Geometrical Pruning of the First Order Regular Perturbation Kernels of the Manakov Equation

Astrid Barreiro, *Student Member, IEEE,* Gabriele Liga, *Member, IEEE,* and Alex Alvarado, *Senior Member, IEEE*

***Abstract*—We propose an approach for constraining the set of nonlinear coefficients of the conventional first-order regular perturbation (FRP) model of the Manakov Equation. We identify the largest contributions in the FRP model and provide geometrical insights into the distribution of their magnitudes in a three-dimensional space. As a result, a multi-plane hyperbolic constraint is introduced. A closed-form upper bound on the constrained set of nonlinear coefficients is given. We also report on the performance characterization of the FRP with multi-plane hyperbolic constraint and show that it reduces the overall complexity with minimal penalties in accuracy. For a 120 km standard single-mode fiber transmission, at 60 Gbaud with DP-16QAM, a complexity reduction of 93% is achieved with a performance penalty below 0.1 dB.**

***Index Terms*—Channel modeling, perturbation-based models, fiber nonlinearities.**

I. Introduction

EFFECTIVE channel models are fundamental in the design and performance assessment of fibre-optic transmission systems [1]. In coherent systems utilizing dual polarisation signals, most analytical models are underpinned by the Manakov equation (ME) [2]. Since the ME does not have a closed-form solution, it is conventionally approximated using diverse techniques [1, Ch. 9]. A widely-used method, particularly for systems operating in the linear and pseudo-linear regimes, is the first-order regular perturbation (FRP) [3]. While the FRP method provides a reliable and accurate approximation to the ME, it requires the computation of nonlinear perturbation coefficients, also called *kernels*. This computation is cumbersome for transmission scenarios involving long distances or large signal bandwidths, for which a large number of kernels is required [1], [4]–[6].

In the context of the design of nonlinearity compensation/mitigation algorithms, several works in the literature have targeted a reduction of the FRP computational complexity. In [7], [8], the authors designed the pulse shape to simplify the computation of kernels. Other approaches, e.g., [9], [10], reduce the number of kernels by considering the temporal phase-matching symmetry. This avoids computing coefficients with zero contribution due to the isotropic phase distribution of the transmitted symbols [11, Sec. VIII]. The authors in [12]–[14] propose quantizing the kernels to substantially reduce the number of kernels used to compute the FRP model. More recently, a growing number of data-driven approaches have emerged to optimize the FRP coefficients [15]–[19]. However, to the best of our knowledge, very little to no attention has been given to the selection of subsets based on the geometric properties of the magnitudes of the kernels.

In this paper, we present a novel approach to constrain the set of FRP kernels based on the geometric characteristics of the spatial distribution of their magnitudes. Unlike [9], [13], [20], [21], our approach does not require precomputation of large sets of kernels nor magnitude-based pruning. Our method introduces a systematic way to select subsets of large kernels that lead us to the definition of a *multi-plane hyperbolic constraint*. We show that using the multi-plane hyperbolic constraint we can significantly reduce the computational complexity, without sacrificing the model's accuracy.

This paper is organized as follows. In Sec. II we introduce the transmission system model and brief key aspects of the time-domain FRP, followed by an identification of the geometric characteristics of the magnitudes of the kernels. In Sec. III we formalize the multi-plane hyperbolic constraint and report an upper bound for the cardinality of the subset of kernels after the multi-plane hyperbolic constraint is applied. In Sec. IV, we evaluate the performance of this constraint for a 120 km standard single-mode fiber transmission, operating at 60 Gbaud with DP-16QAM, and Sec. V is devoted to conclusions.

II. Preliminaries

A. System Model and Characteristics of FRP

In this work, we consider a single-channel, single-span fiber-optic transmission system using a polarisation-multiplexed signal, as shown in Fig. 1. The transmitter generates a sequence of two-dimensional complex symbols with unit energy per complex dimension $\underline{\boldsymbol{a}} = \ldots, \boldsymbol{a}_{n-1}, \boldsymbol{a}_n, \boldsymbol{a}_{n+1}, \ldots$, with n denoting a discrete time instance[1]. This sequence is linearly modulated (MOD) and energy-scaled by E_s to generate the transmitted signal $\boldsymbol{Q}(t,0)$. $E_s = P/2R_s$ is the average energy per transmitted symbol, where P is the launch power, $R_s = 1/T$ is the symbol rate, and T the is the symbol duration. This signal is then propagated through a length-L (baseband equivalent)

A. Barreiro, G. Liga, and A. Alvarado are with the Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands. e-mail: a.barreiro.berrio@tue.nl.

The work of A. Barreiro and A. Alvarado has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 757791).

[1]*Notation convention:* Bold letters denote two-dimensional complex symbols, underlined bold letters denote sequences of symbols, and sets are denoted by calligraphic letters. $(t_1, t_2, \ldots, t_n)$ denote n-tuples and if not stated, the vector $x = [x_1, x_2, \ldots, x_n]^\top$ of dimension n is a column vector. $(\cdot)^\top$ denotes transposition, $(\cdot)^\dagger$ denotes Hermitian transposition, and $|\cdot|$ denotes absolute value when the argument is a scalar and cardinality when it is a set.

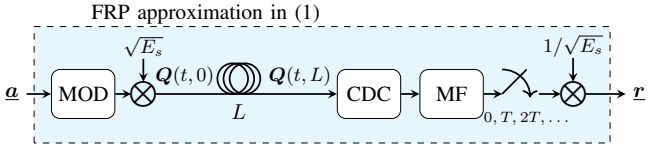

Fig. 1. System model under consideration in this work. The transmitter uses linear modulation (MOD). The receiver applies ideal chromatic dispersion compensation (CDC), matched filtering (MF), and sampling. The coefficient $\sqrt{E_s}$ tunes the launch power P. The optical channel considers a fiber of length L, modeled by ME.

optical channel modeled by the attenuation-normalized ME [2, eq. (3.1.3)]. At the receiver side, ideal chromatic dispersion compensation is performed, followed by a matched filter and sampling at the symbol rate, yielding a sequence of received 2D complex symbols $\underline{\boldsymbol{r}} = \ldots, \boldsymbol{r}_{n-1}, \boldsymbol{r}_n, \boldsymbol{r}_{n+1}, \ldots$

For small enough values of the input signal power, FRP approximates the ME yielding the following input-output relation at time instant 0 [22, Eqs. (5) and (6)]

$$\boldsymbol{y}_0 \approx \boldsymbol{a}_0 + \jmath\frac{8}{9}\gamma E_s \sum_{(k,l,m)\in\mathbb{Z}^3} \left(\boldsymbol{a}_k^\dagger \boldsymbol{a}_l\right) \boldsymbol{a}_m \; S_{klm}. \tag{1}$$

In (1), $\jmath$ is the imaginary unit, γ is the fiber nonlinear coefficient, and $\mathbb{Z}$ is the set of integers. The kernels S_{klm} in (1) are complex-valued coefficients modeling self-phase modulation and are defined as [23, eq. (4)]

$$S_{klm} \triangleq \int_0^L e^{(-\alpha z)} \int_{-\infty}^{\infty} h^*(z,t)h^*(z,t-kT) \; h(z,t-lT)h(z,t-mT)\,\mathrm{d}t\mathrm{d}z, \tag{2}$$

where with a slight abuse of notation, $h(z,t)$ is used to denote the solution of the ME when $\gamma = 0$ and $\boldsymbol{Q}(t,0) = h(t)$.

B. Finite-memory FRP

The mathematical expression in (1) implies that an infinite number of kernels must be numerically computed. In practice, however, pulses are always time-limited, and thus, the triple summation is typically truncated. A typical way to do so is

$$\boldsymbol{y}_0 \approx \boldsymbol{a}_0 + \jmath\frac{8}{9}\gamma E_s \sum_{(k,l,m)\in\mathcal{M}} \left(\boldsymbol{a}_k^\dagger \boldsymbol{a}_l\right) \boldsymbol{a}_m \; S_{klm}, \tag{3}$$

where

$$\mathcal{M} \triangleq \{(k,l,m) \in \mathbb{Z}^3 : -M \le k,l,m \le M\}. \tag{4}$$

The mathematical form of the FRP model after the truncation specified in (3)–(4) yields a *finite-memory* FRP model. Accordingly, M in (4) can be understood as the model's memory length, as $(2M+1)$ symbols around $\boldsymbol{a}_0$ are needed to compute the model's output (3). Henceforth, we refer to finite-memory FRP in (3)–(4) as simply FRP. Although this truncation to the domain of FRP is the common approach to tackle the task of computing infinite kernels in (1), the computation of $\mathcal{M}$ in (4) remains challenging for large M, given that $|\mathcal{M}| = (2M+1)^3$.

TABLE I
FIBER AND PULSE SHAPE PARAMETERS

Nonlinear parameter γ	$1.2\ \mathrm{W}^{-1}\mathrm{km}^{-1}$
Fiber attenuation α	0.2 dB/km
Group velocity dispersion β_2	$-21.7\ \mathrm{ps}^2/\mathrm{km}$
Fiber length L	120 km
Symbol rate R_s	60 Gbaud
Pulse shape $h(t)$	Root-raised-cosine (RRC)
RRC roll-off factor	0.01

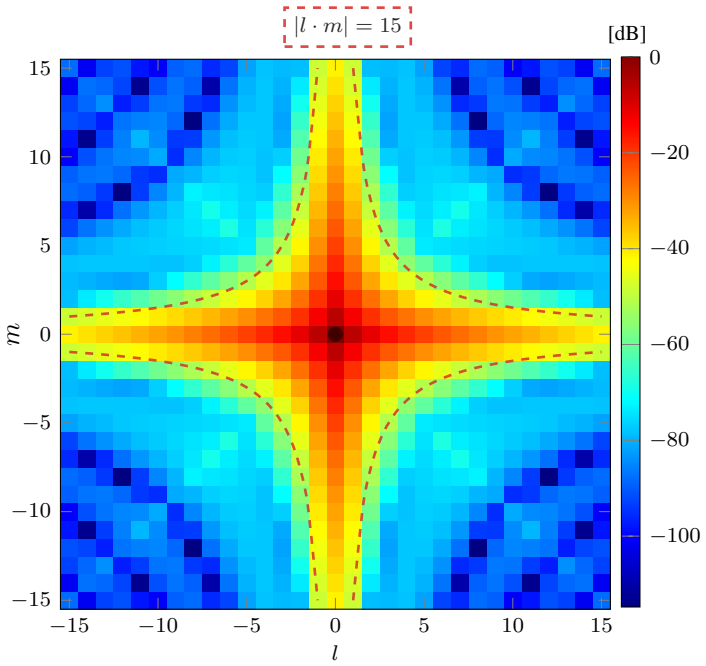

Fig. 2. Heatmap of $10\log\{|S_{klm}|/|S_{000}|\}$ for the plane $k-l-m=0$ when $-15 \le l,m \le 15$. The hyperbolic star defined by $|l\cdot m| = 15$ is shown with dashed red lines.

C. Geometric Characteristics of the Kernels

Let's consider the collection of magnitudes of the kernels, for which $|S_{000}|$ is the largest magnitude. To depict the spatial distribution of the magnitudes of the kernels in a three-dimensional space, we map $|S_{klm}|$ to the (k,l,m) coordinate in $\mathbb{Z}^3$. $|S_{000}|$ is then located in the origin of the coordinate system. From this point forward, the geometric characteristics of the kernels are analyzed by taking this depiction for the subset $\mathcal{M}$ as a frame of reference.

The first aspect to highlight from the spatial distribution of the magnitudes of the kernels is that the kernels with large magnitude exhibit a hyperbolic decay relative to $|S_{000}|$. As previously reported in [9], [13], [18], [24], the plane $k-l-m=0$ contains the largest kernels within $\mathcal{M}$. This is an argument frequently utilized to reduce complexity in perturbation-based methods for nonlinearity compensation. To highlight the characteristics of the plane $k-l-m=0$, we display in Fig. 2 a heatmap of $10\log\{|S_{klm}|/|S_{000}|\}$ dB for the system parameters specified in Table I. The heatmaps are shown for $-15 \le l,m \le 15$.

Fig. 2 shows that the largest kernels in the plane $k-l-m=0$ fall within a *hyperbolic star* domain. This observation confirms the results in [9, Fig. 5], [20, Fig. 3] and [15, Fig. 6]. In

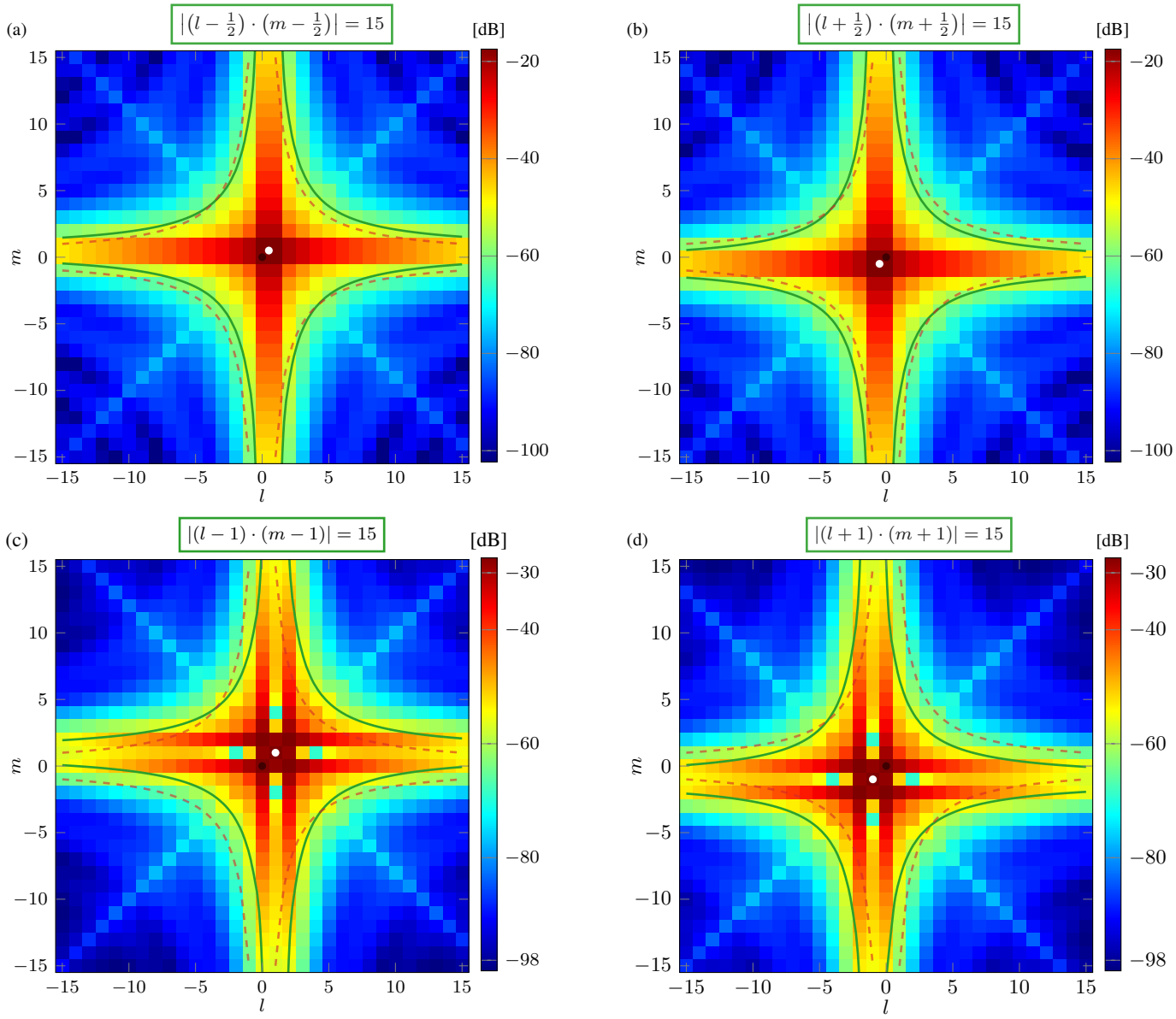

Fig. 3. Heatmap of $10\log\{|S_{klm}|/|S_{000}|\}$ for the planes $k-l-m=j$ when $-15\leq l,m\leq 15$ for (a) $j=1$, (b) $j=-1$, (c) $j=-2$, and (d) $j=2$. The hyperbolic stars given by $|(l-\frac{j}{2})\cdot(m-\frac{j}{2})|\leq 15$ are drawn in green continuous lines, while the hyperbolic star $|l\cdot m|\leq 15$ is highlighted in red. The center of the star $(0,0)$ is emphasized with a black circle.

Figure 2, we highlight the hyperbolic star domain containing a subset of magnitudes above -40 dB. The analytical expression for this hyperbolic star is $|l\cdot m| = M$. The plot of this expression is depicted by dashed red curves in the same figure, with its center located at $(0,0)$, which is emphasized by a black circle.

The second important aspect is that large kernels in the 3D depiction align along parallel diagonal planes characterized by the normal vector $[1,-1,-1]^\top$. These planes, which we call *principal planes*, can be expressed mathematically as $[1,-1,-1]\cdot[k,l,m]^\top = j$, with $j\in\mathbb{Z}$. To illustrate the key geometric characteristics of these principal planes for $j\neq 0$, we present the heatmaps of $10\log\{|S_{klm}|/|S_{000}|\}$ dB for $j=-2,-1,1,2$ in Fig. 3 (a)-(d).

For all the cases displayed in Fig. 3, there is a reduction of the magnitudes of the kernels with respect to $j=0$ for increasing j. Note that the initial value of the scale bar changes from 0 dB for $j=0$ in Fig. 2, to -20 dB and -30 dB for $j=-1,1$ and $j=-2,2$ in Fig. 3, respectively. Additionally, the largest magnitudes in the principal planes for $j\neq 0$ also fall within hyperbolic star domains, for which the centers are highlighted by white circles. As shown in Fig. 3, the white circles are displaced and the hyperbolas are no longer centered at the origin (black circle). Also, notice that pairs j and $-j$ are equal in magnitude but their centers are shifted in opposite directions. The general mathematical expression for the hyperbolic star for any principal plane j is given by $|(l-\frac{j}{2})\cdot(m-\frac{j}{2})| = \hat{M}$. After drawing these hyperbolas in green for $\hat{M}=M$, Fig. 3 (a)-(d) shows that large magnitudes are within the boundaries set by $|(l-\frac{j}{2})\cdot(m-\frac{j}{2})| = M$. In

the following section, the properties identified in Sec. II-C are used to define geometric-based constraints on $\mathcal{M}$.

III. Multi-plane Hyperbolic Constraint FRP

In this section, following the results in Sec. II-C, we propose a geometry-based constraint over $\mathcal{M}$ in (4) that we call *multi-plane hyperbolic constraint*. This constraint effectively decreases the complexity of the FRP model with negligible accuracy.

The analysis in Sec. II-C (see Fig. 2) shows that the principal plane $j = 0$ is the one containing a large fraction of the most significant kernels and that they are located within a hyperbolic star centered in the origin. Furthermore, the analysis shows that principal planes with $j \neq 0$ also contain significant contributions and that the largest are located inside *shifted* hyperbolic stars (see Fig. 3). These observations lead us to define the subset $\mathcal{H}_i(\hat{M}) \subset \mathcal{M}$, where $i \in \mathbb{Z}^+$ will indicate that $2i + 1$ principal planes are included, and $\hat{M}$ determines the size of the hyperbolic star, as follows

$$\begin{aligned}\mathcal{H}_i(\hat{M}) \triangleq \big\{(k,l,m) \in \mathcal{M} : k-l-m = j \ \wedge \\ |(l - j/2)\cdot(m - j/2)| \leq \hat{M}, \\ j \in \mathbb{Z} \ \wedge \ j = -i, -i+1, ..., i\big\}.\end{aligned} \tag{5}$$

Taking $\hat{M} = M^2$ in (5) removes the hyperbolic constraint and taking $i = 0$ collapses the set to a single principal plane. Observe that the set $\mathcal{H}_0(M^2)$ is the case used by the authors of [13], [18]. Other authors have considered a variant of $\mathcal{H}_0(\hat{M})$, for which $\hat{M}$ is selected such that kernels with magnitudes below a user-defined threshold are pruned [9], [20]. Both cases, $\mathcal{H}_0(M^2)$ and $\mathcal{H}_0(\hat{M})$ with magnitude threshold, have been extensively explored as a strategy to filter out negligible contributions to the FRP model.

Considering $i = 0$ in (5) and changing the value of $\hat{M}$, allows us to capture the kernels within *one* hyperbolic star. When (5) is used with nonzero values of i, other principal planes are included (see Fig. 2), while also constraining them via *shifted* hyperbolic stars. Our proposal in (5) considers for the first time principal planes additional to $j = 0$ but also introduces a second tunable parameter ($\hat{M}$). Together, i and $\hat{M}$, set the cardinality of the set $\mathcal{H}_i(\hat{M})$, and adjusting them directly impacts the performance of FRP, as we will see in the next section.

The cardinality of $\mathcal{H}_i(\hat{M})$ measures the complexity of the channel model in (3) when the summation is constrained to $\mathcal{H}_i(\hat{M})$ instead of $\mathcal{M}$. By definition, when $i = 2M$ and $\hat{M} = M^2$, $|\mathcal{H}_{3M}(M^2)| = |\mathcal{M}| = (2M+1)^3$, since that is the case where the hyperbolic constraint degenerates in the entire plane and that the same time includes all the principal planes that $\mathcal{M}$ can accommodate. The most interesting cases are when i is small and $\hat{M} = M$. The assumption on $\hat{M} = M$ will become clear in the next section.

The following proposition provides an upper bound on the number of elements of the proposed set in (5).

Proposition 1. *The cardinality of $\mathcal{H}_i(M)$ is upper bounded as*

$$|\mathcal{H}_i(M)| \leq (2i+1)\left(4\sum_{k=1}^{M}\left\lfloor\frac{M}{k}\right\rfloor + 4(M-1) + 1\right) \tag{6}$$

Proof: Computing $|\mathcal{H}_0(M)|$ requires: (i) solving a discrete integral to count all the elements under half a hyperbola excluding the asymptote $k = 0$, and (ii) summing all the elements lying on $k = 0$ and $l = 0$.

The result for (i) is

$$C_1 = \sum_{k=1}^{M}\left\lfloor\frac{M}{k}\right\rfloor, \tag{7}$$

and (ii) is

$$C_2 = 4M + 1. \tag{8}$$

Adding up (7) and (8) and subtracting the contributions (l, m) in the hyperbolic stars outside of the plane $k - l - m = 0$, results in $|\mathcal{H}_0(M)| = 4C_1 + C_2 - 4$. This result accounts for a full star. Multiplying the number of elements in a full star by the number of principal planes considered, i.e., $(2i + 1)$ when $i \neq 0$, sets an upper bound on $|\mathcal{H}_i(M)|$. This result is an upper bound because the neighboring planes to $i = 0$ do not contain full but cut stars. See for an example Fig. 3 (a). The hyperbolic stars are cut by the boundary conditions $k - l - m = j$ and $-M \leq k, l, m \leq M$. ∎

An important advantage of geometrically constraining the set of FRP kernels is that it does not require precomputation of large sets of kernels nor magnitude-based pruning. Traditionally, most approaches in the literature first compute a set of kernels with large cardinality and then perform a magnitude-based pruning on it. On the other hand, our approach points at the indices (k, l, m) needed to compute the FRP model in the definition of $\mathcal{H}_i(\hat{M})$. Only those kernels are to be calculated using the integral from (2). This contributes to significantly reducing the overall complexity of the computation of the kernels, as well as the model as we will see in the next section. It is important to note that the reduction in the cardinality of the set is relative to $\mathcal{M}$ for a fixed M.

In the following section, we present an assessment of the performance and complexity of the multi-plane hyperbolic approach defined in (5). We establish a comparison with the benchmark set by FRP as specified in (3)–(4).

IV. Numerical Results

Similar to [21], we introduce the nonlinear signal-to-noise ratio (NSNR) as a performance metric. NSNR is computed as

$$\text{NSNR} \triangleq \frac{P}{P_{\text{NLI}}} = \frac{\mathbb{E}[|\boldsymbol{A}|^2]}{\mathbb{E}[|\boldsymbol{Y} - \boldsymbol{A}|^2]}, \tag{9}$$

where P is the transmit power and P_{NLI} is the nonlinear interference power, $\boldsymbol{A}$ and $\boldsymbol{Y}$ are the random variables with samples $\boldsymbol{a}_0$ and $\boldsymbol{y}_0$, respectively, and $\mathbb{E}[\cdot]$ denotes the expectation operator. In this work, the performance of FRP is assessed using the gap to the split-step Fourier method (SSFM). For simplicity, the NSNR gap to SSFM is henceforth denoted as Δ.

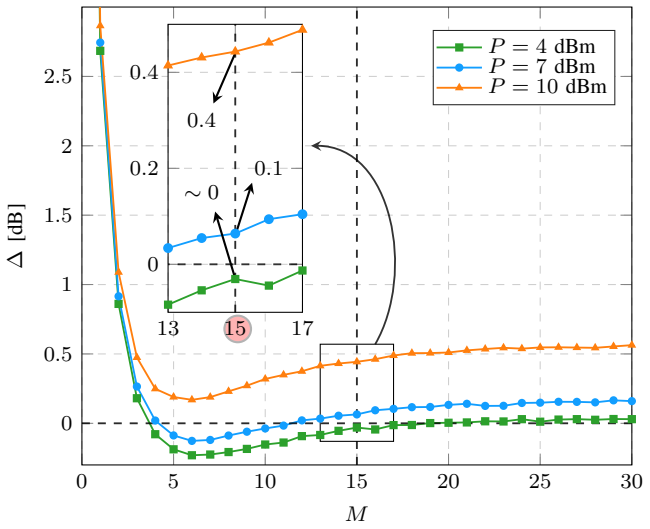

Fig. 4. Δ against M for the set $\mathcal{M}$ in (4). $P = 7$ dBm is the optimum launch power for $L = 120$ km standard single-mode fiber transmission, operating at 60 Gbaud using DP-16QAM when amplified spontaneous emission (ASE) noise is not taken into account, and a 5 dB noise figure is assumed.

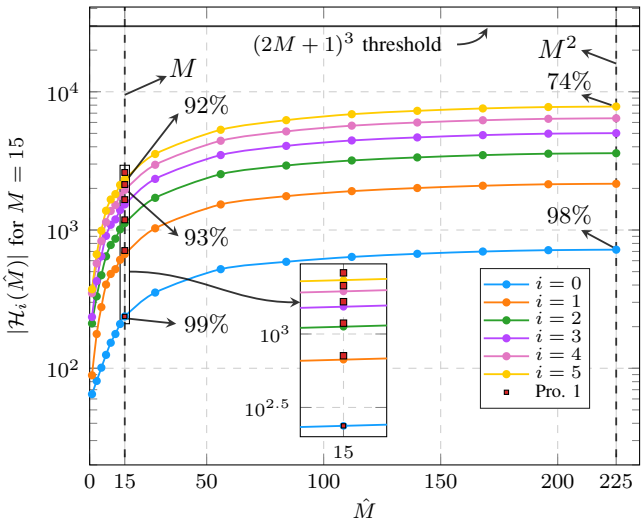

Fig. 5. Cardinality of $\mathcal{H}_i(\hat{M})$ as a function of $\hat{M}$ for $i = [0, 5]$. The definition of $\mathcal{H}_i(\hat{M})$ generates a reduction in the cardinality w.r.t. $\mathcal{M}$. This reduction is indicated for $i = 0$ and $i = 5$ with $\hat{M} = M$ and $\hat{M} = M^2$. For $\hat{M} = 15$, the upper bound in Proposition 1 is highlighted with black squares filled with red.

The results of Δ presented in this section are for the system with parameters specified in Table I. The kernels used to compute the FRP model, constrained by either $\mathcal{M}$ in (4) or $\mathcal{H}_i(\hat{M})$ in (5), are numerically evaluated using (2). Fig. 4 shows Δ as a function of M for $P = 4, 7$ and 10 dBm for the set $\mathcal{M}$. The reference value $P = 7$ dBm is the optimum launch power when amplified spontaneous emission (ASE) noise is not taken into account. In such a case, a 5 dB noise figure is assumed. As shown in [25, Fig. 2] for powers above $P = 10$ dBm, the curves for FRP and SSFM begin to diverge. Fig. 4 shows that for $M > 3$, the NSNR prediction of FRP matches SSFM within a 0.5 dB gap for all cases. For increasing M, all Δ curves show saturation preceded by a small deflection. Henceforth, $\mathcal{M}$ with $M = 15$ is taken as the reference to be constrained to determine the set $\mathcal{H}_i(\hat{M})$. The size $M = 15$ guarantees gaps < 0.5 dB for all powers up to 10 dBm. See the inset of Fig. 4.

Fig. 5 shows the cardinality of $\mathcal{H}_i(\hat{M})$ as a function of $\hat{M}$ for $i = 0, 1, \ldots, 5$. The threshold $|\mathcal{M}| = (2M + 1)^3$ is indicated as a reference. As expected, all cases $(i, \hat{M})$ satisfy $|\mathcal{H}_i(\hat{M})| < |\mathcal{M}|$. Fig. 5 also shows that $\mathcal{H}_0(M)$ and $\mathcal{H}_5(M)$ yield reductions in the cardinality of 99% and 92%, respectively, while $\mathcal{H}_0(M^2)$ and $\mathcal{H}_5(M^2)$ yield 98% and 74%. In Fig. 5 we indicate with red-filled black squares the upper bound on $|\mathcal{H}_i(M)|$ using (6) in Proposition 1. Observe in the inset of Fig. 5, that for $i > 0$, the bound in (6) is not tight.

In Fig. 6 (a) we display the result of Δ against launch power considering the sets $\mathcal{H}_i(M)$ and $\mathcal{H}_i(M^2)$ for $i = 0, 2, 5$. Fig. 6 (a) shows that $i = 0$ is, as expected, the worst-performing case. Increasing i yields improvement for all powers considered. Notice that the differences between the $\{M, M^2\}$ pairs are relatively small (< 0.2 dB for all i). This suggests that kernels outside the hyperbolic star domain of size $\hat{M} = M$, have minimal contributions to (3). This justifies why in Proposition 1 $\hat{M} = M$ is chosen, a case that according to the result shown in Fig. 5, yields a major reduction in the cardinality of the set. Fig. 6 (a) also shows that for $\hat{M} = M^2$ larger values of Δ are observed among different powers compared to $\hat{M} = M$. This can be justified by the fact that removing kernels from the reference set $\mathcal{M}$ induces a displacement of the performance in Δ towards the deflection observed in Fig. 4. Once in the deflected region, it is possible to fall into scenarios where the set with M^2 yields a slightly larger gap than M due to the concavity of the region. Although in Fig. 6 (a) $i = 5$ is the best case we considered, the small gap between $i = 5$ and $i = 2$ suggests that adding principal planes beyond $i = 5$ might not lead to major gains.

Fig. 6 (b) shows Δ against $\hat{M}$ for $P = 7$ dBm, where it is observed that all $\mathcal{H}_i(\hat{M})$ saturate over the set $\mathcal{M}$. Similar to Fig. 6 (a), Fig. 6 (b) shows that increasing i moves the saturation level towards zero, supporting the already highlighted advantage of adding principal planes to $\mathcal{H}_0(\hat{M})$. The same figure also shows that for all $i > 1$, Δ is within 0.5 dB to $\mathcal{M}$. Comparing this result to $\mathcal{H}_0(\hat{M})$ shows that our proposal outperforms the conventional approach found in the literature, i.e., a single plane. Inspecting Fig. 6 (b) for $\hat{M} > M$ reveals a small decrease in performance: the gap to $\mathcal{M}$ opens up as $\hat{M}$ exceeds M. Such a trend confirms that considering $\hat{M}$ larger than M does not necessarily guarantee a better performance. In addition, we observe in the $\hat{M} > M$ region that increasing i does close the gap to $\mathcal{M}$, while in the $\hat{M} < M$ region it does not. This behavior indicates that for $M = 15$ a set $\mathcal{H}_i(\hat{M})$ with $\hat{M} < M$ will not match the performance of another set with $\hat{M} > M$ by simply increasing i. For the case $\hat{M} = M$ it is possible to improve performance by increasing i and it does not have the inconclusiveness previously mentioned for the $\hat{M} < M$ region. This is an additional characteristic of the case $\hat{M} = M$ that motivated us to choose it as a case of interest (see Proposition 1).

Finally, the inset in Fig. 6 (b), shows that $\mathcal{H}_4(M)$ is the

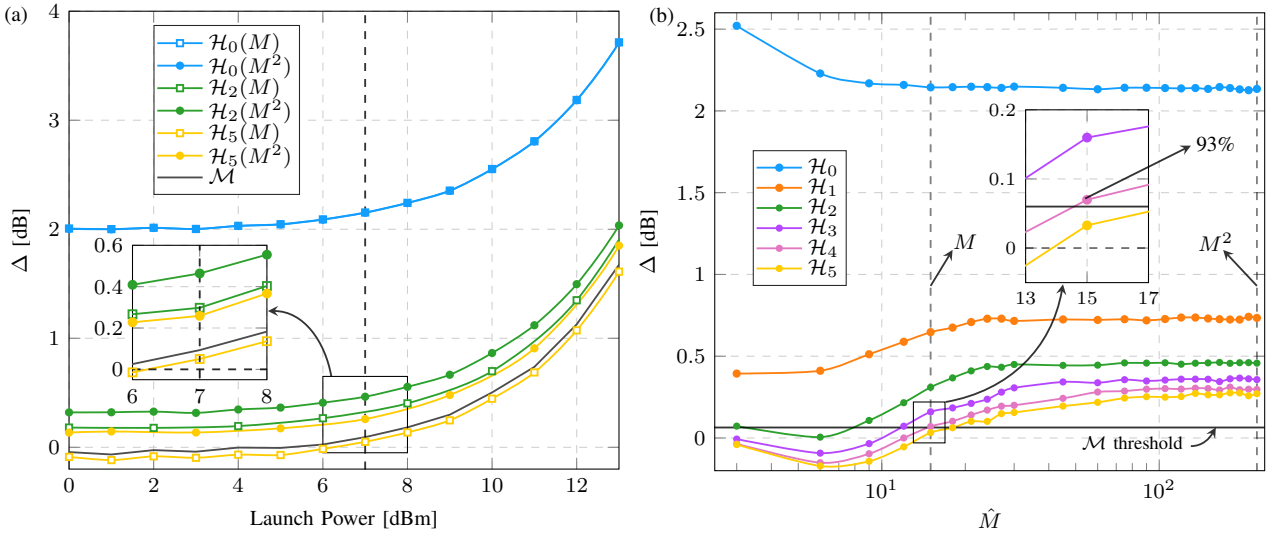

Fig. 6. (a) Δ vs launch power for the sets $\mathcal{H}_0(\hat{M})$, $\mathcal{H}_2(\hat{M})$ and $\mathcal{H}_5(\hat{M})$ for $\hat{M} = M$ and $\hat{M} = M^2$. (b) Δ vs M for the set $\mathcal{H}_i(\hat{M})$ at $P = 7$ dBm for $i = [0, 5]$. In (b) $\mathcal{H}_4(M)$ offers a 93% decrease in complexity with a penalty to the model's accuracy below 0.1 dB,

scenario that yields the closest performance to $\mathcal{M}$. Besides having minimal penalties to the model's accuracy (0.07 dB), $\mathcal{H}_4(M)$ offers a 93% decrease in complexity. Therefore, for the system model under consideration and $P = 7$ dBm, 9 principal planes containing a hyperbolic star of size M, are sufficient to guarantee an equivalent performance to $\mathcal{M}$ with low complexity. Given that Fig. 6 (a) shows a comparable behavior for all the powers considered, we conjecture that the good properties of $\mathcal{H}_4(M)$ extend over other powers with minor penalties in performance.

V. Conclusions

We introduced a geometric constraint to the set of kernels in the first-order perturbative solution to the Manakov equation. The most relevant kernels can be identified via a multi-plane hyperbolic restriction for a given maximum number of kernels. Using the kernels within this set leads to significant reductions in complexity with minimal performance penalties. Future work includes developing perturbation-based nonlinearity compensation/mitigation algorithms considering multi-plane hyperbolic constraints. Furthermore, due to the equivalence between the FRP coefficients and the Volterra kernels, our proposal also has the potential to be applied in the design of low-complexity Volterra equalizers.

References

[1] A. Bononi, R. Dar, M. Secondini, P. Serena, and P. Poggiolini, *Fiber nonlinearity and optical system performance*. Springer, Cham, 2020.

[2] G. Agrawal, *Non-linear fiber optics*, 5th ed. Academic Press, New York, 2012.

[3] A. Vannucci, P. Serena, and A. Bononi, "The RP method: A new tool for the iterative solution of the linear Schrödinger equation," *Journal of Lightwave Technology*, vol. 20, no. 7, pp. 1102–1112, July 2002.

[4] X. Chen and W. Shieh, "Closed-form expressions for nonlinear transmission performance of densely spaced coherent optical OFDM systems," *Opt. Express*, vol. 18, no. 18, pp. 19 039–19 054, Aug. 2010.

[5] P. Poggiolini, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photonics Technology Letters*, vol. 23, no. 11, pp. 742–744, June 2011.

[6] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Opt. Express*, vol. 22, no. 13, pp. 16 335–16 362, June 2014.

[7] A. Ghazisaeidi and R. J. Essiambre, "Calculation of coefficients of perturbative nonlinear pre-compensation for Nyquist pulses," *European Conf. on Optical Communication (ECOC)*, Canes, France, Sep. 2014.

[8] F. Frey, R. Emmerich, C. Schubert, J. K. Fischer, and R. F. Fischer, "Improved perturbation-based fiber nonlinearity compensation," *European Conf. on Optical Communication (ECOC)*, Rome, Italy, Sep. 2018.

[9] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *Journal of Lightwave Technology*, vol. 29, no. 17, pp. 2570–2576, Sep. 2011.

[10] R. Dar and P. J. Winzer, "Nonlinear interference mitigation: Methods and potential gain," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 903–930, Feb. 2017.

[11] A. Mecozzi and R. J. Essiambre, "Nonlinear Shannon limit in pseudolinear coherent systems," *Journal of Lightwave Technology*, vol. 30, no. 12, pp. 2011–2024, June 2012.

[12] W. R. Peng, Z. Li, F. Zhu, and Y. Bai, "Training-based determination of perturbation coefficients for fiber nonlinearity mitigation," *Optical Fiber Communication Conference (OFC)*, Los Angeles, California, Mar. 2015.

[13] M. Malekiha, I. Tselniker, and D. V. Plant, "Efficient nonlinear equalizer for intra-channel nonlinearity compensation for next generation agile and dynamically reconfigurable optical networks," *Opt. Express*, vol. 24, no. 4, pp. 4097–4108, Feb. 2016.

[14] M. Sorokina, S. Sygletos, and S. Turitsyn, "Sparse identification for nonlinear optical communication systems," *International Conference on Transparent Optical Networks (ICTON)*, Girona, Spain, July 2017.

[15] S. Zhang, F. Yaman, K. Nakamura, T. Inoue, V. Kamalov, L. Jovanovski, V. Vusirikala, E. Mateo, Y. Inada, and T. Wang, "Field and lab experimental demonstration of nonlinear impairment compensation using neural networks," *Nature Communications*, vol. 10, no. 1, pp. 1–8, July 2019.

[16] Y. Gao, Z. A. El-Sahn, A. Awadalla, D. Yao, H. Sun, P. Mertz, and K.-T. Wu, "Reduced complexity nonlinearity compensation via principal component analysis and deep neural networks," *Optical Fiber Communication Conference (OFC)*, San Diego, California, Mar. 2019.

[17] M. M. Melek and D. Yevick, "Nonlinearity mitigation with a perturbation based neural network receiver," *Optical and Quantum Electronics*, vol. 52, no. 10, pp. 1–10, Oct. 2020.

[18] A. Redyuk, E. Averyanov, O. Sidelnikov, M. Fedoruk, and S. Turitsyn, "Compensation of nonlinear impairments using inverse perturbation theory with reduced complexity," *Journal of Lightwave Technology*, vol. 38, no. 6, pp. 1250–1257, Feb. 2020.

[19] A. Barreiro, G. Liga, and A. Alvarado, "Data-driven enhancement of the time-domain first-order regular perturbation model," *Journal of Lightwave Technology*, vol. 41, no. 9, pp. 2691–2706, May. 2023.

[20] X. Liang and S. Kumar, "Multi-stage perturbation theory for compensating intra-channel nonlinear impairments in fiber-optic links," *Opt. Express*, vol. 22, no. 24, pp. 29 733–29 745, Dec. 2014.
[21] A. Barreiro, G. Liga, and A. Alvarado, "A phase recovery-aware algorithm for kernel estimation of the Manakov equation," *2023 Asia Communications and Photonics Conference/2023 International Photonics and Optoelectronics Meetings (ACP/POEM)*, Wuhan, China, Nov. 2023.
[22] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Properties of nonlinear noise in long, dispersion-uncompensated fiber links," *Opt. Express*, vol. 21, no. 22, pp. 25 685–25 699, Oct. 2013.
[23] ——, "Inter-channel nonlinear interference noise in wdm systems: Modeling and mitigation," *Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1044–1053, Mar. 2015.
[24] O. S. S. Kumar, A. Amari, O. A. Dobre, and R. Venkatesan, "Enhanced regular perturbation-based nonlinearity compensation technique for optical transmission systems," *IEEE Photonics Journal*, vol. 11, no. 4, pp. 1–12, June 2019.
[25] A. Barreiro, G. Liga, and A. Alvarado, "A data-driven optimization of first-order regular perturbation coefficients for fiber nonlinearities," in *2022 IEEE Photonics Conference (IPC), Vancouver, Canada*, Nov. 2022.