Gain-Layer-on-Waveguides for All Optical Spiking Dynamics

R. Otupiri, A. Joseph, and R. Stabile

Technische Universiteit Eindhoven, EHCI, Den Dolech 2, 5612 AZ, Eindhoven, Netherlands

Neuromorphic photonics promises to revolutionize AI by providing systems that adapt to computational demands with exceptional speed, efficiency, and precision. Traditional electronic models struggle with scalability and energy efficiency, prompting the use of bio-inspired, spiking synapses via gain-layer-on-waveguides. These utilize flexible sol-gel fabrication for on-chip amplification within spiking neural networks. Our research investigates the dynamic responses of an erbium-doped gain-layer-on-waveguide activated by a pulsed pump, focusing on non-linear pulse transformations crucial for mimicking neuronal behavior and enhancing signal propagation in SNNs.

Introduction

The rapid increase in computational and energy demands of neural networks, doubling approximately every 3.4 months [1,2], far exceeds the biennial doubling of hardware density observed under Moore's law. This gap highlights the limitations of traditional electronic hardware, struggling with bandwidth and connectivity issues as Moore's law slows[2]. Consequently, there is a significant push towards developing neuromorphic photonics, which merges optical technology with neuromorphic architectures[3], offering potential solutions with high efficiency and extensive bandwidth capabilities for computing platforms. Neural networks are mainly divided into artificial neural networks (ANNs) and spiking neural networks (SNNs) [4]. ANNs use continuous, non-linear activations to process analog inputs, while SNNs handle discrete, time-based spikes, improving computational power and capacity[5,6,7]. SNNs also have lower power needs because neurons only activate during spikes, reducing activity [1]. Current photonic spiking hardware mainly uses excitable semiconductor lasers on III-V platforms [8], suffering significant optical losses from poor light confinement and high material absorption, challenging scalability [1,6]. Recent strides in silicon-based spiking neurons have utilized phase-change materials on silicon and silicon nitride to mimic spiking behaviors via nonlinear pulse transformations. However, these systems require precise synchronization between output spikes and input data, limiting temporal encoding, optical gain, and asynchronicity, thus restricting network design flexibility [1,8]. This paper introduces an erbium-doped gain-layer on waveguide synapse utilizing optical pulse transformation for spiking. We analyze the dynamic responses of a rare-earth gain-layer activated by pulsed pumps, essential for neuromorphic computing, and explore the temporal dynamics of ions across energy levels using both traditional and innovative pulsed pump and probe techniques.

Gain-layer-on-Waveguides

The Gain-Layer-On-Waveguide Synapse (GLOWS) represents our pioneering effort to develop spiking neurons using a rare-earth gain-layer-on-waveguide. We are employing the sol-gel fabrication method to create these devices on a chip. This technique, both advanced and straightforward, provides precise doping control, low-cost production, and compatibility with various substrates, enhancing light-matter interactions crucial for next-generation photonic applications

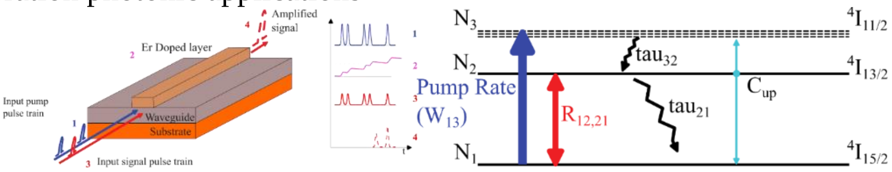

Figure 1. On the left, is a sketch of a Gain layer on waveguide node (left) and an illustration of the spike generation process; a train of input pump pulses (1-blue) is injected into the node causing a change in the state potential of the doped layer (2-puple) of the doped layer to rise above a threshold resulting in the registration of an enhance signal at the output (4 red dash) of the input probe (3- red solid). On the right, Energy-level transitions for the Er^{3+}- doped three-level system at dual wavelength transition of 980nm and 1550nm.

Our approach aims to unravel the complex interactions of spiking pump techniques and ion densities in excited states, creating event-driven responses that mimic natural neuronal functions in a gain-enhanced environment. By incorporating these thin gain-layers into photonic circuits, we enhance light-matter interactions, significantly boosting light modulation efficiency—a key factor for scaling and accurately mimicking the asynchronous spike communications seen in neural networks. The intrinsic gain properties of these layers also promote the self-sustainability of spikes, paving the way for dynamic, autonomous photonic neural networks. This integration marks a substantial advancement in neuromorphic photonics, promising revolutionary changes in how neural networks are designed and implemented.

Modelling temporal dynamics

The analysis of temporal dynamics employs an Erbium-doped (Er^{3+}) three-level model depicted in Figure 2, with a chosen dual pump and probe signal wavelengths of 980nm and 1550 nm. This wavelength selection is due to the high excited state absorption (ESA) of Erbium at 980 nm and stimulated absorption at 1550nm. Factors such as cooperative upconversion, and various spontaneous emission rates are considered for a more precise model. The functioning of the amplifier begins when Er^{3+} ions at the ground level absorb photons, exciting from the $^4I_{15/2}$ state to higher Stark split sub-levels of $^4I_{11/2}$. These ions then quickly transition to lower, metastable sub-levels due to thermal relaxation. Once the pump power surpasses the loss threshold and transparency, a population inversion is established between the $^4I_{15/2}$ and $4I_{13/2}$ levels, facilitating amplification.
The rate equations can be expressed by:

Equation 1. System of equations for a three level system of an erbium gain-layer-on-waveguide describing the population densities N_1, N_2, N_3 and the photon number S as a result of largely spontaneous emission.

$$\begin{aligned}
\frac{dN_1}{dt} &= -W_{13}N_1 + \frac{N_2}{\tau_2} + C_{\rm up}N_2^2 + W_{31}N_3 + R_{21}N_2 - R_{12}N_1 \\
\frac{dN_2}{dt} &= -\frac{N_2}{\tau_2} + \frac{N_3}{\tau_3} - 2C_{\rm up}N_2^2 + R_{12}N_1 - R_{21}N_2 \\
\frac{dN_3}{dt} &= W_{13}N_1 - \frac{N_3}{\tau_3} + \frac{N_4}{\tau_4} - W_{31}N_3 + C_{\rm up}N_2^2 \\
\frac{dS}{dt} &= (R_{21}N_2 - R_{12}N_1) - \text{loss_coefficient} \cdot S + (W_{13}N_1 - \frac{N_2}{\tau_2})
\end{aligned}$$

Here, N_1, N_2, and N_3, represent the concentrations of Er^{3+} ions in the energy states $^4I_{15/2}$, $^4I_{13/2}$, and $^4I_{11/2}$ respectively. S denotes the total photon number as result from both stimulated and spontaneous emission of transitioning Er^{3+} ions. Assuming uniform doping and optical sensitization of all Er^{3+} ions, the rate equations are evaluated under steady-state conditions.

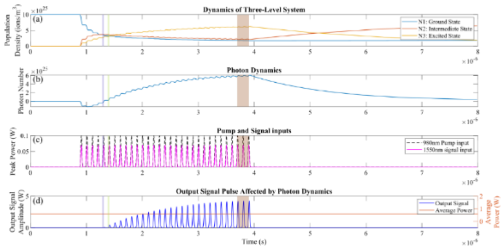

Figure 2. This figure illustrates the interplay of population densities across three energy levels, photon generation, and the modulation of input and output signal pulses within a photonic system. Fig (3a) the evolving ion concentrations in different energy states, Fig (3b) the dynamics of photon numbers reflecting emission processes, Fig (3c) the temporal profiles of pump and signal pulses, and Fig (3d) the resultant output signal strength and average power, showcasing the system's response to Gaussian pulse inputs.

Simulation results

To understand the behavior of the gain-layer-on-waveguide, we inject a series of optical pulses into the waveguide. As depicted in Figure 1(c), we introduce regular Gaussian pump and probe pulses with peak amplitudes of 0.1W and 0.08W respectively, each lasting 20ns and occurring at intervals of 100ns. Figure 1(a) illustrates the resulting dynamics within the three-level ion system, with color-coded regions—blue for transparency, green for inversion, and peach for gain saturation—showing the various states.

Table 1. Simulation parameters used in this work

Description	Parameter & value	Description	Parameter & value
Signal emission cross-section (m^2)	σ = 6.8e-22	$^4I_{13/2}$ state lifetime (s)	τ = 3.6e-6
Signal absorption cross-section (m^2)	σ = 5.8e-22	$^4I_{11/2}$ state lifetime (s)	τ = 7.8e-3
Pump emission cross-section (m^2)	σ = 0	Waveguide loss/ Photon loss coefficient (s^-1)	07e5
Pump absorption cross-section (m^2)	σ = 2.58e-22	Channel waveguide core size (m^2)	25e-12
Cooperative up conversion coefficient (m^3/s)	4.1e-23	Total concentration of ions (ions/m^3)	N_0 = 1e26

Figure 1(b) tracks the evolution of the net photon count, which includes photons from both stimulated and spontaneous emissions as the ions transition between levels. This visualization allows us to pinpoint the precise moments of gain saturation, achieve transparency, and reach inversion. Notably, following inversion, there is a noticeable enhancement of the input pulsed signals, providing real-time insight into the complex dynamics within the waveguide. The parameters used in these simulations are shown in Table 1. The noticeable enhancement of the input pulsed signals, is better illustrated in Fig.3(d); here we initially see the absence of any output pulsed signals prior to inversion, a single output pulsed signal at transparency and enhanced signal pulses post inversion.

Furthermore, a pivotal feature of the Gain-Layer-On-Waveguide that is particularly advantageous for neuro-inspired spiking applications is its short-term memory capability. This property allows the waveguide to temporarily maintain its state potential, which can be utilized in the computation and processing of information, similar to how neural synapses function in the brain. Our study, illustrated in Figure 5, delves into this by highlighting the time window during which the waveguide sustains a state of population inversion.

In Figure 5 (panels a1-a4), we introduce two sequences of input pulse trains. The initial pulse train surpasses the transparency and inversion thresholds, setting the stage for the subsequent train. This second set of pulses, falling within the decay envelope of the first, immediately experiences signal enhancement. Conversely, in Figure 5 (panels b1-b4), pulses outside the designated memory window require additional input to first overcome the thresholds before any signal enhancement is observed. This disparity in timing for the observation of an amplified spike underscores the waveguide's capability to differentially respond based on timing, aligning with the Time to First Spike (TTFS) encoding technique we intend to employ in our systems. This approach provides a nuanced insight into how GLOWS can effectively mimic neural processing dynamics.

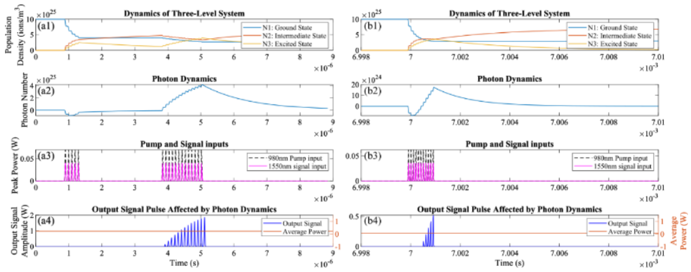

Figure 3. This figure illustrates the critical time window of population inversion within the GLOWS, demonstrating its short-term memory capabilities. Panels a1-a4 demonstrate the enhancement of signal strength following two pulse trains within the decay envelope, showing immediate signal amplification. Panels b1-b4 depict the requirement of additional pulses to achieve inversion and subsequent signal enhancement outside the memory window, illustrating the critical timing for effective neuromorphic encoding based on the Time to First Spike (TTFS) technique.

This memory window corresponds to the upper state life time of erbium ions (~10ms). During this period, there is a predominant ion population in the upper energy levels. By understanding and optimizing this inversion duration, we aim to enhance the waveguide's functionality in neuromorphic computing systems, where rapid and transient information processing is crucial.

Conclusion

In conclusion, the Erbium-doped Gain-Layer-On-Waveguide stands out as a prime choice for neuromorphic computing. It mimics neuronal functions with features like thresholding, temporary memory, asynchronous spiking, and time-encoded signaling with the added benefit of signal amplification, overcoming the inefficiencies of electronic neural hardware. This technology not only aligns with biological neural processes but also holds promise for advancing neural network design, offering a powerful solution for energy-efficient, high-speed computing.

References

[1] Jha, Aashu, Chaoran Huang, Hsuan-Tung Peng, Bhavin Shastri, and Paul R. Prucnal. "Photonic spiking neural networks and graphene-on-silicon spiking neurons." *Journal of Lightwave Technology* 40, no. 9 (2022): 2901-2914.

[2] Amodei, Dario, Danny Hernandez, Girish Sastry, Jack Clark, Greg Brockman, and Ilya Sutskever. "AI and Compute." (2018).

[3] Nahmias, Mitchell A., Alexander N. Tait, Bhavin J. Shastri, Thomas Ferreira De Lima, and Paul R. Prucnal. "Excitable laser processing network node in hybrid silicon: analysis and simulation." *Optics express* 23, no. 20 (2015): 26800-26813.

[4] Maass, Wolfgang. "Networks of spiking neurons: the third generation of neural network models." *Neural networks* 10, no. 9 (1997): 1659-1671.

[5] Wu, Yujie, Lei Deng, Guoqi Li, Jun Zhu, and Luping Shi. "Spatio-temporal backpropagation for training high-performance spiking neural networks." *Frontiers in neuroscience* 12 (2018): 331.

[6] Maass, Wolfgang. "Lower bounds for the computational power of networks of spiking neurons." *Neural computation* 8, no. 1 (1996): 1-40.

[7] Maass, Wolfgang, and Henry Markram. "On the computational power of circuits of spiking neurons." *Journal of computer and system sciences* 69, no. 4 (2004): 593-616.

[8] Peng, Hsuan-Tung, Gerasimos Angelatos, Thomas Ferreira de Lima, Mitchell A. Nahmias, Alexander N. Tait, Siamak Abbaslou, Bhavin J. Shastri, and Paul R. Prucnal. "Temporal information processing with an integrated laser neuron." *IEEE Journal of Selected Topics in Quantum Electronics* 26, no. 1 (2019): 1-9.