Rate Splitting for General Multicast

Presenter: Lingzhi Zhao
Author: Lingzhi Zhao*, Ying Cui*, Sheng Yang†,
Shlomo Shamai (Shitz)‡, Yunbo Han§, and Yunfei Zhang§

Shanghai Jiao Tong University* Paris-Saclay University†

Technion–Israel Institute of Technology‡ Tencent Technology§

IEEE ICC 16–20, May 2022

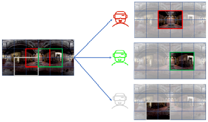
Outline

Outline

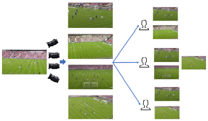
1. Introduction
2. System Model
3. Problem Formulation and Solution
4. Numerical Results
5. Conclusion

Motivation

- Conventional mobile Internet services can be supported by unicast, single-group multicast, and multi-group multicast
    - e.g., (traditional) video, audio, web browsing, social networking, etc.
- Immersive video cannot perfectly adapt to the conventional transmission schemes
    - e.g., 360 video [TWC'21] and multi-view video [TCOM'20]
    - Play an important role in online gaming and cloud meeting, etc.
    - Multiple messages are transmitted to each user, and one message may be intended for multiple users [TWC'21,TCOM'20]
- This motivates us to consider general multicast
    - One message can be intended for any user
    - Include the three conventional transmission schemes as special cases
    - Play a central role for future 6G and beyond networks



(a) 360 video



(b) MVV

Previous Work

- Adopt orthogonal transmission to convert general multicast in immersive video streaming to per resource block single-group multicast
    - Wireless streaming of a 360 video [TWC'21] and an MVV [TCOM'20]
    - Easy to implement, but spatial multiplexing gain is not exploited
- Non-orthogonal transmission achieves higher transmission efficiency
    - The cost to suppress interference in SDMA can be high
    - Decoding interference in NOMA may not be possible
- Rate splitting partially suppresss interference and partially decodes interference
    - Unicast [TIT'81,TIT'13,JSAC'21], unicast together with a multicast message [TCOM'19], multi-group multicast [TVT'20]
    - Optimization of rate splitting for unicast and its slight generalizations cannot apply to general multicast
    - Rate splitting for general multicast for discrete memoryless broadcast channels [ISIT'17] from an information theory perspecitve
- Optimize general rate splitting for general multicast with linear beamforming

Outline

1. Introduction
2. System Model
3. Problem Formulation and Solution
4. Numerical Results
5. Conclusion

General Multicast

- Consider a single-cell wireless network consisting of one BS and K users, where the BS has I independent messages
  - Let $\mathcal{K} \triangleq \{1, \dots, K\}$ denote the set of user indices
  - Let $\mathcal{I} \triangleq \{1, \dots, I\}$ denote the set of I messages
- Consider general multicast
  - Each user $k \in \mathcal{K}$ can request arbitrary I_k messages in $\mathcal{I}$, denoted by $\mathcal{I}_k \subseteq \mathcal{I}$, from the BS
  - Each message in $\mathcal{I}$ is requested by at least one user, i.e., $\cup_{k \in \mathcal{K}} \mathcal{I}_k = \mathcal{I}$
- Partition the message set $\mathcal{I}$ according to the requests from K users
  - $\mathcal{P}_{\mathcal{S}} \triangleq \left(\bigcap_{k \in \mathcal{S}} \mathcal{I}_k\right) \bigcap \left(\mathcal{I} - \bigcup_{k \in \mathcal{K} \setminus \mathcal{S}} \mathcal{I}_k\right)$ denotes the set of messages that is requested by each user in $\mathcal{S}$ and not requested by any user in $\mathcal{K} \setminus \mathcal{S}$
  - $\boldsymbol{\mathcal{P}} \triangleq \{\mathcal{P}_{\mathcal{S}} | \mathcal{P}_{\mathcal{S}} \neq \emptyset, \mathcal{S} \subseteq \mathcal{K}, \mathcal{S} \neq \emptyset\}$ forms a partition of $\mathcal{I}$
    - Refer to each element in $\boldsymbol{\mathcal{P}}$ as a message unit
  - $\boldsymbol{\mathcal{S}} \triangleq \{\mathcal{S} | \mathcal{P}_{\mathcal{S}} \neq \emptyset, \mathcal{S} \subseteq \mathcal{K}, \mathcal{S} \neq \emptyset\}$ specifies the user groups corresponding to $\boldsymbol{\mathcal{P}}$

Illustration Example

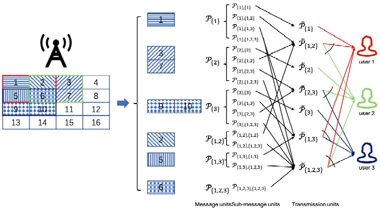


Figure: $K = 3$, $I = 8$, $\mathcal{I}_1 = \{1, 2, 5, 6\}$, $\mathcal{I}_2 = \{2, 3, 6, 7\}$, $\mathcal{I}_3 = \{5, 6, 9, 10\}$, $\mathcal{P}_{\{1\}} = \{1\}$, $\mathcal{P}_{\{2\}} = \{3, 7\}$, $\mathcal{P}_{\{3\}} = \{9, 10\}$, $\mathcal{P}_{\{1,2\}} = \{2\}$, $\mathcal{P}_{\{1,3\}} = \{5\}$, $\mathcal{P}_{\{1,2,3\}} = \{6\}$,
$\boldsymbol{\mathcal{P}} = \{\mathcal{P}_{\{1\}}, \mathcal{P}_{\{2\}}, \mathcal{P}_{\{3\}}, \mathcal{P}_{\{1,2\}}, \mathcal{P}_{\{1,3\}}, \mathcal{P}_{\{1,2,3\}}\}$,
$\boldsymbol{\mathcal{S}} = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}$

Remark (Connection with Unicast and Multicast)

- When $I = K, I_k = 1$, and $\mathcal{I}_k \neq \mathcal{I}_{k'}, k \neq k'$, general multicast reduces to unicast
- When $I = 1$, implying $I_k = 1$, and $\mathcal{I}_k = \mathcal{I}_{k'}, k \neq k'$, general multicast becomes single-group multicast
- When $1 < I < K$ and $I_k = 1$, general multicast reduces to multi-group (I-group) multicast

General Rate Splitting

- Consider rate splitting in the most general form for general multicast
  - Each user group decodes not only the desired message unit $\mathcal{P}_{\mathcal{S}}$ but also part of the message unit of any other user group $\mathcal{P}_{\mathcal{S}'}$, $\mathcal{S}' \neq \mathcal{S}, \mathcal{S}' \in \boldsymbol{\mathcal{S}}$
- Split each message unit $\mathcal{P}_{\mathcal{S}}$ into $2^{K-|\mathcal{S}|}$ sub-message units

$$\mathcal{P}_{\mathcal{S}} = \prod\nolimits_{\mathcal{G} \in \boldsymbol{\mathcal{G}}_{\mathcal{S}}} \mathcal{P}_{\mathcal{S},\mathcal{G}}, \ \mathcal{S} \in \boldsymbol{\mathcal{S}}$$

  - $\boldsymbol{\mathcal{G}}_{\mathcal{S}} \triangleq \{\mathcal{X} | \mathcal{S} \subseteq \mathcal{X} \subseteq \mathcal{K}\}$ collects all $2^{K-|\mathcal{S}|}$ subsets of $\mathcal{K}$ that contain $\mathcal{S}$
  - Define $\boldsymbol{\mathcal{G}} \triangleq \bigcup_{\mathcal{S} \in \boldsymbol{\mathcal{S}}} \boldsymbol{\mathcal{G}}_{\mathcal{S}}$
- The rate of the message unit $\mathcal{P}_{\mathcal{S}}$ is split into the rates of the $2^{K-|\mathcal{S}|}$ sub-message units $\mathcal{P}_{\mathcal{S},\mathcal{G}}, \mathcal{G} \in \boldsymbol{\mathcal{G}}_{\mathcal{S}}$

$$R_{\mathcal{S}} = \sum\nolimits_{\mathcal{G} \in \boldsymbol{\mathcal{G}}_{\mathcal{S}}} R_{\mathcal{S},\mathcal{G}}, \ \mathcal{S} \in \boldsymbol{\mathcal{S}}$$

- Re-assemble the sub-message units $\mathcal{P}_{\mathcal{S},\mathcal{G}}$ to form a transmission unit $\widetilde{\mathcal{P}}_{\mathcal{G}}$ with rate:

$$\widetilde{R}_{\mathcal{G}} = \sum\nolimits_{\mathcal{S} \in \boldsymbol{\mathcal{S}}_{\mathcal{G}}} R_{\mathcal{S},\mathcal{G}}, \ \mathcal{G} \in \boldsymbol{\mathcal{G}}$$

  - $\boldsymbol{\mathcal{S}}_{\mathcal{G}} \triangleq \{\mathcal{S} \in \boldsymbol{\mathcal{S}} | \mathcal{S} \subseteq \mathcal{G}\}$

Illustration Example

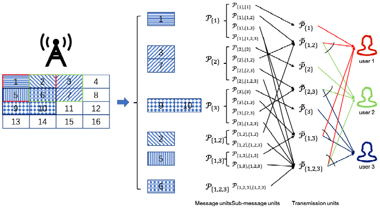


Figure: $\mathcal{G}_{\{1\}} = \{\{1\}, \{1,2\}, \{1,3\}, \{1,2,3\}\}$, $\mathcal{G}_{\{2\}} = \{\{2\}, \{1,2\}, \{2,3\}, \{1,2,3\}\}$, $\mathcal{G}_{\{3\}} = \{\{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$, $\mathcal{G}_{\{1,2\}} = \{\{1,2\}, \{1,2,3\}\}$, $\mathcal{G}_{\{1,3\}} = \{\{1,3\}, \{1,2,3\}\}$, $\boldsymbol{\mathcal{G}} = \{\{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$

Remark (Connection with Rate Splitting for Unicast and Multicast)

- When general multicast degrades to unicast, the proposed general rate splitting reduces to the general rate splitting for unicast [JSAC'21]
- When general multicast degrades to single-group multicast, the proposed general rate splitting reduces to the conventional single-group multicast transmission
- When general multicast degrades to multi-group multicast, the proposed general rate splitting reduces to 1-layer rate splitting for multi-group multicast [TVT'20]

Physical Layer Model and Implementation

- The BS is equipped with M antennas and each user has one antenna.
- Consider a multi-carrier system
  - N and $\mathcal{N} \triangleq \{1, 2, \ldots, N\}$ denote the number of subcarriers and the set of subcarrier indices, respectively
  - The bandwidth of each subcarrier is B (in Hz)
- Consider a discrete-time system
  - Time is divided into fixed-length slots
- Adopt the block fading model
  - For each user and subcarrier, the channel remains constant within each slot and is i.i.d. over slots
- Consider slow fading and study an arbitrary slot
- $\mathbf{h} \triangleq (\mathbf{h}_{k,n})_{k\in\mathcal{K},n\in\mathcal{N}} \in \mathbb{C}^{M\times 1}$ denotes the system channel state
  - Assume that user $k \in \mathcal{K}$ knows his channel state $\mathbf{h}_k \triangleq (\mathbf{h}_{k,n})_{n\in\mathcal{N}}$ and the system channel state $\mathbf{h}$ is known to the BS

Physical Layer Model and Implementation

- For all $\mathcal{G} \in \boldsymbol{\mathcal{G}}$, transmission unit $\widetilde{\mathcal{P}}_{\mathcal{G}}$ is encoded (channel coding) into codewords that span over the N subcarriers
- Consider linear beamforming and use superposition coding
- The transmitted signal on subcarrier n

$$\mathbf{x}_n = \sum\nolimits_{\mathcal{G} \in \boldsymbol{\mathcal{G}}} \mathbf{w}_{\mathcal{G},n} s_{\mathcal{G},n}, \ \ n \in \mathcal{N}$$

  - $s_{\mathcal{G},n}$ denotes a symbol for $\widetilde{\mathcal{P}}_{\mathcal{G}}$ that is transmitted on the n-th subcarrier
    - For all $n \in \mathcal{N}$, let $\mathbf{s}_n \triangleq (s_{\mathcal{G},n})_{\mathcal{G} \in \boldsymbol{\mathcal{G}}}$ and assume that $\mathbb{E}[\mathbf{s}_n \mathbf{s}_n^H] = \mathbf{I}$
  - $\mathbf{w}_{\mathcal{G},n} \in \mathbb{C}^{M \times 1}$ denotes the beamforming vector for transmitting $\widetilde{\mathcal{P}}_{\mathcal{G}}$ on subcarrier n
- The total transmission power constraint

$$\sum_{n \in \mathcal{N}} \sum_{\mathcal{G} \in \boldsymbol{\mathcal{G}}} \|\mathbf{w}_{\mathcal{G},n}\|_2^2 \leq P$$

  - P denotes the transmission power budget

Physical Layer Model and Implementation

- The received signal at user $k \in \mathcal{K}$ on subcarrier $n \in \mathcal{N}$

$$y_{k,n} = \mathbf{h}_{k,n}^H \mathbf{x}_n + z_{k,n} = \mathbf{h}_{k,n}^H \sum\nolimits_{\mathcal{G} \in \boldsymbol{\mathcal{G}}^{(k)}} \mathbf{w}_{\mathcal{G},n} s_{\mathcal{G},n} + \mathbf{h}_{k,n}^H \sum\nolimits_{\mathcal{G}' \in \boldsymbol{\mathcal{G}} \setminus \boldsymbol{\mathcal{G}}^{(k)}} \mathbf{w}_{\mathcal{G}',n} s_{\mathcal{G}',n} + z_{k,n}, \ k \in \mathcal{K}, \ n \in \mathcal{N}$$

  - $\boldsymbol{\mathcal{G}}^{(k)} \triangleq \{\mathcal{G} \in \boldsymbol{\mathcal{G}} | k \in \mathcal{G}\}, k \in \mathcal{K}$
  - $z_{k,n} \sim \mathcal{CN}(0, \sigma^2)$ is AWGN
- Consider joint decoding at each user
  - Each user $k \in \mathcal{K}$ jointly decodes the desired transmission units $\widetilde{\mathcal{P}}_{\mathcal{G}}$
- The achievable rate region of the transmission units

$$\sum_{\mathcal{G} \in \boldsymbol{\mathcal{X}}} \widetilde{R}_{\mathcal{G}} \leq B \sum_{n \in \mathcal{N}} \log_2 \left(1 + \frac{\sum_{\mathcal{G} \in \boldsymbol{\mathcal{X}}} |\mathbf{h}_{k,n}^H \mathbf{w}_{\mathcal{G},n}|^2}{\sigma^2 + \sum_{\mathcal{G}' \in \boldsymbol{\mathcal{G}} \setminus \boldsymbol{\mathcal{G}}^{(k)}} |\mathbf{h}_{k,n}^H \mathbf{w}_{\mathcal{G}',n}|^2} \right), \quad \boldsymbol{\mathcal{X}} \subseteq \boldsymbol{\mathcal{G}}^{(k)}, k \in \mathcal{K}$$

Outline

1. Introduction
2. System Model
3. Problem Formulation and Solution
4. Numerical Results
5. Conclusion

Weighted Sum Rate Maximization

- Optimization variables:
  - Transmission beamforming vectors $\mathbf{w} \triangleq (\mathbf{w}_{\mathcal{G},n})_{\mathcal{G}\in\boldsymbol{\mathcal{G}},n\in\mathcal{N}}$
  - Rates of the sub-message units $\mathbf{R} \triangleq (R_{\mathcal{S},\mathcal{G}})_{\mathcal{S}\in\boldsymbol{\mathcal{S}},\mathcal{G}\in\boldsymbol{\mathcal{G}}}$
- Objective function:
  - Weighted sum rate $\sum_{\mathcal{S}\in\boldsymbol{\mathcal{S}}} \alpha_{\mathcal{S}} R_{\mathcal{S}}$
- Optimization constraints:
  - Total transmission power constraint
  - Achievable rate constraints

Problem 1 (Weighted Sum Rate Maximization)

$$\max_{\mathbf{w},\mathbf{R}\succeq 0} \quad \sum_{\mathcal{S}\in\boldsymbol{\mathcal{S}}} \alpha_{\mathcal{S}} R_{\mathcal{S}}$$

$$\text{s.t.} \quad \sum_{n\in\mathcal{N}} \sum_{\mathcal{G}\in\boldsymbol{\mathcal{G}}} \|\mathbf{w}_{\mathcal{G},n}\|_2^2 \le P$$

$$\sum_{\mathcal{G}\in\boldsymbol{\mathcal{X}}} \widetilde{R}_{\mathcal{G}} \le B \sum_{n\in\mathcal{N}} \log_2\left(1 + \frac{\sum_{\mathcal{G}\in\boldsymbol{\mathcal{X}}} |\mathbf{h}_{k,n}^H \mathbf{w}_{\mathcal{G},n}|^2}{\sigma^2 + \sum_{\mathcal{G}'\in\boldsymbol{\mathcal{G}}\setminus\boldsymbol{\mathcal{G}}^{(k)}} |\mathbf{h}_{k,n}^H \mathbf{w}_{\mathcal{G}',n}|^2}\right), \ \boldsymbol{\mathcal{X}} \subseteq \boldsymbol{\mathcal{G}}^{(k)}, k \in \mathcal{K}$$

Weighted Sum Rate Maximization

Remark (Connection with Rate Splitting for Unicast and Multicast)

- When general multicast degrades to unicast, Problem 1 reduces to the weighted sum rate maximization problem for general rate splitting for unicast in [JSAC'21]
- When general multicast degrades to single-group multicast, Problem 1 reduces to the rate maximization problem for single-group multicast in [TSP'06]
- When general multicast degrades to multi-group multicast, Problem 1 can be viewed as a generalization of the weighted sum rate maximization for multi-group multicast in [TVT'20]

- Problem 1 is a challenging nonconvex problem
- Solutions
  - Transform the nonconvex problem into an equivalent DC problem by introducing auxiliary variables and extra constraints
  - Obtain a KKT point using CCCP

Outline

1. Introduction
2. System Model
3. Problem Formulation and Solution
4. **Numerical Results**
5. Conclusion

Simulation Setup

- Baselines:
  - 1L-RS: extends 1-layer rate splitting [TSP'16] for unicast to general multicast
  - NoRS: extends SDMA for unicast to general multicast
  - OFDMA: considers the maximum ratio transmission (MRT) on each subcarrier and optimizes the subcarrier and power allocation [TWC'21]
- The proposed solution is termed Prop-RS
- Set $K = 3$, $I = 7$, $I_1 = \{1, 4, 5, 7\}$, $I_2 = \{2, 4, 6, 7\}$, $I_3 = \{3, 5, 6, 7\}$
- Set $\mathcal{P}_{\{1\}} = \{1\}$, $\mathcal{P}_{\{2\}} = \{2\}$, $\mathcal{P}_{\{3\}} = \{3\}$, $\mathcal{P}_{\{1,2\}} = \{4\}$, $\mathcal{P}_{\{1,3\}} = \{5\}$, $\mathcal{P}_{\{2,3\}} = \{6\}$, and $\mathcal{P}_{\{1,2,3\}} = \{7\}$
- Set $\alpha_{\mathcal{S}} = 1/7, \mathcal{S} \in \boldsymbol{\mathcal{S}}$, $B = 30$ kHz, $N = 72$, and $\sigma^2 = 10^{-9}$ W
- Consider spatially correlated channel with the correlation following the one-ring scattering model as in [JSAC'21]
  - G denotes the number of user groups
- Evaluate the average of the weighted sum rate of each scheme over the 100 random realizations

Numerical comparisons with baseline schemes

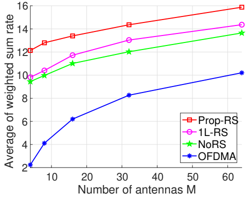


Figure: Weighted sum rate versus M.

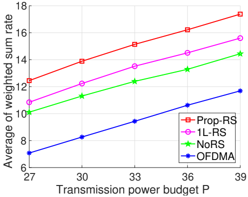


Figure: Weighted sum rate versus P.

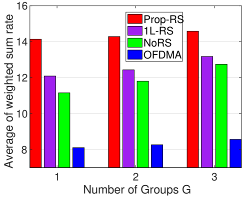


Figure: Weighted sum rate versus G.

- Weighted sum average rate of each scheme increases with M, P, G
- Prop-RS outperforms 1L-RS
  - Prop-RS unleashes the full potential of the flexibility of rate splitting
- Prop-RS outperforms NoRS
  - The cost for NoRS to suppress interference is high
  - Rate splitting together with joint decoding partially decodes interference and partially treats interference as noise

Numerical comparisons with baseline schemes

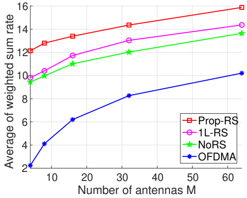


Figure: Weighted sum rate versus M.

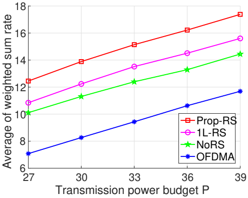


Figure: Weighted sum rate versus P.

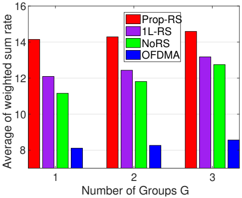


Figure: Weighted sum rate versus G.

- Prop-RS outperforms OFDMA
  - The gain comes from the effective nonorthogonal transmission design
- The gains of Prop-RS over 1L-RS and NoRS increase as G decreases
  - Prop-RS deals with interference in the presence of channel correlation among users flexibly

Numerical comparisons with baseline schemes

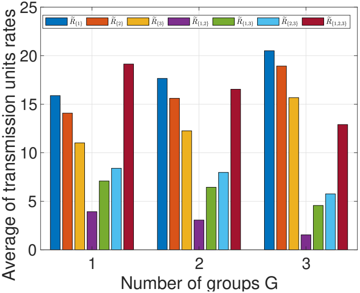


Figure: Rates of transmission units of Prop-RS versus G.

- $\widetilde{R}_{\{1\}}$, $\widetilde{R}_{\{2\}}$, and $\widetilde{R}_{\{3\}}$ increase with G, whereas $\widetilde{R}_{\{1,2\}}$, $\widetilde{R}_{\{1,3\}}$, $\widetilde{R}_{\{2,3\}}$, and $\widetilde{R}_{\{1,2,3\}}$ decrease with G
  - As channel correlation among the users decreases, it is efficient to decode less interference and treat more interference as noise

Outline

1. Introduction
2. System Model
3. Problem Formulation and Solution
4. Numerical Results
5. **Conclusion**

Conclusion

- Investigate the optimization of rate splitting for general multicast
- Adopt linear beamforming at BS and joint decoding at each user
- Maximize the weighted sum rate under the achievable rate region constraints and power constraint and propose an iterative algorithm to obtain a KKT point
  - The proposed optimization framework generalizes the existing ones for rate splitting for unicast and multicast
- Numerically show the notable gains of the proposed solution over existing schemes
- Future work
  - Go beyond linear approaches and investigate nonlinear precoders such as binning
  - Investigate general multicast with partial channel state information at the transmitter side
- Long version was submitted to IEEE Trans. Wireless Commun. [Link]

Reference I

[1] C. Guo, L. Zhao, Y. Cui, Z. Liu, and D. Ng "Power-efficient transmission of multi-quality tiled 360 VR video in MIMO-OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5408-5422, Aug. 2021.

[2] W. Xu, Y. Cui, and Z. Liu, "Optimal multi-view video transmission in multiuser wireless networks by exploiting natural and view synthesisenabled multicast opportunities," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1494-1507, Mar. 2020.

[3] Y. Cui, M. Médard, E. Yeh, D. Leith, F. Lai, and K. R. Duffy, "A linear network code construction for general integer connections based on the constraint satisfaction problem," *IEEE/ACM Trans. Netw.*, vol. 25, no. 6, pp. 3441-3454, Dec. 2017.

[4] H. P. Romero and M. K. Varanasi, "Rate splitting, superposition coding and binning for groupcasting over the broadcast channel: A general framework," *arXiv preprint arXiv:2011.04745*, Nov. 2020.

[5] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49-60, Jan. 1981.

Reference II

[6] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315-328, Jan. 2013.

[7] Z. Li, C. Ye, Y. Cui, S. Yang, and S. Shamai, "Rate splitting for multi-antenna downlink: Precoder design and practical implementation," *IEEE J. Select. Areas Commun.*, vol. 38, no. 8, pp. 1910–1924, Jun. 2020.

[8] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-splitting for multi-antenna non-orthogonal unicast and multicast transmission: spectral and energy efficiency analysis," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8754-8770, Dec. 2019.

[9] H. Chen, D. Mi, B. Clerckx, Z. Chu, J. Shi, and P. Xiao, "Joint power and subcarrier allocation optimization for multigroup multicast systems with rate splitting," *IEEE Trans Veh. Technol.*, vol. 69, no. 2, pp. 2306-2310, Feb. 2020.

[10] N. D. Sidiropoulos, T. N. Davidson and Z. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239-2251, Jun. 2006.

[11] H. Joudeh and B. Clerckx, "Robust transmission in downlink multiuser MISO systems: A rate-splitting approach," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6227–6242, Dec. 2016.