Online kernel-based phase recovery for parametrically amplified optical transmission

Long Hoang Nguyen
AiPT, Aston University
Birmingham B4 7ET, UK
l.nguyen7@aston.ac.uk

Sonia Boscolo
AiPT, Aston University
Birmingham B4 7ET, UK
s.a.boscolo@aston.ac.uk

Stylianos Sygletos
AiPT, Aston University
Birmingham B4 7ET, UK
s.sygletos@aston.ac.uk

***Abstract*—We present a kernel adaptive filtering-based phase compensation method for transmission links with multiple cascaded fibre-optical parametric amplifiers (FOPAs). Our proposed algorithm predicts and cancels the phase distortions induced by pump-phase modulation and laser line-width across all amplification stages. Through numerical simulations, we show effective correction of phase errors in 16-quadrature-amplitude modulation signal transmission, substantially surpassing the performance of conventional carrier phase recovery.**

***Index Terms*—Fibre-optical parametric amplifier, digital signal processing, phase recovery, kernel adaptive filter**

I. Introduction

Fibre-optical parametric amplifiers (FOPAs) [1] utilising four-wave mixing for signal amplification have recently attracted significant research interest owing to the multiple advantages that they can offer in optical communications, including broad gain bandwidth [2] and ultra-fast response. However, stimulated Brillouin scattering (SBS) poses a significant challenge to integrating FOPAs into next-generation optical systems as it limits the pump power that can be delivered to the highly nonlinear fibre (HNLF) [3] and, consequently, the achievable signal gain. A commonly used approach to counteract the SBS effect is to broaden the pump source's line-width, thereby reducing the power spectral density across the Brillouin bandwidth. This technique typically relies on external modulation of the pump phase [4] driven by an electrical signal with multiple radio-frequency (RF) tones. While effective in suppressing SBS and hence enabling high gain [5], this approach introduces undesired temporal fluctuations in the complex parametric gain through modulation of the pump's instantaneous frequency [6]. These fluctuations are then transferred to the output signal phase, becoming a source of distortion in coherent-detection optical systems.

Integrating machine learning with digital signal processing is increasingly recognised for its potential to address performance challenges in optical networks. Kernel-based regression methods have proven to be a robust approach for a range of signal processing applications [7], [8]. These methods draw on the theory of reproducing kernel Hilbert spaces to model nonlinear systems by implementing a nonlinear transformation of the input into a high-dimensional feature space where inner products can be calculated using a positive-definite kernel function satisfying Mercer's condition. This way, the solution obtained as a linear functional in the feature space, corresponds to the solution of the nonlinear problem in the input space. Kernel-based adaptive filtering (KAF) is an online version of these methods where the learned parameters are adapted as the dictionary of observation data evolves over time, with the goal of minimising a least-squares (LS) cost function [8]. As the functional representation in the feature space grows with increasing amount of observations, a naïve implementation of an online kernel method would require growing computational resources during operation. This issue has been addressed in the literature by various dictionary learning strategies to control the dictionary's size within manageable limits. In this paper, we deploy the sliding-window kernel recursive least-squares (SWKRLS) algorithm introduced in [9], which, in every iteration, accepts the new datum and discards the oldest basis, thereby maintaining a dictionary of fixed size.

Our previous work [10] has employed the SWKRLS algorithm to mitigate the phase distortion caused by the pump-phase modulation of an optical phase conjugation device after a stage of conventional carrier phase recovery. In this paper, we introduce a fully online, single-stage phase recovery method based on the SWKRLS algorithm for cascaded FOPA transmission links. The proposed algorithm is capable to estimate and then cancel both the phase distortion caused by pump-phase modulation within each FOPA in the link and the random phase noise induced by non-zero spectral width of the laser sources. The performance of the algorithm is verified numerically in 28-Gbaud single-polarisation 16 quadrature-amplitude modulation (QAM) signal transmission over a cascade of ten FOPAs, achieving substantial bit-error-rate (BER) improvement over conventional phase recovery.

II. Kernel-based phase recovery algorithm

Figure 2 shows the block diagram of our proposed phase recovery scheme. The total phase distortion of the signal at the input of the compensation unit can be written as $\phi_n = \delta\phi_n + \varphi_n$, where n is the symbol index, $\delta\phi_n$ includes the total Wiener random laser phase noise of the system and the contribution to phase noise from the amplified spontaneous

This work was supported by the H2020 MSCA ETN project POST-DIGITAL (EC 263 GA 860360), and the UK EPSRC grants TRANSNET (EP/R035342/1) and CREATE (EP/X019241/1).

emission of the FOPA, and φ_n is the phase distortion caused by the pump-phase modulation. A SWKRLS-based phase estimation block aimed to predict the total phase distortion $\hat{\phi}_n$, and the received symbol y_n was then rotated by this amount, i.e., $z_n = y_n e^{-i\hat{\phi}_n}$, and passed through a decision circuit where a decision-directed symbol, $[z_n]_{\mathrm{D}}$, was extracted. The phase difference between the symbol before phase rotation and the decision-directed symbol, $\tilde{\phi}_n = \angle\{y_n\} - \angle\{[z_n]_{\mathrm{D}}\}$, was used as a phase distortion teacher for the training of the phase estimation algorithm in the decision-directed mode. The algorithm used reference symbols as a training teacher substitution before convergence.

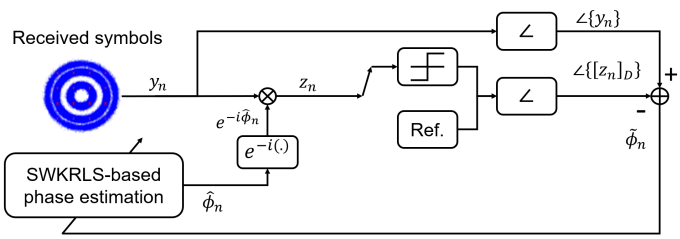

Fig. 1. Block diagram of the proposed kernel-based carrier phase recovery scheme.

The kernel function of the phase estimation algorithm was designed based on the distinctive characteristics of the phase distortion. Our prior understanding indicated that the total phase distortion can be represented by the superposition of a periodic and a long-trend time series, where the former results from the impact of the pump-phase modulation and the latter is contributed by the laser phase noise. The periodic pattern, oscillating at the frequencies of the RF tones used within the FOPA's dithering scheme, can be captured by a summation of exponential sine squared kernels,

$$k_{\mathrm{per}}(t_m, t_n) = \sum_{j=1}^{N_t} \exp[-2\sin^2[\pi d(t_m, t_n)/p_j]/l_{\mathrm{per}}^2], \quad (1)$$

where t_n is the time corresponding to the nth symbol, $d(a, b)$ is the Euclidean distance between two points, p_j ($j = \overline{1, N_t}$) represent the kernel's periods, N_t is the number of dithering tones, and l_{per} is the kernel's length-scale. As the RF tone locations f_j can be extracted at the receiver, the associated periods are given by $p_j = 1/f_j$. The long-term trend, induced by the relatively slower variation of the laser phase noise, can be modelled with a Gaussian kernel (also known as radial basis function (RBF) kernel),

$$k_{\mathrm{RBF}}(t_m, t_n) = \exp[-d(t_m, t_n)^2/(2l_{\mathrm{RBF}}^2)], \quad (2)$$

where l_{RBF} is the length-scale. The total kernel can be then constructed as

$$k_{\mathrm{tot}}(t_m, t_n) = k_{\mathrm{per}}(t_m, t_n) + k_{\mathrm{RBF}}(t_m, t_n). \quad (3)$$

to include both types of patterns into the phase distortion prediction.

We employed the SWKRLS algorithm [9] and our customised kernel design for the prediction of the total phase distortion $\hat{\phi}_n$. In the notation used below, the capitalised bold letters denote vectorised versions of the corresponding variables. The kernel method here attempts to minimise the LS cost function: $|\tilde{\mathbf{\Phi}} - \hat{\mathbf{\Phi}}|^2$, which measures the difference between the phase distortion driven by the decision-directed circuit and the predicted phase distortion. The method maps the input data – the time vector $\mathbf{T}$ – into a feature space where the predicted output can be linearly represented in terms of the transformed data $\dot{\mathbf{T}}$, i.e., $\hat{\mathbf{\Phi}} = \dot{\mathbf{T}}\dot{\mathbf{\Theta}}$. The transformed solution $\dot{\mathbf{\Theta}}$ can also be represented in terms of the transformed data as $\dot{\mathbf{\Theta}} = \dot{\mathbf{T}}^{\mathrm{T}}\alpha$, where the superscript T denotes the transpose operation. Moreover, introducing the kernel matrix $\mathbf{K} = \dot{\mathbf{T}}\dot{\mathbf{T}}^{\mathrm{T}}$, the LS function in feature space can be rewritten as $|\tilde{\mathbf{\Phi}} - \mathbf{K}\alpha|^2$. The solution α of the LS problem can now be found by only computing the kernel matrix, which in our implementation of the method, was selected from the customised design of (3), i.e., $\mathbf{K}(m, n) = k_{\mathrm{tot}}(t_m, t_n)$.

In an online prediction setup, the algorithm is given a stream of time-phase distortion pairs $\{(t_1, \tilde{\phi}_1), (t_2, \tilde{\phi}_2), \dots\}$, but in the SW approach, only the last N input-output pairs are taken into account. At the symbol index n, the training input $\mathbf{T}_n = [t_n, t_{n-1}, \dots, t_{n-N+1}]^{\mathrm{T}}$ and the training output $\tilde{\mathbf{\Phi}}_n = [\tilde{\phi}_n, \tilde{\phi}_{n-1}, \dots, \tilde{\phi}_{n-N+1}]^{\mathrm{T}}$ are formed, and the corresponding regularised kernel matrix $\mathbf{K}_n = \dot{\mathbf{T}}_n\dot{\mathbf{T}}_n^{\mathrm{T}} + \lambda\mathbf{I}$ can be calculated ($\mathbf{I}$ is the identity matrix and λ is a regularisation constant). The updated solution α_n is then obtained as

$$\alpha_n = \mathbf{K}_n^{-1}\tilde{\mathbf{\Phi}}_n. \quad (4)$$

The calculation of the updated solution requires the calculation of the inverse kernel matrix for each window, which is costly both computationally and memory-wise. However, in the update algorithm developed in [9], the computation of $\mathbf{K}_n$ and $\mathbf{K}_n^{-1}$ is not done explicitly, but these matrices are updated recursively using the previous matrices $\mathbf{K}_{n-1}$ and $\mathbf{K}_{n-1}^{-1}$. Specifically, given the regularised kernel matrix $\mathbf{K}_{n-1}$, the new regularised kernel matrix $\mathbf{K}_n$ is constructed by removing the first row and column of $\mathbf{K}_{n-1}$, referred to as $\bar{\mathbf{K}}_{n-1}$, and adding kernels of the new data as the last row and column:

$$\mathbf{K}_n = \begin{bmatrix} \bar{\mathbf{K}}_{n-1} & \mathbf{b}_n \\ \mathbf{b}_n^{\mathrm{T}} & c_n \end{bmatrix}, \quad (5)$$

where $\mathbf{b}_n = [k_{\mathrm{tot}}(t_{n-N+1}, t_n), ..., k_{\mathrm{tot}}(t_{n-1}, t_n)]^{\mathrm{T}}$, and $c_n = k_{\mathrm{tot}}(t_n, t_n) + \lambda$. The inverse kernel matrix $\mathbf{K}_n^{-1}$ is then updated accordingly as

$$\mathbf{K}_n^{-1} = \begin{bmatrix} \bar{\mathbf{K}}_{n-1}^{-1} - \bar{\mathbf{K}}_{n-1}^{-1}\mathbf{b}_n\mathbf{d}_n^{\mathrm{T}} & \mathbf{d}_n \\ \mathbf{d}_n^{\mathrm{T}} & e_n \end{bmatrix}, \quad (6)$$

where $e_n = (c_n - \mathbf{b}_n^{\mathrm{T}}\bar{\mathbf{K}}_{n-1}^{-1}\mathbf{b}_n)^{-1}$ and $\mathbf{d}_n = -\bar{\mathbf{K}}_{n-1}^{-1}\mathbf{b}_n e_n$. After calculating the updated solution α_n using Eq. (4), our phase estimation algorithm produced the predicted phase for the $(n+1)$th symbol as

$$\hat{\phi}_{n+1} = \kappa_{n+1}^{\mathrm{T}}\alpha_n, \quad (7)$$

where κ_{n+1} contained the calculated kernel values between the dictionary points and the new point, i.e., $\kappa_{n+1} = [k_{\text{tot}}(t_{n-N}, t_{n+1}), \ldots, k_{\text{tot}}(t_n, t_{n+1})]^{\text{T}}$.

The operation of our kernel-based phase estimation block can be summarised as follows:

Initialise $\mathbf{K}_0 = \mathbf{I}(1+\lambda)$ and $\mathbf{K}_0^{-1} = \mathbf{I}/(1+\lambda)$
for n=1,2,... **do**
  Given $\tilde{\phi}_n$, form $\tilde{\mathbf{\Phi}}_n$
  Calculate $\mathbf{K}_n$ from (5)
  Calculate $\mathbf{K}_n^{-1}$ from (6)
  Obtain the solution α_n given $\tilde{\mathbf{\Phi}}_n$ from (4)
  Produce $\hat{\phi}_{n+1}$ for the next symbol from (7)
end for

III. System's Configuration and Numerical Results

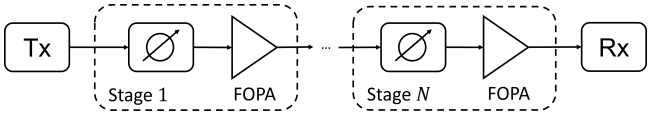

Fig. 2. Schematic diagram of the N-cascaded FOPA system.

We considered a system comprising $N = 10$ identical cascaded stages, each of 20-dB linear loss (corresponding to 100-km standard single-mode fibre) followed by a FOPA, as depicted in Fig. 2. We modelled the FOPA by its complex signal gain calculated as in [11], where the phase mismatch comprised an additional instantaneous term induced by the phase modulation of the pump following [6]. The FOPA was operated at its maximum power gain of 25 dB, accounting for an additional 5-dB insertion loss of the device. We used a four-tone pump-phase modulation ($N_t = 4$) with a base frequency of 100 MHz and a multiple of three spacing between successive tones, i.e., the set of RF tones was $[0.1, 0.3, 0.9, 2.7]$ GHz. The tone's amplitudes and phases were optimised to ensure a close-to-uniform power distribution among the peaks generated across the broadened pump spectrum by implementing a stochastic gradient descent method in TensorFlow [12]. The optimised pump-phase modulation scheme enabled significant increase in the SBS power threshold, thereby satisfying the requirement for a SBS-limited 25-dB gain.

We performed numerical simulations of the transmission of a single-polarisation 28-Gbaud 16-QAM Nyquist shaped signal with a roll-off factor of 0.1. The laser line-widths were 50 kHz and 30 kHz for the transmitter and receiver units and the FOPA pumps, respectively. To avoid the signal symbols experiencing exactly the same phase distortion along the FOPA link, we included a random time shift in the pump-phase modulation sinusoidal waveform at each FOPA stage. The amplifer's noise figure was 4.5 dB. We employed the directly-counted bit-error-rate (BER) as a system's performance metric measured over 50×2^{16} symbols.

In our kernel selection (Eqs. (1)–(3)), since the periods p_j corresponding to the dithering frequencies can be extracted

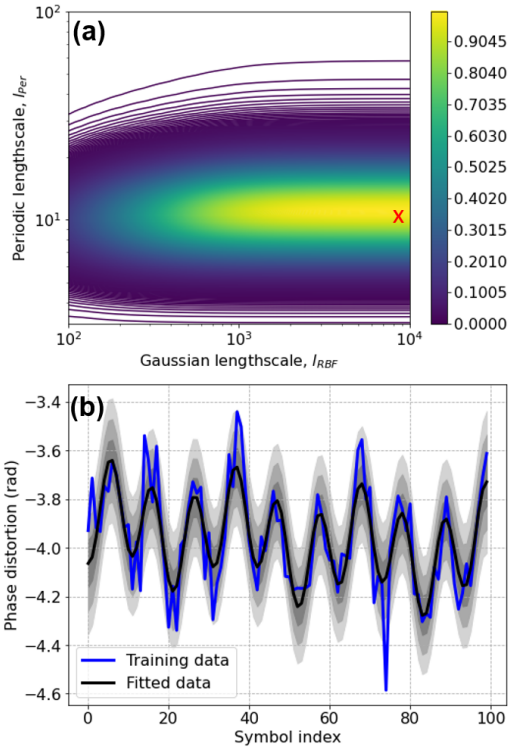

Fig. 3. Kernel's hyper-parameter optimisation in the pre-setting phase. (a) Map of normalised marginal likelihood in the plane of the kernel's length-scales l_{per} and l_{RBF}. The red cross indicates the choice of optimum hyper-parameters. (b) Fitted phase distortion using the optimised hyper-parameters (black) compared to the actual one (blue). The gray shaded regions represent the values that lie within one, two, and three standard deviations of the mean in the normal distribution.

at the receiver, we are left with two hyper-parameters to optimise, i.e., the length-scales l_{per} and l_{RBF}. This can done by maximising the marginal likelihood of a Gaussian process [13] using the training data set within a pre-setting stage. Figure 3(a) shows the heat-map of the normalised marginal likelihood, where larger values (represented by brighter colours) correspond to higher fitting confidence. We used a pre-setting set of 100 symbols and the regularisation parameter $\lambda = 0.1$ for this optimisation. We can see that the optimum value for the length-scale of the periodic kernel is 10, while the Gaussian kernel's length-scale needs to be relatively large, and hence we selected the value 10^4. Figure 3(b) illustrates the accuracy of our kernel-based regression using the optimised hyper-parameters. We used these for the SWKRLS-based phase estimation algorithm with a dictionary size of 100.

We compared the performance of our proposed kernel-based phase recovery method as a function of the number of FOPA stages with conventional phase recovery accomplished by the one-tap least-mean-squares (LMS) algorithm [14], as shown in Fig. 4. As expected, the performance of the conventional method deteriorates quickly with increasing number of FOPA stages. Being designed to operate on the slower laser phase noise, the conventional scheme cannot in fact track and sufficiently suppress the accumulation of the high-frequency

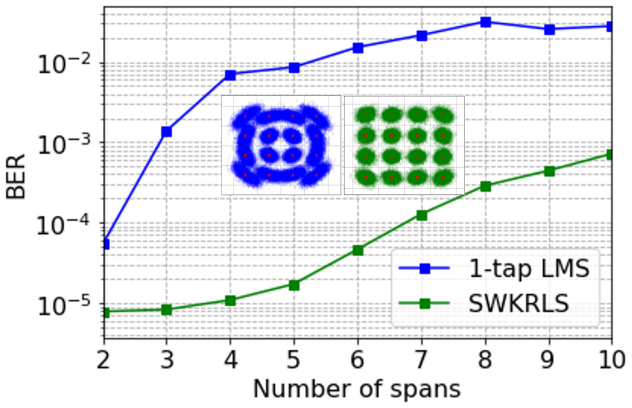

Fig. 4. BER after conventional phase recovery (blue) and the proposed kernel-based phase recovery (green) versus number of cascaded FOPA stages. The inset shows the constellation diagrams after ten stages.

dithering-induced phase distortion along the link. This makes the algorithm prone to phase cycle slips, with consequent increase of phase detection errors. Our proposed scheme, which uses a dithering-induced distortion customised kernel, largely outperforms the one-tap LMS method, bringing about a BER improvement of at least an order of magnitude across the cascaded FOPA system considered.

IV. Conclusion

We developed an online kernel-based approach for carrier phase recovery in systems with multiple cascaded FOPAs. By constructing a customised kernel that captures the physical properties of the phase distortion, especially of the high-frequency distortion caused by phase modulation of the pump sources, our scheme can outperform conventional phase recovery schemes, which typically react slowly to temporal phase variations. The efficacy of our method has been demonstrated in 16-QAM signal transmission, achieving a BER improvement of at least an order of magnitude over the one-tap LMS phase recovery algorithm across a cascade of ten FOPAs.

References

[1] M.E. Marhic, N. Kagi, T.-K. Chiang, and L.G. Kazovsky, "Broadband fiber optical parametric amplifiers," Opt. Lett., vol. 21, pp. 573–575, 1996.

[2] V. Gordienko, M.F.C. Stephens, A.E. El- Taher, and N.J. Doran, "Ultra-flat wideband single-pump Raman-enhanced parametric amplification," Opt. Express, vol. 25, pp. 4810–4818, 2017.

[3] V. Gordienko, A.D. Szabó, M.F.C. Stephens, V. Vassiliev, *et al.*, "Limits of broadband fiber optic parametric devices due to stimulated Brillouin scattering," Opt. Fiber Technol., vol. 66, pp. 102646, 2021.

[4] J.B. Coles, B.P.-P. Kuo, N. Alic, S. Moro, *et al.*, "Bandwidth-efficient phase modulation techniques for stimulated Brillouin scattering suppression in fiber optic parametric amplifiers," Opt. Express, vol. 18, pp. 18138–18150, 2010.

[5] T. Torounidis, P.A. Andrekson, and B.-E. Olsson, "Fiber-optical parametric amplifier with 70-dB gain," IEEE Photon. Technol. Lett., vol. 18, pp. 1194–1196, 2006.

[6] A. Mussot, A. Durécu-Legrand, E. Lantz, C. Simonneau, *et al.*, "Impact of pump phase modulation on the gain of fiber optical parametric amplifier," IEEE Photon. Technol. Lett., vol. 16, pp. 1289–1291, 2004.

[7] B. Scholkopf and A.J. Smola, *Learning with Kernels, Support Vector Machines, Regularization, Optimization and Beyond* (MIT Press, 2001).

[8] J.L. Rojo-Álvarez, M. Martínez-Ramón, J. Muñoz-Marí, and G. Camps-Valls, *Digital Signal Processing with Kernel Methods* (John Wiley & Sons Ltd, 2018).

[9] S.V. Vaerenbergh, J. Vía, and I. Santamaría, "A sliding-window kernel RLS algorithm and its application to nonlinear channel identification," in International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 789–792, 2006.

[10] S. Boscolo, T.T. Nguyen, A.A.I. Ali, S. Sygletos, and A.D. Ellis, "Kernel adaptive filtering-based phase noise compensation for pilot-free optical phase conjugated coherent systems," Opt. Express, vol. 30, pp. 19479–19493, 2022.

[11] M.E. Marhic, *Fiber Optical Parametric Amplifiers, Oscillators and Related Devices* (Cambridge University Press, 2007).

[12] L.H. Nguyen, S. Boscolo, and S. Sygletos, "Online digital compensation of pump dithering induced phase and amplitude distortions in transmission links with cascaded fibre-optical parametric amplifiers," Opt. Express, vol. 32, pp. 13467–13477, 2024.

[13] C.E. Rasmussen and C.K.I. Williams, *Gaussian Processes for Machine Learning* (The MIT Press, 2006).

[14] I. Fatadin, D. Ives, and S.J. Savory, "Blind equalization and carrier phase recovery in a 16-QAM optical coherent system," J. Lightwave Technol., vol. 27, pp. 3042–3049, 2009.