

양자화기 기반 전송률분할 다중접속 프리코더 및 전송 전력 최적화

성지원, 박석준, 최진석

한국과학기술원 전기및전자공학부

Quantized Rate-Splitting Multiple Access Precoding and Power Optimization

Jiwon Sung, Seokjun Park, and Jinseok Choi

School of Electrical Engineering, KAIST

{jiwonsung, sj.park, jinseok}@kaist.ac.kr

요 약

본 논문은 이중 해상도 DAC 다중사용자-다중안테나 통신 시스템에서 기지국에게 주어진 전체 전력 제약조건으로 특정 안테나들을 선택하면서 스펙트럼 효율을 최대화하는 전송률분할 프리코딩 알고리즘을 제안한다. 전력 스케일링과 안테나 선택을 동시에 고려하여 프리코더를 최적화한다는 점에서 높은 의의가 존재한다.

1. 서론

6G 통신에서 스펙트럼 효율을 극대화하는 기술의 필요성은 기하급수적으로 증가하고 있으나, 구현에 있어서 사용자간 간섭 문제와 기지국에서의 전력 소비 문제가 있다. 높은 사용자간 간섭을 줄일 수 있는 물리계층 전송 방법 중에 무선통신 다중사용자-다중안테나 (MU-MIMO) 시스템에서 높은 공간적 자유도와 스펙트럼 효율 (SE)를 가질 수 있는 전송률분할 방법 (RSMA)이 있다. 반면에, 기지국에서 주어진 전력 소비량 조건을 만족하기 위해 최대 전송 전력 스케일링 방법과 특정 안테나를 선택적으로 끄는 안테나 선택 방법이 있다 [1]. 본 논문에서는 둘 다 고려하였으며, DAC의 양자화도 고려하여 전송률분할 방법을 통해 sum SE를 극대화하는 프리코딩 알고리즘을 새롭게 제안한다.

2. 본론

본 논문에서는 다운링크 MU-MIMO 통신 시스템을 고려하였다. 전체 N 의 이중 해상도 access point (AP) 안테나와 K 명의 유저들로 이루어지며, i 번째 안테나는 $b_{\text{DAC},i}$ 의 비트 수를 가진다. 최대 전송 전력을 P , 공용 및 개인 스트림을 s_c, s_k 로 정의하면, 프리코더를 거친 전송 신호는 $\mathbf{x} = \sqrt{P}\mathbf{f}_0 s_c + \sqrt{P} \sum_{k=1}^K \mathbf{f}_k s_k$ 이다. DAC에서 양자화된 신호는 $Q(\mathbf{x}) \approx \mathbf{x}_q = \sqrt{P}\Phi_{\text{DAC}} \mathbf{f}_0 s_c + \sqrt{P}\Phi_{\text{DAC}} \sum_{k=1}^K \mathbf{f}_k s_k + \mathbf{q}_{\text{DAC}}$ 이며, 여기서 $Q(\cdot)$, $\Phi_{\text{DAC}} \in \mathbb{C}^{N \times N}$, $\mathbf{q}_{\text{DAC}} \in \mathbb{C}^{N \times 1}$ 는 각각 스칼라 양자화기, 양자화 손실 행렬, 그리고 양자화 잡음 벡터를 가르킨다. 공통 스트림은 모든 유저들에게 해독 가능하기에, 각 유저는 successive interference cancellation (SIC)으로 공통 스트림을 먼저 해독한 뒤에 개인 스트림을 해독한다. 송수신단에서 완벽한 채널 정보를 가지고 있다고 가정하였으며, 이때 AP와 유저들간 채널 매트릭스를 $\mathbf{H}^H \in$

$\mathbb{C}^{K \times N}$, k 번째 유저가 가지는 채널을 \mathbf{h}_k 로 정의한다. 이때, 수신단에서 받는 baseband 신호는 $\mathbf{y} = \mathbf{H}^H \mathbf{x}_q + \mathbf{n}$ 이며, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_K, \sigma^2 \mathbf{I}_K)$ 은 AWGN이다.

공통 스트림 s_c 가 모든 유저들로부터 해독 가능해야 하기에, s_c 의 $\text{SE}(R_c)$ 는 아래처럼 모든 유저들의 공통 스트림 $s_{c,k}$ 의 $\text{SE}(R_{c,k})$ 의 최솟값이 되어야한다.

$$R_c = \min_{k \in \mathcal{K}} \left\{ \log_2 \left(1 + \frac{P |\mathbf{h}_k^H \Phi_{\text{DAC}} \mathbf{f}_0|^2}{P \sum_{k=1}^K |\mathbf{h}_k^H \Phi_{\text{DAC}} \mathbf{f}_k|^2 + \mathbf{h}_k^H \mathbf{R}_{\text{qDAC}} \mathbf{q}_{\text{DAC}} \mathbf{h}_k + \sigma^2} \right) \right\}.$$

여기서 $\mathbf{R}_{\text{qDAC}} = \Phi_{\text{DAC}} \Phi_{\beta \text{DAC}} \text{diag}(\mathbb{E}[\mathbf{x}\mathbf{x}^H])$ 이다.

공통 스트림 s_c 를 해독해서 제거한 후 k 번째 유저의 $\text{SE}(R_k)$ 는 다음과 같다:

$$R_k = \log_2 \left(1 + \frac{P |\mathbf{h}_k^H \Phi_{\text{DAC}} \mathbf{f}_k|^2}{\sum_{i=1, i \neq k}^K |\mathbf{h}_k^H \Phi_{\text{DAC}} \mathbf{f}_i|^2 + \mathbf{h}_k^H \mathbf{R}_{\text{qDAC}} \mathbf{q}_{\text{DAC}} \mathbf{h}_k + \sigma^2} \right).$$

AP에서 할당된 총 전력을 P_{tot} , 샘플 rate을 f_s , 그리고 안테나 선택을 위한 지시 함수를 $\mathbb{1}_{i \in A}$ 로 정의한다. DAC, 라디오 주파수 체인, 로컬 오실레이터 전력 소모량을 각각 $P_{\text{DAC}}(b_{\text{DAC},i}, f_s)$, P_{RF} , P_{LO} 로 정의하면, i 번째 안테나의 전력 소모량은 $P_{\text{ant},i} = 2P_{\text{DAC}}(b_{\text{DAC},i}, f_s) + P_{\text{RF}}$ 이며 전체 회로 전력 소비량은 $P_{\text{cir}} = P_{\text{LO}} + \sum_{i=1}^N \mathbb{1}_{i \in A} \{P_{\text{ant},i}\}$ 이다.

최적화 문제를 다음과 같이 구성한다.

$$\begin{aligned} & \underset{\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_K}{\text{maximize}} \quad R_c + \sum_{k=1}^K R_k = R_{\Sigma} \\ & \text{subject to} \quad \kappa^{-1} P_{\text{tx}} + P_{\text{cir}} \leq P_{\text{tot}}. \end{aligned}$$

여기서 κ 는 전력 증폭기 효율을 가르킨다.

이 문제의 형태를 바꿔서 generalized power iteration (GPI) 방법을 사용한다. 우선 프리코더가 전력 스케일링을 고려하도록 가중치된 프리코딩 행렬을 $\mathbf{W} = \frac{1}{\sqrt{\tau}} \Phi_{\text{DAC}}^{1/2} \mathbf{F}$ 로 정의하고, 그것을 벡터화 시켜 $\bar{\mathbf{w}} = \text{vec}(\mathbf{W})$ 로 구성한다. $R_{c,k}$, R_k 는 각각 $\log_2 \left(\frac{\bar{\mathbf{w}}^H \mathbf{A}_{c,k} \bar{\mathbf{w}}}{\bar{\mathbf{w}}^H \mathbf{B}_{c,k} \bar{\mathbf{w}}} \right)$, $\log_2 \left(\frac{\bar{\mathbf{w}}^H \mathbf{A}_k \bar{\mathbf{w}}}{\bar{\mathbf{w}}^H \mathbf{B}_k \bar{\mathbf{w}}} \right)$ 와 같은 Rayleigh quotient 꼴로 바꾼다. 더불어, $\min_{k \in \mathcal{K}} \{R_{c,k}\}$ 의 non-smooth 문제를 해결

Proposed Algorithm

1. **initialize:** $\bar{\mathbf{w}}^{(0)}, \tau^{(0)},$ and $\mu^{(0)}$
2. Set the iteration count $t_0 = 0$
3. **while** $t_0 \leq t_{\mu, \max}$ **do**
4. Set the iteration count $t_1 = 0$
5. **while** $\frac{\|\mathbf{F}^{(t_1+1)} - \mathbf{F}^{(t_1)}\|_F}{\|\mathbf{F}^{(t_1)}\|_F} > \epsilon_1$ & $t_1 \leq t_{F, \max}$ **do**
6. Set the iteration count $t_2 = 0$
7. **while** $\frac{|\tau^{(t_2+1)}|}{|\tau^{(t_2)}|} > \epsilon_2$ & $t_2 \leq t_{\tau, \max}$ **do**
8. Update $\tau^{(t_2)}$ using Gradient Descent
9. $t_2 \leftarrow t_2 + 1$
10. $\bar{\mathbf{w}} = \text{Q-GPI-RS}(\mathbf{A}_{\text{KKT}}(\bar{\mathbf{w}}, \tau^{(t_2)}), \mathbf{B}_{\text{KKT}}(\bar{\mathbf{w}}, \tau^{(t_2)}))$
11. Compute $\mathbf{F}^{(t_1)} = \sqrt{\tau^{(t_2)}} \Phi_{\alpha_{\text{DAC}}}^{-1/2} [\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_K]$
12. $t_1 \leftarrow t_1 + 1$
13. Update $\mu^{(t_0)}$ using Binary Search
14. $t_0 \leftarrow t_0 + 1$
15. Compute $\hat{\mathbf{W}} = \tilde{\mathbf{W}} / \max_i \left\| \frac{\tilde{\mathbf{w}}_i}{\sqrt{\alpha_{\text{DAC}, i}}} \right\|$
16. Set $\tilde{\mathbf{f}}_i^{(t_1)} = \mathbf{0}_{1 \times (K+1)}$ if $\left\| \frac{\tilde{\mathbf{w}}_i}{\sqrt{\alpha_{\text{DAC}, i}}} \right\| < \epsilon_{\text{as}}, i = 1, \dots, N$
17. **return** $\mathbf{F}^{(t_1)}$

하기 위해 LogSumExp 방식[2]을 이용해 $\min_{k \in \mathcal{K}} \{R_{c,k}\} \approx -a \ln \left(\sum_{k=1}^K \exp \left(-\frac{1}{a} R_{c,k} \right) \right)$ 로 근사시키며 여기서 $a \rightarrow 0$ 일수록 정밀해진다. 또한, $\tilde{\mathbf{w}}_i = \mathbf{e}_i^H \Phi_{\alpha_{\text{DAC}}}^{-1/2} \mathbf{W}$ 일 때, $\mathbb{1}_{i \in \mathcal{A}}$ 도 $\mathbb{1}_{i \in \mathcal{A}} \approx \log_2 \left(1 + \rho^{-1} \left\| \sqrt{\alpha_{\text{DAC}, i}} \tilde{\mathbf{w}}_i \right\|^2 \right)^{1/\log_2(1+\rho^{-1})}$ 로 근사할 수 있으며 $\rho \rightarrow 0$ 일수록 정밀해진다. 결론적으로, 최적화 문제는 다음과 같이 재구성한다.

$$\underset{\mathbf{w}}{\text{maximize}} \quad -a \ln \left(\sum_{k=1}^K \left(\frac{\bar{\mathbf{w}}^H \mathbf{A}_{c,k} \bar{\mathbf{w}}}{\bar{\mathbf{w}}^H \mathbf{B}_{c,k} \bar{\mathbf{w}}} \right)^{-\frac{1}{a \ln 2}} \right) + \frac{1}{\ln 2} \sum_{k=1}^K \ln \left(\frac{\bar{\mathbf{w}}^H \mathbf{A}_k \bar{\mathbf{w}}}{\bar{\mathbf{w}}^H \mathbf{B}_k \bar{\mathbf{w}}} \right)$$

$$\text{subject to} \quad \|\bar{\mathbf{w}}\| = 1 \quad \text{and} \quad \frac{\tau}{K} P + P_{\text{cir}} \leq P_{\text{tot}}$$

목적함수를 $\mathcal{L} = R_c + \sum_{k=1}^K R_k - \mu(\kappa^{-1} P_{\text{tx}} + P_{\text{cir}} - P_{\text{tot}}) = \ln \lambda(\bar{\mathbf{w}})$ 로 재정의한다. 이 함수는 비볼록 특성 때문에 기존의 프리코딩 방법으로는 최고의 국소 최적해를 도출할 수 없기에, 최고의 최적해를 찾기 위해 GPI 기법을 적용하는 것을 목표로 삼았다. 목적함수를 $\bar{\mathbf{w}}$ 에 대해서 일차 편미분하여 최적화 조건을 구하면 $\mathbf{B}_{\text{KKT}}^{-1} \mathbf{A}_{\text{KKT}}(\bar{\mathbf{w}}) \bar{\mathbf{w}} = \lambda(\bar{\mathbf{w}}) \bar{\mathbf{w}}$ 꼴로 나타낼 수 있다. 이것을 고유값 문제로 해석하면, 주요 고유벡터를 찾는 것이 목적함수의 최고의 국소 최적해를 찾는 것이다. 우리가 제안한 GPI 기반 알고리즘은 μ, \mathbf{F}, τ 를 순차적으로 최적화한다. μ 는 이진 검색 알고리즘을 이용해서 효율적으로 최적값을 구하며, \mathbf{F} 는 [3]에서 제안한 Q-GPI-RS 알고리즘을 통해서 $\lambda(\bar{\mathbf{w}})$ 를 최대화하는 $\bar{\mathbf{w}}$ 를 반복적으로 찾는다. 마지막으로, τ 는 주어진 μ 와 \mathbf{F} 에 대해서 경사 하강 알고리즘으로 찾는다. 마지막으로, 각 안테나의 정규화된 전력이 임계값 ϵ_{as} 를 못 넘기면 해당 안테나를 끄으로써 전력 소비량을 줄인다.

유저 k 의 공간 공분산 행렬을 생성하기 위해 one-ring channel 모델을 사용한다. 부분적인 CSIT를 가정했으며, AP가 추정된 채널에 대해 알려진 추정 오류 공분산을 가지고 접근한다. AP는 16개의 안테

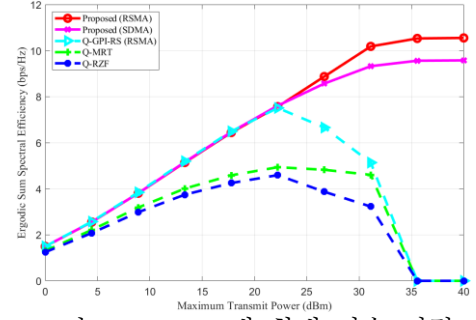


그림 1. Sum SE 대 최대 전송 전력

나를 가지며, 4 등분해서 DAC 비트 수를 각각 4, 8, 12, 16 비트로 설정했다. 4명의 유저들에게 가중치된 프리코딩 신호가 전송되며, ϵ_1 와 ϵ_2 를 각각 0.01, 0.001로 설정하였다. 제안된 알고리즘의 성능을 평가하기 위해 대조군으로 Proposed (SDMA), Q-GPI-RS [3], Q-RZF, 그리고 Q-MRT를 사용했다. 안테나 선택이 없는 알고리즘들은 노름을 기반으로 안테나 전력이 적은 안테나 순서대로 끈다.

그림 1은 최대 전송 전력 P 에 따른 sum SE를 나타낸다. 여기서 전체 전력을 $P_{\text{tot}} = 40$ dBm으로 설정하였다. 최대 전송 전력을 사용했을 때 $P = 22.45$ dBm부터 안테나 선택을 해야하며, $P = 34.30$ dBm부터는 안테나를 다 꺼야한다. 실험 결과를 통해 본 논문에서 제안된 알고리즘이 모든 구간에서 가장 높은 성능을 보유하고 있음을 입증한다.

3. 결론

본 논문에서는 양자화된 다운링크 통신 시스템에서 스펙트럼 효율을 극대화하도록 안테나 선택과 전력 스케일링을 동시에 고려하여 최적화한 프리코딩 알고리즘을 제안하였다. 최적화 문제를 풀기 위해서 min 함수와 지시함수를 smooth한 형태로 근사한 뒤, GPI 방식을 이용해서 최고의 최적해를 찾았다. 안테나 수가 많아짐에 따라 전력 소비량이 증가하더라도 통신 효율을 향상시킬 수 있는 유망한 방안으로 활용될 수 있다.

ACKNOWLEDGEMENT

이 논문은 2024년도 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임 (KRIT-CT-22-078).

4. 참고 문헌

- [1] J. Choi, J. Park, N. Lee, "Energy Efficiency Maximization Precoding for Quantized Massive MIMO Systems," IEEE Trans. Wireless Commun., vol. 21, no. 9, pp. 6803–6817, Feb. 2022.
- [2] C. Shen and H. Li, "On the Dual Formulation of Boosting Algorithms," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 12, pp. 2216–2231, Mar. 2010.
- [3] S. Park, J. Choi, J. Park, W. Shin, and B. Clerckx, "Rate-Splitting Multiple Access for Quantized Multiuser MIMO Communications," IEEE Trans. Wireless Comm., vol. 22, no. 11, pp. 7696–7711, Mar. 2023.