Hybrid Beamforming for Sub-THz MIMO: How Close to Fully Digital?

L. Antonelli, A. A. D'Amico, L. Sanguinetti
Dipartimento Ingegneria dell'Informazione, University of Pisa, Pisa, Italy

***Abstract*—Ultra-massive MIMO (U-MIMO) at sub-THz frequencies is challenging for fully digital systems. This calls for hybrid architectures, which complicates channel estimation, as limited RF chains provide only linear combinations of received pilots. Additionally, near-field effects at sub-THz frequencies demand an optimized sampling grid in both angular and distance domains. This paper aims to design and compare the performance of a hybrid sub-THz U-MIMO system to a fully digital one. We make use of a new metric to show that angular domain oversampling improves the estimation accuracy of state-of-the-art channel estimation schemes for hybrid architectures. Our analysis focuses on both channel estimation and spectral efficiency, and shows that, in low-mobility scenarios, longer pilot sequences can mitigate the limitations imposed by the reduced number of radio-frequency chains. Comparable spectral efficiency of fully digital systems is achieved, at the cost of large dictionaries.**

***Index Terms*—Sub-THz communications, ultra-massive MIMO, channel estimation, hybrid and fully digital architecture.**

I. Introduction and Motivation

Ultra-MIMO (U-MIMO) in the sub-THz band exploits the extremely short wavelengths to pack many antenna elements into a small space, enabling high beam-focusing and spatial multiplexing gains [1]. This is considered a promising technology to meet the high traffic demands of future wireless communications [2]–[4]. However, in an all-digital architecture, where each antenna element is paired with a dedicated radio-frequency (RF) chain, this would be prohibitively power-intensive. To overcome this issue, the number of RF chains must be decreased by adopting hybrid architectures [5].

Although hybrid systems use fewer RF chains M than antennas N, it is known that hybrid architectures can perform as well as fully digital systems, provided that M is equal to or greater than the number of data streams [6]. However, this equivalence holds under the assumption that fully digital and hybrid systems share the same channel state information (CSI). In real-world applications, where the channel is not a-priori known and must be estimated, the equivalence between hybrid and fully digital systems does not hold, because their channel estimators may have very different accuracies.

Unlike sub-6 GHz communications, where user equipments (UEs) are typically in the far field of the BS array and orthogonal codebooks based on discrete Fourier transform matrices are effective, sub-THz frequencies place UEs in the radiating near field [7]. This shift requires the development of near-field codebooks that are different from those used in far-field beamforming. In near-field beamforming, the range domain becomes essential due to the finite beamforming depth, adding complexity beyond the angular considerations of far-field scenarios [8]. Recent polar domain dictionary designs for linear and planar arrays aim to better control codebook column coherence [9], [10]. Since accurate near-field channel estimation typically requires large codebooks, new channel estimation algorithms with overheads independent of dictionary size have emerged. Among them, the polar-domain simultaneous orthogonal matching pursuit (P-SOMP) algorithm [9] efficiently exploits channel sparsity for near-field channel estimation.

This paper presents the design and performance evaluation of a hybrid sub-THz U-MIMO system using the P-SOMP algorithm for channel estimation. We design the dictionary based on the correlation criterion in [9] and introduce a novel metric for its evaluation, showing that the accuracy improves significantly by increasing sampling density in both angular and distance domains. Notably, adding points in the distance domain alone does not yield the same improvement. Our analysis demonstrates that in low-mobility scenarios, using longer pilot sequences can offset the limitations of fewer RF chains, achieving spectral efficiency (SE) comparable to fully digital systems with least-squares (LS) channel estimation, even as the number of UEs grows. However, this comes at the cost of large dictionaries, which may not be practical in real-world scenarios.

II. Channel estimation

We consider the uplink of a U-MIMO system with K single-antenna active UEs. The BS has $M \geq K$ RF chains and is equipped with $N \gg M$ antennas arranged in a uniform linear array (ULA) with an inter-element spacing of δ. Hence, the array aperture is $\Delta = (N-1)\delta$. We assume a fully-connected architecture in which each antenna is connected to each RF chain. Under line-of-sight (LoS) propagation conditions, the channel from UE k to the nth antenna of the BS is

$$h_{kn} = \sqrt{\xi_{kn}} e^{-j\frac{2\pi}{\lambda} d_{kn}} \tag{1}$$

where ξ_{kn} accounts for the path loss and the transmit/receive antenna gains, d_{kn} denotes the Euclidean distance between the centers of the transmitting antenna of UE k and the nth receiving antenna, and λ indicates the wavelength. We call $\mathbf{h}_k = [h_{k1}, \ldots, h_{kN}]^{\mathsf{T}} \in \mathbb{C}^N$ the channel vector of UE k. UEs are co-planar with the ULA and randomly displaced choosing distances from the BS from $\mathcal{U}(\rho_{\min}, \rho_{\max})$ and azimuth angles from $\mathcal{U}(\varphi_{\min}, \varphi_{\max})$. All parameters are chosen according to Tab. I. The boundary between the near-field and far-field of the BS array is typically given by the Rayleigh distance, $\rho_R =$

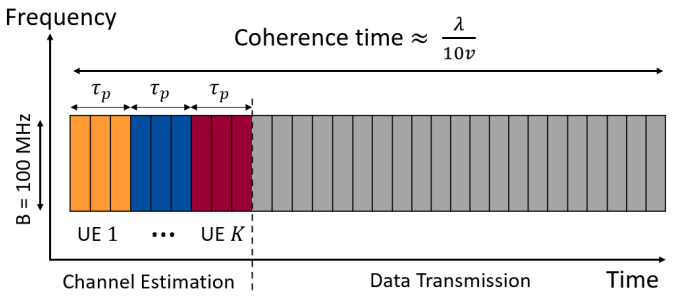


Fig. 1: Uplink transmission protocol for hybrid architectures.

Parameter	Value
Carrier frequency	$f_c = 300\,\text{GHz}$
Wavelength	$\lambda = 1\,\text{mm}$
Number of antennas	$N = 256$
Antenna spacing	$\delta = \lambda/2$
Bandwidth	$B = 100\,\text{MHz}$
Noise power	$\sigma^2 = -86\,\text{dBm}$
UE azimuth	$\varphi_{\min} = -\pi/3, \varphi_{\max} = \pi/3$
UE distance	$\rho_{\min} = 5\,\text{m},\ \rho_{\max} = 25\,\text{m}$
Transmit Power P	15 dBm

Table I: System parameters.

$\frac{2\Delta^2}{\lambda}$. Since UE distances range from $\rho_{\min} = 5\,\text{m}$ to $\rho_{\max} = 25\,\text{m}$, all UEs are in the near-field of the the BS array.

The standard time division duplex protocol is used [11, Sec. 2.3.2], where τ_c available channel uses are divided into τ for the UL channel estimation phase and $\tau_c - \tau$ for data transmission. Next, we review the channel estimation with hybrid and fully digital architectures. In the former case, the P-SOMP channel estimation algorithm proposed in [9] is used.

A. Hybrid Architecture

During the training phase, UEs transmit pilot sequences of length τ_p, following to a time-division multiple access scheme shown in Fig. 1. Consequently, the total number of channel uses required by the training phase is $\tau = K\tau_p$. This overhead could be further reduced using different protocols, such as with multi-carrier waveforms [9], but we do not consider them because we operate in a low-mobility scenario, where the overhead is already much smaller than the coherence block size. Without loss of generality, we focus on an arbitrary UE and we drop the user index k. Let $\boldsymbol{\psi} \in \mathbb{C}^{\tau_p}$ represent the pilot sequence, with $|\psi_i|^2 = 1$ for $i = 1, 2, \ldots, \tau_p$, so that $\|\boldsymbol{\psi}\|^2 = \tau_p$, and let $\mathbf{A}_i \in \mathbb{C}^{M\times N}$ denote the analog combining matrix corresponding to the ith pilot symbol, whose entries are randomly chosen in $\{\pm 1/\sqrt{N}\}$ with equal probability [12]. Then, the ith received pilot signal $\mathbf{y}_i \in \mathbb{C}^M$ can be written as

$$\mathbf{y}_i = \mathbf{A}_i \left(\sqrt{P}\psi_i \mathbf{h} + \mathbf{n}_i\right) \tag{2}$$

where P is the transmit power and $\mathbf{n}_i \in \mathcal{N}_{\mathbb{C}}\left(0, \sigma^2 \mathbf{I}_N\right)$ accounts for the noise. Without loss of generality, we assume $\psi_i = 1$ for $i = 1, 2, \ldots, \tau_p$.

The P-SOMP algorithm, introduced in [9], estimates the channel by using the observation vectors $\mathbf{y} = [\mathbf{y}_1^\mathsf{T}\, \mathbf{y}_2^\mathsf{T} \cdots \mathbf{y}_{\tau_p}^\mathsf{T}]^\mathsf{T} \in \mathbb{C}^{M\tau_p}$, collected during the channel estimation phase. It is based on the polar-domain representation $\mathbf{h}^{\mathcal{P}}$ of the channel $\mathbf{h}$, i.e., $\mathbf{h} = \mathbf{W}\mathbf{h}^{\mathcal{P}}$ given in [9, eq. (8)], where $\mathbf{W} \in \mathbb{C}^{N\times Q}$ is a suitable *dictionary*. The Q columns of $\mathbf{W}$ are the array steering vectors $\mathbf{w}(\varphi, \rho)$ [9, eq. (7)] computed over a discrete grid $\mathbf{G}$ of (φ, ρ) values. A good design of $\mathbf{G}$ is essential for the operation of P-SOMP, and will be discussed in detail in Section III.

B. Fully digital Architecture

In a fully digital architecture, the number of RF chains equals the number of antennas, i.e., $M = N$, and each antenna is connected to a dedicated RF chain. During the training phase, the K UEs transmit over the same time-slot using orthogonal pilot sequences of length $\tau_p \geq K$ [11]. Hence, the total number of channel uses during the training phase is $\tau = K$, instead of $\tau = K\tau_p$ as in the hybrid case. Since the UEs' pilot signals do not interfere, we can focus again on an arbitrary UE and we can drop the user index k. Also, we can assume $\psi_i = 1$. The signal $\mathbf{y}_i$ received in the ith pilot interval can be obtained from (2) by setting $\mathbf{A}_i = \mathbf{I}_N$. Here, we consider the well-known LS estimator, which is simply given by

$$\hat{\mathbf{h}}_{\text{LS}} = \frac{1}{\tau}\sum_{i=1}^{\tau} \mathbf{y}_i \tag{3}$$

and represents the arithmetic mean of the vectors received during the channel estimation phase.

III. Polar-domain dictionary design

The P-SOMP algorithm estimates the channel by utilizing a grid of points, denoted as $\mathbf{G}$, where each grid point is represented by the coordinates $[\rho\cos\varphi, \rho\sin\varphi, 0]$. We begin by reviewing the solution presented in [9], followed by an introduction of our proposed alternative approach.

A. Angular and Distance Sampling [9]

In [9], the grid is designed by considering the maximum correlation between the columns of dictionary $\mathbf{W}$, defined as

$$\mu = \max_{i\neq j} \left\{|\mathbf{w}_i^H \mathbf{w}_j|\right\}, \tag{4}$$

where $\mathbf{w}_i$ and $\mathbf{w}_j$ are two columns of $\mathbf{W}$. On one side, the maximum correlation should be made as small as possible for limiting the number of grid points. On the other side, it cannot be too small because this would lead to a very sparse grid $\mathbf{G}$, with very few number of points and a reduced estimation accuracy. The grid design proposed in [9] leads to a sampling of the angular domain with N points such that $\Phi = \sin\varphi$ is computed over a uniform grid as

$$\Phi = \frac{2n - N + 1}{N}, \quad n = 0, 1, ..., N-1. \tag{5}$$

For each angle φ compatible with (5), the distance is non-uniformly sampled as follows [9, Eq. (15)]

$$\rho = \frac{1}{s}\frac{1}{2\lambda}\left(\frac{N\delta}{\beta}\right)^2 (1 - \Phi^2), \quad s = 1, 2, 3... \tag{6}$$

where the design parameter β controls the distance sampling. Values of the parameter s that result in ρ falling outside the range $[\rho_{\min}, \rho_{\max}]$ or points outside the angular sector

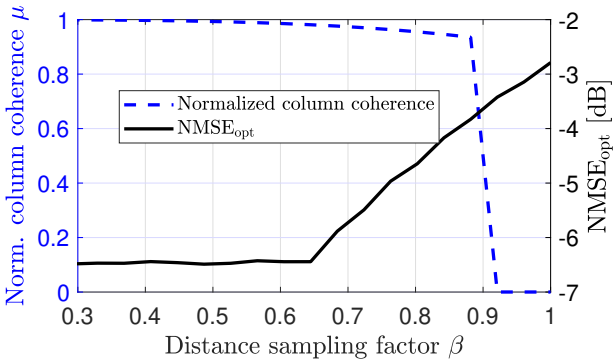


Fig. 2: The maximum normalized column coherence μ and NMSE$_{\text{opt}}$ against β.

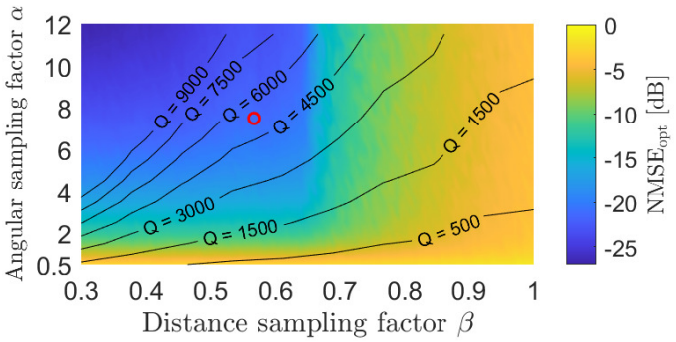


Fig. 3: NMSE$_{\text{opt}}$ and grid size Q for different values of the angular and distance sampling factors

$[\varphi_{\min}, \varphi_{\max}]$ are excluded. This refined grid $\mathbf{G}$ is then used to construct the dictionary $\mathbf{W}$.

Fig. 2 shows μ in (4) as a function of β with parameters in Tab. I. Results show that as β increases, coherence decreases, with a significant drop near $\beta = 0.9$, but this also reduces the dictionary size. A good trade-off between the correlation and the grid size is achieved with $\beta = 0.88$. In this case, the correlation is 0.93 and the grid size is 246.

The correlation metric in (4) does not directly relate to the accuracy of channel estimation. In fact, even though it allows to design the dictionary to improve the localization accuracy [8], the channel acquisition may not be as precise. An alternative is the optimum NMSE[1] defined as

$$\text{NMSE}_{\text{opt}} = 1 - \mathbb{E}\left\{ \min_{\mathbf{g}\in\mathbf{G}} \left[\frac{|\mathbf{h}^H(\mathbf{g})\mathbf{h}(\mathbf{r})|^2}{||\mathbf{h}(\mathbf{g})||^2||\mathbf{h}(\mathbf{r})||^2} \right] \right\} \quad (7)$$

where $\mathbf{h}(\mathbf{r})$ indicates the channel of the UE located in $\mathbf{r}$, and $\mathbf{h}(\mathbf{g})$ is the channel from the point $\mathbf{g} \in \mathbf{G}$. The expectation is computed with respect to the UE's position $\mathbf{r}$. Note that NMSE$_{\text{opt}}$ can be interpreted as a measure of the average error in approximating the channel at an arbitrary point using the channel value at a grid point. From this perspective, lower values of NMSE$_{\text{opt}}$ are expected to improve the estimation accuracy for grid-based algorithms, such as the P-SOMP. This relationship will be further explored in the next section.

Fig. 2 also shows that, despite the correlation μ increasing as β decreases from 0.88 to 0.65, the NMSE$_{\text{opt}}$ reduces to -6.5 dB. We denote as $\mathbf{G}_1$ the grid defined by $\beta = 0.65$, with a total of $Q_1 = 520$ points.

B. Proposed Angular Sampling

The design criterion in [9] allows control over distance sampling through β but not over angular sampling, where points are distributed over N intervals. To address this, we introduce a design parameter $\alpha > 0$ and consider $\lfloor \alpha N \rfloor$ points in the angular domain. This results in $\Phi' = \sin\varphi'$, given by

$$\Phi' = \frac{2n - \lfloor \alpha N \rfloor + 1}{\lfloor \alpha N \rfloor}, \quad n = 0, 1, ..., \lfloor \alpha N \rfloor - 1. \quad (8)$$

[1]It is currently challenging to design the grid G according to (7) instead of (4), but, based on our analysis, we propose to keep designing the grid according to (4) and use (7) as an alternative metric for parameters tuning, such as the distance sampling factor β.

The distance domain is again sampled as

$$\rho' = \frac{1}{s}\frac{1}{2\lambda}\left(\frac{N\delta}{\beta}\right)^2 \left[1 - (\Phi')^2\right] \quad s = 1, 2, 3.... \quad (9)$$

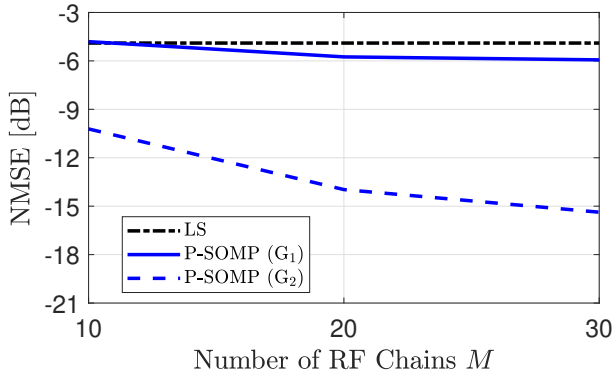
Values of the parameter s that result in ρ' falling outside the range $[\rho_{\min}, \rho_{\max}]$ or points outside the angular sector $[\varphi_{\min}, \varphi_{\max}]$ are excluded. The design in Sec. III-A is a special case of the proposed method, with $\alpha = 1$.

Fig. 3 shows NMSE$_{\text{opt}}$ and Q as a function of α and β, in the same conditions of Fig. 2. With no-oversampling ($\alpha \leq 1$), NMSE$_{\text{opt}} \geq -6.4$ dB for any value of β. However, by oversampling the angular domain ($\alpha > 1$), we can find better trade-off pairs between angular and distance sampling factors (β, α) that improve NMSE$_{\text{opt}}$ without increasing the computational complexity. For example, the curve level with $Q = 1500$ shows that, despite corresponding to grids of 1500 points each, the pair $(\beta, \alpha) = (0.674, 3)$ gives an NMSE$_{\text{opt}} = -12.4$ dB, while the pair $(\beta, \alpha) = (0.384, 1)$ only achieves an NMSE$_{\text{opt}} = -6.4$ dB.

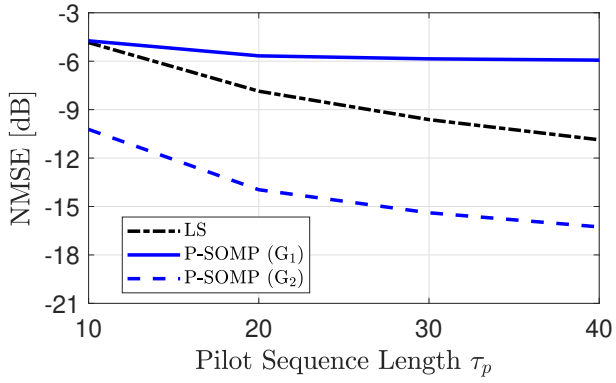
Additionally, either decreasing β or increasing α, NMSE$_{\text{opt}}$ reduces, but the grid size also increases, leading to higher computational complexity. The pair $(\beta, \alpha) = (0.576, 7.5)$ (indicated by the red circle) represents a favorable balance between channel estimation accuracy and complexity. In particular, we denote $\mathbf{G}_2$ the corresponding grid, which yields NMSE$_{\text{opt}} \approx -21$ dB and $Q_2 = 5182$. The analysis in [9] shows that the computational complexity of P-SOMP linearly increases with Q. For example, since $Q_2 \approx 10\,Q_1$, using $\mathbf{G}_2$ instead of $\mathbf{G}_1$ for P-SOMP operations increases its computational complexity by a factor of 10. However, using larger grids does not require additional system resources, such as the number of pilots, which remains independent of Q. Hence, larger grids do not increase communication overhead.

IV. Performance Evaluation

We now evaluate the P-SOMP estimation accuracy with parameters reported in Tab. I, with a specific focus on the effects of pilot length τ_p, number of RF chains M, and grid selection. As a benchmark, we report the estimation accuracy of a fully digital system with $M = N$, considering LS for channel estimation and P-SOMP for hybrid channel estimation. The impact on SE is also quantified.



(a) $\tau_p = 10$.



(b) $M = 10$.

Fig. 4: NMSE against the number of RF chains M and pilot sequence length τ_p.

A. Impact of Radio Frequency Chains

We begin by investigating the impact of M, assuming that $\tau_p = 10$. Fig. 4a shows NMSE of channel estimates as a function of M. Since hybrid architectures are meant to reduce the number of RF chains relative to the number of antennas, we limit our analysis to $M \leq 30$. We observe that P-SOMP estimation improves as M increases. In fact, the hybrid observation vector expands as the number of RF chains grows, enabling P-SOMP to select better grid points and thereby improving channel estimation accuracy. However, P-SOMP performance is limited by the grid. Indeed, when M becomes sufficiently large, P-SOMP consistently selects the optimal grid point, and NMSE stabilizes. The larger the grid, the higher M needs to be for this to happen. For example, NMSE stops improving with $M = 20$ using $\mathbf{G}_1$, while NMSE stabilizes out of the considered range using $\mathbf{G}_2$.

The comparison between hybrid and fully digital estimators shows that P-SOMP performs slightly better than LS using the lower complexity grid $\mathbf{G}_1$. Notably, using the higher complexity grid $\mathbf{G}_2$ results in a much lower NMSE compared to $\mathbf{G}_1$, but the price to pay is an $\frac{Q_2}{Q_1} = 10$ times larger computational complexity. However, P-SOMP outperforms the LS estimator because the latter only achieves an NMSE $= -5$ dB owing to $\tau_p = 10$, so the pilot sequence is relatively short. Therefore, we now evaluate the impact of τ_p.

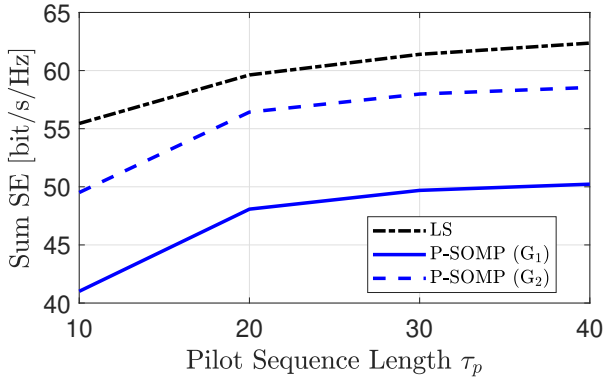


Fig. 5: Average sum SE with MMSE combining as a function of pilot sequence length τ_p with $M = K = 10, \tau_c = 10000$.

B. Impact of the Pilot Sequence Length

We now analyze the impact of τ_p on NMSE for $M = 10$. Increasing τ_p leads to higher SNR for fully digital estimation and proportionally enlarges the observation vector for hybrid estimation. Therefore, the accuracy of all methods improve with higher τ_p. However, Fig. 4b shows that the P-SOMP accuracy is limited by the grid once again. In fact, when τ_p becomes large enough so that P-SOMP consistently selects the best point in the grid for channel estimation, its accuracy stops improving as τ_p increases. The larger the grid, the higher τ_p needs to be before NMSE stabilizes.

Results show that LS performs better than P-SOMP using the lower complexity grid $\mathbf{G}_1$ because the latter stabilizes too quickly and does not improve with $\tau_p \geq 20$. However, we also see that P-SOMP using $\mathbf{G}_2$ performs better than LS in the considered range, and its accuracy keeps improving as τ_p increases. In fact, since $\mathbf{G}_2$ is 10 times larger than $\mathbf{G}_1$, the pilot sequence length after which NMSE stabilizes falls out of the considered range. The price to pay is a 10 times larger computational complexity compared to $\mathbf{G}_1$. Summarizing, LS eventually outperforms P-SOMP using long enough pilot sequences regardless of the grid size. However, if we can afford enough computational complexity, P-SOMP is more accurate than LS with short pilot sequences.

Furthermore, the comparison between Fig. 4a and 4b shows that the impact of M and τ_p on hybrid channel estimation is equivalent. In fact, we observe that the P-SOMP accuracy is approximately the same as long as the hybrid observation vector size $M\tau_p$ stays constant. This suggests that we can potentially compensate for the reduced number of RF chains in hybrid systems by increasing the number of transmitted pilot symbols. Moreover, it suggests that NMSE with P-SOMP stabilizes when $M\tau_p$ is large enough, rather than when M or τ_p are individually high enough.

C. Spectral efficiency evaluation

Channel estimation is the ancillary task to support data detection. Therefore, we evaluate its impact on the achievable uplink SE. This is obtained with the well-known use-and-then-forget bound [11, Sect. 4.2] for the arbitrary user k, which

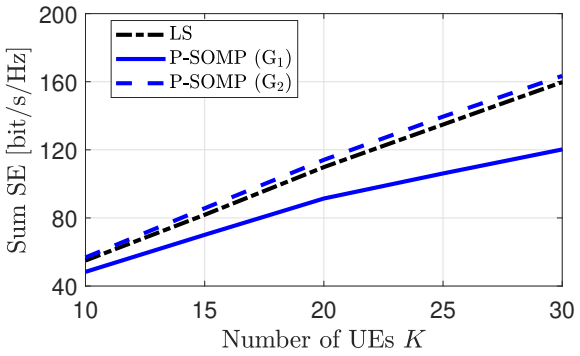


Fig. 6: Average sum SE as a function of the number of UEs K with $M = K$ and $\tau_c = 10000$ and MMSE combining.

yields $\mathsf{SE}_k = (1 - \tau/\tau_c)\,\mathbb{E}\left\{\log_2\left(1 + \gamma_k\right)\right\}$ where the factor τ/τ_c accounts for the overhead loss, and γ_k is given by

$$\frac{\left|\mathbb{E}\left\{\mathbf{v}_k^{\mathsf{H}}\mathbf{h}_k\right\}\right|^2}{\sum_{i=1}^{K}\mathbb{E}\left\{\left|\mathbf{v}_k^{\mathsf{H}}\mathbf{h}_i\right|^2\right\} - \left|\mathbb{E}\left\{\mathbf{v}_k^{\mathsf{H}}\mathbf{h}_k\right\}\right|^2 + \frac{\sigma^2}{\rho}\mathbb{E}\left\{\left\|\mathbf{v}_k\right\|^2\right\}} \tag{10}$$

with $\mathbf{v}_k \in \mathbb{C}^N$ being the receive combining vector of UE k.

We consider a low-mobility scenario with $K = 10$ UEs, where each UE moves with speed $v = 1$ m/s, and compute the coherence time as $\frac{\lambda}{10v}$ [11]. Additionally, we assume that the coherence bandwidth equals B. Therefore, the coherence block size is given by $\tau_c = \frac{B\lambda}{10v} = 10000$. We consider the minimum-mean-square-error (MMSE) combining [11].

1) Impact of Pilot Sequence Length: Fig. 5 shows SE as a function of the pilot sequence length τ_p. We observe that the fully digital system with LS estimator achieves the best SE, but the gap with respect to the hybrid one reduces when using P-SOMP with the higher-complexity grid $\mathbf{G}_2$. The comparison of Fig. 4b and 5 remarks that, from a communicational point of view, it can be more relevant to evaluate the SE instead of the NMSE of channel estimates. In fact, despite the NMSE being significantly lower using P-SOMP with $\mathbf{G}_2$ than using LS, the latter achieves the largest SE.

2) Impact of UEs: We now evaluate the impact of the number of UEs (K) on the SE. We assume $M = K$ because each UE must be paired to at least one RF chain for data detection. However, since hybrid architectures are designed to reduce the number of RF chains, we do not consider $K > 30$. Additionally, we assume $\tau_p = K$ for fully digital estimation and $\tau_p = 20$ for hybrid estimation. Fig. 6 shows sum SE as a function of the number of UEs K. SE for LS and P-SOMP with $\mathbf{G}_2$ grows linearly, which means that effective interference cancellation is possible in the whole considered range. However, when using P-SOMP with $\mathbf{G}_1$, SE linearly grows up to $K = 20$ UEs, but channel estimation becomes not accurate enough to allow for effective interference cancellation with $K > 20$, so SE grows more slowly afterwards.

Notably, we observe that an aggregate SE of 50 bit/s/Hz can be achieved with 10 UEs. Assuming a total bandwidth of 20 GHz split into 200 sub-bands of 100 MHz each; 10 active UEs per sub-band and an average aggregate SE of 50 bit/s/Hz per sub-band, we can theoretically achieve a total aggregate throughput of 1 Tbps (i.e., $200 \times 50 = 1$ Tbps). However, this requires the BS to receive a total power of 30 Watt and 2000 UEs to transmit simultaneously, both of which are impractical with current technology.

V. Conclusions

We investigated channel estimation for a 300 GHz hybrid U-MIMO system using the P-SOMP algorithm [9] and compared it to a fully digital system using LS channel estimation. Our analysis focused on channel estimation accuracy and spectral efficiency, examining the effects of RF chains number (M) and pilot sequence length (τ_p). We found that increasing pilot symbols can offset the reduced RF chains, preserving accuracy and efficiency. Our study also highlighted that current dictionary designs are inadequate for achieving optimal MSE. Numerical results demonstrated that a 300 GHz hybrid U-MIMO system can attain comparable spectral efficiency to a fully digital system, but only when very large dictionaries with angular oversampling are utilized. This underscores the need for new dictionary design criteria, moving beyond the traditional correlation-based metrics used in far-field communications.

Acknowledgment

This work has been performed in the framework of the HORIZON-JU-SNS-2022 project TIMES, cofunded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union.

References

[1] Björnson *et al.*, "Massive MIMO is a reality—what is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 94, 2019.

[2] C. Castro *et al.*, "Long-range high-speed thz-wireless transmission in the 300 ghz band," in *Third International Workshop on Mobile Terahertz Systems (IWMTS)*, 2020.

[3] K. C. Huang and Z. Wang, "Terahertz terabit wireless communication," *IEEE Microwave Magazine*, vol. 12, no. 4, 2011.

[4] T. S. Rappaport *et al.*, "Wireless communications and applications above 100 ghz: Opportunities and challenges for 6g and beyond," *IEEE access*, vol. 7, 2019.

[5] C. Han, L. Yan, and J. Yuan, "Hybrid beamforming for terahertz wireless communications: Challenges, architectures, and open problems," *IEEE Wireless Communications*, vol. 28, no. 4, 2021.

[6] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, 2016.

[7] G. Bacci, L. Sanguinetti, and E. Björnson, "Spherical wavefronts improve MU-MIMO spectral efficiency when using electrically large arrays," *IEEE Wireless Communications Letters*, vol. 12, no. 7, 2023.

[8] E. Björnson, Ö. T. Demir, and L. Sanguinetti, "A primer on near-field beamforming for arrays and reconfigurable intelligent surfaces," in *IEEE Asilomar*, 2021.

[9] M. Cui and L. Dai, "Channel estimation for extremely large-scale mimo: Far-field or near-field?" *IEEE Trans. Commun.*, vol. 70, no. 4, 2022.

[10] Ö. T. Demir and E. Björnson, "A new polar-domain dictionary design for the near-field region of extremely large aperture arrays," in *Proceedings of IEEE CAMSAP*, 2023.

[11] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, 2017.

[12] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, 2010.