Nonlinear Interference Noise Impact for Scalable Optical Network Planning in Horseshoe Scenarios

Tiago Silvério[1,2,*], Mohammad Hosseini[1], Giuseppe Parisi[1], Sasipim Srivallapanondh[1], Luis Velasco[2], Antonio Napoli[1]

[1] Nokia Munich Germany

[2] Universitat Politècnica de Catalunya, Barcelona, Spain

[*] corresponding author: tiago.silverio@nokia.com

***Abstract*—We investigate the impact of Nonlinear Interference (NLI) noise on the scalability of optical networks built using horseshoe-like topologies. By varying key design parameters, such as the number of nodes, per-node capacity, and span length; this study quantifies how nonlinear effects limit achievable reach and aggregate throughput as network scale increases. The analysis provides insight into the balance between traffic growth and physical-layer constraints, highlighting design trade-offs that guide efficient dimensioning of future flexible optical infrastructures.**

***Index Terms*—Network Design, Enhanced Gaussian Noise Model, Response Surface Methodology, Nonlinear Interference, System Optimization, Software Defined Networks.**

I. Introduction

As traffic demands continue to rise, specially due to the recent AI workload exponential growth and the possible impacts on data transport, the evolution of optical networks toward increasingly flexible, cost-effective and high-capacity transmission solutions is required. In this context, Digital Subcarrier Multiplexing (DSCM) has emerged as a promising technique to efficiently exploit spectral resources while enabling fine-grained adaptability to varying link conditions [1]–[4]. However, the deployment of DSCM in horseshoe architectures introduces unique design capabilities with their respective challenges, which are not fully addressed by existing network design approaches [5], [6].

Horseshoe topologies, often encountered in access and aggregation networks, are characterized by their asymmetric structure and operational constraints, notably the limited or absent use of inline optical amplification [7], [8]. This restriction, motivated by cost, energy efficiency, or infrastructure limitations, might force system designers to rely on higher launch powers at the hub node to ensure adequate signal quality mainly across downstream links [5], [6]. As a consequence, increasing launch power beyond a threshold limit will introduce non-negligible NLI impairments, significantly impacting system performance and limiting achievable capacity. Accurate estimation of NLI in such conditions is therefore critical for network planning and optimization. In this work, we compute the aforementioned NLI by utilizing a validated and modified Monte-Carlo based Enhanced Gaussian Noise (EGN) model [7], [9], [10] and analyze the different impacts of system design variables in the overall performance penalties. While the EGN model provides a powerful framework for predicting nonlinear effects in coherent optical systems, its application to large-scale parametric studies remains computationally demanding [11].

To address this challenge, we propose a methodology that combines extensive EGN-based simulations with Response Surface Methodology (RSM) to systematically analyze the impact of key network parameters, including per-node Capacity (C), number of nodes (N), and span length (L), on the accumulated NLI power. By leveraging RSM, we extract practical design constraints and insights that capture the complex relationships between system-level variables and NLI penalties, enabling rapid performance estimation without exhaustive simulations. This approach provides a compact and practical NLI representation for network planning to tackle the trade-offs associated with increasing launch power in horseshoe topologies.

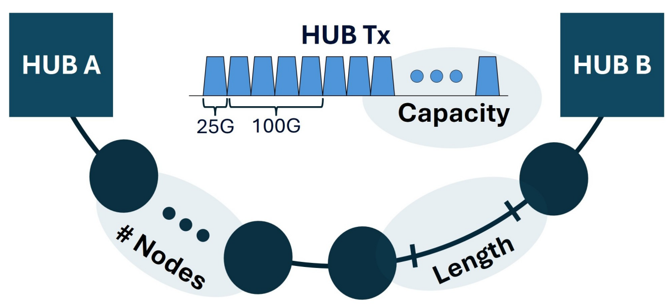

Fig. 1. Representation of the considered horseshoe architecture, highlighting variable per-node capacity (C), number of nodes (N) and span length (L).

II. Simulation Framework

This section presents the simulation framework adopted to study the impact of NLI on horseshoe-based optical networks operating with DSCM. We will, from now on, refer to the space spanned by the sweeping variables $\{C, N, L\}$ as the System Design Space (SDS). To this end, we rely on extensive EGN-based simulations and use the resulting dataset to support

The authors acknowledge the support of the EU Horizon Europe research and innovation programs, EU-funded MSCA-DN NESTOR (GA No.101119983).

979-8-3195-4420-9/26/$31.00 ©2026 IEEE

the RSM analysis, thus enabling us to identify the dominant factors governing system performance and to build practical design guidelines for network dimensioning. The simulations are built upon a set of simplifying assumptions intended to isolate the effect of the sweeping variables on nonlinear performance. Firstly, we assume equal launch powers at every span, so that each segment can be analyzed independently from the standpoint of accumulated NLI impairment. Second, the modeling is power-agnostic, in the sense that the transmitted launch power is not optimized for each individual scenario, but rather explored as a variable affecting the resulting NLI penalties. Third, the central Digital Subcarrier (DSC), or an averaged representation of the central DSCs, is used as the representative Channel Under Test (CUT), since it experiences the strongest NLI interaction and therefore defines the main system bottleneck [7], [10]. Under these assumptions, the analysis focuses on the worst-case nonlinear behavior expected in horseshoe deployments. Table I summarizes the system, sweeping, and output parameters used throughout the study. The physical-layer quantities listed in the table define the transmission scenario deployed in the EGN calculations, while the sweeping variables specify the network configurations explored in the parametric study. The output parameter, η, is used as the metric of interest to characterize the total NLI value calculated by the model and to assess the dependence of NLI on the design space.

Once all simulation scenarios have been evaluated, the analysis proceeds to the RSM stage, which is organized into three phases: (i) data preparation; (ii) model formulation; and (iii) performance assessment. In the data preparation phase, all system sweeping variables belonging to the SDS, are normalized according to:

$$\bar{X} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}, \tag{1}$$

where $X_{\min}$ and $X_{\max}$ denote the minimum and maximum SDS variable values set across the complete simulation. This normalization maps these metrics to the interval $[0, 1]$, improving numerical conditioning and facilitating the comparison of scenarios with different absolute impairment levels. Ultimately, this prevents the imbalanced distribution of errors in the regression procedure.

TABLE I
COMPUTATION PARAMETERS

System Parameters		
Parameter	**Value**	**Units**
α	0.19	dB/km
γ	1.33	$\mathrm{W}^{-1} \cdot \mathrm{km}^{-1}$
GVD	16.8	$\mathrm{ps}^2/\mathrm{km}$
Channel Spa.	6%	/channel
Monte Carlo samples	10^6	points
Optical Band	1550	nm
Symbol Rate	4	GBd/DSC
Pulse Shape	Nyquist	–
Modulation	DP-16QAM	—
CUT	central	DSC
Sweeping Variables		
Number of Nodes	1-10	unitary
Capacity	50-400	Gb/s
Span Length	50-300	km
Output Parameter		
η	–	W^{-2}

The dependence of the NLI on the network parameters is then captured using the RSM formulation, estimated through ordinary least squares regression as implemented in MATLAB's `fitlm` function. Since the modeled response is the NLI metric, the intercept term (i.e. all SDS variables null) is omitted under the assumption that the NLI is null in the absence of system loading. The resulting model is given by

$$\begin{aligned} \eta = \beta_N N + \beta_L L + \beta_C C + \beta_{NN} N^2 + \beta_{LL} L^2 \\ + \beta_{CC} C^2 + \beta_{NL} NL + \beta_{NC} NC + \beta_{LC} LC. \end{aligned} \tag{2}$$

with β_{ii} the weights calculated by the RSM model, where $i = \{L, C, N\}$. The quality of the fitted model is finally assessed using the coefficient of determination, R^2, defined as

$$R^2 = 1 - \frac{\sum_{k=1}^{N} \left(Y_k - \hat{Y}_k\right)^2}{\sum_{k=1}^{N} \left(Y_k - \bar{Y}\right)^2}, \tag{3}$$

where Y_k and $\hat{Y}_k$ are the measured and predicted NLI, respectively, and $\bar{Y}$ is the sample mean of the measured data. Note that there is a trade-off in the complexity of Eq. 2 with respect to the number of higher-order terms included. Incorporating more higher-order terms may improve the fit, but it can also reduce practical interpretability, whereas an overly simplified fitting function may result in poor accuracy.

III. RESULTS

This section presents the numerical evaluation of the proposed framework and discusses the impact of the swept design parameters on the nonlinear response of the considered horseshoe-based network. The results are organized to first illustrate the overall behavior across the full design space, and then to quantify the accuracy of the quadratic regression model and the relative importance of the main and interaction terms.

In Fig. 2, the results of all normalized NLI (according to Eq. 1) computations are represented in a four-dimensional view of the design space. The first three dimensions correspond to the SDS variables, while the fourth dimension is encoded through the size and color of the markers, which are proportional to the computed values of $\bar{\eta}$. This representation provides a compact overview of the nonlinear behavior across the full parameter space. Notably, the computed NLI is increasing with both C and N, while decreasing for larger values of L, as predicted by the literature [9], [11]. This behavior reflects the expected growth of nonlinear penalties as the network becomes more heavily loaded, particularly in scenarios with higher aggregate capacity and a larger number of traversed nodes. Likewise, this behavior also highlights that, with increasing span-length, and therefore fiber accumulated dispersion, the impact of NLI in the system performance decreases.

Given all the computed data, we apply the RSM following the procedure described in Eq. 2. Once the β coefficients are estimated, the regression fitting shown in Fig. 3 compares

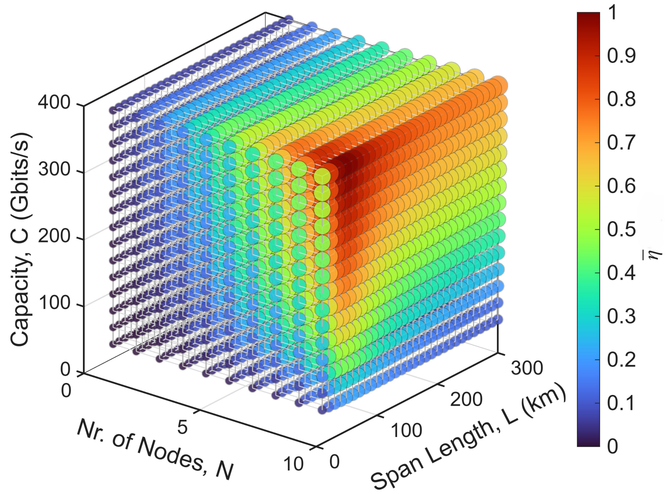

Fig. 2. EGN computation results across all SDS space. Note that the size of the markers per each point is proportional to the size of accumulated NLI in that scenario.

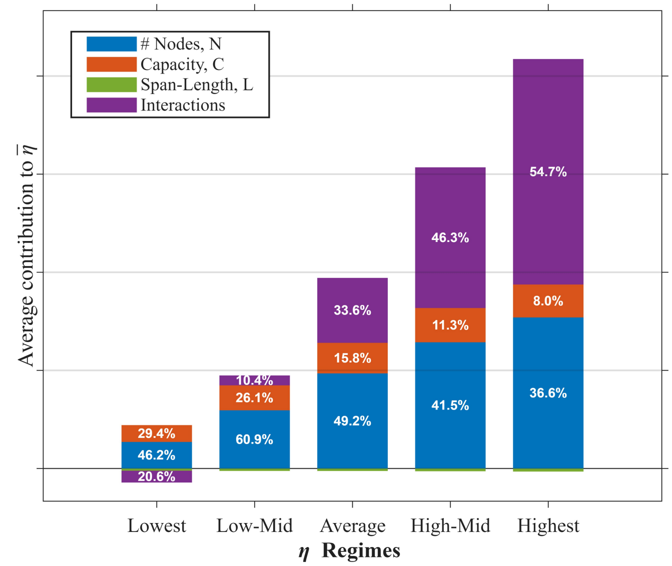

Fig. 4. Breakdown of sweeping variables SDS weights in different NLI regimes, as well as the combined effects of the interaction terms (higher-order parameters).

the predicted (RSM model) $\bar{\eta}$ against the computed (EGN) $\bar{\eta}$, with the fitting error evaluated using (3). The close agreement between the measured and predicted values indicates that the quadratic regression captures the main trends of the simulated SDS with good accuracy ($R^2 = 0.988$).

We then divide the different NLI values into five regimes: lowest, low-middle, average, high-middle, and highest. This division was achieved by partitioning the computed dataset according to the quantiles [0, 0.25, 0.50, 0.80, 0.90, 1.0].

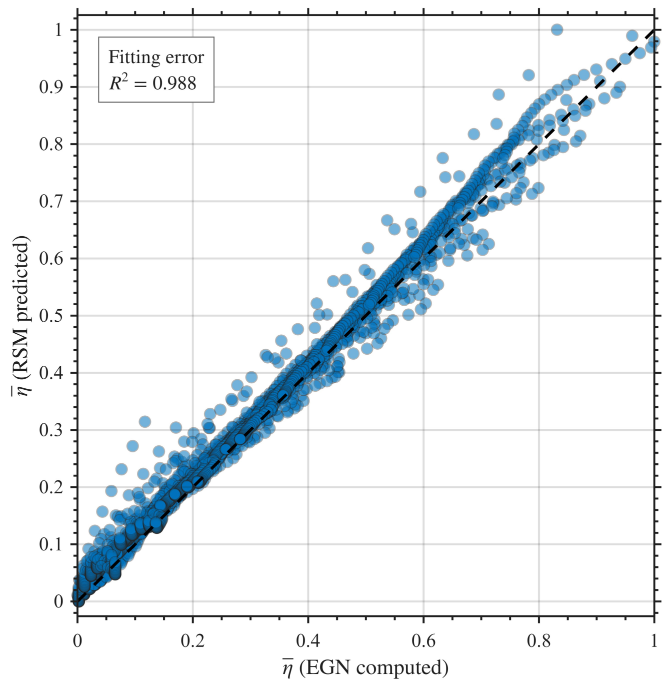

Fig. 3. Normalized predicted (RSM model) versus computed (EGN) NLI values for all scenarios attained by sweeping SDS. The fitting error is also presented in the legend of the plot.

By splitting the full domain of results, it is possible to examine each regime individually and identify which factors are most influential in each operating regime. In Fig. 4, the contribution of each SDS variable is presented, together with the *Interactions* term, which represents the combined effects of more than one variable in the SDS, namely $\{\beta_{NN}, \beta_{LL}, \beta_{CC}, \beta_{NL}, \beta_{NC}, \beta_{LC}\}$. This breakdown reveals how the relative importance of the swept variables changes as the NLI level increases. In the lower regimes, the response is dominated by the direct effects of the design variables, whereas in the higher regimes, the *interactions* terms become increasingly relevant as the multiplicative second-order variables becomes larger, indicating that nonlinear coupling between variables cannot be neglected. Notably, it is possible to see that the isolated effect of per-node Capacity C, decreases with increasing NLI regime, as well as the effect of the number of nodes, N. However, the design variable span length, L, always contributes with a negative value, supporting the theory that increasing L introduces additional accumulated fiber dispersion, decreasing the nonlinear penalty in the system.

Since the *Interactions* term has a strong influence on the final NLI computation, especially in the highest NLI regime, we further breakdown each *interaction* term to determine which variable pairs contribute most strongly to the second-order interactions. In particular, this analysis highlights that these coupled effects are essential for accurate planning of scalable horseshoe optical networks. The most dominant term is the combination of number of nodes N and the system capacity C, evaluated at β_{NC}. This term highly affects NLI, which is in line with the expected behavior already described in this section. Likewise, it is possible to observe the trend that N is responsible for the increasing NLI, where L is related often with a relaxation of the NLI penalties. The influence of the variable SDS C appears neutral in this analysis, however, it produces a strong effect on the simulation runtime, as the frequency bandwidth increase also requires further computation samples to capture accurately the NLI. Note that using the data normalization described in Eq. 1 will always produce higher terms for second-order *interactions* as opposed to the first-order, since the coefficients get multiplied twice by variables that are bounded to [0,1].

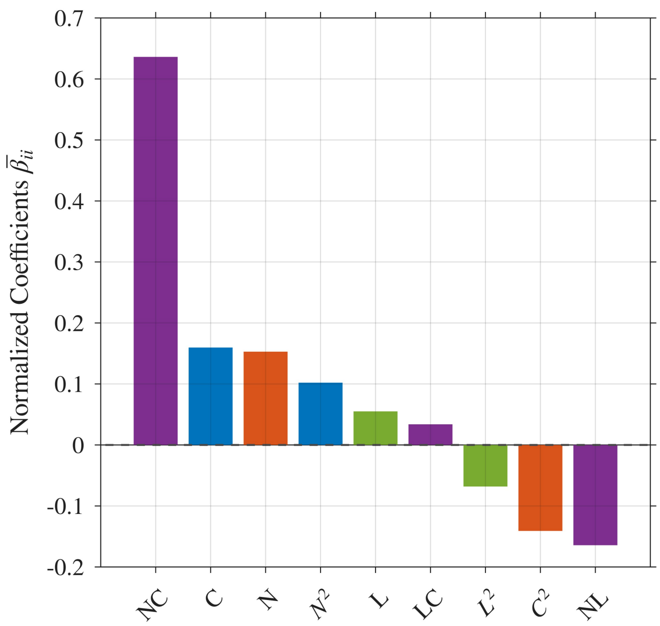

Fig. 5. *Interactions* terms breakdown using normalized coefficient values. Note that first-order terms contribute implicitly with higher magnitude than second-order terms.

The results confirm that the NLI response is strongly dependent on the joint variation of C and N, with interaction terms becoming increasingly relevant in higher-NLI regimes. Moreover, across all the results, it is shown that increasing L, produces an overall decrease in the NLI penalty, consistent with known literature results. Overall, the proposed RSM modeling approach provides both physical insight and a computationally efficient way to assess NLI-limited scalability in horseshoe optical networks.

IV. Conclusions

This paper analyzed the impact of NLI on the scalability of horseshoe optical networks using EGN-based simulations and a quadratic RSM model. By sweeping capacity, number of nodes, and span length, we quantified how these parameters shape the accumulated NLI across the design space. The results show that NLI increases with higher capacity and node count, while interaction terms become increasingly relevant at high-impairment regimes. Nonetheless, this study also captured that the increase of fiber length, i.e. the increase of fiber accumulated dispersion, produced lower NLI scenarios, matching the theory for NLI noise. This novel RSM-based approach regarding the *interaction* terms (i.e. combined effects of C, L and N), revealed which of these interactions most contribute to the accumulated NLI, a valuable understanding towards system design and optimal system planning. Overall, the proposed methodology provided valuable insights using the EGN simulation data, while enabling a fast and compact framework for NLI prediction, making it a practical tool for network dimensioning and planning in scalable horseshoe deployments, which can then be extended to more complex systems, such as Wavelength Division Multiplexing (WDM). Further work will focus on three directions. First, the NLI coefficients can be scaled to launch power values to express these results in terms of power penalties in real deployments. Second, the limits of the RSM approach can be evaluated by comparing its response to data fitting and to data prediction. Finally, dedicated experiments can be used to reduce cross-effects among the SDS variables, enabling a clearer assessment of the individual influence of C, L, and N.

References

[1] D. Welch, A. Napoli, J. Bäck, W. Sande, J. Pedro, F. Masoud, C. Fludger, T. Duthel, H. Sun, S. J. Hand, T.-K. Chiang, A. Chase, A. Mathur, T. A. Eriksson, M. Plantare, M. Olson, S. Voll, and K.-T. Wu, "Point-to-multipoint optical networks using coherent digital subcarriers," *Journal of Lightwave Technology*, vol. 39, no. 16, pp. 5232–5247, 2021.

[2] D. Welch and et al., "Digital subcarrier multiplexing: Enabling software-configurable optical networks," *Journal of Lightwave Technology*, vol. 41, no. 4, pp. 1175–1191, 2023.

[3] E. Paz, G. Parisi, P. Torres-Ferrera, T. Silvério, S. Srivallapanondh, M. Masanas, L. Velasco, M. S. Erkilinç, R. Magri, and A. Napoli, "Analytical qot model for bi-di single-fiber point-to-multipoint dscm coherent systems," in *Optical Fiber Communication Conference (OFC) 2026*. Optica Publishing Group, 2026, p. M2C.4.

[4] M. Radovic, A. Sgambelluri, F. Cugini, R. Kapuscinski, M. Hosseini, S. Parker, R. Berozashvili, F. Paolucci, A. Napoli, and N. Sambo, "Sdn-controlled open ran x-haul with point-to-multipoint transceivers on a horseshoe network," *J. Opt. Commun. Netw.*, vol. 17, no. 11, pp. E109–E116, Nov 2025.

[5] M. M. Hosseini, G. Parisi, J. ao Pedro, P. Torres-Ferrera, and A. Napoli, "Design of extended horseshoe-and-spur optical networks leveraging nonlinearity-awareness," *J. Lightwave Technol.*, vol. 44, no. 10, pp. 3983–3994, May 2026.

[6] M. M. Hosseini, J. Pedro, N. Costa, C. Castro, and A. Napoli, "Optimized design of horseshoe-and-spur filterless networks leveraging point-to-multipoint coherent pluggable transceivers," *Journal of Optical Communications and Networking*, vol. 16, no. 10, pp. 969–980, 2024.

[7] T. Silvério, G. Parisi, P. T. Ferrera, and A. Napoli, "Egn method for nli estimation in digital subcarrier allocation for horseshoe networks," in *2025 IEEE Photonics Society Summer Topicals Meeting Series (SUM)*, 2025, pp. 1–2.

[8] C. Castro, P. Torres-Ferrera, M. Sezer Erkilinc, J. Sime, G. Parisi, J. Pedro, M. Quagliotti, M. Porrega, D. Hillerkuss, C. Fludger, E. Riccardi, and A. Napoli, "Power consumption considerations of coherent transceivers in filterless point-to-multipoint metro-aggregation networks with digital subcarrier multiplexing," *Journal of Optical Communications and Networking*, vol. 17, no. 6, pp. 526–542, 2025.

[9] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, and F. Forghieri, "EGN model of non-linear fiber propagation," *Optics Express*, vol. 22, pp. 16 335–16 362, 06 2014.

[10] G. Parisi, T. Silvério, P. Torres-Ferrera, Y. Rumaldo, N. Costa, C. Castro, M. M. Hosseini, J. Pedro, L. Velasco, and A. Napoli, "Fiber nonlinearities in short-reach systems: Modeling and experimental verification with 400g digital subcarrier multiplexing," *Journal of Lightwave Technology*, pp. 1–10, 2026.

[11] M. R. Zefreh, F. Forghieri, S. Piciaccia, and P. Poggiolini, "Nonlinearity assessment in long haul dispersion managed fiber optic links," in *2020 Italian Conference on Optics and Photonics (ICOP)*, 2020, pp. 1–3.