Bi-Directional Training in Interference Network

Shao-Han Chen

February 25, 2016

1 System Model 1: Private and Common Messages

Under this model, we want to see how cooperation on transmitter(base station) side will improve the channel capacity. One way we implement the cooperation scheme is to make all transmitters for common messages, instead of working separately, a big cooperative array. That is to say, for the backward training part, we **minimize the mean square error(MSE) of total received signals together**. However, we found this method is actually **worse** than the one without cooperation. **How Doesn't this Work?** This is related the way we **measure performance**. Our objective is to maximize the **sum capacity** of individual users, which is equal to minimize $\sum_k MSE_k$. However, our cooperation scheme actually leads to the solution of maximizing total capacity, that is minimize MSE. Please note that we are not claiming cooperations imply worse performance. It's just this particular cooperation scheme doesn't work.

1.1 System Model

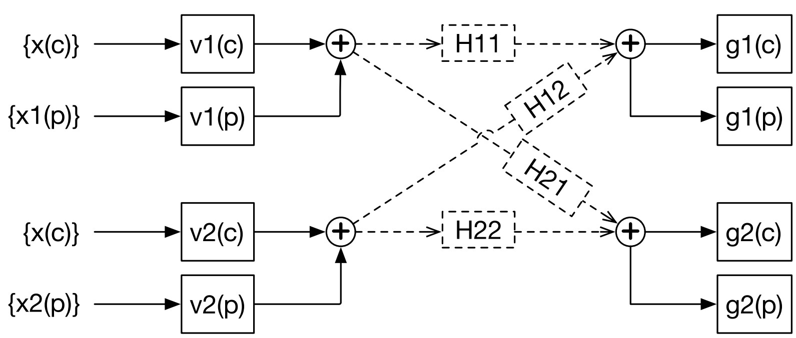


Figure 1: Forward Channel

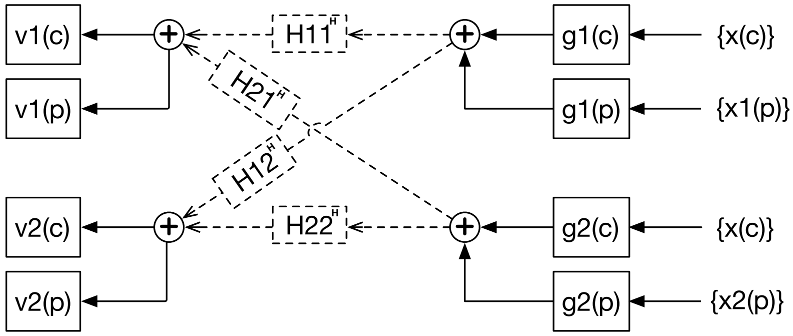


Figure 2: Backward Channel

1.2 Optimization Problem

$$\min_{\mathbf{v}_k,\mathbf{g}_k} \sum_k MSE_k^{(c)} + MSE_k^{(p)}$$

$$\text{subject to } \|\mathbf{v}_k^{(c)}\|^2 + \|\mathbf{v}_k^{(p)}\|^2 = P; \|\mathbf{g}_k\|^2 = P$$

where

$$MSE_k^{(c)} = \mathrm{E}[(x - \mathbf{g}_k^{H(c)}\mathbf{y}_k)(x - \mathbf{g}_k^{H(c)}\mathbf{y}_k)^H]$$

$$MSE_k^{(p)} = \mathrm{E}[(x_k^{(p)} - \mathbf{g}_k^{H(p)}\mathbf{y}_k)(x_k^{(p)} - \mathbf{g}_k^{H(p)}\mathbf{y}_k)^H]$$

1.3 The received signal vector at k-th receiver

$$\mathbf{y}_k = \mathbf{H}_{kk}(\mathbf{v}_k^{(c)}x + \mathbf{v}_k^{(p)}x_k^{(p)}) + \sum_{j\neq k}\mathbf{H}_{kj}(\mathbf{v}_j^{(c)}x + \mathbf{v}_j^{(p)}x_j^{(p)}) + \mathbf{n}_k$$

1.4 SINR Derivation

$$s_k^{(c)} = \mathbf{g}_k^{H(c)}(\sum_i \mathbf{H}_{ki}\mathbf{v}_i^{(c)}x)$$

$$s_k^{(p)} = \mathbf{g}_k^{H(p)}(\mathbf{H}_{kk}\mathbf{v}_k^{(p)}x_k^{(p)})$$

$$n_k^{(c)} = \mathbf{g}_k^{H(c)}(\sum_i \mathbf{H}_{ki}\mathbf{v}_i^{(p)} x_i^{(p)} + \mathbf{n}_k)$$

$$n_k^{(p)} = \mathbf{g}_k^{H(p)}(\sum_i \mathbf{H}_{ki}\mathbf{v}_i^{(c)} x + \sum_{j\neq k} \mathbf{H}_{kj}\mathbf{v}_j^{(p)} x_j^{(p)} + \mathbf{n}_k)$$

$$\frac{|s_k^{(c)}|^2}{|n_k^{(c)}|^2} = \frac{|\mathbf{g}_k^{H(c)} \sum_i \mathbf{H}_{ki}\mathbf{v}_i^{(c)}|^2}{\sum_i |\mathbf{g}_k^{H(c)}\mathbf{H}_{ki}\mathbf{v}_i^{(p)}|^2 + |\mathbf{g}_k^{H(c)}\mathbf{R}_k\mathbf{g}_k^{(c)}|}$$

$$\frac{|s_k^{(p)}|^2}{|n_k^{(p)}|^2} = \frac{|\mathbf{g}_k^{H(p)}\mathbf{H}_{kk}\mathbf{v}_k^{(p)}|^2}{|\mathbf{g}_k^{H(p)} \sum_i \mathbf{H}_{ki}\mathbf{v}_i^{(c)}|^2 + \sum_{j\neq k} |\mathbf{g}_k^{H(p)}\mathbf{H}_{kj}\mathbf{v}_j^{(p)}|^2 + |\mathbf{g}_k^{H(p)}\mathbf{R}_k\mathbf{g}_k^{(p)}|}$$

1.5 Max-SINR Algorithm[Gomadam, 2011]

For Max-SINR algorithm, we assume **all channel state information(CSI) is available** to each user. In this method, the solution for each transceivers are simply Wiener filters. In other words, solving it by Wiener-Hopf equation: $R^{-1}p$ [Adaptive Filter Theory, Simon Haykin]

1.5.1 Forward Training (fix $\mathbf{v}_k^{(c)}$, $\mathbf{v}_k^{(p)}$, $\forall k$)

$$\begin{aligned}\mathbf{g}_k^{*(c)} =& \Big[\mathbf{H}_{kk}\mathbf{v}_k^{(c)}\mathbf{v}_k^{H(c)}\mathbf{H}_{kk}^H + \mathbf{H}_{kk}\mathbf{v}_k^{(p)}\mathbf{v}_k^{H(p)}\mathbf{H}_{kk}^H \\ &+ (\sum_{j\neq k}\mathbf{H}_{kj}\mathbf{v}_j^{(c)})(\sum_{j\neq k}\mathbf{v}_j^{H(c)}\mathbf{H}_{kj}^H) + (\sum_{j\neq k}\mathbf{H}_{kj}\mathbf{v}_j^{(p)}\mathbf{v}_j^{H(p)}\mathbf{H}_{kj}^H) \\ &+ \mathbf{H}_{kk}\mathbf{v}_k^{(c)}(\sum_{j\neq k}\mathbf{v}_j^{H(c)}\mathbf{H}_{kj}^H) + (\sum_{j\neq k}\mathbf{H}_{kj}\mathbf{v}_j^{(c)})\mathbf{v}_k^{H(c)}\mathbf{H}_{kk}^H + \sigma^2\mathbf{I}\Big]^{-1}(\sum_i \mathbf{H}_{ki}\mathbf{v}_i^{(c)})\end{aligned}$$

$$\begin{aligned}\mathbf{g}_k^{*(p)} =& \Big[\mathbf{H}_{kk}\mathbf{v}_k^{(c)}\mathbf{v}_k^{H(c)}\mathbf{H}_{kk}^H + \mathbf{H}_{kk}\mathbf{v}_k^{(p)}\mathbf{v}_k^{H(p)}\mathbf{H}_{kk}^H \\ &+ (\sum_{j\neq k}\mathbf{H}_{kj}\mathbf{v}_j^{(c)})(\sum_{j\neq k}\mathbf{v}_j^{H(c)}\mathbf{H}_{kj}^H) + (\sum_{j\neq k}\mathbf{H}_{kj}\mathbf{v}_j^{(p)}\mathbf{v}_j^{H(p)}\mathbf{H}_{kj}^H) \\ &+ \mathbf{H}_{kk}\mathbf{v}_k^{(c)}(\sum_{j\neq k}\mathbf{v}_j^{H(c)}\mathbf{H}_{kj}^H) + (\sum_{j\neq k}\mathbf{H}_{kj}\mathbf{v}_j^{(c)})\mathbf{v}_k^{H(c)}\mathbf{H}_{kk}^H + \sigma^2\mathbf{I}\Big]^{-1}(\mathbf{H}_{kk}\mathbf{v}_k^{(p)})\end{aligned}$$

1.5.2 Backward Training (fix $\mathbf{g}_k^{(c)}$, $\mathbf{g}_k^{(p)}$, $\forall k$)

$$\mathbf{Z}_{ab} = \mathbf{H}_{ba}^H$$

without cooperation

$$\begin{aligned}\mathbf{v}_k^{*(c)} = \Big[& \mathbf{Z}_{kk}\mathbf{g}_k^{(c)}\mathbf{g}_k^{H(c)}\mathbf{Z}_{kk}^H + \mathbf{Z}_{kk}\mathbf{g}_k^{(p)}\mathbf{g}_k^{H(p)}\mathbf{Z}_{kk}^H \\ & + (\sum_{j\neq k}\mathbf{Z}_{kj}\mathbf{g}_j^{(c)})(\sum_{j\neq k}\mathbf{g}_j^{H(c)}\mathbf{Z}_{kj}^H) + (\sum_{j\neq k}\mathbf{Z}_{kj}\mathbf{g}_j^{(p)}\mathbf{g}_j^{H(p)}\mathbf{Z}_{kj}^H) \\ & + \mathbf{Z}_{kk}\mathbf{g}_k^{(c)}(\sum_{j\neq k}\mathbf{g}_j^{H(c)}\mathbf{Z}_{kj}^H) + (\sum_{j\neq k}\mathbf{Z}_{kj}\mathbf{g}_j^{(c)})\mathbf{g}_k^{H(c)}\mathbf{Z}_{kk}^H + \sigma^2\mathbf{I}\Big]^{-1}(\sum_i \mathbf{Z}_{ki}\mathbf{g}_i^{(c)})\end{aligned}$$

with cooperation

$$\mathbf{V}^{*(c)} = \left\{ [\mathbf{Z}]\mathbf{g}^{(c)}\mathbf{g}^{H(c)}[\mathbf{Z}]^H + [\mathbf{Z}]\begin{bmatrix}\mathbf{g}_1^{(p)}\mathbf{g}_1^{H(p)} & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \mathbf{g}_k^{(p)}\mathbf{g}_k^{H(p)}\end{bmatrix}[\mathbf{Z}]^H + \sigma^2\mathbf{I}\right\}^{-1}([\mathbf{Z}]\mathbf{g}^{(c)})$$

$$\begin{aligned}\mathbf{v}_k^{*(p)} = \Big[& \mathbf{Z}_{kk}\mathbf{g}_k^{(c)}\mathbf{g}_k^{H(c)}\mathbf{Z}_{kk}^H + \mathbf{Z}_{kk}\mathbf{g}_k^{(p)}\mathbf{g}_k^{H(p)}\mathbf{Z}_{kk}^H \\ & + (\sum_{j\neq k}\mathbf{Z}_{kj}\mathbf{g}_j^{(c)})(\sum_{j\neq k}\mathbf{g}_j^{H(c)}\mathbf{Z}_{kj}^H) + (\sum_{j\neq k}\mathbf{Z}_{kj}\mathbf{g}_j^{(p)}\mathbf{g}_j^{H(p)}\mathbf{Z}_{kj}^H) \\ & + \mathbf{Z}_{kk}\mathbf{g}_k^{(c)}(\sum_{j\neq k}\mathbf{g}_j^{H(c)}\mathbf{Z}_{kj}^H) + (\sum_{j\neq k}\mathbf{Z}_{kj}\mathbf{g}_j^{(c)})\mathbf{g}_k^{H(c)}\mathbf{Z}_{kk}^H + \sigma^2\mathbf{I}\Big]^{-1}(\mathbf{Z}_{kk}\mathbf{g}_k^{(p)})\end{aligned}$$

1.6 Bi-Directional Training with LS Algorithm[Shi, 2014]

For Bi-Directional Training, we assume **all channel state information(CSI) is not available** to each user. In this case, we could either adopt Least Mean Square Method(LMS) or Least Square Method(LS) to carry out the solution[Adaptive Filter Theory, Simon Haykin]. **There is actually another interesting finding here**. We all have seen the monotonically decreasing diagram of MSE when designing/studying filters with both methods. We can also show that when the filter converges(achieve minimum MSE), it also maximizes Signal to Interference plus Noise Ratio(SINR). **However, how does the SINR behave before it converges?** For LS, it is monotonically decreasing just as MSE diagram, but for LMS, it is **not!** Therefore, LMS isn't applicable for our application. We need SINR increases after each iteration.

1.6.1 Forward Training (fix $\mathbf{v}_k^{(c)}$, $\mathbf{v}_k^{(p)}$, $\forall k$)

$$\mathbf{g}_k^{(c)}(n+1) = \mathbf{g}_k^{(c)}(n) + \mu \mathbf{y}_k(n)[x(n) - \mathbf{g}_k^{H(c)}(n)\mathbf{y}_k(n)]^*$$

$$\mathbf{g}_k^{(p)}(n+1) = \mathbf{g}_k^{(p)}(n) + \mu \mathbf{y}_k(n)[x_k^{(p)}(n) - \mathbf{g}_k^{H(p)}(n)\mathbf{y}_k(n)]^*$$

1.6.2 Backward Training (fix $\mathbf{g}_k^{(c)}$, $\mathbf{g}_k^{(p)}$, $\forall k$)

without cooperation

$$\mathbf{v}_k^{(c)}(n+1) = \mathbf{v}_k^{(c)}(n) + \mu \mathbf{y}_k(n)[x(n) - \mathbf{v}_k^{H(c)}(n)\mathbf{y}_k(n)]^*$$

$$\mathbf{v}_k^{(p)}(n+1) = \mathbf{v}_k^{(p)}(n) + \mu \mathbf{y}_k(n)[x_k^{(p)}(n) - \mathbf{v}_k^{H(p)}(n)\mathbf{y}_k(n)]^*$$

with cooperation

$$\mathbf{V}^{(c)}(n+1) = \mathbf{V}^{(c)}(n) + \mu \mathbf{Y}(n)[x(n) - \mathbf{V}^{H(c)}(n)\mathbf{Y}(n)]^*$$

$$\mathbf{v}_k^{(p)}(n+1) = \mathbf{v}_k^{(p)}(n) + \mu \mathbf{y}_k(n)[x_k^{(p)}(n) - \mathbf{v}_k^{H(p)}(n)\mathbf{y}_k(n)]^*$$

1.7 Special Case(2 Users, MIMO Channel, Only Common Messages)

$$\mathbf{y}_k = \sum_{i=1}^{2} \mathbf{H}_{ki}\mathbf{v}_i^{(c)} x + \mathbf{n}_k$$

$$\begin{aligned}
MSE_k^{(c)} &= \mathrm{E}[(x - \mathbf{g}_k^{H(c)}\mathbf{y}_k)(x - \mathbf{g}_k^{H(c)}\mathbf{y}_k)^H] \\
&= \mathrm{E}[x^2] - \mathrm{E}[x\mathbf{y}_k^H\mathbf{g}_k^{(c)}] - \mathrm{E}[x\mathbf{g}_k^{H(c)}\mathbf{y}_k] + \mathrm{E}[\mathbf{g}_k^{H(c)}\mathbf{y}_k\mathbf{y}_k^H\mathbf{g}_k^{(c)}] \\
&= 1 - \sum_{i=1}^{2} \mathbf{v}_i^{H(c)}\mathbf{H}_{ki}^H\mathbf{g}_k^{(c)} - \mathbf{g}_k^{H(c)}\sum_{i=1}^{2}\mathbf{H}_{ki}\mathbf{v}_i^{(c)} + \mathbf{g}_k^{H(c)}(\sum_{i=1}^{2}\mathbf{H}_{ki}\mathbf{v}_i^{(c)})(\sum_{i=1}^{2}\mathbf{v}_i^{H(c)}\mathbf{H}_{ki}^H)\mathbf{g}_k^{(c)} + \sigma^2\mathbf{g}_k^{H(c)}\mathbf{g}_k^{(c)}
\end{aligned}$$

$$\begin{aligned}
\mathbf{v}_1^{*(c)} &= \underset{\mathbf{v}_1^{(c)}}{\mathrm{argmin}}(\sum_{k=1}^{2} MSE_k^{(c)}) \\
&= \left[2\mathbf{H}_{11}^H\mathbf{g}_1^{(c)}\mathbf{g}_1^{H(c)}\mathbf{H}_{11} + 2\mathbf{H}_{21}^H\mathbf{g}_2^{(c)}\mathbf{g}_2^{H(c)}\mathbf{H}_{21})\right]^{-1}(2\mathbf{H}_{11}^H\mathbf{g}_1^{(c)} + 2\mathbf{H}_{21}^H\mathbf{g}_2^{(c)} \\
&\quad - \mathbf{g}_1^{H(c)}\mathbf{H}_{12}\mathbf{v}_2^{(c)}\mathbf{H}_{11}^H\mathbf{g}_1^{(c)} - \mathbf{H}_{11}^H\mathbf{g}_1^{(c)}\mathbf{v}_2^{H(c)}\mathbf{H}_{12}^H\mathbf{g}_1^{(c)} - \mathbf{g}_2^{H(c)}\mathbf{H}_{22}\mathbf{v}_2^{(c)}\mathbf{H}_{21}^H\mathbf{g}_2^{(c)} - \mathbf{H}_{21}^H\mathbf{g}_2^{(c)}\mathbf{v}_2^{H(c)}\mathbf{H}_{22}^H\mathbf{g}_2^{(c)})
\end{aligned}$$

$$
\begin{aligned}
\mathbf{v}_2^{*(c)} &= \underset{\mathbf{v}_2^{(c)}}{\operatorname{argmin}}\Big(\sum_{k=1}^{2} MSE_k^{(c)}\Big) \\
&= \Big[2\mathbf{H}_{12}^H \mathbf{g}_1^{(c)} \mathbf{g}_1^{H(c)} \mathbf{H}_{12} + 2\mathbf{H}_{22}^H \mathbf{g}_2^{(c)} \mathbf{g}_2^{H(c)} \mathbf{H}_{22})\Big]^{-1} (2\mathbf{H}_{12}^H \mathbf{g}_1^{(c)} + 2\mathbf{H}_{22}^H \mathbf{g}_2^{(c)} \\
&\quad - \mathbf{g}_1^{H(c)} \mathbf{H}_{11} \mathbf{v}_1^{(c)} \mathbf{H}_{12}^H \mathbf{g}_1^{(c)} - \mathbf{H}_{12}^H \mathbf{g}_1^{(c)} \mathbf{v}_1^{H(c)} \mathbf{H}_{11}^H \mathbf{g}_1^{(c)} - \mathbf{g}_2^{H(c)} \mathbf{H}_{21} \mathbf{v}_1^{(c)} \mathbf{H}_{22}^H \mathbf{g}_2^{(c)} - \mathbf{H}_{22}^H \mathbf{g}_2^{(c)} \mathbf{v}_1^{H(c)} \mathbf{H}_{21}^H \mathbf{g}_2^{(c)})
\end{aligned}
$$

2 System Model 2 : Cooperative Transmitters

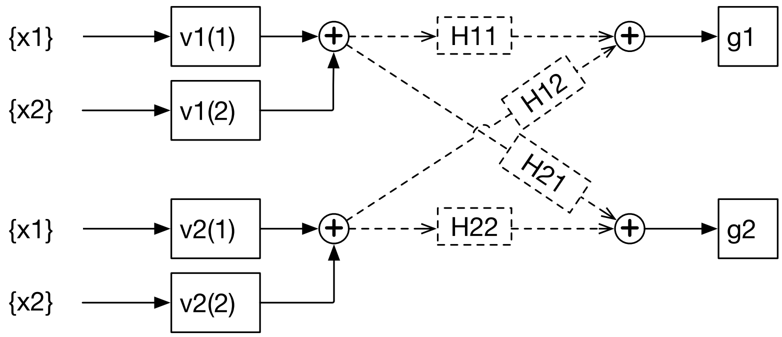


Figure 3: Forward Channel

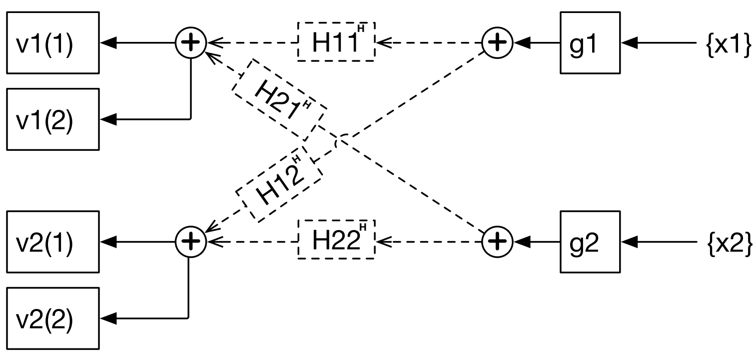


Figure 4: Backward Channel

2.1 Optimization Problem

$$\min_{\mathbf{v}_k^{(j)},\mathbf{g}_k} \sum_k w_k MSE_k$$

$$\text{subject to } \sum_j \|\mathbf{v}_k^{(j)}\|^2 = P; \|\mathbf{g}_k\|^2 = P$$

2.2 The received signal vector at k-th receiver

$$\mathbf{y}_k = \sum_i \left[\mathbf{H}_{ki} \sum_j (\mathbf{v}_i^{(j)} x_j) \right] + \mathbf{n}_k$$

2.3 SINR Derivation

$$s_k = \mathbf{g}_k^H (\sum_i \mathbf{H}_{ki} \mathbf{v}_i^{(k)} x_k)$$

$$n_k = \mathbf{g}_k^H \left[\sum_i \big(\mathbf{H}_{ki} \sum_{j \neq k} \mathbf{v}_i^{(j)} x_j \big) + \mathbf{n}_k \right]$$

$$\frac{|s_k|^2}{|n_k|^2} = \frac{|\mathbf{g}_k^H \sum_i \mathbf{H}_{ki} \mathbf{v}_i^{(k)}|^2}{\sum_{j \neq k} |\mathbf{g}_k^H \sum_i \mathbf{H}_{ki} \mathbf{v}_i^{(j)}|^2 + |\mathbf{g}_k^H \mathbf{R}_k \mathbf{g}_k|}$$

2.4 Forward Direction

The solution is simply $R^{-1}p$.

$$\begin{aligned}
\mathbf{g}_k^* &= \underset{\mathbf{g}_k}{\operatorname{argmin}} (\sum_k \left[MSE_k + \lambda_i (\sum_j \|\mathbf{v}_i^{(j)}\| - P) \right]) \\
&= \left[\sum_j \left[(\sum_i \mathbf{H}_{ki} \mathbf{v}_i^{(j)}) (\sum_i \mathbf{v}_i^{H(j)} \mathbf{H}_{ki}^H) \right] + \sigma^2 \mathbf{I} \right]^{-1} (\sum_i \mathbf{H}_{ki} \mathbf{v}_i^{(k)})
\end{aligned}$$

2.5 Backward Direction: Iterative Method which Doesn't Work

In this method, we found a solution for each $\mathbf{v}_k^{(j)}$ depending on some $\mathbf{v}_x^{(j)}$, where k $\neq$ x. For example, $\mathbf{v}_1^{(1)}$ depends on $\mathbf{v}_2^{(1)}$. Then, we try to solve the optimization problem by iterating between $\mathbf{v}_k^{(j)}$ and $\mathbf{v}_x^{(j)}$, and hope it will work. However, our numerical experiment shows that it will not converge for some channels.

$$\begin{aligned}
MSE_k &= \mathrm{E}[(x_k - \mathbf{g}_k^H \mathbf{y}_k)(x_k - \mathbf{g}_k^H \mathbf{y}_k)^H] \\
&= \mathrm{E}[x_k^2] - \mathrm{E}[x_k \mathbf{y}_k^H \mathbf{g}_k] - \mathrm{E}[x_k \mathbf{g}_k^H \mathbf{y}_k] + \mathrm{E}[\mathbf{g}_k^H \mathbf{y}_k \mathbf{y}_k^H \mathbf{g}_k] \\
&= 1 - \sum_i \mathbf{v}_i^{H(k)} \mathbf{H}_{ki}^H \mathbf{g}_k - \mathbf{g}_k^H \sum_i \mathbf{H}_{ki} \mathbf{v}_i^{(k)} + \mathbf{g}_k^H \sum_j \left[(\sum_i \mathbf{H}_{ki} \mathbf{v}_i^{(j)}) (\sum_i \mathbf{v}_i^{H(j)} \mathbf{H}_{ki}^H) \right] \mathbf{g}_k + \sigma^2 \mathbf{g}_k^H \mathbf{g}_k
\end{aligned}$$

$$\begin{aligned}
\mathbf{v}_1^{*(1)} &= \underset{\mathbf{v}_1^{(1)}}{\operatorname{argmin}}\Big(\sum_{k=1,2}\Big[MSE_k + \lambda_k\big(\sum_{j=1,2}\|\mathbf{v}_k^{(j)}\| - P\big)\Big]\Big) \\
&= \Big[2\mathbf{H}_{11}^H\mathbf{g}_1\mathbf{g}_1^H\mathbf{H}_{11}w_1 + 2\mathbf{H}_{21}^H\mathbf{g}_2\mathbf{g}_2\mathbf{H}_{21}w_2 + 2\lambda_1^* I\Big]^{-1}(2\mathbf{H}_{11}^H\mathbf{g}_1 w_1 - \mathbf{g}_1^H\mathbf{H}_{12}\mathbf{v}_2^{(1)}\mathbf{H}_{11}^H\mathbf{g}_1 w_1 \\
&\quad - \mathbf{H}_{11}^H\mathbf{g}_1\mathbf{v}_2^{H(1)}\mathbf{H}_{12}^H\mathbf{g}_1 w_1 - \mathbf{g}_2^H\mathbf{H}_{22}\mathbf{v}_2^{(1)}\mathbf{H}_{21}^H\mathbf{g}_2 w_2 - \mathbf{H}_{21}^H\mathbf{g}_2\mathbf{v}_2^{H(1)}\mathbf{H}_{22}^H\mathbf{g}_2 w_2) \\
&= \Big[2\mathrm{E}[x_1^*\mathbf{y}_1]\mathrm{E}[x_1^*\mathbf{y}_1]^H w_1 + 2\mathrm{E}[x_2^*\mathbf{y}_1]\mathrm{E}[x_2^*\mathbf{y}_1]^H w_2 + 2\lambda_1^* I\Big]^{-1}(2\mathrm{E}[x_1^*\mathbf{y}_1]w_1 - \mathrm{E}[x_1^*\mathbf{y}_2]^H\mathbf{v}_2^{(1)}\mathrm{E}[x_1^*\mathbf{y}_1]w_1 \\
&\quad - \mathrm{E}[x_1^*\mathbf{y}_1]\mathbf{v}_2^{H(1)}\mathrm{E}[x_1^*\mathbf{y}_2]w_1 - \mathrm{E}[x_2^*\mathbf{y}_2]^H\mathbf{v}_2^{(1)}\mathrm{E}[x_2^*\mathbf{y}_1]w_2 - \mathrm{E}[x_2^*\mathbf{y}_1]\mathbf{v}_2^{H(1)}\mathrm{E}[x_2^*\mathbf{y}_2]w_2)
\end{aligned}$$

$$\begin{aligned}
\mathbf{v}_2^{*(1)} &= \underset{\mathbf{v}_2^{(1)}}{\operatorname{argmin}}\Big(\sum_{k=1,2}\Big[MSE_k + \lambda_k\big(\sum_{j=1,2}\|\mathbf{v}_k^{(j)}\| - P\big)\Big]\Big) \\
&= \Big[2\mathbf{H}_{12}^H\mathbf{g}_1\mathbf{g}_1^H\mathbf{H}_{12}w_1 + 2\mathbf{H}_{22}^H\mathbf{g}_2\mathbf{g}_2^H\mathbf{H}_{22}w_2 + 2\lambda_2^* I)\Big]^{-1}(2\mathbf{H}_{12}^H\mathbf{g}_1 w_1 - \mathbf{g}_1^H\mathbf{H}_{11}\mathbf{v}_1^{(1)}\mathbf{H}_{12}^H\mathbf{g}_1 w_1 \\
&\quad - \mathbf{H}_{12}^H\mathbf{g}_1\mathbf{v}_1^{H(1)}\mathbf{H}_{11}^H\mathbf{g}_1 w_1 - \mathbf{g}_2^H\mathbf{H}_{21}\mathbf{v}_1^{(1)}\mathbf{H}_{22}^H\mathbf{g}_2 w_2 - \mathbf{H}_{22}^H\mathbf{g}_2\mathbf{v}_1^{H(1)}\mathbf{H}_{21}^H\mathbf{g}_2 w_2) \\
&= \Big[2\mathrm{E}[x_1^*\mathbf{y}_2]\mathrm{E}[x_1^*\mathbf{y}_2]^H w_1 + 2\mathrm{E}[x_2^*\mathbf{y}_2]\mathrm{E}[x_2^*\mathbf{y}_2]^H w_2 + 2\lambda_2^* I\Big]^{-1}(2\mathrm{E}[x_1^*\mathbf{y}_2]w_1 - \mathrm{E}[x_1^*\mathbf{y}_1]^H\mathbf{v}_1^{(1)}\mathrm{E}[x_1^*\mathbf{y}_2]w_1 \\
&\quad - \mathrm{E}[x_1^*\mathbf{y}_2]\mathbf{v}_1^{H(1)}\mathrm{E}[x_1^*\mathbf{y}_1]w_1 - \mathrm{E}[x_2^*\mathbf{y}_1]^H\mathbf{v}_1^{(1)}\mathrm{E}[x_2^*\mathbf{y}_2]w_2 - \mathrm{E}[x_2^*\mathbf{y}_2]\mathbf{v}_1^{H(1)}\mathrm{E}[x_2^*\mathbf{y}_2]w_2)
\end{aligned}$$

$$\begin{aligned}
\mathbf{v}_1^{*(2)} &= \underset{\mathbf{v}_1^{(2)}}{\operatorname{argmin}}\Big(\sum_{k=1,2}\Big[MSE_k + \lambda_k\big(\sum_{j=1,2}\|\mathbf{v}_k^{(j)}\| - P\big)\Big]\Big) \\
&= \Big[2\mathbf{H}_{11}^H\mathbf{g}_1\mathbf{g}_1^H\mathbf{H}_{11}w_1 + 2\mathbf{H}_{21}^H\mathbf{g}_2\mathbf{g}_2\mathbf{H}_{21}w_2 + 2\lambda_1^* I\Big]^{-1}(2\mathbf{H}_{21}^H\mathbf{g}_2 w_2 - \mathbf{g}_1^H\mathbf{H}_{12}\mathbf{v}_2^{(2)}\mathbf{H}_{11}^H\mathbf{g}_1 w_1 \\
&\quad - \mathbf{H}_{11}^H\mathbf{g}_1\mathbf{v}_2^{H(2)}\mathbf{H}_{12}^H\mathbf{g}_1 w_1 - \mathbf{g}_2^H\mathbf{H}_{22}\mathbf{v}_2^{(2)}\mathbf{H}_{21}^H\mathbf{g}_2 w_2 - \mathbf{H}_{21}^H\mathbf{g}_2\mathbf{v}_2^{H(2)}\mathbf{H}_{22}^H\mathbf{g}_2 w_2) \\
&= \Big[2\mathrm{E}[x_1^*\mathbf{y}_1]\mathrm{E}[x_1^*\mathbf{y}_1]^H w_1 + 2\mathrm{E}[x_2^*\mathbf{y}_1]\mathrm{E}[x_2^*\mathbf{y}_1]^H w_2 + 2\lambda_1^* I\Big]^{-1}(2\mathrm{E}[x_2^*\mathbf{y}_1]w_2 - \mathrm{E}[x_1^*\mathbf{y}_2]^H\mathbf{v}_2^{(2)}\mathrm{E}[x_1^*\mathbf{y}_1]w_1 \\
&\quad - \mathrm{E}[x_1^*\mathbf{y}_1]\mathbf{v}_2^{H(2)}\mathrm{E}[x_1^*\mathbf{y}_2]w_1 - \mathrm{E}[x_2^*\mathbf{y}_2]^H\mathbf{v}_2^{(2)}\mathrm{E}[x_2^*\mathbf{y}_1]w_2 - \mathrm{E}[x_2^*\mathbf{y}_1]\mathbf{v}_2^{H(2)}\mathrm{E}[x_2^*\mathbf{y}_2]w_2)
\end{aligned}$$

$$
\begin{aligned}
\mathbf{v}_2^{*(2)} &= \underset{\mathbf{v}_2^{(2)}}{\text{argmin}}(\sum_{k=1,2}\left[MSE_k + \lambda_k(\sum_{j=1,2}\|\mathbf{v}_k^{(j)}\| - P)\right]) \\
&= \left[2\mathbf{H}_{12}^H\mathbf{g}_1\mathbf{g}_1^H\mathbf{H}_{12}w_1 + 2\mathbf{H}_{22}^H\mathbf{g}_2\mathbf{g}_2^H\mathbf{H}_{22}w_2 + 2\lambda_2^* I)\right]^{-1}(2\mathbf{H}_{22}^H\mathbf{g}_2 w_2 - \mathbf{g}_1^H\mathbf{H}_{11}\mathbf{v}_1^{(2)}\mathbf{H}_{12}^H\mathbf{g}_1 w_1 \\
&\quad - \mathbf{H}_{12}^H\mathbf{g}_1\mathbf{v}_1^{H(2)}\mathbf{H}_{11}^H\mathbf{g}_1 w_1 - \mathbf{g}_2^H\mathbf{H}_{21}\mathbf{v}_1^{(2)}\mathbf{H}_{22}^H\mathbf{g}_2 w_2 - \mathbf{H}_{22}^H\mathbf{g}_2\mathbf{v}_1^{H(2)}\mathbf{H}_{21}^H\mathbf{g}_2 w_2) \\
&= \left[2\mathrm{E}[x_1^*\mathbf{y}_2]\mathrm{E}[x_1^*\mathbf{y}_2]^H w_1 + 2\mathrm{E}[x_2^*\mathbf{y}_2]\mathrm{E}[x_2^*\mathbf{y}_2]^H w_2 + 2\lambda_2^* I\right]^{-1}(2\mathrm{E}[x_2^*\mathbf{y}_2]w_2 - \mathrm{E}[x_1^*\mathbf{y}_1]^H\mathbf{v}_1^{(2)}\mathrm{E}[x_1^*\mathbf{y}_2]w_1 \\
&\quad - \mathrm{E}[x_1^*\mathbf{y}_2]^H\mathbf{v}_1^{H(2)}\mathrm{E}[x_1^*\mathbf{y}_1]w_1 - \mathrm{E}[x_2^*\mathbf{y}_1]^H\mathbf{v}_1^{(2)}\mathrm{E}[x_2^*\mathbf{y}_2]w_2 - \mathrm{E}[x_2^*\mathbf{y}_2]^H\mathbf{v}_1^{H(2)}\mathrm{E}[x_2^*\mathbf{y}_1]w_2)
\end{aligned}
$$

2.6 Backward Direction: Duality Method

For the duality method, if the optimization problem is convex, then the optimal solution is guaranteed. If it is not, there will exist a **duality gap** which will not give the optimal solution[Convex Optimization, Boyd].

2.6.1 Primal Problem

$$
\min_{\mathbf{v}_k^{(j)}} L = \sum_k \left[w_k MSE_k + \lambda_k(\sum_j \|\mathbf{v}_k^{(j)}\| - P)\right]
$$

We get

$$
\mathbf{v}^{*H} = \mathbf{k}[\mathbf{A} + \boldsymbol{\lambda}]^{-1}
$$

where $\mathbf{k}$ is a constant row vector with channel information, $\mathbf{A}$ is a constant matrix with channel information, and $\boldsymbol{\lambda}$ is a diagonal matrix with each $\boldsymbol{\lambda}_{ii}$ to be some λ_k

2.6.2 Dual Problem

$$
\max_{\boldsymbol{\lambda}} G(\mathbf{v}^*, \boldsymbol{\lambda}) \ = \max_{\boldsymbol{\lambda}} \ (w1 + w2 + w3) - \mathbf{v}^{*H}\mathbf{k}^H - P(\lambda_1 + \lambda_2 + \lambda_3)
$$

We could actually find the **analytic solution** for $\boldsymbol{\lambda}$. However, for 3 users case, the process includes finding the inverse of an 6x6 symbolic matrix which is **solvable** but **complicated**, let alone cases above 3 users. Thus, we choose to solve the dual problem with **numerical method**. By the optimization theory, dual problems are **always** convex. Therefore, we could get the optimal $\boldsymbol{\lambda}$ simply with gradient algorithm. The gradient is shown below.

$$\frac{\partial G(\mathbf{v}^*, \boldsymbol{\lambda})}{\partial \lambda_k} = -\operatorname{Tr}[-P\|\mathbf{k}\|^2 (A + \lambda_k \mathbf{I})^{-T} (A + \lambda_k \mathbf{I})^{-T}]$$

The partial derivatives of each λ_k are mutually independent, and have the same formula. This implies that it's easy to generalize the result for more users.

3 System Model 3 : Augmented Receivers

In this model, we make each receivers(mobile station) possess the ability to assist channel estimation in backward direction. Since it is rather easy to add a filter to receivers, if this method does improve the capacity, it will be extremely fascinating. We solve each transceivers with Wiener-Hopf equation: $R^{-1}p$.

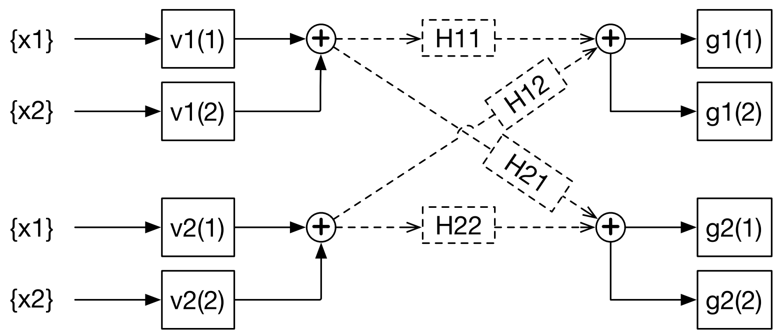


Figure 5: Forward Channel

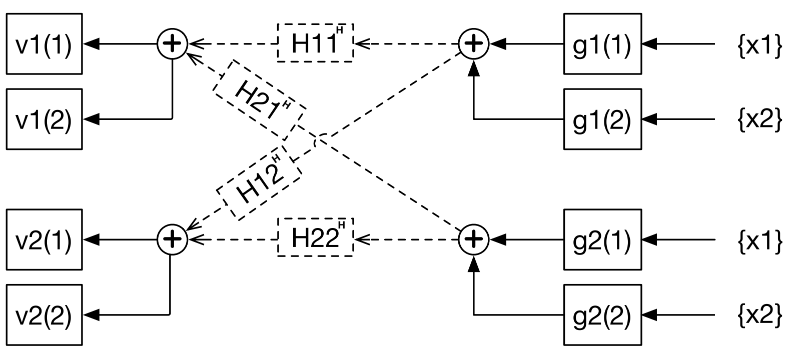


Figure 6: Backward Channel

3.1 Optimization Problem

$$\min_{\mathbf{v}_k^{(j)}, \mathbf{g}_k} \sum_k w_k MSE_k$$

$$\text{subject to } \sum_j \|\mathbf{v}_k^{(j)}\|^2 = P; \|\mathbf{g}_k\|^2 = P$$

3.2 Numerical Results

This is a 3 users, and 2X2 MIMO interference channel, with SNR equals to 20dB. The results are averaged over 1000 complex gaussian channels. X-axis represents the number of iterations: start from backward direction and end in forward direction. Y-axis represents the sum capacity of three users. The first thing we can tell is the **augmented one** is a lot worse than the **simple one**(Only one filter at each transceivers. For example, for 2 users case, only have $v_1^{(1)}$,$v_2^{(2)}$,$g_1^{(1)}$,$g_2^{(2)}$). **This is because transmitters distribute all messages to all users.** In the experiment, transmitter 1 sends 2 bits of message 1 to user 1, 2, and 3. However user 2 and 3 simply drop message 1 off(they don't need it). Thus, the sum capacity of augmented receivers only have 1/3 of capacity compared with the simple one.

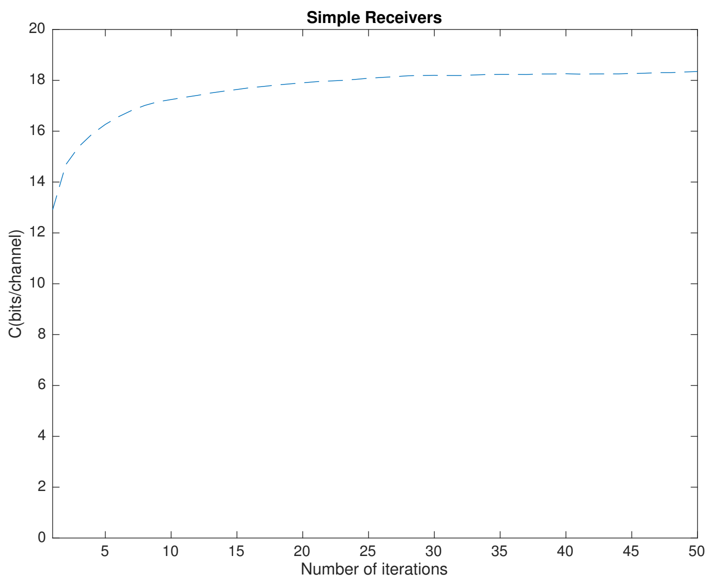


Figure 7: Simple Receivers

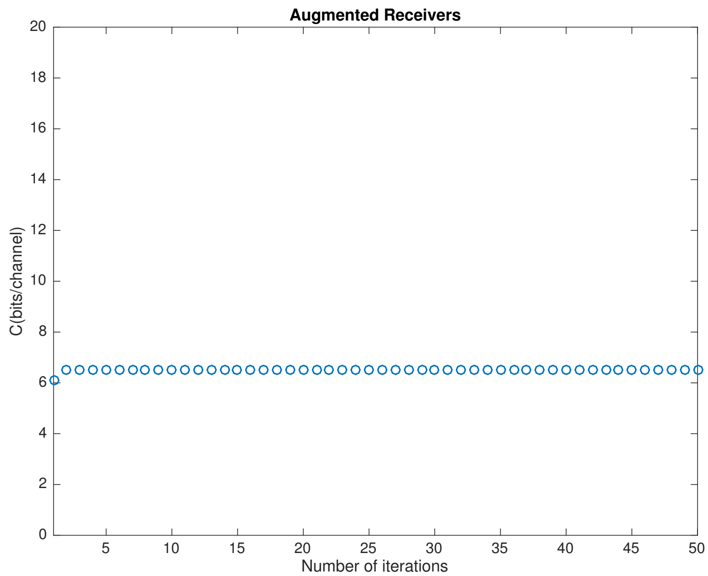


Figure 8: Augmented Receivers

9. Numerical Simulation(2 Users, 2X2 MIMO Channel)

$$Rayleigh\ Fading\ Channel$$

$$Cross\ Channel\ Gain = 0.8 * Direct\ Channel\ Gain$$

$$SNR = \frac{1}{\sigma^2} = 10^3 = 30dB$$

$$\begin{aligned}Observation1: &If\ the\ training\ length\ is\ long\ enough,\ each\ LMS\ filter\ (\mathbf{v}_k^{(c)}, \mathbf{v}_k^{(p)}, \mathbf{g}_k^{(c)}, \mathbf{g}_k^{(p)},)\\ &will\ converge\ to\ Wiener\ filter\end{aligned}$$

$$Observation2:\ Use\ Wiener\ filters,\ and\ only\ send\ common\ messages.\ Sum\ rate\ C\ =\ 11.6\ bit/channel$$

$$Observation3: \ Use\ Wiener\ filters,\ and\ only\ send\ private\ messages.\ Sum\ rate\ C\ =\ 2.63\ bit/channel$$

$$Observation4: \ Use\ Wiener\ filters,\ and\ send\ both\ messages.\ Sum\ rate\ C\ =\ 3.35\ bit/channel$$

$$Observation5: \ Under\ the\ cooperation\ scheme,\ transmitters\ don't\ converge\ to\ Wiener\ filters$$

$$Observation5: \ Under\ the\ cooperation\ scheme,\ C = 3.15\ bit/channel$$

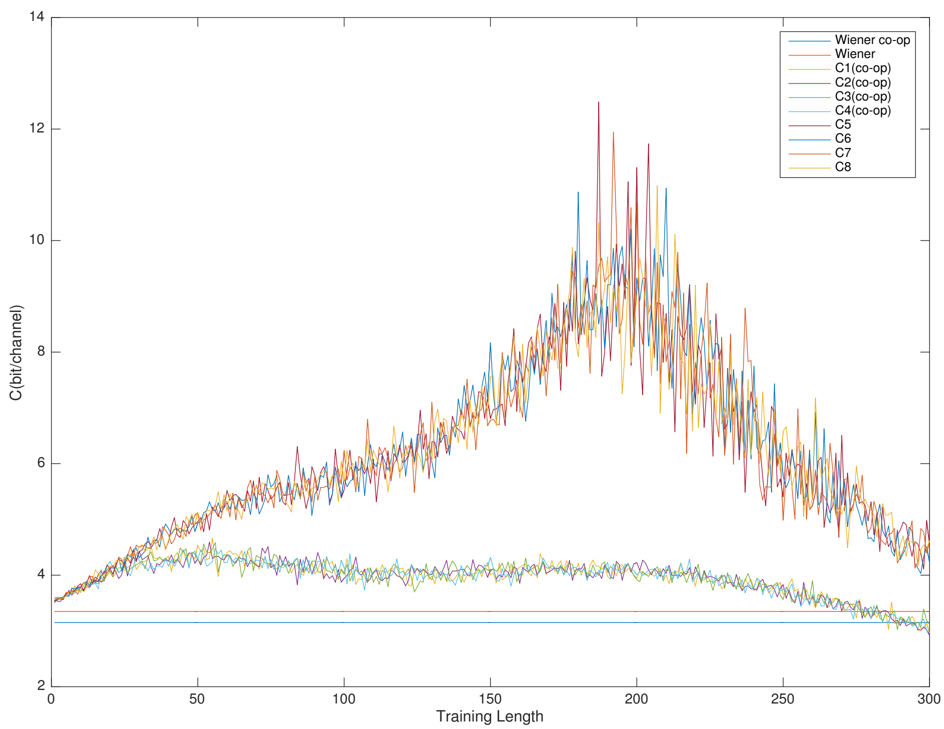


Figure 9: Insert caption

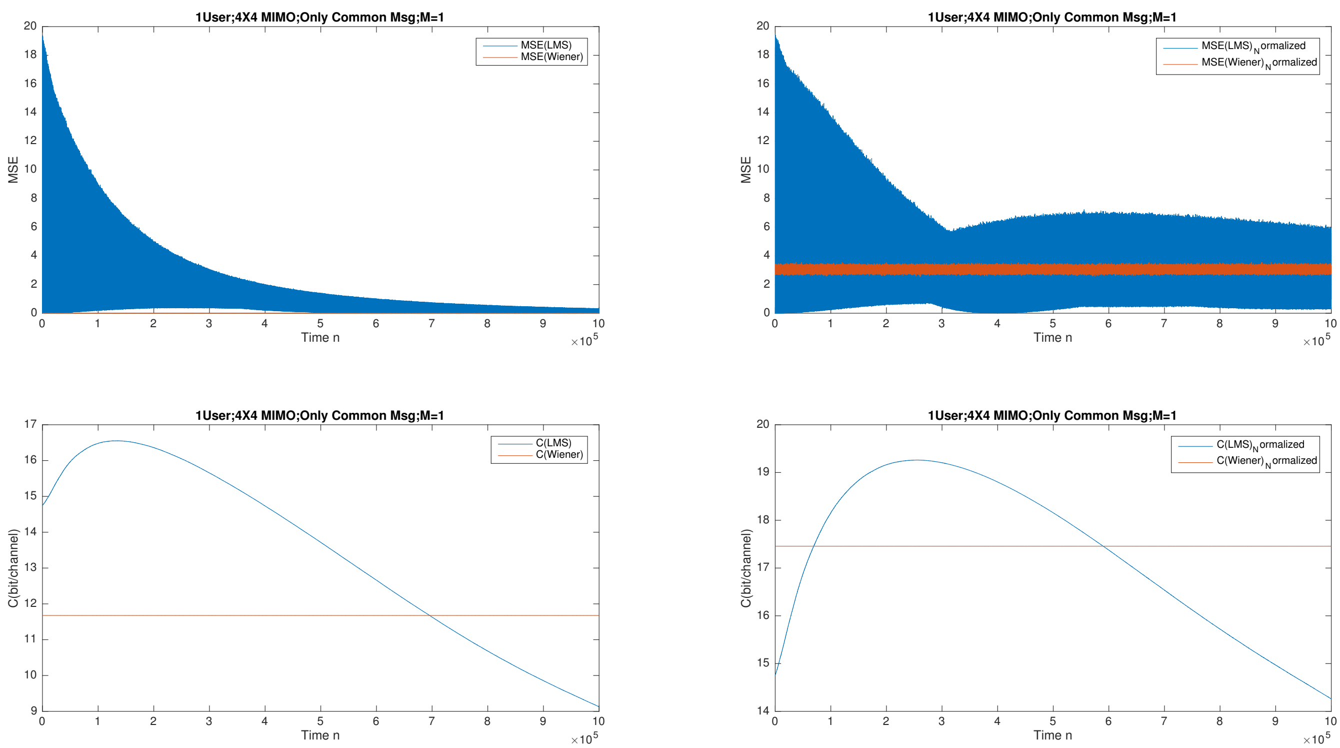


Figure 10: Insert caption

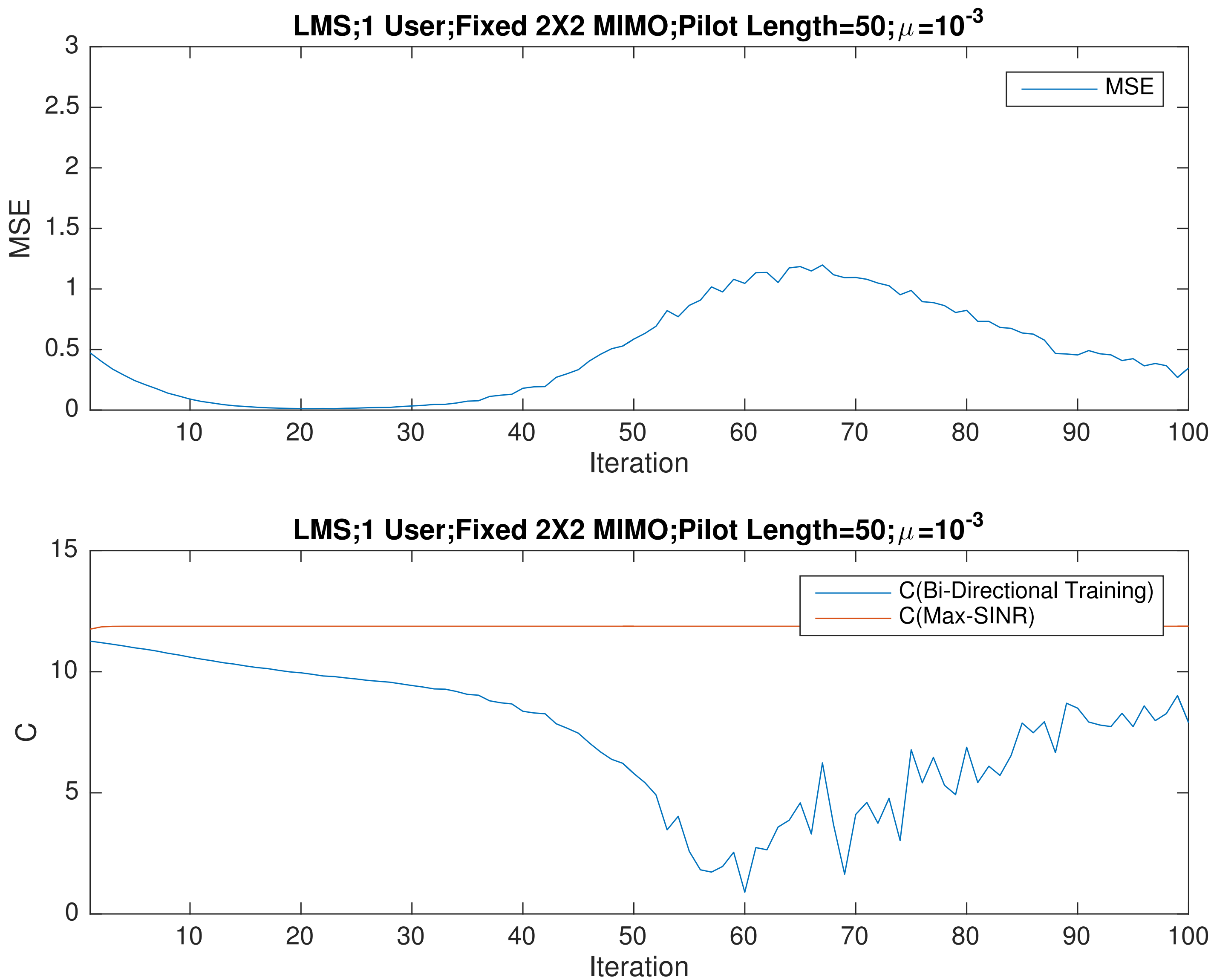


Figure 11: Insert caption

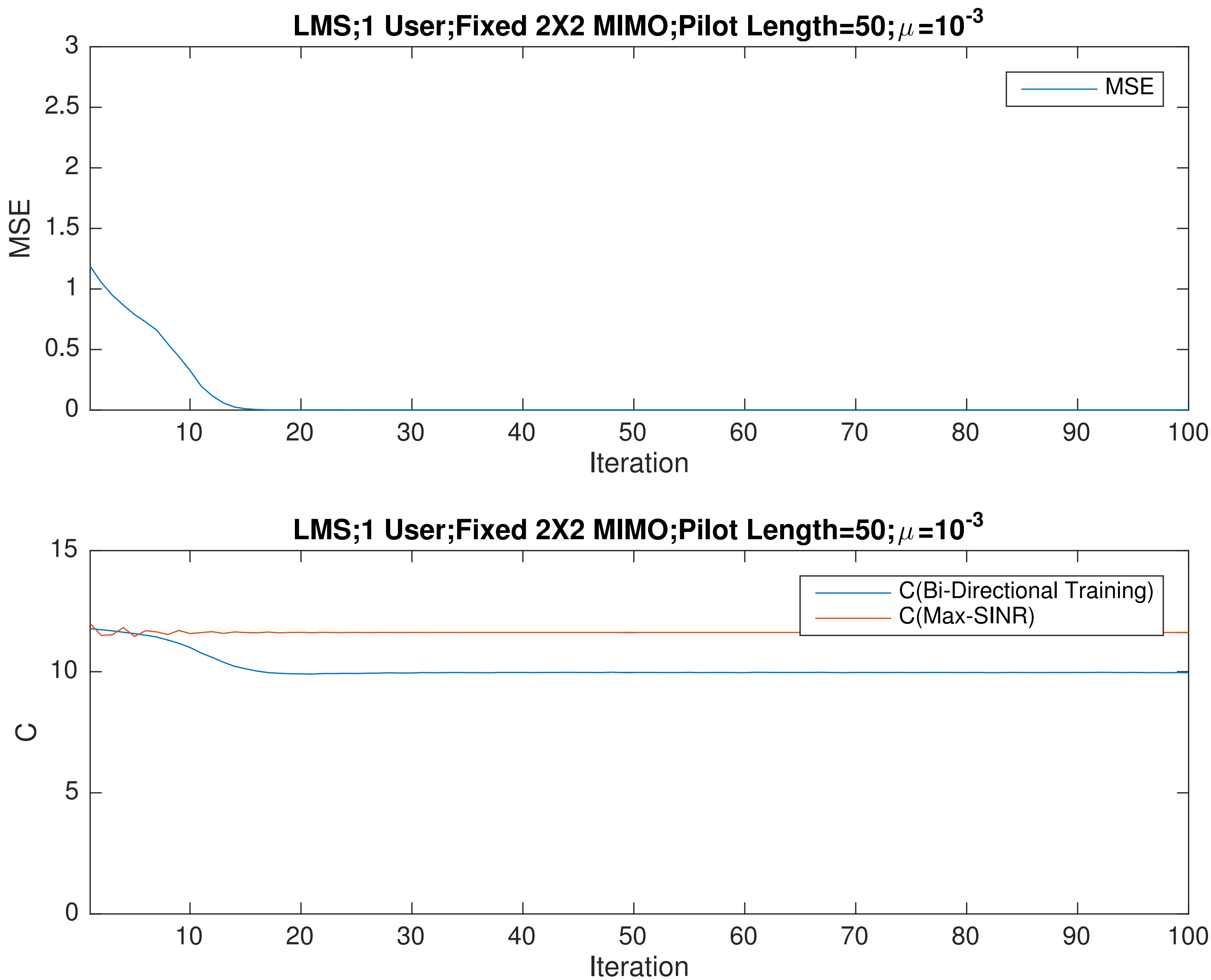
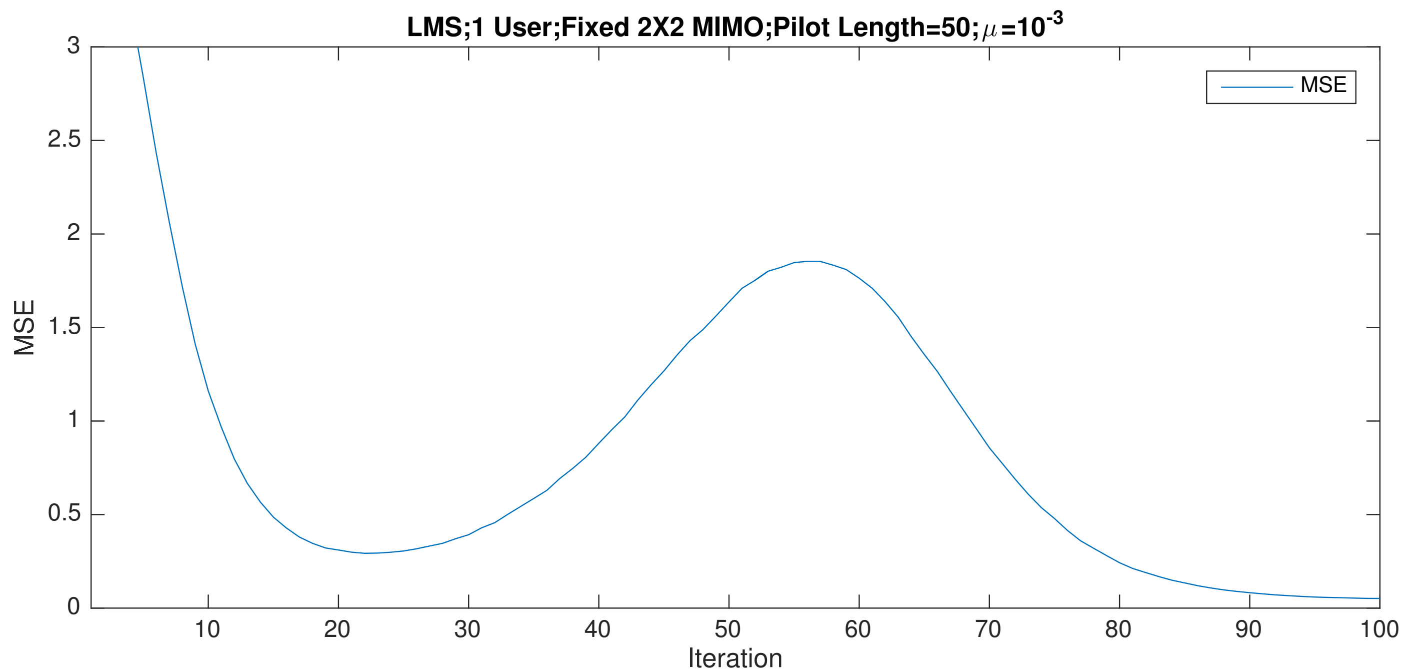


Figure 12: Insert caption



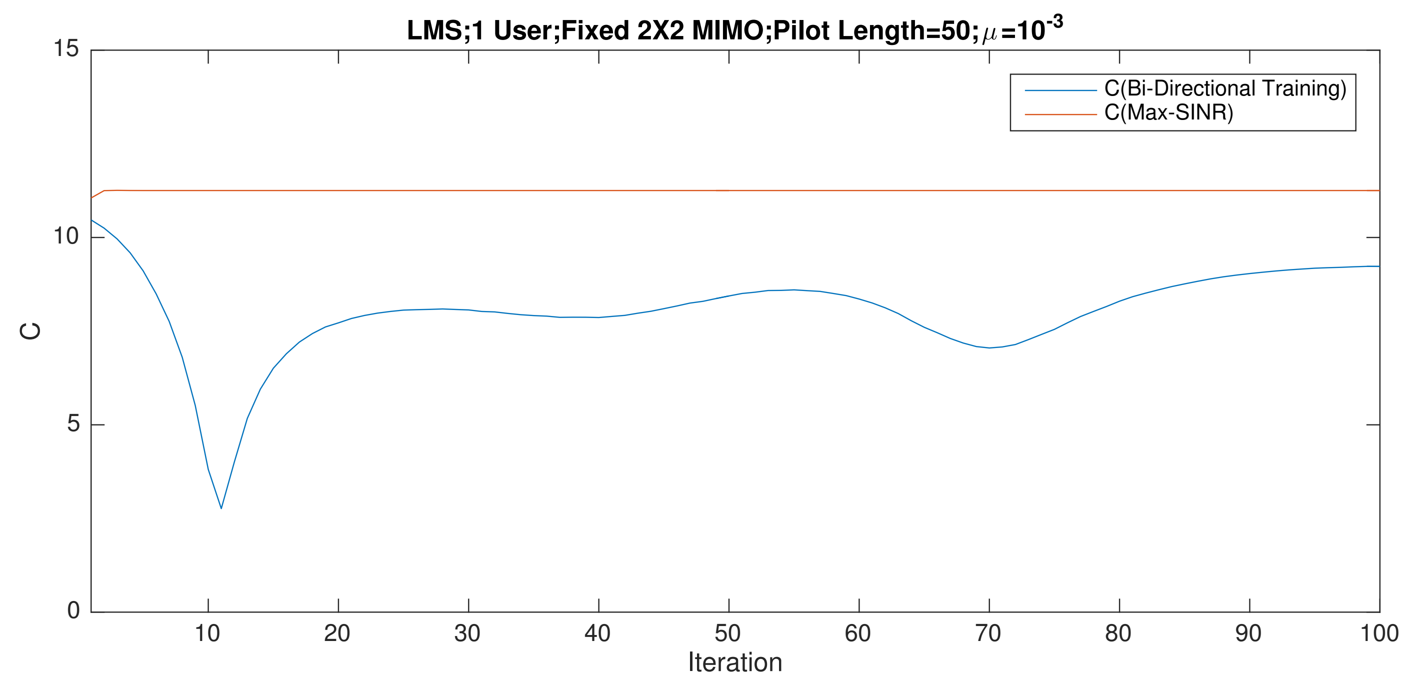
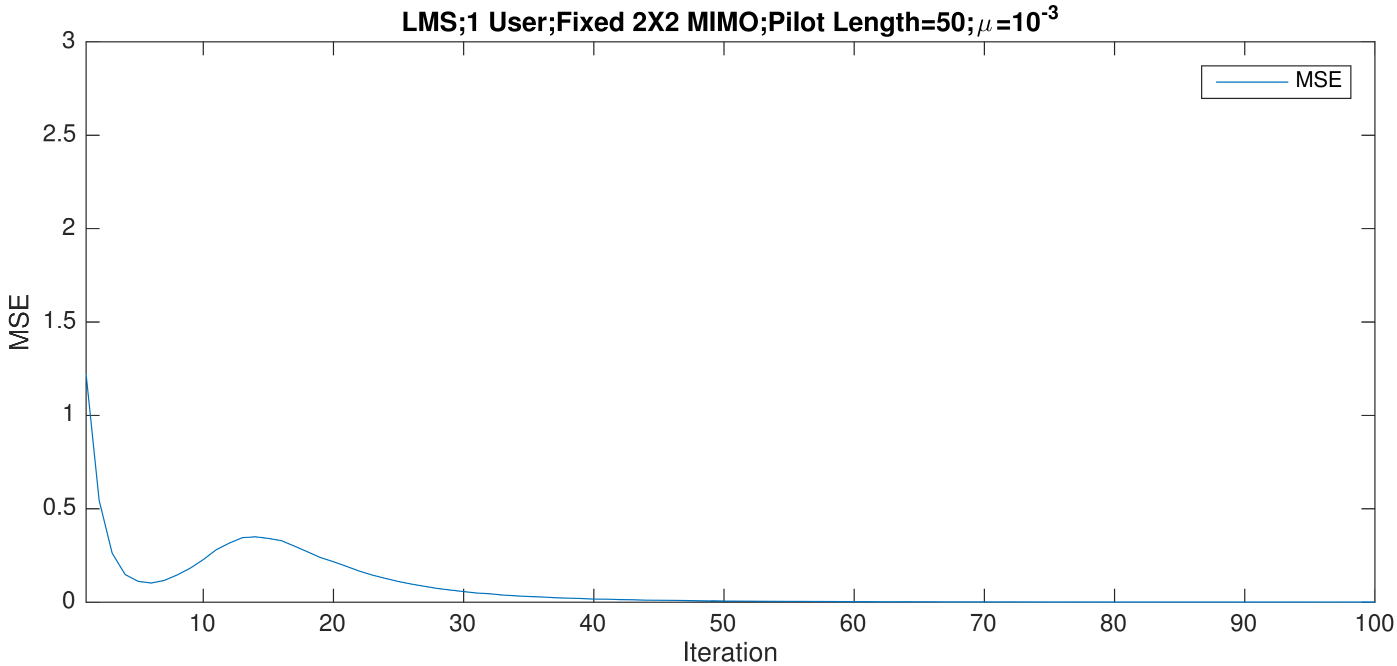


Figure 13: Insert caption



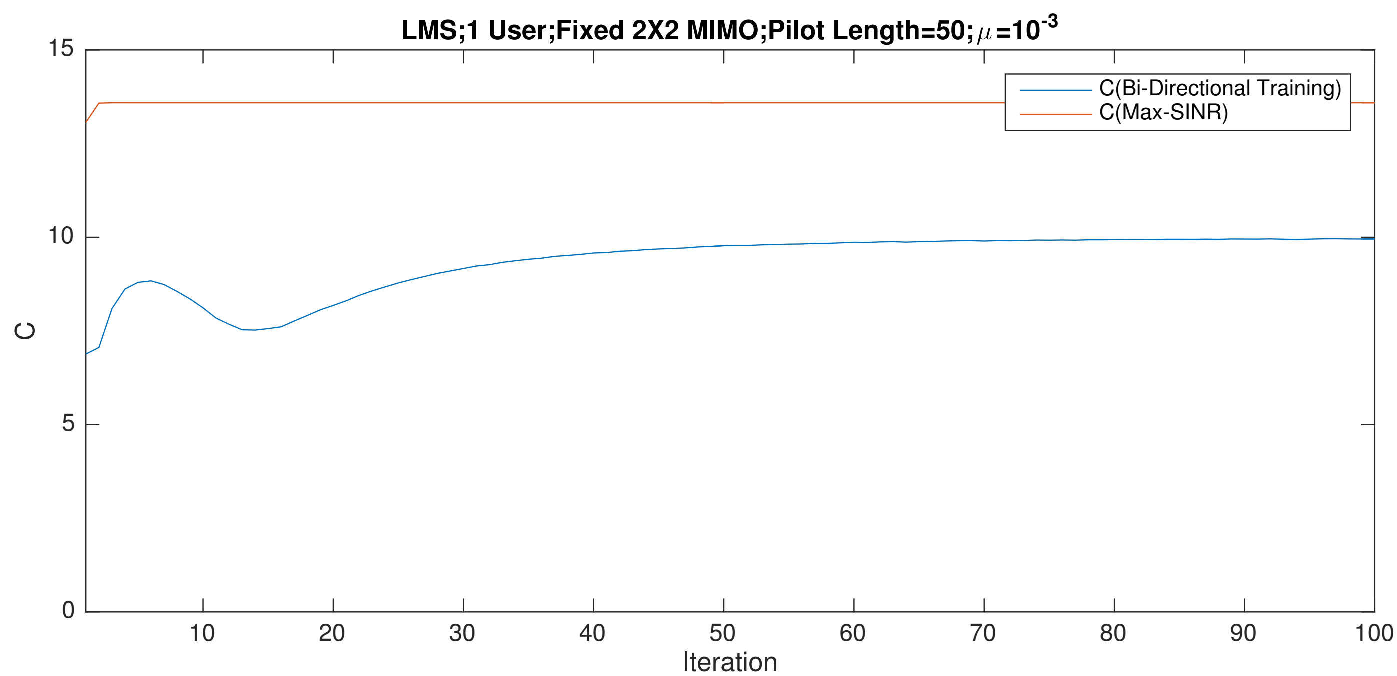


Figure 14: Insert caption

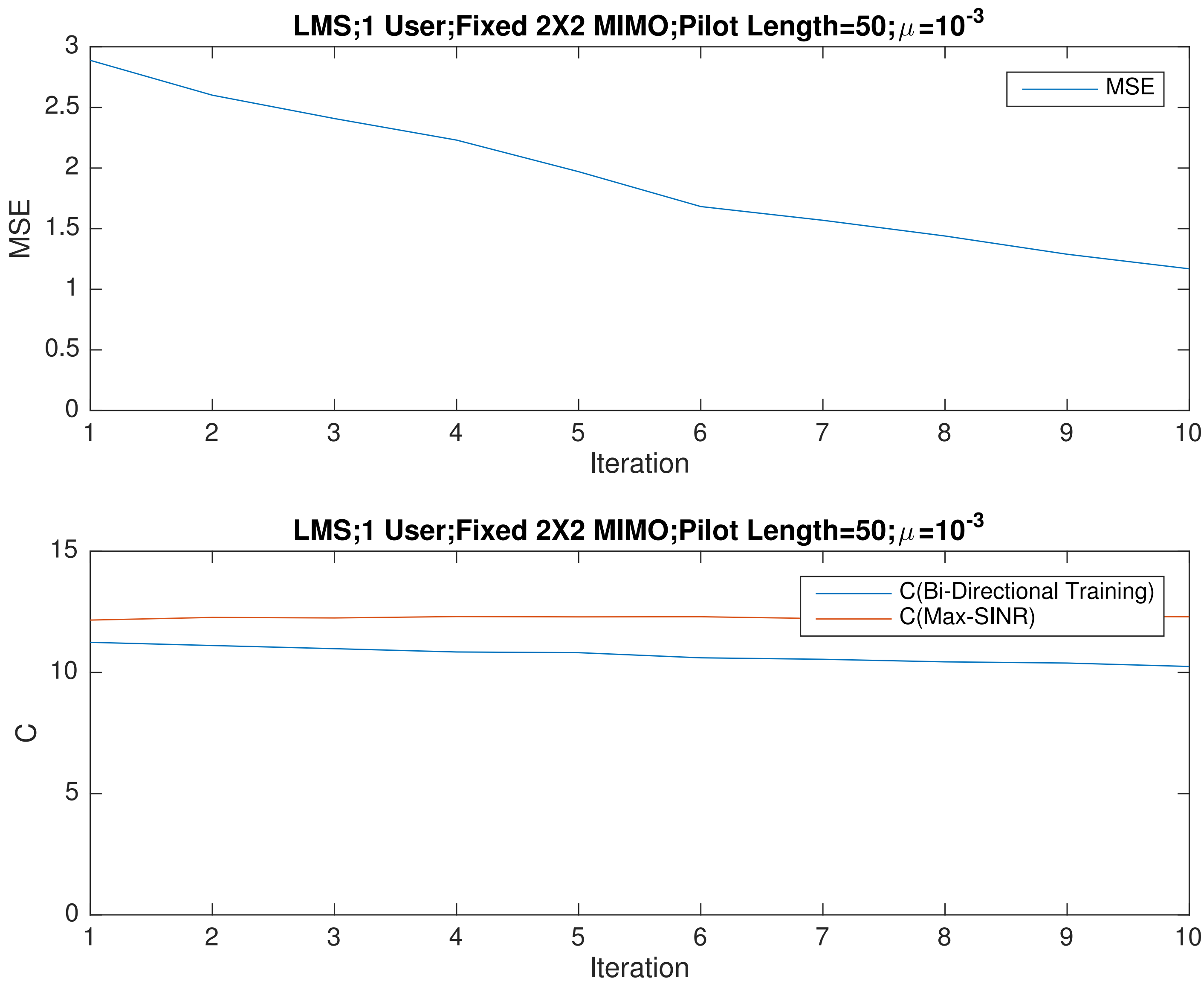


Figure 15: Insert caption

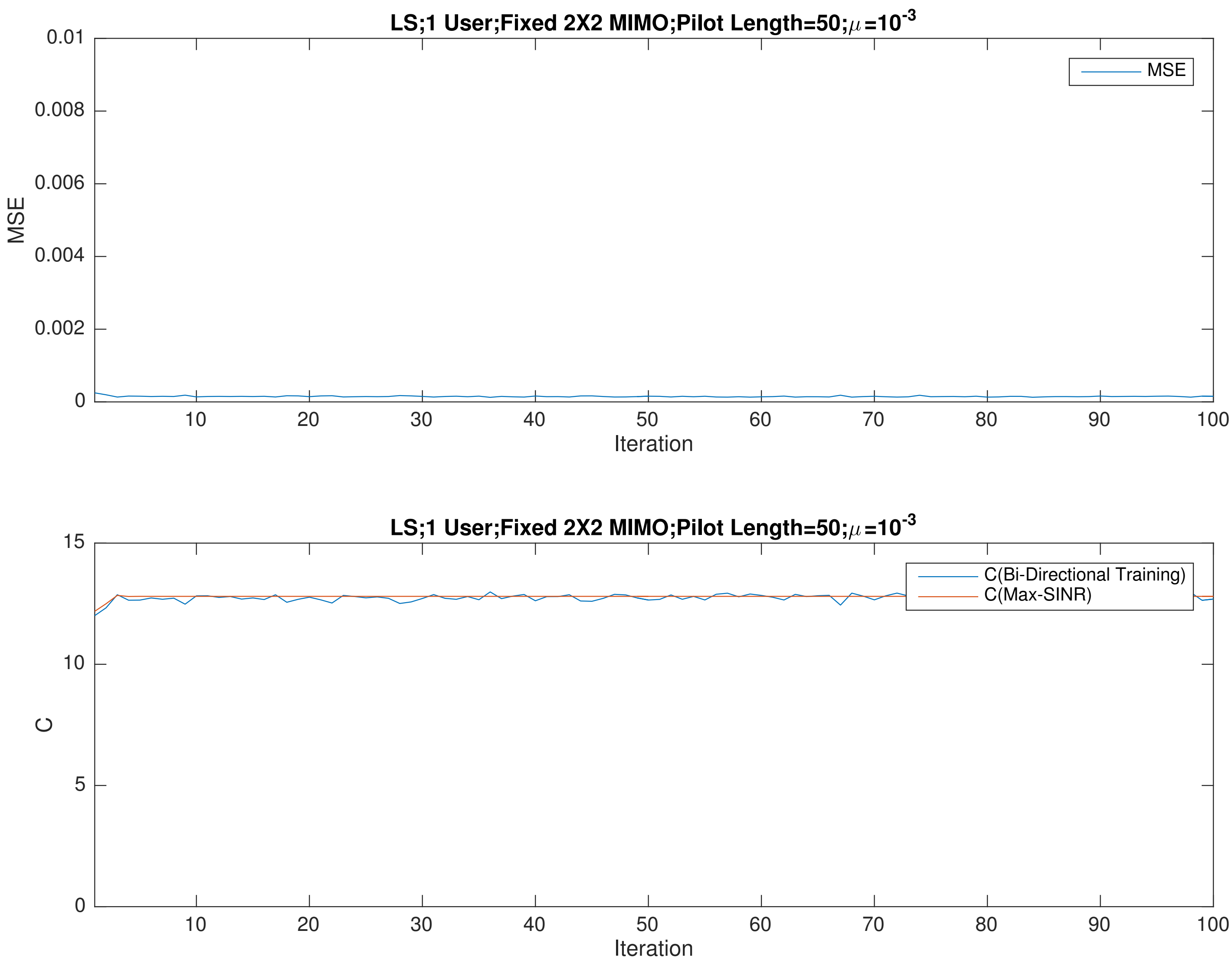


Figure 16: Insert caption

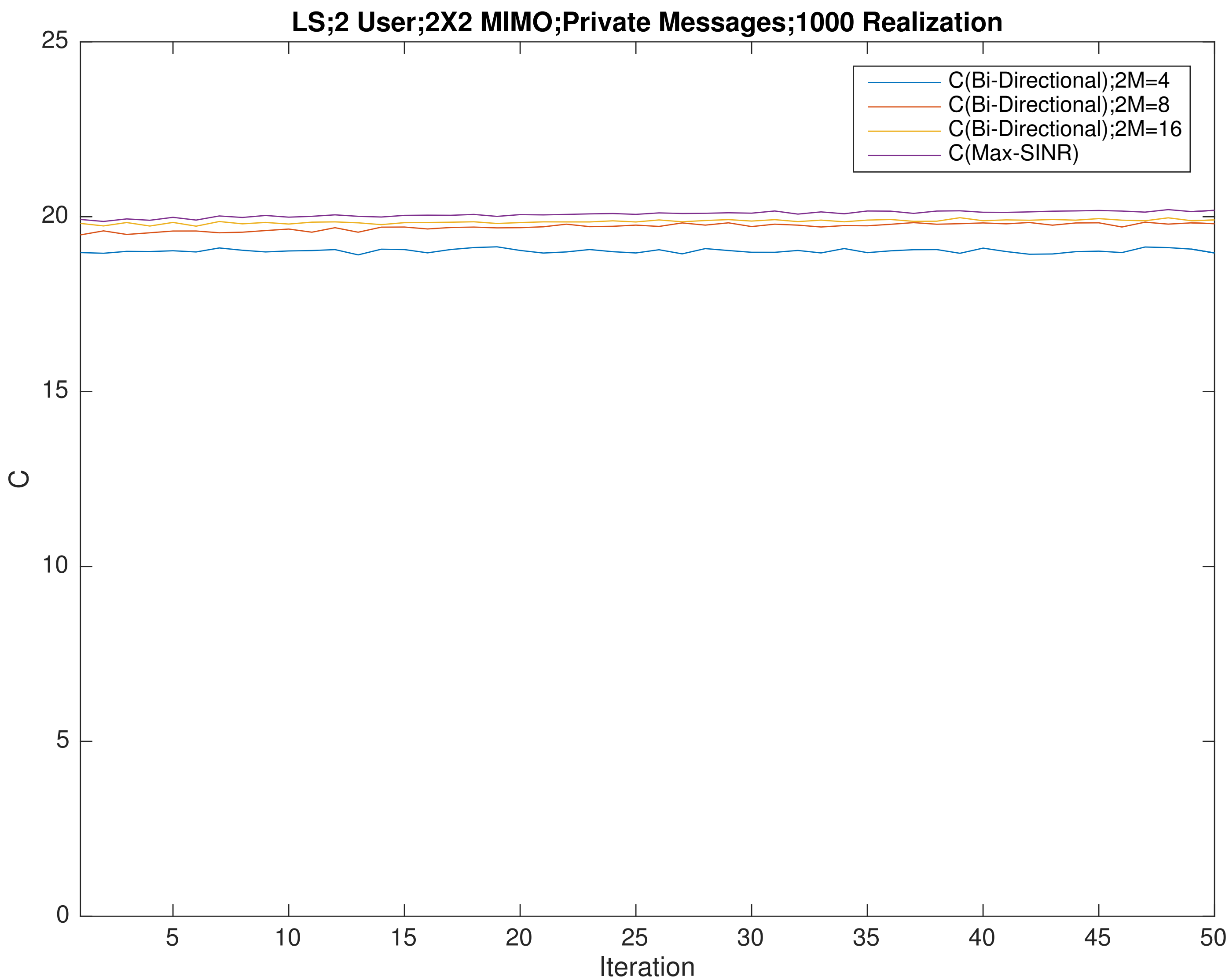


Figure 17: Insert caption

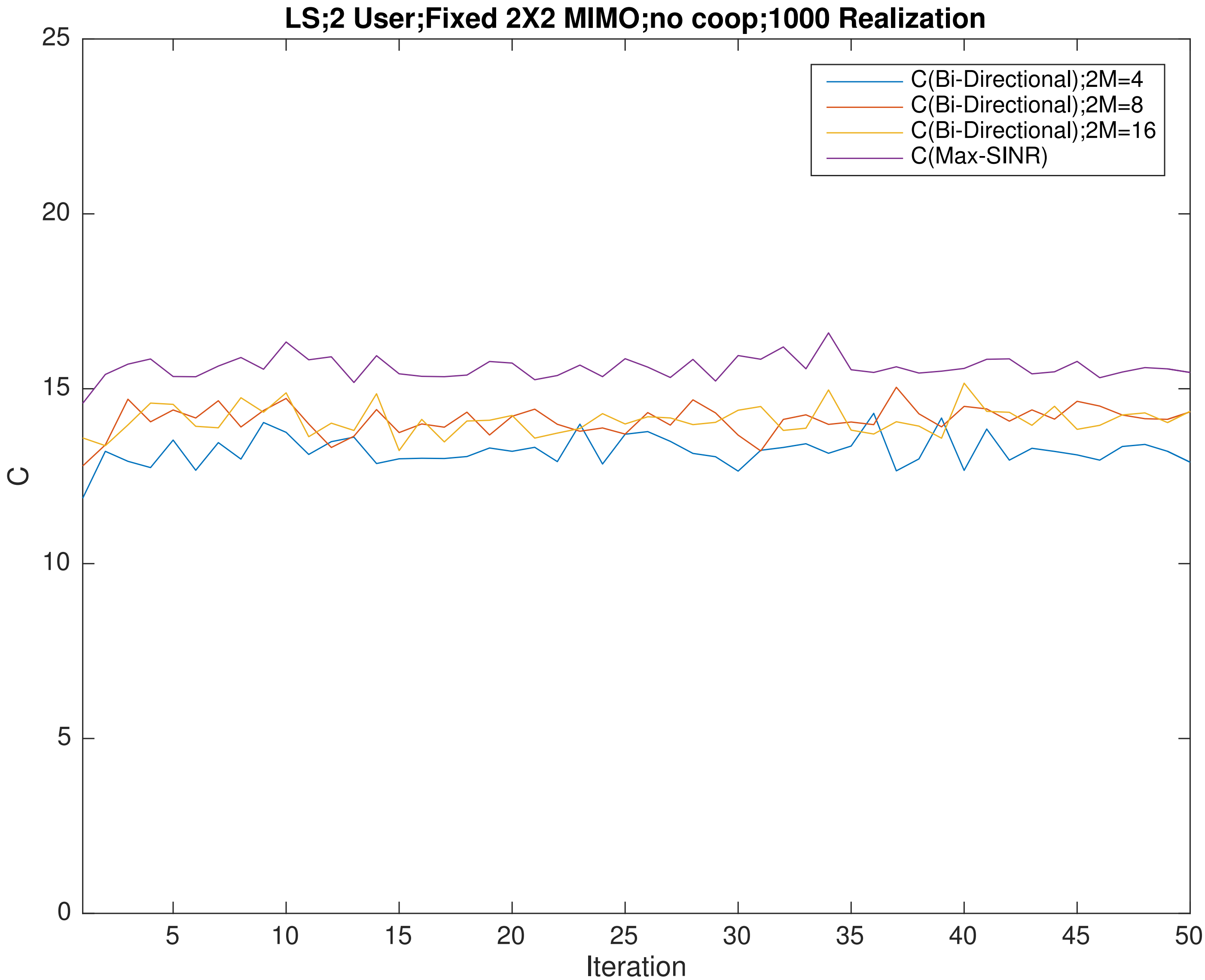


Figure 18: Insert caption

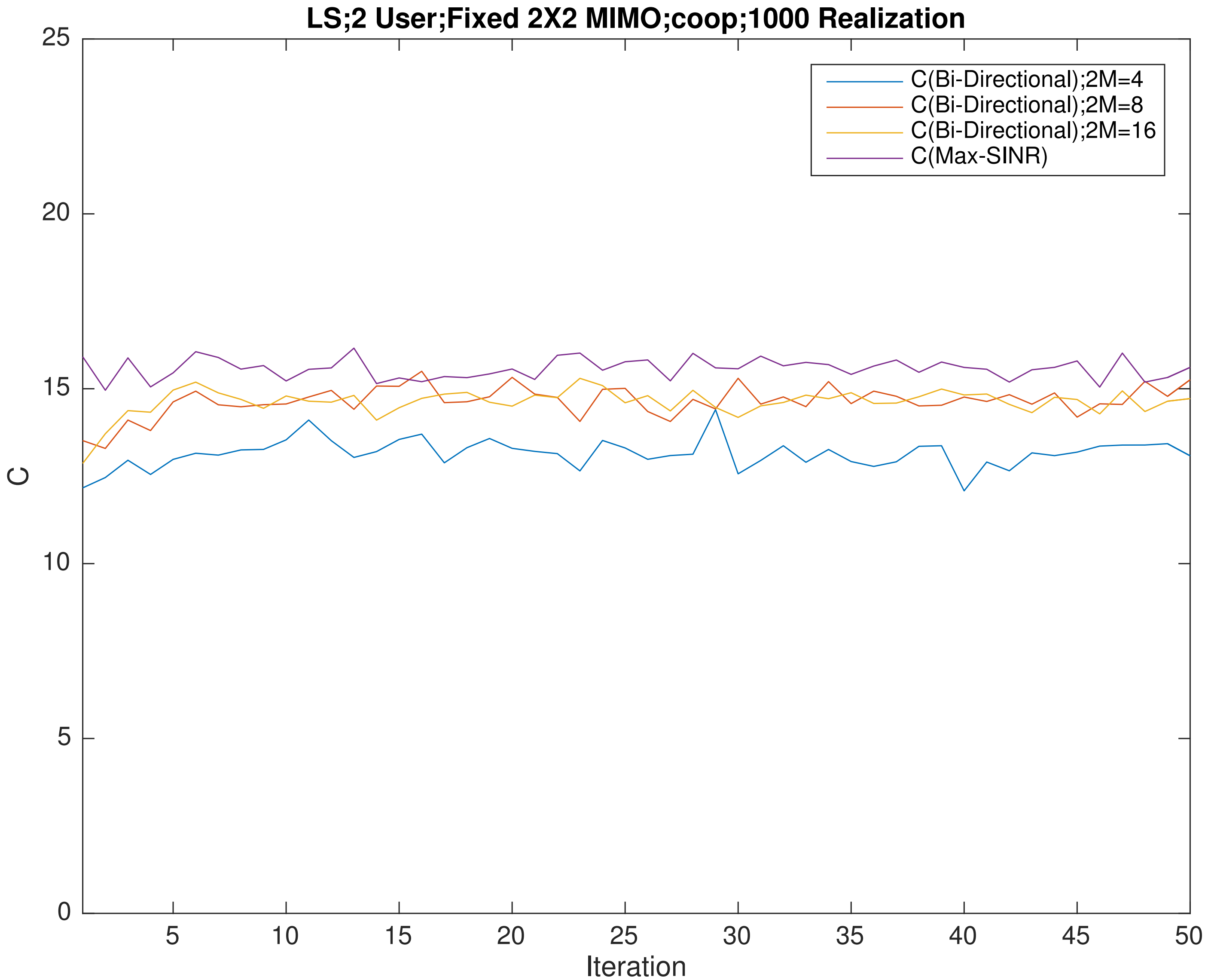


Figure 19: Insert caption