Distributed Precoding Systems in Multi-Gateway Multibeam Satellites

Carlos Mosquera [*] , Roberto López-Valcarce [†] and Tomás Ramirez [‡]
University of Vigo, Vigo, Galicia, 36310, Spain

This paper deals with the coexistence of different groups of terminals -clusters-, each served by an Earth station or gateway through a multibeam satellite. A high co-channel interference results from the reuse of the spectral resources across all beams and clusters. In this context, each gateway precodes the symbols addressed to its respective users. The design follows an MMSE criterion, and channel statistics are seen to satisfy the required knowledge about the inter-cluster interference, thus avoiding the exchange of instantaneous channel state information among gateways through backhaul links.

I. Introduction

The object of this study is a multibeam satellite which relays the signals coming from different ground stations (gateways) to convey their communication with single-antenna terminals. Bent-pipe communication satellites can be considered as non-regenerative relays,[1] essentially filtering and amplifying signals, although they are very complex communication systems and handle simultaneously many streams of information. The foot-print of a multibeam satellite is made of many spot-beams, hundreds in some specific commercial cases, which are synthesized by the on-board beamforming network (BFN) in combination with the antennas radiation pattern. Two implementation approaches are possible:[1] single feed per beam and multiple feeds per beam. For the purpose of this paper, it is of specific interest the case of multiple feeds per beam, for which small subarrays are used for each spot, and adjacent spots share some of the array elements.

Figure 1 depicts the conceptual abstraction of the multibeam satellite operation, with the following features to be highlighted: (i) the feeder links, from the gateways to the satellites, can be assumed transparent, whereas the user links are frequency non-selective; (ii) there is no interference between feeder links and user links, since the communication takes place on different frequency bands; (iii) a given cluster is made of several beams, with one user per beam served at a time by a given frequency carrier; (iv) the user link frequency carriers are made available to all beams and clusters, in what it is known as full-frequency reuse.

One major challenge for multibeam satellites is the large spectral demand on the feeder link between the satellite and the operator stations, since it has to aggregate the traffic from all beams. The use of different gateways can generate several parallel channels provided that the antennas guarantee the required spatial isolation, which is the case for frequencies in Ku-band or Ka-band. Thus, the different feeder links can reuse the whole available bandwidth.

The preprocessing of signals to communicate multiple-antennas in one end -the satellite in our case- with many users simultaneously is supported by theoretical bounds and practical schemes presented in many references. Precoding for multibeam satellites has been extensively explored in the literature to fight interbeam interference in the case of a single gateway, see, e.g.,[2] and[3] among others. As opposed, results for multiple gateways are still incomplete; a centralized multi-gateway resource management, which for mathematical purposes can be assumed, is far from being realistic in practice,[4] due to the backhaul links that would be needed to connect all the gateways. Some precoding schemes for multiple gateways without BFN are

[*]atlanTTic Research Center, mosquera@gts.uvigo.es, and Not AIAA Member
[†]atlanTTic Research Center, valcarce@gts.uvigo.es, and Not AIAA Member
[‡]atlanTTic Research Center, tramirez@gts.uvigo.es, and Not AIAA Member

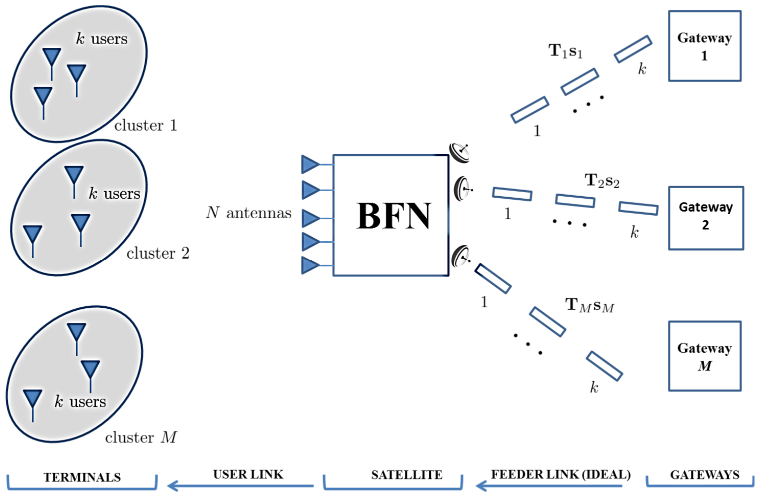

Figure 1: Satellite shared by a number of ground stations.

presented in,[5] which assume the exchange of information for the design of their respective precoders. In general, the mapping of groups of beams to different gateways prevents from a centralized management, with inter-cluster interference more difficult to control.

In this work we try to keep the inter-gateway cooperation at a minimum, so that no information symbols are exchanged among the terrestrial gateways, each communicating with the terminals operating on its cluster. Initially each gateway is expected to know the channel state information (CSI) of the links originating from itself, including inter-cluster links, although the use of channel statistics is shown to predict quite conveniently the required information so the interaction among different clusters can be completely avoided.

After detailing the satellite relaying operation in Section II, we derive the inter-cluster aware procoders in Section III, with performance tested in the simulations shown before the conclusions.

Notation: Upper (lower) boldface letters denote matrices (vectors). $(.)^H$, $(.)^T$, $\mathrm{tr}\{\cdot\}$, $\mathbf{I}_N$, $\mathrm{diag}\{\cdot\}$ denote Hermitian transpose, transpose, matrix trace operator, $N \times N$ identity matrix, and diagonal matrix, respectively. $\mathbb{E}[\cdot]$ is the expected value operator.

II. Satellite Relaying Operation

At each time instant the satellite relays the information symbols to K users. Each user makes use of the whole available bandwidth, thus giving rise to both intra-cluster and inter-cluster interference. The satellite has N radiation elements or feeds, with $N \geq K$. As shown in Figure 1, the number of transmit ground stations is M, each sending k signal streams simultaneously (in different frequency slots, for example) to the satellite, which makes use of n antennas (feeds) to send those symbols to the k users in the mth cluster, with $k \leq n \leq N$. The groups of n feeds are not necessarily disjoint. The information transmitted from each ground station is written as $\mathbf{x}_m = \mathbf{T}_m \mathbf{s}_m$, with $\mathbf{T}_m \in \mathbb{C}^{k \times k}, m = 1, \ldots, M$, a set of distributed precoding matrices. The input energy is normalized as $\mathbb{E}\left[\mathbf{s}_m \mathbf{s}_m^H\right] = \mathbf{I}_k$. The goal of the precoder at each transmitter is mainly the mitigation of the intra-cluster interference, while trying to reduce the negative impact of its interference on other clusters. The BFN is made of the tall submatrices $\mathbf{B}_m \in \mathbb{C}^{n \times k}, m = 1, \ldots, M$, with $k \leq n \leq N$.

If we denote by $\mathbf{y}_m \in \mathbb{C}^{n \times k}, m = k, \ldots, 1$ the vector of received values by terminals in cluster m, and $\mathbf{H}_{mp} \in \mathbb{C}^{k \times n}$ is the channel between the n feeds operated by the pth gateway and the mth cluster, then the signal model reads as

$$\begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_M \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11} & \dots & \mathbf{H}_{1M} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{M1} & \dots & \mathbf{H}_{MM} \end{pmatrix} \begin{pmatrix} \mathbf{B}_1\mathbf{T}_1 s_1 \\ \vdots \\ \mathbf{B}_M\mathbf{T}_M\mathbf{s}_M \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_M \end{pmatrix}, \tag{1}$$

with noise vectors $\mathbf{n}_m \in \mathbb{C}^{k\times 1}, m = 1, \dots, M$ zero-mean unit variance Additive White Gaussian Noise (AWGN) such that $\mathbb{E}\left[\mathbf{n}_m\mathbf{n}_m^H\right] = \mathbf{I}_k$. The vector of samples received by users in cluster m is decomposed as

$$\mathbf{y}_m = \mathbf{H}_{mm}\mathbf{B}_m\mathbf{T}_m\mathbf{s}_m + \sum_{p\neq m} \mathbf{H}_{mp}\mathbf{B}_p\mathbf{T}_p\mathbf{s}_p + \mathbf{n}_m. \tag{2}$$

End users cannot cooperate, and we consider the particular case in which the scaling is the same for all users across the cluster, with $\hat{\mathbf{s}}_m = \frac{1}{\sqrt{t_m}}\mathbf{y}_m$.

As performance metric we use the aggregated MSE, or Sum MSE (SMSE), given by

$$\mathrm{SMSE} = \sum_{m=1}^{M} \mathrm{tr}\{\mathbf{E}_m\}, \tag{3}$$

with

$$\mathbf{E}_m \triangleq \mathbb{E}\left[(\mathbf{s}_m - \hat{\mathbf{s}}_m)(\mathbf{s}_m - \hat{\mathbf{s}}_m)^H\right]. \tag{4}$$

The expectation is computed with respect to the symbols and the noise for a fixed channel, and reads as

$$\mathrm{SMSE} = \sum_{m=1}^{M} \mathrm{tr}\left\{\mathbf{I}_k - \frac{1}{\sqrt{t_m}}\mathbf{H}_{mm}\mathbf{B}_m\mathbf{T}_m - \frac{1}{\sqrt{t_m}}\mathbf{T}_m^H\mathbf{B}_m^H\mathbf{H}_{mm}^H + \mathbf{T}_m^H\mathbf{B}_m^H\left(\sum_{p=1}^{M}\frac{1}{t_p}\mathbf{H}_{pm}^H\mathbf{H}_{pm}\right)\mathbf{B}_m\mathbf{T}_m + \frac{1}{t_m}\mathbf{I}_k\right\}, \tag{5}$$

written in such a way that the impact of $\mathbf{T}_m$ and $\mathbf{B}_m$ on the overall error is limited to the mth term of the summation. This way of dealing together with the interference posed on the same cluster and leaked to other clusters have been explored in other works such as.[6] SMSE is commonly used as a design criterion in the beamforming literature, since the sum of MSEs facilitates the derivation of optimal filters; see, e.g.,[7] and references therein. It is important to remark that the minimization of SMSE and maximization of sum capacity are related, although they can suffer from lack of fairness issues with less favored users.[8]

Under the proposed global MSE framework, the involved coefficients in the transmission process would be the result of decomposing the overall problem into M minimization subproblems:

$$\begin{aligned} \{\mathbf{T}_m, t_m\} &= \arg\min \mathrm{tr}\{\mathbf{E}_m\} \\ \text{s. to } \mathrm{tr}\{\mathbf{B}_m\mathbf{T}_m\mathbf{T}_m^H\mathbf{B}_m^H\} &\leq P_m, m = 1, \dots, M, \end{aligned} \tag{6}$$

with P_m the power allocated to the m-th cluster. This is the power limit for each group of m antennas; note that we are not considering per-antenna-constraints at the satellite. Additionally, we have an overall power constraint on the satellite of the form

$$\sum_{m=1}^{M} P_m = P. \tag{7}$$

All the M problems in (6) are coupled through the scalars $\{t_m\}_{m=1}^{M}$. Nevertheless, if the power constraints are active, closed-form solutions can be obtained as shown in the next section, and small degradation is expected unless the inter-cluster interference becomes highly dominant over the noise and the intra-cluster interference. Initially we will assume that the gateways have access to all channels, although this constraint will be alleviated by using average statistics with a small penalty.

III. Precoder Design

We consider a fixed BFN on-board the satellite, so the only adaptation to the channel variations takes place at the gateway precoders. By defining the Lagrangian

$$\mathcal{L}\left(\mathbf{T}_m, \gamma_m\right) = \operatorname{tr}\{\mathbf{E}_m\} + \gamma_m \left(\operatorname{tr}\{\mathbf{B}_m \mathbf{T}_m \mathbf{T}_m^H \mathbf{B}_m^H\} - P_m\right), \tag{8}$$

and equating to zero $\partial\mathcal{L}\left(\{\mathbf{T}_m\}, \{\gamma_m\}\right)/\partial\mathbf{T}_m = 0$, we get, for $m = 1, \ldots, M$,

$$\mathbf{T}_m = \sqrt{t_m}\left(\mathbf{B}_m^H \mathbf{A}_m \mathbf{B}_m + \gamma_m \mathbf{B}_m^H \mathbf{B}_m\right)^{-1} \mathbf{B}_m^H \mathbf{H}_{mm}^H, \tag{9}$$

with

$$\mathbf{A}_m \triangleq \mathbf{H}_{mm}^H \mathbf{H}_{mm} + \mathbf{\Sigma}_m, \tag{10}$$

$$\mathbf{\Sigma}_m \triangleq \sum_{\substack{p=1 \\ p \neq m}}^{M} \frac{t_m}{t_p} \mathbf{H}_{pm}^H \mathbf{H}_{pm}. \tag{11}$$

Without intercluster interference, this is the solution corresponding to the intracluster MMSE precoder:[2]

$$\mathbf{T}_m = \sqrt{t_m}\left(\mathbf{B}_m^H \mathbf{H}_{mm}^H \mathbf{H}_{mm} \mathbf{B}_m + \frac{k}{P_m}\mathbf{B}_m^H \mathbf{B}_m\right)^{-1} \mathbf{B}_m^H \mathbf{H}_{mm}^H. \tag{12}$$

The use of the precoder expression (9) at the different gateways makes it necessary to learn $\mathbf{\Sigma}_m$, built from the intercluster channels. Even in the case that $\mathbf{H}_{pm}$ were known, the scaling factors $\{t_m\}$ present in $\mathbf{\Sigma}_m$ would need coordination for their computation; a sequential process, for instance, would obtain $\{\gamma_m\}$ for an initial set of $\{t_m\}$, set which would be recomputed for the obtained values of $\{\gamma_m\}$, and so on. Message passing algorithms such as in[9] could be devised for this process. To avoid the exchange of signalling information among different cluster, we approximate $t_m/t_p \approx 1$. We have checked that an iterative exchange process to find $\{\mathbf{T}_m\}$ and $\{t_m\}$ does not provide any significant gain taking into account the complexity of the cooperative system to exchange signalling information. Assuming $t_m/t_p \approx 1$, the optimal regularization factor is obtained when power constraint in (6) is active and is proved to be $\gamma_m = k/P_m$ in the Appendix. Different solution will be tested in next section. In particular, the impact of using the expected leakage channel Gramians $\mathbf{H}_{pm}^H \mathbf{H}_{pm}$ will be evaluated.

IV. Numerical Results

We test the performance of the proposed scheme in a Monte Carlo simulation for the specifications of a multibeam satellite antenna which uses a fed reflector antenna array with $N = 155$ feeds to exchange signals with the users. In particular, we test the distributed MMSE precoders together with the unique gateway solution as reference. As representative example we have chosen the radiation pattern provided by the European Space Agency (ESA) and used in different projects and publications by researchers cooperating in Europe with ESA, see, e.g.,[2] and.[5] This radiation pattern is designed to limit the level of interference among users in systems with conservative frequency reuse and a single gateway. As opposed to this, we assume that the whole available bandwidth is used by all beams, resulting in high intra-cluster and inter-cluster interference levels. For the BFN provided by ESA, Figure 2 shows the histogram of the SIR without precoding for full-frequency reuse, obtained from evaluating the interference for the difference users and 100 realizations. As expected, many users suffer from high interference, given that this BFN is suited for a unique gateway and low co-channel interference associated to a conservative frequency reuse across beams. In the setting under study, the feeder link is shared by $M = 10$ gateways, with the corresponding clusters shown in Figure 3. Clusters are groups of ten spot-beams ($k = 10$). Each gateway uses only a subset of n feeds, which is chosen by maximizing the average gain for all users in the cluster. The allocated power to all clusters is the same, $P_m = P/M$, with P the satellite available transmit power. We assume that the different feeder links are transparent, neglecting the possible impairments in the communication between the gateways and the satellite. The randomness of the Monte Carlo simulation comes from the location of the users at the $K = 100$ spot beams; these locations are chosen from independent uniform distributions inside the different beams, with 100 users being served at each realization, and independently across realizations.

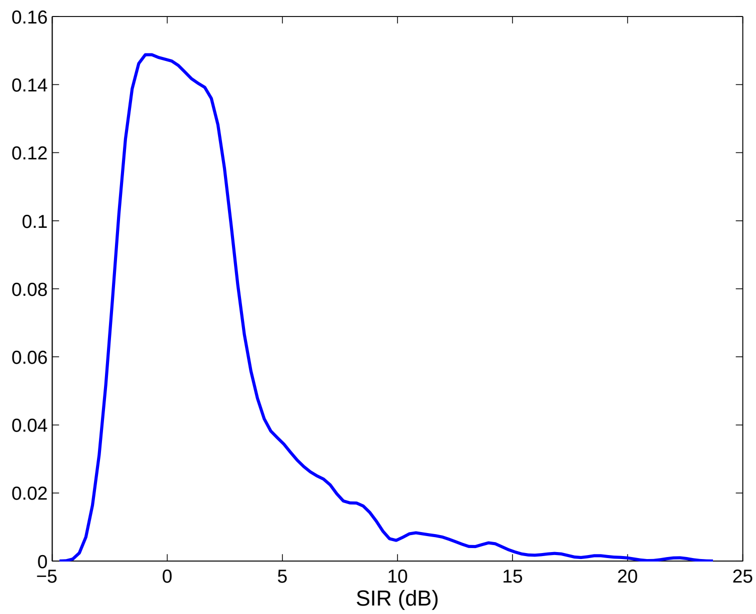

Figure 2: SIR histogram without precoding.

We test the performance of the inter-cluster aware precoders (9) for the following approximations of (11):

$$\mathbf{\Sigma}_m \approx \sum_{\substack{p=1 \\ p\neq m}}^{M} \mathbf{H}_{pm}^H \mathbf{H}_{pm}, \tag{13}$$

$$\mathbf{\Sigma}_m \approx \sum_{\substack{p=1 \\ p\neq m}}^{M} \mathbb{E}\left[\mathbf{H}_{pm}^H \mathbf{H}_{pm}\right]. \tag{14}$$

As we stated in last section, scaling factors are avoided assuming $t_m/t_p \approx 1$. The precoder making use of the second approximation allows the autonomous operation of the gateways, and only the expected Gramians of the leakage channels are required. This expectation is computed empirically.

Figures 4 and 5 present the average SINR for all users after 200 Monte Carlo realizations. Results have been obtained for two different number of feeds, $n = 10$ and $n = 30$, respectively. The performance is upper bounded by the single gateway case ($M = 1$), which serves to illustrate the loss due to the lack of data exchange among gateways. The schemes labeled as *inter* and *inter (average)* address the intercluster interference with the approximations (13) and (14). The scheme *intra* employs the intra-cluster precoder (12), which sets the lower bound. As expected, performance improves if more feeds are assigned to each gateway, keeping in mind that feeds can be shared by different gateways. This is reflected more clearly in Figure 6, where the number of feeds per cluster n is tested from 10 to 30.

Remarkably, the use of instantaneous channel matrices does not improve significantly the performance, and gives additional merit to the autonomous operation of gateways without permanent exchange of information; only average inter-cluster channel Gramians are needed for the implementation of the proposed scheme, yielding an edge with respect to the intra-cluster precoder.

V. Conclusions

The mitigation of co-channel interference in multibeam satellite settings has been addressed in this paper, for the case of several ground stations using the satellite to relay their signals to their respective clusters of beams. Both sources of interference, intra-cluster and inter-cluster, are attenuated by considering the

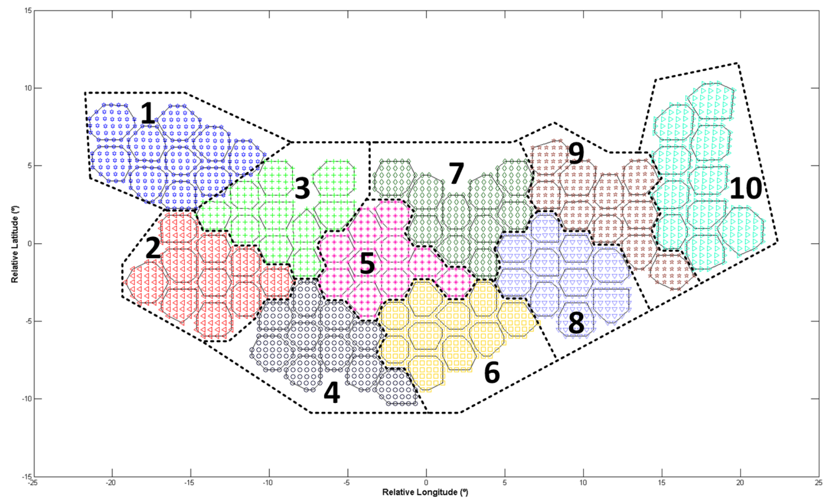

Figure 3: Spot-beams are grouped into clusters.

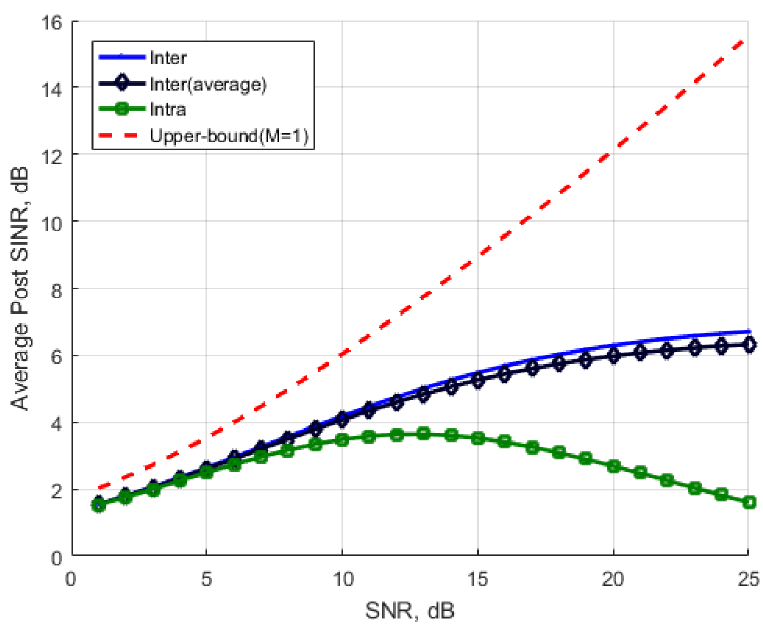

Figure 4: $M = 10, k = 10, n = 10$, 200 realizations. Upper bound: one central gateway. Lower bound: intra-cluster precoder.

global MSE and deriving inter-cluster aware precoders. The need for the instantaneous knowledge of the inter-cluster channels is avoided by using the expected Gramian of the leakage channels with satisfactory results. The optimization of the power transmitted by the satellite is a subject for further study.

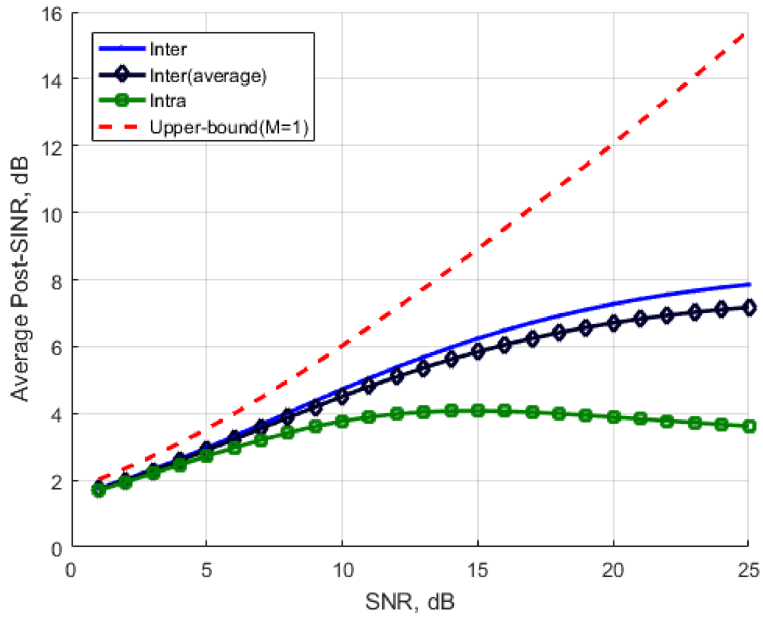

Figure 5: $M = 10, k = 10, n = 30$, 200 realizations. Upper bound: one central gateway. Lower bound: intra-cluster precoder.

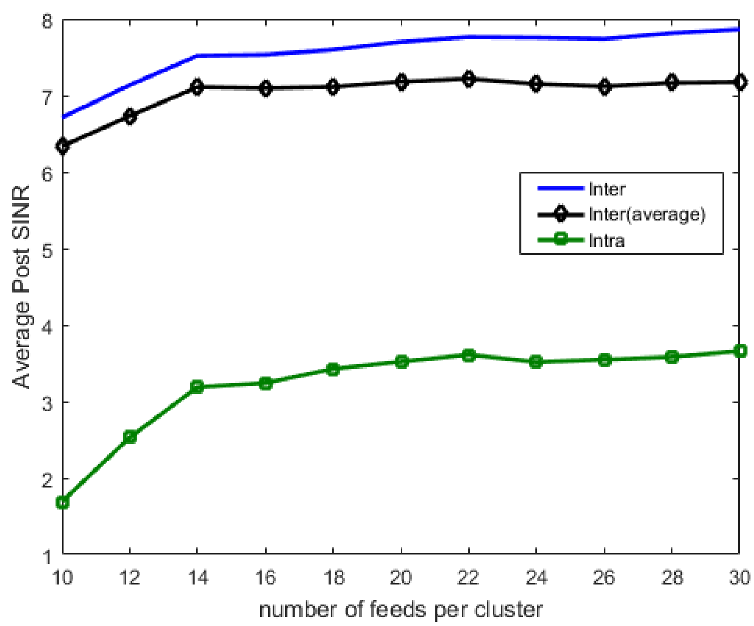

Figure 6: $M = 10, k = 10$,SNR=25 dB, 200 realizations. Performance with respect to the number of feeds per cluster.

Appendix:
Optimization of Regularization Factor

We define

$$\mathbf{F} \triangleq \mathbf{B}_m \mathbf{H}_{mm}^H \mathbf{H}_{mm} \mathbf{B}_m, \tag{15}$$
$$\mathbf{G} \triangleq \mathbf{B}_m^H (\mathbf{H}_{mm}^H \mathbf{H}_{mm} + \mathbf{\Sigma}_m) \mathbf{B}_m, \tag{16}$$

and make use of the *generalized eigenvalue decomposition*, or GEVD. The GEVD of the $k \times k$ matrices $\boldsymbol{F}$ and $\boldsymbol{G}$, with $\mathbf{G}$ positive definite, is given by

$$\boldsymbol{F}\boldsymbol{Z} = \boldsymbol{G}\boldsymbol{Z}\mathbf{\Gamma}, \qquad \text{with} \quad \mathbf{\Gamma} = \operatorname{diag}\{ \begin{matrix} \rho_1 & \rho_2 & \cdots & \rho_k \end{matrix} \}, \tag{17}$$

where the columns of $\boldsymbol{Z} \in \mathbb{C}^{k\times k}$ are the generalized eigenvectors, and $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_k$ are the generalized eigenvalues, in descending order. The number of nonzero generalized eigenvalues equals the rank of $\boldsymbol{F}$.
The GEVD is such that

$$\mathbf{Z}^H \mathbf{G} \mathbf{Z} = \mathbf{I}_k, \tag{18}$$
$$\mathbf{Z}^H \mathbf{F} \mathbf{Z} = \mathbf{\Gamma}. \tag{19}$$

For the precoder $\mathbf{T}_m$ in (9), we write $\operatorname{tr}\{\mathbf{E}_m\}$ in (5) as

$$\begin{aligned}\operatorname{tr}\{\mathbf{E}_m\} &= \operatorname{tr}\{\mathbf{I}_k\} - 2\operatorname{tr}\{(\mathbf{B}_m^H\mathbf{A}_m\mathbf{B}_m + \gamma_m\mathbf{B}_m^H\mathbf{B}_m)^{-1}\mathbf{B}_m^H\mathbf{H}_{mm}^H\mathbf{H}_{mm}\mathbf{B}_m\} + \\ &\operatorname{tr}\{(\mathbf{B}_m^H\mathbf{A}_m\mathbf{B}_m + \gamma_m\mathbf{B}_m^H\mathbf{B}_m)^{-1}\mathbf{B}_m^H\mathbf{A}_m\mathbf{B}_m\cdot(\mathbf{B}_m^H\mathbf{A}_m\mathbf{B}_m + \gamma_m\mathbf{B}_m^H\mathbf{B}_m)^{-1}\mathbf{B}_m^H\mathbf{H}_{mm}^H\mathbf{H}_{mm}\mathbf{B}_m\} + \operatorname{tr}\{(1/t_m)\mathbf{I}_k\}\end{aligned} \tag{20}$$

where t_m is obtained from (9) in such way power restriction is satifiesd, $\operatorname{tr}\{\mathbf{B}_m\mathbf{T}_m\mathbf{T}_m^H\mathbf{B}_m^H\} = P \leq P_m$. If we use the GEVD relations (18) and (19), then (20) is rewritten as

$$\begin{aligned}\operatorname{tr}\{\mathbf{E}_m\} &= k - 2\operatorname{tr}\{(\mathbf{I}_k + \gamma_m\mathbf{S})^{-1}\mathbf{U}^H\mathbf{\Gamma}\mathbf{U}\} + \\ &\operatorname{tr}\{(\mathbf{I}_k + \gamma_m\mathbf{S})^{-1}(\mathbf{I}_k + \gamma_m\mathbf{S})^{-1}\mathbf{U}^H\mathbf{\Gamma}\mathbf{U}\} + \frac{k}{P}\operatorname{tr}\{(\mathbf{I}_k + \gamma_m\mathbf{S})^{-1}\mathbf{U}^H\mathbf{\Gamma}\mathbf{U}(\mathbf{I}_k + \gamma_m\mathbf{S})^{-1}\mathbf{S}\}\end{aligned} \tag{21}$$

where we have used the eigenvalue decomposition $\mathbf{Z}^H\mathbf{B}_m^H\mathbf{B}_m\mathbf{Z} = \mathbf{U}\mathbf{S}\mathbf{U}^H$, with unitary $\mathbf{U}$ and diagonal $\mathbf{S}$. With σ_{ii} the ith diagonal element of $\mathbf{U}^H\mathbf{\Gamma}\mathbf{U}$, and λ_i the ith eigenvalue of $\mathbf{Z}^H\mathbf{B}_m^H\mathbf{B}_m\mathbf{Z}$, we can optimize γ_m to minimize $\operatorname{tr}\{\mathbf{E}_m\}$, alternatively expressed as

$$\operatorname{tr}\{\mathbf{E}_m\} = \sum_{i=1}^{k} \frac{-2\sigma_{ii}}{1+\gamma_m\lambda_i} + \frac{\sigma_{ii}}{(1+\gamma_m\lambda_i)^2} + \frac{k}{P}\frac{\sigma_{ii}\lambda_i}{(1+\gamma_m\lambda_i)^2}. \tag{22}$$

If we compute the derivative and make it equal to zero, we get to

$$\sum_{i=1}^{k} \frac{\sigma_{ii}\lambda_i^2}{(1+\gamma_m\lambda_i)^3}(\gamma_m - k/P) = 0, \tag{23}$$

from which $\gamma_m = k/P$. Now if we substitute the optimal regularization factor in (22), we obtain

$$\begin{aligned}\operatorname{tr}\{\mathbf{E}_m\} &= \sum_{i=1}^{k} \frac{-2\sigma_{ii}}{1+\frac{k}{P}\lambda_i} + \frac{\sigma_{ii}}{(1+\frac{k}{P}\lambda_i)^2} + \frac{k}{P}\frac{\sigma_{ii}\lambda_i}{(1+\frac{k}{P}\lambda_i)^2}. \\ &= \sum_{i=1}^{k} \frac{-\sigma_{ii} - \frac{k}{P}\sigma_{ii}\lambda_i}{(1+\frac{k}{P}\lambda_i)^2} = \sum_{i=1}^{k} \frac{-\sigma_{ii}(1+\frac{k}{P}\lambda_i)}{(1+\frac{k}{P}\lambda_i)^2} = \sum_{i=1}^{k} \frac{-\sigma_{ii}}{(1+\frac{k}{P}\lambda_i)}\end{aligned} \tag{24}$$

which is monotonically dreceasing function of P. Therefore, the optimal value it is obtained for the largest feasible value of P and the power restriction is active, $P = P_m$.

Acknowledgment

This work was partially funded by the Agencia Estatal de Investigación (Spain) and the European Regional Development Fund (ERDF) through the Projects COMPASS under Grant TEC2013-47020-C2-1-R, MYRADA under Grant TEC2016-75103-C2-2-R, WINTER under Grant TEC2016-76409-C2- 2-R, and COMONSENS under Grant TEC2015-69648-REDC, in part by the Xunta de Galicia (Agrupación Estratéxica Consolidada de Galicia accreditation 2016-2019, Red Temática 2017-2018), and in part by ERDF.

Discussions with Vahid Joroughi are gratefully acknowledged.

References

[1]D. Minoli, *Innovations in Satellite Communications Technology.* John Wiley & Sons, Inc., Hoboken, NJ, USA, 2015.

[2]B. Devillers, A. Pérez-Neira, and C. Mosquera, "Joint linear precoding and beamforming for the forward link of multi-beam broadband satellite systems," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE.*

[3]S. Chatzinotas, B. Ottersten, and R. De Gaudenzi, *Cooperative and cognitive satellite systems.* Academic Press, 2015.

[4]G. Zheng, S. Chatzinotas, and B. Ottersten, "Multi-gateway cooperation in multibeam satellite systems," in *2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications-(PIMRC).* IEEE, 2012, pp. 1360–1364.

[5]V. Joroughi, M. A. Vázquez, and A. I. Pérez-Neira, "Precoding in multigateway multibeam satellite systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4944–4956, July 2016.

[6]M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 5, pp. 1711–1721, May 2007.

[7]Y. Zhang, E. Dall'Anese, and G. B. Giannakis, "Distributed optimal beamformers for cognitive radios robust to channel uncertainties," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6495–6508, 2012.

[8]B. K. Chalise and L. Vandendorpe, "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Transactions on Signal Processing*, vol. 58, no. 12, pp. 6355–6368, 2010.

[9]C. K. Wen, J. C. Chen, K. K. Wong, and P. Ting, "Message passing algorithm for distributed downlink regularized zero-forcing beamforming with cooperative base stations," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2920–2930, May 2014.