Task-Based Analog-to-Digital Converters for Bandlimited Systems

Peter Neuhaus, Nir Shlezinger, Meik Dörpinghaus, Yonina C. Eldar, and Gerhard Fettweis

***Abstract*—Acquiring digital representations of multivariate continuous-time (CT) signals is a challenge encountered in many signal processing systems. In practice, these signals are often obtained in order to extract some underlying information, i.e., for a specific task. Employing conventional task-agnostic analog-to-digital converters (ADCs) can be inefficient in such cases. In this work, we study task-based ADCs designed to obtain a digital representation of a multivariate CT input process to recover an underlying random parameter vector, referred to as the *task*. The proposed system employs analog filtering, uniform sampling, and scalar uniform quantization of the input process before recovering the task vector using a linear filter. We optimize the analog and digital filters and derive closed-form expressions for the achievable MSE in recovering a task vector from a set of bandlimited signals when utilizing a fixed quantizer resolution and sampling rate satisfying the Shannon-Nyquist sampling theorem. Guidelines for the design of practical acquisition systems are obtained from the structure of the MSE minimizing analog filter. Our numerical results, which consider the recovery of a set of matched filter outputs under a rate budget, demonstrate that the proposed approach substantially outperforms both, implementing the matched filter solely in the analog or digital domain.**

***Index Terms*—quantization, sampling, estimation, analog-to-digital converter**

I. Introduction

Analog-to-digital converters (ADCs) allow physical signals to be processed using digital hardware. ADCs perform two operations: sampling, i.e., converting a continuous-time (CT) signal into a discrete set of samples, and quantization, where the continuous-amplitude samples are converted into a finite-bit representation. Conventional ADCs are designed to recover the input signal, where the sampling rate is typically chosen matched to the bandwidth of the input signal, while the quantizer resolution is chosen such that the quantization distortion is sufficiently small [1]. If the task of the system is not to reconstruct the CT input, but, e.g., to extract some information from it, this approach can be inefficient. Since the power consumption of ADCs grows with the sampling rate and the quantization resolution [2], such inefficiency directly leads to increased power consumption. The power consumption of conventional ADCs is considered a significant challenge in beyond 5G systems, which are foreseen to utilize a large number of antennas, i.e., massive multiple-input multiple-output (MIMO), as well as large bandwidths in the millimeter wave bands, to meet the ever-increasing demand for higher data rates.

Recently, it was demonstrated that a-priori knowledge about the system task may be utilized in order to design *task-based quantizers* [3]. The works [4]–[7] focused on the quantization aspect of analog-to-digital conversion, showing that the distortion induced by low-resolution quantization can be mitigated by accounting for a task in the system. This was achieved by introducing pre-quantization processing, resulting in hybrid analog-digital systems, as commonly utilized in MIMO communications systems for the purpose of reducing the number of RF chains [8]–[10]. For the sampling operation, such analog processing was shown to facilitate the reconstruction of sub-Nyquist sampled, frequency-sparse analog signals [11], [12], as well as to exploit spatial correlation via joint sampling of multivariate CT signals [13]. The works [11]–[13] all focused on the reconstruction of the sampled signals, and thus did not consider the presence of a task.

Joint sampling and quantization has been investigated in [14]–[16], where the minimum achievable reconstruction distortion under a given rate budget is studied as an indirect source coding problem. However, the performance of those systems can only be achieved by vector quantizers, which are challenging for implementation. Moreover, none of these works accounts for a task. Finally, the recent work [17] used deep learning to design task-based ADCs including both sampling and quantization, empirically demonstrating the potential gains of such joint designs for MIMO systems without providing a theoretical analysis.

In this work, we consider the design and analysis of ADCs for the task of recovering a linear function of the observed signals. Such tasks can represent, e.g., channel estimation as studied in [5] or matched filtering, as considered in our numerical analysis. We focus on bandlimited signals with sampling rates satisfying the Shannon-Nyquist sampling theorem. Following [3]–[7], [17], we consider a hybrid system with pre-acquisition analog combining while utilizing uniform samplers and quantizers, and optimize the system in light of the task.

We analytically characterize the minimum achievable mean squared error (MSE) in recovering the desired task from the analog input signals under a given sampling rate and quantizer resolution for the considered system model. Furthermore, we

This work has received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 164481002 – SFB 912, HAEC, from the German Federal Ministry of Education and Research (BMBF) (project E4C, contract number 16ME0189), from the Benoziyo Endowment Fund for the Advancement of Science, the Estate of Olga Klein – Astrachan, the European Union's Horizon 2020 research and innovation program under grant No. 646804-ERC-COG-BNYQ and from the Israel Science Foundation under grant No. 0100101. Computations were performed at the Center for Information Services and High Performance Computing (ZIH) at TU Dresden.

P. Neuhaus, M. Dörpinghaus, and G. Fettweis are with the Vodafone Chair Mobile Communications Systems, Technische Universität Dresden, 01062 Dresden, Germany (e-mail: {peter_friedrich.neuhaus, meik.doerpinghaus, gerhard.fettweis}@tu-dresden.de). N. Shlezinger is with the School of ECE, Ben-Gurion University of the Negev, Beer-Sheva, Israel (e-mail: nirshl@bgu.ac.il). Y. C. Eldar is with the Faculty of Math and CS, Weizmann Institute of Science, Rehovot, Israel (e-mail: yonina.eldar@weizmann.ac.il).

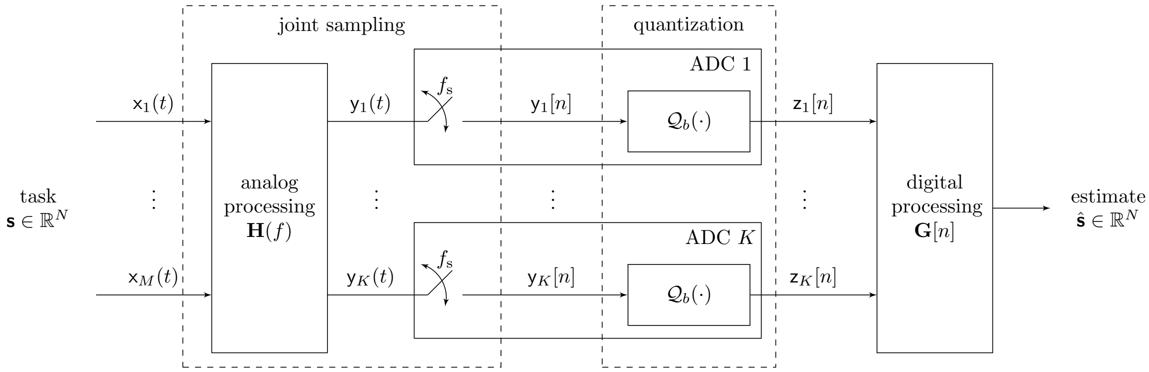

Fig. 1. Overview of the system model. The task is to estimate the random vector $\mathbf{s} \in \mathbb{R}^N$. The digital task estimate is denoted as $\hat{\mathbf{s}} \in \mathbb{R}^N$.

obtain analytical expressions for the linear analog and digital filters, which achieve this distortion. Our results show that neither recovering the task in the analog domain and subsequently sampling and quantizing the analog minimum MSE (MMSE) estimate nor a fully digital architecture, which estimates the task solely in the digital domain, is optimal. In our numerical study, we apply the proposed task-based ADC system to the estimation of the matched filter output in a MIMO system, demonstrating the resulting gains over analog and digital matched filtering.

II. System Model

We consider the hybrid ADC system model illustrated in Fig. 1. Here, a set of M CT zero-mean jointly wide-sense stationary (WSS) random signals $\{\mathsf{x}_m(t)\}_{m=1}^M, t \in \mathbb{R}$, with joint power spectral density (PSD) $\mathbf{C}_{\mathsf{x}}(f) \in \mathbb{C}^{M \times M}, f \in \mathbb{R}$, are acquired for the task of estimating a zero-mean random vector $\mathbf{s} \in \mathbb{R}^N$. The signals $\{\mathsf{x}_m(t)\}$ are bandlimited to $\left(-\frac{f_{\mathrm{nyq}}}{2}, \frac{f_{\mathrm{nyq}}}{2}\right)$. We assume a known statistical relationship between the CT random input signals $\{\mathsf{x}_m(t)\}$ and the random vector of interest $\mathbf{s}$. In particular, we assume that the MMSE estimate of $\mathbf{s}$ from $\{\mathsf{x}_m(t)\}_{m=1}^M$ takes a linear form. By defining $\mathbf{x}(t) = [\mathsf{x}_1(t), \ldots, \mathsf{x}_M(t)]^T$, this assumption implies that there exists a $\mathbf{\Gamma}(t) \in \mathbb{R}^{N \times M}$ such that

$$\tilde{\mathbf{s}} = \mathbb{E}\left\{\mathbf{s} \middle| \{\mathsf{x}_m(t)\}_{m=1}^M\right\} = \int_{\mathbb{R}} \overleftarrow{\mathbf{\Gamma}}(t)\mathbf{x}(t)\mathrm{d}t = (\mathbf{\Gamma} * \mathbf{x})(0), \quad (1)$$

where we use the shorthand notation $\overleftarrow{\mathbf{\Gamma}}(t) = \mathbf{\Gamma}(-t)$ for time reversal. The resulting formulation models tasks which can be expressed as linear functions of the observed signals. Such problems, where the task is a linear function of the observations, often arise in practice, e. g., in channel estimation or matched filtering, as studied in Section IV.

A. Joint Sampling Operation

Our sampling operation implements *joint sampling*, which is a framework for sampling a set of CT signals, allowing to exploit their spatial correlation using analog filters [13]. The M CT input signals are filtered by a multivariate analog filter $\mathbf{H}(f) \in \mathbb{R}^{K \times M}, f \in \mathbb{R}$, and the kth output is given by

$$\mathsf{y}_k(t) = \sum_{m=1}^{M} (h_{k,m} * \mathsf{x}_m)(t), \quad k \in \mathcal{K} = \{1, \ldots, K\}, \quad (2)$$

where $h_{k,m}(t)$ denotes a scalar filter which is the inverse Fourier transform (FT) of $[\mathbf{H}]_{k,m}(f)$. The outputs $\{\mathsf{y}_k(t)\}_{k=1}^K$ of the analog filter $\mathbf{H}(f)$ are sampled by K identical uniform samplers, each with sampling rate f_{s}, i.e., the samples are spaced by $T_{\mathrm{s}} = \frac{1}{f_{\mathrm{s}}}$. Defining the sampling operation as $\mathsf{y}_k[n] = T_{\mathrm{s}}\mathsf{y}_k(nT_{\mathrm{s}})$ yields the discrete-time sequences $\{\mathsf{y}_k[n]\}_{k=1}^K$.

B. Quantization Operation

The sampled signals $\{\mathsf{y}_k[n]\}_{k=1}^K$ are subsequently quantized by K identical uniform scalar mid-rise quantizers with an amplitude resolution of b bits, i.e., each quantizer can produce 2^b distinct output values. The (one-sided) dynamic range of the quantizers is denoted as $\gamma > 0$, such that the quantization step size is given by $\Delta = \frac{2\gamma}{2^b}$. The mid-rise quantization function is then defined as

$$q(x') = \begin{cases} \Delta\left(\left\lfloor \frac{x'}{\Delta} \right\rfloor + \frac{1}{2}\right), & \text{for } |x'| < \gamma \\ \mathrm{sign}(x')\left(\gamma - \frac{\Delta}{2}\right), & \text{else,} \end{cases} \quad (3)$$

where $\lfloor \cdot \rfloor$ denotes rounding to the next smaller integer and $\mathrm{sign}(\cdot)$ is the signum function.

In order to obtain an analytically tractable model for the non-linear quantizers we assume that they are implemented as *nonsubtractive dithered quantizers*, i.e., an additional signal referred to as *dither* is added to the quantizer input prior to quantization [18]. Nonsubtractive denotes therein that the dither is not subtracted, i.e., compensated, after quantization. Hence, the quantizer outputs are given by

$$\mathsf{z}_k[n] = \mathcal{Q}_b(\mathsf{y}_k[n]) = q(\mathsf{y}_k[n] + \mathsf{w}_k[n]), \qquad k \in \mathcal{K}, \quad (4)$$

where $\{\mathsf{w}_k[n]\}_{k=1}^K$ denotes the zero-mean dither random process, which is i.i.d. in time and space, and mutually independent of the input process. Following [18], the probability density function of the dither signal is chosen as a triangular function with a width of 2Δ.

Quantizers are typically required to operate within their dynamic range to yield distinguishable digital representations of different inputs. Consequently, the overload probability, i.e., the probability that the magnitude of the input exceeds the dynamic range, has to be negligible. Hence, the dynamic range

γ is chosen as a multiple η of the largest standard deviation of the dithered input, i.e.,

$$\gamma^2 = \eta^2 \max_{k\in\mathcal{K}} \mathbb{E}\{\tilde{\mathsf{y}}_k^2[n]\}, \tag{5}$$

with $\tilde{\mathsf{y}}_k[n] = \mathsf{y}_k[n] + \mathsf{w}_k[n]$. Using Chebychev's inequality, the overload probability is now upper-bounded by [19, eq. (5-88)]

$$\Pr\{|\tilde{\mathsf{y}}_k[n]| \geq \gamma\} \leq \frac{\mathbb{E}\{\tilde{\mathsf{y}}_k^2[n]\}}{\eta^2 \max_{k'\in\mathcal{K}} \mathbb{E}\{\tilde{\mathsf{y}}_{k'}^2[n]\}} \leq \frac{1}{\eta^2}. \tag{6}$$

For the considered dithered quantizer model and vanishing overload probability, i.e., $\Pr\{|\tilde{\mathsf{y}}_k[n]| \geq \gamma\} = 0$, it follows from [18, Th. 2] that the quantizer output can be written as

$$\mathbf{z}[n] = \mathbf{y}[n] + \mathbf{e}[n]. \tag{7}$$

The quantization error $\mathbf{e}[n]$ in (7) satisfies: *i*) $\mathbf{e}[n]$ is uncorrelated with the input $\mathbf{y}[n]$; *ii*) the entries of $\mathbf{e}[n]$ are uncorrelated in both time and space; and *iii*) the autocorrelation function of the quantization error $\mathbf{e}[n]$ is given by

$$\mathbf{C}_{\mathbf{e}}[l] = \mathbb{E}\left\{\mathbf{e}[n+l]\mathbf{e}^T[n]\right\} = \frac{\Delta^2}{4}\mathbf{I}_K\delta[l], \tag{8}$$

where $\delta[n]$ denotes the Kronecker delta function and $\mathbf{I}_K$ denotes the identity matrix of size $K \times K$. While the resulting model of the quantization error rigorously holds for overload-free nonsubtractive dithered quantizers, it also holds approximately for conventional (non-dithered) uniform quantizers applied to a broad range of inputs [20].

C. Problem Formulation

For a fixed number of ADCs, K, a fixed sampling rate f_s, and a fixed quantizer resolution b, our goal is to design the task-based ADC system of Fig. 1 such that the MSE in recovering the task vector $\mathbf{s}$ is minimized. We focus on linear digital recovery using a digital filter $\mathbf{G}[n] \in \mathbb{R}^{N\times K}$, such that

$$\hat{\mathbf{s}} = \sum_{n\in\mathbb{Z}} \overleftarrow{\mathbf{G}}[n]\mathbf{z}[n] = (\mathbf{G} * \mathbf{z})[0]. \tag{9}$$

The design parameters are thus the analog filter $\mathbf{H}(f)$ and the digital filter $\mathbf{G}[n]$. Hence, we seek to characterize the minimal MSE for the considered setup, given by $\min \mathbb{E}\{\|\mathbf{s} - \hat{\mathbf{s}}\|^2\}$. We focus on sampling rates satisfying the Shannon-Nyquist sampling theorem [1, Th. 4.1], i.e., $f_s \geq f_{\text{nyq}}$. In this case, $\{\mathsf{y}_k[n]\}_{k=1}^K$ is not distorted by aliasing in the frequency domain. An investigation of task-based ADCs with sub-Nyquist sampling w.r.t. the filtered input $\{\mathsf{y}_k(t)\}_{k=1}^K$, i.e., employing $f_s < f_{\text{nyq}}$, is treated in [21].

III. Task-Based ADCs

In this section, we present a theoretical analysis of task-based ADCs. First we derive the minimum achievable MSE and characterize the corresponding system. Then, we discuss the main insights which arise from our derivations.

In order to derive the minimum achievable MSE, we note that the MSE in recovering $\mathbf{s}$, i.e., $\mathbb{E}\{\|\mathbf{s} - \hat{\mathbf{s}}\|^2\}$, can always be decomposed as [22, Appendix]

$$\mathbb{E}\left\{\|\mathbf{s} - \hat{\mathbf{s}}\|^2\right\} = \mathbb{E}\left\{\|\mathbf{s} - \tilde{\mathbf{s}}\|^2\right\} + \mathbb{E}\left\{\|\tilde{\mathbf{s}} - \hat{\mathbf{s}}\|^2\right\}, \tag{10}$$

i.e., as the sum of the MMSE in recovering $\mathbf{s}$ from $\mathbf{x}(t)$ and the MSE of the digital estimate $\hat{\mathbf{s}}$ w.r.t. $\tilde{\mathbf{s}}$. In the following, we thus minimize $\mathbb{E}\{\|\tilde{\mathbf{s}} - \hat{\mathbf{s}}\|^2\}$, as $\mathbb{E}\{\|\mathbf{s} - \tilde{\mathbf{s}}\|^2\}$ is independent of the ADC. The proofs are omitted due to space limitations; they can be found in [21].

First, we identify the MSE minimizing linear digital recovery filter $\mathbf{G}[n]$ for a given analog filter $\mathbf{H}(f)$, as stated in the following proposition:

Proposition 1. *For a given analog filter $\mathbf{H}(f) \in \mathbb{R}^{K\times M}$, a fixed ADC configuration (K, f_s, b), where $f_s \geq f_{\text{nyq}}$, the MSE minimizing linear digital recovery filter is given by*

$$\mathbf{G}^{\text{o}}[n] = T_s \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \mathbf{S}(f)\mathbf{C}_{\mathbf{z}}^{-1}(e^{j2\pi f T_s})e^{j2\pi f n T_s}\text{d}f, \tag{11}$$

where $\mathbf{S}(f) = \mathbf{\Gamma}(f)\mathbf{C}_{\mathbf{x}}(f)\mathbf{H}^H(f)$, $[\mathbf{\Gamma}]_{n,m}(f)$ is given by the FT of $[\mathbf{\Gamma}]_{n,m}(t)$ and $\mathbf{C}_{\mathbf{z}}(e^{j2\pi f T_s}) = T_s\mathbf{H}(f)\mathbf{C}_{\mathbf{x}}(f)\mathbf{H}^H(f) + \frac{\Delta^2}{4}\mathbf{I}_K$. Furthermore, the resulting minimum achievable MSE is given by

$$\text{MSE}(\mathbf{H}(f)) = \mathbb{E}\left\{\|\tilde{\mathbf{s}}\|^2\right\} - \text{tr}\left(T_s \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \mathbf{Q}(f)\text{d}f\right). \tag{12}$$

where $\mathbf{Q}(f) = \mathbf{S}(f)\mathbf{C}_{\mathbf{z}}^{-1}(e^{j2\pi f T_s})\mathbf{S}^H(f)$.

After obtaining the minimum achievable MSE for a given analog filter $\mathbf{H}(f)$ and a fixed ADC configuration (K, f_s, b), our goal is to find the analog filter $\mathbf{H}(f)$ which minimizes (12). The result is summarized in the following theorem:

Theorem 1. *For a fixed ADC configuration (K, f_s, b), where $f_s \geq f_{\text{nyq}}$, the MSE minimizing analog filter $\mathbf{H}^{\text{o}}(f) \in \mathbb{C}^{K\times M}$ is given by*

$$\mathbf{H}^{\text{o}}(f) = \mathbf{U}_{\mathbf{H}}(f)\mathbf{\Sigma}_{\mathbf{H}}(f)\mathbf{V}_{\mathbf{H}}^H(f)\left(\mathbf{C}_{\mathbf{x}}^{1/2}(f)\right)^{\dagger}, \tag{13}$$

where $(\cdot)^{\dagger}$ denotes the pseudo-inverse and

- *$\mathbf{V}_{\mathbf{H}}(f) \in \mathbb{C}^{M\times M}$ is the matrix of right-singular vectors of $\mathbf{\Gamma}(f)\mathbf{C}_{\mathbf{x}}^{1/2}(f)$.*
- *$\mathbf{\Sigma}_{\mathbf{H}}(f) \in \mathbb{R}^{K\times M}$ is a diagonal matrix with diagonal entries $[\mathbf{\Sigma}_{\mathbf{H}}(f)]_{i,i} = 2^{-b}\sqrt{\left(\zeta\sigma_{\bar{\mathbf{\Gamma}},i}(f) - 1\right)^+}$, where $\sigma_{\bar{\mathbf{\Gamma}},i}(f)$ denotes the ith singular value of $\mathbf{\Gamma}(f)\mathbf{C}_{\mathbf{x}}^{1/2}(f)$ and for $\bar{\kappa} = \eta^2(1 - \frac{2\eta^2}{3\,2^{2b}})^{-1}$, ζ is chosen such that*

$$\frac{\bar{\kappa}T_s}{K}\sum_{i=1}^{\min(K,M)}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} [\mathbf{\Sigma}_{\mathbf{H}}(f)]_{i,i}^2\,\text{d}f = 1.$$

- *$\mathbf{U}_{\mathbf{H}}(f) \in \mathbb{C}^{K\times K}$ is a unitary matrix which ensures identical diagonal entries of $\mathbf{U}_{\mathbf{H}}(f)\mathbf{\Sigma}_{\mathbf{H}}(f)\mathbf{\Sigma}_{\mathbf{H}}^H(f)\mathbf{U}_{\mathbf{H}}^H(f)$.*

The dynamic range of the quantizer is set to $\gamma = 1$, and the resulting minimum achievable MSE is given by

$$\text{MSE} = \mathbb{E}\left\{\|\tilde{\mathbf{s}}\|^2\right\} - \sum_{i=1}^{\min(K,M)}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \frac{\sigma_{\bar{\mathbf{\Gamma}},i}^2(f)\,[\mathbf{\Sigma}_{\mathbf{H}}(f)]_{i,i}^2}{[\mathbf{\Sigma}_{\mathbf{H}}(f)]_{i,i}^2 + 2^{-2b}}\,\text{d}f. \tag{14}$$

The MSE minimizing task-based ADC under the considered system model is characterized by the combination of Theorem 1 and Proposition 1.

From Theorem 1 we can obtain guidelines for designing task-based ADCs by examining the structure of the optimal analog filter $\mathbf{H}^{\circ}(f)$: First, a whitening filter, which is matched to the input random process $\mathbf{x}(t)$, is applied. Then, $\mathbf{V}_{\mathbf{H}}(f)$ is the matrix of right-singular vectors of the analog MMSE filter for obtaining $\tilde{\mathbf{s}}$ from the whitened input. However, the remaining terms $\mathbf{U}_{\mathbf{H}}(f)\mathbf{\Sigma}_{\mathbf{H}}(f)$ are different from recovering the MMSE estimate in analog: The 'water-filling'-type expression in $\mathbf{\Sigma}_{\mathbf{H}}(f)$ nullifies the weak eigenmodes, which cannot be resolved with the given quantizer resolution. This corresponds to finding the optimal trade-off between estimation error and quantization distortion. Finally, $\mathbf{U}_{\mathbf{H}}(f)$ rotates the signal such that each ADC processes a signal with an identical variance, i. e., optimally utilizing the dynamic range of all quantizers.

The form of the resulting task-based ADC of Theorem 1 bears some similarity to the task-based quantizer derived in [4]. This follows from focusing on Nyquist rate sampling of bandlimited signals, which mitigates the distortion expected to be induced by arbitrary sampling. However, in this work, the sampling operation can mitigate the error induced by quantization, which is not the case in [4].

The proposed task-based ADC combines both analog and digital processing, tuning the overall system in light of the task. The results in Theorem 1 are quite different from either of the two more intuitive approaches: *i*) *Analog recovery*, where the task is estimated in the analog domain, i.e., $\mathbf{H}(f) = \mathbf{\Gamma}(f)$, and the analog MMSE estimate $\tilde{\mathbf{s}}$ is subsequently quantized by $K = N$ ADCs; *ii*) *Digital recovery*, where no analog processing is employed, i.e., $\mathbf{H}(f) = \mathbf{I}_M$, the CT random process $\mathbf{x}(t)$ is digitized by $K = M$ ADCs and the task is subsequently estimated fully in the digital domain. Since in general $\mathbf{H}^{\circ}(f) \neq \mathbf{\Gamma}(f)$ and $\mathbf{H}^{\circ}(f) \neq \mathbf{I}_M$, it can be concluded that none of the alternative designs is generally optimal in minimizing the MSE. This observation is numerically verified in Section IV. Even larger gains can be obtained when employing sub-Nyquist sampling, which is treated in [21].

IV. Numerical Results

In this section, we evaluate the performance of the proposed task-based ADC in terms of its *normalized MSE*, given by $\frac{\mathbb{E}\left\{\|\tilde{\mathbf{s}}-\hat{\mathbf{s}}\|^2\right\}}{\mathbb{E}\left\{\|\tilde{\mathbf{s}}\|^2\right\}}$. As an example, we consider the problem of estimating the output of a set of matched filters. Digital matched filtering, as typically implemented [23, Sec. 4.3.3], is expected to be sub-optimal for systems with low-resolution ADCs. Let $\mathbf{x}(t)$ be an observed multivariate random process at a MIMO receiver, which is given by

$$\mathbf{x}(t) = \left(\mathbf{F} * \tilde{\mathbf{x}}\right)(t) + \mathbf{n}(t), \tag{15}$$

where $\tilde{\mathbf{x}}(t)$ denotes a zero-mean Gaussian transmit signal with autocorrelation function $\mathbf{C}_{\tilde{\mathbf{x}}}(\tau) = \mathbf{I}_N\delta(\tau)$, $\mathbf{F}(t) = \tilde{\mathbf{F}}\,\frac{1}{T_{\text{nyq}}}\text{sinc}(\frac{t}{T_{\text{nyq}}})$, $\tilde{\mathbf{F}} \in \mathbb{R}^{M\times N}$ denotes the channel, and $\mathbf{n}(t)$ is an additive white Gaussian noise process with autocorrelation function $\mathbf{C}_{\mathbf{n}}(\tau) = \frac{N_0}{2}\frac{1}{T_{\text{nyq}}}\text{sinc}(\frac{t}{T_{\text{nyq}}})\mathbf{I}_M$. Note that the channel $\mathbf{F}(t)$ and noise $\mathbf{n}(t)$ are bandlimited to $f \in (-\frac{f_{\text{nyq}}}{2}, \frac{f_{\text{nyq}}}{2})$ with $f_{\text{nyq}} = \frac{1}{T_{\text{nyq}}}$. We assume our task is to estimate the noiseless matched filter output, i.e.,

$$\mathbf{s} = \left(\overleftarrow{\mathbf{F}}^T * \mathbf{F} * \tilde{\mathbf{x}}\right)(0). \tag{16}$$

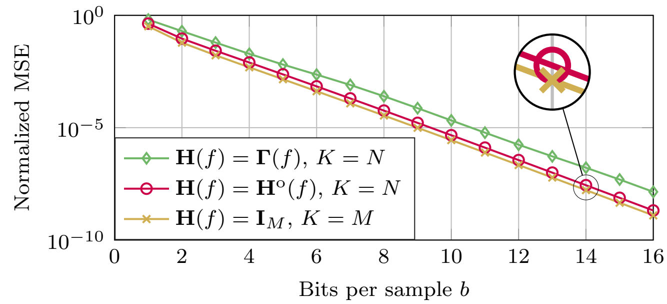

Fig. 2. Normalized MSE vs. number of bits per sample b, $f_{\text{s}} = f_{\text{nyq}}$. The task-based $\mathbf{H}(f) = \mathbf{H}^{\circ}(f)$ achieves nearly the same performance as digital recovery, i.e., $\mathbf{H}(f) = \mathbf{I}_M$, using only $K = 4$ instead of $K = 16$ ADCs.

Hence, the analog MMSE filter $\mathbf{\Gamma}(f)$ is given by [19]

$$\mathbf{\Gamma}(f) = \mathbf{F}^H(f)\mathbf{F}(f)\mathbf{F}^H(f)\left(\mathbf{F}(f)\mathbf{F}^H(f) + \mathbf{C}_{\mathbf{n}}(f)\right)^{-1}, \tag{17}$$

where $[\mathbf{F}]_{n,m}(f)$ is given by the FT of $[\mathbf{F}]_{n,m}(t)$.

With the above, we evaluate the performance for $N = 4$ and $M = 16$, i.e., 4 transmit and 16 receive antennas. The filter bandwidth is chosen as $f_{\text{nyq}} = 400\,\text{MHz}$. The channel is modeled as $\tilde{\mathbf{F}} = \mathbf{C}_{\text{Rx}}^{1/2}\mathbf{H}_{\text{ch}}$, where $\mathbf{C}_{\text{Rx}} \in \mathbb{R}^{M\times M}$ models the spatial correlation at the receiver; it is given by (cf. [24, eq. (19)])

$$[\mathbf{C}_{\text{Rx}}]_{m,n} = \frac{\left(1 - e^{-\sqrt{2}\pi/\sigma_\phi}\right)^{-1}}{1 + \frac{\sigma_\phi^2}{2}(\pi(m-n))^2}, \tag{18}$$

where we use $\sigma_\phi = 1°$. Moreover, $\mathbf{H}_{\text{ch}} \in \mathbb{R}^{M\times N}$ is assumed to be known and fixed; it contains independent entries which are generated randomly as $[\mathbf{H}_{\text{ch}}]_{m,n} \sim \mathcal{N}(0,1)$. We set $\eta(b) = 0.25b + 1.75$, i. e., we increase the dynamic range with the amplitude resolution (cf. (5)), in order to ensure the validity of the model, which has been verified numerically. Furthermore, we define $\text{SNR} = \frac{2}{N_0}\text{tr}\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^T\right) = 10\,\text{dB}$.

First, in Fig. 2, we evaluate the minimum achievable MSE of the proposed system given by (14), which employs the MSE minimizing analog filter $\mathbf{H}^{\circ}(f)$ given by (13) for Nyquist rate sampling, i.e., $f_{\text{s}} = f_{\text{nyq}}$, when varying the number of bits per sample b. For this evaluation, we set the number of ADCs of the proposed system to $K = N = 4$. We compare the performance to the alternative strategies of analog and digital recovery, where we evaluate the corresponding MSEs using (12). Observing Fig. 2, we note that digital recovery, which utilizes 16 ADCs with b bits each, yields the lowest MSE. The proposed task-based ADC, which uses only 4 ADCs of the same resolution, namely, 75 % fewer bits, yields a marginally higher MSE, while analog recovery results in the highest MSE. This indicates that substantial resource savings can be achieved by task-based acquisition with minimal effect on the overall system accuracy.

Next we compare the MSE of the proposed system to analog and digital recovery under a fixed rate budget $R = K \cdot b \cdot f_{\text{s}}$. For the proposed system we perform an exhaustive search over all feasible combinations of $b \in \{1, \ldots, 16\}$ and $K \in \{1, \ldots, M\}$. For the analog and digital recovery systems, we

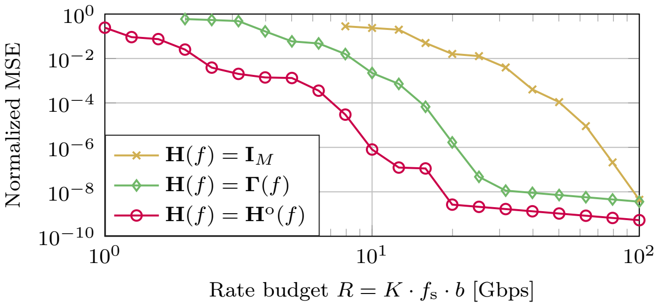

Fig. 3. Achievable MSE for a fixed rate budget $R = K \cdot b \cdot f_{\mathrm{s}}$.

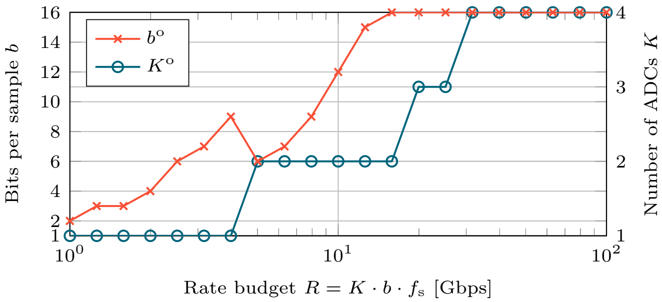

Fig. 4. Evaluation of the MSE minimizing ADC configuration b^{o} and K^{o}. The optimal strategy increases the number of bits b first, before increasing the number of ADCs K when increasing R.

only perform an exhaustive search over b, because K is fixed to N and M, respectively. The sampling rate is then chosen as $f_{\mathrm{s}} = \frac{R}{K \cdot b}$, where combinations of R, b and K which result in $f_{\mathrm{s}} < f_{\mathrm{nyq}}$ are not considered. In Fig. 3 it can be observed that the proposed task-based ADC system outperforms the two competing architectures significantly over a wide range of rates R. At a normalized MSE of 10^{-4}, the proposed system requires only approx. $47\,\%$ and approx. $14\,\%$ of the rate as compared to analog and digital recovery systems, respectively.

Finally, in Fig. 4, we evaluate the combinations of b and K, denoted as b^{o} and K^{o}, which result in the minimum achievable MSE for the proposed task-based ADC system, under a fixed rate budget R, as depicted in Fig. 3. For the considered task, the MSE is minimized for low rates by employing just a single ADC, i.e., $K = 1$, and by choosing the amplitude resolution b as high as possible. The optimal strategy mostly increases the number of bits b first, before increasing the number of ADCs K when increasing R. This result indicates that the quantization error is the major source of distortion in the considered system.

V. Conclusions

In this work, we proposed a task-based acquisition system for linear tasks under rate constraints. The proposed system performs joint sampling using an analog filter, subsequent scalar uniform quantization, and linear digital recovery of the task vector. We obtained closed-form expressions for the minimum achievable MSE and also for the optimal linear processing in the analog and digital domain for a fixed number of ADCs, a fixed sampling rate above the Nyquist rate, and a fixed quantizer resolution. We have demonstrated significant gains in terms of a digital rate budget for the proposed system compared to completely analog or digital recovery strategies. An extension to sub-Nyquist sampling is treated in [21].

References

[1] Y. C. Eldar, *Sampling theory: Beyond Bandlimited Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[2] S. Krone and G. Fettweis, “Energy-efficient A/D conversion in wideband communications receivers,” in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, San Francisco, CA, USA, Sep. 2011, pp. 1–5.

[3] N. Shlezinger and Y. C. Eldar, “Task-based quantization with application to MIMO receivers,” *arXiv preprint arXiv:2002.04290*, 2020.

[4] N. Shlezinger, Y. C. Eldar, and M. R. D. Rodrigues, “Hardware-limited task-based quantization,” *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5223–5238, Oct. 2019.

[5] ——, “Asymptotic task-based quantization with application to massive MIMO,” *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 3995–4012, Aug. 2019.

[6] S. Salamatian, N. Shlezinger, Y. C. Eldar, and M. Médard, “Task-based quantization for recovering quadratic functions using principal inertia components,” in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 390–394.

[7] N. Shlezinger and Y. C. Eldar, “Deep task-based quantization,” *Entropy*, vol. 23, no. 1, p. 104, Jan. 2021.

[8] P. Neuhaus, M. Dörpinghaus, H. Halbauer, S. Wesemann, M. Schlüter, F. Gast, and G. Fettweis, “Sub-THz wideband system employing 1-bit quantization and temporal oversampling,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, Jun. 2020.

[9] S. S. Ioushua and Y. C. Eldar, “A family of hybrid analog–digital beamforming methods for massive MIMO systems,” *IEEE Trans. Signal Process.*, vol. 67, no. 12, pp. 3243–3257, Apr. 2019.

[10] T. Gong, N. Shlezinger, S. S. Ioushua, M. Namer, Z. Yang, and Y. C. Eldar, “RF chain reduction for MIMO systems: A hardware prototype,” *IEEE Syst. J.*, pp. 1–12, May 2020.

[11] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, “Beyond Nyquist: Efficient sampling of sparse bandlimited signals,” *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010.

[12] M. Mishali and Y. C. Eldar, “From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals,” *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 375–391, Apr. 2010.

[13] N. Shlezinger, S. Salamatian, Y. C. Eldar, and M. Médard, “Joint sampling and recovery of correlated sources,” in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 385–389.

[14] A. Kipnis, A. J. Goldsmith, Y. C. Eldar, and T. Weissman, “Distortion rate function of sub-Nyquist sampled Gaussian sources,” *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 401–429, Jan. 2016.

[15] A. Kipnis, Y. C. Eldar, and A. J. Goldsmith, “Fundamental distortion limits of analog-to-digital compression,” *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6013–6033, Sep. 2018.

[16] ——, “Analog-to-digital compression: A new paradigm for converting signals to bits,” *IEEE Signal Process. Mag.*, vol. 35, no. 3, pp. 16–39, May 2018.

[17] N. Shlezinger, R. J. G. van Sloun, I. A. M. Hujiben, G. Tsintsadze, and Y. C. Eldar, “Learning task-based analog-to-digital conversion for MIMO receivers,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Barcelona, Spain, May 2020, pp. 9125–9129.

[18] R. M. Gray and T. G. Stockham, “Dithered quantizers,” *IEEE Trans. Inf. Theory*, vol. 39, pp. 805–812, May 1993.

[19] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. New York, NY: McGraw-Hill, 2002.

[20] B. Widrow, I. Kollar, and M.-C. Liu, “Statistical theory of quantization,” *IEEE Trans. Instrum. Meas*, vol. 45, no. 2, pp. 353–361, Apr. 1996.

[21] P. Neuhaus, N. Shlezinger, M. Dörpinghaus, Y. C. Eldar, and G. Fettweis, “Task-based analog-to-digital converters,” *arXiv preprint arXiv:2009.14088*, 2020.

[22] J. Wolf and J. Ziv, “Transmission of noisy information to a noisy receiver with minimum distortion,” *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 406–411, Jul. 1970.

[23] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. New York, NY, USA: John Wiley & Sons, Inc., 1997.

[24] A. Forenza, D. J. Love, and R. W. Heath, “Simplified spatial correlation models for clustered MIMO channels with different array configurations,” *IEEE Trans. Veh. Commun.*, vol. 56, no. 4, pp. 1924–1934, Jul. 2007.