Low Complexity Precoder Design for Delay Sensitive Multi-stream MIMO Systems

Vincent K. N. Lau†, Yan Chen∗†, Peiliang Qiu∗, and Zhaoyang Zhang∗

†Department of Electrical and Electronic Engineering
HKUST, Clear Water Bay, Kowloon, Hong Kong
eeknlau@ee.ust.hk

∗Institute of Information and Communication Engineering
Zhejiang University, Hangzhou, China
{qiupl418, qiupl, ning_ming}@zju.edu.cn

Abstract—In this paper, we consider delay-optimal MIMO precoder and power allocation design for a MIMO Link in wireless fading channels. There are \( L \) data streams spatially multiplexed onto the MIMO link with heterogeneous packet arrivals and delay requirements. The transmitter is assumed to have knowledge of outdated channel state information (CSIT) as well as the joint queue state information (QSI) of the \( L \) buffers. Using static sorting of the \( L \) eigenchannels, we decompose the \( L \) dimensional MDP into \( L \) independent 1-dimensional MDP and derived low complexity precoding and power control policies (with linear complexity) to minimize average delays of the \( L \) application streams.\(^1\)

I. INTRODUCTION

Multiple Input Multiple Output (MIMO) communication is well-known to boost the wireless spectral efficiency through spatial multiplexing. When channel state information is available at the transmitter (CSIT), substantial performance gain could be obtained by power and precoder adaptation. In [1], [2], a linear MIMO precoder design framework is proposed to minimize the weighted sum of mean square errors (MSE) assuming knowledge of perfect CSIT: However, perfect knowledge of CSIT is difficult to obtain in practice. In [3] and [4], MIMO precoder design utilizing either limited feedback or outdated CSIT is proposed. Yet, all these works assumed that the transmitter has infinite buffer and the information flow is delay insensitive, and focused on optimizing the physical layer performance (such as capacity, throughpout or MSE). In practice, it is very important to consider the delay performance in addition to the conventional physical layer performance in MIMO transceiver.

A combined framework taking into account of both queueing delay and physical layer performance is not trivial as it involves both the queueing theory (to model the queue dynamics) and information theory (to model the physical layer dynamics). In general, there are two approaches to deal with delay problems. The first approach converts the delay constraint into average rate constraint using tail probability at large delay regime [5]. While this approach allows potentially simple solution, the control policy will be a function of CSIT only and such control will be good only for large delay regime. In general, the delay-optimal power and precoder adaptation will be a function of both the CSI and the queue state information (QSI). In the second approach, the problem of finding the optimal control policy (to minimize delay) is cast into a Markov Decision Problem (MDP) or stochastic control problem. Unfortunately, it is well-known that there is no easy solution (e.g. value iteration and policy iteration) to MDP in general. In addition, it is usually very complex to evaluate the optimal solution even numerically. In [6], the authors considered delay optimal power control of a SISO fading channel with perfect CSIT using MDP. To get around the problem, the authors considered asymptotic analysis at large delay regime and obtained interesting tradeoff results as well as insight into the structure of the optimal control policy at large delay regime. In [7], the authors showed that the longest queue highest possible rate (LQHPR) policy is delay-optimal for multi-access fading channels. While all the above works addressed different aspects of the delay sensitive resource allocation problem, there are still a number of first order issues to be addressed.

• Low complexity optimal control policy for delay sensitive resource allocation problem in general delay regime
   Most of the existing works considered large delay asymptotic solutions. However, practical operating region for delay sensitive traffics are usually on the low delay regime and the asymptotic simplifications cannot be applied. Hence, it is important to obtain low complexity control policy for general delay regime.

• Delay sensitive MIMO precoder design with outdated CSIT
   As far as we are aware, there is no work on how to design MIMO precoder and power adaptation for multi-stream delay-sensitive applications. The issue is further complicated by the imperfect CSIT in which there will be spatial interference between the MIMO channels.

• Coupling among multiple delay-sensitive heterogeneous data streams
   While [7] considered multi-user systems, the framework applies to situations with sym-
The total average transmitted power has to satisfy a constraint is the MIMO channel state information (CSI) and 

\[ z \]

is given by:

\[ z = \text{channel outputs given by:} \]

\[ x = Hx + z, \quad H \in C^{N_r \times N_t}, \quad z \in C^{N_r} \]

is the MIMO channel state information (CSI) and \( z \in C^{N_r} \) is a zero-mean circularly symmetric complex Gaussian noise vector with normalized covariance \( I \).

We assume the receiver has perfect knowledge of instantaneous CSIR for detection and decoding, however, the CSIT is imperfect due to the estimation noise on the reverse link pilot in TDD system. The MMSE estimator of the CSIT \( \hat{H} \) at the transmitter is thus given by [8]:

\[ \hat{H} = H + \Delta H \]

(2)

where the CSIT error \( \Delta H \) is a zero-mean complex Gaussian distributed with covariance \( \sigma^2 I \). Furthermore, \( E[\Delta H^H \hat{H}] = 0 \) due to the orthogonality principle of MMSE. Hence, \( \sigma^2 \) is a parameter which represents the CSIT quality. When \( \sigma^2 = 0 \), we have perfect CSIT. When \( \sigma^2 = 1 \), we have \( E[\hat{H}^H \hat{H}] = 0 \) and this is equivalent to no CSIT.

As a result, the equivalent channel (with precoder, MIMO channel and the equalizer) for the \( L \) data stream is given by:

\[ s = W^H H P s + W^H z \]

(3)

And the average SINR of the \( i \)-th data stream (conditioned on \( \hat{H} \)) is thus given by:

\[ SINR_i(P) = E \left[ \frac{|w_i^H H P s|^2}{|w_i^H A_i w_i|^2} \left| \hat{H} \right| \right] \]

(4)

where \( A_i = \sum_{j \neq i} H_{pj} P_{ji} H^H + I \) and \( \{ p_i \} \) is the \( i \)-th column of the precoding matrix \( P \). For sufficiently small symbol error probability, the conditional symbol error probability for QAM constellation (conditioned on \( \hat{H} \)) is given by [9]:

\[ P_e(\hat{H}) \leq \kappa_1 Q \left( \frac{3SINR_i}{2R_i - 1} \right) \leq \frac{\kappa_1}{2} \exp \left( \frac{3SINR_i}{2(2R_i - 1)} \right) \]

(5)

for some constant \( \kappa_1 \). Hence, given a sufficiently small target symbol error probability (SER) \( P_e(\hat{H}) = \epsilon \), the data rate \( R_i \) (bits per symbol) of the \( i \)-th data stream is related to the \( SINR_i(P) \) as

\[ R_i = \log_2(1 + \alpha(\epsilon)SINR_i(P)) \]

(6)

where \( \alpha(\epsilon) \) is some constant depending on the target SEP \( \epsilon \). Since the receiver has perfect CSIR and the data rate is an increasing function of \( SINR_i \), it is shown that for any precoder \( P \), Wiener filter \( W = (H^H H + I)^{-1} H \) can simultaneously maximize \( \{ SINR_1, ..., SINR_L \} \). As a result, the conditional average SINR of the \( i \)-th data stream after Wiener filtering is given by:

\[ SINR_i(P) = E[|w_i^H H A^{-1}_i H P s|^2] \]

(7)

Define the instantaneous MSE matrix as follows:

\[ E(P) = E \left[ (s - \hat{s})(s - \hat{s})^H \right] = (I + P^H H^H H P)^{-1} \]

(8)

Note that the diagonal elements of \( E \) contains the instantaneous MSEs of the \( L \) data streams. Using matrix inversion lemma, it can be shown that

\[ SINR_i(P) \geq E[|w_i|^2] - 1 \]

(9)

where the CSIT error \( \Delta H \) is a zero-mean complex Gaussian distributed with covariance \( \sigma^2 I \). Furthermore, \( E[\Delta H^H \hat{H}] = 0 \) due to the orthogonality principle of MMSE. Hence, \( \sigma^2 \) is a parameter which represents the CSIT quality. When \( \sigma^2 = 0 \), we have perfect CSIT. When \( \sigma^2 = 1 \), we have \( E[\hat{H}^H \hat{H}] = 0 \) and this is equivalent to no CSIT.

As a result, the equivalent channel (with precoder, MIMO channel and the equalizer) for the \( L \) data stream is given by:

\[ s = W^H H P s + W^H z \]

(3)

And the average SINR of the \( i \)-th data stream (conditioned on \( \hat{H} \)) is thus given by:

\[ SINR_i(P) = E \left[ \frac{|w_i^H H P s|^2}{|w_i^H A_i w_i|^2} \left| \hat{H} \right| \right] \]

(4)

where \( A_i = \sum_{j \neq i} H_{pj} P_{ji} H^H + I \) and \( \{ p_i \} \) is the \( i \)-th column of the precoding matrix \( P \). For sufficiently small symbol error probability, the conditional symbol error probability for QAM constellation (conditioned on \( \hat{H} \)) is given by [9]:

\[ P_e(\hat{H}) \leq \kappa_1 Q \left( \frac{3SINR_i}{2R_i - 1} \right) \leq \frac{\kappa_1}{2} \exp \left( \frac{3SINR_i}{2(2R_i - 1)} \right) \]

(5)

for some constant \( \kappa_1 \). Hence, given a sufficiently small target symbol error probability (SER) \( P_e(\hat{H}) = \epsilon \), the data rate \( R_i \) (bits per symbol) of the \( i \)-th data stream is related to the \( SINR_i(P) \) as

\[ R_i = \log_2(1 + \alpha(\epsilon)SINR_i(P)) \]

(6)

where \( \alpha(\epsilon) \) is some constant depending on the target SEP \( \epsilon \). Since the receiver has perfect CSIR and the data rate is an increasing function of \( SINR_i \), it is shown that for any precoder \( P \), Wiener filter \( W = (H^H H + I)^{-1} H \) can simultaneously maximize \( \{ SINR_1, ..., SINR_L \} \). As a result, the conditional average SINR of the \( i \)-th data stream after Wiener filtering is given by:

\[ SINR_i(P) = E[|w_i^H H A^{-1}_i H P s|^2] \]

(7)

Define the instantaneous MSE matrix as follows:

\[ E(P) = E \left[ (s - \hat{s})(s - \hat{s})^H \right] = (I + P^H H^H H P)^{-1} \]

(8)

Note that the diagonal elements of \( E \) contains the instantaneous MSEs of the \( L \) data streams. Using matrix inversion lemma, it can be shown that

\[ SINR_i(P) \geq E[|w_i|^2] - 1 \]

(9)
Hence, we have a lower bound for the average supported data rate (conditioned on $\mathbf{H}$) at the target SER $\epsilon$ given by:

$$R_t \geq \log_2(1 + \alpha(\epsilon)(\mathbf{E}_{ii^{-1}}(\mathbf{P}) - 1)) \quad (9)$$

where $\mathbf{E}_{ii}(\mathbf{P}) = \mathbb{E} \left[ \mathbf{E}(\mathbf{P}) | \mathbf{H} \right]$.

### B. Queue Model, System States and Control Policy

In this paper, the time dimension is partitioned into scheduling slots (each slot has $\tau$ channel uses) and we assume that the CSI $\mathbf{H}$ remains quasi-static within a scheduling slot and i.i.d. between scheduling slots. There are $L$ buffers (of length $N$) at the transmitter for the $L$ application streams respectively. For simplicity, we assume the $L$ application sources follow Poisson arrival with arrival rates $\lambda_1, ..., \lambda_L$ (number of packets per channel use). The packet length of the $i$-th data source, $N_i$, follows exponential distribution with mean packet size $\bar{N}$ (bits per packet). The transmitter is assumed to have knowledge of the queue state information (QSI) of the $L$ buffers. Specifically, the QSI at time $t$ is denoted by $Q(t) = (Q_1(t), ..., Q_L(t)) \in \{0, ..., N\}^L$ where $Q_i(t)$ is the number of packets in the $i$-th buffer at time $t$. As a result, the observed system state at the transmitter, $\chi = (\mathbf{H}, Q)$, consists of both the outdated CSIT and the joint QSI. Given an observed system state realization $\chi$, the transmitter may adjust the precoding matrix $\mathbf{P}$ according to a stationary precoding policy $\pi = \{\mathbf{P}(\chi)\}$ defined below.

**Definition 1 (Stationary Precoding Policy):** A stationary precoding policy $\pi : \{0, 1, ..., N\}^L \times \mathbb{N}_+ \times \mathbb{N} \rightarrow \mathbb{N}_+ \times \mathbb{N}_+$ is defined as the mapping from the currently observed system state $\chi = (\mathbf{H}, Q)$ to a linear precoder $\pi(\chi) = \mathbf{P}(\chi)$. The set of all feasible stationary policies is defined as $\mathcal{P} = \{\pi : \text{Tr} (\mathbb{E} [\pi(\chi)\mathbf{P}(\chi) | \mathbf{Q})] > 0 \forall \mathbf{Q} \in \{0, 1, ..., N\}^L\}$.

Note that the precoding policy $\pi = \{\mathbf{P}(\chi)\}$ here includes both unitary precoding and power allocation, i.e. the linear precoder $\mathbf{P}$ can be decomposed into $\mathbf{U}_P\Sigma_P\mathbf{V}_P$, where $\mathbf{U}_P$ and $\mathbf{V}_P$ are unitary matrices (denoting the precoding actions) and $\Sigma_P$ is a diagonal matrix (denoting the power allocation action). And the transmitter may adjust the precoding and power control actions only at the beginning of scheduling slots and the control action remains unchanged in between the scheduling slots.

Since the packet length is exponentially distributed with mean packet length $\bar{N}$, from (9), the packet service time follows exponential distribution with mean service rate (conditioned on system state $\chi$) [packets per channel use]:

$$\mu_1(\chi) = \frac{1}{\bar{N}} \log_2(1 + \alpha(\epsilon)(\mathbf{E}_{ii^{-1}}(\mathbf{P}) - 1)). \quad (10)$$

The overall delay dynamics of the $L$-stream multiplexed MIMO system can thus be modeled by a $L/M/M/1$ queue, which are coupled together via the precoding policy $\mathbf{P}$. The average delay of the $i$-th data stream is given by:

$$T_i(\pi) \leq \lim_{M \rightarrow \infty} \mathbb{E}\left[ \frac{\sum_{m=1}^{M} Q_i(m)}{M} \right] \quad (11)$$

where $Q_{i,m} = Q_i(m\tau)$ is the QSI of the $i$-th buffer observed at $t = m\tau$, and the average transmit power constraint is

$$\mathcal{P}_{tx}(\pi) = \lim_{M \rightarrow \infty} \mathbb{E}\left[ \sum_{m=1}^{M} \text{Tr}(\pi(\chi_m)\mathbf{P}(\chi_m)) \right] \leq P_0 \quad (12)$$

where $\pi(\chi_m)$ denotes the precoder applied at $t = m\tau$. The optimization problem to minimize the average delay of the $L$ data streams corresponds to a multi-objective optimization problem. The problem is convex and we could study the Pareto optimal delay tradeoff among the $L$ data streams which can be obtained by solving the following scalarized dual problem.

**Problem 1 (Delay Optimal Policy):** For some $\beta = (\beta_1, \beta_2, ..., \beta_L)$ (such that $\beta_i > 0$ for all $i$), we seek to find a stationary policy $\pi \in \mathcal{P}$ that minimizes

$$J^\pi(\chi_0) = \sum_{i=1}^{L} \beta_i T_i(\pi) + \gamma \mathcal{P}_{tx}(\pi). \quad (13)$$

where $\chi_0$ denotes the initial system state. The positive weighting factors $\beta_i$ indicate the relative importance of buffer delay among the $L$ data streams and for each given $\beta$, the solution to (13) corresponds to a point on the Pareto optimal delay tradeoff boundary. The constant $\gamma > 0$ is the Lagrange multiplier for the average transmit power constraint in (12).

### III. MARKOV DECISION PROBLEM FORMULATION

#### A. Embedded Markov Chain and MDP Formulation

Recall that $\{Q(t)\}$ is the continuous time random process (denoting the joint queue state of the $L$ data streams) and $\{Q_m\}$ is the corresponding induced discrete time random process (denoting the joint queue states at observation epochs $\{0, \tau, 2\tau, ...\}$) with $Q_m = Q(m\tau)$. We consider the case where the scheduling slot duration $\tau$ is substantially smaller than the average packet interarrival time as well as the average packet service time, i.e. $\tau \ll \frac{1}{\lambda}$ and $\tau \ll \frac{1}{\mu_1(\chi)}$. Suppose the system state at the $m$-th observation epoch is $\chi_m$. At the
The solution to (17) is very complex due to the coupled $L$-dimensional recursions and brute-force solutions have exponential order of complexity, i.e. $O((N+1)^L)$. We now give the following theorem to show the underlying structure of the optimal precoding structure for $\pi(\chi)$, which provides useful insight for designing the low complexity solution.

**Theorem 1 (Optimal Precoding Matrix):** For any realization of system state $\chi$ (QSI $\mathbf{Q}$, CSIT $\mathbf{H}$), the optimal precoding action $\pi(\chi) = \mathbf{P}$ is given by:

$$\pi(\chi) = \mathbf{P} = \mathbf{U} \Sigma_p$$  \hspace{1cm} (18)

where $\mathbf{U} \in \mathbb{C}^{N \times L}$ is a unitary matrix consisting of $L$ eigenvectors of $\hat{\mathbf{H}}^H \mathbf{H} + N_0 \mathbf{I}$ corresponding to the $L$ largest eigenvalues and $\Sigma_p = \text{diag}\{\sqrt{\nu_1}, \ldots, \sqrt{\nu_L}\}$ is a diagonal matrix containing the power allocations over the $L$ spatial channels. Note that the $L$ largest eigenvalues $\{
u_1, \ldots, \nu_L\}$ are sorted in the same order as $\eta_i = V(q_1, \ldots, q_L) - V(q_1, \ldots, [q_i - 1]^+, \ldots, q_L)$.

**Proof:** Please refer to our full version in [11].

As we can see from this theorem, the main reason that the $L$ data streams are coupled together comes from the demand of sorting the $L$ largest eigenvalues in the same ordering as $\{V(q_1, \ldots, q_L) - V(q_1, \ldots, [q_i - 1]^+, \ldots, q_L)\}$, which is a function of the QSI $\mathbf{Q}$. Therefore, the key idea of our low complexity solution lies in relaxing the sorting requirement in a reasonable way so that the $L$-dimensional MDP problem can be decomposed into $L$ one-dimensional subproblems.

### IV. LOW COMPLEXITY SOLUTION

#### A. Decomposition of the MDP

In the low complexity solution, we propose a static eigenchannel sorting scheme. In particular, we sort the $L$ largest eigenvalues $\xi_1, \ldots, \xi_L$ in the same ordering as $\beta_1, \ldots, \beta_L$ (which represents the relative importance of the $L$ data streams)$^3$. Given a stationary power control policy, $\varphi = \{\varphi_1, \ldots, \varphi_L\}$, the MDP state transition probabilities under the proposed static sorting scheme are thus decomposable among the $L$ data streams. The average cost per stage in (14) under $\varphi = \{\varphi_1, \ldots, \varphi_L\}$ can be decomposed as $J^\varphi_\beta = \sum_{l=1}^L J^\varphi_{\beta,i}$, where

$$J^\varphi_{\beta,i} = \lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^M g_i(Q_i,m,\varphi_i(Q_i,m))$$  \hspace{1cm} (19)

and

$$g_i(Q_i,m,\varphi_i(Q_i,m)) = \beta_i Q_i,m + \gamma \varphi_i(Q_i,m)$$  \hspace{1cm} (20)

Hence, the original “minimal average cost per stage” problem $J^\varphi_\beta = \inf_{\varphi} J^\varphi_\beta$ can be decomposed into $L$ individual subproblems $J^\varphi_{\beta,i} = \inf_{\varphi_i} J^{\varphi_i}_{\beta,i}$ for $i = 1, \ldots, L$. Consider the $i$-th subproblem, the Bellman equation is given by:

$$\theta_i + V_i(q) = \inf_{\varphi_i(q)} \left\{ g_i(q,\varphi_i(q)) + \tau \lambda_i V_i(q + 1) \right\}$$  \hspace{1cm} (22)

$^3$While this is suboptimal in strict sense, it will not cause too much performance loss especially for highly asymmetric cases ($\beta_1 \gg \beta_2 \gg \ldots \gg \beta_L$) or highly symmetric case $\beta_1 \approx \beta_2 \approx \ldots \approx \beta_L$. 

$(m + 1)$--th observation epoch $t = (m + 1)\tau$, one of the following events may happen: 1) packet arrival from the $i$-th data source; 2) Packet departure from the $i$-th data buffer; 3) No change in the $i$-th buffer state. The state transition probability is illustrated in Fig. 2, where

$$\bar{p}_i(Q_m) = \mathbb{E}[\mu_i(X_m)|Q_m = (q_1, \ldots, q_i = q_i, \ldots, q_L)]$$

is the average (over CSIT) service rate which depends on the current QSI only. Furthermore, since the slot duration $\tau$ is small, the probability of multiple packet arrivals or packet departures among the $L$ data sources is negligible and hence $p_{i,p}^{(1)} = 0$ for $|p - q| > 1$.

From the above, the embedded discrete time random variables $\{Q_m\}$ is an irreducible Markov chain induced by a stationary policy $\pi \in \mathcal{P}$. The optimization objective function (average cost per stage) $J^\pi_\beta(\chi_0)$ evaluated at the discrete time observation epochs can be expressed as:

$$J^\pi_\beta(\chi_0) = \limsup_M \frac{1}{M} \sum_{m=1}^M \mathbb{E}_Q [g(Q_m,\pi(Q_m))]$$  \hspace{1cm} (14)

where

$$g(Q_m,\pi(Q_m)) = \sum_{i=1}^L \beta_i Q_{i,m} + \gamma \mathbb{E}[\pi(Q_m)]$$  \hspace{1cm} (15)

and

$$\pi(Q_m) = \mathbb{E}[\pi(\chi_m)\pi^H(\chi_m)|Q_m]$$  \hspace{1cm} (16)

Hence, the delay-optimization problem in Problem 1 could be completely characterized by a multi-dimensional infinite horizon Markov Decision Process (MDP) with partial system state $Q$, per-stage reward function $g(Q,\pi(Q))$, and the conditional average precoding action $\pi(Q)$. The state transition probability of the embedded MDP $\bar{p}_i(Q_{m+1},Q_m,\pi(Q_m))$ is illustrated in Fig. 2.

#### B. Bellman Condition and Optimal Precoding Structure

For the infinite horizon MDP described by Fig. 2, the optimizing policy can be obtained by solving the Bellman equation [10] recursively w.r.t. $(\theta, \{V(q_1, \ldots, q_L)\})$ as follows:

$$\theta + V(q_1, \ldots, q_L) = \inf_{\pi \in \mathcal{P}} \left\{ g(q_1, \ldots, q_L, \pi(q_1, \ldots, q_L)) \right\}$$

$$+ \tau \sum_{i=1}^L \lambda_i V(q_1, \ldots, [q_i + 1] \Lambda N, \ldots, q_L)$$

$$+ \tau \sum_{i=1}^L \bar{p}_i(q_1, \ldots, q_L) V(q_1, \ldots, [q_i - 1]^+, \ldots, q_L)$$

$$+ V(q_1, \ldots, q_L) \left( 1 - \sum_{i=1}^L \tau \lambda_i - \tau \sum_{i=1}^L \bar{p}_i(q_1, \ldots, q_L) \right)$$  \hspace{1cm} (17)

for all $(q_1, \ldots, q_L) \in \{0, 1, \ldots, N\}^L$ where $x_\Lambda y = \min\{x,y\}$. If there is a $(\theta, \{V(q_1, \ldots, q_L)\})$ satisfying (17), then $\theta = \inf_{\pi \in \mathcal{P}} J^\pi_\beta$ is the optimal average reward per stage. The existence and uniqueness of the solution is guaranteed since the induced Markov chain $\{Q_m\}$ is irreducible for any stationary policy $\pi \in \mathcal{P}$.
for all $q \in \{0, 1, \ldots, N\}$, in which $\mu_i(q)$ is given by
\[
\mu_i(q) = \frac{1}{N} E \left[ \log_2(1 + \alpha(q) \varphi_i(\lambda)) \right] |_{\lambda=q, m=q},
\]
$\xi_i$ is the $i$-th eigenvalue of $\hat{H}^H \hat{H} + N \mathbf{I}$ (sorted in the same order as $\{\beta_1, \ldots, \beta_L\}$) and $\varphi_i(q) = \{p_i = \varphi_i(H, Q_i = q) : H \in \mathbb{C}^{N \times N}\}$ denotes the set of power allocation actions for all CSIT realizations at a given QSI $Q_i = q$. Since the embedded Markov chain $\{Q_{i,m}\}$ is irreducible, there is a unique solution $\theta_i$, $i=1,\ldots, L$ and $\delta V_i(0) = 0$, and for all $\lambda = 0, \ldots, N-1$ with two boundary conditions that $\beta_1 = \delta V_i(N) + \theta_1$ and $\delta V_i(0) = 0$.

The solution to the decoupled Bellman equation is given by
\[
\delta V_i(1, \theta) = \frac{\theta}{\beta_1},
\]
\[
\delta V_i(q + 1, \theta) = \frac{(\theta + \phi(\delta V_i(q, \theta))) - \beta_i q}{\lambda_i},
\]
for $q = 1, 2, \ldots, N-1$. Define
\[
f_i(\theta) = \frac{\phi(\delta V_i(N, \theta)) + \theta}{\beta_1}.
\]
The tuple $(\theta, \delta V_i(1, \theta), \ldots, \delta V_i(N, \theta))$ is a solution to the Bellman equation in (24) if and only if $f_i(\theta) = N$. Note that $f_i(\theta)$ is continuous, strictly increasing in $\theta$ and convex, so the inverse function of $f_i(\theta)$ exists and is also continuous, strictly increasing. Hence, there exists a unique $\theta_i^* = f_i^{-1}(N)$ such that $f_i(\theta_i^*) = N$. In other words, $(\theta_i^*, \delta V_i(1, \theta_i^*), \ldots, \delta V_i(N, \theta_i^*))$ is the unique solution satisfying the Bellman equation in (24), and $\theta_i^*$ can be obtained easily by one-dimensional bisection method whose computational complexity could be treated as a known constant.

Using standard optimization technique to solve for (25), the optimal power allocation policy $\varphi_i(q)^*$ (for a given QSI $Q_i = q$) is given by $\varphi_i(q)^* = \{p_i^*(H, q)\}$ where
\[
p_i^*(H, q) = \left( \frac{\delta V_i(q, \theta_i^*)}{\gamma_N} - \frac{1}{\alpha(q) \xi_i} \right) 
\]
for $n = 1, 2, \ldots, N$ and $p_i^*(H, 0) = 0$. As a result, the power allocation solution depends on the QSI only via the equivalent water-level $\gamma^{-1} = \delta V_i(q, \theta_i^*)/\gamma N$. For larger queue size, the equivalent water-level $\gamma^{-1}$ is increased. This result is also consistent with the asymptotic delay-optimal solution for point-to-point single-stream system in [6].

Using the optimal power allocation policy $\varphi_i^*(q)$ for $q = 0, 1, 2, \ldots, N$, the embedded Markov chain $\{Q_{i,m}\}$ of the $i$-th data stream is ergodic and time reversible. The steady state distribution $\Omega(\varphi_i^*) = (\omega_0(\varphi_i^*), \omega_1(\varphi_i^*), \ldots, \omega_N(\varphi_i^*))$ of the queue lengths under the optimal policy $\varphi_i^*$ can be obtained by solving the $L$ one-dimensional detailed balance equations:
\[
\lambda_i \omega_q(\varphi_i^*) = \frac{\mu_i^*(q + 1)}{\mu_i^*(q)}, \quad \forall q = 0, 1, \ldots, N-1
\]
with the overall constraint $\sum_{q=0}^{N} \omega_q(\varphi_i^*) = 1$. As a result, the steady state distribution $\Omega(\varphi_i^*)$ is given by:
\[
\omega_q(\varphi_i^*) = \frac{\prod_{j=0}^{N-1} \frac{\mu_j^*(j)}{\lambda_j}}{\sum_{q=0}^{N} \omega_q(\varphi_i^*)}.
\]
The average delay of the $i$-th data stream is thus given by $T_i^*(\varphi_i^*) = \sum_{q=0}^{N} q \omega_q(\varphi_i^*)$. And finally, the common Lagrange multiplier $\gamma$ among the $L$ data streams can be determined by substituting (26) for $i = 1, \ldots, L$ into (12).

C. Summary of the Low Complexity Solution

The optimal precoding policy consists of an online procedure and an offline procedure, which are summarized below.

**Offline Procedure**

- **Step 1)** Bellman Solution: For $i = 1, \ldots, L$, and a $\gamma$, determine $\{\theta_1^*(\gamma), \ldots, \theta_L^*(\gamma)\}$ according to (26) as well as $\{\delta V_i(q, \theta_1^*(\gamma)), \ldots, \delta V_i(q, \theta_L^*(\gamma))\}$ according to (26).
- **Step 2)** Transmit Power Constraint: Solve for $\gamma$ that satisfies the transmit power constraint in (12) using one dimensional root-finding numerical algorithm.

**Output:** $\{\delta V_i(q, \theta_1^*(\gamma(P_0))), \ldots, \delta V_i(q, \theta_L^*(\gamma(P_0)))\}$, $\gamma(P_0)$, and $\theta_1^*(\gamma(P_0)), \ldots, \theta_L^*(\gamma(P_0))$, which shall serve as inputs to the online procedure.

**Online Procedure**

- **Step 1)** SVD on CSIT: Given the current CSIT $\hat{H}$, obtain the largest $L$ eigenvalues $\xi_1 \leq \xi_2 \leq \cdots \leq \xi_L$ of the matrix $\hat{H}^H \hat{H} + N \mathbf{I}$ and the corresponding eigenvectors (as columns of $U$).
- **Step 2)** Precoder and Data Stream Mapping: The optimal precoder $P = U \Sigma_p$ where $\Sigma_p = diag(\sqrt{p_1}, \ldots, \sqrt{p_L})$ and $U \in \mathbb{C}^{N \times L}$ contains the $L$ eigenvectors obtained in Step 1 as columns. In addition, the $L$ largest eigenvalues are sorted dynamically$^4$ in the same order as $\{\delta V_i(q, \theta_1^*), \ldots, \delta V_i(q, \theta_L^*)\}$ which is a function of the current QSI.

$^4$Note that the static sorting applies to the offline procedure only to obtain simple solution for $\{\delta V_i(q, \theta_1^*)\}$. In the online algorithm, we can follow dynamic sorting of eigenvalues as in theorem 1.
• **Step 3) Optimal Power Allocation:** Based on the precoder and data stream index association in step 2, the power allocation of the \(i\)-th data stream is given by \(p^*_i(\hat{H}, Q) = p^*_i(\hat{H}, q_i)\) according to (26).

### V. Numerical Results and Discussions

We consider \(L = 2\) data streams, each with buffer size \(N = 4\). The mean packet size \(\bar{N}\) is 100 bits per packet, and the arrival rate of each streams \(\lambda_i = 0.05\) packets per channel use, \(i = 1, 2\). The scheduling time unit \(\tau\) and the target SER \(\varepsilon\) are fixed at 5ms and 1%, respectively.

Fig. 3 depicts the average delay of the two streams under different configurations of transmit and receive antennas. Obviously, increasing the number of either transmit or receive antennas results in decreased average delay. Moreover, since only \(L \leq \min\{N_t, N_r\}\) streams are considered, increasing the number of transmit antenna when \(N_t \geq L\) is not as effective as increasing the number receive antennas.

In Fig. 4, the two streams are set to have equal importance, i.e., \(\beta_1 = \beta_2 = 1\), and we compare the sum average delay of both streams for a 2-by-2 MIMO system. Two baselines are taken into comparison: Round-Robin scheme and CSIT only scheme. In the former case, the two streams are serviced in TDMA fashion with equally allocated time slots, while in the latter one, the precoder is designed purely based on the outdated CSIT available, but with no QSI. Great improvement of our proposed scheme over the two baselines can be observed: about 10dB gain compared with Round-Robin scheme and even larger gain compared with CSIT only scheme. This fact also shows that spatial multiplexing may not help effectively without adapting to the QSI. Another message delivered by Fig. 4 is that, the effect of CSIT error variance on CSIT only scheme is much larger than that on the proposed scheme, which verifies the robustness of our proposed scheme to channel estimation error.

### VI. Summary

We considered delay sensitive MIMO system with \(L\) heterogeneous data streams spatially multiplexed together. The design of precoding policy achieving Pareto optimal delay tradeoff is formulated into an \(L\)-dimensional MDP problem. A low complexity solution with worst case complexity \(O(N L)\) is proposed by decomposing the original problem into \(L\) one-dimensional subproblems based on static sorting. Numerical results verify the advantages of taking both QSI and CSIT error into dynamic precoder design.

### References