

Robust cell-free mmWave/sub-THz access using minimal coordination and coarse synchronization

Lorenzo Miretti Member, IEEE, Giuseppe Caire Fellow, IEEE, Sławomir Stańczak Senior Member, IEEE

Abstract—This study investigates simple alternatives to coherent joint transmission for supporting robust connectivity against signal blockage in mmWave/sub-THz access networks. By taking an information-theoretic viewpoint, we demonstrate analytically that with careful design, full macrodiversity gains can be significantly achieved through canonical receiver and channel coordination and synchronization requirements on the infrastructure side. Our proposed scheme extends non-coherent joint transmission by employing a specific form of diversity to combat artificially induced fading that would otherwise make this technique often less favorable against standard throughput selection schemes. Additionally, the inclusion of distributed link-space-time coding day files shows in recovering significant fraction of the optimal performance. Our conclusions are based on one in eight multi-point interface link block fading channel model that enables bit-for-bit engineering and outage rate analysis, while also considering timing offset and its imperfections, compensation of the high simplified approach capturing the features of modern mmWave/sub-THz communications, thereby providing practical design guidelines for realistic systems.

Index Terms—mmWave, sub-THz, cell-free, non-coherent joint transmission, space-time coding, macrodiversity, synchronization, joint transmission, space-time coding, macrodiversity, synchronization.

I. INTRODUCTION

I. Introduction

A well-known major problem in mobile access networks operating at very high carrier frequencies, such as in the upper mmWave or sub-THz bands, is that they are highly sensitive to signal blockage [2], [21], [2]. The reason is due to unavoidable sensitivity to signal blockage [1], [11], [1]. The reason is due physical characteristics of the propagation medium as well as to unavoidable physical characteristics of the propagation expected mode of operation, which relies on highly directional as well as the expected mode of operation, which relies on highly directional line-of-sight transmission in the noise limited regime on highly directional line-of-sight transmission in the [2], [21], [2]. This sensitivity causes connection instability and noise limited regime [1], [11], [1]. This sensitivity causes severely degrades the overall system performance.

In order to mitigate the signal blockage and provide robust system performance,

In order to mitigate the signal blockage and provide connectivity, several multi-connectivity concepts are advocated by academia, industry, and standardization bodies [2], [21], [2], robust connectivity, several multi-connectivity concepts [2]. In fact, a similar trend is also followed in the context are advocated by academia, industry, and standardization of visible-light communication [2], where signal blockage is an evident issue. Essentially, all these concepts attempt to follow in the context of visible-light communication [2], capitalize on the so-called *macrodiversity* gains offered by, where signal blockage is an evident issue. Essentially, simultaneous and possibly coordinated connections to multiple access points. The proposed technologies range from relatively macrodiversity gains offered by simultaneous and possibly simple control plane approaches for fast and efficient access

L. Miretti, G. Caire, and S. Stańczak are with the Technische Universität Berlin, Berlin 10587, Germany, with the Institute für Informationstechnik, Berlin 10587, Germany, and S. Stańczak is also with the Fraunhofer Institute for Signal and Information Processing IIS, Berlin 10587, Germany.

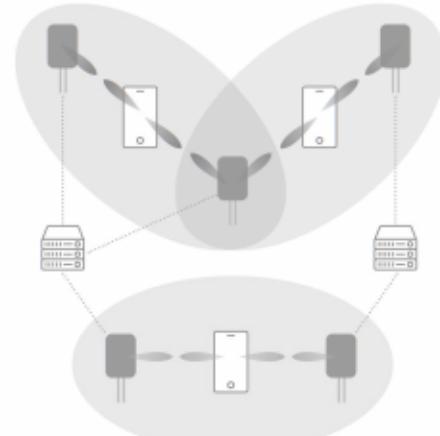


Fig. 1. Network of multiple stations connected to a central base station via multiple received multiple high-rate and fair high-rate mmWave/sub-THz links. The network is shown in three clusters of stations, each with a central base station and multiple antennas. Dashed lines represent direct links between stations and base stations, while solid lines represent relayed links through other stations. The diagram illustrates the concept of joint transmission over multiple points, where a single base station serves multiple stations simultaneously, and stations can relay signals to each other. The network is designed to be robust against signal blockage, as it can switch to alternative paths if one is blocked. The diagram also shows the use of multiple antennas at both the base stations and the stations to achieve high rates and fairness. The network is shown in three clusters of stations, each with a central base station and multiple antennas. Dashed lines represent direct links between stations and base stations, while solid lines represent relayed links through other stations. The diagram illustrates the concept of joint transmission over multiple points, where a single base station serves multiple stations simultaneously, and stations can relay signals to each other. The network is designed to be robust against signal blockage, as it can switch to alternative paths if one is blocked. The diagram also shows the use of multiple antennas at both the base stations and the stations to achieve high rates and fairness.

(theoretical) As up to will total number of transmits are typically introduced as idle gaps used to be, quiet selected code words. Practically, in [7] and in this study, this problem that may be even more difficult to solve is performed in one. However, for [7] it does with the simplest alternatives of shortening this in the configuration $\{3, 1\}, \{2, 1\}$, $\{2, 2\}$. Does not this suggest consideration of the optimal gain using statistically independent with practical transmission of possibly very dangerous signals (where the number of the total (overall frequency chains) in this case is given by the number of transmitted data streams). As shown in processing in a short study [8] this is the data streams. Furthermore, [8] also shows that the performance of the proposed scheme is comparable to the simple adaptive schemes using as the technique Fourier sampler [8]. The discussion of these aspects are achieved due to the generalizing extending digamma algorithm with the fractional unitary system especially complex topologies regime of primary the number of antennas (or radio frequency chains) at the receiver is smaller than the number of transmitted data streams, i.e., where C . Contributions of this study include the following:

C. Contributions of this study

linear processing cannot demultiplex the data streams. In this study, we give the first comprehensive investigation of alternative techniques to coherent joint transmission of the proposed technique against other simpler alternatives tailored to mmWave/sub-THz network MIMO systems. In such as the ones based on transmitter selection. These particular, we review and extend several available techniques aspects are crucial because, given the extremely high and compare them, in terms of macrodiversity gains, SNR rates of mmWave/sub-THz communication systems, low-gains, complexity, and signaling overhead. In contrast to prior complexity processing, is of primary importance.

studies, we adopt an information theoretic viewpoint and, starting from first principles, we study the problem of reliable communication over a multi-point intermittent block

communications of a simple multi-point *intermittent* block fading channel. Despite being considerably simplified, we

fading channel. Despite being considerably simplified, we stress that the proposed model still captures the most essential features of alternative techniques to coherent joint transmission tailored to mmWave/sub-THz systems such as line-of-sight transmission in the noise limited regime and sensitivity to signal blockages. In particular, we review and extend several available techniques and compare them in terms of macrodiversity

Our main result is the identification of a family of macrodiversity gains, SNR gains, complexity, and signaling overhead reduction techniques that extends non-coherent joint transmission as In contrast to prior studies, we adopt an information theoretic viewpoint and, starting from first principles, significantly lower complexity than competing techniques. we study the problem of reliable communication over More precisely, we demonstrate analytically that with a careful simple multi-point intermittent block fading channel design, full macrodiversity gains and significant SNR gains Despite being considerably simplified, we stress that the can be achieved through the use of linear receiver processing proposed model still captures the most essential features standard scalar point-to-point coding, minimal coordination of the envisioned mmWave/sub-THz systems, such as across the transmitters, and coarse synchronization. To this line-of-sight transmission in the noise limited regime and end, a first contribution is the study of a special form of sensitivity to signal blockages. diversity, called phase diversity [2], which is particularly

Our main result is the identification of a family of suitable transmission techniques that extends non-coherent joint model and makes the effective channel fluctuations essentially transmission as a promising candidate for supporting driven by blockage events only. A second contribution is the burst transmission with significantly lower complexity than study of a simple, space-time coding layer that extends the competing techniques. More precisely, we demonstrate well-known Alamouti scheme [?] to provide an additional analytically that with a careful design, full macrodiversity performance boost. Finally, a third contribution is the study gains and significant SNR gains can be achieved through of the impact of unknown timing offsets based on a worst- the use of linear receiver processing, standard scalar case approach similar to the information-theoretical literature point-to-point coding, minimal coordination across the on asynchronous [?] or arbitrarily varying [?] channels. transmitters, and coarse synchronization. To this end, a first contribution is the study of a special form of diversity, called phase diversity [?], which is particularly suitable for

study of a simple example of the fading layer that extends the design guidelines in Section ?? in how they provide the obtained insights to implement. Finally, the motivation and definition of the study of the impact of unknown results on its based on the promising approach described to the information theoretical literature. In this section, we present the details. [3] denote arbitrary vectors and boldfaced upper case letters, A , denote matrices. The n th entry of a is denoted by a_n . We use $\hat{\cdot}$ for definitions. We denote D , \mathbf{Q}^T , \mathbf{Q}^H the transpose and Hermitian transpose of a , respectively. We use $\Pr(\cdot)$ for the probability of an event, and $E[\cdot]$ for the expectation operator. Given a stationary random process $\{a_t\}_{t \in \mathbb{Z}}$, we denote an arbitrary realization by a .

Section ?? studies and extends several suboptimal transmission schemes for the considered channel, and provides practical design guidelines. Section ?? shows how to apply the insights presented in this section to the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

B. Outline and notation In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions. Before going into the details, we also provide an informal description of the overall mmWave/sub-THz system architecture that serves as motivation for the proposed channel model. We use $\hat{\cdot}$ for definitions. We denote by a^T and a^H the transpose and Hermitian transpose of a , respectively. We use $\Pr(\cdot)$ for the probability of an event, and $E[\cdot]$ for the expectation operator.

We consider the downlink of a mmWave/sub-THz network composed of multiple transmitters that jointly serve multiple receivers in the same time-frequency resource, possibly in a user-centric cell-free fashion (Figure ??). We assume a hybrid beamforming architecture in which each transmitter is equipped with a small number of digital antenna ports (DAPs) or receive ports. In this section, we present and study the fundamental limits of the proposed multi-point intermittent block fading channel model that forms the basis of our analysis. We further assume a hybrid beamforming architecture in which each transmitter is equipped with a small number of digital antenna ports (DAPs) or receive ports. Although more advanced receivers may be equipped with multiple digital antenna ports, the single receiver port case is an insightful example of the likely scenario where the number of receive ports is much smaller than the total.

We consider the downlink of a mmWave/sub-THz network composed of multiple transmitters that jointly serve multiple receivers in the same time-frequency resource. As customary, line-of-sight directional transmission is employed to counteract the large path-loss at mmWave/sub-THz frequencies. Since sub-THz networks are typically noise limited, multi-user interference is not a primary issue, provided that the scheduled receivers are sufficiently separated in the spatial domain. Therefore, we omit digital precoding and assume that each digital antenna port is allocated for transmission to a single receiver. The main advantage of the considered architecture is that each transmit digital antenna port need only to maintain coarse mutual synchronization with other transmit ports. We remark that our model is technology agnostic in that we do not specify the underlying beamforming implementation. In fact, an insightful example of the likely scenario where the number of receive ports is much smaller than the total number of transmit ports. As customary, line-of-sight directional transmission is employed to counteract the large path-loss at mmWave/sub-THz frequencies. Since sub-THz networks are typically noise limited, multi-user interference is a problem primarily due to the fact that the subpath differences between users are sufficiently represented in the spatial

domain. A simple example of the fading layer that extends the design guidelines in Section ?? in how they provide the obtained insights to implement. Finally, the motivation and definition of the study of the impact of unknown results on its based on the promising approach described to the information theoretical literature.

C. System architecture [3] denote arbitrary vectors and boldfaced upper case letters, A , denote matrices. The n th entry of a is denoted by a_n . We use $\hat{\cdot}$ for definitions. We denote D , \mathbf{Q}^T , \mathbf{Q}^H the transpose and Hermitian transpose of a , respectively. We use $\Pr(\cdot)$ for the probability of an event, and $E[\cdot]$ for the expectation operator. Given a stationary random process $\{a_t\}_{t \in \mathbb{Z}}$, we denote an arbitrary realization by a .

Section ?? studies and extends several suboptimal transmission schemes for the considered channel, and provides practical design guidelines. Section ?? shows how to apply the insights presented in this section to the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

D. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

E. Outline and notation In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

F. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

G. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

H. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

I. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

J. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

K. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

L. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

M. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

N. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

O. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

P. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

Q. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

R. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

S. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

T. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

U. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

V. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

W. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

X. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

Y. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

Z. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

A. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

B. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

C. Channel model In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

D. System architecture In this section, we present and study the fundamental limits and delay compensation. Section ?? summarizes the results and discusses future promising research directions.

system performance in terms of the insights obtained with blockage probability approach to the realistic system. Our primary focus is to study such insights in time blockage-practical path loss variations induced by blockage events are often so large that the useful signal power becomes negligible in the noise limited regime.

The assumption of independence between $\{\beta_t\}_{t \in \mathbb{Z}} = \{(\beta_{1,t}, \dots, \beta_{L,t})\}_{t \in \mathbb{Z}}$ is reasonable when blockages are caused by events close to the transmitters. Multi-state models and/or correlated blockages may also be investigated, but such analysis is left for future work.

We use random phase shifts realization of β . Since the joint blockage pattern is governed by channel binary Markov chains, the lack of phase synchronization is readily given in terms of the above quantities as $\Pr[\beta = (b_1, \dots, b_L)] = \prod_{l=1}^L \Pr[\beta_l = b_l]$. We also define the number of non-blocked transmitters $\alpha_t := \sum_{l=1}^L \beta_{l,t}$ and observe that drifts $\beta_{l,t+1} - \beta_{l,t}$ for $l = 1, \dots, L$. We readily have that $\Pr[\alpha_t = 1] = \Pr[\beta_l = 0]$ and delay which compensates one key benefit of macro diversity, realistic the probability of a late receiver experiencing complete blockage decreases with L .

also to more realistic systems where the transmitters can only achieve coarse time synchronization and

C. Capacity and optimal transmission

In this section we study the ergodic and outage capacity of the considered channel, defined as follows. We first assume perfect channel state information at the receiver (CSIR) (β, θ) , centralized partial channel state information at the transmitters (CSIT) β , and an instantaneous power constraint $P \in \mathbb{R}_+$ per transmitter. In absence of latency constraints, standard arguments show that a well-defined ergodic capacity in the classical Shannon sense exists [2] and can be expressed in bits/symbol. Blockage process is governed by L i.i.d. binary Markov chains; its step Energy distribution $Q(\beta, h)$ is readily given in terms of the above quantities as $\Pr[\beta = (b_1, \dots, b_L)] = \prod_{l \in \mathcal{Q}} \Pr[\beta_l = b_l]$, where \mathcal{Q} denotes the set of functions mapping each CSIT realization β to a complex-valued symmetric positive semidefinite matrix $Q(\beta)_l$ of size $L \times L$, with diagonal entries satisfying $\alpha \sim \text{Binomial}(L, 1 - p_B)$. We readily have that $\Pr[\alpha \geq 1] = \Pr[\beta \neq 0] = 1 - p_B$, which formalizes one key benefit of macrodiversity; that is, the probability of a receiver to experience complete blockage decreases with L .

$$\begin{bmatrix} x_1[m] \\ \vdots \\ x_L[m] \end{bmatrix} = Q(\beta_l)^{\frac{1}{2}} u[m], \quad t = \left\lfloor \frac{m}{T} \right\rfloor, \quad (1)$$

where this section we study the ergodic and outage capacity of the considered channel, defined as follows. We first assume perfect channel state information at the receiver (CSIR) (β, θ) , centralized partial channel state information assuming partial CSIT β and an instantaneous power constraint $P \in \mathbb{R}_+$ per transmitter is given by power constraint $P \in \mathbb{R}_+$ per transmitter. In absence of latency constraints, standard arguments show that a well-defined ergodic capacity in the classical Shannon sense exists [2], and it can be expressed (in bits/symbol) as

Proof. By the law of total expectation and Jensen's inequality, we obtain the upper bound $\mathbb{E}[\log(1 + h^\top Q(\beta)h)]$,

$$\mathbb{E}[\log(1 + h^\top Q(\beta)h)] \leq \mathbb{E}[\log(1 + \mathbb{E}[h^\top Q(\beta)h])],$$

where \mathcal{Q} denotes the set of functions mapping each CSIT realization β to a complex-valued symmetric positive semidefinite matrix $Q(\beta)$ of size $L \times L$, with diagonal entries satisfying

$$(\forall l \in \{1, \dots, L\}) \quad [Q(\beta)_l \log \left(P + \sum_{i=1}^L \alpha_i(\beta_l) \right)]$$

For a given $Q \in \mathcal{Q}$, the rate $R = \mathbb{E}[\log(1 + h^\top Q(\beta)h)]$ is achievable by vector Gaussian codes with conditional

outputs $Q(\beta) = Q(\beta)$ almost surely (this concludes the proof). \square

$$\begin{bmatrix} x_1[m] \\ \vdots \\ x_L[m] \end{bmatrix}$$

Corollary 1. The proof of Proposition ?? also shows that the ergodic capacity of the considered channel assuming no CSIT and an instantaneous power constraint $P \in \mathbb{R}_+$ per transmitter bearing signal is given by

$$C = \mathbb{E}[\log(1 + \alpha P)].$$

Proposition 1. The ergodic capacity of the considered channel assuming knowledge of the blockage process β at the transmitters does not increase capacity.

The above analysis shows the ergodic capacity (2) where α is achieved by a Binomial($L, 1 - p_B$) scheme in $(?)$ with fixed diagonal input covariance independent of β . Note that this is essentially the same transmission scheme considered in [?], left apart a power allocation step. However, although

$$\mathbb{E}[\log(1 + h^\top Q(\beta)h)] \leq \mathbb{E}[\log(1 + \mathbb{E}[h^\top Q(\beta)h])]$$

optimal from an information-theoretic viewpoint, this scheme presents several practical drawbacks. First and foremost, being equipped with a single antenna, the receiver cannot demultiplex the vector-valued information bearing signal using linear processing, and hence it must perform complex algorithms such as maximum-likelihood vector decoding or approximate versions based, e.g., on successive interference cancellation.

Second, although $(?)$ can be theoretically achieved without CSIT, this would require codewords spanning a very large number of blockage realizations, hence largely exceeding practicality. The constraints of the Proposition induce a less practical system, which should typically be the block-coded channel maximizing CSIT. By anticipating an hour argument of constraint $(?)$ per transmitter, the given achievable rate in a given block-fading realization, the ergodic capacity C can be then approached by means of rate adaptation mechanisms based on CSIT. Note that an intermediate solution in terms of latency is to use hybrid automatic retransmission-request (H-ARQ).

The above analysis shows that the ergodic capacity $(?)$ can be achieved by the transmission scheme in $(?)$ with fixed diagonal input covariance independent of β . Note that this is the same transmission scheme considered in [?], left apart a power allocation step. However,

although optimal from an information-theoretic viewpoint, this scheme presents several practical drawbacks. First and foremost, being equipped with a single antenna, the receiver cannot demultiplex the vector-valued information bearing signal using linear processing, and hence it must

perform complex algorithms such as maximum-likelihood vector decoding or approximate versions based, e.g., on successive interference cancellation. Second, although $(?)$

can be theoretically achieved without CSIT, this would require codewords spanning a very large number of block-

age realizations, which is impractical due to the presence of strict latency constraints and when latency adaptation mechanisms based on CSIT are employed. Two cases of study are considered: (i) a flat fading channel model of a block fading channel, and (ii) a block fading channel model of a block fading channel.

For a given $Q \in \mathcal{Q}$, the rate $R = \mathbb{E}[\log(1 + \mathbb{E}[h^\top Q(\beta)h])]$ is achievable by vector Gaussian codes with conditional

outputs $Q(\beta) = Q(\beta)$ almost surely (this concludes the proof). \square

Corollary 1. The proof of Proposition ?? also shows that the ergodic capacity of the considered channel assuming no CSIT and an instantaneous power constraint $P \in \mathbb{R}_+$ per transmitter bearing signal is given by

$$C = \mathbb{E}[\log(1 + \alpha P)].$$

Proposition 1. The ergodic capacity of the considered channel assuming knowledge of the blockage process β at the transmitters does not increase capacity.

The above analysis shows that the ergodic capacity (2) where α is achieved by a Binomial($L, 1 - p_B$) scheme in $(?)$ with fixed diagonal input covariance independent of β . Note that this is essentially the same transmission scheme considered in [?], left apart a power allocation step. However,

although optimal from an information-theoretic viewpoint, this scheme presents several practical drawbacks. First and foremost, being equipped with a single antenna, the receiver cannot demultiplex the vector-valued information bearing signal using linear processing, and hence it must

perform complex algorithms such as maximum-likelihood vector decoding or approximate versions based, e.g., on successive interference cancellation. Second, although $(?)$

realization, the ergodic capacity can be then approached by means of rate adaptation mechanisms based on CSIT. Note that an intermediate solution in terms of latency is to use hybrid automatic retransmission-request (H-ARQ) mechanisms [?], which also require some form of CSIT.

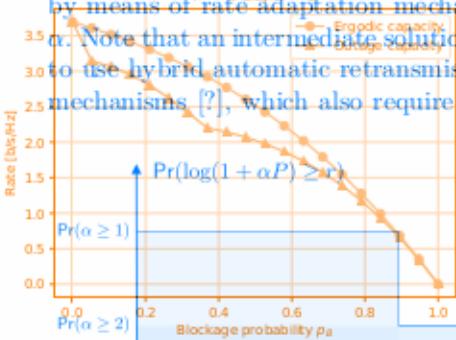


Fig. 3. Ergodic and outage capacity versus blockage probability, SNR, and number of transmitters.

Proposition 2. The outage capacity of the considered channel with power constraint $P \in \mathbb{R}_+$ per transmitter is given by

As a more appropriate performance metric in the presence of strict latency constraints and when rate adaption mechanisms based on CSIT are prohibitive, we *Proof.* The proof readily follows by observing that now study the outage capacity of the considered channel. In particular, by considering fixed-rate block codes of maximum length T we measure capacity as

$$C_{\text{out}} := \sup_{r \in \mathbb{R}} r \cdot \Pr(\log(1 + \alpha P) \geq r).$$

Fig. 2. Pictorial representation of the CCDF of the instantaneous capacity.

Proposition 2. The outage capacity of the considered channel with power constraint $P \in \mathbb{R}_+$ per transmitter is given by

As a more appropriate performance metric in the presence of strict latency constraints and when rate adaption mechanisms based on CSIT are prohibitive, we *Proof.* The proof readily follows by observing that now study the outage capacity of the considered channel. In particular, by considering fixed-rate block codes of maximum length T we measure capacity as

$$C_{\text{out}} := \sup_{r \in \mathbb{R}} r \cdot \Pr(\log(1 + \alpha P) \geq r).$$

The quantitative difference between the ergodic and outage capacity can be visualized in Figure ??, which depicts the complementary cumulative density function (CCDF) of the instantaneous capacity $\log(1 + \alpha P)$. The ergodic capacity is given by the total area under the CCDF, while the outage capacity is given by the largest rectangle. The two metrics are compared in terms of different parameters in Figure ??.

III. PRACTICAL TRANSMISSION SCHEMES

In this section we revisit, compare, and extend a set of practical transmission schemes that do not require the receiver to perform complex decoding algorithms. We will measure the complementary cumulative density function (CCDF) of the instantaneous capacity $\log(1 + \alpha P)$. The ergodic capacity is given by the total area underlying the CCDF, while the outage capacity is given by the largest rectangle. The two metrics are also compared in terms of different parameters in Figure ??.

As first main baseline, we consider a transmission scheme where, for all block fading realizations with at least one non-blocked transmitters, i.e., such that $\alpha > 0$, the network chooses one non-blocked transmitter $l^*(\beta_t)$ and lets ($\forall m \in \mathbb{Z}$)

of practical transmission schemes that do not require the receiver to perform complex decoding algorithms. We will measure performance in terms of ergodic and outage rates, depending on the context and applicability, and discuss crucial implementation aspects. By mapping this scheme to some $Q \in \mathcal{Q}$ in the signal model (??), it can be readily seen that the ergodic rate

As first main baseline we consider a transmission scheme where, for all block fading realizations with at least one non-blocked transmitters, Note that the block fading realization with at least one non-blocked transmitter has the outage such that known

the network chooses one non-blocked transmitter $l^*(\beta_t)$ and lets ($\forall m \in \mathbb{Z}$) ($\forall l \in \{1, \dots, L\}$)

$$x_l[m] = \begin{cases} u[m] & \text{if } l = l^*(\beta_t) \text{ and } \alpha > 0, \\ 0 & \text{otherwise,} \end{cases} \quad t = \left[\frac{m}{T} \right],$$

where $u[m] \sim \mathcal{CN}(0, P)$ is a scalar i.i.d. information bearing signal. By mapping this scheme to some $Q \in \mathcal{Q}$ in the signal model (??), it can be readily seen that the ergodic rate

$$R = (1 - p_B^L) \log(1 + P)$$

is achievable. Note that the concept of outage rate is not meaningful for this scheme, since the outage events are known by the network. The main advantage of transmitter selection is that it can mitigate the effect of blockage using simple scalar codes at fixed rate $\log(1 + P)$. However, this simplicity comes with two major drawbacks. First, that it can mitigate the effect of blockage using simple scalar codes at fixed rate $\log(1 + P)$. However, this simplicity comes with two major drawbacks. First, no SNR gain is provided, i.e., R saturates to $\log(1 + P)$ with two major drawbacks. Second, causal knowledge of β_t at the transmitters is assumed. Alternatively, the choice of the transmitter can be performed by the receiver, but this knowledge of β_t at the transmitters is assumed. Alternatively, the choice of the transmitter can be performed by the receiver, the network. Overall, the required resources for timely blockage estimation and network coordination may be significant, especially in case of frequent blockages.

B. Non-coherent joint transmission

As second main baseline, we consider non-coherent joint transmission of a single scalar codeword from all transmitters. As a third main baseline, we let (consider) non-coherent joint transmission of a single scalar codeword from all transmitters simultaneously, i.e., we let ($\forall m \in \mathbb{Z}$) ($\forall l \in \{1, \dots, L\}$)

where $u[m] = u[m]$ is a scalar i.i.d. information bearing signal. Using (??), we observe that this scheme achieves the ergodic rate

$$\Pr(u[m] \sim \mathcal{CN}(0, P))$$

where $\Pr(u[m] \sim \mathcal{CN}(0, P))$ denotes an effective small-scale fading coefficient which is artificially induced by the non-coherent superposition of signals at the receiver.

Non-coherent joint transmission always outperforms coherent superposition of signals in terms of ergodic rates, and it provides SNR gains as L grows. This trend is clearly visible in Figure ??.

Non-coherent joint transmission always outperforms linear selection in terms of ergodic rates, and it provides SNR gains as L grows. This trend is clearly visible in Figure ??.

For a general proof, we exploit the following iterative property, which essentially states that, on average, adding non-blocked transmitters is beneficial:

Proposition 3. Let $\{\theta_i\}_{i=0}^\infty$ be an i.i.d. random process with first order distribution $\text{Uniform}(0, 2\pi)$. Then, the sequence

Proposition 3. Let $\{\theta_i\}_{i=0}^\infty$ be an i.i.d. random process with first order distribution $\text{Uniform}(0, 2\pi)$. Then, the sequence

$$\{R(i)\}_{i=0}^\infty$$
 given by ($\forall i \geq 0$)

is strictly increasing and unbounded above, i.e.,

$$R(i) := \mathbb{E}\left[\log\left(1 + \sum_{l=1}^L e^{j\theta_l}\right)^2 | P\right].$$

($\forall i \geq 0$) $R(i+1) > R(i)$, and $\lim_{i \rightarrow \infty} R(i) = \infty$.

Proof. The proof is given in Appendix ??.

The gains of non-coherent joint transmission can be then formalized as follows.

The gains of non-coherent joint transmission can be then formalized as follows.

Proposition 4. The ergodic rate in (??) achieved by non-coherent joint transmission satisfies:

- (i) $(\forall L \geq 1) R \geq (1 - \frac{1}{L^2}) \log(1 + P);$
 - (ii) R is a strictly increasing sequence in $L;$
 - (iii) $\lim_{L \rightarrow \infty} R = \infty.$

Proof. The proof is given in Appendix

However, these potential gains may be very difficult to achieve in practice. In fact, although no CSIT is theoretically needed, practical coding schemes based on rate adaptation or HARQ mechanisms need to track and adapt to the instantaneous rate fluctuations $\log(1 + |h|^2 P)$. Unfortunately, Fig. 3. Ergodic and outage capacity versus blockage probability. The instantaneous rate $\log(1 + |h|^2 P)$ fluctuates at a much faster pace and over a much larger dynamic range than for the case of transmitter selection, where the fluctuations are driven by the blockage process only. Therefore, it is not clear whether non-coherent joint transmission is indeed preferable over transmitter selection when the resources consumed by these mechanisms must be taken into account.

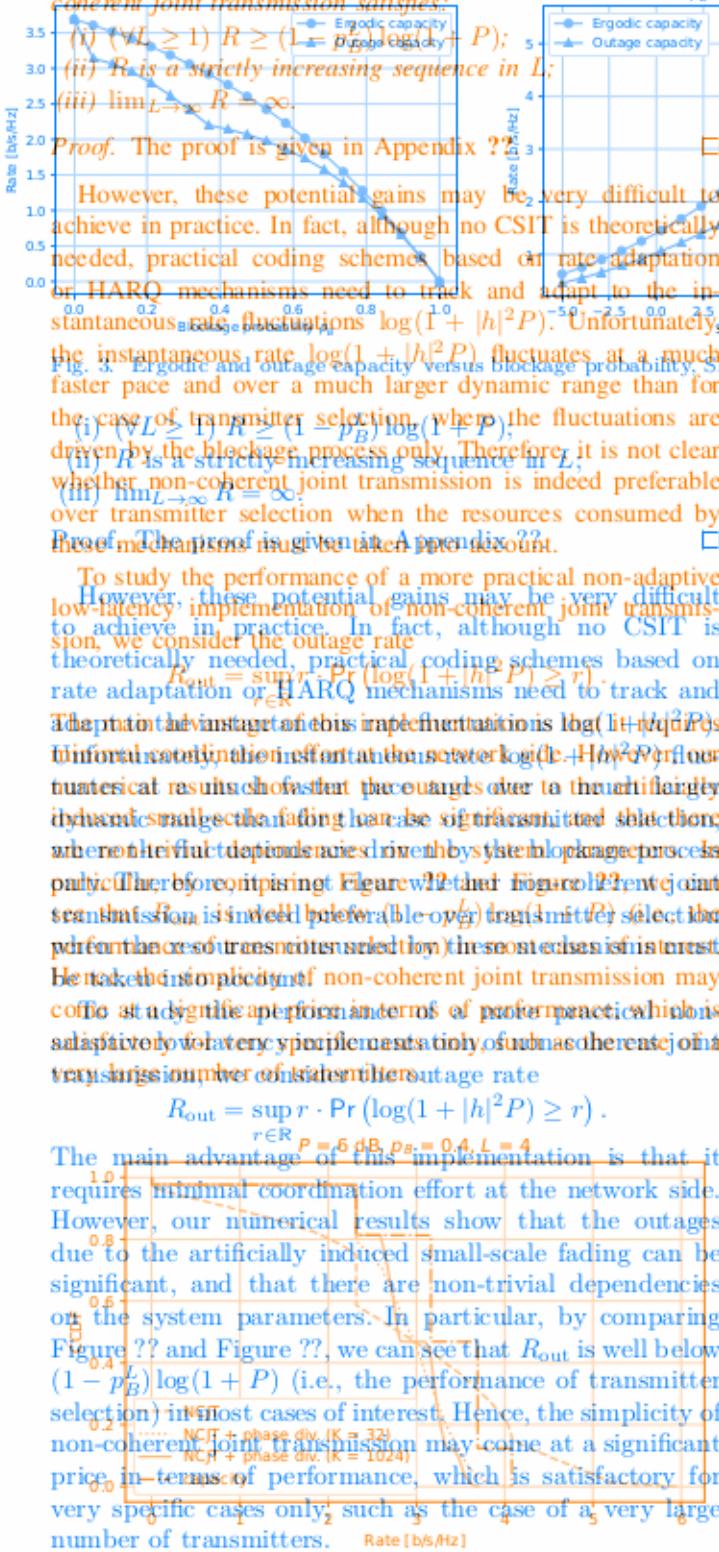
To study the performance of a more practical non-adaptive low-latency implementation of non-coherent joint transmission, we consider the outage rate

$$R_{\text{out}} = \sum_{r \in \mathbb{R}} r \cdot \Pr_{\text{user}}(\log(1 + h^{\text{user}} P) \geq r).$$

The main advantage of this implementation is $\log(1 + h^{\text{user}} P)$ is uniformly bounded in the interval $[0, \log(1 + H^{\text{user}} P)]$. Our numerical results show that the outages due to channel fading are significantly smaller than fading alone, so it sacrifices transmit selectivity, where the transmit dependencies are driven by the blocking of users. Therefore, comparing Figure 12 and Figure 13, we can see that the joint transmission is indeed preferable over the transmit selectivity. The lack of simplicity of non-coherent joint transmission may affect its practicability in terms of performance which is adaptively latency requirements. In fact, although no CSIT is theoretically needed, practical coding schemes based on rate adaptation or HARQ mechanisms need to track and

$$R_{\text{out}} = \sup_{r \in \mathbb{R}} r \cdot \Pr \left(\log(1 + |h|^2 P) \geq r \right).$$

The main advantage of this implementation is that it requires minimal coordination effort at the network side. However, our numerical results show that the outages due to the artificially induced small-scale fading can be significant, and that there are non-trivial dependencies on the system parameters. In particular, by comparing Figure ?? and Figure ??, we can see that R_{out} is well below $(1 - p_B^L) \log(1 + P)$ (i.e., the performance of transmitter selection) in most cases of interest. Hence, the simplicity of non-coherent joint transmission may come at a significant price in terms of performance, which is satisfactory for very specific cases only, such as the case of a very large number of transmitters.



C. Phase diversity Fig. 4: CCDF of the instantaneous capacity, and of the instantaneous rate of non-coherent joint transmission (NCJT) with and without phase diversity.

To address the limitations of non-coherent joint transmission, in this section we propose an extension (based on so-called phase diversity [?]) that mitigates the detrimental impacts of the limitations of incoherent joint transmission. The main idea is to propose controlled fast-faded receiver

called *phase diversity* [?]) that mitigates the detrimental impact of the artificially induced small-scale fading. The main idea is to induce a controlled fast-fading regime for escaping deep fading bursts. We split each fading block of length T into an integer number $M = T/K_{\text{of}}^{\text{f}}$ frames composed by K symbols each. For convenience, we rearrange the time-domain channel model onto a grid model (similar to an OFDM grid) as follows:

$$(\forall m \in \mathbb{Z}) (\forall k \in \{1, \dots, K\})$$

$$y[m, k] = \sum_{t=1}^L h_{t,k} x_t[m, k] + z[m, k], \quad t = \left\lfloor \frac{m}{M} \right\rfloor. \quad (6)$$

We then let $(\forall m \in \mathbb{Z})$ $(\forall k \in \{1, \dots, K\})$ $(\forall l \in \{1, \dots, L\})$
 $x_l[m, k] = e^{j\phi_{l,k}} u[m, k],$ (7)

R, and number of transmitters.
 $\text{where } u[m, k] \sim \mathcal{CN}(0, P)$ is a scalar i.i.d. information bearing signal, and where $(\forall l \in \{1, \dots, L\})$ $(\phi_{l,1}, \dots, \phi_{l,K})$ is a vector of random phases independently and uniformly distributed in $[0, 2\pi]$, which is shared among the transmitters and receiver as a common source of randomness. Provided that the number of frames M is large enough, standard arguments (as the ones used, e.g., for coding in OFDM systems) show that the following outage rate is achievable:

$$P_{\text{out}} = \sup_{r \in \mathbb{R}} r \cdot \Pr \left(\frac{1}{K} \sum_{k=1}^K \log(1 + |h[k]|^2 P) \geq r \right),$$

where $h[k] = \frac{1}{NCfT} \sum_{l=1}^{NCfT} \text{phase div}(K=32) \cdot e^{j(\theta_l + \phi_{l,k})}$ denotes an effective small-scale fading coefficient that fluctuates within the same frame of K symbols, according to the given vector of random phases. The proposed extension is motivated by the following asymptotic property:

Proposition 5 For all $i \in [0, \dots, I]$, and of the instantaneous rate of non-coherent joint transmission (NCJT) with and without phase diversity,

$$\lim_{K \rightarrow \infty} \Pr \left(\left| \frac{1}{K} \sum_{k=1}^K \log(1 + |h[k]|^2 P) - \bar{R}(\alpha) \right| \geq \epsilon \right) = 0,$$

Proof. The proof is based on the weak law of large numbers. We rearrange the time-domain channel model onto a grid (similar to an OFDM grid) as follows: ($\forall m \in \mathbb{Z}$)

The above proposition suggests that the proposed extension of non-coherent joint transmission can reduce the effect of the artificially induced small-scale fading, and make the instantaneous rate fluctuations be essentially driven by the aggregate blockage process $e^{-\rho t}$. This effect is clearly visible in Figure 7, where the PCDF of the scalar instantaneous bit-error signal \bar{B}_t (and where, $t \in [0, T]$) is a step-wise function similar to above CDF of random implants independently and uniformly distributed. The difference between these two figures illustrates the penalty for the discontinuous index k and m , since provided that the number P of frames is usually large enough, standard arguments (quite the ones used here for coding in QFDAM systems) show that the following (so-called) rate is achievable asymptotically larger or equal than the ergodic rate achieved by transmitter selection. This property can be easily formalized as follows:

Proposition 6. Let $\sum_{k=1}^L \text{Re}^{(k)}(\theta) \pm \phi_k$ denote an effective small shakeout fading rate efficient of that fluctuates with the θ 's and

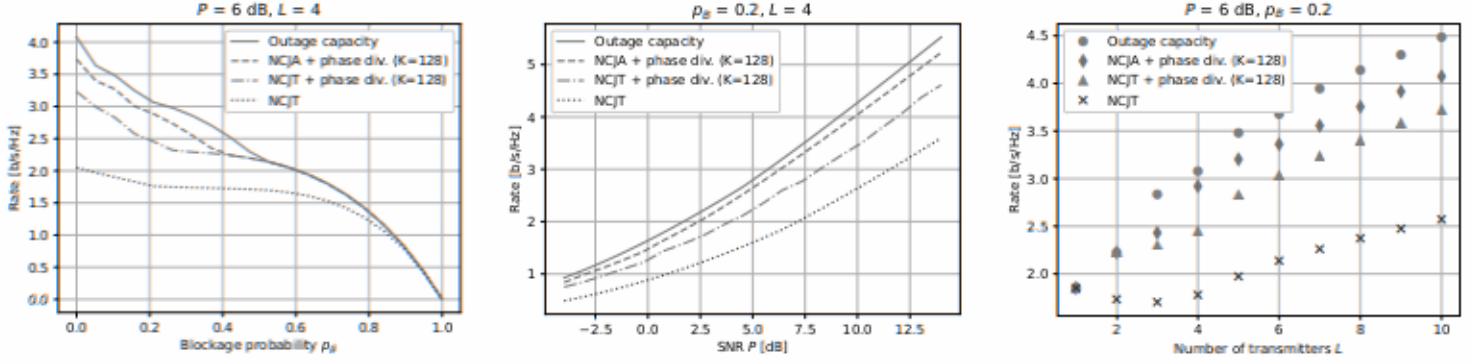


Fig. 5. Outage rate achieved by non-coherent joint transmission (NCJT) with phase diversity, NCJT with phase diversity, joint Alamouti space-time coding (NCJA) and NCJA with phase diversity, achieving system performance under different system parameters.

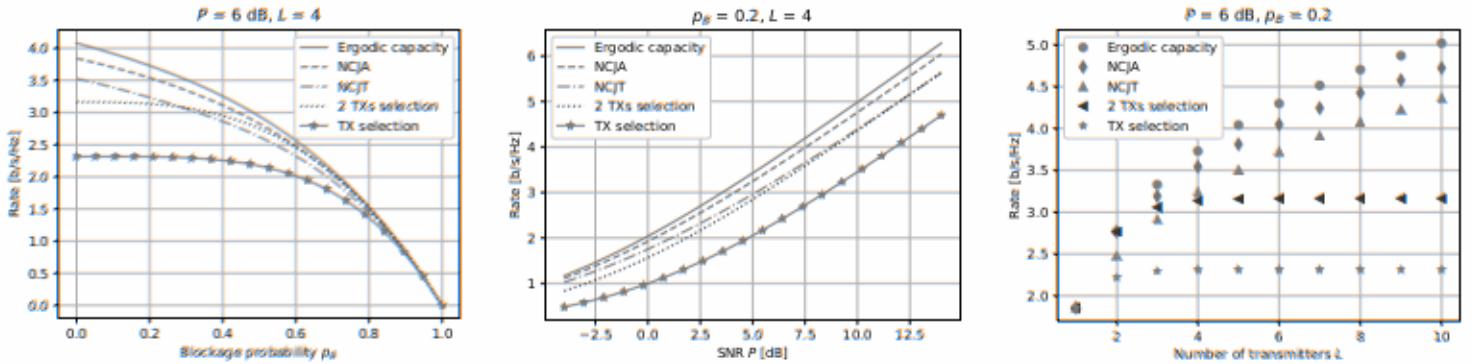


Fig. 6. Ergodic rate achieved by transmitter selection, coherent joint transmission (NCJT), joint transmission (NCJT), joint Alamouti space-time coding (NCJA) and NCJA with phase diversity under different system parameters.

extension of non-coherent joint transmission. Then: frame of K symbols, according to the given vector of random phases. The proposed extension is motivated by the following asymptotic property:

Proof. Same steps as in the proof of Proposition ??.

Proposition 5. For all $i \in \{0, \dots, L\}$, let $R(i)$ as in (??). Corollary 2. The following inequality holds:

Then, $(\forall \epsilon > 0)$

$$R_{\text{out}} \geq \Pr(R \geq 1)R(1) = (1 - p_B^L)\log(1 + R).$$

Despite the asymptotic nature of the above property, our numerical results in Figure ?? and Figure ?? show that i.e., the instantaneous rate converges in probability to $R(i)$ as K grows large. Therefore, at the price of a slight increase in channel coding complexity (comparable to OFDM systems), low propagation numbers. The details are given in Appendix ?? rates transmitter selection, without requiring fast network coordination and rate adaptation mechanisms based on CSIT.

The above proposition suggests that the proposed extension of non-coherent joint transmission achieves extension of the artificially induced small-scale fading, and make the instantaneous rate fluctuations be essentially driven by the channel gain. This respect is simply visible in Figure ?? and Figure ??, where the postulated extension rate stays very close to the ergodic regime, while it significantly reduces the CDF of the instantaneous rate approaching the ergodic rate. The difference between the large of fading is the Jensen's inequality for the discrete

out of L coding rates, based on casual knowledge of the quantities indexed by $i \geq 2$ since $(\forall i \geq 1) R(i) \leq \log(1 + iP)$ holds with equality for $i = 1$.

An important consequence of this property is that, in contrast to simple non-coherent joint transmission, Remark 1. The proposed extension is based on the so-called phase diversity scheme [2], which can be interpreted as a generalization of other known forms of diversity. In particular, the specific realization $(\forall l \in \{1, \dots, L\}) (\forall k \in \{1, \dots, K\})$ Proposition 6. Let \bar{R}_l/K for some $l \in \{1, \dots, L\} \in \Pr(\bar{R}_l)$ be the same performance rate asymptotically achievable by the proposed extension of non-coherent joint transmission. Then applying at each l th transmitter max $\Pr(\bar{R}_l \geq i)R(i)$ each frame of K symbols, and by performing frequency-domain processing. Proof. Same steps as in the proof of Proposition ??, although the original channel is not frequency selective. We Corollary 2. The following diversity property holds mostly studied under classical multi antenna fading models with Path scattering, such as the i.i.d. Rayleigh fading model, for which the effect of small-scale fading cannot be significantly mitigated. Despite the asymptotic nature of the above property, our numerical results in Figure ?? and Figure ?? show that similar conclusions may hold also for moderate values of K . Therefore, at the price of a slight increase in channel Space-time coding complexity (comparable to OFDM systems), the proposed extension can achieve better effective rate. At the same time, it is interesting to note that the proposed extension can achieve better effective rate than the simple space-time coding schemes, such as the Alamouti space-time block coding [3] which uses both

work good and outage capacity hypothesis implementation schemes based on CSIT and scalar decoding, since it converts the original channel model into a more effective thing proposed in next output achieves ($\forall m \in \mathbb{Z}$)

$$R = \mathbb{E} \left[\frac{y[m]}{K} \sum_{k=1}^K \log(1 + \|h[k]\|^2 P) \right] = \mathbb{E}[\log(1 + \|h\|^2 P)],$$

where $y[m]$ is a scalar information bearing signal, which we assume does not provide UoF performance, the gain with respect to simple non-coherent joint transmission. However, the proposed extension is still very useful, for the space-time coding schemes, since it significantly reduces the required complexity for approaching the ergodic capacity via rate adaptation. Especially, it would be enough to choose L in the entitle studies only to choose one out of L coding rates based on causal knowledge, such as realization of orthogonal packing to blockage process, linear as for dispersion codes [?], or similar alternatives. Similarly, in light of the fluctuation blocky fading significantly simplified the design of HARQ schemes.

However, we leave this line of research for future work, and focus on the following simple enhancements of the Remark 1. The proposed extension is based on the so-called phase diversity scheme [?], which can be interpreted as a generalization of other known forms of diversity.

In particular, the specific realization ($\forall l \in \{1, \dots, L\}$) ($\forall l \in \{1, \dots, K\}$) of a set $\{h_l\}_{l=1}^L$ for each $d_l \in \mathbb{N}$ gives the same performance as the grid-transmitter scheme [?] i.e., such that $B_l > 0$, the network chooses a pair of transmitters implemented such applying at each l th transmitter a cyclic shift d_l to each frame of K symbols and by performing Alamouti space-time coding. This effectively corresponds to introducing frequency diversity although the original channel is not frequency selective. We emphasize that this form of diversity channel model must be studied under classical multi-output channel model ($\forall m \in \mathbb{Z}$)

$$y[m] = \sqrt{\alpha_l} u[m] + z[m], \quad t = \lfloor \frac{m}{T} \rfloor,$$

where $u[m]$ is a scalar information bearing signal, which we assume i.i.d. $\mathcal{CN}(0, P)$. The following ergodic rate is achievable using standard techniques:

D. Space-time coding

$R = \Pr(\alpha \geq 2) \log(1 + 2P) + \Pr(\alpha = 1) \log(1 + P).$

For the special case of $L = 2$ transmitters, the well-known Alamouti space-time block coding scheme [?] exploits the SNR gains offered by simultaneous transmission from two transmitters. In contrast, the comparison against non-coherent joint transmission is non-trivial and highly dependent on the system parameters, as also shown in Figure ??.

single-input single-output model ($\forall m \in \mathbb{Z}$)

2) Non-coherent joint Alamouti space-time coding with phase diversity. We consider the same grid channel model in (??), and cluster the transmitters in pairs. For each pair of transmitters, we modify the transmit signals in (??) by replacing the iid information bearing signal $u[m]$, with its time coding schemes for $L > 2$ transmitters with reasonable complexity performance tradeoff are still worthy of investigation. One possibility would be to revisit the available studies on some specific classes of high-rate low-complexity space-time codes, such as quasi-orthogonal space-time block codes [?], linear dispersion codes [?], or similar alternatives, in light of the intermittent block fading model considered in this study. However, we leave this line of research for future work, and focus on the

following simple achievements of the two transmitter selection and phase diversity schemes by means of an Alamouti space-time coding stage $\Pr \left(\frac{1}{K} \sum_{k=1}^K \log(1 + \|h[k]\|^2 P) \geq r \right)$,

and) Two transmitters selection: For each t th block-fading realization with at least one non-blocked transmitters, i.e., such that $\log(1 + \|h[t]\|^2 P) > \mathbb{E}[\log(1 + \|h\|^2 P)]$ of transmitters $\{l^*, l'^*\}_{l=1}^L$ such that $(\beta_{l^*,t}, \beta_{l',t}) \neq (0,0)$, and let them transmit the same scalar information bearing channel using Alamouti space-time coding. The other joint transmission in the case of phase diversity is motivated by the significant benefit it offers in terms of outage rate and simplification of rate adaptation / HARQ mechanisms for the ergodic regime. In particular, similar to non-coherent joint transmission, as K grows large, the instantaneous rate becomes essentially driven by the blockage process:

$$y[m] = \sqrt{\min(2, \alpha_l)} u[m] + z[m], \quad t = \lfloor \frac{m}{T} \rfloor,$$

Proposition 7 Let $y[i]$ is information bearing signal, which we assume i.i.d. $\mathcal{CN}(0, P)^2$. $P + \sum_{l=L/2+1}^L e^{j\theta_l} P$ is achievable using standard techniques:

$$R = \Pr(\alpha \geq 2) \log(1 + 2P) + \Pr(\alpha = 1) \log(1 + P).$$

This scheme always outperforms transmitter selection, since K exploits the SNR gains offered by simultaneous transmission from two transmitters. In contrast, the comparison against non-coherent joint transmission is non-trivial and highly dependent on the system parameters as also shown in Figure ??.

Proof. (Sketch) The proof is based on the weak law of large numbers and it is similar to the proof of Proposition ??.

2) Non-coherent joint Alamouti space-time coding with phase diversity: We consider the same grid channel model in (??) and cluster the transmitters in pairs. For each pair of transmitters, we modify the transmit signals in (??) by replacing the i.i.d. information bearing signal $u[m, k]$ with its Alamouti space-time coded version (spanning two consecutive subchannels independently to compensate the different physical propagation delay of the signals from different subchannels only up to (some τ_l)(more thinking of set τ_{\max}))

$$h_t[k] := \begin{bmatrix} h_t^{1/2} \beta_{1,t} e^{j\theta_{1,t}} + z[m], \\ h_t^{1/2} \beta_{1,t} e^{j(\theta_{1,t} + \phi_{1,k})} + z[m], \end{bmatrix}$$

where the impulse responses $h_t[m]$ for $t \in \{1, \dots, L\}$ model the intermittent block fading process and the inter-symbol interference originating from the residual delay. Specifically, we let:

$$h_t[m] = u_t g(\tau_t), \quad h_{l,t} = \beta_{l,t} e^{j\theta_{l,t}}, \quad t = \lfloor \frac{m}{T} \rfloor,$$

where g is some continuous-time causal impulse response which models the convolution between the pulse shape and receiver filters, and $\tau_l \leq j \tau_{\max}$ is an unknown and possibly non-integer residual delay for the l th transmitter. We assume for simplicity standard square pulses at the transmitter side and matched filtering at the receiver side, i.e., a triangular impulse response

of outage rate and simplification of rate adaptation / HARQ mechanisms for the ergodic regime. In particular, similar to non-coherent joint transmission, as K grows large basic OFDM implementations. Different pulse modulation by the peak coverage but we leave their analysis to future work.

Proposition 7 In OFDM system with K subcarriers and a cyclic prefix of length $D \geq \lceil \tau_{\max} \rceil^2 + 1$, such that an integer number $M = T/(K+D)$ of OFDM symbols are transmitted in each fading block, the channel model in the time-frequency domain is given by $\forall m \in \mathbb{Z}$ $(\forall k \in \{0, \dots, K-1\})$

$$\Pr \left(\frac{1}{K} \sum_{t=0}^{K-1} \log(1 + \|h_t[k]\|^2 P) - R(\alpha_1, \alpha_2) \geq \epsilon \right) \xrightarrow{R \rightarrow \infty} 0,$$

$$Y[m, k] = \sum_{t=0}^{M-1} H_{l,t}[k] X_t[m, k] + Z[m, k], \quad t = \left[\frac{m}{M} \right],$$

where $(\alpha_1, \alpha_2) := (\sum_{l=1}^{L/2} \beta_l, \sum_{l=L/2+1}^L \beta_l)$.

where $Z[m, k] \sim \mathcal{CN}(0, 1)$ is a sample of a white Gaussian noise process (in the time and frequency), and where $H_{l,t}[k]$ of large amplitude and its time-frequency proof of (Proposition ??). By splitting the delay τ_l into an integer part $d_l \in \mathbb{N}$ and a fractional part $\delta_l \in [0, 1)$ such that $\tau_l = d_l + \delta_l$, we obtain

Interestingly, Figure ?? and Figure ?? show that the above scheme is able to recover a significant fraction of both the outage and the ergodic capacity in most regimes where $H_{l,t}[k] = \beta_{l,t} e^{j\theta_{l,t}} e^{-j2\pi \frac{k}{K} d_l} G_l[k]$.

A. Capacity IV. Coarse time synchronization

We capture the impact of the network and jointly elaborate different physical layer propagation models of the signals from different transmitters in asynchronous channel by varying offset channels. As discussed earlier, the capacity of OFDM may provide a robust ergodic link under unknown delays. In particular, we consider the implications of an upper bound on the ergodic capacity of the considered channel is readily given by the ergodic capacity R under perfect time synchronization $C = \mathbb{E}[\log(1 + y/P)] = \sum_i (h_i[m] + z[m])$.

In addition, by applying for each subcarrier the signaling where the impulse responses $h_l[m]$ for $l \in \{1, \dots, L\}$ scheme achieving C , we obtain the following lower bound on the ergodic capacity:

$$R = \inf_{\tau \in [0, \tau_{\max}]^L} \frac{1}{K+D} \sum_{k=0}^{K-1} \mathbb{E} \left[\log \left(1 + \sum_{l=1}^L |H_{l,k}|^2 P \right) \right] \quad (8)$$

where g is some continuous-time causal impulse response which models the convolution between the pulse shape $H_{l,k}$ and receiver filters, and τ_{\max} is an unknown and possibly non-integer residual delay for transmitter capacity loss only through the cyclic prefix length D , i.e., the system is insensitive to their actual values. On the other hand, fractional parts $\delta_l = (\delta_{l,1}, \dots, \delta_{l,L})$ contribute to capacity loss in a non-trivial manner through the induced frequency selectivity, i.e., through the fluctuations of $|G_l[k]|^2$ across the subcarriers.

It turns out that the ergodic rate R in (??) admits a closed form expression given by the intuitive case of completely off-grid sampling ($\forall l \in \{1, \dots, L\}$) $\delta_l = 0.5$. Future work.

Proposition 8 An OFDM system with K subcarriers. The following prefix of length $D \geq \lceil \tau_{\max} \rceil + 1$, such that an integer number $M = T/(K+D)$ of OFDM symbols are transmitted in each fading block, the channel model in the time-frequency domain is given by $(\forall m \in \mathbb{Z})$ $(\forall k \in \{0, \dots, K-1\})$

$$\lim_{K \rightarrow \infty} R = \mathbb{E} \left[\log \left(1 + \alpha \frac{P}{4} + \frac{1}{2} (\sqrt{1 + \alpha P} - 1) \right) \right] \quad (9)$$

$$Y[m, k] = \sum_{t=0}^{M-1} H_{l,t}[k] X_t[m, k] + Z[m, k], \quad t = \left[\frac{m}{M} \right],$$

Proof. The proof is given in Appendix ??.

where $Z[m, k] \sim \mathcal{CN}(0, 1)$ is a sample of a white Gaussian noise process (similar to the case of frequency-hopping where $H_{l,t}[k]$ this channel is characterized by the time delays of the

outage capacity $\text{outage}_m[m]$. By perfecting the synchronization integer part $d_l \in \mathbb{N}$ and a fractional part $\log \in [0, P]$, such that $\tau_l = d_l + \delta_l$, we obtain and lower bounded by

$$R_{\text{out}} = \inf_{\tau \in [0, \tau_{\max}]^L} \sup_{r \in \mathbb{R}} \Pr \left(\frac{1}{K+D} \sum_{k=0}^{K-1} \log \left(1 + \sum_{l=1}^L |H_{l,k}|^2 P \right) \geq r \right), \quad (9)$$

which captures the impact of clustered forms stated controllable residual delays, we follow a worst-case approach similar to the information theoretical literature on asynchronous [?].

Proposition 9. Consider the achievable outage rate (??). The following equality holds:

$R_{\text{out}} = \max_{\tau \in [0, \tau_{\max}]^L} \text{Capacity}$ for all possible delays $\tau = (\tau_1, \dots, \tau_L) \in [0, \tau_{\max}]^L$. This implies $\Pr(\alpha \geq i) \leq \frac{1}{K+D} \sum_{k=0}^{K-1} \log \left(\frac{P}{2} \left(\frac{2\pi k}{K} \right) \right)$ the considered channel is readily given by the ergodic capacity R under perfect time synchronization $C = \mathbb{E}[\log(1 + \alpha P)]$, studied in Proposition ?? $R = \frac{1}{4} \log \left(1 + i \frac{P}{4} + \frac{1}{2} (\sqrt{1 + iP} - 1) \right)$.

In addition, by applying for each subcarrier the signaling scheme (achieving C) we obtain the following lower bound on the ergodic capacity:

Proposition ?? and Proposition ?? show that the capacity of the considered system follows similar trends as in the synchronous case. In addition, they characterize the price of enforcing robustness against uncontrollable residual delays. More specifically, the above results show that robust integer parts $d_l = (d_{l,1}, \dots, d_{l,L})$ of the delays τ_l contribute to capacity loss only through the cyclic prefix length D , i.e., the system is insensitive to their actual values. On the other hand, the fractional parts $\delta_l = (\delta_{l,1}, \dots, \delta_{l,L})$ contribute to capacity loss in a non-trivial manner through the induced frequency selectivity, i.e., through the fluctuations of $|G_l[k]|^2$ across the subcarriers.

It turns out that the ergodic rate R in (??) admits a closed form expression given by the intuitive case of completely off-grid sampling ($\forall l \in \{1, \dots, L\}$) $\delta_l = 0.5$.

Proposition 8. Consider the achievable ergodic rate (??). The following equality holds:

$$R = \frac{1}{K+D} \sum_{k=0}^{K-1} \mathbb{E} \left[\log \left(1 + \alpha \frac{P}{2} \left(1 + \cos \left(\frac{2\pi k}{K} \right) \right) \right) \right].$$

B. Non-coherent joint transmission and phase diversity

Furthermore, As a more practical alternative to the near-optimal scheme studied in the previous section, we now revisit non-coherent joint transmission, i.e., the transmission of a single scalar code word from all transmitters simultaneously, with and without phase diversity. By taking the same approach, we achieve the capacity of the considered channel can be upper bounded by the tight joint with phase diversity.

$$C_{\text{out}} = \max_{\tau \in [0, \tau_{\max}]^L} \Pr(\alpha \geq i) \log(1 + iP)^2$$

$$\text{and lower bounded by } R_{\text{out}} = \inf_{\tau \in [0, \tau_{\max}]^L} \sup_{r \in \mathbb{R}} \Pr \left(\frac{1}{K+D} \sum_{k=0}^{K-1} \log \left(1 + \left| \sum_{l=1}^L H_{l,k} \right|^2 P \right) \geq r \right).$$

The above expression can be evaluated as stated next. Similarly to the capacity analysis, the resulting expression corresponds to the case of completely off-grid sampling, which can be characterized in closed form as stated next.

Proposition 10. Consider the achievable outage rate (??). The following equality holds:

$$R = \frac{1}{K+D} \sum_{k=0}^{K-1} \mathbb{E} \left[\log \left(1 + |h|^2 \frac{P}{2} \left(1 + \cos \left(\frac{2\pi k}{K} \right) \right) \right) \right],$$

where $|h| := \sqrt{\sum_{l=1}^L b_l e^{j\theta_l}} \log \left(1 + \cos \left(\frac{2\pi k}{K} \right) \right)$. Furthermore,

$$\text{Furthermore, } \lim_{K \rightarrow \infty} R = \max_{\alpha \geq i} \Pr(\alpha \geq i) \log \left(1 + i \frac{P}{4} + \frac{1}{2} (\sqrt{1 + |h|^2 P} - 1) \right).$$

Proof. The proof is given in Appendix ??.

Proof. (Sketch) The proof follows from the same techniques as in the proof of Proposition ??.

The above proposition shows that the ergodic rate achieved by non-coherent joint transmission follows similar trends as in the synchronous case, and experience similar penalties due to cyclic prefix overhead and off-grid sampling as in the worst-case price of enforcing robustness against uncontrollable residual delays. More specifically, the above results show that robust transmission can be theoretically achieved. We know that if phase diversity is not used, the outage at the price of a cyclic prefix overhead $\approx K/(K+D)$ performance must be at least as bad as in the synchronous case which depends on the maximum uncontrollable integer delay and a SNR loss factor $0.5 + 0.5 \cos(2\pi k/K)$ for against uncontrollable delays. Phase diversity is expected to each subcarrier k which stems from the uncontrollable improve performance significantly, but we provide a formal fractional justifications only in an asymptotic sense. Clearly, the price of the multiplicative overhead can be mitigated if the number of subcarriers K .

Building on the OFDM grid model, we apply the phase diversity technique directly in the frequency domain, i.e., we let $(\forall m \in \mathbb{Z}) (\forall k \in \{0, \dots, K-1\}) (\forall l \in \{1, \dots, L\})$, the effective SNR for a given number of non-blocked transmitters $\alpha \in \mathbb{N}$ approaches

where $U[m, k] \sim \mathcal{CN}(0, P)$ is a scalar i.i.d. information bearing signal, and where $(\forall l \in \{1, \dots, L\}) (\phi_{l,0}, \dots, \phi_{l,K-1})$ is a vector of random phases uniformly distributed in $[0, 2\pi]$. For given delays τ , the following instantaneous rate is achievable on subcarrier $k \in \{0, \dots, K-1\}$:

B. Non-coherent joint transmission and phase diversity

As a more practical alternative to the near-optimal scheme studied in the previous section, we now revisit Similar to the synchronous case, we then observe that non-coherent joint transmission, i.e., the transmission of a single scalar codeword from all transmitters simultaneously with and without phase diversity. By taking the same worst-case approach as in the capacity analysis, we first consider the achievable ergodic rate (with or without phase diversity) R_k :

$$R_k(\beta) := \mathbb{E} \left[\log \left(1 + \left| \sum_{l=1}^L b_l e^{j\theta_l} G_l[k] \right|^2 P \right) \right].$$

$$\text{Then, } \inf_{\tau \in [0, \tau_{\max}]} \frac{1}{K+D} \sum_{k=0}^{K-1} \mathbb{E} \left[\log \left(1 + \left| \sum_{l=1}^L H_l[k] \right|^2 P \right) \right].$$

$$\lim_{K \rightarrow \infty} \Pr \left(\left| \frac{1}{K+D} \sum_{k=0}^{K-1} (R_k - R_k(\beta)) \right| \geq \epsilon \right) = 0, \quad (10)$$

The above expression can be evaluated as stated next. Similarly to the capacity analysis, the resulting expression corresponds to the case of completely off-grid sampling.

Proof. The proof is based on Hoeffding's inequality. The details are given in Appendix ??.

The following equality holds:

$$R = \frac{1}{K+D} \sum_{k=0}^{K-1} \mathbb{E} \left[\log \left(1 + |h|^2 \frac{P}{2} \left(1 + \cos \left(\frac{2\pi k}{K} \right) \right) \right) \right],$$

On top of simplifying practical implementations of rate adaptation and HARQ mechanisms for the ergodic regime, the above property allows us to approximately evaluate the (worst-case) outage performance by focusing on the asymptotically

$$\lim_{K \rightarrow \infty} R = \mathbb{E} \left[\log \left(1 + |h|^2 \frac{P}{4} + \frac{1}{2} (\sqrt{1 + |h|^2 P} - 1) \right) \right].$$

Reaching the proof given in Appendix ??.

□

The above proposition shows that the ergodic rate achieved by non-coherent joint transmission follows similar trends as in the synchronous case, and experience

which are characterized by cyclic prefix overhead and off-grid sampling as in the worst-case capacity analysis. **Proposition 12.** Consider the achievable outage rate (??). The following equality holds:

Unfortunately, the exact (worst-case) achievable outage rate seems difficult to characterize for finite K . We know that, if phase diversity is not used, the outage performance must be at least as bad as in the synchronous case and include the aforementioned penalties to ensure robustness against uncontrollable delays. Phase diversity is expected to improve performance significantly, but we provide a formal justification only in an asymptotic sense.

Building on the OFDM grid model, we apply the phase diversity technique directly in the frequency domain, i.e., we let $(\forall i \in \{0, \dots, K-1\}) (\forall l \in \{1, \dots, L\})$

$$\text{where } X_l[m, k] = e^{j\phi_{l,k}} U[m, k] e^{j\theta_l} \left| \sum_{l=1}^L e^{j\phi_{l,k}} P - 1 \right|$$

where $U[m, k] \sim \mathcal{CN}(0, P)$ is a scalar i.i.d. information bearing signal, and where $(\forall l \in \{1, \dots, L\}) (\phi_{l,0}, \dots, \phi_{l,K-1})$ is a vector of random phases independently (and uniformly) distributed in $[0, 2\pi]$. The proof follows from the proof of Proposition ??.

Given by the proof of Proposition ??, the instantaneous rate is achievable on subcarrier $k \in \{0, \dots, K-1\}$:

As expected, we observe that the asymptotically achievable outage rate follows similar trends as in the synchronous case and experience similar penalties as in the worst-case capacity analysis. We conclude this section by pointing out that the above discussion can be extended to a related frequency domain version of the non-coherent joint Alamouti scheme with phase diversity studied in Section ??.

However, we omit the details due to space limitation. For the same reason, we do not provide additional numerical results. Nevertheless, we report that, as expected, the numerical results are similar to the synchronous case up to a SNR penalty of roughly 6dB. Then ($\forall \epsilon > 0$)

$$\lim_{K \rightarrow \infty} \Pr \left(\left| \frac{1}{K+D} \sum_{k=0}^{K-1} (R_k - R_k(\beta)) \right| \geq \epsilon \right) = 0,$$

Our results provide a set of practical design guidelines for realizing full macrodiversity gains and significant SNR gains in wave/sub THz downlink access networks with low-complexity receivers and minimal coordination and synchronization requirements at the infrastructure side. In particular, we illuminate the importance of transmit diversity techniques to make non-coherent joint transmission a competitive alternative to standard transmitter selection. Specifically, the proposed extension based on phase diversity significantly improves the outage performance of non-adaptive (fixed-rate) coding schemes, and it allows to design rate adaptation or HARQ mechanisms focusing only on the large-scale fluctuations given by the relatively slow blockage process. Furthermore, we illuminate the potential of simple space-time coding schemes which can be significantly faster than the optimal performance.

Finally, we show that robustness against timing offsets can be achieved using OFDM and a simple SNR margin. The following equality holds:

One main limitation of this study is that, in order to gain analytical insights, the considered channel model is significantly simplified. Hence, an interesting research direction is to test or

where (analytically or via simulation) the results of this study by considering more realistic models covering 2-dg., asymmetric path loss, correlated blockages, multi-path propagation, and further interference. Another main limitation is that the considered solution against timing offsets may be quite poor in the context of energy efficiency. In particular, the sensitivity of OFDM to hardware non-linearities is known to be a major cause of reduced energy efficiency. In addition, the reported SNR margin of 6dB may be too conservative, since based on a worst-case approach and on a simple pulse shape which is known to be quite sensitive to off-grid sampling. Hence, the extension of our results to alternative waveforms, pulse shapes, and less conservative approaches is of great interest. Other interesting research directions include the design of better performing space-time codes tailored to the considered multi-point intermittent block fading model, the study of the outage versus latency tradeoff using the properties of the blockage process, and the extension to multi-antenna receivers.

Proof. (Sketch) The proof follows from the same techniques as in the proof of Proposition 1. The proof is of great interest. Other interesting research directions include the design of better performing space-time codes tailored to the considered multi-point intermittent block fading model, the study of the outage versus latency tradeoff using the properties of the blockage process, and the extension to multi-antenna receivers.

We first state two auxiliary lemmas. As expected, we observe that the asymptotically achievable multi-point outage rate follows similar trends as in the synchronous case, and experience similar penalties as in the worst-case capacity analysis. We conclude this section by pointing out that the above discussion can be extended to a related frequency-domain version of the non-coherent joint Alamouti scheme with phase diversity studied in Section ??.

Lemma 1. Let $\theta \sim \text{Uniform}(0, 2\pi)$. Then $(\forall x \in \mathbb{C}) (\forall y \in \mathbb{C})$

$$\mathbb{E}[\log(1 + |x + e^{j\theta}y|^2)] = \log\left(\frac{1 + a + \sqrt{(1+a)^2 - b^2}}{2}\right),$$

where $a := |x| + |y|$ and $b := 2|x||y|$.

Our results provide a set of practical design guidelines. **Proof.** We define the shorthand $c := \frac{\theta}{2\pi} \in [0, 1]$, and let: $\mathbb{E}[\log(1 + |x + e^{j\theta}y|^2)]$ gains in mmWave/sub-THz downlink access networks with low-complexity receivers and minimal coordination and synchronization requirements at the infrastructure side. In particular, we illuminate the importance of transmit diversity techniques to make non-coherent joint transmission a competitive alternative to standard transmitter selection. Specifically, the proposed extension based on phase diversity significantly improves the outage performance of where adaptive (flexible rate) coding, scheduling and rate adaptation mechanisms focusing only on the large-scale fluctuations given by the relatively slow blockage process. Furthermore, we illuminate the potential of simple space-time coding schemes to recover a significant fraction of the optimal performance. Finally, we show that robustness against timing offsets can be achieved using OFDM and a simple SNR margin.

One main limitation of this study is that, in order to gain analytical insights, the considered channel model is significantly simplified. Hence, an interesting research direction is to test or extend (analytically or via simulation) the results of this study by considering more realistic models covering, e.g., asymmetric path loss, correlated blockages, multi-path propagation, and multi-user interference. Another main limitation is that $R(i)$ is strictly increasing in i , i.e., $\Pr(\alpha_L \geq L) > \Pr(\alpha_{L-1} \geq L)$. We are therefore ready to prove that $R(i)$ is strictly increasing in i , i.e., $R(i+1) > R(i)$. We focus on the case of $i \geq 0$, since the case $i \leq 0$ is trivial. We define the random variable $X := \sum_{l=1}^L e^{j\theta_l}$ independent of θ_{l+1} , and observe that

We are therefore ready to prove that $R(i)$ is strictly increasing in i , i.e., $R(i+1) > R(i)$. We focus on the case of $i \geq 0$, since the case $i \leq 0$ is trivial. We define the random variable $X := \sum_{l=1}^L e^{j\theta_l}$ independent of θ_{l+1} , and observe that

In addition, the reported SNR margin of 6dB may be too conservative, since based on a worst-case approach and on a simple pulse shape which is known to be quite sensitive to off-grid sampling. Hence, the extension of our results to alternative waveforms, pulse shapes, and less conservative approaches is of great interest. Other interesting research directions include the design of better performing space-time codes tailored to the considered multi-point intermittent block fading model, the study of the outage versus latency tradeoff using the properties of the blockage process, and the extension to multi-antenna receivers.

We finally show that $R(i)$ is unbounded above, as a direct implication of the following inequality:

$$\lim_{i \rightarrow \infty} \mathbb{E}[\log(1 + iP)] = \infty,$$

To prove the above inequality, we let $X_i := \sum_{l=1}^i e^{j\theta_l}$ and observe that by Markov's inequality ($\forall i \geq 1$)

$$\Pr(|X_i| \geq iP) \geq \Pr(\log(1 + iP) \geq \log(1 + iP))$$

$$= \Pr(|X_i|^2 \geq 1).$$

By the central limit theorem (Appendix A), X_i converges in distribution to $X \sim \mathcal{CN}(0, 1)$ (Proposition ??). Hence, by the continuous mapping theorem, $|X_i|^2$ converges in distribution to $|X|^2 \sim \text{Exp}(1)$. We first state two auxiliary lemmas.

Lemma 1. $R(\theta) \sim \text{Uniform}(0, 2\pi)$. Then $(\forall x \in \mathbb{C}) (\forall y \in \mathbb{C})$

$$\lim_{c \rightarrow \infty} \frac{\mathbb{E}[\log(1 + |x + e^{j\theta}y|^2)]}{\log(1 + iP)} \geq \lim_{i \rightarrow \infty} \Pr(|X_i|^2 \geq 1) = \Pr(|X|^2 \geq 1),$$

B. Proof of Proposition ?? where $a := |x|^2 + |y|^2$ and $b := 2|x||y|$.

We write $R(L)$ to highlight the dependency of R on L . Proof. We define the shorthand $c := \frac{\theta}{2\pi} \in [0, 1]$, and let:

The proof relies on the following identity, which follows by the law of total expectation ($L \geq 1$)

$$R(L) = \mathbb{E}\left[\mathbb{E}\left[\log\left(1 + \sum_{l=1}^L \beta_l e^{j\theta_l} P\right)\right] \middle| \beta_1, \dots, \beta_L\right]$$

$$= \log(1 + a) + \mathbb{E}[\log(1 + c \cos(\theta))],$$

where $\alpha_L := \sum_{l=1}^L \beta_l$. To prove (i), we observe that ($\forall L \geq 1$)

$$\mathbb{E}[\bar{R}(\alpha_L)] \geq \Pr(\alpha_L \geq iP) \frac{1 + a + \sqrt{(1+a)^2 - b^2}}{2},$$

where the last equality follows from algebraic manipulations and the known definition of integral $\Pr(\alpha_L \geq iP) = \int_{iP}^{\pi} \Pr(\alpha_L \geq t) dt$.

which the inequality follows from Proposition ??.

To prove (ii), we define the shorthand $\Delta R(i) := R(i) - R(i-1) > 0$,

where the inequality follows from Proposition ??.

We then observe that for $(\forall L \geq 2) \in \mathbb{C}$, the function $f : \mathbb{R}_+ \rightarrow \mathbb{R}$ given by

$$\mathbb{E}[\bar{R}(\alpha_L)] = \sum_{i=1}^L \Pr(\alpha_L \geq iP) \frac{1 + a + \sqrt{(1+a)^2 - b^2}}{2^i},$$

where $a := |x|^2 + |y|^2$ and $\bar{R}(L) := \sum_{i=1}^L \Pr(\alpha_L \geq iP)$ is strictly increasing within its domain.

Proof. (Sketch) Standard calculus shows that the argument of the logarithm has strictly positive derivative for all $x \in \mathbb{C}$. $= \Pr(\alpha_L \geq L) \Delta R(L) + \mathbb{E}[\bar{R}(\alpha_{L-1})]$

We are now ready to prove that $R(i)$ is strictly increasing, i.e., that $(\forall i \geq 0) \bar{R}(i+1) > \bar{R}(i)$. We focus on $i \geq 1$, since the case $i \geq 0$ is trivial. We define the random variable $X := \sum_{l=1}^L e^{j\theta_l}$ independent of θ_{l+1} , and observe that

$(\forall i \geq 0) \mathbb{E}[\bar{R}(\alpha_L)] \geq \Pr(\alpha_L \geq i) \bar{R}(i).$

Taking the limit on both sides gives

$$\lim_{i \rightarrow \infty} \mathbb{E}[\bar{R}(\alpha_L)] \geq \lim_{i \rightarrow \infty} \Pr(\alpha_L \geq i) \bar{R}(i) = \bar{R}(i),$$

which, by Proposition ??, implies $\lim_{L \rightarrow \infty} \mathbb{E}[\bar{R}(\alpha_L)] = \infty$.

Where the inequality follows by Lemma ?? and Lemma ??.

We finally show that $\bar{R}(i)$ is unbounded above, as a direct implication of the following inequality:

$$\Pr \left(\left| \frac{1}{K} \sum_{k=1}^K \log(1 + |h[k]|^2 P) - \bar{R}(\alpha) \right| \geq \epsilon \right).$$

To prove the above inequality, we let $X_i := \frac{1}{K} \sum_{k=1}^K e^{j\theta_k}$, and observe that by Markov's inequality ($\forall i \geq 1$) $\Pr[X_i \geq \bar{R}(i)] \leq e^{-\bar{R}(i)}$.

We then observe that $\Pr[\log(1 + |h[k]|^2 P) \geq \bar{R}(\alpha)] \geq e^{-\bar{R}(\alpha)}$ and $\Pr[\sum_{k=1}^K \log(1 + |h[k]|^2 P) \geq K\bar{R}(\alpha)] \geq e^{-K\bar{R}(\alpha)}$. Since $\{h[k]\}_{k=1}^K$ is an i.i.d. sequence with mean $\bar{R}(\alpha)$, we have $\Pr[\sum_{k=1}^K \log(1 + |h[k]|^2 P) \geq K\bar{R}(\alpha)] \geq e^{-K\bar{R}(\alpha)}$.

By the central limit theorem, X_i converges in distribution to $\mathcal{N}(0, 1 + \bar{R}(\alpha)^2 P)$. Hence, by the continuous mapping theorem, $|X_i|^2$ converges in distribution to $\chi_1^2 \sim \text{Exp}(1)$. Taking limits on both sides concludes the proof, e.g., on the dominated convergence theorem (probabilities are trivially bounded).

$$\lim_{i \rightarrow \infty} \frac{\bar{R}(i)}{\log(1 + iP)} \geq \lim_{i \rightarrow \infty} \Pr(|X_i|^2 \geq 1) = \Pr(|X|^2 \geq 1).$$

D. Proof of Proposition ??

B. Proof of Proposition ??

We write $R(L)$ to highlight the dependency of R on L . The proof relies on the following identity, which follows by the law of total expectation: $(\forall L \geq 1)$

$$R(L) = \inf_{\beta \in \mathbb{R}} \mathbb{E} \left[\log \left(\frac{1}{K+D} \sum_{k=0}^{K-1} \mathbb{E} \left[\log \left(1 + \sum_{l=1}^L \beta_l |G_l[k]|^2 P \right) \middle| \beta_1, \dots, \beta_L \right] \right) \right].$$

Then, we notice that $|G_l[k]|^2 = \delta_l^2 + (1 - \delta_l)^2 + 2\delta_l(1 - \delta_l)\cos(2\pi k/K)$ is an upward facing parabola in δ_l , with minimum at $\delta_l = 0.5$, for all subcarriers $k \in \{0, \dots, K-1\}$.

Since the argument of the expectation is monotonic increasing in $|G_l[k]|^2$, the desired expression follows by letting $(\forall l \in \{1, \dots, L\}) \delta_l = 0.5$ and from simple algebraic manipulations. For the second part of the statement, we observe that, by standard properties of Riemann sums, we have $(\forall n \in \mathbb{N})$,

$$\sum_{k=0}^{K-1} \Pr(\alpha_L = k) R(1) = (1 - p_B) \log(1 + P),$$

where the inequality follows from Proposition ?? To prove (ii), we define the shorthand $\Delta \bar{R}(i) := \bar{R}(i) - \bar{R}(i-1) > 0$, where the inequality follows from Proposition ?? We then observe that $(\forall L \geq 2) \sum_{i=1}^L \Pr(\alpha_L = i) \Delta \bar{R}(i) > 0$.

The integral can be evaluated in closed form using the same technique as in the proof of Lemma ?? in Appendix ??, concluding the proof. We remark that the limit and the expectation can be exchanged since α is a discrete random variable. $= \Pr(\alpha_L \geq L) \Delta \bar{R}(L) + \sum_{i=1}^{L-1} \Pr(\alpha_L \geq i) \Delta \bar{R}(i)$

E. Proof of Proposition ?? $\Delta \bar{R}(L) + \sum_{i=1}^{L-1} \Pr(\alpha_{L-1} \geq i) \Delta \bar{R}(i)$

By the law of total probability, we have

$$\Pr \left(\left| \frac{1}{K+D} \sum_{k=0}^{K-1} (R_k - \bar{R}_k(\beta)) \right| \geq \epsilon \right)$$

where the first inequality follows from the property $(\forall L \geq 1) (\forall i \geq 0) \Pr(\alpha_L \geq i) \geq \Pr(\alpha_{L-1} \geq i)$.

To prove (iii) we observe that, by Markov's inequality:

$$\Pr \left(\left| \frac{1}{K+D} \sum_{k=0}^{K-1} (R_k - \bar{R}_k(\beta)) \right| \geq \epsilon \right) \leq e^{-\epsilon^2}.$$

We then observe that, conditioned on (β, θ) , each R_k is an independent random variable with mean $\bar{R}_k(\beta)$. Furthermore, we observe that $\sum_{k=0}^{K-1} \Pr(\alpha_L \geq i) R(k) = R(i)$ which (by Proposition ?? we imply by Hoeffding's inequality)

On Proof of Proposition ??

which gives

$$\Pr \left(\left| \frac{1}{K+D} \sum_{k=0}^{K-1} (R_k - \bar{R}_k(\beta)) \right| \geq \epsilon \right) \leq 2e^{-\frac{2K\epsilon^2}{\log^2(1+L^2P)}}.$$

The proof is concluded by exchanging limit and expectation based, e.g., on the dominated convergence theorem (probabilities are trivially bounded). $= \Pr \left(\left| \frac{1}{K+D} \sum_{k=0}^{K-1} \log(1 + |h[k]|^2 P) - R(\alpha) \right| \geq \epsilon \right) \leq e^{-\epsilon^2}.$

We then observe that, conditioned on (β, θ) , $\{\log(1 + |h[k]|^2 P)\}_{k=1}^K$ is an i.i.d. sequence with mean $R(\alpha)$. Hence, the weak law of large numbers applies and

$$X := \sum_{k=1}^K H_k \text{ and write } \Pr \left(\left| \frac{1}{K+D} \sum_{k=1}^K \log(1 + |h[k]|^2 P) - R(\alpha) \right| \geq \epsilon \right) \xrightarrow{K \rightarrow \infty} 0.$$

The proof is concluded by exchanging limit and expectation based, e.g., on the dominated convergence theorem (probabilities are trivially bounded). $= \mathbb{E} \left[\mathbb{E} \left[\log \left(1 + |X + H_L[k]|^2 P \right) \middle| X, \beta_L \right] \right].$

Using Lemma ?? and Lemma ?? in Appendix ??, we notice that the inner expectation is an increasing function of β_L . For the first part of the statement, we observe that

as in the proof of Proposition ?? in Appendix ?? The second part follows similar steps as the second part of the proof of Proposition ??, by replacing α with $|h|^2$. The limit and expectation can be exchanged based, e.g., on the dominated convergence theorem $\lim_{\delta \in [0, 1]} \sum_{k=0}^{K+D-1} \mathbb{E} \left[\log \left(1 + \sum_{l=1}^L \beta_l |G_l[k]|^2 P \right) \right]$.

Then, we notice that $|G_l[k]|^2 = \delta_l^2 + (1 - \delta_l)^2 + 2\delta_l(1 - \delta_l)\cos(2\pi k/K)$ is an upward facing parabola in δ_l , with minimum at $\delta_l = 0.5$, for all subcarriers $k \in \{0, \dots, K-1\}$. Since the argument of the expectation is monotonic increasing in $|G_l[k]|^2$, the desired expression follows by letting $(\forall l \in \{1, \dots, L\}) \delta_l = 0.5$ and from simple algebraic manipulations. For the second part of the statement, we observe that, by standard properties of Riemann sums, we have $(\forall n \in \mathbb{N})$

$$\begin{aligned} & \lim_{K \rightarrow \infty} \frac{1}{K+D} \sum_{k=0}^{K-1} \log \left(1 + \alpha \frac{P}{2} \left(1 + \cos \left(\frac{2\pi k}{K} \right) \right) \right) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left(1 + \alpha \frac{P}{2} (1 + \cos(\theta)) \right) d\theta. \end{aligned}$$

The integral can be evaluated in closed form using the same technique as in the proof of Lemma ?? in Appendix ??, concluding the proof. We remark that the limit and the expectation can be exchanged since α is a discrete random variable.

E. Proof of Proposition ??

By the law of total probability, we have

$$\begin{aligned} & \Pr \left(\left| \frac{1}{K+D} \sum_{k=0}^{K-1} (R_k - \bar{R}_k(\beta)) \right| \geq \epsilon \right) \\ &= \mathbb{E} \left[\Pr \left(\left| \frac{1}{K+D} \sum_{k=0}^{K-1} (R_k - \bar{R}_k(\beta)) \right| \geq \epsilon \right) \right]. \end{aligned}$$

We then observe that, conditioned on (β, θ) , each R_k is an independent random variable with mean $\bar{R}_k(\beta)$. Furthermore, we observe that $\frac{1}{K+D} \sum_{k=0}^{K-1} R_k$ is readily upper bounded