Preemphasis-Aware Semiconductor Optical Amplifier Model

Hartmut Hafermann,[1,*] Xiaohui Zhao,[1] Shuqi Yu[1] and Yann Frignac[1]

[1] *Huawei Technologies France, Paris Research Center, Optical Communication Technology Lab*

[*] *hartmut.hafermann@huawei.com*

Abstract: A preemphasis-aware model for SOAs with non-flat WDM inputs yielding a root-mean-square error of less than 0.05 dB is presented. It outperforms generic neural-network models while using a fraction of the training data.

1. Introduction

The linear scaling of system capacity with optical bandwidth is driving efforts to realize Ultra-Wide Band (UWB) transmission beyond the C band. SOAs are receiving increased attention in this context [1]. The interplay between wavelength-dependent attenuation, stimulated Raman scattering and the responses of multiple cascaded amplifiers across bands becomes particularly relevant in UWB systems and calls for an optimization of the input power spectral load. Taking into account the SOA response beyond an idealized flat or tilted gain is hence desirable, but physical models such as the Connelly model [2] are computationally expensive. The various physical parameters also require elaborate measurement techniques.

Neural network (NN) models have shown promise because of their speed and differentiability aiding gradient-based optimization. The amount of training data required however can be substantial. They also lack explainability and control of error. The error of cascaded predictions needed for optical links typically increases exponentially, imposing strict constraints on the accuracy of individual models.

Recently a new type of model to predict the gain profiles of Erbium-Doped Optical Amplifiers (EDFAs) has been presented [3]. A regression-based ML model incorporates physical knowledge in the form of an analytical formula that accounts for the impact of a preemphasis of the input spectral load. The hybrid approach offers high accuracy, is parameter agnostic and requires minimal data for training. For WDM inputs with moderate preemphasis, a root-mean-square error (RMSE) of 0.05 dB was reported.

Here we show that the preemphasis-aware approach accurately models the steady-state gain response of an SOA with non-flat WDM inputs. The approach outperforms generic neural network models, despite being trained only on a small subset of the data available for training.

2. Experimental setup and measurements

We are interested in the static gain response of the SOA and ignore dynamic gain effects. WDM inputs are simulated by an ASE noise source with flat output power spectral density (Fig. 1 a). A wave shaper (WS) generates a WDM comb consisting of 40 selectively attenuated channels. Total input power is controlled by the VOA. Both total input and output power are measured by the power meter (PM) following calibration. The optical spectrum analyzer (OSA) output is integrated to obtain the power per channel $P_i = P(\lambda_i)$ centered at wavelength λ_i. Per-channel power is normalized to sum to the measured total power.

We considered 6 injection currents evenly spaced between 1000 and 1500 mA. The generated inputs are divided into two groups. The first group has linear power variation (flat or tilted) across the band and 7 distinct tilt values between 0 and ± 3 dB and varying in 1 dB steps, as shown in Fig. 1 b). Total power was varied in 1dB steps from -6 to 7 dBm, totaling 588 samples. Corresponding gains show a small dependence on preemphasis, with up to 0.2 dB gain variation, see Fig. 1 c). Another group with 600 samples (100 for each current) contains inputs with random preemphasis, with per-channel power excursions $P_i/\bar{P}$ (positive and negative) spanning a range of 12 dB. Total power was varied in the range from -6 to 7 dBm, ranging from small-signal well into the saturated regime.

3. Preemphasis-aware SOA model

Combining the rate equation for the carrier density N in steady-state ($dN/dt = 0$) with the evolution equations of the signal and ASE photon densities of the Connelly model [2] and neglecting waveguide attenuation leads to:

$$\frac{I}{e} - R(x) - \sum_{i=1}^{N_{\mathrm{ch}}} Q^{\mathrm{in}}(\lambda_i)(G(\lambda_i, x) - 1) - Q_{\mathrm{tot}}^{\mathrm{out,ASE}}(x) = 0. \tag{1}$$

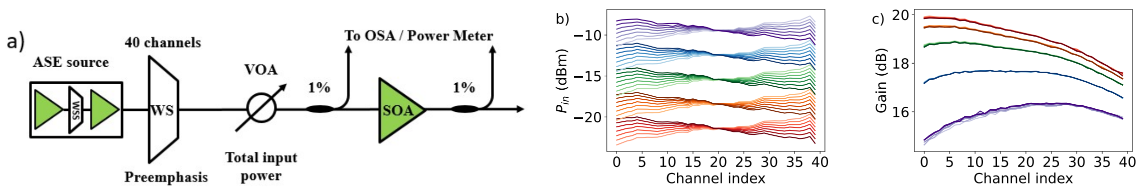

Fig. 1. a) The experimental setup for the characterization of the SOA under test, b) example flat/tilted power spectral load for 5 values of total input powers. c) Corresponding gains for fixed current.

This equation is akin to the extended Saleh model for EDFAs [4]. It balances carrier generation through the injection current I with carrier depletion through spontaneous and stimulated recombination. Here e is the electron charge and $Q^{\text{in}}(\lambda_i)$ the input photon flux (photons/s) in the channel centered around wavelength λ_i. $R(x)$ and $G(\lambda_i, x)$ are the in general *unknown* total carrier recombination rate (excluding guided spontaneous emission) and per-channel gain at the *unknown* population inversion x. $Q_{\text{tot}}^{\text{out,ASE}}(x)$ comprises forward and backward components of the total output flux due to amplified spontaneous emission. The model hence accounts for ASE self-saturation that cannot be neglected at large gain.

For two WDM inputs (e.g. one flat and one with random preemphasis) with same total input power, the carrier densities (population inversions) of the device differ, for two reasons: (i) for fixed power, the photon flux coupled into the device depends on wavelength. More importantly, (ii) the number of carriers converted into output photon flux depends on spectral power allocation because the gain is not flat.

Under a homogenous gain assumption, a gain response can be associated to the population inversion. Following [3], for a given preemphasis, we seek an *equivalent* total input power of a *flat* input spectral load that produces the same population inversion x (at the same injection current). To this end, we equate two versions of (1), one for each type of input. Substituting $Q^{\text{in}}(\lambda_i) = \lambda_i P_i/(hc)$, expressing the per-channel input power of the random load as $P_i = \bar{P} + \Delta P_i$ and that of the equivalent flat load as $P_i = \bar{P} + \Delta P_{\text{eq}}$, we equate, all other terms being equal, $\sum_{i=1}^{N_{\text{ch}}} \lambda_i(\bar{P} + \Delta P_i)(G_i(x) - 1) = \sum_{i=1}^{N_{\text{ch}}} \lambda_i(\bar{P} + \Delta P_{\text{eq}})(G(\lambda_i, x) - 1)$. Hence the per-channel input power correction is

$$\Delta P_{\text{eq}} = \sum_{i=1}^{N_{\text{ch}}} \lambda_i \Delta P_i (G(\lambda_i, N_{\text{ch}}\bar{P}) - 1) \Big/ \sum_{i=1}^{N_{\text{ch}}} \lambda_i (G(\lambda_i, N_{\text{ch}}\bar{P}) - 1). \tag{2}$$

Here λ_i accounts for (i) and the gain appears due to reason (ii). The unknown gain $G_i(x)$ at the unknown population inversion x is *approximated* by the gain of the *flat* input $G(\lambda_i, N_{\text{ch}}\bar{P})$ evaluated at the total input power of the *random* load $P_{\text{tot}}^{\text{in}} = \sum_i^{N_{\text{ch}}} P_i = N_{\text{ch}}\bar{P}$. The gain for a random load characterized by average per-channel power $\bar{P}$ and deviations ΔP_i, respectively, is approximated by the gain of a flat input as $G(\lambda_i, \bar{P}, \Delta P_i) \approx G(\lambda_i, N_{\text{ch}}(\bar{P} + \Delta P_{\text{eq}}))$. We tested iterative improvement of the gain as in [5]. The refined gain can be reinserted into (2), giving rise to an iteration where the correction $\Delta P_{\text{eq}}^{(n)}$ at iteration n is determined from the gain $G(\lambda_i, N_{\text{ch}}(\bar{P} + \Delta P_{\text{eq}}^{(n-1)}))$ of the previous iteration. $\Delta P_{\text{eq}}^{(0)} = 0$ corresponds to a preemphasis-unaware model.

In practice, measured gain profiles for a fixed current and for flat input at various total input power levels are stored in a lookup table. The gain at intermediate powers is obtained by linear regression for given current.

4. Accuracy of the preemphasis-(un)aware model

We evaluate the preemphasis-aware model for both groups using the same lookup table built from flat inputs (zero tilt) for the 14 power levels and 6 injection currents, 84 spectral loads in total. These samples are removed from the test set. Fig. 2 shows histograms of the prediction errors in the per-channel input power obtained with the preemphasis-(un)aware models for WDM inputs with (a) tilted and (b) random preemphasis. In terms of $\text{RMSE} = \sqrt{(1/N_{\text{ch}}) \sum_{i=1}^{N_{\text{ch}}} (P_i^{\text{pred,dB}} - P_i^{\text{true,dB}})^2}$, the former distribution corresponds to 0.025 dB and the latter to 0.048 dB, respectively. As expected, the distribution for random inputs is wider because of a larger preemphasis (12 dB vs. 6 dB range). The RMSE for flat/tilted inputs increases from 0.017 to 0.021 and 0.025 dB as the tilt increases from ± 1 to ± 2 and ± 3 dB. The largest errors are found on samples with the largest tilt (± 3 dB).

In the preemphasis-unaware model ($\Delta P_{\text{eq}}^{(0)} = 0$), the gain is determined by the injection current and total input power alone. This approximation is already quite accurate. Nevertheless, the correction (2) with iteration usually improves the error. For example, for tilted inputs with up to 3 dB tilt, the RMSE is improved from 0.035 to 0.025 dB. The impact is more significant for outliers: the correction decreases the 99-percentile of the error distribution from 0.15 to 0.10 dB and the maximum error from 0.34 to 0.21 dB. Typically these are the cases with high input power and largest preemphasis (largest tilt). In agreement with [5] we find that the improvements from iteration 2 and beyond are negligible.

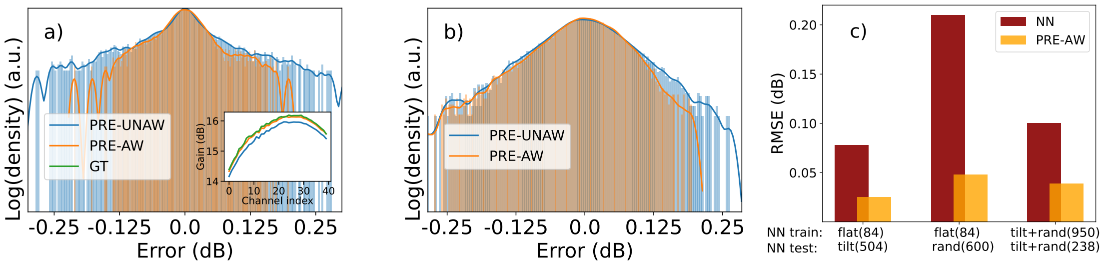

Fig. 2. Error distributions of predicted per-channel power for the preemphasis-aware (PRE-AW) and unaware (PRE-UNAW) model for (a) tilted and (b) random preemphasis. Corrections from the preemphasis-aware model are more significant for loadings with larger input power and preemphasis (or tilt). The inset shows an example of ground truth (GT) and predicted gain from the two models. (c) Comparison of prediction errors of the neural network (NN) and PRE-AW models. Number of NN train/test samples are indicated. The PRE-AW model is tested on the same data as the NNs, but trained on flat inputs only (84 samples). It requires less data while being more accurate.

It is further possible to obtain predictions for a given current I, given a model trained at a similar reference current I_{ref} [5]. Using (1), $P_{\mathrm{tot}}^{\mathrm{in}}$ of a flat WDM input at current I can be related to a flat input at $P_{\mathrm{tot}}^{\mathrm{in,ref}}$ and I^{ref} at the same population inversion x: $P_{\mathrm{tot}}^{\mathrm{in,ref}} = P_{\mathrm{tot}}^{\mathrm{in}}(I^{\mathrm{ref}} - I_0)/(I - I_0)$. I_0 can be viewed as the portion of the current that is effectively lost for signal amplification due to spontaneous recombination and generation of ASE. We can thus map the gain $G(\lambda_i, P_{\mathrm{tot}}^{\mathrm{in}}, I)$ in (2) to $G(\lambda_i, P_{\mathrm{tot}}^{\mathrm{in,ref}}, I^{\mathrm{ref}})$, which can be retrieved from the lookup table. For flat/tilted inputs, the prediction error degrades from 0.025 to about 0.05 dB for the closest measured current (100 mA apart).

5. Comparison to neural network models

To illustrate the advantage of the preemphasis-aware model, we trained generic NN models of the gain consisting of 1 up to 4 hidden layers with 32 up to 256 hidden units per layer on different subsets of the available data. The per-channel as well as total input power and the injection current serve as NN inputs.

To put the accuracy of the preemphasis-aware model into perspective, we first trained NNs using the same fixed training and test sets. Training on the 84 flat WDM inputs only and testing on flat/tilted inputs yields an RMSE of 0.078 dB for the best NN model, see Fig. 2 (c). Evaluating the same NN on inputs with random preemphasis yields even larger RMSE in excess of 0.2 dB depending on NN size. This should be compared to 0.048 dB for the preemphasis-aware model. The NNs do not discover (2), neither the approximate relation of the gain to the total input power. For most reliable predictions, the NNs should be trained with samples generated from the same data distribution as the test set. As expected, the prediction accuracy improves when training and test sets contain both flat/tilted and random inputs. In general, the deeper and wider networks tend to give more accurate results. 10-fold cross validation with random splits of the joint datasets with 80% of the 1188 samples used for training yields a best RMSE of 0.100 dB, still significantly larger than that of the preemphasis-aware model (0.039 dB).

6. Conclusion

Using a simple gain model, the SOA gain response can be obtained from knowledge of the total input power and injection current to within 0.05 dB in RMSE. The preemphasis-aware SOA model reduces the maximum error by 0.1 dB. Compared to generic NN models, the RMSE is reduced from 0.1 to 0.05 dB while requiring less than 10% of the available training data. The larger amounts of data required by NNs may not always be available. The results illustrate the advantage of incorporating physical knowledge in the preemphasis-aware approach.

References

1. X. Zhao et al. Real-time 59.2 Tb/s unrepeated transmission over 201.6 km using ultra-wideband SOA as high power booster. In *2022 European Conference on Optical Communication (ECOC)*, pages 1–4, 2022. Paper We4D.5.
2. M. J. Connelly. Wideband SOA steady-state numerical model. *IEEE J. Quantum Electron.*, 37, 2001.
3. A. Meseguer et al. Highly accurate measurement-based gain model for constant-pump edfa for non-flat wdm inputs. In *2021 Optical Fiber Communications Conference and Exhibition (OFC)*, pages 1–3, 2021. Paper M5C.4.
4. A. A. M. Saleh et al. Modeling of gain in erbium-doped fiber amplifiers. *IEEE Photon. Technol. Lett.*, 2, 1990.
5. V. M. C. et al. Mathias. Extension of the measurement-based gain model for non-flat wdm inputs and various pump currents. In *2021 European Conference on Optical Communication (ECOC)*, pages 1–4, 2021. Paper Th1A.5.