

Traffic Aware Precoder Design for Space Frequency Resource Allocation

Ganesh Venkatraman, Antti Tölli[†], Le-Nam Tran and Markku Juntti

Email: {gvenkatr, antti.tolli, le.nam.tran, markku.juntti}@ee.oulu.fi

Centre for Wireless Communications (CWC),
Department of Communications Engineering (DCE),
University of Oulu, Oulu, FI-90014

Outline

Introduction

System Model & Problem Formulation

Centralized Solutions

- Existing Q-WSRM Formulation
- JSFRA Formulation (SINR Relaxation)
- JSFRA Formulation (MSE Reformulation)
- Simulation Results

Distributed Solutions

- Primal & ADMM based decompositions
- KKT based Distributed Solution
- Simulation Results

Time Correlated Fading Performance

- Methods to Update Precoders via OTA
- Simulation Results

Conclusions

Section 1

Introduction

KKT Expressions

$$\mathbf{m}_{l,k,n}^{(i)} = \left(\sum_{x \in \mathcal{U}} \sum_{y=1}^L \alpha_{y,x,n}^{(i-1)} \mathbf{H}_{b_k,x,n}^H \mathbf{w}_{y,x,n}^{(i-1)} \mathbf{w}_{y,x,n}^{H(i-1)} \mathbf{H}_{b_k,x,n} + \delta_b \mathbf{I}_{N_T} \right)^{-1} \alpha_{l,k,n}^{(i-1)} \mathbf{H}_{b_k,k,n}^H \mathbf{w}_{l,k,n}^{(i-1)}$$

$$\mathbf{w}_{l,k,n}^{(i)} = \left(\sum_{x \in \mathcal{U}} \sum_{y=1}^L \mathbf{H}_{b_x,k,n} \mathbf{m}_{y,x,n}^{(i)} \mathbf{m}_{y,x,n}^{H(i)} \mathbf{H}_{b_x,k,n}^H + N_0 \mathbf{I}_{N_R} \right)^{-1} \mathbf{H}_{b_k,k,n} \mathbf{m}_{l,k,n}^{(i)}$$

$$\epsilon_{l,k,n}^{(i)} = \left| 1 - \mathbf{w}_{l,k,n}^{H(i)} \mathbf{H}_{b_k,k,n} \mathbf{m}_{l,k,n}^{(i)} \right|^2 + \sum_{(x,y) \neq (l,k)} \left| \mathbf{w}_{l,k,n}^{H(i)} \mathbf{H}_{b_y,k,n} \mathbf{m}_{x,y,n}^{(i)} \right|^2 + \|\mathbf{w}_{l,k,n}\|^2 N_0$$

$$t_{l,k,n}^{(i)} = -\log_2(\epsilon_{l,k,n}^{(i-1)}) - \frac{(\epsilon_{l,k,n}^{(i)} - \epsilon_{l,k,n}^{(i-1)})}{\log(2) \epsilon_{l,k,n}^{(i-1)}}$$

$$\sigma_{l,k,n}^{(i)} = \left[\frac{a_k q}{\log(2)} \left(Q_k - \sum_{n=1}^N \sum_{l=1}^L t_{l,k,n}^{(i)} \right)^{(q-1)} \right]^+$$

$$\alpha_{l,k,n}^{(i)} = \alpha_{l,k,n}^{(i-1)} + \rho^{(i)} \left(\frac{\sigma_{l,k,n}^{(i)}}{\epsilon_{l,k,n}^{(i)}} - \alpha_{l,k,n}^{(i-1)} \right)$$

Introduction and Motivation

- Complex algorithms are adopted to maximize throughput to satisfy the data requirements from higher layers
- Available wireless resources are to be utilized efficiently to minimize the backlogged packets
- Spatial and Frequency resources are exploited to empty the packets waiting at the BSs
- We discuss precoder designs for multi-user MIMO-OFDM setup to minimize the number of queued packets

Section 2

System Model & Problem Formulation

Symbols used

- OFDM system with N sub-channels and N_B BSs, each equipped with N_T transmit antennas
- Let K be the total number of users with N_R antennas
- Let \mathcal{B} and \mathcal{U} denote the set of coordinating BSs and users in the system
- Let L be the total available spatial streams for a user k , given by $\min(N_T, N_R)$

System Model

- The l th spatial signal received on sub-channel n of user k is given by

$$\hat{d}_{l,k,n} = \mathbf{w}_{l,k,n}^H \mathbf{H}_{b_k,k,n} \mathbf{m}_{l,k,n} d_{l,k,n} + \mathbf{w}_{l,k,n}^H \mathbf{n}_{l,k,n} + \mathbf{w}_{l,k,n}^H \sum_{i \in \mathcal{U} \setminus \{k\}} \mathbf{H}_{b_i,k,n} \sum_{j=1}^L \mathbf{m}_{j,i,n} d_{j,i,n} \quad (1)$$

- where $\mathbf{m}_{l,k,n}$ and $\mathbf{w}_{l,k,n}$ are transmit and receive beamformers corresponding to the l th spatial stream on the n th sub-channel of user k

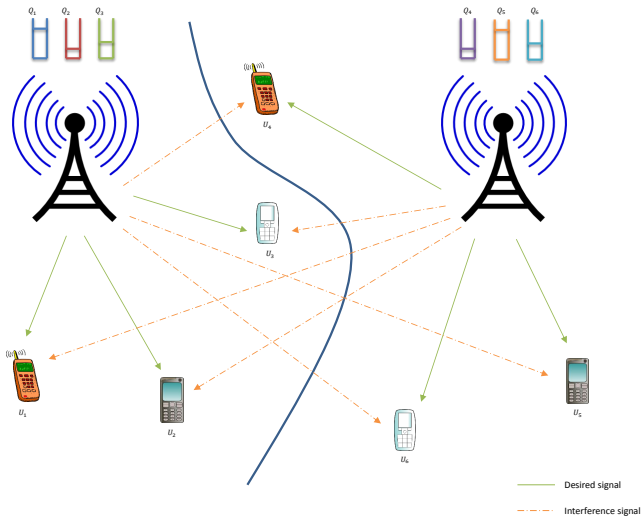
System Model

- $\mathbf{H}_{b_k,k,n} \in \mathbb{C}^{N_R \times N_T}$ denotes the channel between BS b_k and user k
- $d_{l,k,n}$ and $n_{l,k,n}$ correspond to data symbol and equivalent noise on l th spatial stream of user k
- Using the above notations, the SINR seen by the l th spatial stream on the n th sub-channel for user k is given by

$$\gamma_{l,k,n} = \frac{|\mathbf{w}_{l,k,n}^H \mathbf{H}_{b_k,k,n} \mathbf{m}_{l,k,n}|^2}{\hat{N}_0 + \sum_{(j,i) \neq (l,k)} |\mathbf{w}_{l,k,n}^H \mathbf{H}_{b_i,k,n} \mathbf{m}_{j,i,n}|^2} \quad (2)$$

- where $\hat{N}_0 = \|\mathbf{w}_{l,k,n}^H \mathbf{n}_{l,k,n}\|^2$

System Model with User Queues



Queueing Model

- Each user is associated with backlogged packets of size Q_k packets.
- Queued packets Q_k of each user follows dynamic equation at the i th instant as

$$Q_k(i+1) = \left[Q_k(i) - t_k(i) \right]^+ + \lambda_k(i) \quad (3)$$

- where $t_k = \sum_{n=1}^N \sum_{l=1}^L t_{l,k,n}$ denotes the total number of transmitted packets corresponding to user k in the previous i th instant
- λ_k represents the fresh arrivals of user k at BS b_k

Problem Formulation

- **Objective** - to minimize the number of backlogged packets waiting at BSs
- **Optimization variables** - transmit precoders and receive beamformers
- **MIMO-OFDM** - scheduling of users across sub-channels is inherently performed by precoders

Section 3

Centralized Solutions

Queue-Weighted Sum Rate Maximization (Q-WSRM)

- Q-WSRM formulation is the result of **minimizing the conditional Lyapunov drift**¹
- Q-WSRM formulation is also called as **back pressure algorithm**, since it acts greedily in minimizing the backlogged packets at each instant

$$\underset{t_{l,k,n}}{\text{minimize}} \quad \sum_{k \in \mathcal{U}} \left\{ Q_k(i)^2 - Q_k(i-1)^2 \right\},$$

- where Q_k follows the dynamic Queue expression in (3)

¹Neely, Michael J. "Stochastic network optimization with application to communication and queueing systems." Synthesis Lectures on Communication Networks 3.1 (2010): 1-211.

Queue-Weighted Sum Rate Maximization (Q-WSRM)

- Q-WSRM formulation, which is obtained by solving Lyapunov drift, is given by

$$\underset{t_{l,k,n}}{\text{maximize}} \quad \sum_{k \in \mathcal{U}} Q_k \left(\sum_{n=1}^N \sum_{l=1}^L t_{l,k,n} \right) \quad (4a)$$

$$\sum_{n=1}^N \sum_{l=1}^L t_{l,k,n} \leq Q_k / Q_{k,n} \quad (4b)$$

- Queue-Rate product is maximized
- Users with more number of backlogged packets are favored over good channel users

JSFRA Formulation (SINR Relaxation)

- Centralized Design - precoders are designed by a controller, which are then used by all BSs in \mathcal{B}
- The objective used to design transmit precoders is

$$v_k = \left| Q_k - \sum_{n=1}^N \sum_{l=1}^L t_{l,k,n} \right|^q \quad (5)$$

- To generalize the objective, we use $\tilde{v}_k \triangleq a_k^{\frac{1}{q}} v_k$, where a_k is arbitrary weights used control the priorities
- Exponent q plays different role based on the value it assumes
 - ▶ $\ell_{q=1}$ results in greedy allocation
 - ▶ $\ell_{q=2}$ ideal for the delay or buffer size limited scenarios
 - ▶ $\ell_{q=\infty}$ provides fair resource allocation in each transmission instant

JSFRA Formulation (SINR Relaxation)

- Optimization problem with queue difference objective is nonconvex due to the constraint

$$\underset{\mathbf{m}_{l,k,n}}{\text{minimize}} \quad \left| Q_k - \sum_{n=1}^N \sum_{l=1}^L \log(1 + \gamma_{l,k,n}) \right|^q \quad (6a)$$

$$\text{subject to} \quad \gamma_{l,k,n} \leq \frac{\left| \mathbf{w}_{l,k,n}^H \mathbf{H}_{b_k,k,n} \mathbf{m}_{l,k,n} \right|^2}{\beta_{l,k,n}} \quad (6b)$$

$$\dot{N}_0 + \sum_{(j,i) \neq (l,k)} |\mathbf{w}_{l,k,n}^H \mathbf{H}_{b_l,k,n} \mathbf{m}_{j,i,n}|^2 \leq \beta_{l,k,n} \quad (6c)$$

- The nonconvex constraints are approximated by sequence of convex subsets and solved iteratively by **successive convex approximation (SCA) method**
- Receive beamformers are designed by the MMSE receivers using the converged transmit precoders

JSFRA Formulation (MSE Reformulation)

- Alternatively, we solve the queue minimization problem by utilizing the relation between the MSE and the SINR as

$$\epsilon_{l,k,n} = (1 + \gamma_{l,k,n})^{-1} \quad (7)$$

- Equivalence is valid only when the receivers are designed with the mean squared error (MSE) objective, *i.e.*, using MMSE receivers
- Problem involves nonconvex constraint

$$t_{l,k,n} \leq -\log_2(\epsilon_{l,k,n}) \quad (8a)$$

$$\begin{aligned} \epsilon_{l,k,n} &= \mathbb{E}[(d_{l,k,n} - \hat{d}_{l,k,n})^2] = |1 - \mathbf{w}_{l,k,n}^H \mathbf{H}_{b_k,k,n} \mathbf{m}_{l,k,n}|^2 \\ &\quad + \sum_{(j,i) \neq (l,k)} |\mathbf{w}_{l,k,n}^H \mathbf{H}_{b_i,k,n} \mathbf{m}_{j,i,n}|^2 + \hat{N}_0 \end{aligned} \quad (8b)$$

JSFRA Formulation (MSE Reformulation)

- The nonconvex constraint is approximated by a sequence of convex constraints, which is performed using SCA technique
- The iterative procedure is carried out until convergence or for suitable number of iterations
- The above reformulation works only with the MMSE receiver

Centralized Solutions

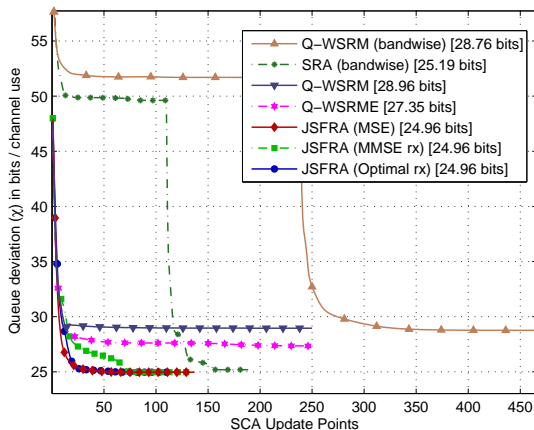


Figure: System Model - $\{N, N_B, K, N_T, N_R\} = \{2, 3, 9, 4, 2\}$

Time Correlated Fading - Centralized Performance

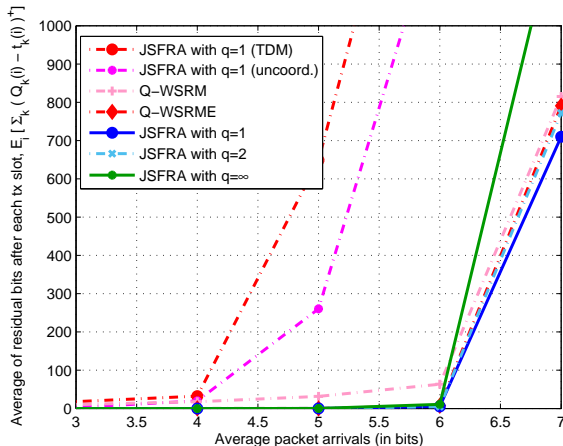


Figure: System Model - $\{N, N_B, K, N_R\} = \{3, 2, 8, 1\}$ after 250 transmissions

Section 4

Distributed Solutions

Distributed Methods

- **Small System - centralized approach is viable**, provided channel remains constant for multiple transmission slots
- However, overhead involved in the centralized design scales up significantly as the network size grows
- Distributed approaches based on primal decomposition or ADMM can be used to reduce the signaling requirements
- Signaling involved in the design of precoders are only **scalar interference variables**
- Approximated convex subproblem in each SCA step is performed via distributed methods

Primal Decomposition Method

- Precoder design is based on master-slave approach
- Interference to the neighboring BS users are **bounded by a scalar variable, treated as a constant in subproblems**
- The interference thresholds are determined by the master problem and used in each slave subproblem constraint as

$$\zeta_{l,k,n,b} \geq \sum_{i \in \mathcal{U}_b} \sum_{j=1}^L |\mathbf{w}_{l,k,n}^H \mathbf{H}_{b,k,n} \mathbf{m}_{j,i,n}|^2 \quad \forall b \in \bar{\mathcal{B}}_{b_k}. \quad (9)$$

ADMM based Decomposition Method

- The ADMM is superior to other distributed schemes in terms of the convergence speed
- ADMM includes an additional quadratic term

$$\|v_k\|_q + \sum_{l,k,n} \nu_{l,k,n}^{(j)} \left(\zeta_b - \zeta_b^{(j)} \right) + \frac{\rho}{2} \|\zeta_b - \zeta_b^{(j)}\|^2 \quad (10)$$

in objective, where $\zeta_b^{(j)}$ is global consensus variable

- $\zeta_{l,k,n,b}$ in (9) is **treated as an optimization variable in ADMM**
- Consensus variables are updated upon exchanging corresponding local $\zeta_{l,k,n,b}$'s among coordinating BSs

KKT based Distributed Solution

- Decentralization methods involve significant signaling exchanges via backhaul
- **Overhead is large for multi-antenna receivers** - iterative design should reduce the backlogged packets significantly in first few iterations
- To achieve that, we design precoders by solving the Karush-Kuhn-Tucker (KKT) equations of the JSFRA problem via MSE reformulation
- **Group update of all involved optimization variables** - to speed up the convergence of precoder design

Performance of Distributed Solutions

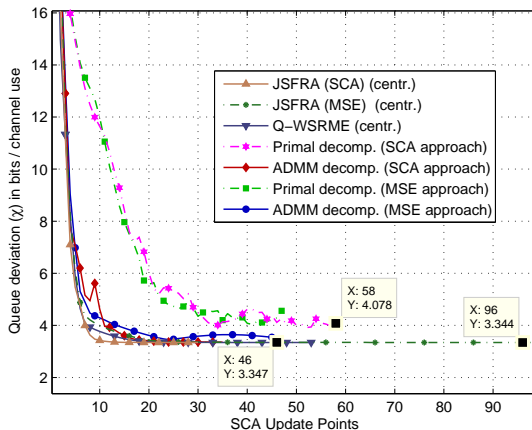


Figure: System Model - $\{N, N_B, K, N_T, N_R\} = \{3, 2, 8, 4, 1\}$

Performance of KKT based Approach

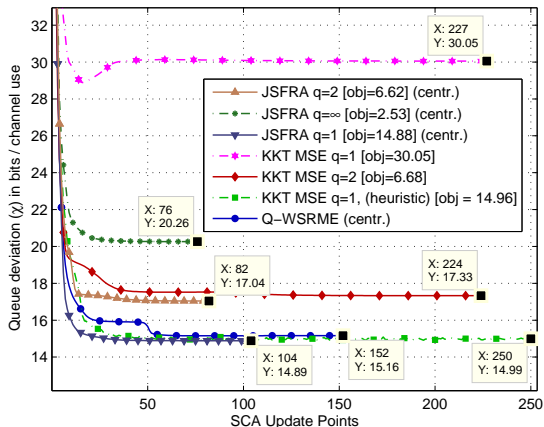


Figure: System Model - $\{N, N_B, K, N_T, N_R\} = \{5, 2, 8, 4, 1\}$

Section 5

Time Correlated Fading Performance

Time Correlated Precoder Design - Introduction

- So far, we have discussed few methods to design precoders in a distributed manner
- Now, we discuss a **practical approach to design precoders with minimal backhaul usage**
- To discuss further, we consider KKT based design for following reasons
 - ▶ Due to closed form solution, **each SCA update requires only single exchange of coupling variables**
 - ▶ Moreover, transmit and receive precoders expressions can be formed by using precoded pilot transmissions
- We **adopt Bi-directional training (BiT)** to update transmit and receive beamformers ²

²P. Komulainen, A. Tölli & M. Juntti, Effective CSI Signaling and Decentralized Beam Coordination in TDD Multi-Cell MIMO Systems, IEEE Transactions on Signal Processing, vol. 61, no. 9, pp. 2204 – 2218, May 2013

Time Correlated Precoder Design - Strategy A

- To devise over-the-air (OTA) based iterative algorithm - let us consider transmit and receive beamformer expressions

$$\mathbf{m}_{l,k,n}^{(i)} = \left(\sum_{x \in \mathcal{U}} \sum_{y=1}^L \alpha_{y,x,n}^{(i-1)} \mathbf{H}_{b_k,x,n}^H \mathbf{w}_{y,x,n}^{(i-1)} \mathbf{w}_{y,x,n}^{H(i-1)} \mathbf{H}_{b_k,x,n} + \delta_b \mathbf{I}_{N_T} \right)^{-1} \alpha_{l,k,n}^{(i-1)} \mathbf{H}_{b_k,k,n}^H \mathbf{w}_{l,k,n}^{(i-1)}$$

$$\mathbf{w}_{l,k,n}^{(i)} = \left(\sum_{x \in \mathcal{U}} \sum_{y=1}^L \mathbf{H}_{b_x,k,n} \mathbf{m}_{y,x,n}^{(i)} \mathbf{m}_{y,x,n}^{H(i)} \mathbf{H}_{b_x,k,n}^H + N_0 \mathbf{I}_{N_R} \right)^{-1} \mathbf{H}_{b_k,k,n} \mathbf{m}_{l,k,n}^{(i)}$$

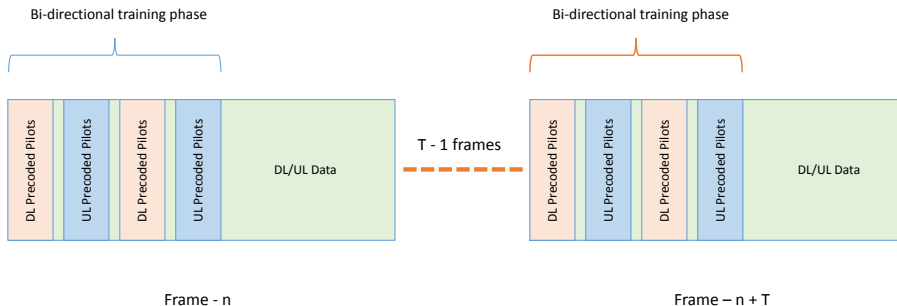
- Note that transmit precoders $\mathbf{m}_{l,k,n}$ depend on $\mathbf{H}_{b_k,i,n}^T \mathbf{u}_{l,i,n}$, i.e., **effective uplink channel**
- Similarly, receive beamformers $\mathbf{w}_{l,k,n}$ depend on $\mathbf{H}_{b_k,i,n} \mathbf{m}_{l,k,n}$, i.e., **effective downlink channel**

Time Correlated Precoder Design - Strategy A

- To update receive beamformers at the user terminals, **each BS transmits precoded downlink pilots orthogonally** by using $\mathbf{m}_{l,k,n}^{(i)}$
- Upon receiving effective channel, each user update the respective receive beamformer
- Similarly, to update transmit beamformers at BSs, **each user transmits two precoded pilot transmission**, namely, $\sqrt{\alpha_{l,k,n}^{(i-1)}} \mathbf{w}_{l,k,n}^*$ and $\alpha_{l,k,n} \mathbf{w}_{l,k,n}^*$
- Upon receiving two uplink precoded pilots, each BS then evaluates the respective transmit beamformers of all users
- The above iterative procedure is performed until convergence or for predetermined number of updates - **Strategy A**

Bi-directional Frame Structure

- To perform iterative OTA based update, we **consider time division duplexing (TDD) frame format**
- Using OTA transmissions, **effective channel can be conveyed to BSs and users by precoded pilots**



Time Correlated Precoder Design - Strategy B

- To **speed up the convergence** of iterative algorithm, we update the precoders of desired users **locally by fixing neighboring BSs transmit beamformers**
- To do so, in addition to the signaling requirements in strategy A, **inter-cell interference covariance seen by desired users** is also fed back to serving BS, *i.e.*,

$$\mathbf{Z}_{l,k,n} = \sum_{i \in \mathcal{U} \setminus \mathcal{U}_b} \mathbf{H}_{b_i,i,n} \mathbf{m}_{l,i,n} \mathbf{m}_{l,i,n}^H \mathbf{H}_{b_i,i,n}^H + N_0 \mathbf{I}_{N_R}. \quad (11)$$

- Since inter-cell interference covariance is not required explicitly, **effective whitened channel is enough to evaluate precoders**
- The whitening filter $\mathbf{Q}_{l,k,n}$ is given by $\mathbf{Z}_{l,k,n}^{-1} = \mathbf{Q}_{l,k,n}^H \mathbf{Q}_{l,k,n}$
- Whitened channel is constructed at BS by **transmitting columns of $\mathbf{Q}_{l,k,n}$ orthogonally from each user**

Time Correlated Precoder Design - Strategy B

- Upon receiving $\mathbf{Q}_{l,k,n}$, each BS can construct whitened channel as

$$\bar{\mathbf{H}}_{b_k,k} = \mathbf{Q}_{l,k,n} \mathbf{H}_{b_k,k}$$

- The BS specific local iteration j to update transmit and receive beamformer is given as

$$\mathbf{m}_{l,k,n}^{(j)} = \left(\sum_{x \in \mathcal{U}_b} \sum_{y=1}^L \alpha_{y,x,n}^{(j-1)} \bar{\mathbf{H}}_{b_k,x,n}^H \mathbf{w}_{y,x,n}^{(j-1)} \mathbf{w}_{y,x,n}^{H(j-1)} \bar{\mathbf{H}}_{b_k,x,n} + \delta_b \mathbf{I}_{N_T} \right)^{-1} \alpha_{l,k,n}^{(j-1)} \bar{\mathbf{H}}_{b_k,k,n}^H \mathbf{w}_{l,k,n}^{(j-1)}$$

$$\mathbf{w}_{l,k,n}^{(j)} = \left(\sum_{x \in \mathcal{U}_b} \sum_{y=1}^L \bar{\mathbf{H}}_{b_x,k,n} \mathbf{m}_{y,x,n}^{(j)} \mathbf{m}_{y,x,n}^{H(j)} \bar{\mathbf{H}}_{b_x,k,n}^H + \mathbf{I}_{N_R} \right)^{-1} \bar{\mathbf{H}}_{b_k,k,n} \mathbf{m}_{l,k,n}^{(j)}$$

- The above iterative procedure is performed until convergence or for predetermined number of updates - Strategy B

Summary - Signaling Requirement for OTA based Updates

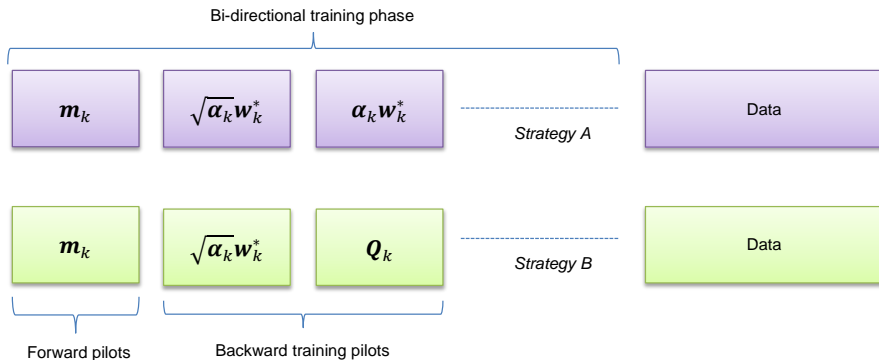


Figure: TDD frame structure with BiT signaling

Assumptions and Evaluation Model

- Every BS and user terminal uses **orthogonal pilots in UL and DL OTA signaling**
- For simplicity, pilot transmissions used to convey the equivalent channel information in one BiT iteration - **consume η resource share**.³
- Under this assumption, the effective rate by considering the signaling overhead is given as

$$\tilde{t}_{l,k,n} = (1 - l_{\max} \eta) \times t_{l,k,n} \quad (12)$$

- Total number of backlogged packets is evaluated as - $\chi = \sum_{k=1}^K [Q_k - \tilde{t}_k]^+$
- In all simulations, **we consider $\eta = 1\%$**

³In practice, the performance depends on the amount of available pilots and the size of coherence block.

Average Backlogged Packets - Distributed Design

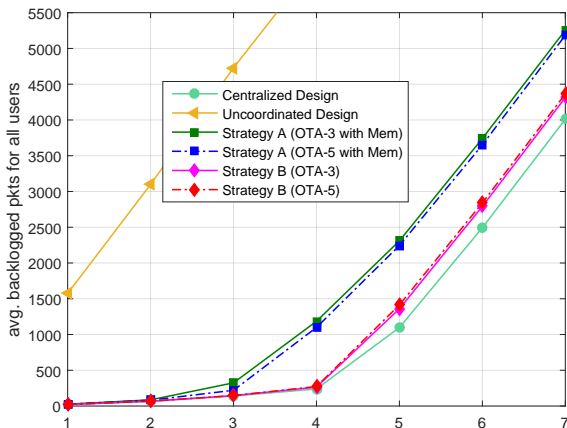


Figure: Average backlogged packets for $\{N, N_B, K, N_T, N_R\} = \{3, 2, 12, 4, 2\}$ evaluated over 250 slots with $f_d T_s \approx 0.1$

Average Backlogged Packets - Distributed Design

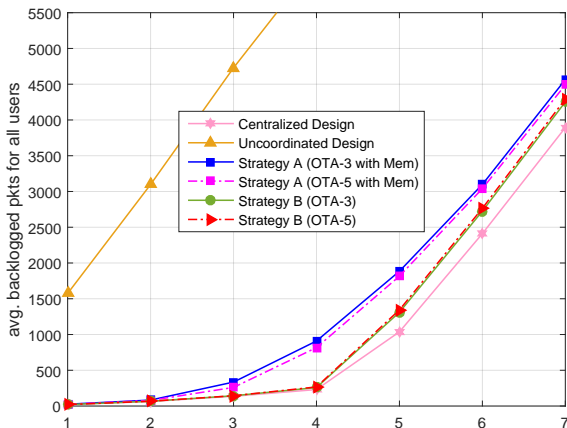


Figure: Average backlogged packets for $\{N, N_B, K, N_T, N_R\} = \{3, 2, 12, 4, 2\}$ evaluated over 250 slots with $f_d T_s \approx 0.01$

Section 6

Conclusions

Conclusions

- We studied cross layer problem of designing transmit and receive beamformers based on the number of residual packets
- Since the problem is nonconvex, we solve the problem iteratively by solving convex subproblems in each iteration
- We also proposed a practical way of implementing the precoder design in a distributed manner by solving the KKT expressions
- Extensions of the proposed work in time-correlated fading scenario with limited number of information exchange cycle is also presented

Questions !