Neural network assisted design of plasmonic nanostructures on superconducting transition-edge-sensors for single photon detectors

SERGIO G. RODRIGO,[1,2,4] 🆔 CARLOS POBES,[2,3,5] 🆔 MARTA SÁNCHEZ CASI,[2,3] LUIS MARTÍN-MORENO,[2,3] AND AGUSTÍN CAMÓN LASHERAS[2,3]

[1]*Departamento de Física Aplicada, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain*
[2]*Instituto de Nanociencia y Materiales de Aragón (INMA),CSIC-Universidad de Zaragoza, 50009 Zaragoza, Spain*
[3]*Departamento de Física de la Materia Condensada, Universidad de Zaragoza, Zaragoza 50009, Spain*
[4]*sergut@unizar.es*
[5]*cpobes@unizar.es*

Abstract: Transition edge sensors (TESs) are extremely sensitive thermometers made of superconducting materials operating at their transition temperature, where small variations in temperature give rise to a measurable increase in electrical resistance. Coupled to suitable absorbers, they are used as radiation detectors with very good energy resolution in several experiments. Particularly interesting are the applications that TESs may bring to single photon detection in the visible and infrared regimes. In this work, we propose a method to enhance absorption efficiency at these wavelengths. The operation principle exploits the generation of highly absorbing plasmons on the metallic surface. Following this approach, we report nanostructures featuring theoretical values of absorption reaching 98%, at the telecom design frequency (λ = 1550 nm). The optimization process takes into account the TES requirements in terms of heat capacity, critical temperature and energy resolution leading to a promising design for an operating device. Neural networks were first trained and then used as solvers of the optical properties of the nanostructures. The neural network topology takes the geometrical parameters, the properties of materials and the wavelength of light as input, predicting the absorption spectrum at single wavelength as output. The incorporation of the material properties and the dependence with frequency was crucial to reduce the number of required spectra for training. The results are almost indistinguishable from those calculated with a commonly used numerical method in computational electromagnetism, the finite-difference time-domain algorithm, but up to 10^6 times faster than the numerical simulation.

1. Introduction

Transition Edge Sensors (TES) consist of thin films of superconducting materials that translate very small changes in temperature into measurable variations of electrical resistance when polarized right in their superconducting transition [1]. TESs, coupled to proper absorbers, can thus be used as very sensitive thermometers in microcalorimeter devices to provide excellent energy resolution and photon resolving capabilities. When a photon is absorbed, the device temperature changes by a tiny amount, which causes a measurable change in TES resistance. Small differences in the absorbed energy translate into appreciable differences in the resistance change, which provide the excellent resolving power of these sensors. The device can be made

sensitive to a certain energy range with a defined quantum efficiency by selecting a proper material and thickness for the absorber.

The temperature increase caused by the incoming photon can be expressed as $\delta T = E/C$, where E is the absorbed energy and C is the heat capacity of the material. After that, the temperature jumps to $T' = T_0 + \delta T$, where T_0 is a certain base temperature maintained by a thermal bath. The heat capacity C will determine the minimum measurable energy. On the other hand, the device is coupled to the thermal bath through a thermal conductance G. This needs to be also properly chosen in order to set the natural time constant of the system $\tau_0 = C/G$, as this will limit the counting rate. If G is too low, the device will work as an integrator of power, namely, as a bolometer, instead of as a photon counting device. TESs are used in this bolometer mode for instance in Cosmic Microwave Background experiments [2–5]. On the other hand, TESs are being developed in photon counting mode for the detection of radiation in the soft X-ray range, see for instance the Athena (Advanced Telescope for High Energy Astrophysics) project founded by the European Space Agency, whose aim is the study of the Hot and Energetic Universe [6].

TESs can also be interesting for photonic applications [7–9], in the ultraviolet (UV), visible (VIS) and particularly the infrared (IR) range, where optical fibers have very low attenuation at 1550 nm wavelength [10], allowing the exchange of information through very long distances, which is key for many applications in telecommunications. Another possible application is the exchange of quantum keys [11–13], which allows to safely exchange information by having a classical and a quantum channel through which photons will travel. This requires detecting photons individually with the highest possible quantum efficiency.

In this work, we propose a novel TES design for photon counting at the telecom wavelength by use of thin sheets of molybdenum and gold. Although different materials (W, Ir, Ti) are usually used in this energy range [14–18], we have made use of our expertise in the development of Mo/Au TESs for the soft X-ray range [19,20] by simply adapting the absorbing element of the device. In this new design, the high absorption efficiency is achieved by nanostructuring the metallic Mo/Au bilayer, avoiding the use of an extra absorber to reduce the total heat capacity and to simplify the fabrication. Here we show how the corrugation of the superconductor enhances the TES absorbing capabilities thanks to the excitation of Surface Plasmon Polaritons (SPPs) [21].

2. Methods

The nanostructure chosen for this work is a gold sheet placed on a molybdenum layer (both on top of a Si_3N_4 substrate), as shown in Fig. 1(a). The gold sheet is periodically corrugated by grooves (period p). If the gold was not perforated, the Au/Mo bilayer would behave as an almost perfect mirror. However, the grating can allow the coupling of incident light with SPPs, otherwise inhibited due to the mismatch in momentum at the metal surface.

For illustration, Fig. 1(b) shows the amplitude of a SPP excited at 1465 nm, corresponding to the absorption resonance of panel (c) (the geometrical parameters can be found in the figure caption). The excitation of SPPs occurs near $\lambda \approx p$, because the coupling of the incident light with a SPP mode with momentum $k = 2\pi/p$ (provided by reciprocal k-vector of the lattice as stated by the Bloch's theorem) [22]. The local electric field amplitude is several times more intense than that of the incident light, which in simulations is set as 1 kV/cm. The pattern is dominated by the SPP field and the absence of horizontal interference stripes is a clear sign that, at that incident wavelength, virtually no light is reflected. However, as shown in Fig. 1(c) transmission is strongly impaired due to the presence of the molybdenum layer, which is larger than the penetration length of the EM fields inside the metal in all the TES investigated throughout this work. Therefore, the substrate is optically isolated from the bilayer. The calculations were done with the numerical method Finite-Difference Time-Domain (FDTD) [23].
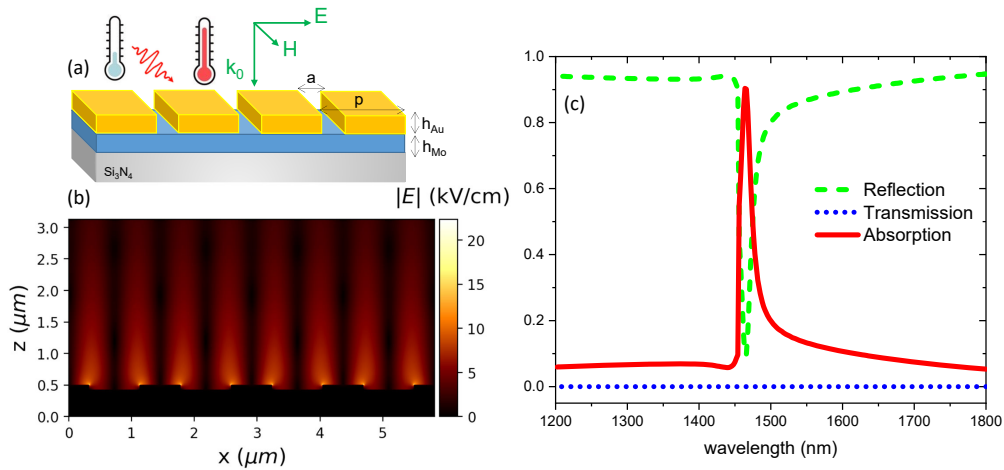
Fig. 1. (a) Schematic visualization of the structure: the radiation impinges on a groove array able to efficiently absorb light at a given wavelength. The nanostructure consists of a gold layer of thickness h_{Au}, perforated periodically with grooves of width a and period p, everything on top of a molybdenum layer (thickness h_{Mo}). The substrate is silicon nitride, which does not play any role in the optical response for molybdenum thicknesses larger than the penetration length of the EM fields inside the metal. (b) This nanostructure supports the excitation of surface plasmons, seen as high electric field amplitude lobes near the slits in the example, corresponding to $\lambda = 1465$ nm, the energy of maximum absorption in the spectra of panel (c). The geometry is defined by $p = 1455$ nm, $a = 795$ nm, $h_{Au} = 70$ nm and $h_{Mo} = 190$ nm. The calculations were conducted with the FDTD method.

The FDTD method is one of the most widely used methods in computational electromagnetism. The FDTD algorithm solves Maxwell's equations numerically and, therefore, provides the entire optical response of a system, including the temporal evolution of the electromagnetic field. This method has the ability to treat different materials, from dielectrics to metals, and different types of sources. Through this algorithm the transmission and reflection coefficients of the system can be obtained, and varying its geometrical parameters we can maximize the absorption at the desired energy. Therefore the whole process of designing the TES could be done with this method, but this is a "brute force" quite inefficient approach when applied in complex systems like the investigated here, as shown later on.

Specifically, the calculation of the nanostructures can be approximated as a 2D system due to traslational invariance along the grooves direction. The nanostructures are discretized with a 5 nm cubic mesh. The system is illuminated by a Gaussian pulse at normal incidence. The polarization is such that the electric field oscillates perpendicular to the face of the grooves (x-direction). This choice is important because such a groove array acts as a polarizer. If the electric field is oriented along the grooves the optical response of the corrugated system does not differ from that of the unmodified system. The material properties of gold and molybdenum are incorporated into FDTD by means of a Drude-Lorentz model, which fitting parameters were obtained from the experimental dielectric constants of Au and Mo in the UV-VIS-IR range and tabulated in Ref. [24,25], respectively. For further details about this particular implementation of the FDTD method see Ref. [26].

Note that we take room temperature values of the dielectric constant for both metals. For temperatures $T>T_c$, up to 4K, a slightly more negative value of the real part of the dielectric constant and a reduction of the imaginary component has been observed experimentally [27]. This effect would be beneficial, increasing the optical quality factor of the nanostructures. For

$T < T_c$ the electrodynamics of metallic superconductors strongly depends on both the temperature and the photon energy [28]. However, metal reflection, transmission and absorption are still the same as in the case of T>Tc, for photon energies higher than the superconducting gap $\Delta(T = 0) = 1.764 \, k_B T_c$. This quantity is $\approx 8.6x10^{-6}$ eV for the expected $T_c \approx 100 \, mK$ operation temperature of the Au/Mo bilayer, which is orders of magnitude lower than 0.8 eV, energy corresponding to $\lambda = 1550$ nm. Therefore, the effect on the surface plasmon dynamics would be well described by the optical response of metals at room temperature

The space of parameters to investigate in the optimization process is large (4 geometric parameters) and the time required to calculate the spectra of a single 2D nanostructure by FDTD is relatively long (5-6 hours per simulation in a Intel Xeon E5-2680 v4 @ 2.40 GHz processor). It is not thus efficient trying to find a set of optimal parameters by trial and error. Note that this is not just a limitation of our particular implementation of the FDTD method; simulation time scales as usual for this kind of system, that is: $A \times (\lambda/q)^3$, where A is the 2D area simulated, λ the wavelength of light and q the mesh size.

Since a few years ago Neural Networks (NN) are catching great attention in all scientific fields, including Nanophotonics, with the potential to boost the solution to several problems. For example NN has been demonstrated to perform better that other methods in inverse design problems dealing with light scattering from multilayer dielectric nanoparticles [29]. In this work we made use of NNs in the process of optimization. We have designed and trained a NN able to calculate optical spectra 10^6 times faster than our FDTD code. Once trained, we used the NN in combination with a classical Gradient Descent (GD) optimizer [30] to find the best choice of parameters that produce the maximum absorption at 1550 nm wavelength.

One of the main advantages of this process is that with the NN, once trained, it can be used in any optimization process with the same speed factor with respect the numerical method used to get the training data. In our case, the computational time will substantially increase if 3D simulations were conducted instead. However, a NN will do the calculation at the same speed that in the 2D case.

Our NN implementation uses *Tensorflow* [31], which consists of a Python deep-learning work environment to define and train almost any type of architecture. In this case, we used a specialized library for the tensor calculations based on Keras [32].

3. Results

In order to train the NN, 2600 spectra (transmission and reflection) were generated with the FDTD algorithm. The wavelength range goes from 1200 nm to 1800 nm in steps of 2 nm (300 points per spectrum). These spectra were obtained from groove arrays defined by randomly selecting the geometrical parameters within these intervals: period $p \in [1300$ nm, 1700 nm];groove width $a \in [0, p]$; gold thickness $h_{Au} \in [20$ nm, 400 nm] and molybdenum thickness $h_{Mo} \in [20$ nm, 200 nm]. This parameter space is compatible with fabrication requirements.

The NN is trained with 2400 spectra of the 2600 generated by the FDTD, the rest are left for test and validation (one hundred in each case, which is small as compared to typical datasets). The NN topology (see schematics in Fig. 2) which better performed in our investigations consists of an *input* layer having 9 neurons, corresponding to the 4 geometric parameters (h_{Au}, h_{Mo}, a and p), the dielectric constants of the materials (ϵ_{Au}^i, ϵ_{Au}^r, ϵ_{Mo}^i, ϵ_{Mo}^r) and the wavelength λ. The *output* layer consists of one neuron that provide absorption (A) at λ. The rest of the NN consists of 2 hidden layers of 40 neurons each, all dense, which in this environment means that they are completely connected. Sigmoid function is applied for neuron activation for all the layers. The cost function *binary cross-entropy* was used. To evaluate the good functioning of the network we used the mean squared error (MSE). Finally the stochastic gradient descent (SGD) was the chosen optimizer [30].
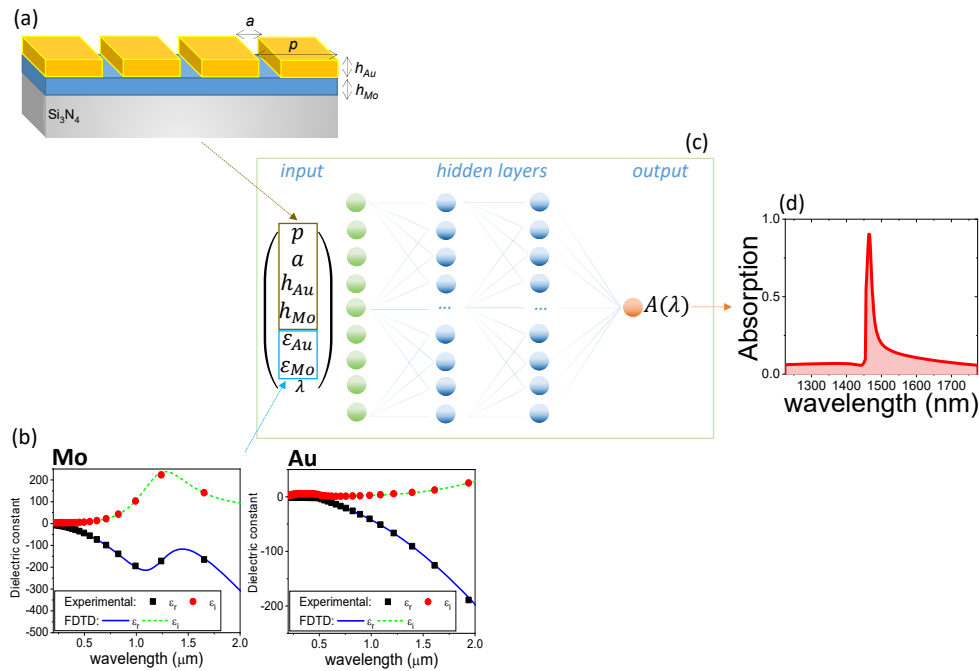
Fig. 2. The NN model takes as *input*: i) the geometrical parameters of the nanostructure (h_{Au}, h_{Mo}, a and p), shown in (a); the dielectric constants of the materials (ϵ_{Au}^i, ϵ_{Au}^r, ϵ_{Mo}^i, ϵ_{Mo}^r), which experimental tabulated values are compared with those used in calculations in (b); and iii), the wavelength of light λ. c) The NN topology in this work consist in two hidden layers with 40 neurons each, fully connected. The sigmoid function is applied for neuron activation for all the layers. The cost function *binary cross-entropy* was chosen. Finally, the *output* layer consists of one neuron that provides absorption at λ, illustrated in (d) with an spectrum obtained by iteration of the trained NN in λ.

The following optimal values for the hyperparameters were found: *mini-batch size* = 64, *learning rate* $\eta = 1.0$, regularization parameter $\lambda = 0.1$, momentum = 0.45, *decay* = $1 \cdot 10^{-6}$ and *Nesterov* = True. The momentum, together with the *decay* and *Nesterov*, are characteristic parameters of the SGD optimizer that offers *Keras*, to improve the convergence process towards the minimum of the function to be optimized [31].

In Fig. 3(a) the evolution of the MSE function can be seen against the number of epochs for the 2600 spectra training and test datasets. Values very close to zero are reached in both cases, with even lower values for the test data set, showing that no overfitting is taking place. The inset of Fig. 3(a) depicts MSE as a function of the size of the dataset, revealing that only a few hundred spectra would be enough to achieve comparable MSE that of the trained NN, used herein. The chosen training dataset does not imply a minimum size for it. In fact, our approach based on the use of NN would properly work with less data. The good performance of our trained NN can be seen in Fig. 3(b) where the dispersion curve of absorption at $\lambda = 1550$ nm (predicted by the NN vs true results from FDTD numerical calculations) is depicted for the 2400 training data spectra. Interestingly, a randomly chosen set of geometrical parameters will most likely lead to low absorption values, which reinforces the idea of using a NN as a Maxwell's equation "solver" in the optimization process instead of a canonical numerical method. To get some insight into how a given MSE value represents the accuracy of our trained NN, Fig. 3(c) shows the best and the worst FDTD versus NN comparisons as defined by MSE. In the best situation the agreement

between both results is excellent, but even in the *worst* case the resemblance between the results of both methods is remarkable.
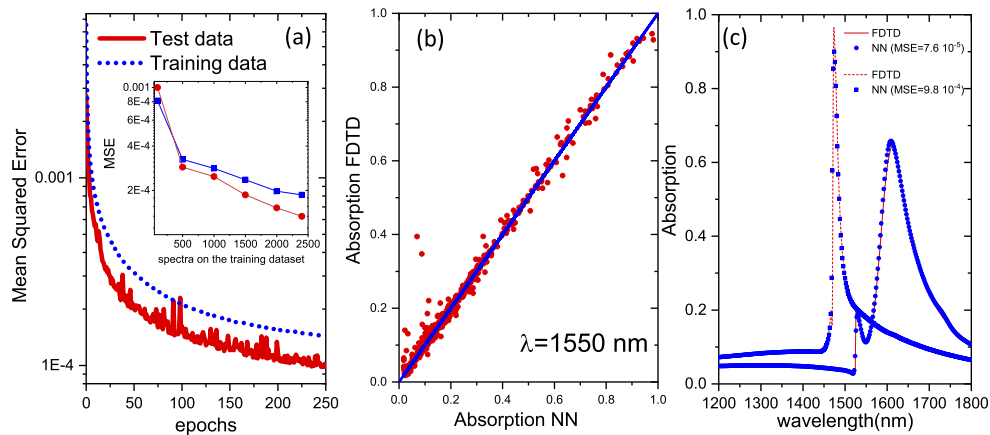

Fig. 3. (a) MSE evolution versus the number of epochs for training data (blue curve) and test data (red curve). This is obtained through the NN, implemented in the Keras environment, described in the text. The inset shows MSE as a function of the number of spectra on the training dataset. (b) Dispersion representation of training data showing NN results for absorption as compared to those obtained with the FDTD method. (c) Spectra obtained by the NN for two specific cases of the test data set. The lowest MSE corresponds to $p = 1525$ nm, $a = 615$ nm, $h_{Au} = 145$ nm, $h_{Mo} = 150$ nm, and the highest to $p = 1470$ nm, $a = 940$ nm, $h_{Au} = 115$ nm, $h_{Mo} = 100$ nm. The best agreement between the NN prediction and the FDTD result is obtained for the case with the lower MSE, as expected.

With the GD algorithm a great number of iterations were necessary to obtain an optimized design. We seeded the GD algorithm with fifty sets of geometrical parameters. On average, 1000 iterations were needed to converge. The whole process was possible only thanks to the trained NN; it would have been unpractical with the numerical method.

Optimization of the absorption is not the only issue to be considered for designing a functional TES. The operating temperature, the heating capacity and the spectral resolution are parameters that have to be finely tuned so that a given TES design eventually works, because both quantities depend on the Mo and Au thicknesses.

To achieve the appropriate heat capacity and optimal sensitivity, TESs work at very low temperatures. Normally, it is not possible to find superconducting materials with the required operating temperature, so an effect known as the proximity effect [33] is used, in which a normal metal is deposited on a superconductor reducing its critical temperature T_c. Thus, the final T_c of the bilayer depends on the relative proportion of thicknesses.

In addition, to detect a photon absorbed by the sensor, it has to produce a measurable temperature change as dictated by the expression $\delta T = E/C$. A TES can be sensitive to temperature variations of up to a few μK, and for that values, the heat capacity must be around $10\ fJ/K$. So in order to observe the photon absorption event at the IR ($E < 1\ eV$) there are constrains on the lateral size of the device. There must be a balance between the plasmonic response (the more periods, the better) and the heat capacity, that has to be low enough. We estimated that a compromise is achieved taking this size equal to $20 \times p$.

On the other hand, although the nanostructure is tuned and optimized for a single wavelength (1550 nm), energy resolution is still important. The Poisson nature of sources leads to the possibility to detect multiphoton events within the time resolution of the device. For applications where resolving the number of photons is important (and where TESs can actually provide an

added value), the device needs to guarantee a good enough energy resolution to disentangle the number of photons in each detection (the expected energy detected will be $E = mE_0$ with m the number of photons and E_0 the energy observed when only one photon is detected). This is not as restrictive as in spectroscopic applications, at least for low m, but needs to be considered. We use a common expression to estimate the expected energy resolution for our devices:

$$\Delta E = 2.355 \times \left[4 \times Kb \times T_c^2 \times C/\alpha \times \sqrt{n/2} \right]^{1/2}$$

where T_c and C are the critical temperature and heat capacity of the device. α and n are parameters related to the TES model [1]. In order to get an estimate of T_c we use Martinis expression [33]:

$$T_c = T_{c0} \left[\frac{d_{Mo}}{d_0 \times 1.13 \times (1 + 1/r) \times t} \right]^r$$

where T_{c0} is the bulk critical temperature of molybdenum, d_{Mo} is the molybdenum thickness, d_0 is a constant which depends on the density of states of molybdenum, r is proportional to the ratio of thicknesses of molybdenum and gold, and t is the transparency of the interface, a free parameter with values in the range (0,1).

From this, we can also make an estimate for the electronic heat capacity of the sensor in the normal state for every combination of thicknesses obtained from the neural network and use the obtained values of T_c and C to estimate the expected resolution:

$$C = T_c \times Area \times [d_{Mo} \times \gamma_{Mo} \times (\rho/A)_{Mo} + d_{Au} \times \gamma_{Au} \times (\rho/A)_{Au}]$$

where γ is the molar specific heat, and ρ/A is the mass density divided by the atomic weight [1] .

Figure 4 summarized the optimization process. In Fig. 4(a), the left axis shows the absorption of the groove array at $\lambda = 1550$ nm and the rest of vertical axes are for the geometrical parameters, T_c, C and ΔE, which ultimately determines the TES performance. The GD algorithm was powered by the NN corresponding to optimized groove arrays to operate at $\lambda = 1550$ nm.

From the results of Fig. 4, it is clear that there is a range of geometrical parameters which provides high absorption ($\approx 98\%$). To illustrate and demonstrate the validity of our approach using a NN, we compared in Fig. 4(b) the absorption spectrum obtained with both our trained NN and the FDTD method, in this case for a grating with $p = 1490$ nm, $a = 260$ nm, $h_{Au} =110$ nm and $h_{Mo} =30$ nm, showing excellent agreement. Note that these values are obtained by rounding off those returned by the optimizer and that best fit the discretized grid used in the FDTD method (5 nm). It has been proven that the change in the absorption curve is minimal when passing from nominal values obtained by the optimizer to the values used in the FDTD calculations. The chosen value for h_{Mo} is also compatible with fabrication and is the lower limit to avoid 2D effects [34].

We also verified that the value obtained for the absorption is not very sensitive to the variation of the parameters of the TES. We have found that a change of ± 5 nm/10 nm in the thickness of the gold or in the width of the period, causes a shift of the position of the peak in, approximately, the same amount.

From the geometrical parameters of the example in Fig. 4(b) we can calculate the expected T_c, C and ΔE:

$$T_C = 108 \; mK; C = 1.4 \; fJ/K; \Delta E = 0.08 \; eV$$

This calculated values are very promising for an operating device both with respect operating temperature and expected energy resolution, which would allow photon resolving capabilities at the design wavelength of 1550nm ($E_0 = 0.8eV$ versus $\Delta E = 0.08 \; eV$). Our group plans to fabricate and test the first devices guided by this design procedure in a near future.

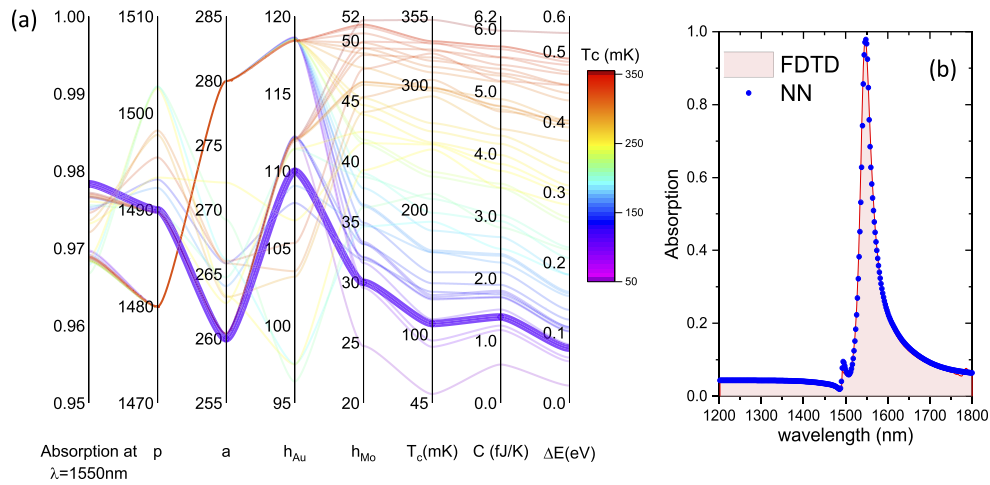**Fig. 4.** (a) The parallel plot summarizes the results of the best trials in terms of absorption, T_c, C and ΔE out of the series of optimization processes conducted. Left axis shows the absorption of the groove array at $\lambda = 1550$ nm and the rest of vertical axis are for the geometrical parameters (in nm), T_c, C and ΔE, variables that ultimately determine TES performance. (b) Optimized groove array (highlighted in (a) with thick line): absorption spectra predicted by the trained NN (blue dots) and the calculated with the FDTD method (red line) for $p = 1490$ nm, $a = 260$ nm, $h_{Au} = 110$ nm and $h_{Mo} = 30$ nm. Absorption reaches almost 98% at the peak.

4. Conclusions

A plasmonic nanostructure based on a periodically corrugated metallic bilayer of gold and molybdenum has been proposed as the absorbing component of a TES sensitive to telecom radiation. The nanostructure achieves a theoretical absorption efficiency of 98% at $\lambda = 1550$ nm and constitutes a promising design to use TESs in this important spectral range. The optimization steps were enabled by the use of NNs. We trained a NN to calculate the light absorbed by the nanostructure speeding the process of calculating each absorption spectrum up to 10^6 times over a classical numerical algorithm like the one used in this work based on the FDTD method. With the rise of quantum applications, single photon detectors with high absorption efficiency in the telecom are needed. TESs emerge as a promising technology. Their energy resolving capabilities and the possibility to tune the sensitivity to other wavelengths make them interesting also in spectroscopic applications in the infrared and visible range.

References

1. K. Irwin and G. Hilton, *Transition-Edge Sensors* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005), pp. 63–150.
2. G. Wang, V. Yefremenko, V. Novosad, L. Bleem, C. L. Chang, J. McMahon, A. Datesman, J. Pearson, R. Divan, T. Downes, A. T. Crites, S. S. Meyer, and J. E. Carlstrom, "Development of absorber coupled TES polarimeter at millimeter wavelengths," IEEE Trans. Appl. Supercond. **19**(3), 544–547 (2009).
3. J. Ding, P. A. R. Ade, A. J. Anderson, J. Avva, Z. Ahmed, K. Arnold, J. E. Austermann, A. N. Bender, B. A. Benson, L. E. Bleem, K. Byrum, J. E. Carlstrom, F. W. Carter, C. L. Chang, H. M. Cho, J. F. Cliche, A. Cukierman, D.

Czaplewski, R. Divan, T. de Haan, M. A. Dobbs, D. Dutcher, W. Everett, A. Gilbert, R. Gannon, R. Guyser, N. W. Halverson, N. L. Harrington, K. Hattori, J. W. Henning, G. C. Hilton, W. L. Holzapfel, J. Hubmayr, N. Huang, K. D. Irwin, O. Jeong, T. Khaire, D. Kubik, C. L. Kuo, A. T. Lee, E. M. Leitch, S. S. Meyer, C. S. Miller, J. Montgomery, A. Nadolski, T. Natoli, H. Nguyen, V. Novosad, S. Padin, Z. Pan, J. Pearson, C. M. Posada, A. Rahlin, C. L. Reichardt, J. E. Ruhl, B. R. Saliwanchik, J. T. Sayre, J. A. Shariff, I. Shirley, E. Shirokoff, G. Smecher, J. Sobrin, I. Stan, A. A. Stark, K. Story, A. Suzuki, Q. Y. Tang, R. B. Thakur, K. L. Thompson, C. Tucker, K. Vanderlinde, J. D. Vieira, G. Wang, N. Whitehorn, W. L. K. Wu, V. Yefremenko, and K. W. Yoon, "Optimization of transition edge sensor arrays for cosmic microwave background observations with the south pole telescope," IEEE Trans. Appl. Supercond. **27**(4), 1–4 (2017).

4. C. M. Posada, P. A. Ade, Z. Ahmed, A. J. Anderson, J. E. Austermann, J. S. Avva, R. B. Thakur, A. N. Bender, B. A. Benson, J. E. Carlstrom, F. W. Carter, T. Cecil, C. L. Chang, J. F. Cliche, A. Cukierman, E. V. Denison, T. de Haan, J. Ding, R. Divan, M. A. D. lessD. Dutcher, W. Everett, A. Foster, R. N. Gannon, A. Gilbert, J. C. Groh, N. W. Halverson, A. H. Harke-Hosemann, N. L. Harrington, J. W. Henning, G. C. Hilton, W. L. Holzapfel, N. Huang, K. D. Irwin, O. B. Jeong, M. Jonas, T. Khaire, A. M. Kofman, M. Korman, D. Kubik, S. Kuhlmann, C. L. Kuo, A. T. Lee, A. E. Lowitz, S. S. Meyer, D. Michalik, C. S. Miller, J. Montgomery, A. Nadolski, T. Natoli, H. Nguyen, G. I. Noble, V. Novosad, S. Padin, Z. Pan, J. Pearson, A. Rahlin, J. E. Ruhl, L. J. Saunders, J. T. Sayre, I. Shirley, E. Shirokoff, G. Smecher, J. A. Sobrin, L. Stan, A. A. Stark, K. T. Story, A. Suzuki, Q. Y. Tang, K. L. Thompson, C. Tucker, L. R. Vale, K. Vanderlinde, J. D. Vieira, G. Wang, N. Whitehorn, V. Yefremenko, K. W. Yoon, and M. R. Young, "Fabrication of detector arrays for the SPT-3G receiver," J. Low Temp. Phys. **193**(5-6), 703–711 (2018).

5. W. B. Everett, P. A. R. Ade, Z. Ahmed, A. J. Anderson, J. E. Austermann, J. S. Avva, R. B. Thakur, A. N. Bender, B. A. Benson, J. E. Carlstrom, F. W. Carter, T. W. Cecil, C. L. Chang, J. F. Cliche, A. Cukierman, E. V. Denison, T. de Haan, J. Ding, M. A. Dobbs, D. Dutcher, A. Foster, R. N. Gannon, A. J. Gilbert, J. C. Groh, N. W. Halverson, A. H. Harke-Hosemann, N. L. Harrington, J. W. Henning, G. C. Hilton, W. L. Holzapfel, N. Huang, K. D. Irwin, O. B. Jeong, M. Jonas, T. S. Khaire, A. M. Kofman, M. M. Korman, D. L. Kubik, S. E. Kuhlmann, C. lin Kuo, A. T. Lee, A. E. Lowitz, S. S. Meyer, D. Michalik, J. Montgomery, A. Nadolski, T. J. Natoli, H. T. Nguyen, G. I. Noble, V. Novosad, S. Padin, Z. Pan, J. E. Pearson, C. M. Posada, A. S. Rahlin, J. E. Ruhl, L. J. Saunders, J. T. Sayre, I. Shirley, E. Shirokoff, G. Smecher, J. A. Sobrin, A. A. Stark, K. T. Story, A. Suzuki, Q. Y. Tang, K. L. Thompson, C. E. Tucker, L. R. Vale, K. Vanderlinde, J. D. Vieira, G. Wang, N. Whitehorn, V. Yefremenko, K. W. Yoon, and M. R. Young, "Design and bolometer characterization of the SPT-3G first-year focal plane," J. Low Temp. Phys. **193**(5-6), 1085–1093 (2018).

6. "The x-ray integral field unit (x-IFU) for athena," in *Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray*, T. Takahashi, J.-W. A. den Herder, and M. Bautz, eds. (SPIE, 2014).

7. B. Cabrera, R. M. Clarke, P. Colling, A. J. Miller, S. Nam, and R. W. Romani, "Detection of single infrared, optical, and ultraviolet photons using superconducting transition edge sensors," Appl. Phys. Lett. **73**(6), 735–737 (1998).

8. B. Cabrera, J. M. Martinis, A. J. Miller, S. W. Nam, and R. Romani, "TES spectrophotometers for near IR/optical/UV," AIP Conf. Proc. **605**, 565–570 (2002).

9. T. Gerrits, A. Lita, B. Calkins, and S. W. Nam, "Superconducting transition edge sensors for quantum optics," in *Superconducting Devices in Quantum Optics, Quantum Science and Technology* (Springer, 2016).

10. D. Rosenberg, A. E. Lita, A. J. Miller, S. Nam, and R. E. Schwall, "Performance of photon-number resolving transition-edge sensors with integrated 1550 nm resonant cavities," IEEE Trans. Appl. Supercond. **15**(2), 575–578 (2005).

11. R. H. Hadfield, "Single-photon detectors for optical quantum information applications," Nat. Photonics **3**(12), 696–705 (2009).

12. R. H. Hadfield, J. L. Habif, J. Schlafer, R. E. Schwall, and S. W. Nam, "Quantum key distribution at 1550nm with twin superconducting single-photon detectors," Appl. Phys. Lett. **89**(24), 241129 (2006).

13. P. A. Hiskett, D. Rosenberg, C. G. Peterson, R. J. Hughes, S. Nam, A. E. Lita, A. J. Miller, and J. E. Nordholt, "Long-distance quantum key distribution in optical fibre," New J. Phys. **8**(9), 193 (2006).

14. A. Miller, S. Nam, J. Martinis, and A. Sergienko, *Tungsten Transition-Edge Sensors for IR/Optical/UV Photon Counting* (Universal Academy Press, Inc., Tokyo, JA, 2003).

15. A. E. Lita, B. Calkins, L. A. Pellochoud, A. J. Miller, and S. Nam, "High-efficiency photon-number-resolving detectors based on hafnium transition-edge sensors," AIP Conf. Proc. **1185**, 351–354 (2009).

16. L. Lolli, E. Taralli, C. Portesi, M. Rajteri, and E. Monticone, "Aluminum-titanium bilayer for near-infrared transition edge sensors," Sensors **16**(7), 953 (2016).

17. D. Fukuda, G. Fujii, T. Numata, K. Amemiya, A. Yoshizawa, H. Tsuchida, H. Fujino, H. Ishii, T. Itatani, S. Inoue, and T. Zama, "Titanium-based transition-edge photon number resolving detector with 98% detection efficiency with index-matched small-gap fiber coupling," Opt. Express **19**(2), 870–875 (2011).

18. M. Rajteri, C. Portesi, E. Taralli, and E. Monticone, "How to avoid reflection losses in superconducting light detectors," J. Mod. Opt. **56**(2-3), 385–389 (2009).

19. C. Pobes, L. Fàbrega, A. Camón, N. Casañ-Pastor, P. Strichovanec, J. Sesé, J. Moral-Vico, and R. M. J. Calleja, "Development of cryogenic x-ray detectors based on Mo/Au transition edge sensors," IEEE Trans. Appl. Supercond. **27**(4), 1–5 (2017).

20. C. Pobes, L. Fàbrega, A. Camón, P. Strichovanec, J. Moral-Vico, N. Casañ-Pastor, R. M. Jáudenes, and J. Sesé, "Comparison of different Mo/Au TES designs for radiation detectors," J. Low Temp. Phys. **193**(3-4), 282–287 (2018).

21. W. Barnes, A. Dereux, and T. Ebbesen, "Surface plasmon subwavelength optics," Nature **424**(6950), 824–830 (2003).
22. F. J. Garcia-Vidal, L. Martin-Moreno, T. W. Ebbesen, and L. Kuipers, "Light passing through subwavelength apertures," Rev. Mod. Phys. **82**(1), 729–787 (2010).
23. A. Taflove and S. C. Hagness, *Computational electrodynamics: the finite-difference time-domain method*; 3rd ed., Artech House antennas and propagation library (Artech House, Boston, MA, 2005).
24. F. Hao and P. Nordlander, "Efficient dielectric function for FDTD simulation of the optical properties of silver and gold nanoparticles," Chem. Phys. Lett. **446**(1-3), 115–118 (2007).
25. W. S. M. Werner, K. Glantschnig, and C. Ambrosch-Draxl, "Optical constants and inelastic electron-scattering data for 17 elemental metals," J. Phys. Chem. Ref. Data **38**(4), 1013–1092 (2009).
26. S. G. Rodrigo, *Optical properties of nanostructured metallic systems* (Springer Berlin Heidelberg, 2012).
27. S. V. Jayanti, J. H. Park, A. Dejneka, D. Chvostova, K. M. McPeak, X. Chen, S.-H. Oh, and D. J. Norris, "Low-temperature enhancement of plasmonic performance in silver films," Opt. Mater. Express **5**(5), 1147–1155 (2015).
28. M. Dressel, "Electrodynamics of metallic superconductors," Adv. Condens. Matter Phys. **2013**, 1–25 (2013).
29. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljacic, "Nanophotonic particle simulation and inverse design using artificial neural networks," Sci. Adv. **4**(6), eaar4206 (2018).
30. M. Nielsen, *Neural networks and deep learning* (Springer, 2013).
31. "TensorFlow: Large-scale machine learning on heterogeneous systems," (2015). Software available from tensorflow.org.
32. F. Chollet, *Deep learning with python* (Manning, 2017).
33. J. M. Martinis, G. C. Hilton, K. D. Irwin, and D. A. Wollman, "Calculation of Tc in a normal-superconductor bilayer using the microscopic-based Usadel theory," Nucl. Instrum. Methods Phys. Res., Sect. A **444**(1-2), 23–27 (2000).
34. L. Fàbrega, A. Camon, I. Fernandez-Martinez, J. Sesé, M. Parra-Borderias, O. Gil, R. Gonzalez-Arrabal, J. Costa-Krämer, and F. Briones, "Size and dimensionality effects in superconducting Mo thin films," Supercond. Sci. Technol. **24**(7), 075014 (2011).