Nano Communication Networks 25 (2020) 100312

Contents lists available at ScienceDirect

Nano Communication Networks

journal homepage: www.elsevier.com/locate/nanocomnet

Machine learning enables design of on-chip integrated silicon T-junctions with footprint of 1.2 μm× 1.2 μm

Sourangsu Banerji [a], Apratim Majumder [a], Alexander Hamrick [b], Rajesh Menon [a], Berardi Sensale-Rodriguez [a,*]

[a] Department of Electrical and Computer Engineering, The University of Utah, Salt Lake City, UT 84112, USA
[b] School of Computing, The University of Utah, Salt Lake City, UT 84112, USA

ARTICLE INFO

Article history:
Received 19 March 2020
Received in revised form 16 June 2020
Accepted 1 July 2020
Available online 10 July 2020

Keywords:
Nanophotonics
Silicon photonics
Machine learning

ABSTRACT

To date, various optimization algorithms have been employed to design and improve the performance of nanophotonic structures. Here, we propose to utilize a machine-learning algorithm viz. binary-Additive Reinforcement Learning Algorithm (b-ARLA) coupled with finite-difference time-domain (FDTD) simulations to design ultra-compact and efficient on-chip integrated nanophotonic 50:50 beam splitters (T-junctions). The T-junctions reported in this paper have a footprint of only 1.2 μm × 1.2 μm. To the best of the authors' knowledge, these designs are amongst the smallest ever reported to date across either simulations or experiments. For all the designs, the simulated net power transmission efficiency is ∼80%, corresponding to insertion loss < 1 dB, at λ = 1.55 μm. We envision that the design methodology, as reported herein, would be useful in general for designing any efficient integrated-photonic device for optical communications systems.

1. Introduction

In the last decade, the field of silicon nanophotonics has witnessed major breakthroughs [1]. The critical enabler for its unprecedented success can be attributed to the development of advanced foundry services [1,2]. In addition to this, nanophotonic designs have been demonstrated with very efficient structures that can be implemented to obtain favourable characteristics like high sensitivity, low-loss, and high index contrast in dielectric distribution [3–5]. The integration of such all-dielectric passive nanophotonic components such as multiplexing couplers, waveguides, and so on with active devices such as lasers, LEDs, etc. onto a single chip will ultimately lead to the miniaturization of optical circuits with high data processing capability, very similar to what we see in silicon chips used for integrated electronics as of today. However, contrary to electronic circuits, there is still a lack of effective design methodologies in nanophotonics [6,7].

Traditional nanophotonic design strategies are based upon theoretical and scientific intuitions [8–11]. However, most of the time, it does not provide analytical solutions for complex nanophotonic structures and light manipulation behaviour [12]. In addition to this, device designs based on analytical methods may also not satisfy performance requirements like compactness, efficiency, bandwidth, and power transmission. For this reason, a wide variety of numerical approaches such as evolutionary algorithm [13], objective-first inverse-design algorithm [14–17], topology optimization [18], nonlinear-search algorithm [19–23], and direct-binary-search algorithm [24,25] have been implemented to design integrated-nanophotonic structures. Amongst all, inverse optimization, or objective-first inverse-design algorithms have been shown to deliver the best performing nanophotonic structures with adequate computational trade-offs [14–17,26–32]. From such a perspective, we can see that inverse-design algorithms (one such example is the adjoint method) are highly suitable for developing next-generation, compact nanophotonic devices with novel functionalities and features.

Machine learning has also recently emerged and attracted a great deal of attention from both academia and industry alike as a viable design methodology. In all areas of physics itself, ranging from gravitational wave analysis [33], to materials designs [34,35], to phase transitions in quantum physics [36,37], machine learning have successfully been leveraged to provide for performance comparable to some of the most advanced design methodologies in a natural and straightforward manner. In all these previous examples, we observe that the advantage of machine learning lies in the accurate modelling and characterization of complex relationships within the underlying systems. To summarize, the advantages of machine learning are four-fold. First, machine learning algorithms allow for *hardware parallelization*. For example, if we consider popular evolutionary algorithms, we will observe that they heavily depend on two important

* Corresponding author.
E-mail address: berardi.sensale@utah.edu (B. Sensale-Rodriguez).

https://doi.org/10.1016/j.nancom.2020.100312
1878-7789/

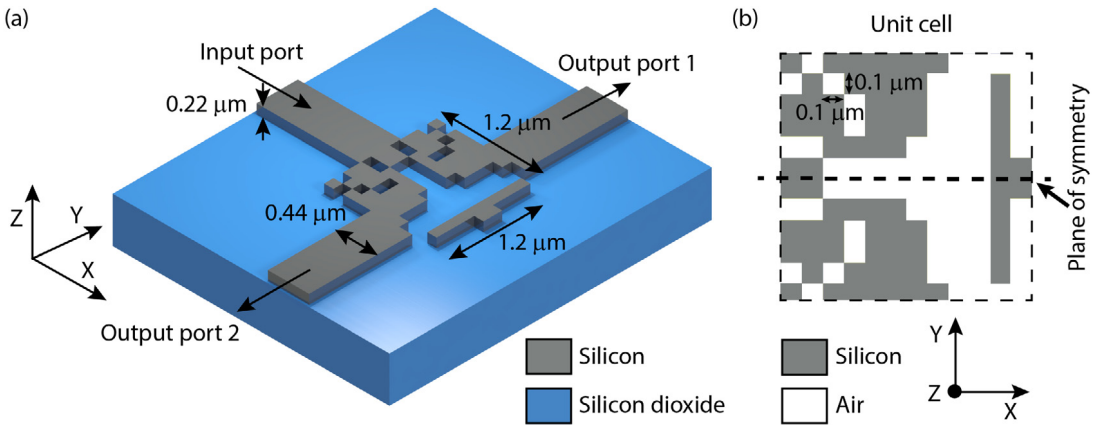

Fig. 1. Schematic of a pixelated nanophotonic structure. (a) 3D representation of the T-junction splitter. (b) Top view of the unit cell.

factors: (a) the number of generations, and (b) diversity of initial solutions. In fact, their computational complexity rises with each necessary operation like reproduction, mutation, recombination, selection, and survival of the fittest [38]. Combination of these operations adversely affects the performance, probabilistic transition, and convergence for these algorithms.

In contrast, machine-learning algorithms, even the one described in this paper, do not require any operation of such sort. The algorithm described in this paper is fully parallelizable. The training phase can be divided, and the simulations can be distributed arbitrarily across multiple computers, i.e., the data is generated and evaluated independently of each other [12,38]. *This is true not just in case of genetic algorithms only but for all class of optimization algorithms, including inverse optimization or even topology optimization. By transitioning from an optimization-based design methodology to a prediction-based one, we gain computational advantage via hardware parallelization* [39].

Second, machine learning also does not depend on the quality of initial solutions to guarantee reasonable solutions. Third, in contrast to an inverse-design algorithm like the adjoint method, machine learning can solve the forward design problem much faster with a neural network (deep learning). Even though this advantage is not big enough when comparing against the adjoint method, which requires only two forward simulations for the entire optimization, machine learning still has a marginal advantage in the sense that one can restrict the design space to manufacturable devices and physical solutions, which are harder to find with adjoint methods. Fourth, in machine learning (especially in deep learning), the model is trained to "***intelligently learn***" the non-linear relationships between the input and output over a large dataset. The model in this way can also "***intelligently learn***", for example, Maxwell's equations and solve them, without explicitly knowing about them. This allows for possible discovery of solutions outside of the boundaries of the training data, and the ability to transfer knowledge between problems by a method known as "*transfer learning*". This approach represents a complete paradigm shift in thinking of how nanophotonics research has been understood till date and what it could lead to in the time to come; to enable equally disruptive series of novel findings in nanophotonics.

Out of various machine learning algorithms, there are many advantages of using reinforcement learning algorithms. First, reinforcement learning algorithms are generally used to solve very complex problems that cannot be solved by conventional techniques therefore it can achieve long-term results and can outperform humans in many tasks (e.g. Go strategy game [40]). Second, it can correct for the errors that occurred during the training process; therefore, once an error is corrected by the model, the chances of occurring the same error are very less [41,42]. Third, in the absence of a training dataset, it is bound to learn from its experience, hence, it is intended to achieve the ideal behaviour (similar to a real human being) within a specific context, to maximize its performance [43]. Finally, reinforcement learning algorithms strikes a balance between exploration and exploitation. Exploration is the process of trying different things to check if they are better than what has been tried before whereas, exploitation is the process of trying the things that have worked best in the past. Other learning algorithms do not perform this balance [41,42].

Considering the advantages stated above, researchers working in the field of optics and photonics have started harnessing machine learning to develop foundry compatible optical components for large scale industrial rollout [39,44–49]. In this work, we utilized a machine learning algorithm, namely binary-Additive Reinforcement Learning Algorithm (b-ARLA) coupled with a finite-difference time-domain (FDTD) method to demonstrate efficient and ultra-compact 50:50 beam splitters (T-junctions) as shown in Fig. 1(a). The top view of the "unit cell", as shown in Fig. 1(b) consists of square sub-unit pixels of either silicon or air. Square pixels have an advantage from the fabrication point-of-view. Although electron-beam-lithography can be used to fabricate the structures presented in this paper (minimum feature size $\sim$ 100 nm), for large scale fabrication of a large number of devices on wafers, typically the industry standard is to rely on immersion projection lithography. It is well-known that optical proximity correction is required to ensure correct size and shape of the features [50]. Since the optical lithography industry has for a long time perfected this process for square and rectangular features which are the most abundant geometrical shapes present in ICs, we believe that our structures can be readily adapted to multi-wafer processes. Hence, square pixels have an advantage during fabrication. In addition to this, circular pixels, i.e. holes may also be used to design similar digital metamaterial devices, as have been demonstrated elsewhere [51].

Finally, Table 1 provides a more detailed analysis of power splitters reported in the literature to date and compares it with the work as described in this paper. A few noteworthy implementations of ultra-compact Y- or T-junction 50:50 power splitters reported in the scientific literature have had area footprints > 2 μm^2 with < 1 dB insertion loss at the telecom wavelength of 1.55 μm [49,52–59]. Therefore, it is evident that with a footprint of only 1.2 $\mu m \times 1.2$ μm, the designs reported in this paper are amongst the smallest ever reported across either simulations or experiments as of date.

Table 1
Summary of symmetric and asymmetric nanophotonic power splitters.

Reference	Splitter shape	Imbalance (split ratio)	Insertion loss	Bandwidth (centre wavelength = 1550 nm)	Polarization	Footprint
Tahersima et al. [45]	Y	50:50	–	200 nm	TE	6.76 μm^2 (2.6 $\mu m \times 2.6$ μm)
Tahersima et al. [45]	Y	25:75	–	200 nm	TE	6.76 μm^2 (2.6 $\mu m \times 2.6$ μm)
Zhang et al. [52]	Y	50:50	0.13 dB	~80 nm	TE	2.4 μm^2 (1.2 $\mu m \times 2$ μm)
Kurt et al. [53]	T	50:50	<0.3 dB	~200 nm	TM	16.77 μm^2 (4.1 $\mu m \times 4.1$ μm)
Alpkilic et al. [54]	T	100:0	<0.3 dB	–	TM	7.84 μm^2 (2.8 $\mu m \times 2.8$ μm)
Alpkilic et al. [54]	T	100:0	0.73 dB	–	TE	7.84 μm^2 (2.8 $\mu m \times 2.8$ μm)
Xu et al. [55]	Y	50:50	~1 dB	~30 nm	–	13 μm^2 (3.6 $\mu m \times 3.6$ μm)
Xu et al. [55]	Y	40:60	~1 dB	~30 nm	–	13 μm^2 (3.6 $\mu m \times 3.6$ μm)
Xu et al. [55]	Y	25:75	~1 dB	~30 nm	–	13 μm^2 (3.6 $\mu m \times 3.6$ μm)
Lin et al. [56]	Y	50:50	0.36 dB	~100 nm	TE	3.2 μm^2 (1.4 $\mu m \times 2.3$ μm)
Ren et al. [57]	Y	50:50	0.33 dB	~40 nm	TE	7.2 μm^2 (2.4 $\mu m \times 3$ μm)
Xie et al. [58]	Y	50:50	<1.5 dB	~40 nm	TE	18 μm^2 (4 $\mu m \times 4.5$ μm)
Chang et al. [59]	Y	50:50	<1.5 dB	~80 nm	TE	8.3 μm^2 (2.9 $\mu m \times 2.9$ μm)
This work	T	50:50	0.86 dB	~40 nm	TE	1.44 μm^2 (1.2 $\mu m \times 1.2$ μm)
This work	T	50:50	0.95 dB	~40 nm	TE	1.44 μm^2 (1.2 $\mu m \times 1.2$ μm)

2. Design and optimization

The perceptron-like machine learning algorithm [55] used in this study is implemented to reduce the insertion loss (splitting the input power with minimal loss) of the power splitter at an operating wavelength of 1.55 μm. The algorithm developed herein combines both the "*additive updates*" feature of a perceptron algorithm [54,60] as well as the "*reward for state idea*" of reinforcement learning [60]. The flowchart of the algorithm is shown in Fig. 2, which depicts that the algorithm consists of two phases: training and inference.

2.1. Training phase

The training phase starts with creating a photonic structure where each constituent sub-unit of the "unit cell" is randomly distributed. Essentially, the entire "unit cell" consists of 12 x 12 randomly distribution pixels. However, due to the inherent symmetry of the structure itself, i.e., 50:50 split ratio, the generated random structure now consists of 12 x 6 binary pixels (flip symmetry across the *y*-direction), where "1" denotes the high refractive index Si-sub-units and low refractive index air-sub-units are represented by "0" in the binary "unit cell". Keeping in mind the capability of current fabrication technologies, the design parameters of each sub-unit within the "unit cell" are taken very conservatively (e.g., 100 nm minimum features). Therefore, the initial design parameters are fixed as follows: square-shaped sub-units with a size of 0.1 $\mu m \times 0.1$ μm, the height of the structure is 0.22 μm (typical in SOI), and the material refractive index is $n_{Si} = 3.46$. The refractive index of air is $n_{air} = 1$. Therefore, the size of the complete photonic structure is 1.2 $\mu m \times 1.2$ $\mu m \times 0.22$ μm. The size of the input and output waveguides is fixed at 1 μm (length) $\times$ 0.44 $\mu m \times 0.22$ μm (cross-section). The substrate thickness is assumed to be 0.6 μm. Following the creation of the "unit cell" structure along with the input and the output waveguide ports, a 2.5D varFDTD (variational FDTD) method [61] is incorporated to analyse the time-domain response of the photonic structure at $\lambda = 1.55$ μm. Common desktop CPUs were employed to perform the 2.5D varFDTD simulations (Lumerical Inc.'s Lumerical MODE solutions). A more elaborate description of the full-wave simulation is provided in the following section. While extracting the response of the structure, the insertion loss (in dB) was extracted for the nanophotonic structure. The rationale behind considering the insertion loss as the metric for optimization is inspired by the fact that photonic integrated circuit designers would typically be interested in the total amount of light that is transmitted into the fundamental mode. To accommodate these requirements, the insertion loss is defined as:

$$\text{Insertion Loss (I. L.)} = -10\log_{10}(\text{mean}(T_{net})) \quad (1)$$

where T_{net} is the net transmission into the fundamental mode of the output waveguides. The difference between the numerical value of the insertion loss in the worst possible scenario and the insertion loss extracted from the steady state response of the *i*th randomly generated photonic layout is defined as the reward function R_i, which is then defined as follows:

$$R_i = |\text{I.L.}_{worst} - \text{I.L.}_i| \quad (2)$$

In the above expression, I.L._{worst} as stated earlier of the insertion loss in the worst possible scenario, which essentially refers to two different scenarios: a binary "unit cell" all of whose individual sub-unit elements are either (a) "0" (air) or (b) "1" (silicon). The reward function R_i approaches to I.L._{worst} only when $\text{I.L.}_i \sim 0$, i.e., when any of the *i*th iteration results in a structure where the splitting of the input power is ideally lossless.

The modulus of the reward function R_i generates a single valued positive number which is then multiplied with the *i*th binary square matrix (B_i). One can think of this step as assigning a

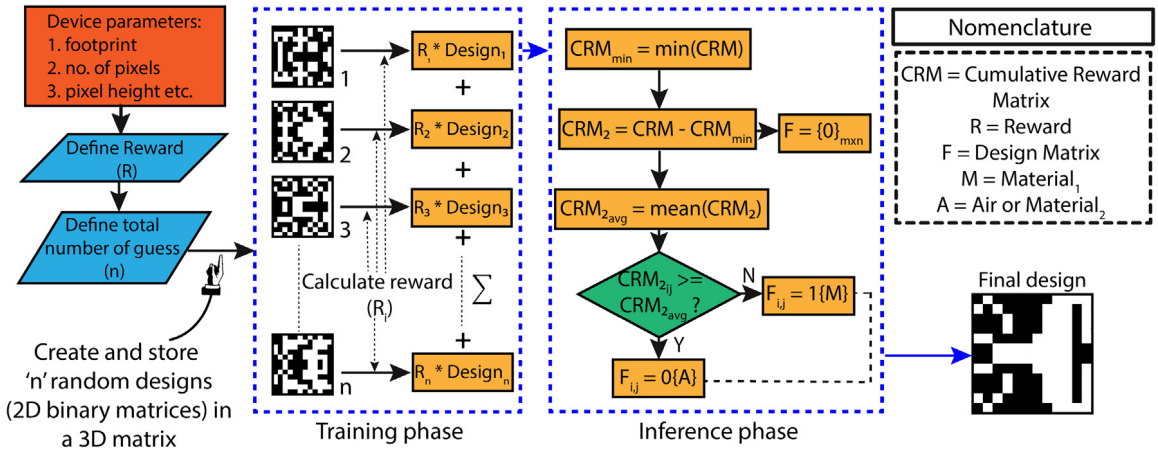

Fig. 2. Flow diagram of the binary-Additive Reinforcement Learning Algorithm (b-ARLA).

"*confidence*" score to the all the pixels within B_i. Mathematically, this can be expressed as,

$$R_i \times \begin{bmatrix} a_{11} & \cdots & a_{j1} \\ \vdots & \ddots & \vdots \\ a_{1k} & \cdots & a_{jk} \end{bmatrix} \tag{3}$$

where, $B_i = \begin{bmatrix} a_{11} & \cdots & a_{j1} \\ \vdots & \ddots & \vdots \\ a_{1k} & \cdots & a_{jk} \end{bmatrix}$, such that $(a_{11}, \ldots, a_{j1}, \ldots, a_{1k}, \ldots, a_{jk}) \in (0, 1)$. For the sake of this problem, we consider $j = k$. This will simplify to a single matrix, say reward matrix RM_i since R_i is a scalar as shown below,

$$RM_i = \begin{bmatrix} R_i a_{11} & \cdots & R_i a_{j1} \\ \vdots & \ddots & \vdots \\ R_i a_{1k} & \cdots & R_i a_{jk} \end{bmatrix} \tag{4}$$

or, as $RM_i = R_i B_i$. This is done for all the "n" number of randomly generated binary square matrices as shown within the training phase block of Fig. 2. Next, all the "n" number of such matrices are summed together into a single matrix which is known as the cumulative reward matrix (CRM). The CRM can be now written as,

$$CRM = RM_1 + RM_2 + \cdots + RM_n = \sum_{i}^{n} RM_i \tag{5}$$

The CRM is of the same dimension as the reward matrix *RM* or the binary square matrix *B*. It contains the accumulated rewards for each pixel i.e. the algorithm now has the information of the specific positions within the square binary matrix which has one of the highest "*confidence*" scores. This will be beneficial later. The CRM is then saved and passed onto the final phase: The inference phase.

2.2. Inference phase

The CRM helps the algorithm to generalize the problem. A simple way to understand this is as follows. There are two matrices B_i and $B_{i+1}$with reward value of R_i and R_{i+1}. Here, for simplicity $R_i \geq R_{i+1}$. The *j*th position of both B_i and B_{i+1}can contain (1) both "'1", (2) both "0" and (3) combination of "1" and "0". The accumulated reward for the *j*th position in the CRM will be (a) $(1 \times R_i + 1 \times R_{i+1})$, (b) $(0 \times R_i + 0 \times R_{i+1})$, and (c) $(0 \times R_i + 1 \times R_{i+1})$ *or* $(1 \times R_i + 0 \times R_{i+1})$. Simple analysis tells that (a) > (c) > (b). The algorithm understands the positions of "1" and "0" based on this "*confidence*" scores in the CRM. Finally, what remains is just to filter and separate it out. Therefore, the first step is to ascertain this minimum. This is then subtracted from the CRM. In, parallel, a final binary matrix F is created which is initialized with all "0"'s. This takes care of "0" valued pixels in the matrix. In the subsequent step, the mean of the CRM is taken to handle for cases like case (1) and case (3) as shown in the former example because a value greater than this mean only guarantees that specific pixel to be "1" without fail as the cumulative reward values for such pixel positions will be amongst the highest. Nonetheless, the rationale behind evaluating the mean here can also be understand from the perspective of signal processing where signal averaging is carried out to extract the real signal from a noisy channel. Hence, the final decision-making step involves this check. If the value is found to be greater than the mean, the matrix F is updated with a value of "1" else is kept as it is i.e. "0".

3. Results and discussion

The designed nanophotonic structure comprises of square pixelated Si-sub-units that are intelligently distributed in an air medium, i.e., the algorithm predicts the location of the Si-sub-units to obtain the desired 50:50 power splitting of the input power with minimal loss. The T-junctions are designed for a light source with the fundamental TE polarization (with non-zero components of E_x, E_y, and H_z). To avoid any undesired back reflections, perfectly matched layers (PML) surrounded the boundaries of the computational domain.

Two different commercially available software from Lumerical Inc. were employed during the entire process. During the training as well as the inference phase, Lumerical MODE solutions was utilized to extract the response (value of insertion loss) of the nanophotonics structure via 2.5D varFDTD. An additional post-validation after the inference phase was employed to cross-check the obtained results across a full 3D FDTD with Lumerical FDTD solutions. The reason behind using 2.5D varFDTD was to speed up the "*learning*" process. The 2.5D varFDTD, as well as the 3D FDTD, had a mesh accuracy of 1/35 of the free space wavelength to ensure accurate modelling. Parallelization during the training phase for the 2.5D varFDTD simulations was carried out on 10 Intel Pentium i7 CPUs with 16 GB RAM each. The complete numerical simulations took $\sim$ 67 h (almost three days). The number of matrices used in the learning phase (i.e., N rounds) was empirically set to 10,000. Care was also taken to make sure that this randomly generated binary "unit cell" matrix is unique

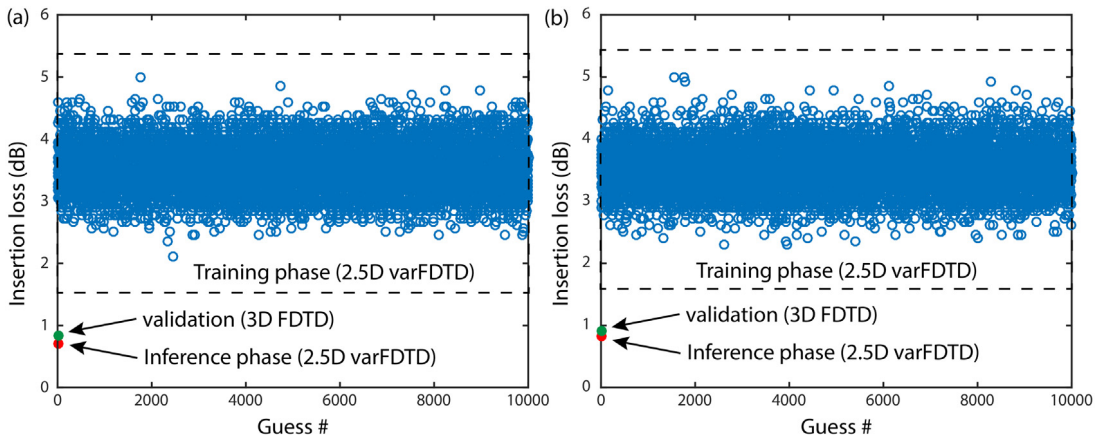

Fig. 3. The insertion loss (in dB) for each guess (hollow blue dots) and final prediction (solid red dot) using 10000 guesses across both designs. The minimum possible insertion loss achieved in the final prediction from the inference phase across both designs are (a) ∼0.86 dB and (b) ∼0.95 dB respectively, although the insertion loss (in dB) for nearly all guesses during the training phase for both (a) and (b) are in the range ∼2 to 5 dB.

across all the CPUs such that the same structure is not running on two different CPUs at any given instant of time.

Two different designs for the on-chip integrated nanophotonic in-plane incidence 50:50 T-junctions were made, each corresponding to a different worst possible insertion loss ($I.L._{worst}$) scenario as derived in Eq. (2). The 2.5D FDTD simulated $I.L._{worst}$ value for a binary "unit cell", all of whose individual sub-unit elements are "0" (air) was ∼9.77 dB and a binary "unit cell" all of whose individual unit elements are "1" (silicon) was ∼16.62 dB. Theoretically speaking, even though this change in $I.L._{worst}$ value would neither have changed anything in relation to the working principle of the algorithm nor the final prediction, it was certainly interesting to cross-check this before making any conclusive statements. As observed in Fig. 3, the insertion loss for each guess and final prediction using 10000 guesses follows a very similar trend. For nearly all guesses during the training phase, the average insertion loss was in the range of ∼2 to 5 dB. The minimum insertion loss was achieved in the final prediction from the inference phase across both the designs. The insertion loss (efficiency in terms of total power-out as % of power-in) was determined to be ∼0.82 dB and ∼0.87 dB across 2.5D varFDTD for both the designs, respectively. The full 3D FDTD gives a slightly worse but more accurate value for the insertion loss at ∼0.86 dB and ∼0.95 dB for the operational wavelength of 1.55 μm for each design, respectively. The subsequent "unit cell" structure, steady-state electric field distribution at $\lambda = 1.55$ μm, and the insertion loss for each design under broadband operation (1.45–1.65 μm) are plotted in Fig. 4(a–b). It can be observed that for both the designs that the insertion loss was to some extent, virtually wavelength insensitive with variations below 10% over the wavelength range from 1.52 μm to 1.57 μm. To be specific, the bandwidth was from ∼1.53 μm to ∼1.57 μm (∼40 nm bandwidth) for the first design in Fig. 4(a). For the other design in Fig. 4(b), this bandwidth was in the range of 1.52 μm to 1.56 μm (∼40 nm bandwidth). One must keep in mind here that the machine learning algorithm was trained on and the final inference was made at only a single wavelength of 1.55 μm. There is no crosstalk amongst the output waveguides. Scattering was negligible, as evidenced by the steady-state response plots in Fig. 4(a–b) with the appearance of an interference pattern at the input end, indicating the existence of very weak back-reflection. Furthermore, upon even more close inspection of the steady-state intensity profile in Fig. 4(a–b) one can further conclude : (a) an efficient splitting of the fundamental mode and (b) a strong modal match at the output port (waveguide) with excellent coupling efficiency for both the splitter designs.

Finally, an unbiased comparison in relation to large conventional integrated beam splitters needs to be carried out to highlight the true significance of this work, apart from the design methodology discussed in the previous sections. We acknowledge the fact here that the designs discussed in this study are certainly not the best in terms of power efficiency (insertion loss) compared to what has already been reported in the literature. An insertion loss of 0.8 to 1 dB (as in the case of the designs reported herein) in general corresponds to a net power transmission efficiency of ∼80%–83%. Conventional integrated beam splitters or even optimized 50:50 splitters (Y-shape or T-shape) report insertion losses of < 0.3 dB (∼90%–93% in terms of net power transmission) [2,38,49,52–54]. However, we believe that the advantage of a small area footprint outweighs such a marginal ∼10% reduction in efficiency from the perspective that now a greater number of devices can be integrated together in a single photonic chip than what has been previously possible (analogous to what has been seen for transistors in electronic circuits over the last decades). This will eventually lead designers to design complicated photonic logic circuits with more flexibility. We can proclaim this as a "***Photonic Moore's law***" [62]. This trade-off between efficiency and footprint can be handled by utilizing "unit cell" structures with a larger number of sub-unit pixels in the same total area, that is a larger sub-unit density (say for example 30 x 30 or even 60 x 60) to approximate sharp bends more gradually than what has been done with a coarse structure (12 x 12) in this paper. The geometry constraints assumed in this work were taken so to represent what can be demonstrated at a standard university-level fabrication facility without much difficulty. However, at industrial foundries, one can expect to exploit this to create even more efficient beam splitter structures with such an ultra-compact area footprint. Apart from this, the use of higher refractive index materials [21,23] has also been previously shown to improve device performance and is expected to provide the same advantages here as well.

In addition to this, a smaller structure will also have a lower heat generation "*per device*" in contrast to a conventional one [63]. Now, if one is interested in lowering the overall heat generation in the photonic circuit keeping the total number of individual structures same, a smaller footprint will have the advantage that heat sinks can be easily accommodated within the same area along with the device in place of a conventional larger structure taking up the same amount of space. This will eventually lower the *operational cost* and provide cost-effective solutions [64,65].

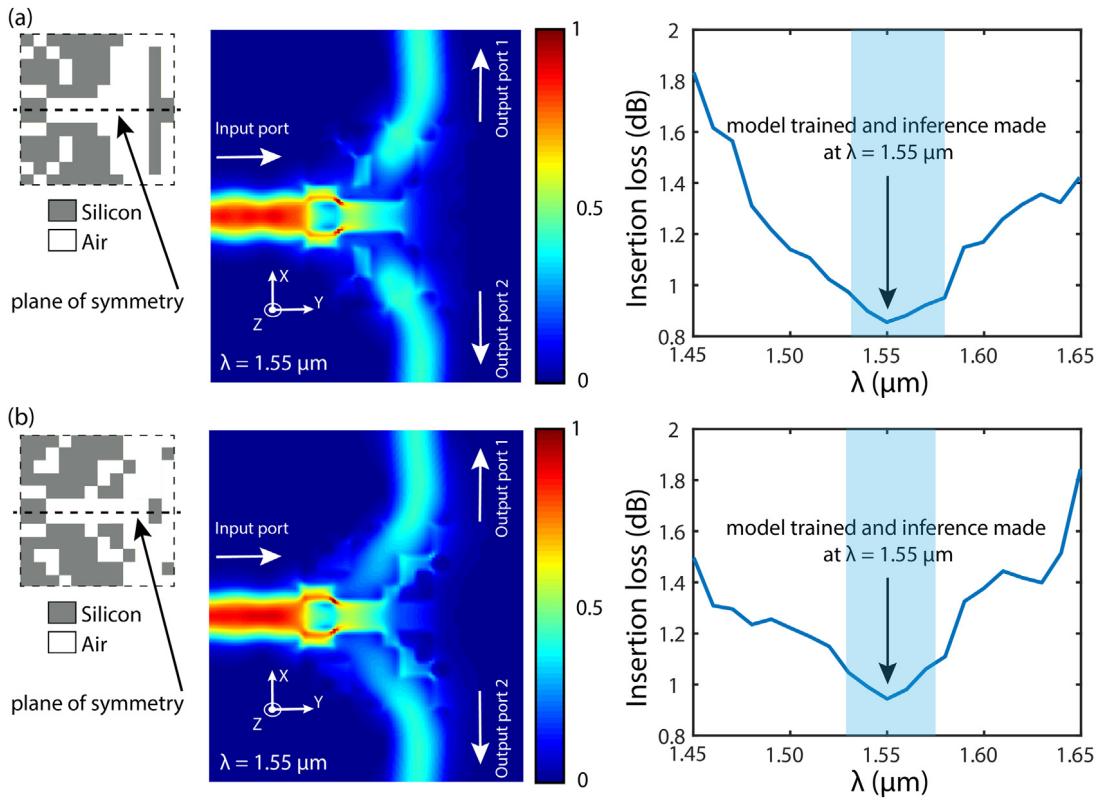

Fig. 4. The predicted pixel profile, steady-state response at the operational wavelength of 1.55 μm and the insertion loss of the structure under broadband operation (1.50–1.60 μm) for the design where $I.L._{worst}$ refers to a binary "unit cell" all of whose individual sub-unit elements are either (a) "0" (air) or (b) "1" (silicon).

4. Conclusion

We introduced the design of subwavelength ultra-compact and efficient on-chip integrated nanophotonic 50:50 beam splitters (T-junctions) via a machine learning algorithm. We numerically investigated its power splitting effect at an operating wavelength of $\lambda = 1.55$ μm by using FDTD simulations. Despite its low insertion loss, as indicated from the full-wave simulations, we would like to point out that imperfections and impurities during the fabrication step will inevitably decrease the device efficiency and degrade its performance. However, these problems will not have any detrimental effect in relation to the functionality of the designed structure. From such perspective, using the algorithm as described in this paper one can now extend this to realize situation dependent nanophotonic splitter designs such as asymmetric power splitters (which are optimized for unbalanced power splitting i.e. 40:60, 20:80 or 100:0), splitters based on different geometries, for example Y-branch splitters or even polarization beam splitters by defining an appropriate reward function. Nonetheless, reconfigurability can also be achieved with the use of active materials in these pixelated "unit cells". Overall, the method employed and our results evidence that the use of machine learning algorithms is a promising technique for the inverse design of wide variety of efficient passive and active integrated photonics devices.

CRediT authorship contribution statement

Sourangsu Banerji: Conceptualization, Methodology, Software, Visualization, Writing - original draft. **Apratim Majumder:** Software. **Alexander Hamrick:** Software. **Rajesh Menon:** Conceptualization, Writing - review & editing. **Berardi Sensale-Rodriguez:** Supervision, Conceptualization, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the NSF awards: ECCS #1936729 and MRI #1828480.

References

[1] G.T. Reed, A.P. Knights, Silicon Photonics: An Introduction, John Wiley & Sons, 2004.
[2] L. Chrostowski, M. Hochberg, Silicon Photonics Design: From Devices To Systems, Cambridge University Press, 2015.
[3] H. Jia, T. Zhou, L. Zhang, J. Ding, X. Fu, L. Yang, Optical switch compatible with wavelength division multiplexing and mode division multiplexing for photonic networks-on-chip, Opt. Express 25 (17) (2017) 20698–20707.
[4] F. Ren, J. Li, Z. Wu, T. Hu, J. Yu, Q. Mo, Z. Li, et al., Three-mode mode division-multiplexing passive optical network over 12-km low mode-crosstalk FMF using all-fiber mode MUX/DEMUX, Opt. Commun. 383 (2017) 525–530.
[5] I. Cerutti, N. Andriolli, P. Velha, Engineering of closely packed silicon-on-isolator waveguide arrays for mode division multiplexing applications, J. Opt. Soc. Amer. B 34 (2) (2017) 497–506.
[6] IEEE Standard for Verilog Hardware Description Language, IEEE, 2006.
[7] P.J. Ashenden, The Designer's Guide To VHDL, Morgan Kaufmann, 2008.
[8] Q. Xu, B. Schmidt, S. Pradhan, M. Lipson, Micrometre-scale silicon electro-optic modulator, Nature 435 (7040) (2005) 325–327.
[9] L.W. Luo, G.S. Wiederhecker, J. Cardenas, C. Poitras, M. Lipson, High quality factor etchless silicon photonic ring resonators, Opt. Express 19 (7) (2011) 6284–6289.
[10] S.A. Miller, Y.C. Chang, C.T. Phare, M.C. Shin, M. Zadka, S.P. Roberts, B. Stern, X. Ji, A. Mohanty, O.A.J. Gordillo, U.D. Dave, M. Lipson, Large-scale optical phased array using a low-power multi-pass silicon photonic platform, Optica 7 (1) (2020) 3–6.

[11] M. Yu, Y. Okawachi, R. Cheng, C. Wang, M. Zhang, A.L. Gaeta, M. Lončar, Raman Lasing and soliton mode-locking in lithium niobate microresonators, Light: Sci. Appl. 9 (1) (2020) 1–7.

[12] J. Leuthold, C. Koos, W. Freude, Nonlinear silicon photonics, Nat. Photonics 4 (8) (2010) 535.

[13] E. Bor, M. Turduev, H. Kurt, Differential evolution algorithm based photonic structure design: Numerical and experimental verification of subwavelength $\lambda/5$ focusing of light, Sci. Rep. 6 (2016) 30871.

[14] J. Lu, J. Vuckovic, Nanophotonic computational design, Opt. Express 21 (11) (2013) 13351–13367.

[15] A.Y. Piggott, J. Lu, T.M. Babinec, K.G. Lagoudakis, J. Petykiewicz, J. Vučković, Inverse design and implementation of a wavelength demultiplexing grating coupler, Sci. Rep. 4 (2014) 1–5.

[16] A.Y. Piggott, J. Lu, K.G. Lagoudakis, J. Petykiewicz, T.M. Babinec, J. Vučković, Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer, Nat. Photon. 9 (6) (2015) 374–377.

[17] L. Su, A.Y. Piggott, N.V. Sapra, J. Petykiewicz, J. Vučković, Inverse design and demonstration of a compact on-chip narrowband three-channel wavelength demultiplexer, ACS Photon. 5 (2) (2017) 301–305.

[18] P.I. Borel, A. Harpøth, L.H. Frandsen, M. Kristensen, P. Shi, J.S. Jensen, O. Sigmund, Topology optimization and fabrication of photonic crystal structures, Opt. Express 12 (9) (2004) 1996–2001.

[19] B. Shen, P. Wang, R. Polson, R. Menon, Integrated metamaterials for efficient and compact free-space-to-waveguide coupling, Opt. Express 22 (22) (2014) 27175–27182.

[20] B. Shen, R. Polson, R. Menon, Integrated digital metamaterials enables ultra-compact optical diodes, Opt. Express 23 (8) (2015) 10847–10855.

[21] B. Shen, P. Wang, R. Polson, R. Menon, An integrated-nanophotonics polarization beamsplitter with $2.4\times2.4\mu$m2 footprint, Nat. Photonics 9 (6) (2015) 378–382.

[22] A. Majumder, S. Banerji, K. Miyagawa, M. Meem, M. Mondol, B. Sensale-Rodriguez, R. Menon, Programmable metamaterials & metasurfaces for ultra-compact multi-functional photonics, in: CLEO: Applications and Technology (AM4K-5), Optical Society of America, 2019.

[23] A. Majumder, B. Shen, R. Polson, R. Menon, Ultra-compact polarization rotation in integrated silicon photonics using digital metamaterials, Opt. Express 25 (2017) 19721–19731.

[24] A. Majumder, B. Shen, R. Polson, T. Andrew, R. Menon, An ultra-compact nanophotonic optical modulator using multi-state topological optimization, 2017, arXiv preprint arXiv:1712.02835.

[25] B. Shen, R. Polson, R. Menon, Integrated digital metamaterials enables ultra-compact optical diodes, Opt. Express 23 (2015) 10847–10855.

[26] H. Jia, T. Zhou, X. Fu, J. Ding, L. Yang, Inverse-design and demonstration of ultracompact silicon meta-structure mode exchange device, ACS Photonics 5 (5) (2018) 1833–1838.

[27] A.Y. Piggott, J. Petykiewicz, L. Su, J. Vučković, Fabrication-constrained nanophotonic inverse design, Sci. Rep. 7 (1) (2017) 1–7.

[28] F. Callewaert, S. Butun, Z. Li, K. Aydin, Inverse design of an ultra-compact broadband optical diode based on asymmetric spatial mode conversion, Sci. Rep. 6 (2016) 1–10.

[29] L. Su, R. Trivedi, N.V. Sapra, A.Y. Piggott, D. Vercruysse, J. Vučković, Fully automated optimization of grating couplers, Opt. Express 26, 4023–4034.

[30] C.M. Lalau-Keraly, S. Bhargava, O.D. Miller, E. Yablonovitch, Adjoint shape optimization applied to electromagnetic design, Opt. Express 21 (18) (2013) 21693–21701.

[31] J. Wang, Y. Shi, T. Hughes, Z. Zhao, S. Fan, Adjoint-based optimization of active nanophotonic devices, Opt. Express 26 (3) (2018) 3236–3248.

[32] M. Meem, S. Banerji, A. Majumder, F.G. Vasquez, B. Sensale-Rodriguez, R. Menon, Broadband lightweight flat lenses for long-wave infrared imaging, Proc. Nat. Acad. Sci. 116 (43) (2019) 21375–21378.

[33] R. Biswas, et al., Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data, Phys. Rev. D 88 (6) (2013) 062003.

[34] S.V. Kalinin, B.G. Sumpter, R.K. Archibald, Big-deep smart data in imaging for guiding materials design, Nat. Mater. 14 (100) (2015) 973–980.

[35] J. Carrasquilla, R.G. Melko, Machine learning phases of matter, Nat. Phys. 13 (2017) 431–434.

[36] L. Wang, Discovering phase transitions with unsupervised learning, Phys. Rev. B 94 (19) (2016) 195105.

[37] D. Deng, X. Li, S. Das Sarma, Machine learning topological states, Phys. Rev. B 96 (19) (2017) 195145.

[38] M. Turduev, E. Bor, C. Latifoglu, I. Halil Giden, Y. Sinan Hanay, H. Kurt, Ultracompact photonic structure design for strong light confinement and coupling into nanowaveguide, J. Lightwave Technol. 36 (2018) 2812–2819.

[39] K. Yao, R. Unni, Y. Zheng, Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale, Nanophotonics 8 (3) (2019) 339–366.

[40] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, Mastering the game of go without human knowledge, Nature 550 (7676) (2017) 354–359.

[41] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, MIT press, 2018.

[42] L.P. Kaelbling, M.L. Littman, A.W. Moore, Reinforcement learning: A survey, J. Artif. Intell. Res. 4 (1996) 237–285.

[43] D.Silver T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, A general reinforcement learning algorithm that masters chess, shogi, and go through self-play, Science 362 (6419) (2018) 1140–1144.

[44] D. Liu, Y. Tan, E. Khoram, Z. Yu, Training deep neural networks for the inverse design of nanophotonic structures, ACS Photon. 5 (4) (2018) 1365–1369.

[45] Y. Kiarashinejad, M. Zandehshahvar, S. Abdollahramezani, O. Hemmatyar, R. Pourabolghasem, A. Adibi, Knowledge discovery in nanophotonics using geometric deep learning, Adv. Intell. Syst. (2019) 1900132.

[46] S. So, J. Rho, Designing nanophotonic structures using conditional deep convolutional generative adversarial networks, Nanophotonics 8 (7) (2019) 1255–1261.

[47] S. Chugh, S. Ghosh, A. Gulistan, B.M.A. Rahman, Machine learning regression approach to the nanophotonic waveguide analyses, J. Lightwave Technol. 37 (24) (2019) 6080–6089.

[48] S. So, T. Badloe, J. Noh, J. Rho, J. Bravo-Abad, Deep learning enabled inverse design in nanophotonics, Nanophotonics (2020).

[49] M.H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, K. Parsons, Deep neural network inverse design of integrated photonic power splitters, Sci. Rep. 9 (1) (2019) 1–9.

[50] Y. Cai, Q. Zhou, X. Hong, R. Shi, Y. Wang, Application of optical proximity correction technology, Sci. China Ser. F-Inf. Sci. 51 (2008) 213–224.

[51] W. Chang, L. Lu, X. Ren, D. Li, Z. Pan, M. Cheng, D. Liu, M. Zhang, Ultracompact dual-mode waveguide crossing based on subwavelength multimode-interference couplers, Photon. Res. 6 (2018) 660–665.

[52] Y. Zhang, S. Yang, A.E.-J. Lim, G.-Q. Lo, C. Galland, T. B.-Jones, M. Hochberg, A compact and low loss Y-junction for submicron silicon waveguide, Opt. Express 21 (2013) 1310–1316.

[53] H. Kurt, I.H. Giden, D.S. Citrin, Design of T-shaped nanophotonic wire waveguide for optical interconnection in H-tree network, Opt. Express 19 (2011) 26827–26838.

[54] A.M. Alpkılıç, Y.A. Yılmaz, H. Kurt, Parametric study of multi-outputs T-junction spatial mode demultiplexers design with an objective-first algorithm, in: Nanoengineering: Fabrication, Properties, Optics, Thin Films, and Devices XVI, 11089, 110890M, International Society for Optics and Photonics, 2019.

[55] K. Xu, L. Liu, X. Wen, W. Sun, N. Zhang, N. Yi, S. Sun, S. Xiao, Q. Song, Integrated photonic power divider with arbitrary power ratios, Optim. Lett. 42 (4) (2017) 855–858.

[56] Z.. Lin, W. Shi, Broadband, low-loss silicon photonic y-junction with an arbitrary power splitting ratio, Opt. Express 27 (10) (2019) 14338–14343.

[57] X. Ren, W. Chang, L. Lu, M. Yan, D. Liu, M. Zhang, Digitized adjoint method for inverse design of digital nanophotonic devices, 2019, arXiv preprint arXiv:1902.00654.

[58] H. Xie, Y. Liu, Y. Wang, Y. Wang, Y. Yao, Q. Song, J. Du, Z. He, K. Xu, An ultra-compact 3-dB power splitter for three modes based on pixelated meta-structure, IEEE Photon. Technol. Lett. 32 (6) (2020) 341–344.

[59] W. Chang, X. Ren, Y. Ao, L. Lu, M. Cheng, L. Deng, D. Liu, M. Zhang, Inverse design and demonstration of an ultracompact broadband dual-mode 3 dB power splitter, Opt. Express 26 (18) (2018) 24135–24144.

[60] C. Latifoğlu, Binary matrix guessing problem, 2017, arXiv preprint, arXiv:1701.06167.

[61] https://www.lumerical.com/learn/whitepapers/lumericals-2-5d-fdtd-propagation-method/.

[62] M. M. Smit, J.J.G.M. Van der Tol, M. Hill, Moore's law in photonics, Laser Photonics Rev. 6 (1) (2012) 1–13.

[63] J. Punch, Thermal challenges in photonic integrated circuits, in: 2012 13th International Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems, IEEE, 2012, pp. 1–6.

[64] J.A. Hudgings, K.P. Pipe, R.J. Ram, Thermal profiling for optical characterization of waveguide devices, Appl. Phys. Lett. 83 (19) (2003) 3882–3884.

[65] G. Gilardi, W. Yao, H.R. Haghighi, X.J. Leijtens, M.K. Smit, M.J. Wale, Deep trenches for thermal crosstalk reduction in InP-based photonic integrated circuits, J. Lightwave Technol. 32 (24) (2014) 4864–4870.

Sourangsu Banerji is a fourth year Ph.D. student in the Department of Electrical and Computer Engineering at the University of Utah, currently advised by Prof. Berardi Sensale-Rodriguez and Prof. Rajesh Menon. His current research interests are in the areas of diffractive optics, metamaterials, plasmonics, nanotechnology and nanophotonics with an emphasis on utilizing computational approaches towards their design. He has extensively worked on translating and exploiting well-established methods and concepts from computer science and engineering right up to the realm of optics and photonics by demonstrating the computational design of planar passive and reconfigurable nano-scale devices in the THz/visible regime. His work also includes looking into several different optimization algorithms as well as machine learning techniques for possible use in designing on-chip integrated nanophotonic and free space optical devices. His work caters to the growing need for designing challenging and complicated devices/systems by employing computational approaches which are neither time consuming not memory intensive yet readily scalable.

Apratim Majumder is a Postdoctoral Scholar in the Department of Electrical and Computer Engineering at the University of Utah. He completed his Ph.D. in Electrical and Computer Engineering also at the University of Utah in December of 2016 with a focus on optics and photonics-based nanotechnology.

Alexander Hamrick is an undergraduate student studying Computer Science at the University of Utah. His interests include software engineering, machine learning, and big data.

Rajesh Menon is an Associate Professor of Electrical and Computer Engineering at the University of Utah. Prior to joining the University of Utah in August 2009, Prof. Menon was a research engineer and a post-doctoral scientist in the Research Laboratory of Electronics at MIT. He received the S.M and Ph.D. degrees, both from MIT. From 2005 to 2009, Prof. Menon was the Chief Technology Officer of LumArray, Inc, a company he co-founded with colleagues at MIT. Prof. Menon has pioneered several technologies that will enable far-field optics to manipulate and image matter with nanoscale resolution. His research has spawned numerous publications (some which have had extensive media coverage), patents, and a spin-off company. He has led several projects in nanopatterning and nanoscopy with support from DARPA, the NSF and the MIT Deshpande Center for Technological Innovation.

Berardi Sensale-Rodriguez is a tenured Associate Professor at the University of Utah, with an appointment with the Department of Electrical & Computer Engineering. He joined the faculty at the University of Utah in 2013, after earning his Ph.D. in Electrical Engineering from the University of Notre Dame (UND). During his research career, he has received the National Science Foundation (NSF) CAREER Award, the Eli J. and Helen Shaheen Graduate School Award in Engineering at UND, and the Best Student Paper Award at the 37th International Conference on Infrared, Millimeter and Terahertz Waves (IRMMW-THz), the best paper in Imaging Systems and Applications at the 2019 OSA Imaging and Applied Optics Congress, and the 2019 ECE department Outstanding Teaching award. Sensale-Rodriguez's research and teaching interests are in the area of (opto)electronic devices and materials. His research projects involve (i) simulation and design of electronic and photonic devices, in particular employing emerging materials, (ii) growth, fabrication and characterization of electronic and optical materials and devices, (iii) system integration of these devices.