Two-level Precoding for High Throughput Satellites with non-Cooperative Gateways

Tomás Ramírez, Carlos Mosquera, Roberto López-Valcarce

***Abstract*—This paper considers the design of an on-board beamforming network for a multibeam satellite managed by several gateways. This design, based on the SLNR (Signal-to-Leakage and Noise Ratio) criterion, can be fixed based on second-order statistics of the channel. The mitigation of the co-channel interference caused by the reuse of the spectral resources across beams is complemented by multi-cluster on-ground precoders, designed to fight the intra-cluster multiuser interference. These multi-cluster on-ground precoders are also able to limit some of the intra-cluster interference posed into the network in an autonomous way. Fruitful connections with recent two-level interference mitigation schemes for terrestrial systems can be made.**

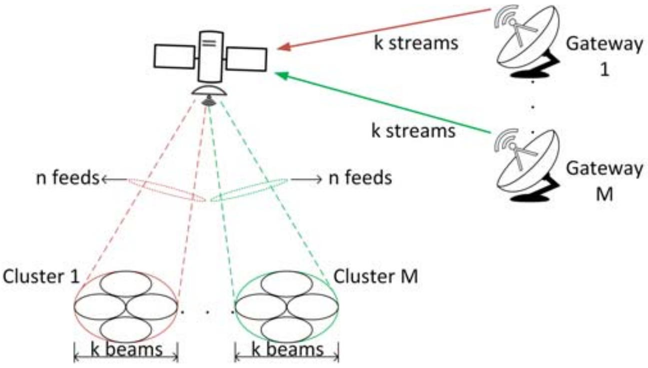

Fig. 1. Satellite shared by a number of ground stations.

I. INTRODUCTION

Ground terminals served by multibeam satellites can suffer from significant interference, especially those near the edge of the beams, if aggressive frequency reuse schemes across beams are enforced. This setting can be characterized by a multi-user multiple antenna formulation, and baseband interference mitigation schemes can be used to decrease the interference. The design of the corresponding algorithms becomes more complex when more than one Earth station or gateway serve their intented terminals through the same satellite and by making use of the same frequency resources. The use of different gateways generate several parallel channels for the feeder link which reuse the whole bandwidth while avoiding mutual interference. In contrast, the user links from satellite to user terminals are significantly affected by cross-interference caused by the side-lobes of the different radiating elements together with the adoption of full-frequency reuse.

Figure 1 shows the operation of the multibeam satellite. 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 radiation pattern of the antennas. If the same spectrum is reused across all beams, users will suffer a high level of co-channel interference. By assuming an ideal feeder link, this operation can be characterized by a Multi-User Multiple-Input Single-Output (MU-MISO) model, with multiple antennas in one site (satellite) communicating with many single-antenna terminals. The user signals transmitted by the gateways can be precoded to mitigate the co-channel interference, at least among the users under the management of each gateway; the corresponding group of beams is labeled as *cluster*. The ultimate goal is to fix the on-board BFN, since its configurability is usually quite limited, and to let the precoders at the gateways adapt to the channel changes, preferably by avoiding the interaction among the different gateways. This problem has been partially addressed in recent years for satellite scenarios, see, e.g., [1] and [2]. To some extent, fruitful connections with recent two-level interference mitigation schemes for terrestrial systems can be made. This concept has been pursued in the literature under different names:

Authors are with the atlanTTic research center, University of Vigo, Galicia, Spain. (e-mail: {tramirez, mosquera, valcarce}@gts.uvigo.es). 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. Also funded by the Xunta de Galicia (Agrupación Estratéxica Consolidada de Galicia accreditation 2016-2019; Red Temática RedTEIC 2017-2018) and the European Union (European Regional Development Fund - ERDF).

two-stage beamforming [3], hierarchical interference mitigation [4], or two-tier precoding [5]. The common ground is to design an outer beamformer to address the inter-group interference, whereas an inner precoder performs the intra-group multiplexing; in our setting, the groups are represented by the clusters. These ideas can be traced back to [6] and [7], where channel dimensionality reduction techniques are proposed for massive MIMO systems: a pre-beamforming stage applies an approximate block-diagonalization based on second-order statistics. Thus, no real-time feedback of the channel values is required to track the slow variation of the spatial correlation. Caire *et al.* applied these ideas to a single cell setting, generalizing the concept of sectorization commonly employed in cellular deployments, to grouping based on second-order statistics. Other works such as [4], [5] applied similar ideas to multiple interfering base stations.

The contribution of this paper is the design of an on-board beamforming network based on the SLNR (Signal-to-Leakage and Noise Ratio) criterion, which can be fixed based on second-order statistics of the channel. The two-level approach is complemented by the application of adaptive multi-cluster minimum mean-square error (MMSE) precoders [8], which operate with instantaneous local channel information and global statistics.

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. System Model

The satellite serves K terminals at each channel use. All K users get access to the same frequency spectrum, 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 Fig. 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 antenna 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, and $\mathbf{s}_m \in \mathbb{C}^{k \times 1}, m = 1, \ldots, M$ the symbols transmitted by the mth gateway. The transmit power is normalized as $\mathbb{E}\left[\mathbf{s}_m \mathbf{s}_m^H\right] = \mathbf{I}_k$. Only a subset of n feeds is used to give service to any given cluster, so that the weights with content in the BFN can be collected by the tall submatrices $\mathbf{B}_m \in \mathbb{C}^{n \times k}, m = 1, \ldots, M$, with $k \leq n \leq N$.

If we decompose the received signal and noise vectors into their respective vectors per cluster, $\mathbf{y}_m \in \mathbb{C}^k$ and $\mathbf{n}_m \in \mathbb{C}^k$, respectively, and denoting by $\mathbf{H}_{m,p} \in \mathbb{C}^{k \times n}$ the channel between the n feeds operated by the pth gateway and the mth cluster, then the input-output relation can be written as Eq. (1), with $\mathbf{F}_m = \mathbf{B}_m \mathbf{T}_m$ the overall matrix encompassing the precoder $\mathbf{T}_m$ and the BFN $\mathbf{B}_m$:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_M \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{1,1}\mathbf{F}_1 & \cdots & \mathbf{H}_{1,M}\mathbf{F}_M \\ \mathbf{H}_{2,1}\mathbf{F}_1 & \cdots & \mathbf{H}_{2,M}\mathbf{F}_M \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{M,1}\mathbf{F}_1 & \cdots & \mathbf{H}_{M,M}\mathbf{F}_M \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_M \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_M \end{bmatrix} \tag{1}$$

The noise $\mathbf{n}_m \in \mathbb{C}^k$ is zero-mean unit variance Additive White Gaussian Noise (AWGN), such that $\mathbb{E}\left[\mathbf{n}\mathbf{n}^H\right] = \mathbf{I}_k$.

III. BFN and Precoder Design

The proposed two-level strategy for a common BFN and separated M gateways is illustrated in Figure 2. $\widetilde{\mathbf{H}}_m$ is the $(M-1)k \times n$ matrix defined block-wise which comprises the channels from the n feeds operated by gateway m to the users served by all the other gateways:

$$\widetilde{\mathbf{H}}_m \triangleq \begin{bmatrix} \mathbf{H}_{1,m} \\ \vdots \\ \mathbf{H}_{m-1,m} \\ \mathbf{H}_{m+1,m} \\ \vdots \\ \mathbf{H}_{M,m} \end{bmatrix}. \tag{2}$$

In this way, the power of all the inter-cluster interference caused by gateway m is given by

$$\sum_{i \neq m} \mathbb{E}\left[\|\mathbf{H}_{i,m}\mathbf{F}_m\mathbf{s}_m\|^2\right] = \mathrm{tr}\{\mathbf{F}_m^H \widetilde{\mathbf{H}}_m^H \widetilde{\mathbf{H}}_m \mathbf{F}_m\}. \tag{3}$$

On the other hand, the total power received by users in the mth cluster from gateway m (including intra-cluster interference) is

$$\mathbb{E}\left[\|\mathbf{H}_{m,m}\mathbf{F}_m\mathbf{s}_m\|^2\right] = \mathrm{tr}\{\mathbf{F}_m^H \mathbf{H}_{m,m}^H \mathbf{H}_{m,m} \mathbf{F}_m\}. \tag{4}$$

A reasonable metric to quantify precoder performance is the ratio of the intra-cluster signal power (4) to the off-cluster leaked interference (3) plus noise, termed *Signal-to-Leakage plus Noise Ratio:*

$$\mathrm{SLNR}_m = \frac{\mathrm{tr}\{\mathbf{F}_m^H \mathbf{H}_{m,m}^H \mathbf{H}_{m,m} \mathbf{F}_m\}}{\mathrm{tr}\{\mathbf{F}_m^H \widetilde{\mathbf{H}}_m^H \widetilde{\mathbf{H}}_m \mathbf{F}_m\} + \mathrm{tr}\{\mathbb{E}\left[\mathbf{n}_m^H \mathbf{n}_m\right]\}}. \tag{5}$$

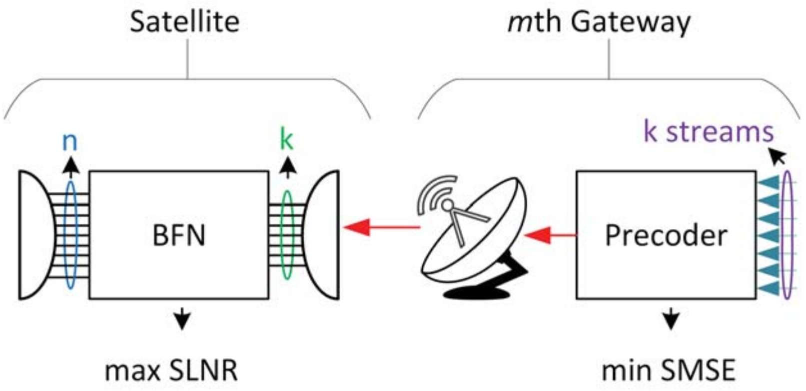

Fig. 2. Two-level precoding.

In practice it may not be feasible to optimize $\mathbf{F}_m$ to the specific channel realization, due to the limited flexibility of the satellite BFN in most cases, so we turn to the average SLNR as design metric. Since $\mathbb{E}[X/Y] \geq \mathbb{E}[X]/\mathbb{E}[Y]$ for independent random variables X and positive Y [9], then

$$\mathbb{E}[\mathrm{SLNR}_m] \geq \frac{\mathrm{tr}\{\mathbf{F}_m^H \mathbb{E}\left[\mathbf{H}_{m,m}^H \mathbf{H}_{m,m}\right] \mathbf{F}_m\}}{\mathrm{tr}\{\mathbf{F}_m^H \mathbb{E}\left[\widetilde{\mathbf{H}}_m^H \widetilde{\mathbf{H}}_m\right] \mathbf{F}_m\} + k}. \tag{6}$$

For convenience, we will denote the expectation of the Gramians as

$$\mathcal{G}_m \triangleq \mathbb{E}\left[\mathbf{H}_{m,m}^H \mathbf{H}_{m,m}\right], \quad \widetilde{\mathcal{G}}_m \triangleq \mathbb{E}\left[\widetilde{\mathbf{H}}_m^H \widetilde{\mathbf{H}}_m\right]. \tag{7}$$

If we work with the lower bound, the optimization problem would read as

$$\begin{aligned} \mathbf{F}_m = \arg\max \frac{\mathrm{tr}\{\mathbf{F}_m^H \mathcal{G}_m \mathbf{F}_m\}}{\mathrm{tr}\{\mathbf{F}_m^H \widetilde{\mathcal{G}}_m \mathbf{F}_m\} + k} \\ \text{s. to } \mathrm{tr}\{\mathbf{F}_m \mathbf{F}_m^H\} \leq P_m. \end{aligned} \tag{8}$$

The optimization problem (9) is such that the power constraint at the solution is not necessarily active, at least for some gateways. However, the identification of such a solution requires the exchange of information among gateways, as will become clear later in the paper. In consequence, we will use the maximum available power at the gateways, with the understanding that the associated performance loss is the price to pay for an autonomous operation. Thus, we can rephrase the problem as

$$\begin{aligned} \mathbf{F}_m = \arg\max \frac{\mathrm{tr}\{\mathbf{F}_m^H \mathcal{G}_m \mathbf{F}_m\}}{\mathrm{tr}\{\mathbf{F}_m^H (\widetilde{\mathcal{G}}_m + \alpha_m \cdot \mathbf{I}_n) \mathbf{F}_m\}} \\ \text{s. to } \quad \mathrm{tr}\{\mathbf{F}_m \mathbf{F}_m^H\} = P_m \end{aligned} \tag{9}$$

where $\alpha_m = \frac{k}{P_m}$. This constrained bound for the SLNR will be used to drive the design of the joint precoding and BFN matrix $\mathbf{F}_m$.

A. Optimization of constrained SLNR

For convenience, we drop the subindex m, and write henceforth $\mathcal{G}_m$, $\widetilde{\mathcal{G}}_m$, $\mathbf{F}_m$ and α_m in (9) simply as $\mathcal{G}$, $\widetilde{\mathcal{G}}$, $\mathbf{F}$ and α respectively. We exploit the Generalized Eigenvalue Decomposition (GEVD):

$$\mathcal{G}\mathbf{Z} = (\widetilde{\mathcal{G}} + \alpha \cdot \mathbf{I}_n)\mathbf{Z}\mathbf{\Gamma}, \tag{10}$$

where $\mathbf{\Gamma}$ is $n \times n$ diagonal with the generalized eigenvalues (in descending order), and the columns of $\mathbf{Z}$ comprise the generalized eigenvectors. Then it holds that $\mathbf{Z}^H \mathcal{G} \mathbf{Z} = \mathbf{\Gamma}$ and $\mathbf{Z}^H(\widetilde{\mathcal{G}} + \alpha \cdot \mathbf{I}_n)\mathbf{Z} = \mathbf{I}_n$. Letting now $\mathbf{P} = \mathbf{Z}^{-1}\mathbf{F}$, with (economy-size[1]) SVD $\mathbf{P} = \mathbf{U}\mathbf{S}\mathbf{V}^H$, one has

$$\frac{\mathrm{tr}\{\mathbf{F}^H \mathcal{G} \mathbf{F}\}}{\mathrm{tr}\{\mathbf{F}^H (\widetilde{\mathcal{G}} + \alpha \cdot \mathbf{I}_n)\mathbf{F}\}} = \frac{\mathrm{tr}\{\mathbf{P}^H \mathbf{\Gamma} \mathbf{P}\}}{\mathrm{tr}\{\mathbf{P}^H \mathbf{P}\}} = \frac{\mathrm{tr}\{\mathbf{S}\mathbf{U}^H \mathbf{\Gamma} \mathbf{U}\mathbf{S}\}}{\mathrm{tr}\{\mathbf{S}^2\}}, \tag{11}$$

which does not depend on $\mathbf{V}$. It can be easily shown that (11) is maximized with respect to $\mathbf{S}$ (diagonal positive semidefinite) and $\mathbf{U}$ (semi-unitary) when $\mathbf{S} = \alpha \mathbf{e}_1 \mathbf{e}_1^H$ and $\mathbf{U} = [\ \mathbf{e}_1 \quad \mathbf{U}_0\ ]$, where $\alpha > 0$ is arbitrary, $\mathbf{e}_1$ is the first column of $\mathbf{I}_k$, and $\mathbf{U}_0$ is an arbitrary $n \times (k-1)$ semi-unitary matrix such that $\mathbf{U}_0^H \mathbf{e}_1 = \mathbf{0}$. However, this yields

$$\mathbf{P} = \alpha \begin{bmatrix} (\mathbf{V}\mathbf{e}_1)^H \\ \mathbf{0} \end{bmatrix} \qquad \Rightarrow \qquad \mathrm{rank}\,\mathbf{P} = 1. \tag{12}$$

Hence, the overall channel $\mathbf{H}_m \mathbf{F} = \mathbf{H}_m \mathbf{Z}\mathbf{P}$ from gateway m to its intended users would also have rank 1, clearly insufficient to support k intra-cluster users requiring k independent streams.

To avoid this undesired effect, we may fix $\mathbf{S} = \alpha \mathbf{I}_k$, in which case (11) becomes simply $\mathrm{tr}\{\mathbf{U}^H \mathbf{\Gamma} \mathbf{U}\}$. Maxi-

---

[1] In other words, $\mathbf{U}$ is $n \times k$ whereas $\mathbf{S}$ and $\mathbf{V}$ are $k \times k$.

mizing this with respect to $\mathbf{U}$ semi-unitary, the optimum is found to be $\mathbf{U} = \begin{bmatrix} \mathbf{I}_k \\ \mathbf{0} \end{bmatrix}$, yielding

$$\mathbf{P} = \alpha \begin{bmatrix} \mathbf{V}^H \\ \mathbf{0} \end{bmatrix}, \tag{13}$$

which now has full rank k. The precoder-beamformer matrix becomes $\mathbf{F} = \mathbf{Z}\mathbf{P} = \alpha \mathbf{Z}_1 \mathbf{V}^H$, where $\mathbf{Z}_1$ is $n \times k$ and comprises the first k columns of $\mathbf{Z}$, and α is determined to satisfy the transmit power restriction.

B. Splitting between Ground and Satellite

We can split $\mathbf{F} = \mathbf{B}\mathbf{T}$ in different ways; for example, if we want $\mathbf{B}$ to be semi-unitary, we use the SVD $\mathbf{Z}_1 = \mathbf{U}_1 \mathbf{S}_1 \mathbf{V}_1^H$ and then take $\mathbf{B} = \mathbf{U}_1$ and $\mathbf{T} = \alpha \mathbf{S}_1 \mathbf{V}_1^H \mathbf{V}^H$. There is no reason why this constrained factorization of the precoding matrix should be able to mitigate the intra-cluster interference -some simulations showed us that this is not the case. In fact, the BFN design is oriented to causing as little inter-cluster interference as possible to other gateways' users, but it does not take into account the issue of intra-cluster interference. In consequence, we keep the BFN design as $\mathbf{B} = \mathbf{U}_1$ to reduce the inter-cluster interference, and drive the precoder design by the sum mean-square error (SMSE) metric, given by

$$\mathrm{SMSE} = \sum_{m=1}^{M} \mathrm{tr}\{\mathbf{E}_m\}, \tag{14}$$

with

$$\mathbf{E}_m \triangleq \mathbb{E}\left[(\mathbf{s}_m - \hat{\mathbf{s}}_m)(\mathbf{s}_m - \hat{\mathbf{s}}_m)^H\right]. \tag{15}$$

In terms of SMSE, the scaling applied by the receive terminals matters. End users cannot cooperate, and we consider the particular case in which the scaling is the same for all users across the cluster, with

$$\hat{\mathbf{s}}_m = \frac{1}{\sqrt{t_m}} \mathbf{y}_m \tag{16}$$

and $1/\sqrt{t_m}$ the scaling applied by users in cluster m [10]. The involved coefficients in the transmission process are the result of decomposing the overall problem into M minimization subproblems:

$$\begin{aligned} \{\mathbf{T}_m, t_m\} &= \arg\min \mathrm{tr}\{\mathbf{E}_m\} \\ \text{s. to } \ \mathrm{tr}\{\mathbf{B}_m \mathbf{T}_m \mathbf{T}_m^H \mathbf{B}_m^H\} &\leq P_m, m = 1, \ldots, M, \end{aligned} \tag{17}$$

with P_m the power allocated to the m-th cluster. The expression of the corresponding inter-cluster MMSE precoders is given, for $m = 1, \ldots, M$, by [8]

$$\mathbf{T}_m = \sqrt{t_m} \left(\mathbf{B}_m^H \mathbf{A}_m \mathbf{B}_m + \gamma_m \mathbf{B}_m^H \mathbf{B}_m\right)^{-1} \mathbf{B}_m^H \mathbf{H}_{mm}^H, \tag{18}$$

with

$$\mathbf{A}_m \triangleq \mathbf{H}_{mm}^H \mathbf{H}_{mm} + \mathbf{\Sigma}_m, \tag{19}$$

$$\mathbf{\Sigma}_m \triangleq \sum_{\substack{p=1 \\ p \neq m}}^{M} \frac{t_m}{t_p} \mathbf{H}_{pm}^H \mathbf{H}_{pm}. \tag{20}$$

The coefficients $\{t_m\}_{m=1}^M$ can be easily obtained if the power constraint in (17) is active; otherwise, all the M problems are entangled and exchange of information among the gateways would be required to find the solution. Interestingly, the optimal regularization factor can be proved to be $\gamma_m = k/P_m$ if the dependence between $\mathbf{\Sigma}_m$ and t_m is avoided, for example, by approximating $\mathbf{\Sigma}_m$ as

$$\mathbf{\Sigma}_m \approx \tilde{\mathbf{H}}_m^H \tilde{\mathbf{H}}_m, \tag{21}$$

or

$$\mathbf{\Sigma}_m \approx \mathbb{E}\left[\tilde{\mathbf{H}}_m^H \tilde{\mathbf{H}}_m\right]. \tag{22}$$

This, together with the use of the expectation of the Gramians in the GEVD decomposition in (10), serves to come up with a practical implementation with fixed BFN and independent gateways which do not need to cooperate.

IV. RESULTS

We have tested the performance of the different schemes in a Monte Carlo simulation for the specifications of a multibeam satellite antenna which uses a fed reflector antenna array with $N = 155$ feeds and $K = 100$ beams. The radiation pattern is provided by the European Space Agency (ESA) and used, among other references, in [2] and [11]. Simulation results are presented in Figs. 3 and 4 for $n = 16$ and $n = 30$ feeds, respectively, and $k = 10$ users per cluster ($M = 10$ clusters). The average Signal to Interference and Noise Ratio, labeled as post-SINR, is shown versus the average SNR of the users[2]. The calibration of the operation point is such that the average SNR is defined as

$$\mathrm{SNR} = \mathbb{E}\left[\mathrm{tr}\{\mathbf{H}\mathbf{M}\mathbf{M}^H\mathbf{H}^H\}\right]/K \tag{23}$$

and $\mathbf{M}$ the transmit beamforming matrix

$$\mathbf{M} = \frac{\sqrt{P}}{\sqrt{\mathrm{tr}\{\mathbf{H}^H\mathbf{H}\}}} \mathbf{H}^H.$$

The upper bound curve, taken from [2], depicts the performance of an on-ground beamforming single-gateway setting, with the matrix $\mathbf{F}$ in Sec. III operating entirely on the ground station, and designed to minimize the MSE.

---

[2] The beam pattern is such that the antenna gain is not uniform across a given beam.

The fully adaptive solution corresponds to the BFN tracking the channel evolution as described in Sec. III-A, following the maximization of (5). If we fix the BFN, the corresponding metric to maximize is (6), which makes use of the expected Gramians $\mathbb{E}[\mathbf{H}^H\mathbf{H}]$ and $\mathbb{E}[\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}]$. In such a case the two *"Robust BFN"* curves are obtained. The ground precoders follow from the minimization of the MSE as per (18). Note that for the choice $\mathbf{B} = \mathbf{U}_1$, $\mathbf{B}_m^H\mathbf{B}_m = \mathbf{I}_m$ in (18). The curves labeled as *"Coordinated"* use global real-time channel information to compute the precoders, by means of (21). On the other side, the curves labeled as *"Independent"* display the performance for independent gateways, which make use of the approximation (22) to compute (18), thus avoiding the exchange of real-time information. As it can be seen, the use of second-order statistics to handle the inter-cluster channels has a very low impact on the SINR. In contrast, the lack of adaptation of the BFN has a significant impact on the performance; the gap could be reduced by allowing some degree of adjustment in the BFN. Nevertheless, the exposed methodology for the design of BFN provides a significant advantage with respect to the baseline BFN, represented by the lower curve. This fixed BFN, provided by ESA, has not been designed for the independent operation of the gateways.

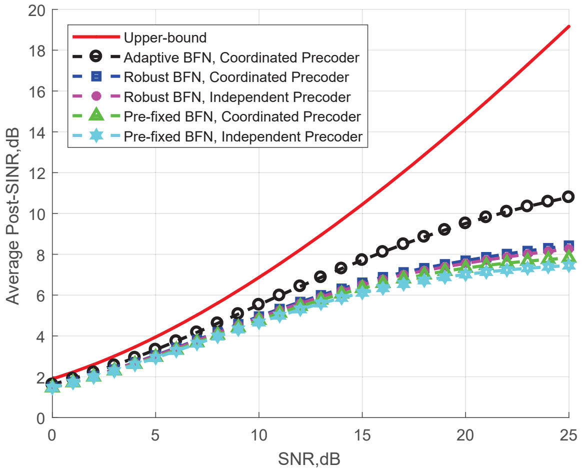

Fig. 3. $n = 16$, $k = 10$, 1000 realizations. Fixed BFN provided by ESA.

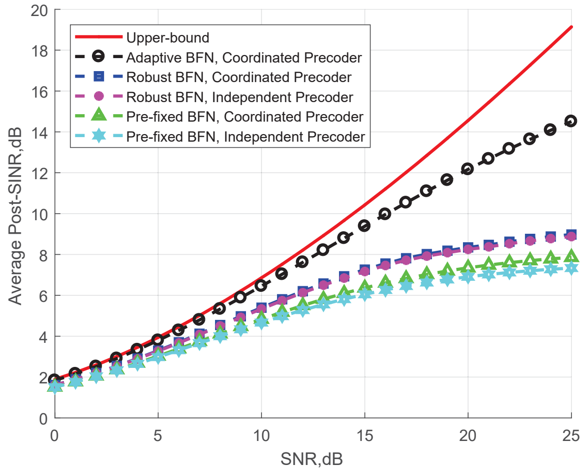

Fig. 4. $n = 30$, $k = 10$, 1000 realizations. Fixed BFN provided by ESA.

V. CONCLUSIONS

A Signal-to-Leakage and Noise Ratio criterion has been employed to design the on-board beamforming network of a multibeam satellite accessed by several gateways, each managing a cluster of beams. This criteria accounts for the inter-cluster interference; the approach is completed by using on-ground multi-cluster precoders based on the MMSE criterion to improve the multiplexing of signals inside the clusters. This two-level approach mimics some recent results devised for cellular terrestrial systems, in a context of multi-user MIMO settings. A non-adaptive BFN and non-cooperative ground precoders are also proposed for practical purposes, and based on the expected Gramians of all the underlying channels.

REFERENCES

[1] 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.

[2] C. Mosquera, R. López-Valcarce, and V. Joroughi, "Distributed precoding systems in multi-gateway multibeam satellites: Regularization and coarse beamforming," *arXiv preprint arXiv:1705.08339*, 2017.

[3] D. Kim, G. Lee, and Y. Sung, "Two-stage beamformer design for massive MIMO downlink by trace quotient formulation," *IEEE Transactions on Communications*, vol. 63, no. 6, pp. 2200–2211, June 2015.

[4] A. Liu and V. Lau, "Hierarchical interference mitigation for massive MIMO cellular networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 18, pp. 4786–4797, Sept 2014.

[5] J. Chen and V. K. N. Lau, "Two-tier precoding for FDD multi-cell massive mimo time-varying interference networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1230–1238, June 2014.

[6] A. Adhikary, J. Nam, J. Y. Ahn, and G. Caire, "Joint spatial division and multiplexing – the large-scale array regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, Oct 2013.

[7] J. Nam, A. Adhikary, J. Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 876–890, Oct 2014.

[8] C. Mosquera, R. López-Valcarce, and T. Ramírez, "Distributed precoding systems in multi-gateway multibeam satellites," in *35th AIAA International Communications Satellite Systems Conference*. American Institute of Aeronautics and Astronautics, 2017.

[9] K. Mullen, "A note on the ratio of two independent random variables," *The American Statistician*, vol. 21, no. 3, pp. 30–31, June 1967.

[10] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, Jan 2005.

[11] 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*.