UC Santa Cruz

UC Santa Cruz Previously Published Works

Title

Modeling of nanohole silicon pin/nip photodetectors: Steady state and transient characteristics

Permalink

https://escholarship.org/uc/item/66f2w9c7

Journal

Nanotechnology, 32(36)

ISSN

0957-4484

Authors

Yamada, Toshishige
Devine, Ekaterina Ponizovskaya
Ghandiparsi, Soroush
et al.

Publication Date

2021-09-03

DOI

10.1088/1361-6528/abfb98

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Nanotechnology

ACCEPTED MANUSCRIPT

Modeling of nanohole silicon pin/nip photodetectors: Steady state and transient characteristics

To cite this article before publication: Toshishige Yamada *et al* 2021 *Nanotechnology* in press https://doi.org/10.1088/1361-6528/abfb98

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2021 IOP Publishing Ltd**.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.
As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 128.114.34.22 on 01/05/2021 at 16:37

Modeling of Nanohole Silicon *pin*/*nip* Photodetectors: Steady State and Transient Characteristics

Authors: Toshishige Yamada[1,2*], Ekaterina Ponizovskaya Devine[2,3], Soroush Ghandiparsi[3], Cesar Bartolo-Perez[3], Ahmed S. Mayet[3], Hilal Cansizoglu[3], Yang Gao[3], Ahasan Ahamed[3], Shih-Yuan Wang[2], and M. Saif Islam[3].

Affiliations:
[1]Electrical and Computer Engineering, Baskin School of Engineering, University of California, Santa Cruz, Santa Cruz, California 95064,
[2]W&WSens Devices Inc., 4546 El Camino Suite 215, Los Altos, California 94022,
[3]Electrical and Computer Engineering, University of California, Davis, Davis, California 95618

Email: *tyamada@soe.ucsc.edu

Abstract:

Theory is proposed for nanohole silicon *pin*/*nip* photodetector (PD) physics, promising devices in the future data communications and lidar applications. Photons and carriers have wavelengths of 1 μm and 5 nm, respectively. We propose vertical nanoholes having 2D periodicity with a feature size of 1 μm will produce photons slower than those in bulk silicon, but carriers are unchanged. Close comparison to experiments validates this view. First, we study steady state nanohole PD current as a function of illumination power, and results are attributed to the voltage drop partitions in the PD and electrodes. Nanohole PD voltage drop *depends* on illumination, but series resistance voltage drop does not, and this explains experiments well. Next, we study transient characteristics for the sudden termination of light illumination. Nanohole PDs are much faster than flat PDs, and this is because the former produces much less slow diffusion minority carriers. In fact, most photons have already been absorbed in the *i*-layer in nanohole PDs, resulting in much less diffusion minority carriers at the bottom highly doped layer. Why diffusion in PDs is slow and that in bipolar junction transistors is quick is discussed in Appendix.

Accepted Manuscript

Introduction

Photodetectors (PDs) are devices to convert an optical signal to an electric signal and are key devices in the future data communications and lidar applications [1,2]. Good PDs must have large external quantum efficiency (EQE) and electron-hole pairs are efficiently generated. EQE is defined by $EQE = \frac{J/q}{P/h\nu}$, where J [A/m^2] is the device current, q [C] is the unit charge, P [W/m^2] is the photon power density, and $h\nu$ [J] is the photon energy, and this definition is widely used. The numerator J/q is equivalent to the product of carrier velocity and density.

We have proposed and reported that the EQE is substantially enhanced experimentally by creating vertical 2D periodic array of nanoholes for the standard *pin* or *nip* PDs [3-11]. The 2D nanoholes have a period of ~1 μm, which is designed to be comparable to the illumination photon wavelength. The role of these nanoholes is to provide a 2D periodic potential in the horizontal directions and make the photon propagation velocity slower than that of control flat devices. Since nanoholes form an array of periodic potentials, they do not cause any damping in horizontal propagation, but instead, modify how waves propagate in the device. Nanoholes are vacancies of silicon atoms in the bulk silicon background. Thus they tend to be potential barriers for both photon and electron waves. From quantum mechanics point of view, 2D periodic nanoholes will create 2D periodic lattice, and waves in there will form Bloch waves. The dispersion relation of these Bloch waves is more *flattened* since nanoholes add more potential barrier. Since the photon wavelength and the nanohole potential period are both ~1 μm, effects are maximized. This is the mechanism for "slow photons." However, electrons have a typical wavelength of ~5 nm in bulk silicon, which is much shorter than the nanohole feature size, these nanoholes do not change the properties of electrons. In fact, quantum interference is important only when the wavelength and barrier period are comparable. Therefore, in nanohole PDs, we have slow photons and the same electrons in bulk silicon. Slow photons can stay longer in silicon materials while carrier properties are unchanged, and this will be the basis of enhanced photon absorption in nanohole PDs.

Similar ideas to add 2D structures and enhance photon absorption are independently proposed and experimentally investigated for various photoelectric devices, where photon absorption plays a critical role. Berkeley studied solar cells with short-circuited *pn* junctions [12]. Sony studied CMOS image sensors with floating photon absorbing silicon layers [13]. Stanford studied silicon single-photon avalanche diodes with reverse biased *nip* junctions [14].

To characterize how electron-hole pair generation is enhanced in these devices, it is important to pay attention to the figure of merit (FOM). Comparing devices with and without 2D nanostructures, the FOM is about 10 percent larger in Berkeley solar cells, several tens percent in Sony image sensors, and several hundred percent or more in our nanohole PDs and Stanford avalanche photodiodes. The 2D nanostructures in these devices surely enhance photon absorption by introducing slow photons.

Solar cells are *pn* junctions and convert photon energy to electric energy, and are operated in the forward bias condition. In fact, open-circuit voltage and short-circuit current often characterize device performance and solar cells are practically operated somewhere between these two extremes in the *forward* bias condition. Then generated electron-hole pairs will be subject to immediate recombination, resulting in the smaller FOM about 10 percent. PDs are *pin/nip* junctions and are operated in the *reverse* bias condition. Generated electron-hole pairs are immediately spatially separated, and recombination is minimized. This makes the PDs' FOM much larger than that of solar cells. Nanohole PDs and nanohole solar cells are essentially

Accepted Manuscript

different in that the former is operated under reverse bias condition preventing electron-hole pair recombination, while the latter is operated in the forward bias condition allowing electron-hole pair recombination. In PDs, recombination is prevented, while in solar cells, recombination is allowed. This will reduce FOM in solar cells compared to that in PDs. Thus, these devices are essentially different and FOM numbers are not transferable between PDs and solar cells.

Image sensors generate electron-hole pairs from photons, separate electrons and holes spatially, and use the designated polarity of charges to CMOS gate voltage input. For this purpose, the built-in electric field in a *pin*/*nip* junction is used in the floating condition. For example, the *n*-side is connected to a DC battery and floating *p*-side voltage, changing its value on the order of semiconductor energy band gap with and without illumination, drives the CMOS at the gate voltage. Because of the built-in electric field in the *pin*/*nip* junction in the floating condition, generated electron-hole pairs are spatially separated and recombination is reduced. The FOM is about several tens percent.

PDs generate electric current in response to illumination. Generated electron-hole pairs need to be significantly accelerated to achieve a large on-off ratio in the device output current. For this purpose, *pin*/*nip* junctions are strongly reverse biased, so that the *i*-layer electric field is maximized. The reverse bias condition is necessary for achieving large drift current but is also critical in spatially separating generated electron-hole pairs. This immediate separation of generated electrons and holes following photon absorption eventually eliminates unwanted recombination and the FOM is maximized about several hundred percent.

Avalanche photodiode also use *pin*/*nip* junctions in the reverse bias condition. The strong electric field in the *i*-layer in the reverse bias condition is intended to induce avalanche breakdown originally, but it also separates generated electron-hole pairs as a biproduct. Because of the additional magnification mechanism, FOM for avalanche diodes need clarification. FOM is defined as (nanohole performance)/(flat performance), and this estimation is done just before the onset of avalanche breakdown. In fact, once avalanche breakdown occurs, the device loses memory for how many original carriers the device had before breakdown. Thus, FOM numbers for PDs and avalanche photodiodes are identical, several hundred percent in both.

Our previous studies [3-11] have experimentally shown that the EQE is in fact substantially enhanced with 2D periodic nanoholes in such a way that the FOM increases by several hundred percent. This is because photons have a slower horizontal velocity in 2D periodic nanoholes, can stay longer in the silicon material, and have more chances for being absorbed in the *i*-layer of the devices. Accordingly, 2D periodic nanohole devices have substantially high photon absorption compared to flat devices. Absorbed photons generate electron-hole pairs. If devices are short-circuited, or they are left in the floating condition, these generated electron–hole pairs will be subject to recombination. Thus, the final EQE is not maximized. However, if devices have a reverse bias condition, generated electron-hole pairs are quickly separated and recombination is substantially suppressed. This explains why the FOM is several hundred percent in reverse bias devices such as our nanohole PDs or avalanche photodiodes with 2D nanostructures. The slow photon effect is also called as photon trapping.

This is our present understanding, and we will examine if our experimental results are consistently understood. For this purpose, two experimental results are tested: steady state output as a function of constant input illumination, and transient output for sudden termination of input illumination.

Nanohole Photodetector Device Structure

Accepted Manuscript

One of the popular PDs are *pin* photodetectors, where semiconducting *p*-, *i*- and *n*-layers are connected in series and are reverse biased [3-4]. In dark, there is practically no current since the device is reverse biased. Under illumination, electron-hole pairs are generated in all *pin* layers, and they are accelerated in the *i*-layer, resulting in large output current. Important figures of merit are (a) the ratio of the steady state device current under illumination and in dark (on-off ratio), and (b) the device transient response from illumination to dark (turn off time).

Our PDs in Fig. 1(a) have a backbone structure made of top *p*-layer, a middle *i*-layer, and a bottom *n*-layer, all made of silicon. In order to enhance photon absorption, 2D periodic nanoholes are placed perpendicularly to the *pin* structure [3-4]. The role of nanoholes is to enhance the photon absorption in the *i*-layer silicon, using vertical nanoholes periodically placed in the 2D *p*- and *i*-layers. Illumination is applied from the top in a practical device layout so that the photon input is vertical (red solid arrows) with respect to the *pin* layer. Due to the presence of nanoholes, there will be a horizontal photon flow (read broken arrows) and this is schematically shown in Fig. 1(a). Similar efforts are made to let the light traveling to the horizontal direction by introducing a 2 D structure at the device surface. The influence of the 2D periodic nanoholes is to significantly enhance EQEs, while leaving photogenerated electron-hole pair physical properties to be unchanged.

Steady State Characteristics with Constant Illumination

Towards the estimation of the on-off ratio, steady state output is studied as a function of illumination strength with device voltage fixed. Since more electron-hole pairs will be generated, device output should increase with illumination. Device voltage provides acceleration for carriers and thus, output is expected to increase with voltage. However, experiments reveal more complication. In this article, a model is developed to explain experimentally observed data, based on the carrier transport physics in the silicon material, under the influence of light illumination and device voltage.

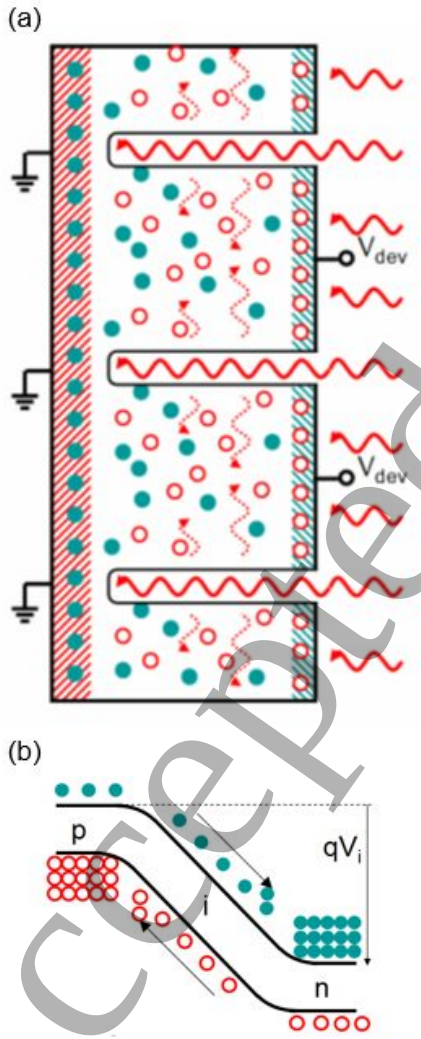
In this work, nanoholes are implicitly considered as an enhancement in electron-hole pair generation, which is supported by quantum efficiency measurements in Ref. 3. Electrons and holes are assumed to be identical in the bulk silicon material. In fact, nanoholes have a feature size of micrometers in our PDs in Fig. 1(a) bottom. Photons also have a wavelength of micrometers and there should be appreciable modulations in photon properties, while electrons and holes have a wavelength of ~5 nm and should not be changed.

The Fig. 2 (a) inset shows the photograph of a fabricated device. Nanoholes are placed on the 2D device planes [3-11]. Input photons are applied perpendicularly to the *pin* silicon place. They travel along the nanoholes and are reflected at the bottom of the nanoholes. There will be an interference of incoming and reflected photons. As a result, photons start moving to the horizontal directions. These horizontal photons feel the presence of 2D periodic potential created by nanoholes and there is a group velocity modulation, i.e., the group velocity decreases compared to that in bulk silicon (consistent with the relativity). Since these photons have slower group velocity, they can stay longer in the silicon material, resulting in higher electron-hole pair generation. This is the physics of PD photon absorption enhancement. The input-output characteristics in Fig. 1 are measured for PDs with nanoholes. Of course, similar behavior is observed for control devices simple *pin* devices without nanoholes. Photons inside the bulk

Accepted Manuscript

silicon will have a designated dispersion (energy E vs. wavenumber k) relation determined by potentials created by silicon atoms, and a designated group velocity $v_{gr} = \frac{1}{\hbar}\frac{dE}{dk}$, where $\hbar$ is the reduced Planck constant, E is the energy, and k is the wavenumber. When 2D nanoholes are added to bilk silicon, a photon wave will see additional 2D periodic potential. This will give a change in the functional form for the bulk silicon dispersion, and this modification is in the direction of slower photon group velocity.

Figure 2 (a) shows the steady state PD output current as a function of input 850 nm illumination power with the device voltage fixed for 250 μm diameter PDs. This is done in a steady state condition and the measurements are performed after all transients have faded away. When illumination is less than 4 mW with $|V_{dev}|$ = 2 V, etc., output current depends linearly on input illumination. There is very little device voltage dependence. We call it the device linear mode. When illumination is larger than 5 mW with $|V_{dev}|$ = 2 V, however, output current tends to saturate with a clear dependence on device voltage. We call it the device saturation mode. The transition from the device linear to saturation modes occur at a designated illumination, and it depends on the device voltage, such that 4 mW with $|V_{dev}|$ = 2 V, 5.5 mW with $|V_{dev}|$ = 3 V, etc.



Accepted Manuscript

Fig. 1. (a) Studied *pin* photodetector with nanoholes (~1 μm wide) in constant device voltage V_{dev} as a parameter in terms of particle view, where blue circles are electrons and red circles are holes with red wavy photon input. (b) Energy band without contact electrodes (series resistance).

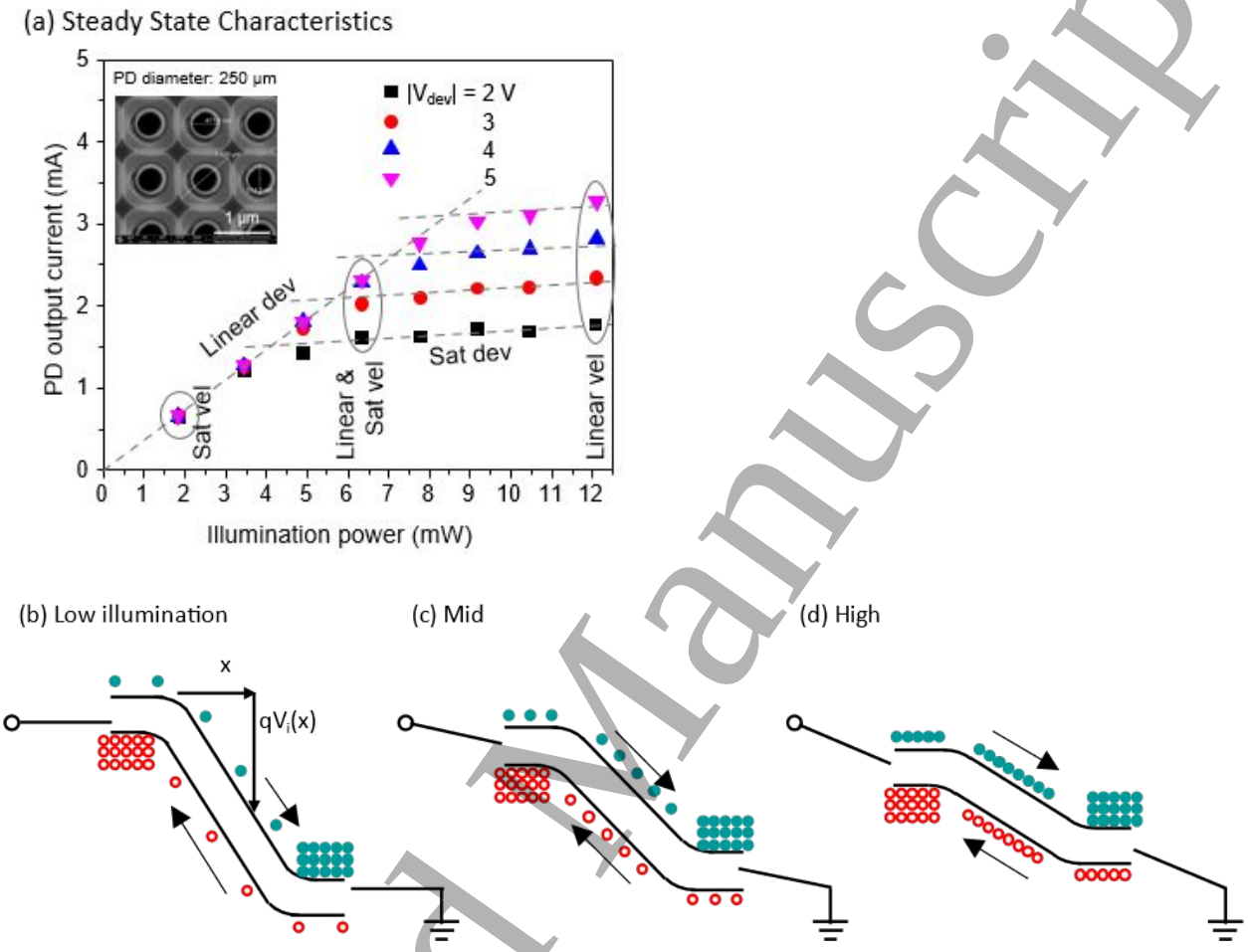


Fig. 2. (a) Experimental steady state results for the PD with a 250 μm diameter output current as a function of input 850 nm illumination power with the device voltage as a parameter. In the figure, "vel" stands for velocity. There are two different modes: (1) Device linear mode. When illumination is weak less than 4 mW, the PD output current depends linearly on the illumination power. The current is independent of the device voltage. This mode is related to transport saturation mode. (2) Device saturation mode. When the illumination power is strong, larger than 5 mW, the output current tends to saturate. The current shows a linear dependence on the device voltage. This is related to transport linear mode. (b), (c), and (d) are energy bands for low, middle, and high illumination. As the illumination increases, the device current increases and there will be larger voltage drop across the series resistance.

The existence of the device linear and saturation modes can be understood as follows. The *pin* photodetector is contacted to conducting wires at both *p*- and *n*-layers. Because of the high doping in *p*- and *n*-layers, the voltage drop mainly occurs in the *i*-layer while the voltage drops in the metallic wire and *p*- and *n*-layers are negligible in dark. Illumination will increase minority carriers in *p*- and *n*-layers, while leaving majority carriers unchanged. In dark, the p^+ and n^+ layers are 10^{18} cm^{-3} and minority carriers are 10^2 cm^{-3} because of the mass action law for bulk silicon at 300 K. Let us assume the illumination generates 10^{16} cm^{-3} e-h pairs. Under illumination,

Accepted Manuscript

majority carriers are $10^{18} + 10^{16} \sim 10^{18}$, and there is no change from the dark. The minority carriers are 10^2 cm^{-3} in dark, but are 10^{16} cm^{-3} under illumination. Thus, only the minority carriers undergo significant changes. Thus, only the minority electron concentration in the *p*-layer and the minority hole concentration in the *n*-layer are increased under illumination. Both electron and hole concentrations in the *i*-layer are increased.

When the illumination is low, most device voltage drop occurs in the *i*-layer as shown in Fig. 2(b). As a result, a strong *E*-field appears in the *i*-layer, and electrons and holes will run with the device saturation velocity as shown in Fig. 2(b). The number of generated electron-hole pairs is linearly dependent on the illumination power as experimentally established for our photodetectors [3-4], and this linearity is widely seen in different semiconductor devices from photodetectors [15]. The device linear mode continues up to the mid illumination in Fig. 2(c). Because of the increased illumination, the voltage drop in the *i*-layer is smaller than that in Fig. 2(d) and the voltage drops in the wires can be visible since the wire resistances are independent of the illumination while the *pin* photodetector resistance is certainly smaller. After illumination is further increased, there are many electron-hole pairs and the *i*-layer resistance will decrease significantly, and it is comparable to the device series resistance. The *i*-layer voltage is only a fraction of the entire device voltage. The *i*-layer *E*-field decreases with increasing illumination, and electrons and holes run in the transport linear mode, i.e., sub saturation velocity. In addition to that, the *i*-layer mobility decreases with increasing the electron-hole pair concentration causing more scattering events. This effective decrease in the *i*-layer *E*-field and decrease in the carrier mobility in the *i*-layer are the mechanisms for the device saturation mode. Thus, the device current behavior in Fig. 1(b) is caused by two effects, that is as illumination is increased, (1) the *E*-field in the *i*-layer decreases (less acceleration), and (2) the mobility in the *i*-layer decreases due to more electron-hole pairs (worse transport). This physical understanding will be backed up with mathematics based on the differential form [16] in the following.

According to the experiment in Fig. 2(a), the output current is independent of the device voltage in the device linear mode, while the current is dependent on the device voltage in the device saturation mode. The change in the device linear current ΔI_{lin} and the device saturation current ΔI_{sat} are expressed by the change in illumination power ΔP and the device voltage ΔV. These relations are empirically expressed with empirical equations by,

$$\Delta I_{lin} = \frac{\partial I_{lin}}{\partial P}\Delta P + \frac{\partial I_{lin}}{\partial V}\Delta V = 3.3\left(\frac{mA}{mW}\right)\Delta P(mW) + 0\Delta V(V) \quad , \tag{1}$$

$$\Delta I_{sat} = \frac{\partial I_{sat}}{\partial P}\Delta P + \frac{\partial I_{sat}}{\partial V}\Delta V = 0.30\left(\frac{mA}{mW}\right)\Delta P(mW) + 1.33\left(\frac{mA}{V}\right)\Delta V(V) \quad . \tag{2}$$

In the device linear mode, ΔI_{lin} changes linearly with ΔP, but has no dependence on ΔV in the device linear mode. According to Fig. 1(b), the device linear current changes by a differential of $\partial I_{lin}/\partial P =$ 3.3 mA/mW. However, there is practically no device voltage dependence. In the device saturation node, however, ΔI_{sat} changes by $\partial I_{sat}/\partial P =$ 0.30 mA/mW << $\partial I_{lin}/\partial P =$ 3.3

Accepted Manuscript

mA/mW, and ΔI_{sat} increases with device voltage by $\partial I_{sat} / \partial V$ = 1.33 mA/V. It is the purpose of this manuscript to explain how we can understand these empirical relations in Fig. 1(b).

Once the energy bands are identified for given illumination power P and i-layer voltage V_i, which is basically equivalent to the device voltage in dark but only a partial fraction of it with a large illumination, it is straightforward to obtain the output device current. All elements p-, i-, and n-layers are connected in series, and we study the steady state current. All current is conduction current [17-19] related to physical movement of charged particles, and there should be no displacement current [17-20] related to time dependence of the electric field. In developing an equivalent circuit for our PDs, it must be created by resistors only, and there should be no capacitors, *as long as steady state properties are concerned*, which is the entire purpose of this manuscript. The current is given by the product of carrier density and drift velocity consistent with the i-layer electric field F. Carrier transport information is included as drift velocity in the i-layer. The carrier density is uniquely determined by illumination P only, and is independent of the i-layer.

In order to understand the experimental results in Eqs. (1) and (2), it is necessary to express the output device current in terms of device voltage V_{dev} in stead of internal i-layer voltage V_i. Experimentally, it is not easy to measure the i-layer voltage V_i and it is necessary to create an appropriate equivalent circuit model since it is not certain how much voltage drops occur in the wires under the given illumination. Figure 3(a) shows one of possible equivalent circuit exclusively for the steady state analysis. Usually, capacitors or sometimes inductors are incorporated to estimate the device turn-off time, which is *transient* analysis. As emphasized si far, we restrict ourselves to understand *steady state* properties in experiments in Fig. 1. Therefore, capacitor- or inductor-elements will not contribute to the final results and they are eliminated from the beginning. It must be emphasized that the equivalent circuit in Fig. 3(a) is intended for the *steady state* analysis only.

When a *pin* PD is connected to the ground and the constant current source at V_{dev}, there are some resistors contributing to voltage drops. The p-, i-, and n-layers have finite resistors, and their resistance values depend on illumination since they will decrease when more electron-hole pairs are created. There are some resistors not linked to illumination. They are device contact resistances between the PD device and wires for the measurement circuit, etc. All of them are collectively identified as series resistance. Figure 3(a) schematically shows resistances at low illumination and high illumination. In low illumination, there are little electron-hole pairs in the i-layer, and the resistance in the i-layer is the most dominant. This means $V_i \sim V_{dev}$ when the illumination is low. In high illumination, however, appreciable electron-hole pairs are created in the i-layer. Minority electrons are created in the p-layer, and minority holes are created in the n-layer. With the same illumination, both electrons and holes will increase in the i-layer, while only the minority carriers will increase in the p- and n-layers. Thus, resistance decrease due to illumination in the i-layer is much more significant than those in p- and n-layers. In high illumination, the resistances in the p-layer, i-layer, n-layer, and series resistance will be on the same order. Thus, it is expected that V_i = 1/4 - 1/3 V_{dev}, or only 1/4 to 1/3 of device voltage V_{dev} can appear as i-layer voltage V_i. This is one physical effect due to high illumination.

Another physical effect due to high illumination is degradation in transport. Figure 3(b) plots the low field electron- and hole- mobility values as a function of carrier density in silicon, where the data is taken from Ref. [1]. It is clearly shown that both electron- and hole- mobility values by a factor of 1/3 as the carrier density increases from 10^{16} cm^{-3} to 10^{19} cm^{-3}. This is due to

Accepted Manuscript

increasing carrier-carrier scattering randomizing carrier momentum in transport [1,20,21]. At the high field, velocity saturation occurs at ~10 kV/cm around 10^{16} cm^{-}, but because of the low field mobility reduction, higher electric field is necessary before velocity saturation occurs. In Fig. 3(c), electron- and hole- drift velocity values are plotted as a function of electric field. By increasing the carrier density from 10^{16} cm^{-3} to 10^{18} cm^{-3}, drift velocity drops by a factor of ~1/3. Actually these numbers are consistent with our experimental device sizes [3-11]. The typical i-layer length is 1 μm and the typical device voltage is a few volts. This corresponds to a few tens of kV/cm and is a typical electric field for the onset of transport saturation in silicon [1].

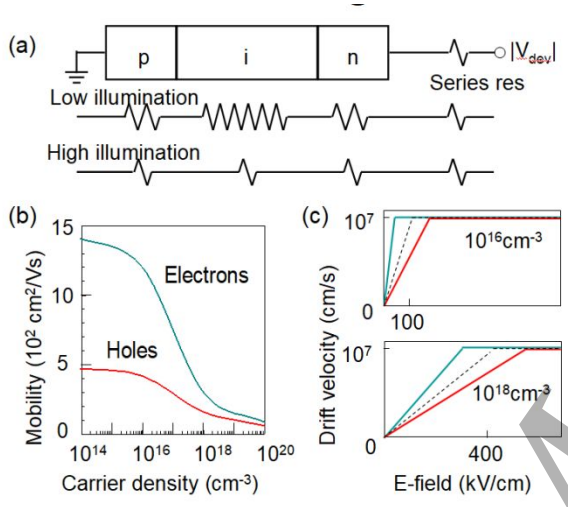


Fig. 3 (a) Equivalent circuit exclusively for steady state analysis (removing any capacitances or even inductances needed for transient studies). (b) Electron and hole mobility as a function of carrier density. (c) Electron and hole drift velocity as a function of electric (E) field as carrier density as a parameter. The broken lines indicate an average which can be used in device modeling.

At the very beginning, it is worth doing order estimation for the experimental data in Fig. 1. Comparing the device linear mode and saturation mode, the response to illumination is an order of magnitude different, which are 3.3 mA/mW vs 0.3 mA/mW. The factor 0.3/33 ~ 1/11 represents current change due to low vs high illumination. The first factor comes from V_i/V_{dev}, which was estimated to be 1/4 to 1/3 in Fig. 3(a). The second factor comes from transport degradation of 1/3 in high illumination in Figs. 3(b) and 3(c).

Next, we will develop mathematics for this physical understanding. Figure 4 compares energy bands in low, mid, and high illumination fixing device voltage V_{dev}. Although V_{dev} is fixed, V_i keeps decreasing as illumination is increased. Due to the decrease in V_i, drift velocity changes from the transport saturation mode in Fig. 4(a) through transition in Fig. 4(b), and end up with the transport linear mode in Fig. 4(c). In Fig. 4(a), illumination is weak and the *i*-layer electric field F is so strong that transport saturation mode results. Since both carriers run with the saturation velocity, there is no V_{dec} dependence in output current. This is the device linear mode with transport saturation. In Fig. 4(c), illumination is strong and V_i is weak. This is the device saturation mode with transport linear.

Accepted Manuscript

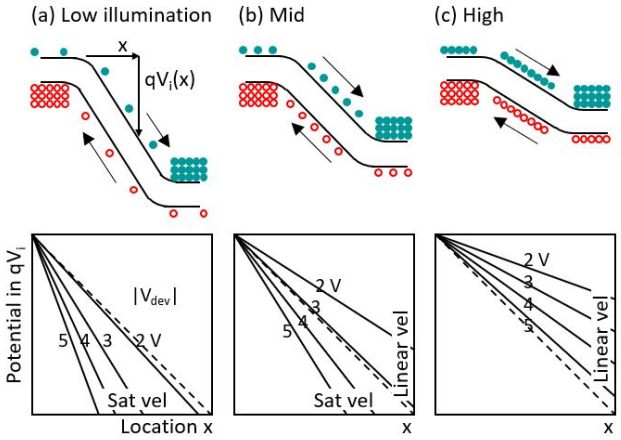


Fig. 4. Energy bands and potentials inside the *i*-layer (excluding metallic wires) with device voltage as a parameter. Illumination is increased from (a) low, (b) mid, to (c) high. In the figure, "vel" stands for velocity. The device linear mode is linked to transport saturation in small illumination in (a), and the device saturation mode is linked to transport linear in large illumination in (c). (b) is an intermediate illumination and both transport saturation and linear modes coexist.

The device linear mode in Fig. 4(a) has a saturation drift velocity of 1×10^7 cm/s for both electrons and holes as in Fig. 3(c). By using the transport relation, the device linear current is given using the unit charge q, the carrier number $\Delta N = \Delta n \times A$ with carrier density n and cross-section area A, and the saturation drift velocity v_{sat}, by

$$\Delta I_{lin}(P) = q\Delta N(P) \times v_{sat} = q\,\Delta n(P) A \times v_{sat} = q\frac{\partial n}{\partial P} A\Delta P \times v_{sat} \quad . \tag{3}$$

Now, $\partial n / \partial P$ is a quantity indicating the density of electrons and holes in cm^{-3} generated per 1 mW, and is experimentally determined by 1×10^{19} cm^{-3} from Fig. 1 for a PD with a round PD with a diameter of 25 μm. Our p and n doping is in the neighborhood of 5×10^{19} cm^{-3}. This is consistent with the threshold illumination power of 4.5 mW in Fig. 1.

$$\frac{\partial n}{\partial P} = 3.3\left(\frac{mA}{mW}\right)\frac{1/A}{qv_{sat}} = 1.0 \times 10^{19}\frac{1}{cm^3 mW} \quad . \tag{4}$$

The device saturation mode is complicated in analysis and extensive discussion is necessary. In the device saturation mode, the output current depends on illumination P and device voltage V_{dev}. Illumination increases the electron-hole pairs and the carrier density N in the *i*-layer is increased directly with illumination power P. The linear velocity v_{dr} depends on two elements. They are (i) the *i*-layer carrier density through illumination power P since mobility depends on

Accepted Manuscript

carrier density as shown in Figs. 3(b) and 3(c), and (ii) the *i*-layer electric field F through the illumination power P and the device voltage V_{dev}.

$$\begin{aligned}\Delta I_{sat}(P,V_{dev}) &= q\Delta\left[N(P)v_{dr}\ (P,V_{dev})\right] = q\Delta\left[N(P)\mu(P)F(P,V_{dev})\right] \\ &= q\frac{\partial N}{\partial P}\Delta P\times\mu F + qN\frac{\partial\mu}{\partial n}\frac{\partial n}{\partial P}\Delta P\times F + qN\mu\frac{\partial F}{\partial P}\Delta P + qN\mu\frac{\partial F}{\partial V_i}\frac{\partial V_i}{\partial V_{dev}}\Delta V_{dev} \\ &= qA\frac{\partial n}{\partial P}\Delta P\times\mu F + qAn\frac{\partial\mu}{\partial n}\frac{\partial n}{\partial P}\Delta P\times F + qAn\mu\frac{\partial F}{\partial P}\Delta P + qAn\mu\frac{\partial F}{\partial V_i}\frac{\partial V_i}{\partial V_{dev}}\Delta V_{dev}.\end{aligned} \quad (5)$$

Therefore, mathematically there are exactly three terms showing the illumination power P dependence and a term showing the device voltage V_{dev} dependence, and it is clearly shown that there are no double counting or no forgotten contribution.

Next we will examine how the PD output current depends on device illumination power and device voltage in the device linear mode and the device saturation mode. In Fig. 4, the device voltage V_{dev} dependence has already been discussed. In the device linear mode in Fig. 4(a), the PD output current does not depend on device voltage V_{dev} since the *i*-layer electric field F is quite strong even with the smallest V_{dev} of -2 V, and carriers reach the saturation drift velocity, even though the field F certainly increases with -2, -3, -4, to -5 V. In the device saturation mode in Fig. 4(c), however, the PD output current depends linearly on device voltage V_{dev} since the *i*-layer electric field is quite weak and the drift velocity is proportional to the *i*-layer electric field, which depends on the device voltage V_{dev}. This is expressed in the fourth term in Eq. (5).

It would be counterintuitive that our PDs show an entangled behavior for the device and transport modes: the device linear mode is linked to transport saturation mode, and the device saturation mode is linked to transport linear mode. In order to clarify this relation, Fig. 4 is prepared. The top figure is in principle a reproduction of Fig. 2(a), where the output device current is shown as a function of illumination strength. Four operating points are shown in the order of illumination strength, light blue, dark blue, red, and yellow for fixed V_{dec} = -2 V. Device output current increases with illumination. Light blue is in the device saturation mode, and bark blue, red, and yellow are in the device saturation mode. In order to show the entangled nature of the device operation, the carrier drift velocity (averaged over electrons and holes) is created for these four operating points as a function of *i*-layer electric field. Thus, depending on the device linear or saturation modes, dv_{dr}/dF behaves as follows.

$$v_{dr}(dev\,lin) = v_{sat}, \quad (6a)$$

$$v_{dr}(dev\,sat) = \mu(n)F. \quad (6b)$$

$$\frac{dv_{dr}}{dF}(dev\,lin) = 0, \quad (7a)$$

$$\frac{dv_{dr}}{dF}(dev\,sat) = \mu(n). \quad (7b)$$

Accepted Manuscript

Light blue is in the device saturation mode, and in the transport saturation mode. Dark blue, red, and yellow are in the device saturation mode. In the device linear mode, dv_{dr}/dF is zero. In the device saturation mode, dv_{dr}/dF is the mobility and is a decreasing function of *i*-layer carrier density as discussed in Fig. 3(b) since carrier-carrier scattering increases with illumination. Accordingly, the Fig. 5 bottom figure is concluded. If there were no mobility degradation with increased illumination, read and yellow would align with the line connecting the origin and dark blue. The experimentally observed gradient in Fig. 2(a) in the device saturation mode is strongly flat, and this mobility degradation and the decreased *i*-layer voltage drop are two independent mechanisms.

Based on the Fig. 5 bottom figure, the change in the normalized drift velocity change $\Delta v_{dr}/v_{dr}$ in the *device saturation and transport linear* modes (dark blue, red, and yellow) can be expressed as a function of *i*-layer carrier density *n* and device voltage V_{dev} by,

$$\frac{\Delta v_{dr}}{v_{dr}} = -\left|\frac{n_0}{\mu_0}\frac{\partial\mu}{\partial n}\right|\frac{\Delta n}{n_0} + \frac{R_{i0}}{R_{i0}+R'_s}\left(-\frac{\Delta n}{n_0} + \frac{\Delta|V_{dev}|}{|V_{dev0}|}\right). \quad (8)$$

Here, the subscript 0 indicates a device operation reference position at $|V_{dev}|$ = 2 V and the change in V_{dev} is measured from there. It has to be noted that V_i is a fraction of V_{dev} reflecting a ratio of *i*-layer resistance and series resistance. It is clearly shown that increasing the carrier density will decrease the drift velocity, while increasing the device voltage (in the device saturation and transport linear mode) will increase the drift velocity.

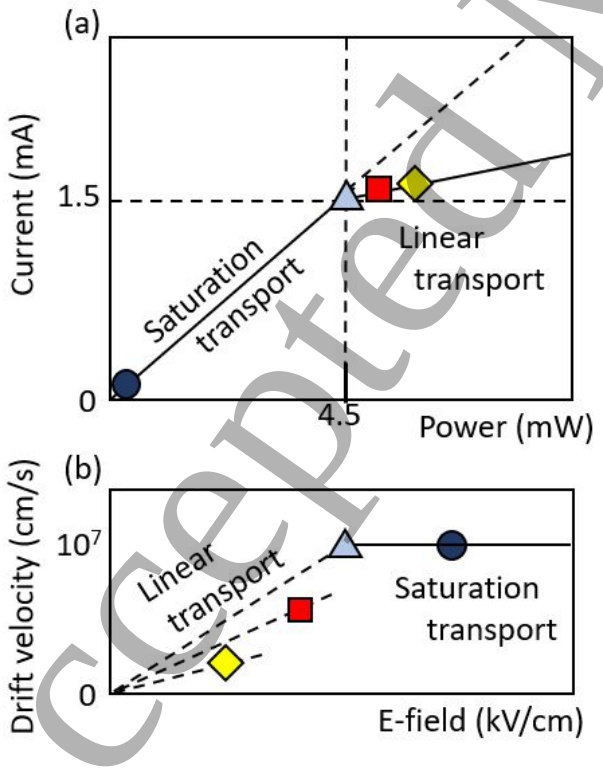


Fig. 5. (a) PD output current with V_{dev} = –2 V as a function of illumination power *P*. In the plot, vel is velocity and dev is device. The device linear mode is realized with *P* < 4.5 mW and the device saturation

Accepted Manuscript

mode is realized with $P > 4.5$ mW. (b) Carrier drift velocity with $V_{dev} = -2$ V as a function of *i*-layer electric field *F*. The operating points in (a) as a function of illumination power *P* appear in the reverse order in (b) as a function of the *i*-layer electric field *F*.

By using Eq. (8), it is possible to drive how the onset of device saturation mode shifts in Fig. 2(a) when the device voltage is changed. By setting the left hand of Eq. (8) to be zero, we have

$$\frac{\Delta n}{n_0} = \frac{1}{1+\alpha}\frac{\Delta|V_{dev}|}{|V_{dec0}|} \text{ or } \Delta n = \frac{1}{1+\alpha}\frac{n_0}{|V_{dev0}|}\Delta|V_{dev}|,$$
$$\text{where, } \alpha \equiv \left|\frac{n_0}{\mu_0}\frac{\partial\mu}{\partial n}\right| \times \frac{R_{i0}+R'_s}{R_{i0}} = \left|\frac{\partial \ln\mu}{\partial \ln n}\right| \times \frac{R_{i0}+R'_s}{R_{i0}} = \left|\frac{\partial \log\mu}{\partial \log n}\right| \times \frac{R_{i0}+R'_s}{R_{i0}}, \quad (9)$$

From Fig. 2(a), we have an experimental relation of $\Delta n(mW) = 1.3\Delta V_{dec}(V)$ for a relatively wide range of device voltage, V_{dev} = -2 to -5 V. Thus, $\frac{1}{1+\alpha}\frac{n_0}{|V_{dev0}|} = 1.3(mW/V)$. Since n_0 = 4 mW and $|V_{dev0}|$ = 2 V, α = 0.54. This experimental α value is in the reasonable range determined by carrier transport for your photodetectors. In fact, according to Fig. 3(b), mobility log μ changes by ~0.5 and log *n* changes by ~1, and the resistance ratio is also close to 1 due to the device design. Thus, α defined in Eq. (9) takes a value ~0.5 for our photodetectors, and it is consistent with our experimental value.

Transient Characteristics with Sudden Termination of Illumination

Next, transient measurements are performed for nanohole devices and flat devices with the device voltage as a parameter. Figure 6(a) and 6(b) show characteristics of nanohole devices and flat devices both with 30 μm diameters, respectively, as a device voltage as a parameter. The continuous 850 nm illumination at 0.5 s is suddenly turned off and the device output (converted to voltage) is shown as a function of time for various applied device voltages. It is immediately noticed that nanohole devices in Fig. 6(a) turn off with much smaller tails than flat devices in Fig. 6(b). When the device voltage is high, there may be an influence of avalanche breakdown, but in this manuscript, we assume they it is simply absent. Systematic investigations between avalanche breakdown and long time tales are left for our future study.

This behavior can be consistently understood by (1) nanohole devices have a significantly enhanced quantum efficiency and (2) they have much less photogenerated minority carriers (electrons) in the bottom *p*-layer, which must obey the slow diffusion process and move from the *p*-layer, through the *i*-layer, and to the *n*-layer.

Let us compare nanohole and flat *nip* devices in Fig. 6(c). In the theory developed below, there is no difference between *pin* and *nip*, and the device schematic figure is prepared for *nip*. The top *n*-layer L_n is very thin (L_n << 1μm), the middle *i*-layer L_i is thick (L_i ~ 1 μm), and the bottom *p*-layer L_p is thin (L_p < 1 μm). The coordinates are defined so that the right end is $x = 0$, and the left end is $x = -L_n - L_i - L_p \equiv -L_{tot} \sim -L_i$. The nanoholes have a depth L_{hn}, about a half of L_i as shown in left figures. The top right shows the photon density as a function of *x* for

Accepted Manuscript

the nanohole (solid) and flat (broken) devices. In the flat device, the absorption occurs slowly, uniformly throughout the device and the exponential decay constant α_0 is small. Thus, the photon density probability function $P_0(x)$ in the flat device is simple exponential with a broken line. In the nanohole device, the exponential decay α_{nh} is fast in the nanohole region and $\alpha_{nh} \gg \alpha_0$. Thus, the photon density $P_{nh}(x)$ in the nanohole device has to gradients. Absorption is obtained by multiplying the absorption probability β_0 and β_{nh} in flat and nanohole, respectively. Between α and β, there is a relation based on the photon velocity c such that $\beta = c\alpha$.

Now a normalized flat-device illumination power density is $P_0(x)$ and that for nanohole device is $P_{nh}(x)$. They are cumulative distribution function. Normalized flat-device absorption probability is $A_0(x)$ and that for nanohole device is $A_{nh}(x)$, with β_0 and β_{nh} are absorption probabilities (between 0 and 1) for flat and nanohole devices, respectively. They satisfy $\beta_0 \sim 0$ and $\beta_{nh} \sim 1$.

$$P_0(x) = \exp(-\alpha_0|x|), \quad A_0(x) = \beta_0 P_0(x) \ . \tag{10a}$$

$$P_{nh}(x) = \exp(-\alpha_{nh}|x|), \quad A_{nh}(x) = \beta_{nh} P_{nh}(x), \ -L_{nh} < x < 0 \ . \tag{10b}$$

$$P_{nh}(x) = \exp(-\alpha_0|x + L_{nh}|)\exp(-a_{nh}L_{nh}), \quad A_{nh}(x) = \beta_0 P_{nh}(x), \ x < -L_{nh} \ . \tag{10c}$$

In flat and nanohole devices, an integral relation holds,

$$P(x) = 1 - \int_{-L_{tot}}^{0} A(x)dx \ . \tag{11}$$

Since $A(x)$ is a functional of $P(x)$, Eq.(13) defines an integral equation and it can be solved numerically via recursion method. The long time tail is related to slow diffusion electrons in the bottom p-layer. Thus, the amount of absorption at the bottom p-layer in flat and nanohole devices are

$$\int_{-L_{tot}}^{-L_{tot}+L_p} dx A_0(x) \sim A_0(-L_i)L_p = \beta_0 P_0(-L_i)L_p = \beta_0 \exp(-\alpha_0 L_i)L_p, \tag{12a}$$

$$\int_{-L_{tot}}^{-L_{tot}+L_p} dx A_{nh}(x) \sim A_{nh}(-L_i)L_p = \beta_0 P_{nh}(-L_i)L_p = \beta_0 \exp(-\alpha_0|L_i - L_{nh}|)\exp(-\alpha_{nh}L_{nh})L_p \ . \tag{12b}$$

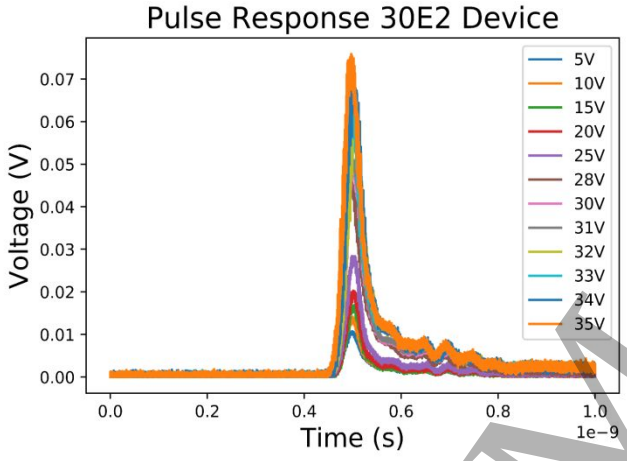
Thus, the low nanohole-device carrier density at the bottom p-layer with respect to the high flat device carrier density is

Accepted Manuscript

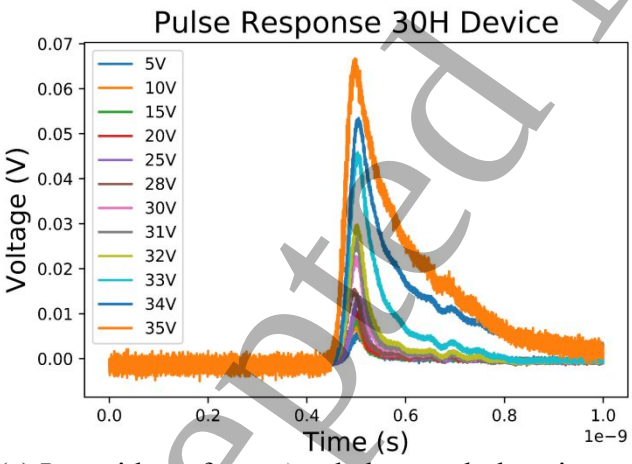
$$\frac{A_{nh}(-L_i)L_p}{A_0(-L_i)L_p} = \exp\left[-(\alpha_{nh} - \alpha_0)L_{nh}\right] \quad (13)$$

The final mathematical is simply a nanohole region depth times the absorption coefficient difference between nanohle and flat devices. The ratio is the long time tail amplitude ratio of nanohole devices with respect to that of flat devices, and this is smaller as the difference in the photon absorption in nanohole and flat devices are larger, as expected.

(a) Nanohole PD turn-off transient characteristics



(b) Flat PD turn-off transient characteristics



(c) Logarithm of generated electron-hole pairs

Accepted Manuscript

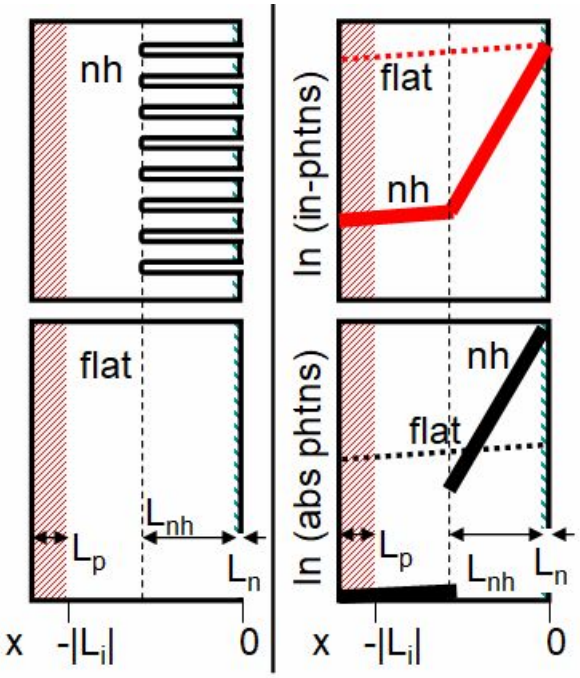


Fig. 6. Time response of device output for sudden termination of 850 nm illumination at 0.5 s for 30 μm diameter PDs: (a) nanohole devices and (b) flat devices. (c) Left top and bottom are nanohole (nh) and flat devices. They are created using nip structures, where the extremely thin top *n*-layer is $x = 0$ to $-L_n$, the next *i*-layer is $x = -L_n$ to $-L_i$, and the bottom *p*-layer is $x = -L_i$ to $-L_i - L_p$. The nanoholes are $x = 0$ to $-L_{nh}$. Right top is logarithm of normalized incoming photon density (cumulative distribution density) $P(x)$ and right bottom is absorption probability (probability distribution density) $A(x)$ as a function of device depth x from the surface. In the flat device, the incoming photon density decreases slowly (broken line) in the flat device, and it decreases quickly up to $-L_{nh}$ in the nanohole device and then slowly (thick solid line). The absorbed photons decrease slowly (broken) in the flat device, and it decreases quickly within the nanohole region followed by the slow absorption region. At $x = -L_{nh}$, the photon absorption drops significantly in the nanohole device.

Transient studies in Fig. 6 show superior transient characteristics in nanohole devices. Nanohole devices have significantly enhanced photon absorption characteristics over controlled flat devices. This is a benefit of nanohole devices, and the same advantage eliminates unwanted long time tales in transient behavior. Long time tails in nanohole and flat devices are, in the end, proportional to the following. Let us numerically estimate using our experimental data in Fig. 4(a) of Ref. 3. For the photon wavelength of 850 nm with the *i*-layer of $L = 1.3$ μm, we have $\alpha_F L = 0.07$ and $\alpha_N L = 0.60$.

$$\text{Nanohole PD long time tail amplitude: } \exp\{-\alpha_N(i)L\} = \exp(-0.60) = 0.54. \quad (14a)$$

$$\text{Flat PD long time tail amplitude: } \exp\{-\alpha_F(i)L\} = \exp(-0.07) = 0.93. \quad (14b)$$

Nanohole devices typically have a factor of ~0.54/0.93 = 0.58 smaller long time tail amplitude compared to flat devices. This is consistent with the experimental results in Figs. 6(a) and 6(b). In fact, transient characteristics are studied experimentally in Figs. 6(a) and 6(b) for various device voltages. The larger the device voltage is, the larger the output is, and consistently,

Accepted Manuscript

nanohole devices show smaller long time tales than flat devices over all device voltages, and the ratios do not have strong device voltage dependence. It is fair to conclude that the enhanced *i*-layer photon absorption in nanohole devices is responsible for the much better transient behavior and this is backed up with our experimental data in Fig. 6. In more detail, the experimental data could include an onset of avalanche breakdown, which is assumed not relevant in the theory here. In the future, we will study how the avalanche breakdown influence long time tails.

In this transient treatment, the skin effect [1,22,23] is not considered at all. When a high frequency signal transmits through a conductor, the signal tends to stay at the vacuum-material boundary, and its characteristic localization length is known as skin depth. The shorter the skin depth is, the more the signal localized at the vacuum-material boundary. The skin depth is inversely proportional to the square root of frequency and electron-hole pair density. When a device is suddenly turned off due to the abrupt termination of input light illumination, the device output will turn off. According to our experimental data, the turn off time is estimated to be about 30 ps, corresponding to a frequency of ~30 GHz. If the light illumination is weak so that the generated electron-hole pair is less than 10^{18} cm^{-3}, the skin depth is over 1 μm and is safely ignored. If we are interested in a very high illumination so that the generated electron-hole pair is over 10^{20} cm^{-3}, then the skin depth is about 0.1 μm and the effect must be considered in the transient analysis.

Summary and Discussions

The physics of *pin* or *nip* nanohole photodetectors in steady state and transient conditions is studied by comparing experimental data and theoretical behavior. This is in the direction of recent progress in nanoscale three-dimensional materials and devices [24-32]. In theory, series resistance, carrier transport with linear and saturation drift velocities, and slow diffusion minority carrier photogeneration in the *p*- or *n*-layers are considered. Experimental results are consistently explained by our model.

We have first studied *pin* photodiodes under steady state illumination with a device voltage as a parameter. It is shown that the device linear and saturation modes as a function of illumination strength are in fact related to transport saturation and linear modes, respectively. This entanglement in modes, i.e., the device linear mode is linked to transport saturation mode and the device saturation mode is linked to transport linear mode, is the key to understand our experimentally observed device behaviors. Experimental characteristics of our photodetectors are well explained with a model based on the carrier transport. The role of series resistance is increasingly important as input illumination increases.

Next, *nip* nanohole and flat devices are compared when illumination is suddenly turned off. Experimentally the former shows negligible long time tails, while the latter does. This difference is attributed to the minority carrier photogeneration difference in the bottom *p*-layer. Nanohle devices show significantly enhanced photon absorption, and this improves steady state characteristics naturally. The same effect can eliminate un-preferred long time tales in nanohole devices.

Although nanohole devices have enhanced photon absorption, electron and hole physical transport properties in nanohole devices are kept identical to those in flat devices. The insights can be used in nanohole PD design principle.

Accepted Manuscript

We have shown that nanoholes can enhance PD device properties significantly. In this article, dark current is not studied systematically. Regarding this issue, dark current due to recombination near nanoholes is related to surface recombination and is closely related to how the surface states are controlled in the device [33,34]. When appropriate surface treatments are done and passivation is achieved, it is possible to minimize surface recombination. Dark current is thermodynamically limited, *i.e.*, given the band gap and lattice temperature, how may electron-hole pairs are created is the same as flat devices and is independent of nanoholes. This means that nanohole PDs can enjoy maximized EQE and still minimize dark current, which is rather an engineering challenge.

Data Availability Statement: The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment: Financial support by W&WSens Devices Inc. is gratefully acknowledged.

Accepted Manuscript

Appendix I: Why diffusion in bipolar transistors (BJTs) is fast while that in *pin/nip* PDs is not.

As discussed above, the slow diffusion process in the *p*- or *n*-layers in the *pin* PDs will result in the long time tail in the sudden turning off of the illumination. However, BJTs are known to be one of the fastest devices although its Forward Active mode used the diffusion process in the base region. It is instructive to compare *pin/nip* PDs and BJTs and understand why diffusion causes long time tail in the former while it does not in the latter.

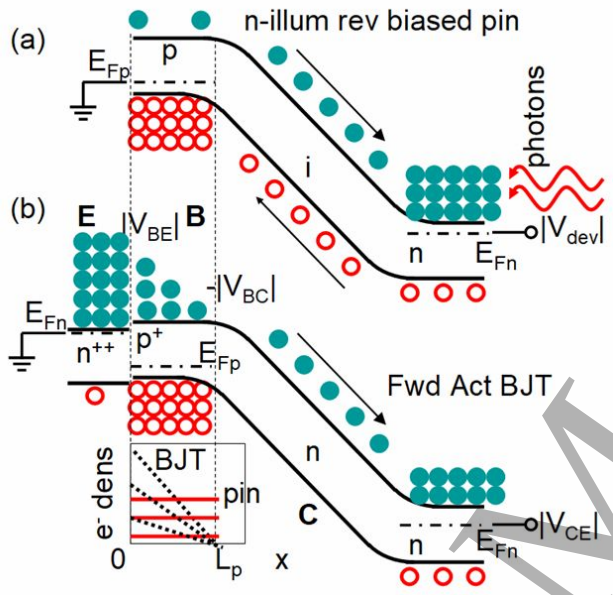


Fig. A1. Energy bands for *nip* PD and *npn* BJT are compared. The PD bottom *p*-layer and the BJT base, both with $x = 0$ to L_p, are considered equivalent, since diffusion is the dominant transport mechanism. (a) Energy band of *nip* PD illuminated from the right *n*-layer direction. Because of illumination, Fermi level generally splitting occurs, and only quasi-Fermi levels for majority carriers, E_{Fp} in the *p*-layer and E_{Fn} in the *n*-layer are shown. Reverse bias voltage $E_{Fp} - E_{Fn} = V_{dev}$, is applied at the *n*-layer with respect to the grounded *p*-layer. (b) Energy band of npn bipolar junction transistor (BJT) in the forward active mode, where again only quasi-Fermi levels for majority carriers are shown. The electron density is shown as a function of location *x*. To be consistent with BJT literature, ground is placed at the left end in Appendix.

First, the energy band of *nip* PD illuminated from the right *n*-layer direction is shown in Fig. A1(a). Because of illumination, Fermi level generally splitting occurs, and only quasi-Fermi levels for majority carriers, E_{Fp} in the *p*-layer and E_{Fn} in the *n*-layer are shown. Reverse bias voltage is applied at the *n*-layer with respect to the grounded *p*-layer. We have

$$-|E_{Fp} - E_{Fe}| = V_{dev} < 0 \ , \tag{A1}$$

$$n(0) \sim n_{illm} \ , \quad n(L_P) \sim n_{illm} \ . \tag{A2}$$

The minority carrier density in the bottom *p*-layer is expected to be flat in the ideal situation of no appreciable photon absorption. The minority carrier electron density in the *p*-layer is independent of location when the photon absorption is negligible in the *p*-layer. When the

Accepted Manuscript

illumination is suddenly turned off, electrons in the *p*-layer must depend on the *diffusion* to be back to the dark state, but this process is quite slow due to the lack of the electron density gradient under illumination.

Next, the energy band of *npn* bipolar junction transistor (BJT) in the forward active mode is shown in Fig. 1(b). Again only quasi-Fermi levels for majority carriers are plotted.

$$V_{BE} = |E_{Fp} - E_{Fe}| > 0 \quad . \tag{A3}$$

The difference of majority quasi-Fermi levels at the *p*-base and *n*-emitter, respectively, is $V_{EB} > 0$ indicating the forward bias.

$$V_{BC} = -|E_{Fn} - E_{Fp}| < 0 \quad . \tag{A4}$$

The difference of majority quasi-Fermi levels at the *n*-collector and *p*-base, respectively, is $V_{CB} < 0$ indicating the reverse bias. (A2) and (A3) show that the BJT is forward active (FA) mode. In the base, there should be a designated density gradient due to device voltages in the FA mode. In fact, the electron density at $x = 0$ (the emitter-base boundary) and at $x = L_P$ (the base-emitter boundary) is as follows with a minority carrier electron density at the base with zero bias n_{B0}.

$$n(0) = n_{B0}[\exp(qV_{EB}/k_BT) - 1] \ ,$$

$$n(L_P) = 0 \quad . \tag{A5}$$

There is no practical electron density gradient in the *nip* PD as in Eq. A2 is the mechanism for the unwanted long time tail. This is in sharp contrast to the designated electron density is guaranteed to be created in the BJT base as in Eq. A5. This makes BJT one of the fastest devices.

The photon absorption in the bottom *p*-layer in the nip PDs will create a long time tail, but the problem can be minimized in nanohole PDs. In nanohole PDs, the photon absorption is enhanced significantly due to the photon trapping or slow photons [3-11], and many photons are absorbed in the *i*-layer. Thus, much less photons can reach the bottom *p*-layer.

Appendix II: Experimental data related to silicon nanohole PDs in Figs. 2 and 6.

Experimental data related to Figs. 2 and 6 are tabulated in Table A1. For further interested readers in experimental details, our previous papers in Refs. 3 and 4 can be checked. In supplementary sections of Ref. s, sample preparation and measurements are discussed in depth.

1	Fig. 2 & Fig. 6	Device type	Silicon-vertical *pin* on an SOI substrate
2		Dark current	< 1 nA at -10 V
3		Series resistance	Nanohole *pin*: 83 Ω
4		Device thickness	Top 0.3 μm contact p^+-layer / middle 2 μm *i*-layer / bottom 0.3 μm contact n^+-layer
5		Nanoholes diameter/period	Diameter 700 nm / Period 1000 nm

Accepted Manuscript

6		Fabrication	Top-down fabrication / CMOS compatible
7	Fig. 2 Steady-state	EQE	Flat *pin*: 9 % at 850 nm (not shown) Nanohole *pin*: 55 % at 850 nm
8		Responsivity	Flat *pin*: 0.11 A/W at reverse bias -3 V (not shown) Nanohole *pin*: 0.33 A/W at reverse bias -3 V
9		Device diameter	250 μm
10	Fig. 6 Transient	Illumination wavelength/raise time	Flat *pin* device: 850 nm / 27 ps in Fig. 6(b) Nanohole *pin* device: 850 nm / 25 ps in Fig. 6(a)
11		Data rate (Gb/s)	> 25 Gb/s
12		Device diameter	30 μm

Table A1: Experimental data are shown for Fig. 2 nanohole PD devices for steady-state device current as a function of illumination power. Although not shown, flat PD device data are shown in comparison. Experimental data are also shown for Fig. 6(a) nanohole and Fig. 6(b) flat PD devices for transient device output as a function of time when illumination is suddenly turned off. In the theory developed in this particle, there is no difference between *pin* and *nip*.

Procedure for obtaining silicon nanohole PDs for steady-state measurements in Fig. 2(a):

Nanoholes (also called microholes as in Ref. 3) utilized help to couple the vertically incident light to the lateral modes through the horizontal plane of the silicon layer. According to an optics view, this property provides the orthogonal bending of the incident light and let the optical power to distribute uniformly throughout the thin layer and eliminate high density of electron-hole pair accumulation in one spot. The power per unit volume remains at a low level, contributing to highly linear behavior for more than 9 dBm input power and > 3 mA of DC photocurrent when reverse biased above 3.5 Volts. More information about this procedure can be found in Ref. 35, which is based on the optics view. According to this view, photons interfere with nanoholes so that absorption is enhanced.

In the main text above, we have developed a quantum mechanical view such that photon waves interact with 2D periodic nanohole potential, resulting in slow photons. These slow photons in nanohole PDs can stay in the silicon material for longer time, and will be absorbed more than photons in the flat device. Slow photons are also called photon trapping, and the scientific meanings are the same as slow photons.

About the procedure of the measurement, the surface of the 30 μm-diameter photodetector contains nanoholes in a funnel shape profile arranged in a square lattice (Fig. 2a, inset). The wavelength of the light is set to 850 nm and the power is varied from 0 to 12mW.

Procedure for obtaining silicon nanohole PDs for transient measurements in Figs. 6(a) and 6(b):

For the high-speed characterizations, we used a mode-locked pulsed fiber laser with a wavelength of 850 nm, with a sub-picosecond pulse width and a repetition rate of 20 MHz. The incident power on the photodiodes during AC measurements was kept at ~100 μW. The laser pulse was focused onto the active region of the photodiodes using a single-mode lensed fiber tip. The photoresponse, in the form of electrical pulses, was observed on a sampling oscilloscope with a 20 GHz electrical sampling module.

Accepted Manuscript

Digging nanoholes in a silicon crystal increases silicon surfaces. At the semiconductor surfaces, there are surface states for electrons, whose energies are inside the bandgap. Since the electron energies are inside the band gap, these electrons cannot propagate in the silicon crystal, and will be localized at the surfaces. All these defects contribute to undesirable surface states [33,34]. These surface states will be unwanted recombination centers and may reduce photon absorption. Nanohole states must be introduced without increasing surface states, otherwise, they can cause persistent photocurrent, contributing to noise and possible signal distortion during high-speed operations. We have applied several methods to inhibit device degradation induced by surface damage, including thermal oxidation and oxide removal, low ion energy etch, solution-based hydrogen passivation and the growth of a nanoscale thin passivating oxide film via a rapid thermal oxidation (RTO) process [34-37]. The technique is often called passivation.

Hydrogen passivation helped to suppress leakage current in devices that were freshly etched and allowed efficient carrier collection even with a low electrical field. Thermal oxidation and hydrogen passivation reduced the leakage current to the nA level, but the leakage current was on the scale of μA after low ion energy etch. In the case of wet oxidation, it causes a decrease in the EQE of hole-integrated photodiodes due to additional dopant diffusion.

Accepted Manuscript

References:

[1] Simon M. Sze, Physics of Semiconductor Devices, ver. 2 (Willey, New York, 1981).
[2] S. B. Alexander, Optical Communication Receiver Design, (SPIE, Bellingham, 1997).
[3] Yang Gao, Hilal Cansizoglu, Kazim G. Polat, Soroush Ghandiparsi, Ahmet Kaya, Hasina H Mamtaz, Ahmed S. Mayet, Yinan Wang, Xinzhi Zhang, Toshishige Yamada, Ekaterina Ponizovskaya Devine, Aly F. Elrefaie, Shih-Yuan Wang, and M. Saif Islam, Nature Photonics 11 (5), 301-308 (2017).
[4] Yang Gao, Hilal Cansizoglu, Soroush Ghandiparsi, Cesar Bartolo-Perez, Ekaterina Ponizovskaya Devine, Toshishige Yamada, Aly F. Elrefaie, Shih-Yuan Wang, and M. Saif Islam, ACS Photonics 4 (8) 2053-2060 (2017).
[5] Hilal Cansizoglu, Ekaterina Ponizovskaya Devine, Yang Gao, Soroush Ghandiparsi, Toshishige Yamada, Aly F. Elrefaie, Shih-Yuan Wang, and M. Saif Islam, IEEE Trans. Elec. Dev. 65 (2), 372-381 (2018).
[6] Hilal Cansizoglu, Aly F. Elrefaie, Cesar Bartolo-Perez, Toshishige Yamada, Yang Gao, Ahmed S. Mayet, Mehmet F Cansizoglu, Ekaterina Ponizovskaya Devine, Shih-Yuan Wang, M. Saif Islam, IEEE Trans. Elec. Dev., 65 (2), 382-391 (2018).
[7] Ahmed S. Mayet, Hilal Cansizoglu, Yang Gao, Soroush Ghandiparsi, Ahmet Kaya, Cesar Bartolo-Perez, Badriyah AlHalaili, Toshishige Yamada, Ekaterina Ponizovskaya Devine, Aly F. Elrefaie, Shih-Yuan Wang, and M. Saif Islam, J. Op. Soc. Am. B 35 (5), 1059-1065 (2018).
[8] Hilal Cansizoglu, Cesar Bartolo-Perez, Yang Gao, Ekaterina Ponizovskaya Devine, Soroush Ghandiparsi, Kazim G Polat, Hasina H. Mamtaz, Toshishige Yamada, Aly F. Elrefaie, Shih-Yuan Wang, and M. Saif Islam, Phot. Res. 6 (7), 734-742 (2018).
[9] Soroush Ghandiparsi, Aly F. Elrefaie, Ahmed S. Mayet, Taha Landolsi, Cesar Bartolo Perez, Hilal Cansizoglu, Yang Gao, Hasina Mamtaz, Hossein Rabiee Golgir, Ekaterina Ponizovskaya Devine, Toshishige Yamada, Shih-Yuan Wang, and M. Saif Islam, IEEE J. Lightwave Tech. 31 (23), 5748 - 5755 (2019).
[10] Jun Gou, Hilal Cansizoglu, Cesar Bartolo-Perez, Soroush Ghandiparsi, Ahmed S Mayet, Hossein Rabiee-Golgir, Yang Gao, Jun Wang, Toshishige Yamada, Ekaterina Ponizovskaya Devine, Aly F. Elrefaie, Shih-Yuan Wang, and M. Saif Islam, Nanophotonics 8 (10), 1747-1756 (2019).
[11] Hilal Cansizoglu, Ahmed S. Mayet, Soroush Ghandiparsi, Yang Gao, Cesar Bartolo-Perez, Hasina H Mamtaz, Ekaterina Ponizovskaya Devine, Toshishige Yamada, Aly F. Elrefaie, Shih-Yuan Wang, and M. Saif Islam, IEEE Phot. Tech. Lett. 31 (20), 1619-1622 (2019).
[12] Erik Garnett and Peidong Yang, Nano Lett. 10 (3), 1082-1087 (2010).
[13] Suguru Saito, Takashi Oinoue, Yoshiya Hagimoto, and Hayato Iwamoto, Nature 7, 3832 (2017).
[14] Kai Zang, Xiao Jiang, Yijie Huo, Xun Ding, Matthew Morea, Xiaochi Chen, Ching-Ying Lu, Jian Ma, Ming Zhou, Zhenyang Xia, Zongfu Yu, Theodore I. Kamins, Qiang Zhang, and James S. Harris, Nature 8, 628 (2017).
[15] Toshishige Yamada, Hidenori Yamada, Andrew J. Lohn, and Nobuhiko P. Kobayashi, Nanotechnology 22 (5), 055201 (2011).
[16] John Bardeen and Walter H. Brattain, Phys. Rev. 75 (8), 1208-1225 (1949).
[17] Toshishige Yamada, Appl. Phys. Lett. 78 (12), 1739-1741 (2001).
[18] Toshishige Yamada, Appl. Phys. Lett. 80 (21), 4027-4029 (2002).

Accepted Manuscript

[19] Antonis Orphanou, Toshishige Yamada, and Cary Y. Yang, J. Appl. Phys. 119 (21), 214311 (2016).
[20] Toshishige Yamada and David K. Ferry, Solid State Electronics 38 (4), 881-890 (1995).
[21] Toshishige Yamada, Hisashi Yabutani, Tsutomu Saito, and Cary Y. Yang, Nanotechnology 21 (26), 265707 (2010).
[22] Harold A. Wheeler, Proc. I. R. E. 20 (9), 412-424 (1942).
[23] Hidetoshi Takahashi, Electromagnetism (in Japanese), (Shokabo, Tokyo, 1959).
[24] Max M. Shulaker, Gage Hills, Rebecca S. Park, Roger T. Howe, Krishna Saraswat, H-S Philip Wong, Subhasish Mitra, Nature 547 (7661), 74-78 (2017).
[25] Xiaodong Yang, Jie Yao, Junsuk Rho, Xiaobo Yin, Xiang Zhang, Nature Photonics 6 (7), 450-454 (2012).
[26] Maria E. Gracheva, Anlin Xiong, Aleksei Aksimentiev, Klaus Schulten, Gregory Timp, Jean-Pierre Leburton, Nanotechnology, 17 (3), 622 (2006).
[27] Clint J. Novotny, Edward T. Yu, Paul KL Yu, Nano Tell. 8(3), 775-779 (2008).
[28] Rubén C. Villarreal, Francisco Mireles, Ernesto E. Marinero, and Bruce A. Gurney, Appl. Phys. Lett. 98, 172102 (2011).
[29] Shonak Bansal, Kuldeep Sharma, Prince Jain, Neha Sardana, Sanjeev Kumar, Neena Gupta and Arun K. Singh, RSC Adv., 8 (69), 39579-39592 (2018).
[30] Shonak Bansal, Avishek Das, Prince Jain, Krishna Prakash, Kuldeep Sharma, Naveen Kumar, Neha Sardana, Neena Gupta, Sanjeev Kumar and Arun K. Singh, IEEE Trans. Nano. 18, 781-789 (2019).
[31] Shonak Bansal, Krishna Prakash, Kuldeep Sharma, Neha Sardana, Sanjeev Kumar, Neena Gupta and Arun K Singh, Nanotechnology 31 (40), 405205 (2020).
[32] Antonis Orphanou, Toshishige Yamada, and Cary Y. Yang, Nanotechnology 23 (9), 095401 (2012).
[33] John Bardeen, Phys. Rev. 71, 717-727 (1947).
[34] Toshishige Yamada, Charles W. Bauschlicher, and Harry Partridge, Phys. Rev. B 59 (23), 15430-15436 (1999).
[35] Hilal Cansizoglu, Yang Gao, Cesar Bartolo Perez, Soroush Ghandiparsi, Ekaterina Ponizovskaya Devine, Mehmet F. Cansizoglu, Toshishige Yamada, Aly F. Elrefaie, Shih-Yuan Wang, M. Saif Islam, Proc. SPIE 10349, 103491C (2017).
[36] Ahmed S. Mayet, Hilal Cansizoglu, Yang Gao, Ahmet Kaya, Soroush Ghandiparsi, Toshishige Yamada, Shih-Yuan Wang, M. Saif Islam, Proc. SPIE 9924, 9924C (2016).\
[37] Ahmed S. Mayet, Hilal Cansizoglu, Yang Gao, Soroush Ghandiparsi, Ahmet Kaya, Cesar Bartolo-Perez, Badriyah AlHalaili, Toshishige Yamada, Ekaterina Ponizovskaya Devine, Aly F. Elrefaie, Shih-Yuan Wang, M. Saif Islam, JOSA B, 35 (5), 1059-1065 (2018).

Accepted Manuscript