In the not-too-distant future, advances in machine learning will spur a new, transformative generation of optical communication and measurement systems.
In a celebrated five-game match in March 2016, AlphaGo—an artificial-intelligence (AI) system from the Google subsidiary DeepMind—decisively defeated the 18-time world champion Lee Sedol in the ancient game of Go, a strategic board game with significantly more combinations of moves than chess. Successor programs to AlphaGo have built on that record, and one of them, AlphaZero, is now considered the strongest player in the complex game’s history. (Indeed, in November 2019, Lee Sedol announced that, in light of AI’s increasing dominance in Go, he was retiring from professional play.)

The well-publicized story of AlphaGo highlights how machine learning, once a niche field within computer science, is going mainstream, and is transforming a range of scientific disciplines and industries. Examples abound well outside of Go—an AI system that manages data-center cooling and helps saves 30% of the energy needed; discoveries of new planets and materials; re-creating Nobel Prize-winning experiments in ways that take a fraction of the time of the original feats.

These and other cases illustrate machine learning’s power to find solutions to highly complex problems in unconventional ways. That power will prove increasingly important with the world’s growing focus on climate action, environment, energy and resource efficiency—all areas in which photonics could play a significant role. And machine learning itself is spurring a new generation of intelligent photonic systems in communications and other areas.

The need for intelligence

The advent of intelligent photonic systems is well timed, as the complexity and performance requirements of optical communication systems and networks will increase substantially in the next 10 to 15 years. Although there has been tremendous progress in terms of signaling baud-rate, flexible channel spacing, modulation formats and coding schemes, the current technology’s linear evolution cannot satisfy future data capacity demands. One viable long-term solution would be to employ systems with an ultra-wideband transmission window (encompassing the O+E+S+C+L bands), coupled with the use of space-division multiplexing (SDM) via multicore and multimode fibers.

The high complexity of these systems will make designing optimal signaling and detection schemes for them challenging. Moreover, optimizing amplification schemes—a key enabler for long-distance transmission—requires fast tuning of a very large number of parameters, as arbitrary gain profiles will be highly desired. And optimizing the increasingly

Where machine learning has an edge

Machine learning could well give rise to a new generation of intelligent photonic systems. But just what is machine learning really good at?

In general, machine learning is good at learning tasks from a given data set without being explicitly programmed. A number of different learning frameworks can guide these approaches:

**SUPERVISED LEARNING**

comprises tasks that involve finding mappings (dependencies) between a system’s inputs and outputs.

**UNSUPERVISED LEARNING**

focuses on learning time-varying dynamics from measured noisy data.

**REINFORCEMENT LEARNING**

can produce fully autonomous agents that interact with the environment to learn optimal behavior that improves with time.

Machine learning can therefore bring significant advantages when the relationships between a system’s inputs and outputs are complex and intractable—or, to express it differently, when no model exists that can describe the relation between the inputs and the outputs.
Machine learning is spurring a new generation of intelligent photonic systems in communications and other areas. And the advent of such systems is well timed.

Complex system’s energy efficiency, traffic routing, channel power and bandwidth allocation, as well as modulation format selection, will become difficult using standard tools that rely on analytical or semi-analytical models. The overall focus on minimizing optical networks’ power consumption, in watts per bit, will require rethinking optical communication systems and network design.

Complicating matters still further will be the increasing future focus on quantum information security, which requires coexistence and management of classical and quantum channels in the same optical network. As quantum signals generally have significantly lower power than classical signals, reception of the quantum signals is more challenging—strengthening the case for intelligent optical receivers that can receive and distinguish between classical and quantum signals.

**Network, measurement, sources**

Beyond the advantages that machine learning might bring to optical communication systems and networks, another great opportunity lies in introducing some degree of intelligence into optical measurement systems. Current optical signal analyzers, for instance, cannot distinguish between different signal impairments, and cannot determine whether the impairments originate from the transmission channel or the components themselves. The ability of signal analyzers to distinguish between different nonlinear impairments such as self-phase modulation (SPM), cross-phase modulation, inter-channel four-wave mixing (IFWM) and stimulated Raman scattering (SRS) would foster the design of more effective transmitter and receiver signal-processing algorithms.

Moreover, as future optical communication systems are pushed to work closer to their theoretical limits, the noise properties of laser and frequency-comb sources will need to steadily improve. Optical frequency combs (OFCs) could play a significant role in next-generation high-speed optical and optical–wireless communication systems, due to OFCs’ ability to provide frequency-stable and low-phase-noise comb lines from a single optical source.

Furthermore, as the baud rate of optical communication systems approaches 100 GHz, noise characterizations in the GHz range will be valuable. Current characterization methods of static parameters (such as correlation matrix, linewidth enhancement factor, carrier lifetime and resonance frequency) and dynamic parameters (such as amplitude and phase noise and timing jitter) are quite cumbersome, requiring multiple setups, and are currently limited to a couple of MHz. Most important, the ability to intelligently explore the knowledge obtained from such characterizations, and use them to enable improved system design, will become extremely valuable.

All of these challenges will require large amounts of effort in which the combination of domain knowledge and machine learning will be crucial. Luckily, an increasing number of research groups and activities are exploring how machine learning can be applied in optical communication systems and networks, each with its own view on the technique’s potential impact. The topic is gaining more and more visibility at leading conferences and in journal special issues—clear signs of an important, emerging cross-disciplinary topic.

In the rest of this feature, we offer our view of some of the machine-learning techniques that we believe may have an impact for next-generation optical communication and measurement systems.

**CASE 1: Ultra-wideband optical amplification**

As noted above, one route to satisfying future bandwidth demands is to widen the wavelength window to cover all five transmission bands. But designing an optical amplification scheme covering all five bands is a challenging task. Solutions now being explored include semiconductor optical amplifiers (SOAs), bismuth-doped fiber amplifiers (BDFAs) and Raman amplifiers (RAs). An important property of future optical amplifiers will be the ability to provide arbitrary gain profiles. This is because the power profile that maximizes the achievable information rate may not be flat (due to nonlinear interactions between sub-bands), and also because arbitrary gain profiles can be used to compensate for
some less-than-ideal properties of optical add-drop multiplexers. The latter would lead to an improved power budget and thereby to energy efficiency, as gain-flattening filters can be avoided.

Optical amplifiers based on the Raman effect offer low-noise properties due to distributed amplification, and gain availability across the entire multi-band when operated in a multi-pump configurations. Most important, RAs offer great flexibility in the gain profile design through adjustments in pump powers and wavelengths, making them an attractive solution for future ultra-wideband optical communication systems. And RAs can be used in hybrid configuration schemes in combination with SOAs or BDFAs to realize certain gain profiles.

The main challenge with RA design is the selection of pump powers and wavelengths that would result in a specific gain profile. This is a typical inverse-system design (ISD) problem, in which one is given a target output \( Y_{\text{targ}} \) and asked to determine the corresponding input \( X_{\text{targ}} \). Solving the problem is challenging because of the nonlinear relationship between Raman gain and pump power and wavelength. And future optical networks will require fast reconfiguration—which translates into fast adjustment of amplifier gains.

Machine-learning techniques, such as auto-encoders that employ multi-layer neural networks, can be used to address the ISD problem, and are already being explored in the design of photonic integrated circuits. We have recently proposed a machine-learning-based framework for ISD and used it for optimizing Raman pump powers and wavelengths, targeting arbitrary gain profiles in C+L band.

ISD using machine learning begins with a data set—achieved, in an experimental environment, by exciting the system with a set of inputs (pump powers and wavelength) and measuring the desired outputs (gain profiles). Once the data have been recorded, a multilayer neural network is trained to learn the inverse mapping between Raman gain and pump powers and wavelengths. After this training set, the targeted gain profile can be presented to an inverse-mapping multilayer neural network to obtain an ultrafast prediction of pump powers and wavelength, performed via matrix multiplications in the network.

If the inverse-mapping multilayer neural network isn’t sufficiently accurate, additional fine optimization of pump powers and wavelengths can be employed. The fine optimization relies on a gradient descent and a forward multilayer neural network that has learned the mapping between pump powers and wavelength and Raman gain profiles. The power of machine learning is that the differential equations describing relationships between pump powers and wavelength and Raman gain is replaced with a network consisting only of matrix multiplications. That, in turn, allows for quick computation of the gradients and optimization of the system—something not possible if differential equations are directly included in the optimizer.

We were able to use this framework for pump power and wavelength allocation that would result in a desired arbitrary gain profile of the Raman amplifier operating in C and C+L bands. The system resulted in a low mean (0.46 and 0.35 dB) and standard deviation (0.20 and 0.17 dB) of the maximum error in numerical

---

**Designing a machine-learning-based inverse system**

1. **The problem**
   A highly complex physical system relates an output, \( Y \), to a given input, \( X \). We would like to determine the input that will result in a specific, targeted output, \( T \).

2. **Learn inverse mapping from output \( Y \) to input \( X \)**
   The nodes in the multilayer neural net perform successive weighted matrix multiplications on \( Y \) to infer \( X \) with increasing precision (left). Once the network is trained, an arbitrary output \( T \) can be mapped to a specific input.
Machine-learning techniques, such as auto-encoders that employ multi-layer neural networks, are already being explored in the design of photonic integrated circuits.

calculations for the C+L band and in experimental results for the C band, respectively.

**CASE 2: Communicating in nonlinear channels**

In linear communication channels, well-established techniques exist for designing optimal signaling and detection schemes that maximize the achievable information rate (AIR). But the fiber optic channel is nonlinear, and maximizing AIR requires operating in a nonlinear regime, making the techniques for linear channels ineffective. As systems move increasingly toward ultra-wideband transmission, nonlinear interactions between different sub-bands will come both from Kerr nonlinearities and from stimulated Raman scattering (SRS), and will cause nonlinear interactions between different sub-bands. The combination will result in a highly complex transmission environment, and will make designing strategies for optimal signaling and detection a nontrivial problem.

One approach for overcoming those problems is to learn directly from the input-output data. The idea is to use multilayer neural networks, at both the transmitter and the receiver, to represent an encoder and decoder, respectively. The transmitter and the receiver neural networks are jointly optimized to learn the encoding and decoding strategies that would minimize the error between the transmitted and received bits, and thus to maximize AIR. This approach, called end-to-end (E2E) learning, is being actively explored in optical communication systems. A significant advantage of E2E learning is that it is agnostic to the channel model and can therefore be useful for ultra-wideband optical communication systems.

To make this approach practical, one challenge that needs to be solved involves the training of the transmitter and receiver multilayer neural networks. Current training methods, which heavily rely on backpropagation of gradients, require that the channel model is differentiable. This is quite cumbersome for designing optimal signaling and detection strategies for ultra-wideband transmission, as it would require computing gradients through the nonlinear Schrödinger equation. Therefore, the state-of-the-art results presented so far use an approximate channel model that is differentiable. Meanwhile, the deficiencies of this approach are also spurring exploration of novel research directions that do not require gradient computation for training of multilayer neural networks.

We have used E2E learning to determine geometrically shaped signal constellations, including bit mapping, that are robust to optical-fiber channel nonlinearities (specifically in the C band). The learned constellations yielded an improved performance
compared with state-of-the-art geometrically shaped constellations, with up to up to 0.13 bit/4D in simulations and 0.12 bit/4D in experiments.

**CASE 3: Characterizing lasers and frequency combs**

As with the previous cases, using machine learning to incorporate intelligence into laser and frequency comb characterization begins by gathering data. This can be achieved by heterodyning the laser or frequency comb under test with the local oscillator (LO) laser or comb; the beat signal is then detected using a balanced receiver and an analog-to-digital converter, and the samples are stored for offline data analysis.

Advances in optical communication systems have enabled balanced receivers operating with up to 100 GHz of analog electrical bandwidth, with sampling rates in the range of 160 GS/s. This is opening up opportunities for machine-learning techniques to enable intelligent ultra-broadband characterization of lasers and frequency combs.

In such an approach, the recorded data represent a sequential time series containing all the necessary information about static and dynamic parameters of interest. The observable data include latent or hidden variables that constitute the dynamic and static parameters of the laser and frequency comb under test. Therefore, parameters of interest, such as amplitude and phase noise, and their subsequent correlation matrices are not directly observable, but need to be learned from the data in which they’re hidden. Complicating this task is the presence of noise that originates from the measurement equipment itself, as well as fundamental (quantum) noise.

Machine-learning methods that rely on Bayesian inference in state-space models, such as Bayesian filtering, offer a powerful approach for learning laser and frequency comb static and dynamic parameters from the measurement data. This approach allows for inferring hidden parameters from observable data that may be sparse and noisy. Bayesian filtering offers statistically optimum and, thereby, the theoretically most accurate estimation of static and dynamic parameters.

Our group has shown that these techniques can yield phase noise measurement of record sensitivity. We were able to measure optical phase noise from signal power of –75 dBm (SNR of –11 dB in 1.1 GHz receiver bandwidth). Practically, that meant that the phase noise measurements were, to a high degree, not limited by the measurement noise floor.
Given the complexity of next-generation optical communication systems, we believe that machine learning will play a central role in enabling spectrally and power-efficient data transport.

Turn allowed measurements down to –200 dB rad²/Hz and up to 10 GHz, which is useful when measuring the Schawlow–Townes (quantum-noise-limited) laser linewidth. Typical, state-of-the-art phase noise measurement equipment can only measure phase noise down to –140 dB rad²/Hz and is limited to a couple of MHz.

The principle behind the laser phase noise measurements can also be expanded to frequency comb phase noise and correlation matrix characterization. Indeed, the framework of Bayesian filtering is well suited for frequency comb characterization, as it allows for joint tracking of phase noise of multiple frequency lines and, thus, for accurate estimation of phase correlation matrix—outperforming standard methods. More generally, the same framework can also be modified to include learning of relative intensity noise (RIN). In this way, laser and frequency comb characterization of RIN, phase noise and noise correlation matrix can be performed from a single measurement.

Vision for the future

Given the complexity of next-generation optical communication systems, we believe that machine learning will play a central role in enabling spectrally and power-efficient data transport. In this feature, we have covered mainly applications of machine learning for optimizing the physical layer. Yet great progress has been made in using machine learning in higher networking layers for network planning, failure prediction and optical performance monitoring.

Having intelligence in all layers will afford a unique opportunity to perform optimization across multiple dimensions. This will provide a path towards intelligent optical networks that are autonomous, self-healing, and able to predict traffic demands and operate with maximum energy efficiency. Those networks will be able not only to provide high data rates, but also to support U.N. sustainability goals by realizing green communication infrastructure.

Reaching these goals, though, will require technological progress in several directions. In particular, a new workforce will need to be trained that can master both photonics in general and machine learning and optical communication in particular. Machine-learning mathematics, programming and algorithm training constitute a powerful skill set—but understanding the underlying physics of the problem, and being able to build and debug experimental setups to gather the necessary data, are just as important. To handle both sides of the problem, the research community will need to increase its investment in creating cross-disciplinary education programs that combine telecommunication or photonics engineering with machine learning.

References and Resources