Modeling the X-ARAPUCA's thin-film dichroic filters using physics informed neural networks

M. J. O. Rodrigues, A. T. Bezerra, G. A. Valdiviesso

gustavo.valdiviesso@unifal-mg.edu.br

Poster #256

NEUTRINO 2024

XXXI International Conference on NEUTRINO PHYSICS and ASTROPHYSICS

June 16-22 2024

Milano - Italy

Unifal-MG Universidade Federal de Alfenas

ICT Instituto de Ciência e Tecnologia

The problem at hand

The **X-ARAPUCA** is the latest iteration of a family of devices capable of detecting single photons from liquid argon scintillation, serving as the building blocks of both DUNE's and SBND's Photo Detection System.

Along with the instrumentation for the device, a full physics simulation called **ArapucaSim** was developed to be able to replicate the observed efficiencies of real devices. This makes it a useful tool for looking for alternatives to both materials and geometries.

However, there's a computational bottleneck when simulating the **dichroic filters**, which are in charge of trapping the photons inside the X-ARAPUCA's. This stage of the algorithm plagues the simulation's efficiency.

Here, we present an alternative model that shifts the computational load to a previous stage, namely the training of a **Generative Adversarial Network** (GAN), informed by a realistic optical model. Our hypothesis is that the GAN could learn the general physical behaviour of a given filter while being able to mimic the realistic limitations of manufacturing, including the material characteristics and the thin film's structure.

By achieving a well trained model, only a forward propagation will be necessary during the ArapucaSim runtime, with inputs based on the photon's wavelength and incidence angle, as well as the real filter transmission curve. **This calculation is faster** than using a transfer matrix model and at least as fast as interpolating a fine-grained table with the filter's characteristics.

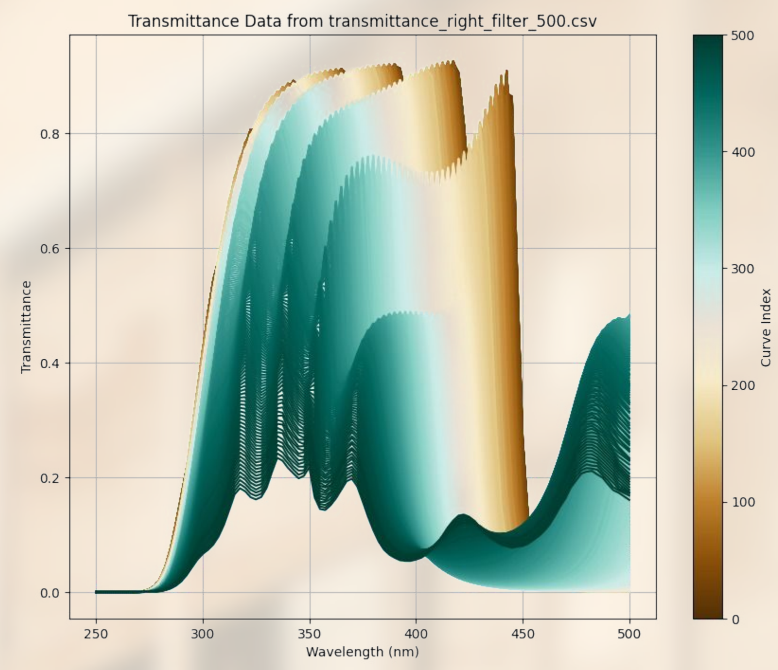
Preliminary Results

The mock filter was represented by a set of 500 transmittance curves, varying in incidence angles from 0 to 85 degrees.

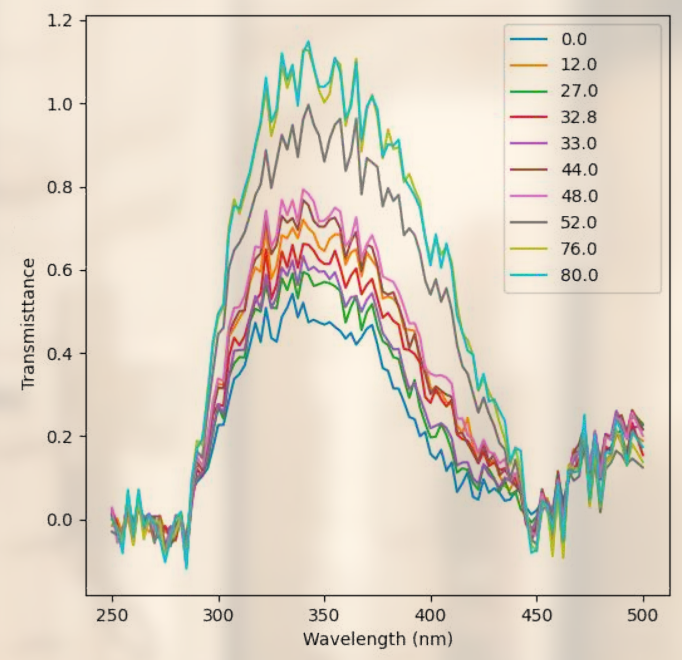
The training section was set so that the only inputs of the model are the wavelength and the incidence angle. In order to probe the efficiency of the model, training sections were 3000 epochs long.

After 3000 epochs of training, the NN model still struggles representing the mock film. Without employing artificial penalties, the model allows for unphysical values for the transmittance, i.e., smaller than zero and larger than one. New training sections are needed, most likely with a different set of hyperparameters.

In order to be successful, our model needs not only to be able to reproduce the mock filter produced by the TTM, but also to do that in meaningful physical way. In the future, this learning will be transferred to a new model where the inputs are expanded to include a limited set of transmission curves from the real filter.

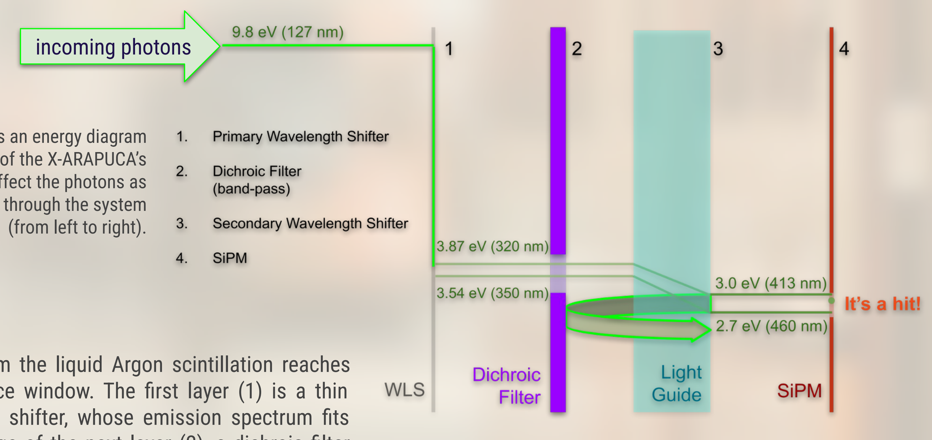


Discrete representation of the mock filter used in this study.



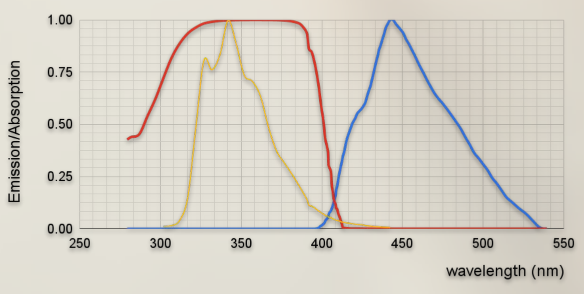
NN model's response as a function of wavelength and incidence angle.

The X-ARAPUCA Light Traps



This is an energy diagram showing how each of the X-ARAPUCA's optical elements affect the photons as they go through the system (from left to right).

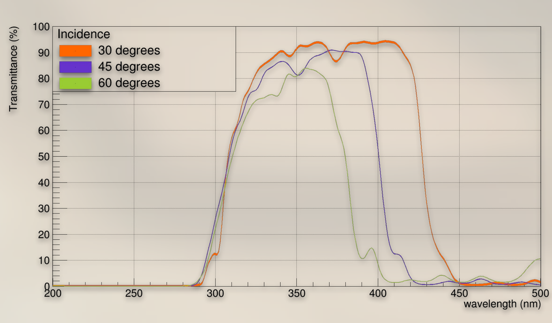
Light coming from the liquid Argon scintillation reaches the device's acceptance window. The first layer (1) is a thin coating of wavelength shifter, whose emission spectrum fits within the passing-range of the next layer (2), a dichroic filter (this work's main object). The third layer is a waveguide (3) doped with a second shifter, which now brings the photons' energies even lower, so that light can't go back through the filter. The converted light if guided to a set of silicon photomultipliers (4).



By design, all the emission and absorption spectra have to match from end to end, connecting the original light signal to the SiPM sensors.

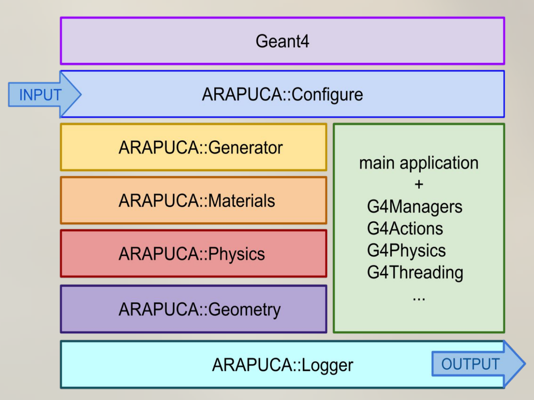
In particular, the first layer (e.g. pTP) emits light isotropically, so that at most 50% is directed inwards. When reaching the second shifter (e.g. EJ286), the doping concentration is calculated so that at least 99% of the light reaching it is converted, following its emission spectrum.

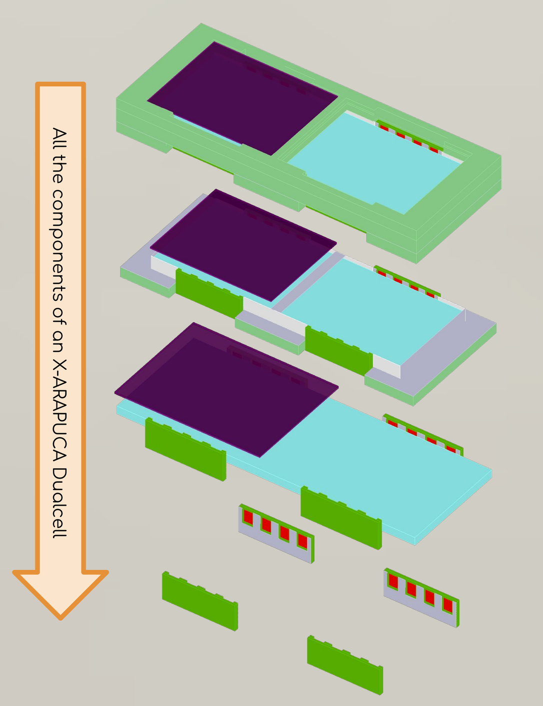
The dichroic filters are composed of a stack of thin films deposited on a glass substrate. The composition of the optical assembly is unknown. The filter's transmission is measured for different angles of incidence. The filter's passing-range is tuned to the previous layer's emission, at 45º of incidence.



The ArapucaSim Software

- It uses the Geant4 framework as its backbone, expanding on its functionality by adding new models and materials specific for the describing the X-ARAPUCA.
- ARAPUCA::Materials is a materials library where all the data from every component is store and organized. Each material representation is validate independently and can be accessible by the simulation during runtime.
- ARAPUCA::Physics is a collection of models designed and validated by experimental data. This library includes a new scintillation model and the model for the dichroic filters.
- ARAPUCA::Geometry is a geometry handler, where each model of the ARAPUCA family is built by calling materials and models from the other libraries. The handler can be used to simulate a single module or place it in more complex setups.

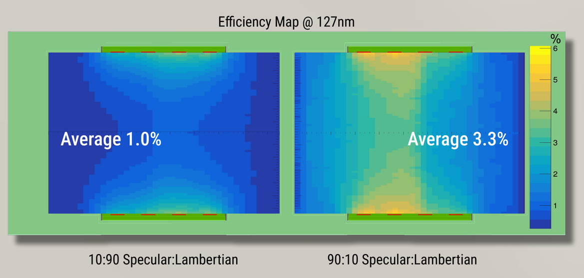




The X-ARAPUCA is a family of devices tailored to each experiment's goals. On the left we see the exploded model of a Dual Cell, the model produced for the Short Baseline Near Detector (SBND).

- G10 frame
- Dichroic filters
- Vikuiti reflectors
- Doped waveguide
- SiPMs

The ArapucaSim is used to study the device's response when materials and geometry are changed, being an integral part of its development cycle.



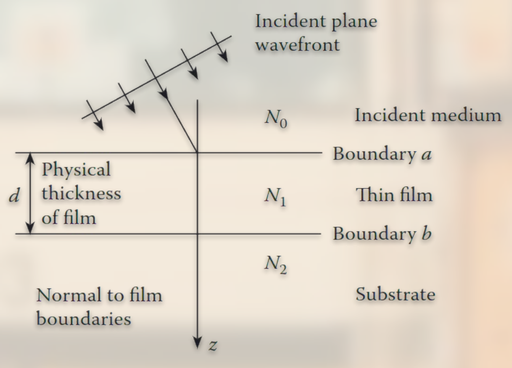
Transfer Matrix Model

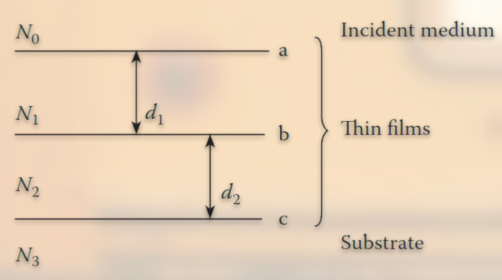
The most accurate physical description of a thin film assembly is the Transfer Matrix Model (TMM). This calculates how the phase of an electromagnetic wave changes when crossing from one medium to another.

Each layer r with complex refraction index $N_r = n + ik$ shifts the wave's phase by δ_r, given by

$$\delta_r = \frac{2\pi N_r d_r \cos\vartheta_r}{\lambda},$$

where the wave λ with incident angle Θ_r crosses a thin layer with thickness d_r.





The incident angle is related to all the internal angles through Snell's law:

$$N_0 \sin\vartheta_0 = N_r \sin\vartheta_r = N_m \sin\vartheta_m$$

The whole assembly with q layers is represented by the product of matrices which transfer the incident wave's phase from the first medium $r=1$ up to the last one $r=q$. This is then applied over the substrate with admittance η_m, resulting on the assembly's characteristic matrix $[B,C]$.

$$\begin{bmatrix} B \\ C \end{bmatrix} = \left\{ \prod_{r=1}^{q} \begin{bmatrix} \cos\delta_r & (i\sin\delta_r)/\eta_r \\ i\eta_r \sin\delta_r & \cos\delta_r \end{bmatrix} \right\} \begin{bmatrix} 1 \\ \eta_m \end{bmatrix}$$

$$R = \left(\frac{\eta_0 B - C}{\eta_0 B + C}\right)\left(\frac{\eta_0 B - C}{\eta_0 B + C}\right)^*,$$

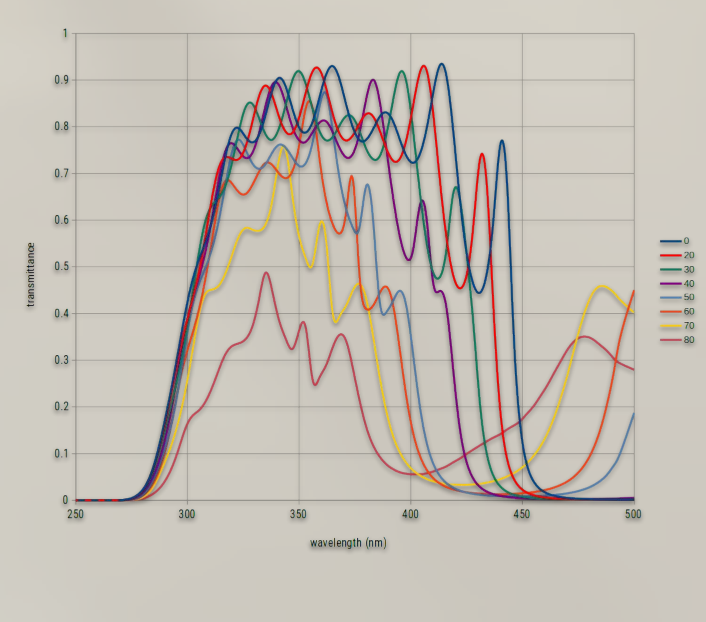
$$T = \frac{4\eta_0 \operatorname{Re}(\eta_m)}{(\eta_0 B + C)(\eta_0 B + C)^*},$$

$$A = \frac{4\eta_0 \operatorname{Re}(BC^* - \eta_m)}{(\eta_0 B + C)(\eta_0 B + C)^*},$$

From the characteristic matrix $[B,C]$ and the admittance of both the incident medium η_0 and the substrate η_m, it is possible to obtain the optical assembly's reflectance R, transmittance T and absorbance A.

Mock Filter Analysis

Here we used the TMM to model a hypothetical filter with only two types of layers, namely silicon dioxide and tantalum pentoxide. The layers' thickness is calculated in order to produce low-pass filter cutting at 450nm. The assembly is 20 layers tall, alternating between the two materials, deposited on top of a boron silicate glass, which has a natural cutoff at 300nm, creating a band-pass filter.

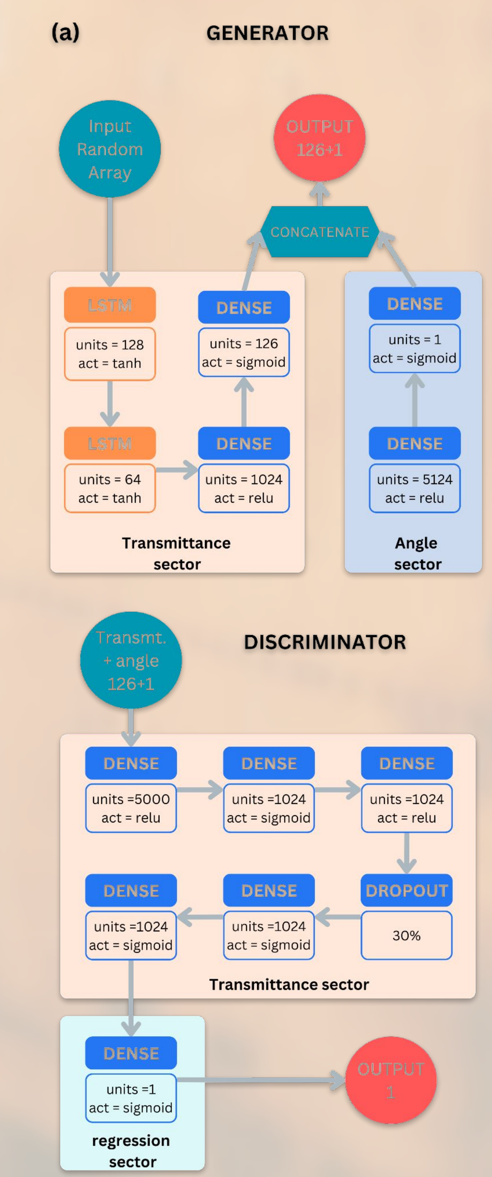


The resulting filter exhibits the desired behaviour. The upper bound of the passing band depends on the incident angle, and so does the total transmittance. The lower bound is independent of incidence since the cutoff is a property of the substrate glass.

Since it is not possible to know the composition of a commercial filter, a mock analysis was carried out, using this hypothetical filter as a model. Random deviations from the ideal design were introduced in order to make it sufficiently different from the model, thus creating a minimal discrepancy to minimise biases coming from the neural network.

Generative Adversarial Networks

Generative Adversarial Neural Networks (GANs) are a specialized architecture used for creating synthetic data that approximates the original samples [4, 3]. They consist of two key neural networks: the generator and the discriminator, both of which are trained together. The discriminator's objective is to differentiate between actual and fake inputs. The generator receives as input a latent array formed of random numbers and is intended to create fake samples that can trick the discriminator. This adversarial process forces both networks to improve their performance iteratively, resulting in the generator creating more realistic data over time.



The GAN model

The generator receives as input the latent random array. It is formed by two main sectors, the first one (Transmittance sector) is responsible for regressing the transmittance and the second sector (angle) is responsible for encoding the angle information. The structured outputs of the sector are concatenated and the generator outputs a 126+1 array, namely, the transmittance curve and the associated angle.

The discriminator receives the concatenated transmission curve and its respective angle and outputs a single classification value.

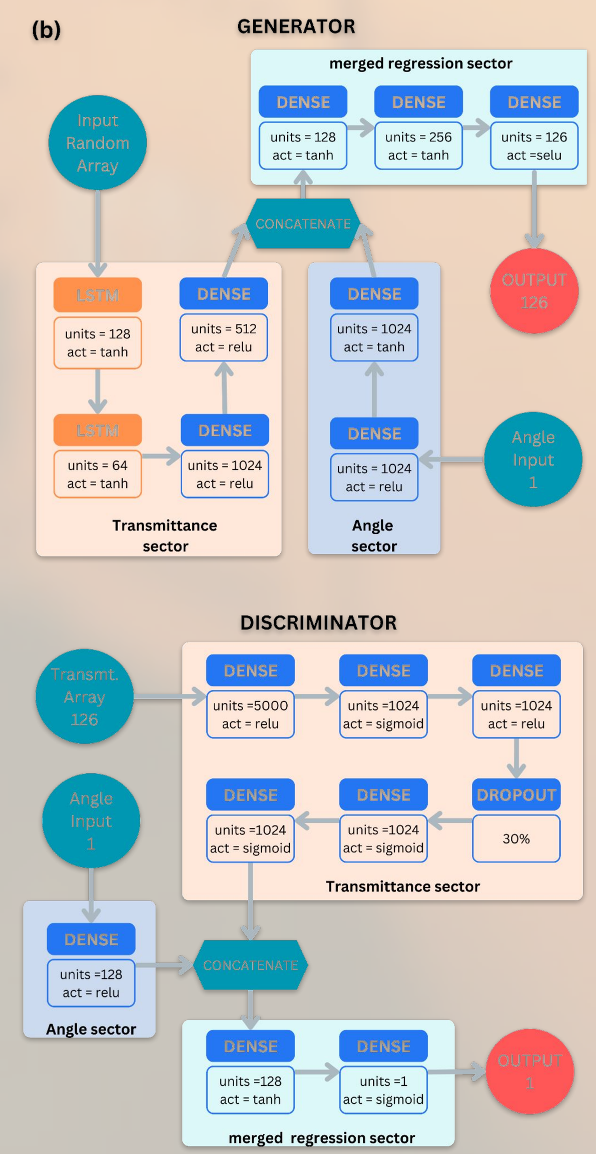
The so-called Wasserstein generative adversarial network (WGAN) is implemented. The WGAN is a specific type of GAN that aims to address the challenges associated with training instabilities and potential collapse as observed in conventional GANs. The Wasserstein distance is employed by WGAN as a metric for quantifying the dissimilarity of probability distributions. The utilization of distance-based techniques ensures the attainment of smoother gradients and, in principle, mitigates the occurrence of vanishing gradients, a common factor contributing to the model collapse.

The WGAN model

In addition to the use of the Wasserstein distance within the loss, also augmentation of the number of layers and neuron count is done. he network consists of two distinct sectors designed to process the transmittance and angle individually. The results for each sector are combined and undergo additional processing layers to regress only the transmittance curve corresponding to the inputted angle. The discriminator is designed to independently monitor the transmittance and angle by utilizing two distinct sectors too. The combined outputs are then forwarded to a final sector, which produces a classification of the inputted transmittance curve and its connection with the inputted incidence angle

The generator receives separately the latent array and the angle as input, and processes them through two split sectors. The sectors' outputs are concatenated together and an additional sequence of dense layers is added for further processing, outputting the respective transmittance curve.

The discriminator also receives the transmittance and angle inputs separately. It processes and merges them. The merged output is further processed by a new sequence of dense layers that output a single classification value.



Acknowledgements

This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES)*, and the *Fundação de Amparo à Pesquisa do Estado de Minas Gerais (APQ-00544-23 FAPEMIG)*.