ from the digital sequence $\{y_s[n]\}$. Then the filter bank given in (10) minimizes the MSE $\mathbb{E}\left(|x(t)-\hat{x}(t)|^2\right)$ over all filter banks for every t.*

Proposition 2 implies that the filter bank optimizing capacity also minimizes the MSE between the original and the reconstructed signals.

4. REFERENCES

[1] R. G. Gallager, *Information theory and reliable communication*, John Wiley & Sons, Inc, New York, 1968.

[2] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. on Information Theory*, vol. 46, no. 3, pp. 933 –946, May. 2000.

[3] Y. C. Eldar and T. Michaeli, "Beyond bandlimited sampling," *IEEE Signal Processing Magazine*, vol. 26, no. 3, pp. 48 –68, May. 2009.

[4] H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Mathematica*, vol. 117, pp. 37–52, 1967.

[5] P.P. Vaidyanathan, "Generalizations of the sampling theorem: Seven decades after Nyquist," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 48, no. 9, pp. 1094 –1109, Sept. 2001.

[6] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. on Info. Theory*, vol. 51, no. 4, pp. 1261 –1282, Apr. 2005.

[7] Y. Chen, Y. C. Eldar, and A. J. Goldsmith, "Shannon meets Nyquist: capacity limits of sampled analog channels," *in preparation.*

[8] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375 –391, Apr. 2010.