MITSUBISHI ELECTRIC
Changes for the Better
MITSUBISHI ELECTRIC RESEARCH LABORATORIES, INC

Finding the Right Deep Neural Network Model for Efficient Design of Tunable Nanophotonic Devices

Minwoo Jung[1,2], Keisuke Kojima[1], Toshiaki Koike-Akino[1], Ye Wang[1], Dayu Zhu[1,3], and Matthew Brand[1]

[1]Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA.

[2]Department of Physics, Cornell University, Ithaca, NY 14853, USA.

[3]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA.

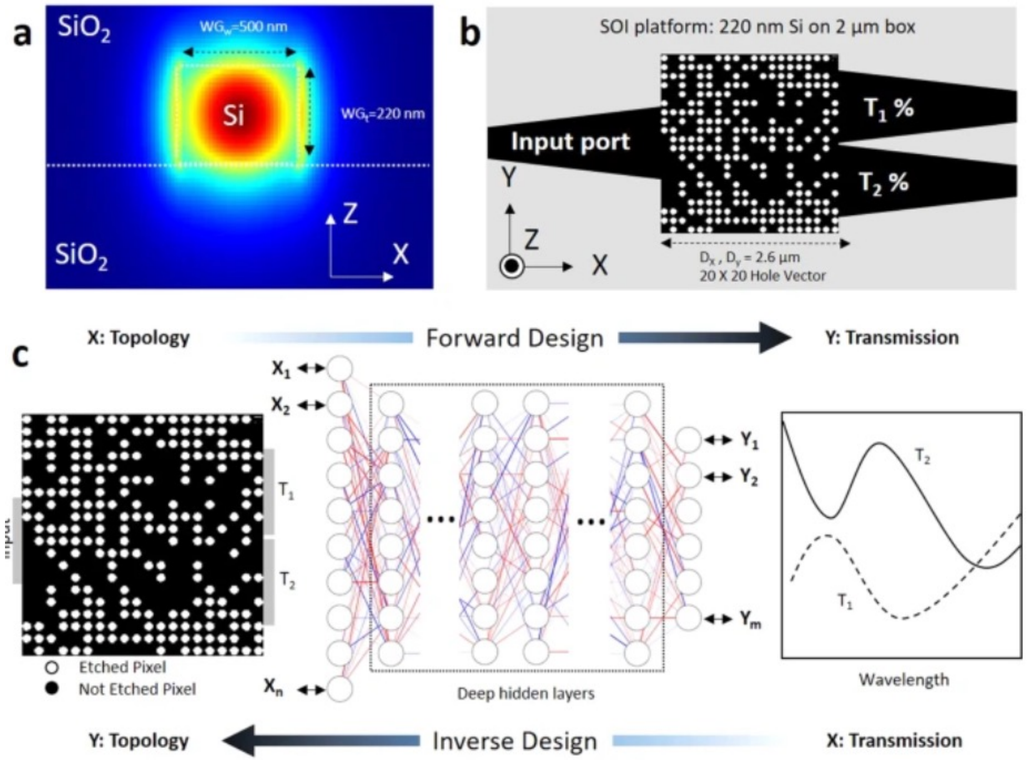
CLEO2022: SW5E.6

Neural Networks for Photonic Device Design Application

Deep Neural Network Inverse Design of Integrated Photonic Power Splitters

Mohammad H. Tahersima, Keisuke Kojima, Toshiaki Koike-Akino, Devesh Jha, Bingnan Wang, Chungwei Lin & Kieran Parsons

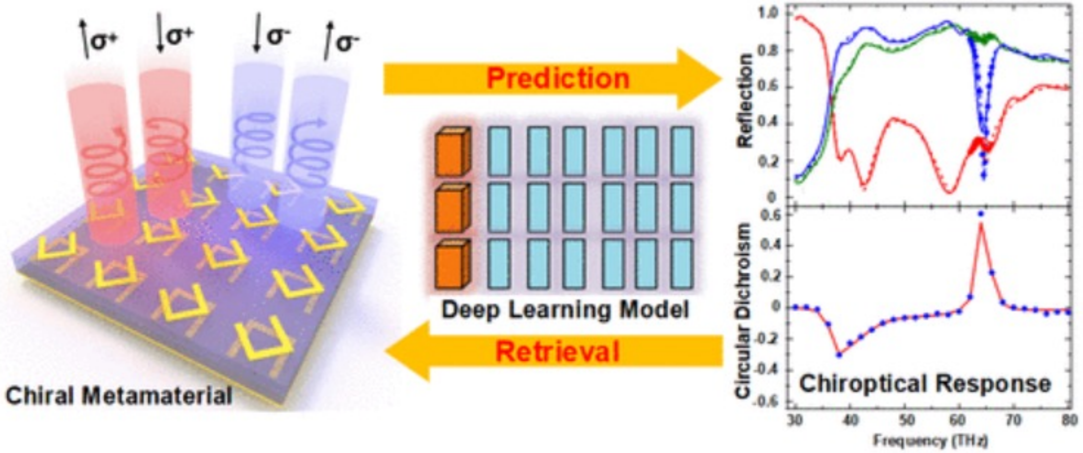
Scientific Reports 9, Article number: 1368 (2019) | Cite this article



Deep-Learning-Enabled On-Demand Design of Chiral Metamaterials

Wei Ma, Feng Cheng, and Yongmin Liu*

Cite this: *ACS Nano* 2018, 12, 6, 6326–6334
Publication Date: June 1, 2018
https://doi.org/10.1021/acsnano.8b03569

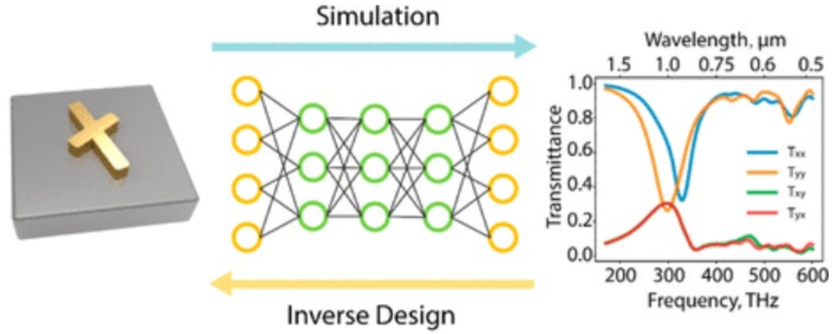


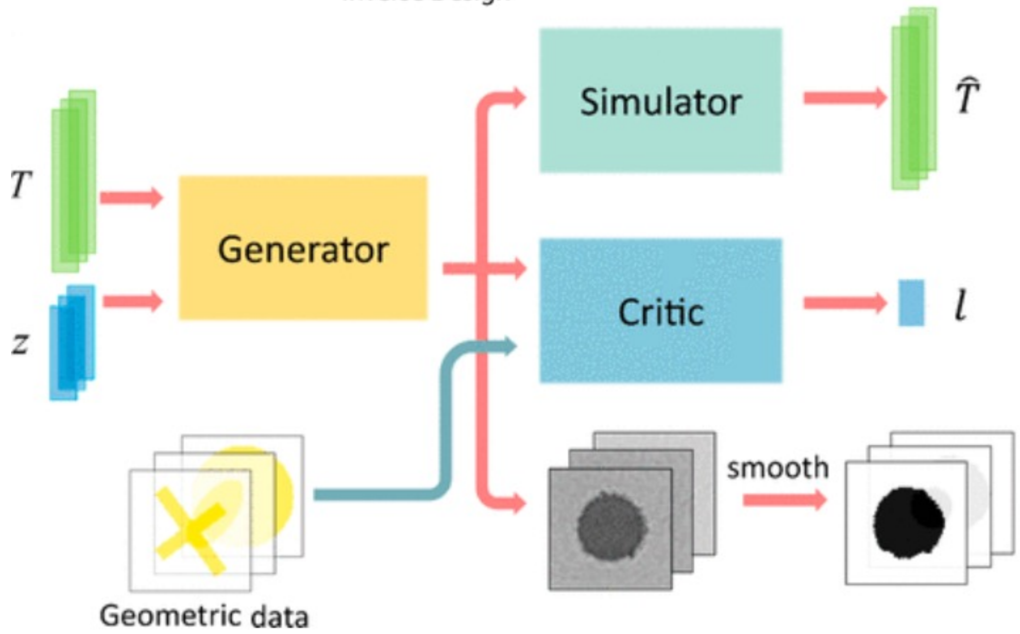
Generative Models : Encoder-Decoder Structure

Generative Model for the Inverse Design of Metasurfaces

Zhaocheng Liu, Dayu Zhu, Sean P. Rodrigues, Kyu-Tae Lee, and Wenshan Cai*

Cite this: *Nano Lett.* 2018, 18, 10, 6570–6576
Publication Date: September 12, 2018
https://doi.org/10.1021/acs.nanolett.8b03171



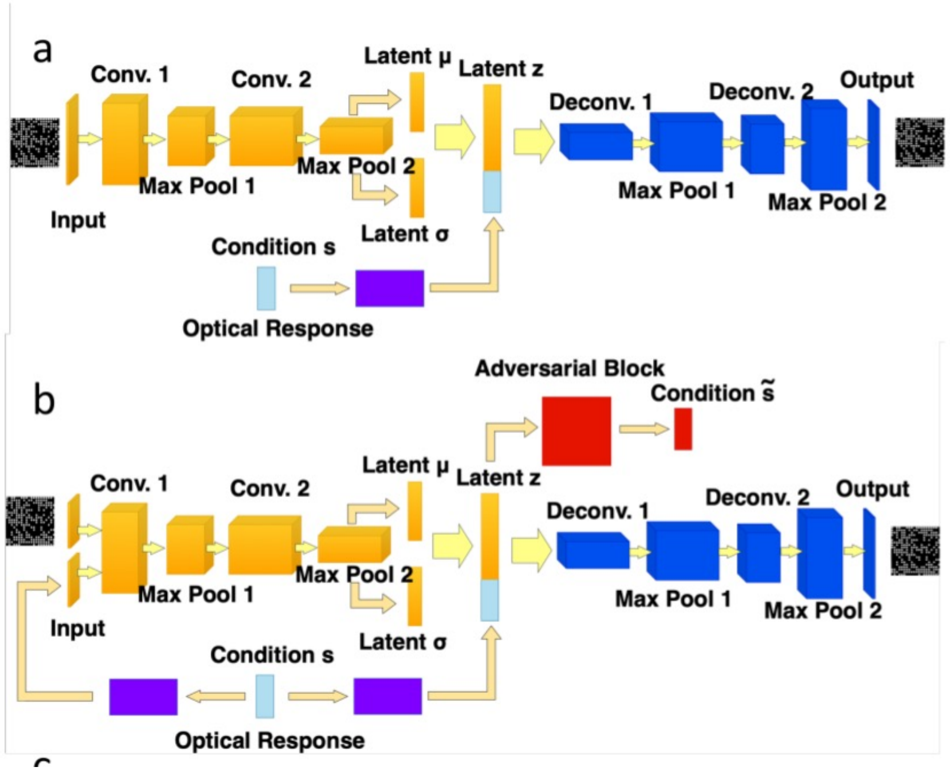


LASER & PHOTONICS REVIEWS

Original Paper

Generative Deep Learning Model for Inverse Design of Integrated Nanophotonic Devices

Yingheng Tang, Keisuke Kojima, Toshiaki Koike-Akino, Ye Wang, Pengxiang Wu, Youye Xie, Mohammad H. Tahersima, Devesh K. Jha, Kieran Parsons, Minghao Qi,

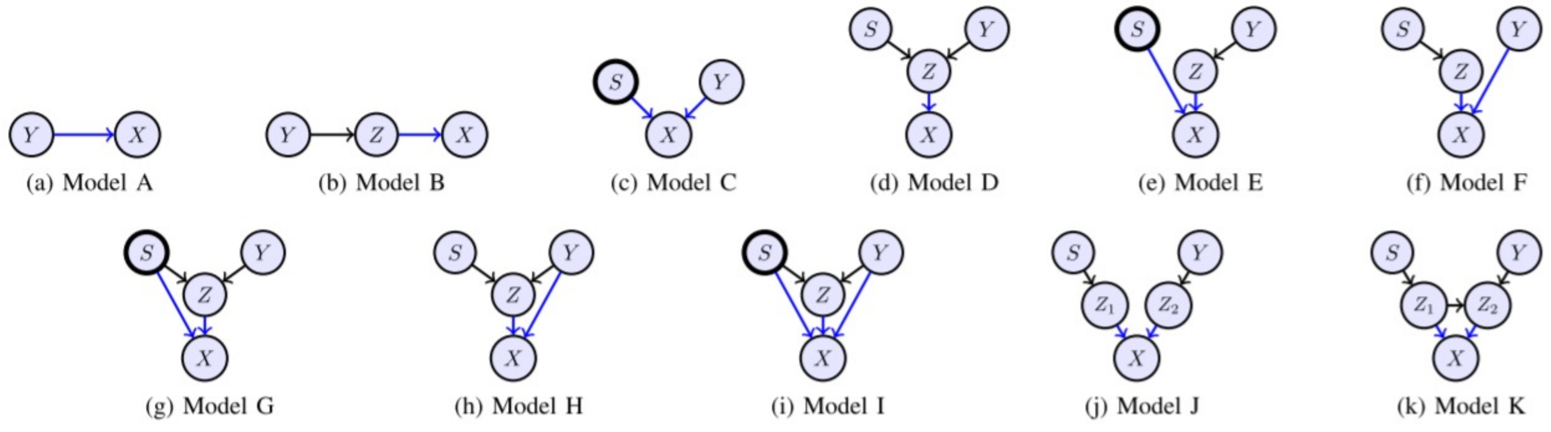


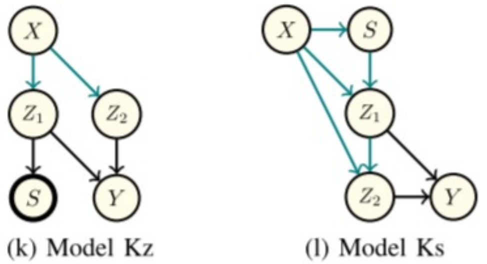
Bayesian Graph Exploration for finding optimal ANN architecture

IEEE Access

AutoBayes: Automated Bayesian Graph Exploration for Nuisance-Robust Inference

ANDAC DEMIR[1], (Student Member, IEEE),
TOSHIAKI KOIKE-AKINO[2], (Senior Member, IEEE),
YE WANG[2], (Senior Member, IEEE), AND
DENIZ ERDOGMUS[1], (Senior Member, IEEE)





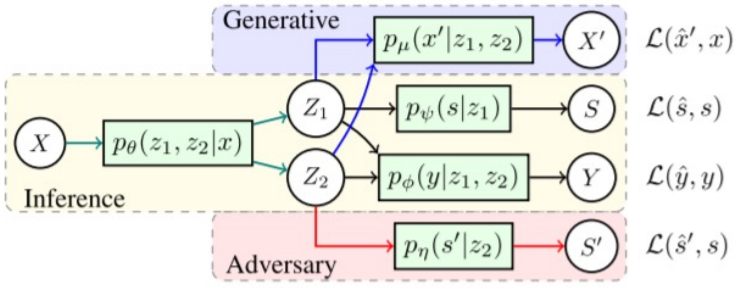


FIGURE 7. Overall network structure for pairing generative model K and inference model Kz.

Depending on the specifics of the system, we need to choose the right ANN architecture.

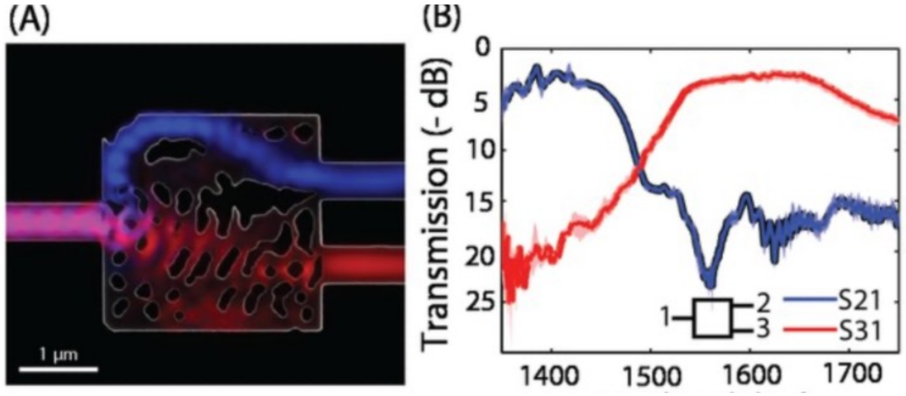
For generative models, we need to come up with the most plausible Bayesian-inference model.

Considered system : tunable SNOI wavelength splitter

Inverse design in nanophotonics

Sean Molesky, Zin Lin, Alexander Y. Piggott, Weiliang Jin, Jelena Vučković & Alejandro W. Rodriguez

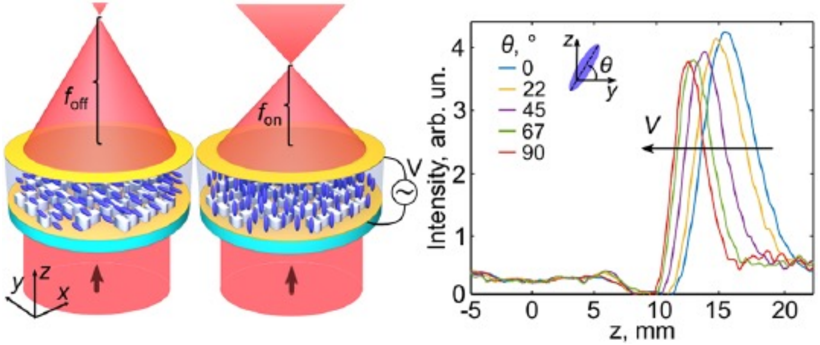
Nature Photonics **12**, 659–670 (2018) | Cite this article

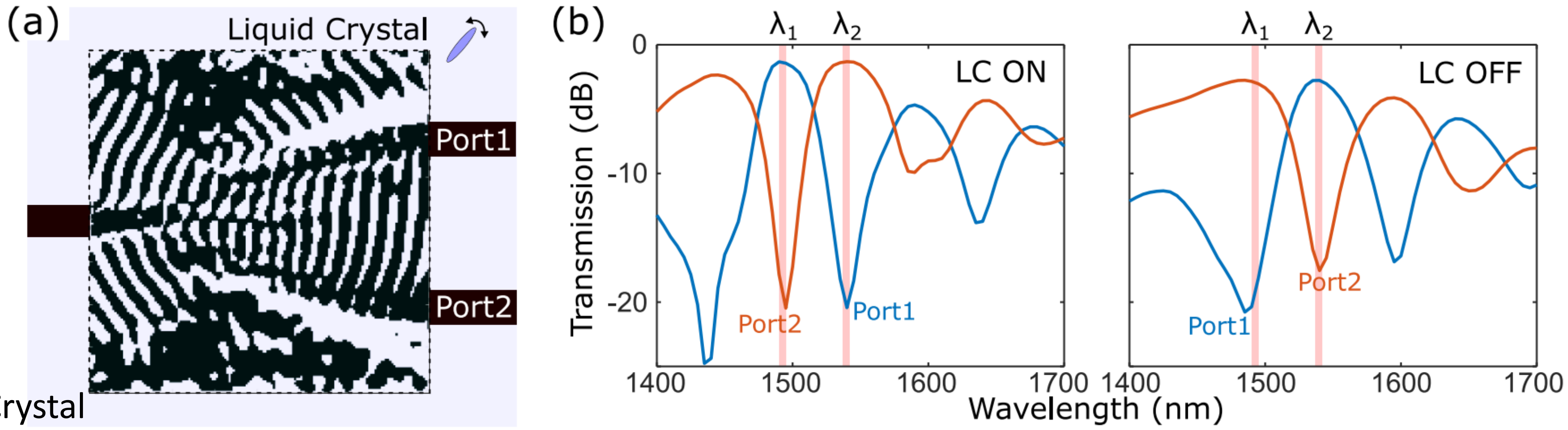


Electrically Actuated Varifocal Lens Based on Liquid-Crystal-Embedded Dielectric Metasurfaces

Melissa Bosch, Maxim R. Shcherbakov,* Kanghee Won, Hong-Seok Lee, Young Kim, and Gennady Shvets

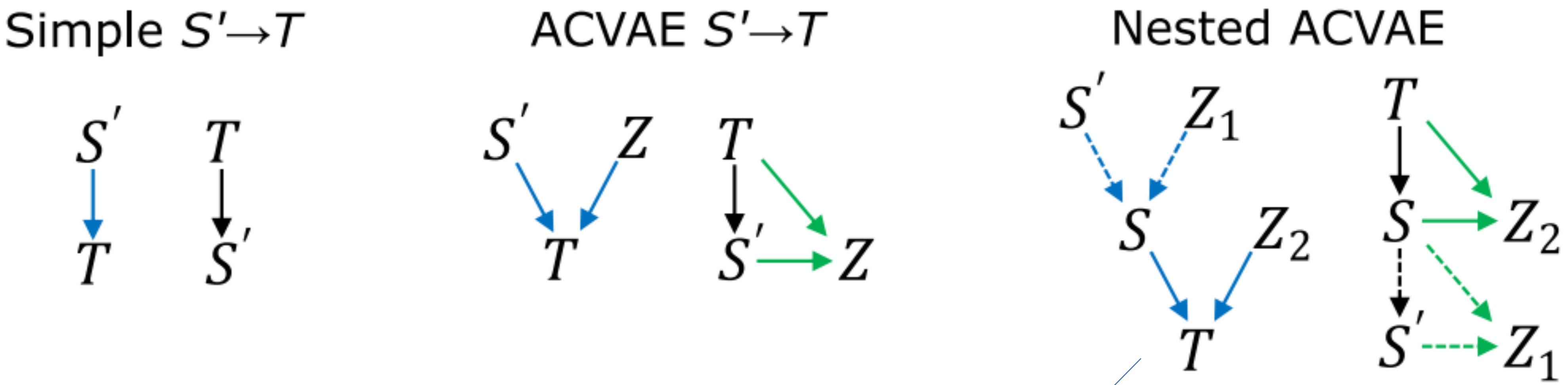
Cite This: *Nano Lett.* 2021, 21, 3849–3856 Read Online

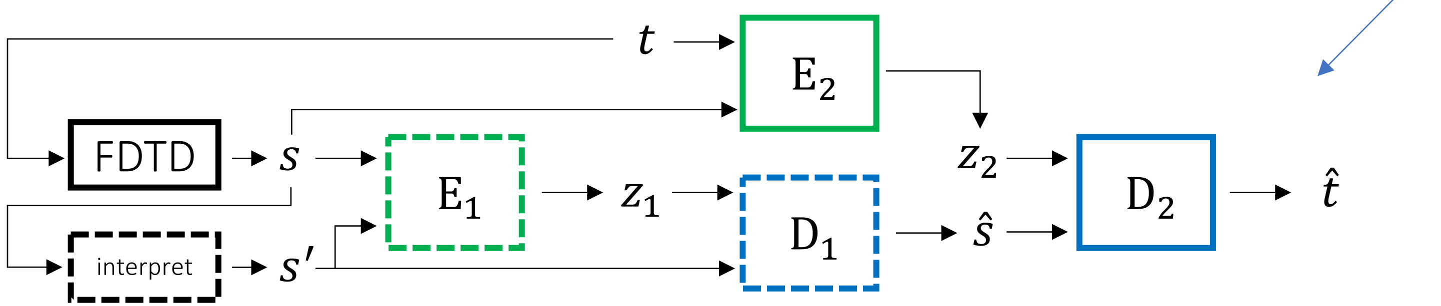




Black : SiN_x (n=2.4)
Light Blue : Liquid Crystal

Considered Bayesian Graphs





$\bar{s'} = A_1(z_1)$

$\bar{s} = A_2(z_2)$

t : device topology; 161×161

s : Full spectrum; $4(\text{LC on, off; port } 1{,}2) \times 61(1.4\mu\text{m}: 5\text{nm}: 1.7\mu\text{m})$ or 244

s': User-friendly spectrum information; $[\bar{\lambda}, \Delta\lambda, \overline{ER}]$

z_1: Latent variable out of the encoder1

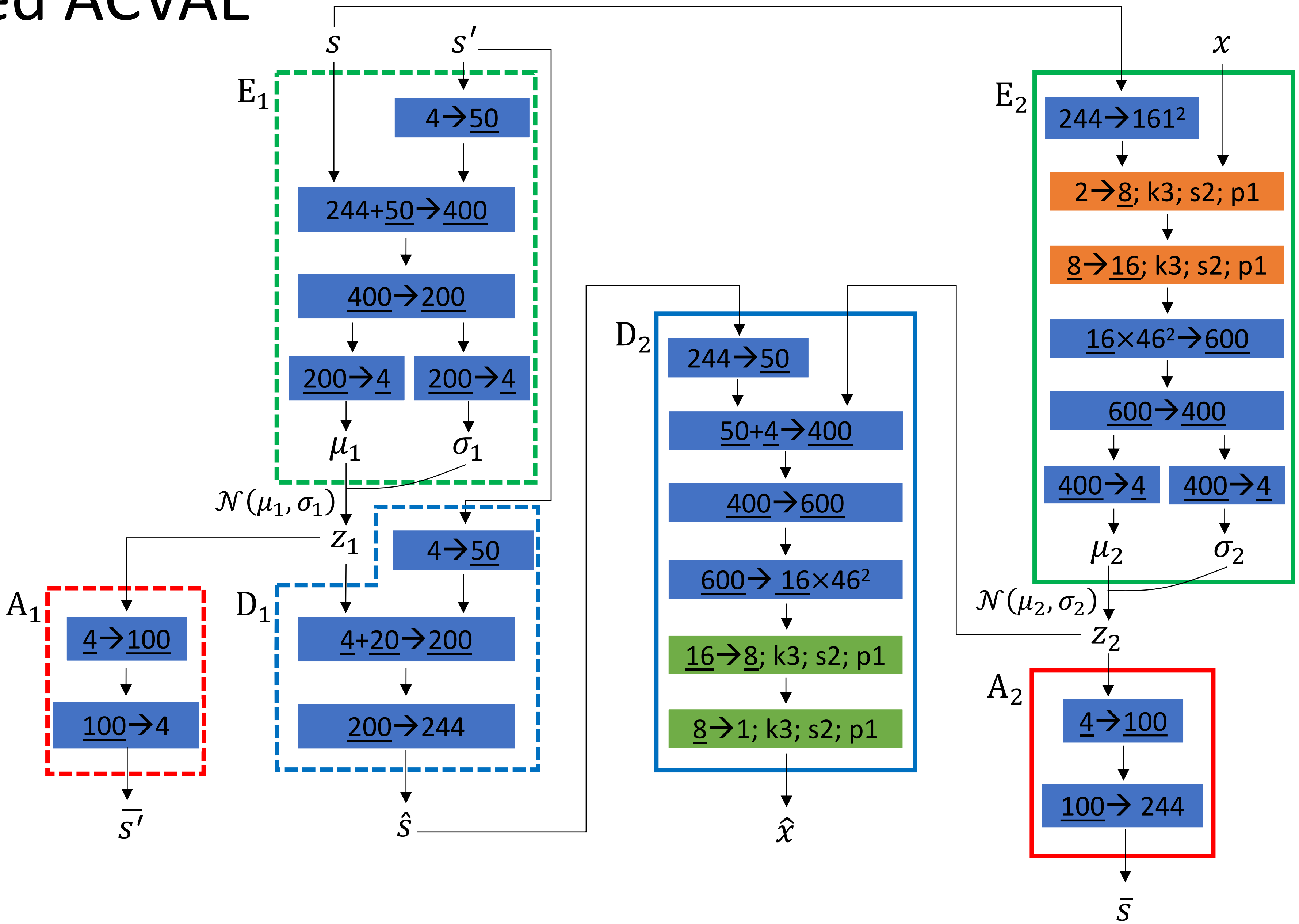
z_2: Latent variable out of the encoder2

$\hat{a}$: Variable a generated out of a decoder

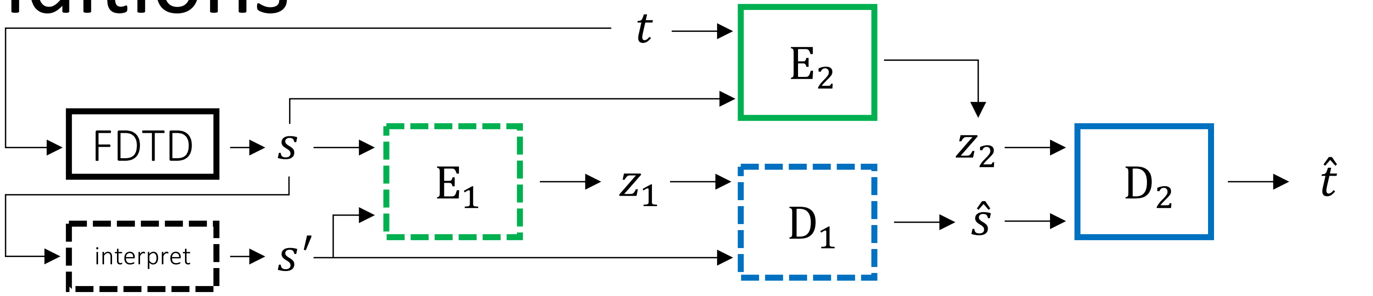
$\bar{a}$: Variable a generated out of an adversarial block

Nested ACVAE

Underscored : Tunable hyperparameters



Loss conditions



$\overline{s'} = \mathrm{A}_1(z_1)$
$\bar{s} = \mathrm{A}_2(z_2)$

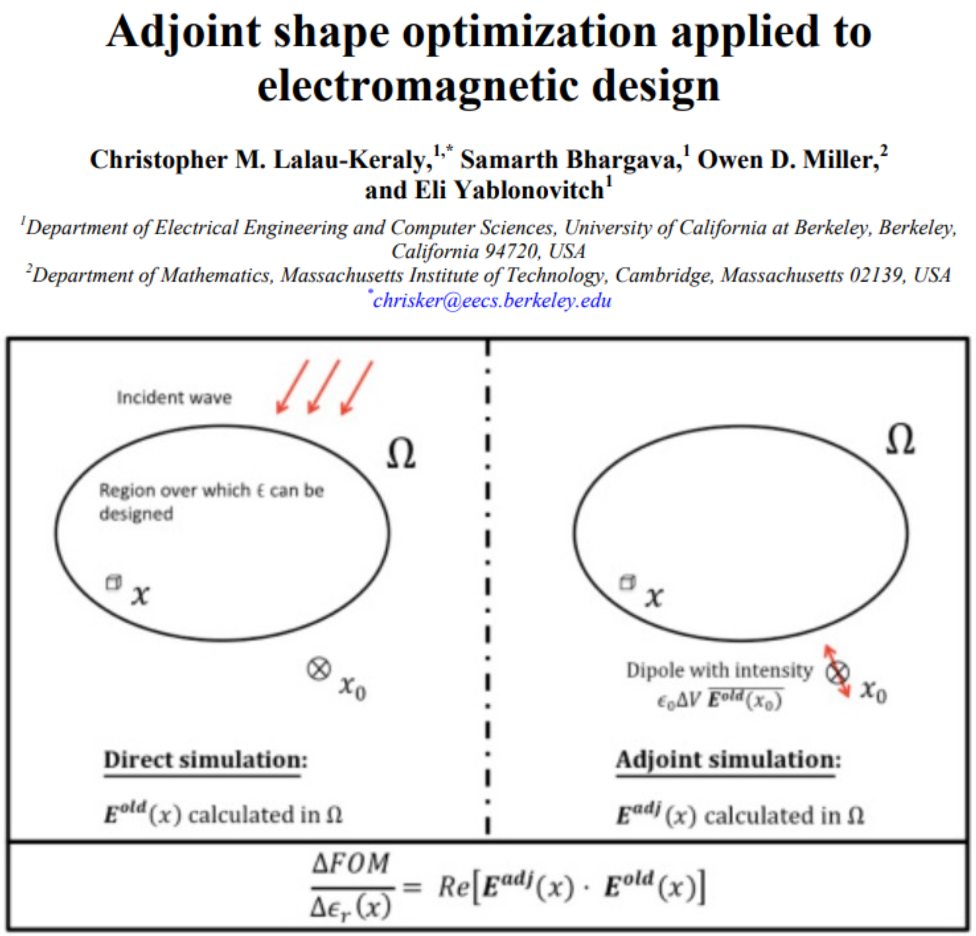
- Basic training losses : $c_1\mathrm{MSE}(x,\hat{x}) + c_2\mathrm{MSE}(s,\hat{s}) + c_3\mathrm{KLD}(z_1) + c_4\mathrm{KLD}(z_2)$
- Adversarial losses : $-c_5\mathrm{MSE}(s',\overline{s'}) - c_6\mathrm{MSE}(\hat{s},\bar{s})$
- Cycle-consistency losses : $c_7\mathrm{MSE}\big(z_1,\mathrm{E}_1(\hat{s};s')\big) + c_8\mathrm{MSE}\big(z_2,\mathrm{E}_2(\hat{t};s)\big)$
- s'-meaning-enforcing loss : $c_9\mathrm{MSE}\big(s',I(\hat{s})\big)$ (I refers to the "interpret" dashed box)

+ higher-order-cycle losses : e.g. $\mathrm{MSE}\big(t,\hat{\hat{t}}\big)$, where $\hat{\hat{t}} = \mathrm{D}_2(z_2 = \mathrm{E}_2(\hat{x},s);\hat{s})$

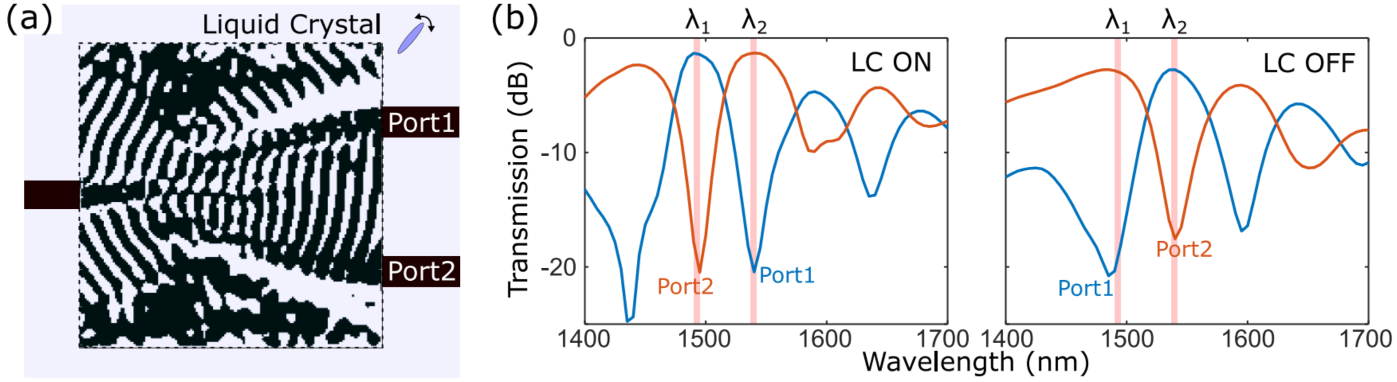
MSE losses can be replaced to any similar losses if necessary. Empirically, I found that $c_2\mathrm{MSE}\big(\sqrt{s},\sqrt{\hat{s}}\big) + c_9\mathrm{MSE}(\log s', \log I(\hat{s}))$ works better than $c_2\mathrm{MSE}(s,\hat{s}) + c_9\mathrm{MSE}\big(s',I(\hat{s})\big)$.

- Ultimate validation loss : $\mathrm{MSE}\Big(s',I\big(F(\hat{t})\big)\Big)$ (F refers to the "FDTD" solid box)

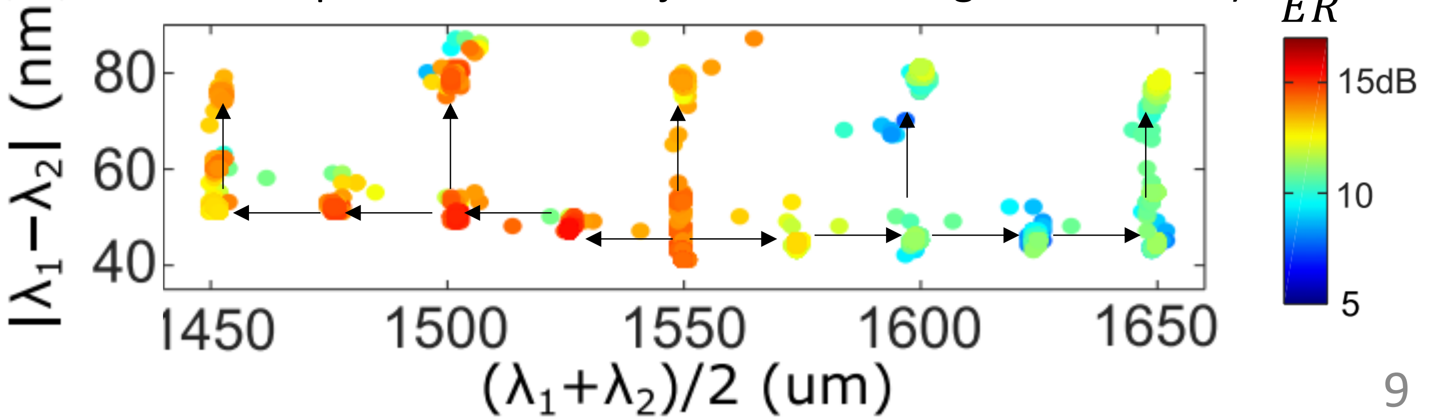
Training dataset preparation


Adjoint shape optimization applied to electromagnetic design

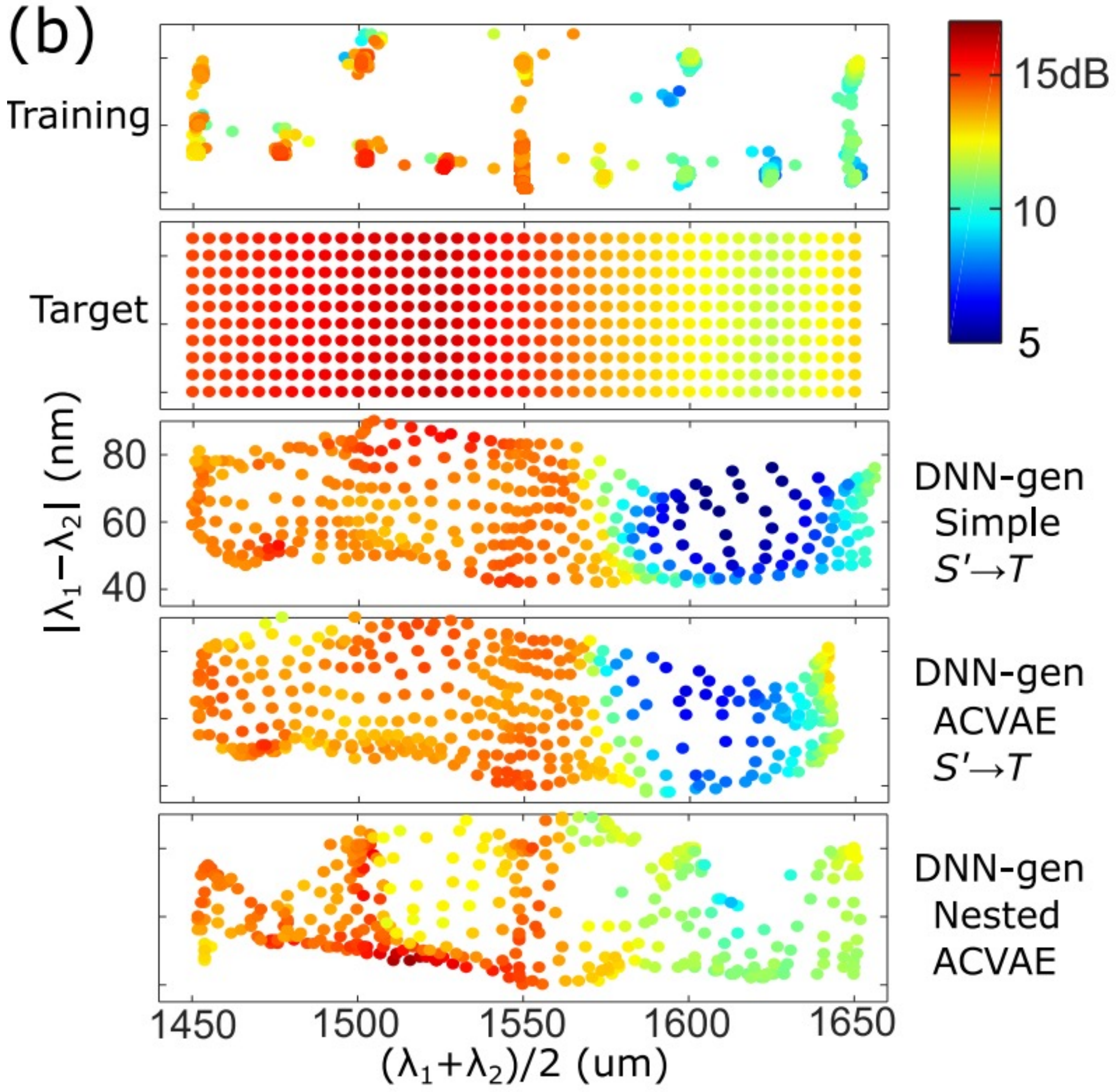
Christopher M. Lalau-Keraly,[1,*] Samarth Bhargava,[1] Owen D. Miller,[2] and Eli Yablonovitch[1]
[1]Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, California 94720, USA
[2]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
*chrisker@eecs.berkeley.edu

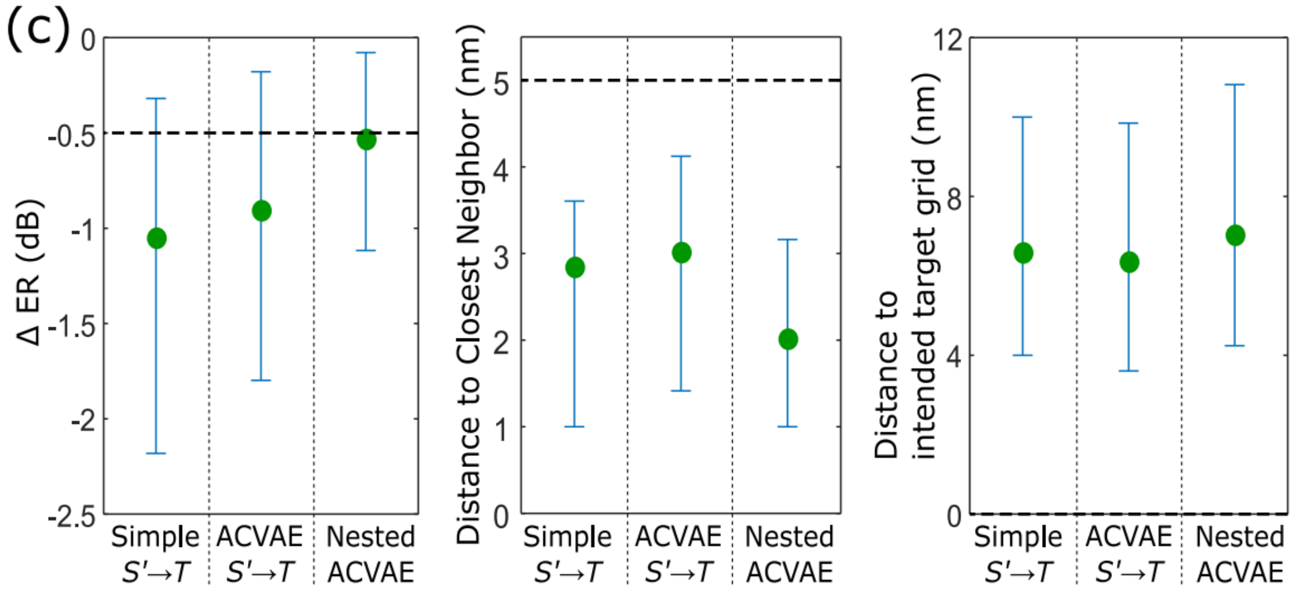


To ensure some interpolatability of dataset,
we cascaded the adjoint optimization
(taking the end result as an initial point
for the optimization for adjacent wavelength conditions)



Network Validation result



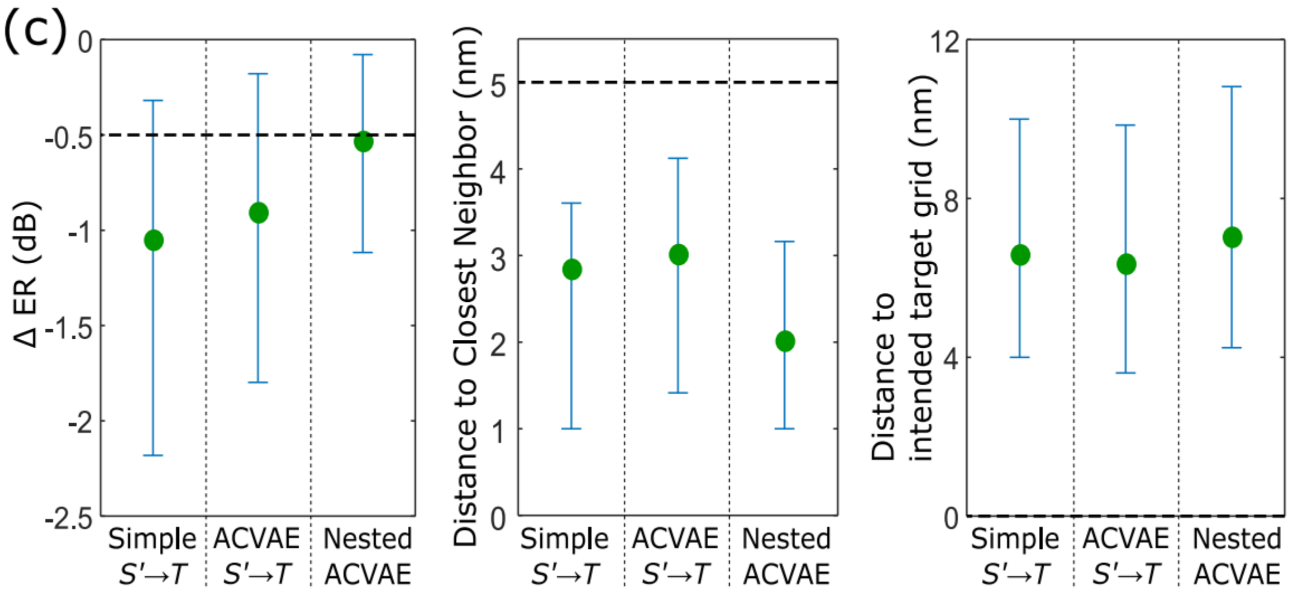


In terms of the extinction ratio of the device,
Nested (*S'*□*S*→*T*) ACAVE performs the best

But, the uniform coverage of optimal device conditions gets a little worse by adding the full spectrum information (*S*)

*Time to run a 3D-adjoint optimization : 10~50 hrs
*Time to train the network : 1~2hrs
*Time to generate a device topology from a trained network, and validate in FDTD : 2~3 mins

Discussion



More validations would be needed to draw conclusive remarks, but

- **We observed that the inclusion of full spectrum information ($\boldsymbol{S}$) helps in terms of better generated ER values**
- **But, it seems that the user-friendly intuitive specs ($\boldsymbol{S'}$; $[\bar{\lambda}, \Delta\lambda, \overline{ER}]$) works better for uniform interpolation**

Conclusion and Outlook

- We demonstrated Auto-Bayes-based network-architecture exploration for optimal design of deep-neural network for complex photonic system (liquid-crystal-tunable wavelength splitter).

- Different architectures show different advantages

- Especially, in a narrow-band wavelength-specific performing photonic devices, the usage of the full spectrum (outside of the wavelength windows that actually matter for the device spec) comes with both plus and minus

- To fully utilize the generative nature of our networks, latent-space optimization (with what values of Z_1 and Z_2 will the generated device performance be optimized?) is an interesting direction to look forward to.