Coding Across Heterogeneous Parallel Erasure Broadcast Channels is Useful

Sunghyun Kim
ETRI, South Korea
koishkim@etri.re.kr

Soheil Mohajer
University of Minnesota, USA
soheil@umn.edu

Changho Suh
KAIST, South Korea
chsuh@kaist.ac.kr

***Abstract*—Motivated by recent efforts to harness millimeter-wave (mmWave) bands, known to have high outage probabilities, we explore a K-user parallel packet-erasure broadcast channel that consists of orthogonal subchannels prone to packet-erasures. Our main result is two-fold. First, in the homogeneous channel where all subchannels have the same erasure probability, we show that the separation principle holds, i.e., coding across subchannels provides no gain. Second, in the heterogeneous channel where the subchannels have different erasure probabilities, we devise a scheme that employs coding across subchannels and show that the principle fails to hold, i.e., coding across subchannels provides a gain. Inspired by this finding, we demonstrate our scheme to be effective in harnessing the mmWave bands. Compared to the current approach in the 4G systems which allocates subchannels to users exclusively, we show that our scheme offers a huge gain. We find the gain to be significant in scenarios where the erasure probabilities are largely different, and importantly to increase with the growth of K. Our result calls for joint coding schemes in future wireless systems to meet growing mobile data demands.**

I. Introduction

Mobile data demands are on the rise at an increasing pace. To meet the growing demands, tapping into unexplored high-frequency bands is being actively considered. One may think that applying existing techniques in the established 4G systems will lead to its efficient use. However, a new challenge arises with high-frequency carriers.

The challenge is that signals conveyed by high-frequency carriers (called millimeter waves, mmWave for short) suffer from high outage probabilities because of their vulnerability to blockage [1]. In dense urban areas such as downtown Manhattan, empirical research on mmWave signals has found that the outage probability is around 0.34 and 0.65 for transmitter-receiver pairs located within 200m and 425m respectively [2]. Worse yet, it could be lower for pairs placed farther and for affordable devices whose signal reception quality is mediocre. Unless addressed, this will lead to disappointing performance.

To see this, let us consider a K-user parallel packet-erasure broadcast channel where the transmitter sends packets in M orthogonal subchannels and the receivers receive each packet with probability p[1]. In the 4G systems, a range of spectrum is divided into orthogonal bands and they are allocated to users exclusively. In the channel we consider, when the M subchannels are divided into K distinct sets and each set is allocated to one user, a packet transmission rate of Mp is achieved. One can observe that its performance is seemingly far from optimal. On the bright side, this approach features low implementation complexity. However, in terms of performance, it clearly does not live up to expectations.

The poor performance comes from the approach in the 4G systems which does not exploit multiuser diversity. In contrast, the scheme proposed in past works [3], [4] in the single-input packet-erasure broadcast channel exploits it. The scheme makes use of channel state feedback. From feedback, the transmitter knows which users have received which packets, and generates coded packets that can benefit multiple users at once (see Fig. 2). Applying it to each subchannel independently, one can verify that a rate of $M\frac{K}{\sum_{i=1}^{K}\frac{1}{1-(1-p)^i}}$ is achieved. There is a significant gain compared to the approach in the 4G systems. Apparently, such coded packets that can benefit multiple users at once are more valuable than packets intended for only one user. Inspired by this observation, we propose a better way to harness the mmWave bands in this work.

Instead of using them alone, we can combine them with the bands in the current 4G systems considered stable [5], building on a notion of carrier aggregation that combines discontiguous bands in serving users. To investigate this channel, we consider a K-user parallel packet-erasure broadcast channel where M and N subchannels are prone to erasures with probability $1-p$ and $1-q$ respectively for $p < q \leq 1$.

As our result, we develop a scheme that employs coding across subchannels, selectively sending packets that are more valuable in subchannels of better stability. We compare it to an extension of the state-of-the-art scheme that does not employ coding across subchannels (per-subchannel scheme). In the homogeneous channel where all subchannels have the same erasure probability, we show that coding across subchannels provides no gain. Interestingly, however, in the heterogeneous channel where the subchannels have different erasure probabilities, it is not the case. Our scheme turns out to outperform the per-subchannel scheme by a noticeable gap. It clearly shows that in broadcast channels with heterogeneity, *the separation principle fails to hold* and coding across subchannels is useful.

Our proposed scheme turns out to be effective in practical settings. In scenarios where erasures occur frequently in the

[1] In practice, forward error correction protocols are implemented to enable the receiving end to recover corrupted data in the received packet. However, it cannot be done if data corruption occurs too severely. To capture this, for simplicity, we introduce binary random states in this work (see Section II).

978-1-5090-4096-4/17/$31.00 ©2017 IEEE

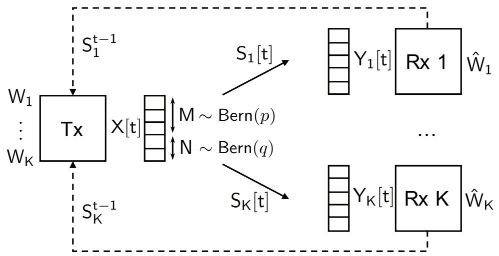

Fig. 1. K-user parallel packet-erasure broadcast channel.

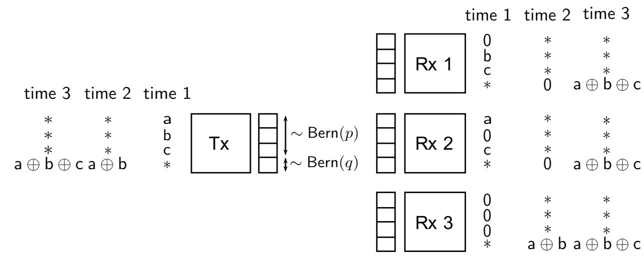

Fig. 2. A three-user example. Three subchannels are prone to erasures with probability $1 - p$ and one subchannel is prone to erasures with probability $1 - q$ for $p < q < 1$. To highlight elements that matter, some packets are not specified. Time 1 corresponds to block-B, and Times 2 and 3 correspond to block-$(B + 1)$ based on the block Markov model (see Fig. 3).

unstable bands and modestly in the stable bands, e.g., $K = 25$, $M = 1000$, $N = 100$, $p = 0.1$ and $q = 0.8$, the relative gain compared to the approach in the 4G systems reaches up to 294%, a four-fold increase (Fig. 5(a)). More importantly, the gain turns out to increase as K grows (Fig. 5(b)).

II. Problem Formulation

Fig. 1 describes the K-user parallel packet-erasure broadcast channel. Transmitter wishes to deliver message W_k reliably to Receiver k, $\forall k = 1, \ldots, K$. The channel consists of $M+N$ orthogonal erasure subchannels. Let $\boldsymbol{S}_k[t]$ be the states of subchannels from Transmitter to Receiver k as an $(M+N)$-by-$(M + N)$ diagonal matrix whose (i, i)-entries follow Bern(p) for $i = 1, \ldots, M$, Bern(q) for $i = M+1, \ldots, M+N$, and $p < q \leq 1$. The diagonal entries are assumed independent over i. Also, $\boldsymbol{S}_k[t]$ are assumed i.i.d. over k and t.

Let us elaborate more on the rationale behind our modeling of packet-erasures using binary random states. The current 4G systems transfer data in packets. When data-carrying signals are transmitted in unstable wireless channels, some data in the received packet can be corrupted. In practice, various mechanisms, such as forward error correction and retransmission, are implemented to recover the corrupted data at the receiving end. However, in case of severe data corruption, the mechanisms do not work and the received packet results in an irrecoverable decoding-failure. For the sake of simplicity, we represent such decoding-failures from severe data corruption with a binary random state with some fixed packet-erasure probability.

Let $X_k[t] \in \mathbb{C}^{M+N}$ be the sent packet of Transmitter at time t and let $Y_k[t] \in \mathbb{C}^{M+N}$ be the received packet of Receiver k at time t: $Y_k[t] = \mathbf{S}_k[t]X_k[t]$. We assume current erasure states are available at the corresponding receivers, and the transmitter gets feedback of *past* erasure states from the receivers.

In this paper, we are interested in the sum packet transmission rate: $\mathsf{Rate} = \sum_{k=1}^{K} R_k$, where R_k's can be simultaneously achieved for $k = 1, \ldots, K$. We aim to maximize Rate.

III. An Illustrative Example

We propose a coding scheme in detail in Section IV. Before we describe it, let us begin with an illustrative example in Fig. 2, which highlights key elements of our proposed scheme.

- Time 1: the transmitter sends three fresh packets, one intended for each user, in the three unstable subchannels. A particular event takes place: none are received by the intended users. a and b are received by one unintended user, and c is received by two. For later use, they store the received packets as side information. Assuming channel state feedback is available, the transmitter knows which packets have been received which users.
- Time 2: the transmitter sends a coded packet $a \oplus b$ in the stabler subchannel. Notice that if the coded packet is received by Users 1 and 2, it will lead to *two packets decoded per transmission*. User 1 decodes a by exploiting its side information b, and User 2 decodes b likewise. Unfortunately, it is received by User 3 only, and not received by the intended Users 1 and 2. For later use, User 3 stores its received packet as side information.
- Time 3: the transmitter sends a different coded packet $a \oplus b \oplus c$ in the stabler subchannel. Notice that if the coded packet is received by Users 1, 2 and 3, it will lead to *three packets* decoded per transmission. Fortunately, the coded packet is received by all users. The users decode their desired packet by exploiting their side information.

Note two key coding strategies at play.

- The first is to seek maximal gains attainable by exploiting side information. At any moment, the transmitter generates coded packets in a way that they can benefit as many users as possible per transmission. At Time 2, it attempts to achieve two packets decoded at once by sending $a \oplus b$. At time 3, it attempts to achieve three packets decoded at once by sending $a \oplus b \oplus c$.
- The second is to send coded packets, which are worth multiple packets in effect, in stabler subchannels. The reason why the coded packets are sent in this manner is clear. It is to maximize the odds of the delivery of the coded packets that are more valuable than fresh packets.

The first strategy can be found in past works [3], [4], [6] by which our scheme is motivated. In particular, the work of [6] has devised a systematic way of coding packets by exploiting channel state feedback at the transmitter, introducing a concept of *order*. By order-i packets, it means that i such packets, each intended for a distinct user, can be coded into a single packet that can benefit up to i users at once. In the example, a and b are once order-2, and later become order-3, together with c. The order of a packet can change over time, and such transitions are analyzed in detail in [6].

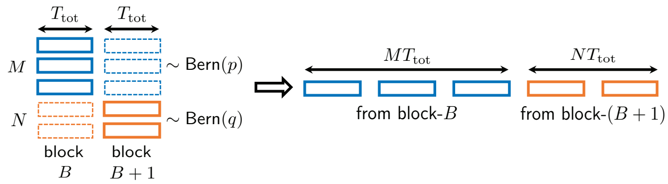

Fig. 3. Graphical illustration that depicts how we interpret channel uses. We consider the block Markov model.

The second strategy can be linked to decentralized caching [7] operating in two phases. In the placement phase, without coordination, the users store random fragments of other users' contents in their cache for later utility. In the delivery phase, encoded contents are transmitted in a shared link and multiple users can decode their desired contents by exploiting the side information (network coding at play [8]–[10]). Our scheme can be viewed as the two phases taking place concurrently across multiple subchannels. While placement (occurring randomly) is done mostly in unstable subchannels, delivery (aimed to exploit side information) is done mostly in stabler subchannels.

Before we end this section, let us leave a final remark. In the literature, the idea to kill multiple birds with one stone is often called the butterfly effect. One coded packet (or symbol) is in effect worth multiple packets (or symbols) per transmission upon reception at multiple users. In all past works [3], [4], [6] that motivate our scheme, the butterfly effect played a central role. The work of [6] has devised a systematic way to achieve such effects building on the concept of order in a Gaussian broadcast channel. The works of [3], [4] have sought to benefit from such effects by applying coding over time in a single-input packet-erasure broadcast channel.

Due to its large influence, the butterfly effect can be found in a variety of contexts: caching [7], [11], cache-aided networks [3], [12], network coding [8]–[10], delayed channel state information at the transmitter [4], [6], feedback [13] and bursty channels with relays [14], [15] to name a few. In all contexts, exploiting side information is at the heart of notable gains.

IV. Coding Scheme and Achievable Rate

In this section, building on the key elements presented in the example, we describe our proposed scheme in detail.

First, we explain how we interpret the channel. We adopt the block Markov model as depicted in Fig. 3. Each block is of length $T_{\rm tot}$, sufficiently large to ensure the law of large numbers to be valid. In our scheme, the transmitter exploits channel state feedback to generate highest-order coded packets possible. Moreover, it sends as many high-order coded packets as possible in stabler subchannels to increase the number of packets decoded per transmission. To perform these two, it exploits feedback from block-B, and uses the packets sent in block-B in generating coded packets to send in block-$(B+1)$.

Now, we explain how coded packets are generated in detail. We revisit the example in Section III. For clarity, we introduce some notation. Let $\mathcal{E}_{k,\mathcal{U}}$ be the set of packets intended for User k but received by the users in $\mathcal{U}$. In block-B, the transmitter divides $MT_{\rm tot}$ p-channel uses (unstable channel uses) equally

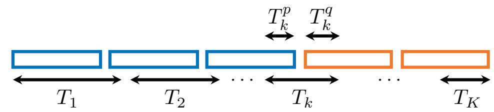

Fig. 4. Graphical illustration of our scheme. We have higher chances of delivering packets using q-channel uses. We send high-order packets using q-channel uses to achieve a large number of packets decoded per transmission.

into three and uses each to send fresh packets for each user. For User 1 (by symmetry, the same applies to Users 2 and 3),

- $(MT_{\rm tot}/3)\times p$ packets are decoded at User 1.
- $(MT_{\rm tot}/3)\times p(1-p)^2$ packets (expected size of $\mathcal{E}_{1,2}$ and $\mathcal{E}_{1,3}$) are received by one unintended user, and not by User 1. They are of order-2.
- $(MT_{\rm tot}/3)\times p^2(1-p)$ packets (expected size of $\mathcal{E}_{1,\{2,3\}}$) are received by two unintended users, and not by User 1. They are of order-3.

From channel state feedback, the transmitter knows which packets have been received by which users in the previous blocks. For the packets in $\mathcal{E}_{1,2}$ and $\mathcal{E}_{2,1}$ sent in block-B, it can identify the packets intended for User 1 but received by User 2 (a's) and vice versa (b's). Then, in block-$(B+1)$, it attempts to send the XORs of the packets in $\mathcal{E}_{1,2}$ and $\mathcal{E}_{2,1}$ ($a\oplus b$'s) in q-channel uses (stabler erasure subchannel uses). If the coded packets are received by Users 1 and 2, both users can decode their desired packets by exploiting side information properly. This leads to *two packets decoded per channel use*. By symmetry, the same can happen for the other pairs of users.

The transmitter can apply the same technique to the packets in $\mathcal{E}_{1,\{2,3\}}$, $\mathcal{E}_{2,\{3,1\}}$ and $\mathcal{E}_{3,\{1,2\}}$. From feedback, it can identify the packets intended for User 1 but received by Users 2 and 3 (a's) and the like (b's and c's) in block-B. In block-$(B+1)$, it attempts to send the XORs of the packets in $\mathcal{E}_{1,\{2,3\}}$, $\mathcal{E}_{2,\{3,1\}}$ and $\mathcal{E}_{3,\{1,2\}}$ ($a\oplus b\oplus c$'s) in q-channel uses. This can lead to *three packets* decoded per channel use.

In Section III, we provide an example where order-2 packets can become order-3. We extend it to a general K-user case.

Suppose we send L_1 fresh packets intended for each user (KL_1 fresh packets in total) during $T_{\rm tot}$ time slots, using $MT_{\rm tot}$ p-channel and $NT_{\rm tot}$ q-channel uses. We send as many high-order coded packets as possible in q-channel uses. Coded packets of some order may be sent in both p-channel and q-channel uses. Let k be this order (see Fig. 4). A fraction λ of order-k coded packets are sent in T_k^p p-channel uses, and the remaining fraction of them in T_k^q q-channel uses. Let $f=\frac{L_1}{T_{\rm tot}}$. We have three parameters we can choose: f, k and λ.

Packets of order-i can be promoted to order-j for $i<j$. Let $\alpha_{i\to j}(p)$ be the probability that an order-i packet is not received by the intended user but received by $j-i$ additional unintended users in a p-channel use, and becomes order-j [3]:

$$\alpha_{i\to j}(p) := p^{j-i}(1-p)^{K-j+1}. \tag{1}$$

Also, let $\beta_j(p)$ be the probability that an order-j packet is either received by the intended user, or promoted to a higher-order packet and handled with the higher-order packets [3]:

$$\beta_j(p) := 1-(1-p)^{K-j+1}. \tag{2}$$

Note that upward order transitions take place. Let $L_{i\to j}$ be the number of order-j packets intended for a subset of j users generated from order-i packets intended for a subset of i users:

$$L_{i\to j} = \begin{cases} t_i\alpha_{i\to j}(p), & i<k; \\ t_i\alpha_{i\to j}(q), & i>k; \\ t_k^p\alpha_{k\to j}(p) + t_k^q\alpha_{k\to j}(q), & i=k. \end{cases} \quad (3)$$

Here, t_j is the number of channel uses required to handle the order-j coded packets for a subset of j users:

$$t_j = \begin{cases} \frac{1}{\beta_j(p)}\sum_{i=1}^{j-1}\binom{j-1}{i-1}L_{i\to j}, & j<k; \\ \frac{1}{\beta_j(q)}\sum_{i=1}^{j-1}\binom{j-1}{i-1}L_{i\to j}, & j>k; \\ t_k^p + t_k^q, & j=k, \end{cases} \quad (4)$$

where t_k^p and t_k^q are given as:

$$t_k^p = \frac{\lambda}{\beta_k(p)}\sum_{i=1}^{k-1}\binom{k-1}{i-1}L_{i\to k}, \quad (5)$$

$$t_k^q = \frac{1-\lambda}{\beta_k(q)}\sum_{i=1}^{k-1}\binom{k-1}{i-1}L_{i\to k}. \quad (6)$$

$\beta_j(p)$ and $\beta_j(q)$ appear because to deliver a certain number of packets, we need a larger number of channel uses due to erasures. Now, let us explain the binomial coefficient and the summation. Fix j and consider a subset $\mathcal{U}_j$ of j users. Pick User $u \in \mathcal{U}_j$. For User u, order-j packets are packets intended for User u that have not been received by User u but stored at the $j-1$ unintended users in $\mathcal{U}_j \setminus u$. To compute how many subsets $\mathcal{U}_i$'s of i users from which order-j packets for User u are generated, we exclude User u for later inclusion and pick $i-1$ users among the remaining $j-1$ users in $\mathcal{U}_j$ for $i<j$.

We need initial conditions to compute (3) and (4). For $j=1$, we set $L_{i\to j}$ as L_1, and t_j as follows:

$$t_1 = \begin{cases} \frac{1}{\beta_1(p)}L_1, & k>1; \\ \frac{\lambda}{\beta_1(p)}L_1 + \frac{1-\lambda}{\beta_1(q)}L_1, & k=1. \end{cases} \quad (7)$$

Let T_j be the number of channel uses required to handle all order-j coded packets (fresh packets for $j=1$). There are $\binom{K}{j}$ ways to construct subsets of j users.

$$T_j = \binom{K}{j}t_j. \quad (8)$$

We set f, k and λ, the three parameters we can choose, in a way that the two conditions are met to use all channel uses:

$$\sum_{j=1}^{k-1}T_j + \binom{K}{k}t_k^p = MT_{\text{tot}}, \ \sum_{j=k+1}^{K}T_j + \binom{K}{k}t_k^q = NT_{\text{tot}}. \quad (9)$$

Let us denote the set parameters by f^*, k^* and λ^*. We send f^*T_{tot} fresh packets for each user during T_{tot} time slots, and all of them are decoded. This gives us the following theorem.

Theorem 1: An achievable sum packet transmission rate in the K-user parallel packet-erasure broadcast channel by applying coding across subchannels is

$$\mathsf{Rate}_{\mathsf{proposed}} = Kf^*,$$

where f^* is the solution of the equations from (1) to (9).

In the homogeneous channel where the erasure probabilities are identical, schemes that code across subchannels does not lead to larger gains than those do not. The following theorem describes the optimal sum rate in the homogeneous channel.

Theorem 2: The optimal sum packet transmission rate in the K-user M-parallel packet-erasure broadcast channel that consists of M orthogonal subchannels susceptible to packet-erasures with probability $1-p$ is

$$\mathsf{Rate}_{\mathsf{optimal}} = M\frac{K}{\sum_{i=1}^{K}\frac{1}{1-(1-p)^i}}.$$

Proof: The complete proof requires detailed arguments and is omitted. We provide a less rigorous one instead. Let C_M be the capacity of the parallel channel and C_1 be that of the single channel ($M=1$). Suppose $C_M > MC_1$. In the parallel channel, all channel uses have the same erasure probability over time and frequency. Thus, over M time slots in the single channel, we could send M packets an optimal scheme in the parallel channel would send across M subchannels, achieving C_M/M. This implies there exists a scheme that achieves a rate above C_1. This is a contradiction. Therefore, $C_M = MC_1$ and $C_1 = K/\sum_{i=1}^{K}\frac{1}{1-(1-p)^i}$ can be shown from [3], [4], [16]. ■

Remark: The works of [3], [4] have developed an optimal scheme in the K-user single-input channel. It generates coded packets of multiorder from feedback [6] to exploit multiuser diversity. Extending it in a per-subchannel manner is optimal in the homogeneous channel (Theorem 2), but in the heterogeneous channel, it is not. Coding *across* subchannels has been shown to achieve optimality for $K=2$ from [17], [18], and more importantly, it is shown to provide a significant scalable gain in this work for $K \geq 2$. The suboptimality of the per-subchannel extension arises because channel characteristics are not exploited to the fullest. Not all packets are equal. As coded packets are in effect worth multiple packets, one can expect it to be more useful to send them in stabler subchannels to increase the number of packets decoded per transmission. Thus, we specifically devise our scheme to send as many coded packets as possible in stabler subchannels, to make it gainful as well as operationally simple. Its optimality, however, is open.

V. Simulation Results

To examine the gains attained by our scheme, we consider two existing schemes. In the naive scheme, the transmitter assigns subchannels to users exclusively, and sends packets for each user only in the subchannels assigned to it. This approach used in the 4G systems does not exploit multiuser diversity.

$$\mathsf{Rate}_{\mathsf{naive}} = Mp + Nq.$$

In the per-subchannel scheme, the transmitter applies coding over time per subchannel to exploit multiuser diversity. This is an extension of the scheme in [3], [4] in a per-subchannel manner (no coding across subchannels), which is an optimal scheme in the homogeneous channel (Theorem 2).

$$\mathsf{Rate}_{\mathsf{per\text{-}sub}} = M\frac{K}{\sum_{i=1}^{K}\frac{1}{1-(1-p)^i}} + N\frac{K}{\sum_{i=1}^{K}\frac{1}{1-(1-q)^i}}.$$

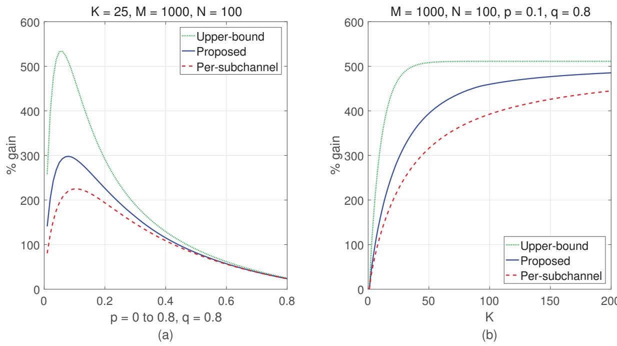

Fig. 5. (a) relative gain v.s. p: $K = 25$, $M = 1000$, $N = 100$, $q = 0.8$; and (b) relative gain v.s. K: $M = 1000$, $N = 100$, $p = 0.1$, $q = 0.8$.

We evaluate the gain attained by the proposed scheme from two different perspectives. One is that we compare it to the naive scheme in order to see the benefit of advanced coding schemes over the approach in the 4G systems. To this end, we consider the following relative gain:

$$\mathsf{Gain}_{\mathsf{proposed}} = \frac{\mathsf{Rate}_{\mathsf{proposed}} - \mathsf{Rate}_{\mathsf{naive}}}{\mathsf{Rate}_{\mathsf{naive}}} \times 100.$$

The other is that we compare it to the per-subchannel scheme that does not apply coding across subchannels in order to see the additional benefit attainable by coding across subchannels. To this end, likewise, we consider the following relative gain:

$$\mathsf{Gain}_{\mathsf{per\text{-}sub}} = \frac{\mathsf{Rate}_{\mathsf{per\text{-}sub}} - \mathsf{Rate}_{\mathsf{naive}}}{\mathsf{Rate}_{\mathsf{naive}}} \times 100.$$

From $\mathsf{Gain}_{\mathsf{proposed}}$ and $\mathsf{Gain}_{\mathsf{per\text{-}sub}}$, we can see (1) how large gains advanced coding schemes offer compared to the approach in the 4G systems; (2) how useful applying coding across subchannels is compared to the per-subchannel scheme.

Fig. 5(a) depicts the relative gains with respect to p for $K = 25$, $M = 1000$, $N = 100$ and $q = 0.8$. As the range of spectrum is far wider in the mmWave bands than in the 4G systems bands, we set M to be ten times N. The (blue) solid curve plots $\mathsf{Gain}_{\mathsf{proposed}}$, the (red) dashed curve $\mathsf{Gain}_{\mathsf{per\text{-}sub}}$, and the (green) dotted curve the cut-set bound on the relative gain. This setting could represent scenarios where cells with small coverage serve a reasonable number of users (e.g., pico cells). One can see that for fairly large p (say 0.6), our scheme can achieve a reasonable gain of 57%. In adverse scenarios, that is for small p (say 0.1), our scheme can achieve a huge gain of 293%, a four-fold increase, with a noticeable gap compared to a gain of 225% achieved by the per-subchannel scheme.

Fig. 5(b) depicts the relative gains with respect to K for adverse scenarios: $M = 1000$, $N = 100$, $p = 0.1$ and $q = 0.8$. One can see that for a handful of users ($K > 5$), our scheme can achieve at least a two-fold gain of 104%. Promisingly, as K increases, the gain enlarges. It reaches a five-fold gain of 394% at $K = 50$. Note a gap compared to the per-subchannel scheme where no coding across subchannels is employed.

VI. Conclusion

We explored the K-user parallel erasure broadcast channel where some subchannels are prone to erasures while some are less so. We found that coding across subchannels can provide significant gains and achieve better performance than the per-subchannel scheme. This result shows that in heterogeneous parallel broadcast channels, the separation principle fails to hold.

Acknowledgment

The work of S. Kim was supported by the Institute for Information & communications Technology Promotion (IITP) grant (No. 2014-0-00282) funded by the Korean government (MSIP; Ministry of Science, ICT & Future Planning). The work of C. Suh was supported by the National Research Foundation of Korea (NRF) grant (No. 2015R1C1A1A02036561) funded by the Korean government (MSIP).

References

[1] R. W. Heath, N. González-Prelcic, and S. Rangan, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, pp. 436–453, Apr. 2016.
[2] T. S. Rappaport, G. R. MacCartney, and M. K. Samimi, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, pp. 3029–3056, Sept. 2015.
[3] A. Ghorbel, M. Kobayashi, and S. Yang, "Cache-enabled broadcast packet erasure channels with state feedback," *Allerton Conf. on Commun., Control, and Computing*, Oct. 2015.
[4] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser broadcast erasure channel with feedback — capacity and algorithms," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5779 – 5804, Sept. 2013.
[5] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, pp. 366–385, Mar. 2014.
[6] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, July 2012.
[7] M. A. Maddah-Ali and U. Niesen, "Decentralized coded caching attains order-optimal memory-rate tradeoff," *IEEE/ACM Trans. Netw.*, vol. 23, pp. 1029–1040, Aug. 2015.
[8] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.
[9] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.
[10] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," *IEEE/ACM Trans. Netw.*, vol. 16, pp. 497–510, June 2008.
[11] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, pp. 2856–2867, May 2014.
[12] K. Shanmugam, M. Ji, A. Tulino, J. Llorca, and A. Dimakis, "Finite-length analysis of caching-aided coded multicasting," *IEEE Trans. on Inf. Theory*, no. 10, pp. 5524–5537, Oct. 2016.
[13] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.
[14] S. Kim, I.-H. Wang, and C. Suh, "A relay can increase degrees of freedom in bursty interference networks," *IEEE Int. Symp. on Inform. Theory*, June 2015.
[15] S. Kim, S. Mohajer, and C. Suh, "Role of a relay in bursty networks with correlated transmissions," *IEEE Int. Symp. on Inform. Theory*, Aug. 2016.
[16] A. El Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 379–381, May 1978.
[17] C.-C. Wang and J. Han, "The capacity region of two-receiver multiple-input broadcast packet erasure channels with channel output feedback," *IEEE Trans. on Inf. Theory*, no. 9, pp. 5597–5626, Sept. 2014.
[18] M. Heindlmaier, N. Reyhanian, and S. S. Bidokhti, "On capacity regions of two-receiver broadcast packet erasure channels with feedback and memory," *Allerton Conf. on Commun., Control, and Computing*, Sept./Oct. 2014.