JOINT PREDISTORTION AND PAPR REDUCTION IN MULTIBEAM SATELLITE SYSTEMS

Alberto Mengali, Bhavani Shankar Mysore R. and Björn Ottersten

Interdisciplinary Centre for Security, Reliability and Trust, (SnT), University of Luxembourg

ABSTRACT

Precoding for multibeam satellites with aggressive frequency reuse has attracted interest of late towards enhancing system capacity. Most of the works on precoding mitigate the linear co-channel interference between the beams caused by frequency reuse. However, the high power amplifier (HPA), an integral part of the satellite payload, is inherently non linear. Non-linear amplification combined with the linear co-channel interference introduces non-linear co-channel distortions at the receiver. Further, signals with very high peak to average power ratios (PAPR), typical of spectrally efficient modulations, are sensitive to the non-linear characteristic of the HPA and necessitate large back-off to have manageable distortion levels. In this work, a novel architecture comprising multistream Crest Factor Reduction (signal pre-clipping) and Signal Predistortion (SPD) in cascade, is devised to counter the non-linearities and co-channel interference in multibeam satellite systems. An iterative algorithm to optimize the parameters of the signal clipping and predistortion is devised taking recourse to analytical derivations. The proposed joint estimation paradigm is shown to compare favorably with state-of-art and provides a framework to combine predistortion and precoding.

***Index Terms*—** Multibeam Systems, Frequency Reuse, Signal predistortion, Clipping, Volterra, Crest Factor Reduction, Direct Estimation, Least Mean Squares

I. INTRODUCTION

Multibeam satellite systems offer a number of advantages including frequency reuse, smaller beams, flexible use of resources among others [1]. Many of the recently launched satellites employ a multibeam payload. The small sized beams offered by such systems are central towards realizing a high throughput broadband satellite [2]. However, an increasing demand in data fuelled by new subscriptions, applications and services as well as the scarcity of bandwidth warrants further exploitation of the multibeam architecture in a manner hitherto not considered. In particular, motivated by the gains offered by multiantenna multiuser terrestrial communications [3], aggressive frequency reuse combined with transmitter processing (precoding) to mitigate the resulting co-channel interference have been considered in multibeam satellite systems [4]. Encouraged by the gains offered by precoding, the recently issued DVB-S2x standard [5] provides frame structures that allow for the implementation such techniques. Several works have devised precoding techniques for these frame structures to enable higher system throughput using reuse factors less than the traditional value of four [6], [7], [8], [9].

While several transmission strategies have been proposed in literature, the underlying radio propagation channel, justifiably, has been assumed linear. The effect of the satellite components, like the high power amplifier (HPA), are typically assumed linear and absorbed into this model. However, the HPA, a key component in satellite processing, actually exhibits non-linear characteristics when operated in the high efficiency region (near saturation) [10]. Further, large excursions of the signal amplitude around the mean results in high Peak to Average Power Ratio (PAPR), potentially leading to severe distortions. The aforementioned non-linear distortions can be reduced by operating the HPA in the quasi-linear region; this entails backing-off the HPA from the saturation by an amount depending on the PAPR of the input signal. A high back-off, typically required for high order modulations, drastically reduces the power efficiency and, hence, results in lower signal output power. Thus the severity of distortion introduced by the HPA depends on its non-linear characteristic, input signal distribution and the required power efficiency. The generated interference is further accentuated by the on-board Input Multiplexing (IMUX) and Output Multiplexing (OMUX) filters which introduce memory.

The traditional approach to counteract the distortion effects while maintaining the required level of output power involves preprocessing the signal prior to amplification. This technique, referred to as signal predistortion (SPD), is usually performed in the baseband and is well studied in literature; polynomial functions [11], [12] and look up tables [13] have been used to implement SPD. However, spectrally efficient modulations suffer from additional distortion due to PAPR which cannot be well-compensated using SPD. This has motivated the use of PAPR reduction techniques, referred to as Crest Factor Reduction (CFR), in conjunction with SPD, see [14] and references therein. However, these works only mitigate the effects of the IMUX-HPA-OMUX cascade, but do not consider the co-channel (inter-beam) interference due to aggressive frequency reuse.

Precoding techniques alter the distribution of the input signal thereby increasing its PAPR [15] and potentially enhancing the non-linear distortions. However, precoding techniques developed for linear channel do not minimize the non-linear distortions explicitly. Joint design of precoding and predistortion has been considered in a few works [16], [17]. These works either consider a linear manipulation of the channel [16] or single channel systems [17]; further, the effect of PAPR is not considered.

Unlike the cited works, we consider a non-linear processing at the gateway (GW) of a multibeam satellite that considers transponder effects in addition to the linear co-channel interference. Building on [14], the GW processing comprises multistream CFR followed by a multistream SPD wherein SPD operates jointly on the different streams [18]. Such an approach is different from the traditional single stream SPD [12]. A novel iterative optimization framework to obtain the parameters of the multistream CFR and SPD mechanisms jointly is then analytically derived. This framework generalizes the single stream CFR and SPD optimization in [14] and serves as tool towards an unified design of key transmission strategies − predistortion, precoding and PAPR reduction.

The reminder of the paper is organized as follows: Section II present the system model emphasizing the scenario considered and models for SPD and CFR, Section III details the iterative algorithm for optimizing SPD and CFR, Section IV illustrates the gains of the joint design while Section V draws conclusions.

Notation: Lower-case bold symbols, $\boldsymbol{a}$, and upper-case bold symbols, $\boldsymbol{A}$, respectively denote vectors and matrices, ^{T} denotes the transposition, ^{H} is the Hermitian operation while $*$ indicates element-wise complex conjugation; $\otimes$ is the Kronecker product.

II. SYSTEM MODEL

II-A. Scenario Description, Space and Ground Segments

We consider a generic multibeam GEO broadband satellite offering services to fixed users in the Ka-band. The forward link of the system employs DVB-S2x as the air-interface [5]; a return channel to the satellite GW is considered towards enabling bidirectional services. A frequency reuse factor of β is used with M co-channel beams leading to a total of βM beams. We consider one user per beam served by a single carrier in this study; this can be achieved by resorting to Time Division Multiplexing (TDM) of multiple users [1]. Based on the aforementioned description, it suffices to consider a M co-channel beam system with M users (one per beam). We further assume users only have a single antenna.

The antenna of the multibeam satellite comprises a parabolic reflector and an array of feeds allowing the generation of multiple beams. We focus on a single feed per beam configuration [1], [15] where each feed signal is processed by a separate transponder before transmission. Each transponder is a cascade of IMUX, Automatic Gain Control (AGC), HPA and OMUX; typical characteristics of these can be obtained from [5].

The GW processing generates M streams for transmission on the M feeds, one per feed. Typically, these streams are multiplexed in frequency on the feeder link and are translated to the appropriate downlink frequency (same for all streams) on-board the satellite. We assume an ideal feeder link and lossless frequency translation on-board the satellite.

II-B. Signal Model

Fig. 1 abstracts the multibeam system under consideration where a different stream is transmitted to each of the M single antenna users. Let the signal on kth stream at time instance n, be

$$u_k(n) = \sum_m a_k(m) p_k(nT_s - mT), \quad k \in [1, M],$$

where $\{a_k(m)\}$ are the constellation symbols, $p_k(\cdot)$ is the pulse-shaping function, T is the symbol duration and T_s is the sample duration with $T_s = \lambda T$ and $\lambda >> 1$ is the oversampling factor. The

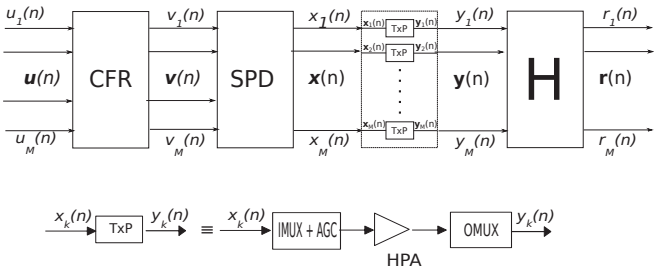

Fig. 1. System Model

signals $\{u_k(n)\}_{k=1}^M$ are then input to a cascade of CFR and SPD blocks, which, in turn, generate M streams, $\{x_k(n)\}_{k=1}^m$. These streams are then frequency multiplexed and transmitted by the GW.

The assumptions of lossless feeder links and on-board frequency translations allow us to consider each stream as being processed by a separate transponder section (TxP). The processing involves an IMUX filter to reject out-of-band noise, an AGC that adjusts the signal power according to the chosen Input Backoff (IBO) level, a memoryless HPA and an OMUX filter rejecting out-of-band emissions. Further, the IMUX/ OMUX filters need not be identical across the TxPs.

[1]The same can be extended to multiple users per beam on a single carrier by multiplexing them in a DVB-S2x frame [6]

The output signal, $\{y_k(n)\}_{k=1}^M$ is then radiated on feed k at instance n. These co-frequency signals undergo mutual interference over the transmission path from the M transmit antennas to the M receivers; the resulting channel is assumed to be frequency flat and is denoted by a $M \times M$ matrix $\boldsymbol{H}$ [6], [8]. Let $r_k(n)$ denote the received signal at user k at instance n and $\boldsymbol{r}(n) = [r_1(n), r_2(n), \ldots, r_M(n)]^T, \boldsymbol{y}(n) = [y_1(n), y_2(n), \ldots, y_M(n)]^T$. With $\boldsymbol{\eta}(n)$ denoting the stacked receiver front-end noise, we have,

$$\boldsymbol{r}(n) = \boldsymbol{H}\boldsymbol{y}(n) + \boldsymbol{\eta}(n). \tag{1}$$

II-C. Transponder Model

We consider a baseband Volterra non-linear system with memory to model each of the transponders [19]. In this paper, we truncate the infinite Volterra series to degree 3 with a memory depth of K_s for all orders. Further, for ease of presentation, we assume identical transponder characteristics in the sequel; incorporation of different characteristics is straightforward, but would require a burdensome and less readable notation. Using the formulation in [14], we have, $\forall k \in [1, M]$,

$$y_k(n) = \boldsymbol{p}_1^T \boldsymbol{G}_1 \boldsymbol{x}_k(n) + \boldsymbol{p}_3^T \boldsymbol{G}_3 \left[\boldsymbol{x}_k(n) \otimes \boldsymbol{x}_k(n) \otimes [\boldsymbol{x}_k(n)]^*\right], \tag{2}$$

where $\boldsymbol{x}_k(n) = [x_k(n), x_k(n+1), \ldots, x_k(n+K_s-1)]^T$ is a $K_s \times 1$ vector and $\boldsymbol{p}_d$ denotes the column vector of the transponder model coefficients. We can write the set of equations in (2) compactly as,

$$\boldsymbol{y}(n) = \left[\boldsymbol{G}_1 \boldsymbol{X}(n)\right]^T \boldsymbol{p}_1 + \left\{\boldsymbol{G}_3 [\boldsymbol{X}(n) \star \boldsymbol{X}(n) \star \boldsymbol{X}^*(n)]\right\}^T \boldsymbol{p}_3, \tag{3}$$

where, $\star$ denotes the column-wise Khatri-Rao matrix product [20] and $\boldsymbol{X}(n) = [\boldsymbol{x}_1(n), \boldsymbol{x}_2(n), \ldots, \boldsymbol{x}_M(n)]$ (a $K_s \times M$ matrix). In (3), $\boldsymbol{G}_d$ represents the $L_d(K_s) \times (K_s)^d$ matrix that selects the $L_d(K_s)$ non-redundant terms of the dth order Khatri-Rao product. Further, $L_d(k) = C\left(k - 1 + \frac{(d+1)}{2}, \frac{(d+1)}{2}\right) \times C\left(k - 1 + \frac{(d-1)}{2}, \frac{(d-1)}{2}\right)$ where $C(n,k) = \frac{n!}{k!(n-k)!}$ and $k!$ is factorial of k. Finally $\boldsymbol{p}_d$ is the $L_d(K_s) \times 1$ Volterra coefficient vector corresponding to the degree d.

Unlike in the multicarrier scenario [21], [22], different streams (columns of $\boldsymbol{X}$) are amplified separately. This effect is captured by the Khatri-Rao product efficiently compared to the widely used Kronecker products [2]

II-D. Crest Factor Reduction Model

Building on the single CFR formulation in [14], a multistream CFR is described below, where the output of the kth CFR, $v_k(n)$, is related to its input, $u_k(n)$, as,

$$v_k(n) = \begin{cases} u_k(n) & \text{if } |u_k(n)| \le |\gamma_k|^2, \\ |\gamma_k|^2 \; u_k(n)/|u_k(n)| & \text{if } |u_k(n)| > |\gamma_k|^2 \end{cases} \tag{4}$$

where γ_k is the clipping parameter for the kth stream. Since a joint design of $\{\gamma_k\}$ is employed, we will use $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_M]^T$.

II-E. Multistream SPD Model

As depicted in Fig. 1, the kth stream has its dedicated HPA; however, the downlink channel $\boldsymbol{H}$ introduces interference from the output of other HPAs onto $r_k(n)$. To cater to this effect, we consider the general architecture of a multistream predistortion in the sample domain where the streams $\{u_k(n)\}_{k=1}^M$ are processed jointly with an aim to reduce the interference on $\{r_k(n)\}_{k=1}^M$. This predistorter has the multiple input/ multiple output character of a multicarrier

[2]It is easy to verify that $\boldsymbol{Q}_1 \star \cdots \star \boldsymbol{Q}_n = (\boldsymbol{Q}_1 \otimes \cdots \otimes \boldsymbol{Q}_n)\boldsymbol{S}_{m^n \times m}$ [20] where $\{\boldsymbol{Q}_i\}$ are $P \times m$ matrices and $\boldsymbol{S}_{m^n \times m}$ is a $m^n \times m$ selection matrix that extracts only m columns out of the m^n columns resulting from the full Kronecker product.

data predistortion [22], but operates in the signal domain [11]; a similar architecture for two channels is presented in [18].

The multistream predistorter is also modelled using a third degree baseband Volterra series with each degree having a memory of K_p. Towards describing the signal model of the predistorter, we let $\boldsymbol{v}_k(n) = [v_k(n), v_k(n+1), \ldots, v_k(n+K_s+K_p-2)]^T$ to be the $(K_s+K_p-1)\times 1$ vector of samples input to the predistorter, $\boldsymbol{V}(n) = [\boldsymbol{v}_1(n), \boldsymbol{v}_2(n), \ldots, \boldsymbol{v}_M(n)]$, $\boldsymbol{w}_k$ to be N_w predistorter coefficients used for generating the kth predistorted stream and $\boldsymbol{\Phi}(n)$ is a $K_s \times N_w$ matrix drawing elements from $\boldsymbol{V}$ and $\boldsymbol{V}\otimes\boldsymbol{V}\otimes\boldsymbol{V}^*$ [12]. Clearly, $\boldsymbol{\Phi}(n)$ contains terms of the form $v_k(n)$ and $v_{k_1}(n-l_1)v_{k_2}(n-l_2)v^*_{k_3}(n-l_3)$. Recalling the notation of $\boldsymbol{x}_k(n)$ from Section II-B, we have [14],

$$\boldsymbol{x}_k(n) = \boldsymbol{\Phi}(n)\boldsymbol{w}_k, k \in [1, M]. \tag{5}$$

The number of predistorter coefficients per stream depends on the number of streams, memory and degree of the predistorter model. It can be enumerated as, $N_w = M(L_1(K_p) + L_3(K_p))$. Further, (6) indicates that the generation of kth predistorted stream uses data from all the streams (available through $\boldsymbol{\Phi}(n)$). For subsequent analysis, we will use

$$\boldsymbol{X}(n) = \boldsymbol{\Phi}(n)\boldsymbol{W}, \tag{6}$$

where $\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_M]$ is the $N_w \times M$ predistorter matrix. Without loss of generality, the number of predistorter coefficients is assumed to be the same for all k.

III. JOINT CFR AND MULTISTREAM PREDISTORTER DESIGN

The objective of this study is to determine the CFR thresholds, $\boldsymbol{\gamma}$, and multistream SPD coefficients, $\{\boldsymbol{w}_k\}$, that minimize the interference at the receiver. In particular, defining $\boldsymbol{u}(n) = [u_1(n), u_2(n), \ldots, u_M(n)]^T$, the cost function reduces to,

$$\boldsymbol{W}_{opt}, \boldsymbol{\gamma}_{opt} = \arg\min_{\{\boldsymbol{W},\boldsymbol{\gamma}\}} \sum_n ||\boldsymbol{r}(n) - \boldsymbol{u}(n)||_2^2, \tag{7}$$

where $||\cdot||_2$ is the l_2 norm. Clearly, SPD affects signal distribution and hence the PAPR, thereby impacting the CFR design. On the other hand, the predistorter design is influenced by the CFR since the latter affects the input distribution to SPD. Thus the design of CFR and SPD are dependent and this section deals with an iterative approach for their design. Fig. 2 illustrates the considered transmitter architecture including the two iterative optimization processes where, CFR and SPD blocks implement the models in Sections II-D and II-E respectively. Further, the *CHANNEL* block in Fig. 2 refers to the cascade of the transponder and $\boldsymbol{H}$.

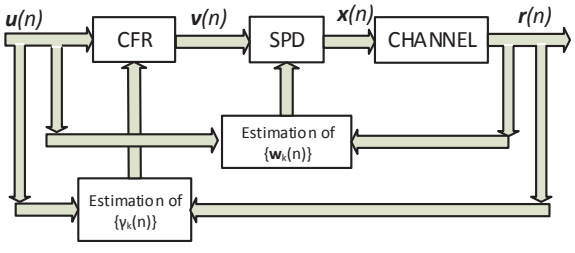

Fig. 2. Schematic depicting Joint Optimization of CFR and SPD

III-A. Estimation of the Predistorter Coefficients

Let $e_k(n) = r_k(n) - u_k(n)$ denote the interference (both non-linear and co-channel) at the receiver k and $\boldsymbol{e}(n) = \boldsymbol{r}(n) - \boldsymbol{u}(n)$. The target of the predistortion function is to reduce interference power, $E\left[||\boldsymbol{e}(n)||_2^2\right] \propto \sum_n ||\boldsymbol{r}(n) - \boldsymbol{u}(n)||_2^2$, by a proper choice of $\{\boldsymbol{w}_k\}$. We employ the *direct learning* paradigm [14] since the SPD has to invert not only the *CHANNEL* block (as in the traditional indirect method [12]) but should also account for the actions of the CFR block. We achieve the objective in (7) by formulating a least mean squares (LMS) algorithm for minimizing $C(\boldsymbol{w}(n)) = \boldsymbol{e}^H(n)\boldsymbol{e}(n)$.

Denoting $D_{\boldsymbol{x}}$ to be the complex partial differentiation operator with regards to vector $\boldsymbol{x}$ [23] and $\boldsymbol{w} = [\boldsymbol{w}_1^T, \boldsymbol{w}_2^T, \ldots, \boldsymbol{w}_M^T]^T$, the update equation for $\boldsymbol{w}$ takes the form,

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \mu D_{\boldsymbol{w}^*(n)} C(\boldsymbol{w}(n)) \tag{8}$$

where μ is the step-size and $\boldsymbol{w}(n)$ is the estimate of $\boldsymbol{w}$ at iteration n. Towards obtaining $D_{\boldsymbol{w}^*(n)}C(\boldsymbol{w}(n))$, we note that the relation between $\boldsymbol{w}(n)$ and $\boldsymbol{e}(n)$ can be obtained through (1), (3) and (6) and involves complex variables and their conjugates. It is therefore natural to apply the chain rule in [23] involving the variables $\boldsymbol{e}(n), \boldsymbol{r}(n), \boldsymbol{y}(n), \mathbf{X}(n)$ in succession. Following the differentiation rules in [23], it can be shown that $D_{\boldsymbol{r}^*(n)}e(n), D_{\boldsymbol{r}(n)}e^*(n), D_{\boldsymbol{y}^*(n)}r(n), D_{\boldsymbol{y}(n)}r^*(n)$ are all zero. Using these, we have,

$$\begin{aligned} D_{\boldsymbol{w}^*}C(\boldsymbol{w}) &= (D_{\boldsymbol{e}}C(\boldsymbol{w}))(D_{\boldsymbol{r}}\boldsymbol{e})(D_{\boldsymbol{y}}\boldsymbol{r})(D_{\boldsymbol{X}^*}\boldsymbol{y})(D_{\boldsymbol{W}^*}\boldsymbol{X}^*) \\ &+(D_{\boldsymbol{e}^*}C(\boldsymbol{w}))(D_{\boldsymbol{r}^*}\boldsymbol{e}^*)(D_{\boldsymbol{y}^*}\boldsymbol{r}^*)(D_{\boldsymbol{X}^*}\boldsymbol{y}^*)(D_{\boldsymbol{W}^*}\boldsymbol{X}^*). \end{aligned} \tag{9}$$

Using the rules in [23] and after manipulations, various components of (9) can be analytically obtained. For ease of comprehension and lack of space, these are delegated to the footnote [3]. Once the gradient is determined, the update step in (9) follows.

It should be noted that the coefficients characterising the transponder $\{\boldsymbol{p}_d\}$ and the linear downlink channel $\boldsymbol{H}$ are assumed to be known at the transmitter. In fact, the estimation $\{\boldsymbol{p}_d\}$ is a part of the direct estimation paradigm [22] and $\boldsymbol{H}$ can be obtained from a training phase where the HPA is driven in a quasi-linear region. Further, the transmitter is also assumed to have perfect information about the received signal at all the receivers, $\boldsymbol{r}(n)$, during the training. This can be achieved by using operator installed dedicated receivers. In order to enhance performance, the vector of predistortion parameters, $\boldsymbol{w}$, is initialized with the result of the off-line indirect estimation where CFR is assumed to be absent.

III-B. Adaptive Crest Factor Reduction

We now optimize the clipping factor of the kth stream, γ_k, to reduce the PAPR and improve system performance. Since a multistream predistorter is considered, a natural choice is to consider a joint design of $\{\gamma_k\}$. Following the methodology in [14], we obtain an equivalent non-linear channel after cascading SPD and the bank of parallel HPAs. Such a system is non-linear, and following (3), the relation between $\boldsymbol{v}(n)$ and $\boldsymbol{y}(n)$ can be written using Kronecker products as,

$$\boldsymbol{y}(n) = [\boldsymbol{\Pi}_1\boldsymbol{V}(n)]^T\boldsymbol{q}_1 + \{\boldsymbol{\Pi}_3[\boldsymbol{V}(n)\otimes\boldsymbol{V}(n)\otimes\boldsymbol{V}^*(n)]\}^T\boldsymbol{q}_3. \tag{10}$$

where $\boldsymbol{V}(n)$ is defined in Section II-E, $\boldsymbol{q}_d$ corresponds to the degree d coefficient of the cascaded model and $\boldsymbol{\Pi}_k$ is the selection matrix for order d. Recalling (4) and the definition of $\boldsymbol{V}$, we note that the

[3] $D_{\boldsymbol{e}^*}C(\boldsymbol{w}) = \boldsymbol{e}^T, D_{\boldsymbol{e}}C(\boldsymbol{w}) = \boldsymbol{e}^H, D_{\boldsymbol{r}^*}\boldsymbol{e}^* = \boldsymbol{I}_M, D_{\boldsymbol{r}}\boldsymbol{e} = \boldsymbol{I}_M$, $D_{\boldsymbol{y}^*}\boldsymbol{r}^* = \boldsymbol{H}^H, D_{\boldsymbol{y}}\boldsymbol{r} = \mathbf{H}$, $D_{\boldsymbol{X}}\boldsymbol{y} = (\boldsymbol{I}_M\otimes\boldsymbol{p}_1^T\boldsymbol{G}_1) + (\boldsymbol{S}^T_{M^3\times M}\otimes \boldsymbol{p}_3^T\boldsymbol{G}_3)\boldsymbol{A}^{(3)}(\boldsymbol{X}^*)$, $D_{\boldsymbol{W}}\boldsymbol{X} = \boldsymbol{I}_M\otimes\boldsymbol{\Phi}, D_{\boldsymbol{W}^*}\boldsymbol{X}^* = \boldsymbol{I}_M\otimes\boldsymbol{\Phi}^*$, where, $\boldsymbol{A}^{(3)}(\boldsymbol{Z}^*) = \boldsymbol{A}(\boldsymbol{Z}^*, \boldsymbol{Z}\otimes\boldsymbol{Z})[\boldsymbol{A}(\boldsymbol{Z},\boldsymbol{Z}) + \boldsymbol{B}(\boldsymbol{Z},\boldsymbol{Z})]$, and the definitions of $\boldsymbol{A}(\cdot,\cdot), \boldsymbol{B}(\cdot,\cdot)$ are available in [23] and $\boldsymbol{S}$ is a selection matrix described in footnote 2.

elements of $\boldsymbol{V}$ contain information about $\{\gamma_k\}$. Hence, we propose a LMS algorithm similar to the previous section as,

$$\boldsymbol{\gamma}(n+1) = \boldsymbol{\gamma}(n) - \epsilon D_{\boldsymbol{\gamma}^*} C(\gamma) \tag{11}$$

where ϵ is the algorithm step and $C(\boldsymbol{\gamma}) = ||\boldsymbol{r}(n) - \boldsymbol{u}(n)||_2^2$. We use the chain rule similar to Section III-A to evaluate the gradient. In particular, the gradient contains 8 terms comprising terms of the form, $(D_{\boldsymbol{e}}C), (D_{\boldsymbol{r}}\boldsymbol{e}), (D_{\boldsymbol{y}}\boldsymbol{r}), (D_{\mathbf{X}}\mathbf{y}), (D_{\boldsymbol{\Phi}}\mathbf{X}), (D_{\boldsymbol{V}}\boldsymbol{\Phi})$ as well as, $(D_{\boldsymbol{e}^*}C), (D_{\boldsymbol{r}^*}\boldsymbol{e}), (D_{\boldsymbol{y}^*}\boldsymbol{r}), (D_{\boldsymbol{X}^*}\boldsymbol{y}), (D_{\boldsymbol{\Phi}^*}\boldsymbol{X}), (D_{\boldsymbol{V}^*}\boldsymbol{\Phi})$. These quantities can be obtained from system equations and are omitted for brevity.

While two options on sequencing the SPD and CFR optimizations are discussed in [14], we follow alternate minimization of CFR and SPD. The number of iterations depend on the termination criteria.

IV. PERFORMANCE EVALUATION

In this section we numerically evaluate the performance of the algorithms for a selected case of study.

IV-A. Simulation Scenario

Fig. 3 illustrates the considered scenario having $M = 7$ co-channel beams with identical coverage radii of 250km and a full frequency reuse ($\beta = 1$). The layout corresponds to the first tier of beams in a classical circular tessellation. We further consider users at the centre of their respective beams in this first study.

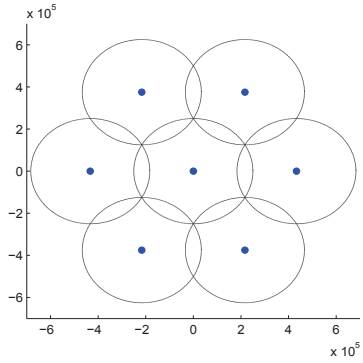

Fig. 3. 7 beam layout considered for simulations.

Following the works of [7] and [15], we obtain $\boldsymbol{H} = \boldsymbol{\Theta}\boldsymbol{B}$, where $\boldsymbol{B}$ denotes the beam gain matrix and the signal phase due to different propagation paths between the users is captured by $\boldsymbol{\Theta}$. Given the ith user position, we define $\theta_{k,i}$ as the off-axis angle of the user i with respect to the bore-sight of the beam k and the 3 dB angle for the kth beam by $\theta_{k3\text{dB}}$. Then the beam gain from kth antenna feed to ith user (the (i,k) element of $\boldsymbol{B}$) is approximated as the square root of $G_{s,k}\left(\frac{J_1(u)}{2u} + 36\frac{J_3(u)}{u^3}\right)^2$ [15]. Here, $G_{s,k}$ is the satellite transmit antenna gain for the kth beam and $u = 2.07123\frac{\sin(\theta_{k,i})}{\sin(\theta_{k3\text{dB}})}$. Further, J_1 and J_3, respectively, are the first and third order Bessel functions of first kind. In this work, we assume $G_{s,k} = 1, \theta_{k3\text{dB}} = 0.4^o \ \ \forall k$ and $\boldsymbol{\Theta} = \boldsymbol{I}$ [15].

All the streams $\{u_k(n)\}$ employ identical modulations (16 or 32 APSK), symbol rates $R_s = 4$ MBaud and roll-off factors $\rho = 0.2$.

Typical responses for the IMUX and OMUX filter are extracted from [5] and frequency scaled to yield wideband filters (500 MHz). This exercise is undertaken to minimize the effect of these filters and focus on the effect of HPA in this study. Further, we model the HPA using the standard memoryless Saleh model with AM/AM and AM/PM functions taking the form $A(r) = \frac{2r}{1+r^2}; \Phi(r) = \frac{\pi}{6}\frac{r^2}{1+r^2}$, respectively [10].

IV-B. Simulation Set-Up

We first choose the SPD and channel model parameters. In particular, we let, $K_s = K_p = 1$ (channel and SPD models are both memoryless). While this choice allows for faster simulations, the results obtained are rather conservative due to model mismatches. The channel parameters, $\boldsymbol{p}_d$ (kindly refer to (3)) and $\boldsymbol{q}_d$ (in (10)) are estimated on-line using standard least squares (LS) techniques [22]. It should be noted that these models are used only for training CFR and SPD; performance evaluations use the transponder per-se.

The SPD and CFR estimation algorithms, described respectively in Sections III-A and III-B, are trained over 10^6 samples corresponding to 10^5 input symbols per stream and an oversampling factor $\lambda = 10$. Step-size parameters are obtained by employing additional simulations for their search. In the sequel, we use, $\mu = 10^{-5}$ in (8), $\epsilon = 10^{-3}$ in (11) and undertake training of CFR and SPD in the absence of noise. It should be noted that the convergence of the component LMS algorithms (SPD/ CFR) is sensitive to transponder, $\boldsymbol{H}$ and the step-size parameters. Further, it suffices to use the schedule involving a first SPD estimation followed by CFR estimation and finally a SPD update due to diminishing gains.

For purposes of comparison with the proposed method, we consider,

- MMSE precoder taking the form $\left(\boldsymbol{H}^*\boldsymbol{H} + \sigma^2\mathbf{I}\right)^{-1}\boldsymbol{H}^*$, where σ^2 is the receiver noise variance. This model caters to co-channel interference and ignores the non-linear distortions.
- Cascade of MMSE precoder and a non-linear compensation technique (in that order), where both are independently designed. We use the single channel CFR and SPD cascade obtained from [22] on each of the streams as the non-linear compensation technique. This reference case reflects the idea of first minimizing the non-linear effects and subsequently using the MMSE precoder to mitigate linear co-channel interference. The non-linear compensation technique is designed without $\boldsymbol{H}$ while the MMSE precoder is obtained without the transponder.
- Single beam using the CFR and SPD cascade of [22]. This case does not suffer from co-channel interference and serves as an upper bound.

IV-C. Performance Evaluation

We evaluate resulting Signal to Interference plus Noise Ratio of the central beam (worst performance) and one of the remaining beams having the best performance as a function of the Output Back-off (OBO). The OBO is defined as $OBO = 10\log_{10}(P_{out}/P_{out}^{\text{S}})$, where P_{out} is the RF power used and P_{out}^{S} is the amplifier saturation output power. OBO needs to be accounted since it depicts the reduction in RF power and a fair comparison is achieved by enumerating the performance of different techniques at identical OBO.

Figs. 4 and 5, respectively, illustrate the worst and best beam SINR as a function of OBO for 32 APSK constellations for $E_s/N_0 = 20$ dB. The achieved SINR values correspond to the region of operation involving 32 APSK [5]. We sweep IBO (Input Back-off) from 2 dB to 9 dB in steps of 1 dB to obtain different OBO and their corresponding SINR. The MMSE precoder and the proposed technique exhibit a larger variation in OBO over this range compared to the other techniques. This indicates the sensitivity of signal distributions to IBO for these methods.

Evidently, the proposed joint multistream CFR and SPD mechanism performs better than the independent design or MMSE precoder. It is about 2 dB (OBO = 3.5 dB) from the single channel bound for central beam (kindly refer to Fig. 4); the loss reduces to 1 dB (OBO = 3.5 dB) for the peripheral beam (Fig. 5). This is a straightforward manifestation of the co-channel interference. Further, the MMSE precoding performs poorly due to the non-linearity and a high IBO is needed to extract meaningful

performance from this scheme. On the other hand, the proposed scheme achieves gains of 0.5 dB in SINR (at OBO $\approx$ 4 dB) and 0.3 dB in OBO (at SINR = 13.5 dB) over the independent precoder and non-linear compensation cascade in Fig. 4. This is slightly reduced in Fig. 5 due to lower co-channel interference.

Similar effects are seen for 16 APSK constellations whose performance is plotted in Figs. 6 and 7 for the central and peripheral beams respectively.

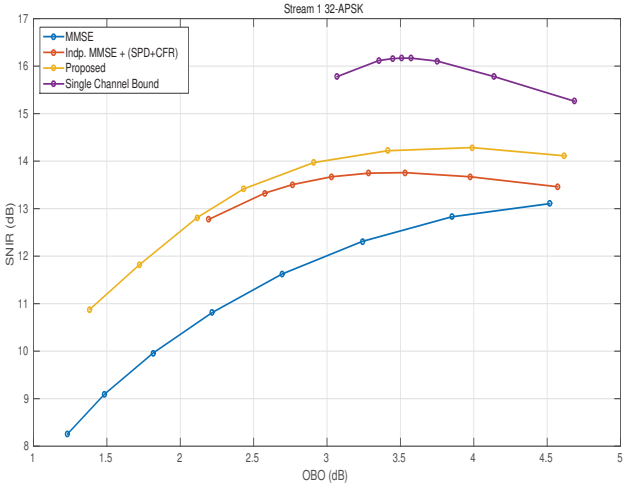

Fig. 4. SINR performance of the central (worst) beam as a function of OBO, SNR = 20 dB, 32 APSK modulation.

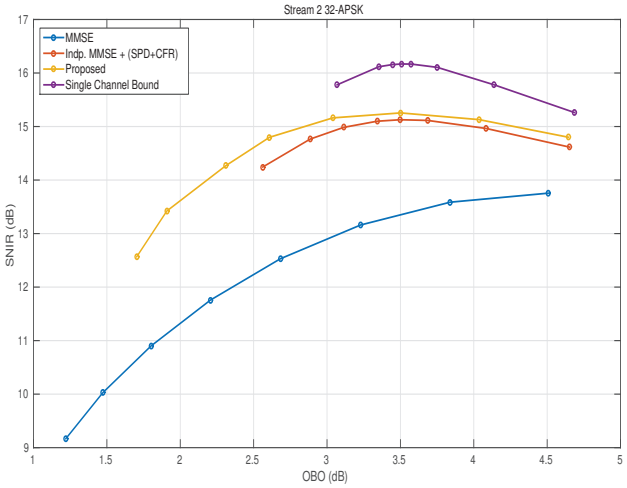

Fig. 5. SINR performance of a peripheral (best) beam as a function of OBO, SNR = 20 dB, 32 APSK modulation.

V. CONCLUSION

We propose a novel GW processing algorithm for minimizing co-channel interference in multibeam systems when the transponder characteristics are non-linear. This framework comprises multi-stream CFR and SPD blocks in cascade. LMS-based direct estimation paradigm is used to determine the optimal predistortion coefficients and clipping thresholds. An analytical approach is pursued to determine the components of the LMS algorithm and a numerical performance evaluation depicts gain in chosen reference

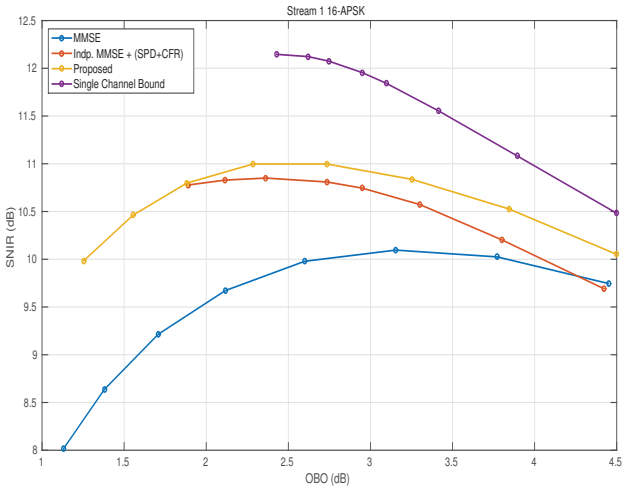

Fig. 6. SINR performance of the central (worst) beam as a function of OBO, SNR = 15 dB, 16 APSK modulation.

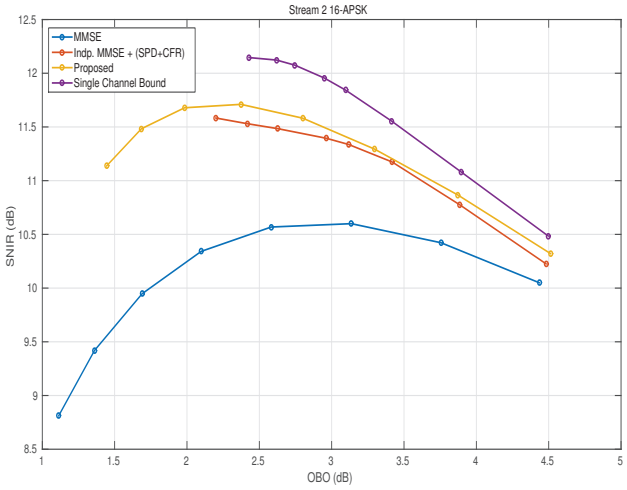

Fig. 7. SINR performance of a peripheral (best) beam as a function of OBO, SNR = 15 dB, 16 APSK modulation.

scenarios. The need to jointly design the key interference mitigation techniques at the transmitter is highlighted.

VI. ACKNOWLEDGEMENT

This work was supported by the National Research Fund, Luxembourg under AFR grant for Ph.D. project (Reference 10064089) on Reliable Communication Techniques for Future Generation Satellite Systems. The authors thank Dr. Nicolò Mazzali for useful comments.

VII. REFERENCES

[1] M. Bousquet G. Maral and Z. Sun (Contributing Editor), *Satellite Communications Systems: Systems, Techniques and Technology, 5th Edition*, Wiley Eastern, 2009.

[2] P. Angeletti, F. Coromina, F. Deborgies, R. De Gaudenzi, A. Ginesi, and A. Vernucci, "Satcoms 2020 R&D challenges: Part I: Broadband fixed communications," in *Proc. 27th AIAA International Communications Satellite Systems Conference*, 2009.

[3] D. Gesbert, M. Kountouris, R. W. Heath, C.-B Chae, and T. Sälzer, "Shifting the MIMO paradigm," *IEEE Signal Processing Magazine*, vol. 24, no. 5, Sept 2007.

[4] G. Zheng, S. Chatzinotas, and B. Ottersten, "Generic optimization of linear precoding in multibeam satellite systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2308 2320, Jun 2012.

[5] "DVB Document A 171-2: Digital video broadcasting (DVB); implementation guidelines for the second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications; Part 2 : DVB-S2 extensions DVB-S2x," March 2015.

[6] D. Christopoulos, P.-D. Arapoglou, S. Chatzinotas, and B. Ottersten, "Linear precoding in multibeam satcoms: Practical constraints," in *Proc. 31st AIAA International Communications Satellite Systems Conference (ICSSC), Florence*, Oct 2013.

[7] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Multicast multigroup precoding and user scheduling for frame-based satellite communications," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4695 – 4707, Sep 2015.

[8] G. Taricco, "Linear precoding methods for multi-beam broadband satellite systems," in *Proc. 20th European Wireless 2014*, May 2014.

[9] M. A. Vazquez, Ana Perez-Neira, D. Christopoulos, S. Chatzinotas, B. Ottersten, P.-D Arapoglou, A. Ginesi, and G. Taricco, "Precoding in multibeam satellite communications: Present and future challenges," *IEEE Wireless Communication Magazine, Submitted*, June 2015.

[10] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. COM-29, no. 11, pp. 1715–1720, Nov. 1981.

[11] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159 – 165, Jan. 2004.

[12] C. Eun and E. J. Powers, "A new Volterra predistorter based on the indirect learning architecture," *IEEE Trans. on Signal Processing,*, vol. 45, no. 1, pp. 223 –227, Jan 1997.

[13] B. Ai, Z.-X. Yang, C.-Y. Pan, S.-G. Tang, and T.-T. Zhang, "Analysis on LUT based predistortion method for HPA with memory," *IEEE Trans. on Broadcasting,*, vol. 53, no. 1, pp. 127–131, March 2007.

[14] R. Piazza, M. R. Bhavani Shankar, and B. Ottersten, "Generalized direct predistortion with adaptive crest factor reduction control," in *Proc. 40th International Conference on Acoustics, Speech and Signal Processing ICASSP*, Brisbane, Australia, April 2015.

[15] D. Spano, D. Christopoulos, S. Andrenacci, S. Chatzinotas, B. Ottersten, and J. Krause, "Total degradation analysis of precoded signals onto non-linear satellite channels," in *Proc. 21st Ka Conference, Bologna, Italy,*, Oct 2015, [Online]: http://hdl.handle.net/10993/21955.

[16] Y. Cheng and M. Pesavento, "Predistortion and precoding vector assignment in codebook-based downlink beamforming," in *2013 IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2013.

[17] M. Alvarez-Diaz, C. Mosquera, M. Neri, and G. Corazza, "Joint precoding and predistortion techniques for satellite telecommunication systems," in *2nd International Symposium on Wireless Communication Systems*, Sept 2005.

[18] S.A. Bassam, M. Helaoui, and F.M. Ghannouchi, "2-D digital predistortion (2-D-DPD) architecture for concurrent dual-band transmitters," *Microw. Theory and Techniques, IEEE Trans. on*, vol. 59, no. 10, pp. 2547 –2553, Oct. 2011.

[19] Martin Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, John Wiley & Sons, Apr. 1980.

[20] L. Shuangzhe and T Gotz, "Hadamard, khatri-rao, kronecker and other matrix products," *International J. Information and Systems Sciences*, vol. 4, pp. 160–177, 2008.

[21] B. F. Beidas, R. I. Seshadri, and N. Becker, "Multicarrier successive predistortion for nonlinear satellite systems," *IEEE Trans. on Commun.,*, vol. 63, pp. 1373–1382, 2014.

[22] R. Piazza, B. Shankar, and B. Ottersten, "Data predistortion for multicarrier satellite channels based on direct learning," *Signal Processing, IEEE Transactions on*, vol. 62, no. 22, pp. 5868 – 5880, Oct 2014.

[23] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *Signal Processing, IEEE Transactions on*, vol. 55, no. 6, pp. 2740 – 2746, June 2007.