Detecting Single Photons Using Capacitive Coupling of Single Quantum Dots

Yang Zhang,† Yang Wu,‡ Xiaoxin Wang,§ Lei Ying,∥ Rahul Kumar,‡ Zongfu Yu,∥ Eric R. Fossum,§ Jifeng Liu,*† Gregory Salamo,‡ and Shui-Qing Yu*†

†Department of Electrical Engineering, University of Arkansas, Fayetteville, Arkansas 72701, United States
‡Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701, United States
§Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire 03755, United States
∥Department of Electrical and Computer Engineering, University of Wisconsin, Madison, Wisconsin 53705, United States

*Supporting Information

ABSTRACT: Capturing single photons through light–matter interactions is a fascinating and important topic for both fundamental research and practical applications. The light–matter interaction enables the transfer of the energy of a single photon (∼1 eV) to a bound electron, making it free to move either in the crystal lattice or in the vacuum. In conventional single photon detectors (e.g., avalanche photodiodes), this free electron triggers a carrier multiplication process which amplifies the ultraweak signal to a detectable level. Despite their popularity, the timing jitter of these conventional detectors is limited to tens of picoseconds, mainly attributed to a finite velocity of carriers drifting through the detectors. Here we propose a new type of single photon detector where a quantum dot, embedded in a single-electron transistor like device structure, traps a photogenerated charge and gives rise to a sizable voltage signal (∼7 mV per electron or hole by simulation) on a nearby sense probe through capacitive coupling (with a capacitance ∼ aF). Possible working modes of the proposed detector are theoretically examined. Owing to a small lateral dimension of the quantum dot, detailed analyses reveal that the intrinsic timing jitter of the proposed detector is in the femtosecond to subpicosecond range, and the intrinsic dark count rate is negligible up to moderately high temperatures. These figures of merit are orders of magnitude superior to those of the state-of-the-art single photon detectors work in the same spectral range, making the proposed detector promising for timing-sensitive and quantum information applications.

KEYWORDS: single photon detector, quantum dot, electron tunneling, hole trapping, timing jitter, dark count rate, single-electron transistor, voltage signal

The ability to detect weak light down to the single photon level is highly desirable in many fields, such as optical quantum information and communication,1–4 biological and medical imaging,5–8 spectroscopy,9,10 astronomical observation,11 light detection and ranging (LIDAR),12 and satellite laser ranging (SLR).13,14 Traditionally, single photons are detected using devices that can amplify optical signals with a carrier multiplication process. Typical examples are photomultiplier tubes (PMTs) and avalanche photodiodes (APDs). Although these detectors are widely used, PMTs and APDs have their limitations. The former is known to be bulky and fragile, needs to operate at high voltages, and suffers from low detection efficiency in the infrared (IR) region; the latter is primarily plagued by high dark counts and afterpulsing problems.

Recently, quantum dot (QD) based devices have emerged as a promising class of single-photon detectors (SPDs). These detectors usually operate at relative low voltage bias, show very low dark count rates,2 and avoid involving the avalanche multiplication process that brings in large noise. In addition, the spectral response of QD-based SPDs (QDSPDs) can be readily tuned from the infrared to the ultraviolet–visible range, by changing the QD material and size.15,16 On the other hand, QDSPDs feature compact size and offer the opportunity for direct integration with key components of emerging quantum information processing (QIP) technologies, such as QD single-photon sources,17 waveguide,18,19 and quantum memory and repeater.2,20 Over the years, two types of QDSPDs have been developed. The first type was characterized by the use of a small number of electrostatically defined QDs as light absorbers, where the current flowing through a quantum point contact transistor integrated with the QD(s) was monitored as the output signal. QDs in this type of SPDs were typically formed...
by applying negative potentials to gates patterned on the top of a two-dimensional electron gas (2DEG). Komiya et al. reported a single-photon detector consisting of an electrostatically defined QD on a GaAs 2DEG that extended the spectral response to the far-infrared region.23 Rao et al. demonstrated the single photoelectron response with a QD electrostatically formed on a GaAs/AlGa1−xAs modulation-doped heterostructure.22 Gustavsson et al. realized single microwave photon detection using double QDs.23 It should be noted that the detector reported by Gustavsson et al. was not a conventional detector, as this detector was fed with microwave photons generated by itself. The second type of QDSPDs consisted of an array of self-assembled QDs integrated with either a detector, as this detector was fed with microwave photons generated by itself. The second type of potential changes has recently been demonstrated on readout signal for the detector. The successful readout of this single-electron transistors (SETs) and SPDs. Here, the potential change caused by trapping a single electron or hole into the QD. When the single-electron charging energy (e/\text{C}) is much larger than the energy spacing between two consecutive nondegenerate single-electron eigenstates, \( E_s \) can be approximated by \( e/\text{C} \), where \( e \) is the elementary charge and \( \text{C} \) is the self-capacitance of the QD. The detector works in the so-called Coulomb blockade regime. The Coulomb blockade effect helps eliminate the detector noise due to charge number fluctuation in the QD. For simplicity of the discussion, we assume that single photons are absorbed solely by the QDs, and only one photon is absorbed during every detection cycle. The six possible working modes of the SQDSPD are summarized in Table 1, where Modes IV and VI are complementary to Modes III and V, respectively. Modes IV and VI are included for the completeness of the paper. In the following, we will describe Modes I, II, III, and V in detail, but restrict ourselves to describing the stationary states of each mode. The transients between stationary states are omitted for simplicity.

**Working Mode I.** Figure 2a is the equilibrium energy band diagram of the device. For simplicity, we assume that the QD is undoped and hence is free of excess carriers. This assumption holds for all the Modes. In this Mode, an excess electron is required to be trapped in the QD. The ground state (GS) of the electron in the QD is essentially the first quantized state, while the next available state is the excited state (ES). Initially, the device is in the Coulomb blockade state, that is, electrons in the source are prohibited from tunneling into the QD.

To prepare the device for detecting a single photon, nonzero biases are applied to the drain and the bias gate to bring the device out of the Coulomb blockade state (Figure 2b), allowing exactly one electron to tunnel from the source into the QD and then to the drain. Figure 2b corresponds to the conduction regime between two Coulomb diamonds in the SET's stability diagram where continuous single electron current flows through the device.34 Increase of the gate bias to \( V_{bg1} \) puts the device again into the Coulomb blockade state and change near the QD upon single photon absorption. The sense probe is located at the opposite side of the QD to the bias gate in Figure 1a. In principle, the sense probe can be anywhere near the QD (e.g., on the top of the QD). The thickness of the tunnel barrier is chosen to ensure that a single electron or hole can be well localized in the QD under certain gate bias. This property of the tunnel barrier is parameterized by a tunnel resistance, which will be discussed in detail in *Tunnel Resistance*. In addition, the device temperature is set to satisfy: \( kT \ll E_s \), where \( k \) is the Boltzmann constant, \( T \) is the temperature, and \( E_s \) is the energy cost to add one electron or hole into the QD. When the single-electron charging energy (e/\text{C}) is much larger than the energy spacing between two consecutive nondegenerate single-electron eigenstates, \( E_s \) can be approximated by e/\text{C}, where e is the elementary charge and C is the self-capacitance of the QD. The detector works in the so-called Coulomb blockade regime. The Coulomb blockade effect helps eliminate the detector noise due to charge number fluctuation in the QD. For simplicity of the discussion, we assume that single photons are absorbed solely by the QDs, and only one photon is absorbed during every detection cycle. The six possible working modes of the SQDSPD are summarized in Table 1, where Modes IV and VI are complementary to Modes III and V, respectively. Modes IV and VI are included for the completeness of the paper. In the following, we will describe Modes I, II, III, and V in detail, but restrict ourselves to describing the stationary states of each mode. The transients between stationary states are omitted for simplicity.

**Working Mode I.** Figure 2a is the equilibrium energy band diagram of the device. For simplicity, we assume that the QD is undoped and hence is free of excess carriers. This assumption holds for all the Modes. In this Mode, an excess electron is required to be trapped in the QD. The ground state (GS) of the electron in the QD is essentially the first quantized state, while the next available state is the excited state (ES). Initially, the device is in the Coulomb blockade state, that is, electrons in the source are prohibited from tunneling into the QD.

To prepare the device for detecting a single photon, nonzero biases are applied to the drain and the bias gate to bring the device out of the Coulomb blockade state (Figure 2b), allowing exactly one electron to tunnel from the source into the QD and then to the drain. Figure 2b corresponds to the conduction regime between two Coulomb diamonds in the SET’s stability diagram where continuous single electron current flows through the device.34 Increase of the gate bias to \( V_{bg1} \) puts the device again into the Coulomb blockade state and leads to the trapping of an electron in the QD (Figure 2c). In practice, drain and gate biases used in Figure 2b and Figure 2c can be determined from the stability diagram of the device.35

---

**Figure 1.** Schematic diagrams of the SQDSPD: top (a) and cross-sectional (b, c) views. BG and SP in (c) stand for bias gate and sense probe, respectively.
Once a single photon with proper energy is absorbed by the QD, the trapped electron is promoted to the ES (Figure 2d). The excited electron can either relax back to the GS with a rate $\Gamma_{er}$ or tunnel out of the QD with a rate $\Gamma_{ed}$ (Figure 2e), depending on which rate is larger. $\Gamma_{er}$ is proportional to the overlap of the ES and GS wave functions. $\Gamma_{ed}$ is determined by the coupling between the ES and the available state in the drain, and can be tuned by changing the thickness of the tunnel barrier. Once the trapped electron tunnels out of the QD and the gate bias is reduced to $V_{bg1}$, the QD would shortly be replenished with exactly one electron from the source (Figure 2f), and eventually the device is brought back to its initial state shown in Figure 2c by increasing the gate bias from $V_{bg1}$ to $V_{bg2}$.

$\Gamma_{es}$ and $\Gamma_{ed}$ are the electron tunnel rates from the source and to the drain, respectively. $\Gamma_{er}$ is the electron relaxation rate. The drain and gate biases are maintained at $V_{ed}$ from (c) to (e), while the gate bias in (f) is set to $V_{bg1}$ to allow the electron to tunnel from the source. (g, h) Sketches of the potential ($V_{sp}$) transients picked up by the sense probe: (g) corresponds to the case where the excited electron prefers to relax back to the GS, while (h) is the case where the electron tunnels out of the QD. The effect of gate bias change on $V_{sp}$ is not shown in (h).

Table 1. Summary of the Working Modes

<table>
<thead>
<tr>
<th>working modes</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV&lt;sup&gt;a&lt;/sup&gt;</th>
<th>V</th>
<th>VI&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>trapped carrier</td>
<td>electron</td>
<td>hole</td>
<td>hole</td>
<td>electron</td>
<td>hole</td>
<td>electron</td>
</tr>
<tr>
<td>band alignment (dot-barrier)</td>
<td>type I</td>
<td>type I</td>
<td>type II</td>
<td>type I</td>
<td>type I</td>
<td></td>
</tr>
<tr>
<td>band alignment (dot-substrate)</td>
<td>type I</td>
<td>type I</td>
<td>type I</td>
<td>type II</td>
<td>type II</td>
<td></td>
</tr>
<tr>
<td>transition type</td>
<td>intraband</td>
<td>interband</td>
<td>interband</td>
<td>interband</td>
<td>interband</td>
<td></td>
</tr>
<tr>
<td>output signal</td>
<td>voltage or current</td>
<td>voltage or current</td>
<td>voltage</td>
<td>voltage</td>
<td>voltage</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Details of these two working modes can be found, respectively, in Figures S1 and S2 in the Supporting Information (SI).

Figure 2. (a) Equilibrium energy band diagram along the path indicated by the dashed line in Figure 1b. $\mu_s$ and $\mu_d$ are the electrochemical potentials of the source and the drain, respectively. $V_{bg}$ and $V_{ds}$ are the bias gate potential and the drain potential, respectively. All the potentials in this paper are referenced to the source, which is connected to the ground. (b–f) Basic operational steps for Mode I, as detailed in the main text. $\Gamma_{es}$ and $\Gamma_{ed}$ are the electron tunnel rates from the source and to the drain, respectively. $\Gamma_{er}$ is the electron relaxation rate. The drain and gate biases are maintained at $V_{ed}$ from (c) to (e), while the gate bias in (f) is set to $V_{bg1}$ to allow the electron to tunnel from the source. (g, h) Sketches of the potential ($V_{sp}$) transients picked up by the sense probe: (g) corresponds to the case where the excited electron prefers to relax back to the GS, while (h) is the case where the electron tunnels out of the QD. The effect of gate bias change on $V_{sp}$ is not shown in (h).

Once a single photon with proper energy is absorbed by the QD, the trapped electron is promoted to the ES (Figure 2d). The excited electron can either relax back to the GS with a rate $\Gamma_{er}$ or tunnel out of the QD with a rate $\Gamma_{ed}$ (Figure 2e), depending on which rate is larger. $\Gamma_{er}$ is proportional to the overlap of the ES and GS wave functions. $\Gamma_{ed}$ is determined by the coupling between the ES and the available state in the drain, and can be tuned by changing the thickness of the tunnel barrier. Once the trapped electron tunnels out of the QD and the gate bias is reduced to $V_{bg1}$, the QD would shortly be replenished with exactly one electron from the source (Figure 2f), and eventually the device is brought back to its initial state shown in Figure 2c by increasing the gate bias from $V_{bg1}$ to $V_{bg2}$.

Figure 2g and h are the potential transients on the sense probe for the extreme cases where $\Gamma_{er} \gg \Gamma_{ed}$ and where $\Gamma_{er} \ll \Gamma_{ed}$, respectively. The potential of the sense probe under conditions shown in Figure 2c is taken as a baseline potential (black horizontal lines in Figure 2g,h), against which the sense probe potential varies over time. In Figure 2g, a potential change ($\Delta V_{sp1}$) occurs on the sense probe upon absorbing a single photon, which is caused by a charge redistribution in the QD due to the single-electron wave function difference between the GS and the ES, as discussed in detail in our previous work. The drain and gate biases are maintained at $V_{ed}$ from (c) to (e), while the gate bias in (f) is set to $V_{bg1}$ to allow the electron to tunnel from the source. (g, h) Sketches of the potential ($V_{sp}$) transients picked up by the sense probe: (g) corresponds to the case where the excited electron prefers to relax back to the GS, while (h) is the case where the electron tunnels out of the QD. The effect of gate bias change on $V_{sp}$ is not shown in (h).

$\Gamma_{es}$ and $\Gamma_{ed}$ are the electron tunnel rates from the source and to the drain, respectively. $\Gamma_{er}$ is the electron relaxation rate. The drain and gate biases are maintained at $V_{ed}$ from (c) to (e), while the gate bias in (f) is set to $V_{bg1}$ to allow the electron to tunnel from the source. (g, h) Sketches of the potential ($V_{sp}$) transients picked up by the sense probe: (g) corresponds to the case where the excited electron prefers to relax back to the GS, while (h) is the case where the electron tunnels out of the QD. The effect of gate bias change on $V_{sp}$ is not shown in (h).

Table 1. Summary of the Working Modes

<table>
<thead>
<tr>
<th>working modes</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV&lt;sup&gt;a&lt;/sup&gt;</th>
<th>V</th>
<th>VI&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>trapped carrier</td>
<td>electron</td>
<td>hole</td>
<td>hole</td>
<td>electron</td>
<td>hole</td>
<td>electron</td>
</tr>
<tr>
<td>band alignment (dot-barrier)</td>
<td>type I</td>
<td>type I</td>
<td>type II</td>
<td>type I</td>
<td>type I</td>
<td></td>
</tr>
<tr>
<td>band alignment (dot-substrate)</td>
<td>type I</td>
<td>type I</td>
<td>type I</td>
<td>type II</td>
<td>type II</td>
<td></td>
</tr>
<tr>
<td>transition type</td>
<td>intraband</td>
<td>interband</td>
<td>interband</td>
<td>interband</td>
<td>interband</td>
<td></td>
</tr>
<tr>
<td>output signal</td>
<td>voltage or current</td>
<td>voltage or current</td>
<td>voltage</td>
<td>voltage</td>
<td>voltage</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Details of these two working modes can be found, respectively, in Figures S1 and S2 in the Supporting Information (SI).
capacitance of the QD is small (~aF), the potential change induced by a single charge can be in the range of mV, which can be readily measured by state-of-the-art electronic instruments. The electron lifetime \( \tau_{\text{e}} \) in Figure 2h is predominantly determined by the tunnel rate through: \( \tau_{\text{e}} \sim 1/\Gamma_{\text{ed}} \) and the retention time \( \tau_{\text{r}} \) of \( \Delta V_{\text{sp}} \) is dependent on the tunnel rate \( \Gamma_{\text{ed}} \) and the electron transport processes in the electrodes. On the other hand, the photoassisted electron tunneling events shown in Figure 2e give rise to spikes on the drain-source current transients. These current spikes may also be used as output signals for the detector.

**Working Mode II.** In this Mode, a negative or zero gate bias is applied to empty the GS of any electrons and put the device in the Coulomb blockade state (Figure 3a). An electron–hole pair is created by interband transition upon absorbing a single photon (Figure 3b). Driven by the drain-source electric field, the electron tunnels out of the QD with a rate \( \Gamma_{\text{ed}} \), and the hole is trapped (Figure 3c). Asymmetric tunnel barriers are adopted on the source and the drain sides to favor the tunneling of the electron and the trapping of the hole. With the device configuration shown in Figure 3c, the electron tunnel rate \( \Gamma_{\text{ed}} \) is usually larger than the hole tunnel rate \( \Gamma_{\text{hs}} \), mainly due to the thinner tunnel barrier on the drain side and the smaller electron effective mass. In addition, we assume that \( \Gamma_{\text{ed}} \) is also larger than the electron–hole recombination rate \( \Gamma_{\text{eh}} \), such that the electron prefers to tunnel out of the QD rather than recombine with the hole immediately after the optical excitation. This condition can be satisfied by selecting an appropriate tunnel barrier thickness on the drain side. After the electron tunnels out, the electrochemical potential of the QD is lowered due to the net positive charge situating in the QD. The potential lowering is indicated by the black arrow in Figure 3d. The dashed lines are the electron/hole energy levels prior to the tunneling event. Lowering of the QD potential brings the device out of the Coulomb blockade state. Hence, electrons in the source can tunnel into the QD with a rate \( \Gamma_{\text{es}} \) during the lifetime of the trapped hole (Figure 3e). The tunnelled-in electron can either tunnel out of the QD to start another loop (Figure 3c–e) or recombine with the hole to reset the detector naturally (Figure 3f). Alternatively, the device can be reset by applying a large positive gate pulse, which helps push the hole out and neutralize the QD.

Figure 3g,h illustrate the evolution of the sense probe potential driven by the events indicated by blue arrows. The baseline potential (black line) is the sense probe potential measured in Figure 3a. \( \Delta V_{\text{sp}} \) and \( \Delta V_{\text{sp4}} \) are potential changes induced by dipole creation in the QD and by electron tunneling event, respectively. \( \Delta V_{\text{sp3}} \) vanishes once the electron recombines with the hole (Figure 3g), while \( \Delta V_{\text{sp4}} \) disappears every time the electron tunnels into the QD from the source or the hole is pushed out of the QD by a gate pulse (Figure 3h). Since the hole has a relatively long lifetime before annihilation, the electron may pass through the QD multiple times. This gives rise to a train of \( \Delta V_{\text{sp4}} \) signals, as shown in Figure 3g,h. The electron lifetime \( \tau_{\text{e}} \) at the GS is determined by \( \Gamma_{\text{ed}} \) through \( \tau_{\text{e}} \sim 1/\Gamma_{\text{ed}} \). Similar to \( \tau_{\text{r}} \), the \( \Delta V_{\text{sp4}} \) retention time \( \tau_{\text{r}} \) is related to \( \Gamma_{\text{es}} \) and the electron transport processes in the electrodes. As will be shown in Simulation of the Potential Change in SQDSPD, \( \Delta V_{\text{sp4}} \) is typically more than an order of magnitude larger than \( \Delta V_{\text{sp3}} \). Therefore, it is expected that \( \Delta V_{\text{sp4}} \) is the major contributor to the voltage output signal of the detector. Like in Mode I, the electron tunneling events may also manifest...
Figure 4. (a−e) Energy band diagrams illustrating the operation of Mode III along the timeline. In (e), only the hole tunneling process is sketched for the resetting of the detector. (f) Sense probe potential transient. The $V_{sp}$ in (a) is used as the baseline potential (black lines). Constant drain and gate biases are applied in (a−e).

Figure 5. Operation of Mode V: (a−e) and (a′−e′) are energy band diagrams along paths indicated by the dashed lines in Figure 1b and c, respectively. $V_s$ is the substrate bias. The epi-layer is not drawn in (a′−e′) for simplicity. For this Mode, the substrate shown in Figure 1a−c is disconnected from the source and biased at a positive potential to facilitate the extraction of the electron through the substrate. (f) Energy band diagram showing the reset of the detector by a gate pulse. (g) Sense probe potential transient. The change of $V_{sp}$ after resetting the detector by a gate pulse is indicated by a yellow dashed line.
themselves as spikes on drain-source current transients. These spikes may be considered as the current output signals of the detector.

**Working Mode III.** Figure 4a shows the initial-state energy band diagram of the detector under applied biases. Unlike in Modes I and II, this Mode adopts a type II band alignment between the QD and the barrier layer. With this configuration, only the hole is confined in the QD while the electron escapes readily from the QD after photoexcitation (Figure 4c). Similar to the situation in Figure 3d, the trapped hole leads to a drop in the electrochemical potential of the QD (Figure 4d), and an electrostatic potential change also exhibits on the sense probe as an output voltage signal. Eventually, the detector is reset by one of the two processes: (i) hole tunneling out of the QD with a rate \( \Gamma_{hs} \) and recombining with an electron at the metal–semiconductor interface; or (ii) a large gate pulse bias pushing the hole out of the QD and then recombining with the electron in the source. In this Mode, the electron has very low chance to enter the QD again from the source side since its injection barrier \( \phi \) is quite high (Figure 4e).

Figure 4f shows a sketch of the sense probe potential transient. \( \Delta V_{sp5} \) has the same origin as \( \Delta V_{sp3} \); \( \Delta V_{sp6} \) denotes the potential change after the escape of the photogenerated electron. \( \tau_{e} \) is related to the electron transit time and is negligibly short for a small QD and a thin barrier layer; \( \tau_{h} \) is essentially the hole lifetime, and can be estimated with 1/\( \Gamma_{hs} \). Since \( \Gamma_{hs} \) is usually small for thick barriers, \( \tau_{h} \) is relatively long and can be of the order of microsecond or even longer.\(^{39,40}\) This long lifetime translates into a long retention time for \( \Delta V_{sp6} \), which helps relax the stringent requirements imposed on the timing and the speed of the measurement electronics connected to the sense probe. To speed up the reset process, a positive pulse can be applied to the gate to eliminate \( \Delta V_{sp6} \) (yellow dashed line in Figure 4f) and hence bring the detector back to the initial state shown in Figure 4a. For this detector, \( \Delta V_{sp6} \) is used as the output signal. Here, a current spike might not be detectable because the electron cannot be reinjected into the QD to finish the loop.

**Working Mode V.** In this Mode, the energy band alignment between the QD and the barrier is type I while it is type II between the QD and the substrate. The energy band diagrams are shown in Figure 5. As in the other Modes, the QD is initially empty of any excess carriers by applying appropriate biases shown in Figure 5a. After absorbing a single photon, an electron–hole pair is created in the QD (Figure 5b,b'). Proper barrier thickness is chosen to prevent the electron and hole
from easy escaping to the drain and source, respectively. Owing to the type II band alignment and the positive substrate bias, the photogenerated electron runs away through the substrate, leaving the hole trapped in the QD (Figure 5e,f). The electrochemical potential of the QD is decreased due to the trapped positive charge (Figure 5d,f). The detector may be reset naturally by hole tunneling out of the QD (red arrows in Figure 5e,f). Instead, it can also be reset by grounding the substrate and applying a large positive gate bias to push the hole out of the QD (Figure 5f). Figure 5g shows the detector response to a single photon, which is similar to that of the Mode III, as shown in Figure 4f.

## SIMULATION OF THE POTENTIAL CHANGE IN SQDSPD

In Working Modes of the Proposed SQDSPD we have described in detail four Modes of the proposed SQDSPD. Upon absorbing a single photon, all these Modes give rise to a potential change (voltage output) on the sense probe, which is triggered by an event of electron escaping from the QD or by a charge redistribution in the QD (Mode I). The voltage output signal originating from the charge redistribution has been discussed in detail in our previous work. Here, using Mode II (type I band alignment) and Mode III (type II band alignment) as examples, we focus on assessing the values of potential changes caused by electron escaping from the QD with COMSOL simulations. Figure 6a is a three-dimensional (3D) schematic diagram of the detector used for the simulation. The QD features a lens shape similar to those reported in the literature. The voltage output signal is extracted by evaluating the average voltage change on a stripe-shaped sense probe positioned on the top of the QD.

### Simulation Results of Working Mode II

The energy band diagrams of the material system used for this simulation are sketched in Figure 6b,c. The simulation procedures were similar to those in ref 36, except that the single-electron and single-hole wave functions were obtained by solving 3D Schrödinger’s equation and Poisson’s equation in a self-consistent manner. Since the applied $V_{bg}$ and $V_a$ broke the rotational symmetry of the structure, 3D modeling in COMSOL was called upon for the simulation. The calculation scheme can be found in Figure S3 in the SI. The parameters used in the simulation for Mode II is summarized in Table 2.

Figure 7 presents the ground-state single-electron and single-hole wave functions from a self-consistent calculation for the dipole case as in Figure 3b. Owing to the type I band alignment, the electron and the hole were both confined in the QD, with the electron wave function spreading out more throughout the QD than the hole wave function doing. Under a relatively large negative gate bias (see Table 2), the hole wave function shifted toward the bias gate, and the shape of the hole wave function departed from a perfect circle to form an egg-like shape on the $y$–$x$ plane (Figure 7d). It seemed that the hole screened out most of the electric field exerted by the bias gate, rendering the electron wave function shape in the $y$–$x$ plane almost undistorted from a circle (Figure 7a), as expected for the wave function under zero field. After the electron’s escape by tunneling, the trapped single-hole wave function was also calculated without using a self-consistent method because it was a single charge problem. The results are shown in Figure S4 in the SI.

The potential distribution of the detector in space was obtained by solving the Poisson’s equation in the presence and absence of charges in the QD. The potential difference was evaluated by subtracting the zero-charge potential distribution (as in Figure 3a) from the dipole (as in Figure 3b) or single hole (as in Figure 3d) potential distribution, and the results are presented in Figure 8. It is not surprising that the positions of maximum potential difference were correlated well with the maxima of the wave functions shown in Figures 7 and S4. The potential differences faded away from the maxima due to dielectric screening. For the electron–hole dipole case (Figure 8a–c), the sign of the potential difference ($\Delta V$) could be either positive or negative, depending on the position where $\Delta V$ was extracted. Figure 8c shows the voltage output signal extracted with an infinitesimal sense probe kept at a constant distance to the top of the QD (see inset in Figure 8b). As the angle ($\theta$) swept from 30.9° to 170.6°, the output signal changed its polarity from positive to negative, and it exhibited a maximum value of 6.3 mV at 39° and a minimum value of −3.5 mV at 136°. In contrast, the output signal for the single-hole case maintained positive across all angles (Figure 8f) and peaked with a value of 12.4 mV at 44°. In practice, it is nontrivial to position the sense probe at individual locations indicated by the red asterisks in Figure 8b. We considered the average output signal on a zero-width stripe formed by connecting all the asterisks. To evaluate the average output signal, we integrated the curves in Figure 8c,f over the angles and then divided the results by the range of angles. The calculated average output signals were 0.11 and 7.0 mV for the dipole and single-hole cases, respectively. The former was more than an order of magnitude smaller than the latter, which could be attributed to the canceling effect of the positive and negative charges in space. Therefore, we conclude that the voltage output signal of the Mode II should mainly originate from the potential change induced by the trapping of a single hole in the QD.

### Simulation Results of Working Mode III

For the simulation of Mode III, GaSb and GaAs were adopted for the QD and barrier materials, respectively. The energy band diagrams of this Mode are displayed in Figure 6d,e. The type II band alignment between the dot and the barrier together with the applied drain bias, facilitated the escape of the electron through the barrier to the drain. The simulation parameters for Mode III were identical to those in Table 2, except that the barrier thickness was 3.2 nm on both sides to provide strong confinement to the hole. In this simulation, only $\Delta V_{q6}$ in Figure 4f was assessed.

Figure 9a–c shows the ground-state wave function of the trapped hole cut through different planes. As expected, this wave function resembled closely the shape of the trapped hole.
wave function obtained in Mode II (Figure S4). The potential difference in space is presented in Figure 9d−f. Similar to Mode II, we assessed the average output signal on a stripe-shaped sense probe by integrating the curve in Figure 9f over the angles. The obtained output signal was 7.3 mV, comparable to the output signal given by the trapped hole in Mode II.

**Figure 7.** Color-coded plots of the ground-state electron and hole wave functions of the Mode II for the dipole case. (a−c) Electron wave function. (d−f) Hole wave function. The directions of x, y, and z axes are defined in Figure 6a, but the absolute coordinate values for the simulations are not shown. The dashed lines in (a) and (d) indicate locations of the cuts for z=x and z=y planes. The y=x plane was cut at z = 0 nm.

**Figure 8.** Potential difference in space for Mode II. Color coded plots: (a, b) electron−hole dipole case calculated using the wave function in Figure S4. Dashed lines indicate the cuts for z=x planes and the possible positions for the sense probes. The y=x plane was cut at z = 1 nm. The inset in (b) illustrates a cross-sectional view of the QD/barrier layer and the sense probe locations. The distance d (marked by white arrows) between the sense probe and the top of the QD was kept at 5 nm. θ is the angle defined as shown in the inset. (c) and (f) are the voltage output signals extracted in (b) and (e), respectively, at different sense probe positions represented by red asterisks drawn in (b).

**DISCUSSION**

In previous sections, we have described the operation principles of the SQDSPD and have evaluated the amplitude of the voltage output signal for Modes II and III by COMSOL simulations. As mentioned in the first section of this paper, one huge advantage of the SQDSPD is that the response of the detector can cover a broad spectrum range that can be easily tuned by using QD of different materials15 or by leveraging the quantum size effect.16 In Mode I, the trapped electron is excited...
by a single photon from the GS to the ES through an intraband transition, which suggests a spectral response ranging from near-infrared (NIR) to far-infrared (FIR) depending on the energy spacing between the ES and the GS. In Modes II–VI, an electron–hole dipole is created in the SQDSPD via an interband transition upon absorbing a single photon. Considering the energy bandgaps of common III–V and II–VI semiconductor materials, the spectral response of Modes II–VI covers the range from ultraviolet (UV) to NIR. In this section, we estimate some important device parameters related to the SQDSPD and discuss possible readout schemes for the detectors.

**Tunnel Resistance.** As elucidated in Working Modes of the Proposed SQDSPD, the operation of the SQDSPD relies on the successful quantum confinement of a single electron or a single hole in the QD. Like in SETs, the strength of the quantum confinement is directly related to the barrier thickness and height, and can be measured by a phenomenological quantity “tunnel resistance” in the small bias limit.\(^3\) To achieve a good confinement for the single charges in the QD, it requires that the tunnel resistance \(R_T\) satisfies the relationship \(R_T \gg \hbar/\epsilon^2 \approx 25.8\, \text{k}\Omega\).\(^4\) where \(\hbar\) is the Planck constant and \(\epsilon\) is the elementary charge. In the SQDSPD, \(R_T\) can be readily tuned by changing the barrier thickness for given material systems. In the simulations of Modes II and III, appropriate barrier thicknesses were selected to give desirable tunnel resistances for the electron or the hole.

Using the Wentzel–Kramers–Brillouin (WKB) approximation, the tunnel resistance is linked to the barrier thickness through the following equations:\(^5\)

\[
R_T = \frac{\hbar}{2\pi\epsilon^2 |\mathcal{I}|^2 D_V D_T}
\]

\[
|\mathcal{I}|^2 = \exp\left[-\frac{2}{\hbar} \int_0^d \sqrt{2m^* (V(y) - E_f) dy}\right]
\]

where \(\hbar\) is the reduced Planck constant, \(\mathcal{I}\) is the tunneling probability through the barrier, \(D_V\) and \(D_T\) are, respectively, the density of states on the initial side and on the final side of the tunnel barrier, \(d\) is the barrier thickness, \(m^*\) is the electron or hole effective mass in the tunnel barrier, \(V(y)\) is the potential energy profile of the barrier for the electron or the hole, and \(E_f\) is the carrier kinetic energy in the \(y\) direction. For the \(R_T\) estimation, constant \(V(y)\)‘s equal to the band offsets shown in Figure 6b–e were used. Under an electric field in the \(y\) direction, \(E_f\) was approximated with the first quantized level energy with respect to the corresponding band edge.

For the operation of Mode II, the photogenerated electron was required to escape from the QD with ease (Figure 3c). Therefore, the drain-side barrier should not be too thick so as to facilitate the tunneling of the electron. On the other hand, the hole was required to be trapped in the QD after photogeneration (Figure 3d), which was achieved by using a relatively thick barrier on the source side. In the simulation of Mode II, we adopted 1.4 and 2 nm barriers for the drain side and the source side, respectively. The former gave an electron \(R_T\) of 28.2 k\Omega\) while the latter resulted in a hole \(R_T\) of 9.6 M\Omega\) and an electron \(R_T\) of 2.4 M\Omega\) for the electron that re-entered the QD from the source. Details of the resistance calculation can be found in the SI. The electron tunnel resistance on the drain side was slightly larger than the resistance quantum (25.8 k\Omega\), indicating that the electron was in the weak confinement regime. In contrast, the hole tunnel resistance was almost 3 orders of magnitude larger than the resistance quantum, suggesting that the hole would be well localized in the QD. The moderate 2.4 M\Omega\) \(R_T\) on the source allowed the electron to re-enter the QD with a finite probability and hence close the loop. This is important if current is to be used as the detector output. In Mode III, a 3.2 nm barrier, which corresponded to a \(R_T\) of 550 M\Omega\), was used to provide a strong confinement for the hole. The strong confinement would help prolong the trapped hole lifetime in the QD, which might be favorable for the readout of the voltage signal (\(\Delta V_{\text{sp6}}\) in Figure 4f).
**Intrinsic Timing Jitter.** Timing jitter is an important figure of merit for a SPD. It is a measure of the timing performance of the SPD and is defined as the statistical variation in the time interval between the single photon arrival and the detection of an output signal. A smaller timing jitter means a more accurate correlation of single photon events on the time axis. A small timing jitter is crucial for many timing-sensitive applications, such as time-correlated single photon counting, LIDAR, SLR, quantum information and communication, and so on. State-of-the-art Si single-photon avalanche diodes (Si-SPADs) have a timing jitter of ~30–50 ps. For these SPDs, further reduction of the timing jitter is hindered by the finite carrier transport time through the detector and other factors.

For the SQDSPD proposed here, the dimension of the active area is only tens of nanometers, which significantly reduces the contribution of the carrier transport time in the QD to the timing jitter. In addition, the potential change through capacitive coupling as the output signal and the absence of the avalanche process help bring the timing jitter close to its intrinsic value ($\Delta \tau_0$). In Mode I (for $\Delta V_{\text{sel}}$) and Mode II, $\Delta \tau_0$ is dominated by the electron transport. Since the electron wave function is well distributed in the QD (see Figure 7a–c for example), $\Delta \tau_0$ is mainly determined by the tunnel rate $\Gamma_{\text{ed}}$ in Figure 2e and Figure 3c through $\Delta \tau_0 \approx 1/\Gamma_{\text{ed}}$. For Modes III and IV, $\Delta \tau_0$ is also governed by the electron transport. However, it should be negligibly small in these two modes because the electron wave function is well extended across the QD and there is no barrier to prevent the electron from exiting the QD.

On the other hand, $\Delta \tau_0$ is controlled by the hole transport in Modes IV and VI. In these Modes, the hole wave function mainly occupies a portion of the QD due to the hole’s relatively large effective mass, similar to the situation shown in Figure 7d–f. In this case, the particle picture of the hole should be involved to estimate the finite transit time across the QD. For Mode IV with an InAs/GaSb (dot/barrier) material system, using a hole drift velocity of $5 \times 10^4$ m/s in InAs and a QD radius of 10 nm, $\Delta \tau_0$ was estimated to be 0.2 ps. This value is still more than 2 orders of magnitude smaller than the best SPD timing jitters reported in the literature and comparable to the estimated intrinsic timing jitter of superconducting nanowire single-photon detectors (SNSPDs).

The timing jitter of a detector system is determined by jitter contributions from various sources such as intrinsic timing jitter, electronic noises, readout circuit, light source and so on. Since the intrinsic timing jitter imposes a fundamental limit on the timing jitter of the detector system, it would be interesting to provide an estimate of it. Below we estimate the intrinsic timing jitter of Mode II at a sufficiently low temperature by calculating the tunnel rate $\Gamma_{\text{ed}}$ using the Fermi’s golden rule. The transition rate $\Gamma_{\text{ed}}$ from an initial state $i$ in the QD to a final state $f$ in the metal contact was expressed as

$$\Gamma_{\text{ed}}(\Delta F) = \frac{2\pi}{\hbar} |T|^2 \int_{E_i}^{E_f} dE \delta(E_i - E_f - \Delta F)$$

where $|T|^2$ was the tunneling probability from the initial state to the final state, $E_i$ and $E_f$ were, respectively, the energies of the initial and final state, and $\Delta F$ was the free energy change of the system caused by the transition. The delta function ensured that the energy was conserved for the transition process.

Assuming a constant tunneling probability $|T|^2$ for all tunneling modes, the total transition rate, involving all the relevant occupied states in the QD and unoccupied states in the metal contact, was given by

$$\Gamma(\Delta F) = \frac{2\pi}{\hbar} |T|^2 \int_{E_i}^{E_f} dE \delta(E_i - E_f - \Delta F)$$

where the density of states (DOS) of the occupied states in QD ($D_i$) and the unoccupied states in metal contact ($D_f$) were introduced to convert the summations over momentum to the integrals over energy. $f(E)$ was the Fermi–Dirac function. $E_i$ and $E_f$ were the conduction band bottom of the QD and the metal contact Fermi level, respectively.

It was assumed that only one energy level in the QD was involved in the transition. Due to its finite width, the DOS of the quantized energy level in the QD was approximated with a Lorentzian function:

$$D_i(E_i) = \frac{\hbar \Gamma_i}{2\pi((E - E_i)^2 + (\Gamma/2)^2)}$$

where $\Gamma_i$ was the electron transition rate. It was essentially the tunnel rate $\Gamma_{\text{ed}}$ in Figure 3c since it was assumed that $\Gamma_{\text{ed}} \gg \Gamma_i$. $E_i$ was the lowest quantized level involved in the transition. On the other hand, the DOS in the metal contact ($D_f$) was assumed to be a constant, and was calculated with the two-dimensional DOS of the contact metal by

$$D_f(E_f) = \frac{m^*_e A}{\pi \hbar^2}$$

where $m^*_e$ was the electron effective mass in the metal contact and $A$ was the contact area. Inserting eq 5 and eq 6 into eq 4 and using the tunneling probability through the barrier under bias, the tunnel rate $\Gamma_{\text{ed}}$ found to be $1.5 \times 10^{14}$ s$^{-1}$. The intrinsic timing jitter was then estimated to be

$$\Delta \tau_0 \approx \frac{1}{\Gamma_{\text{ed}}} = 6.6 \text{ fs}$$

This intrinsic timing jitter is about 4 orders of magnitude smaller than the timing jitters of the state-of-the-art SPDs, making the SQDSPD very promising for timing-sensitive applications.

**Dark Count Rate.** Dark count rate (DCR) is another key performance parameter of a SPD, which is defined as the average number of false counts recorded by a SPD per second in the absence of light input. The DCR should be as low as possible to reduce the error in detecting single photons. The origin of the DCR usually differs for dissimilar types of SPDs, and is related to the materials properties, bias conditions, external disturbances, and so on. For semiconductor-based SPDs, the DCR is associated with excess carriers generated by thermally excitation and/or field-enabled tunneling processes. Since the SQDSPD proposed here operates under moderate electric field (<10$^4$ V/cm), the contribution of field-enabled tunneling processes to the DCR can be neglected. Hence, we focus on the DCR contributed by thermal excitation.

In general, valence band electrons can be thermally excited to the conduction band either directly or indirectly through midgap trap states. The total generation rate ($G_{\text{th}}$) is given by the sum of the net generation rates of the direct excitation and indirect excitation in a QD. The direct excitation rate can be expressed as

$$G_{\text{th}} = \sum \frac{\hbar^2}{2m^*_e} |T|^2 E_{\text{c,i}} \delta(E_{\text{c,i}} - E_f - \Delta F)$$

where $E_{\text{c,i}}$ was the energy of the conduction band in the metal contact.
The total DCRs in Figure 10 were negligible at low temperatures presented in Figure 10a and b, respectively. For both SPDs, the small active volume of the order of $10^{-5}$ cm$^3$ due to the minute thermally generated carrier concentrations is the output voltage signal, $I_{\text{spike}}$ in (b) is the current spike in Modes I and II.

$$G_d = B n_i^3 V$$

where $B$ is the radiative recombination coefficient, which is 1.2 $\times$ $10^{-9}$ cm$^3$s$^{-1}$ for InAs at 77 K. $n_i$ is the intrinsic carrier concentration, and $V$ is the QD volume. The net generation rate due to indirect excitation in a QD is given by the Shockley–Read–Hall (SRH) model with

$$G_{SRH} = \frac{n_i V}{\tau_{SRH}}$$

where $\tau_{SRH}$ is the carrier lifetime associated with SRH processes. Equation 9 is derived by assuming that only a single SRH generation center level close to the intrinsic Fermi level is involved in the generation and the lifetime is identical for electrons and holes. The intrinsic carrier concentration $n_i$ is a function of the bandgap and the temperature and is given by

$$n_i = C T^{1.5} \exp\left( -\frac{E_g}{2kT} \right)$$

where $C$ is a constant determined by electron and hole effective masses, and $E_g$ is the bandgap of the semiconductor material.

The DCR was estimated with $DCR = G_d$ by assuming that every thermally generated carrier results in a dark count, that is, the carrier escape and collection efficiencies were 100%. The DCRs for SPDs with an InAs QD and an GaSb QD were calculated using eqs 8–10 (see details in SI), and the results are presented in Figure 10a and b, respectively. For both SPDs, the total DCRs in Figure 10 were negligible at low temperatures due to the minute thermally generated carrier concentrations and the small active volume of the order of $10^{-10}$ cm$^3$. The DCRs exceeded 1 Hz at 165 and 265 K for the InAs-based and GaSb-based SPDs, respectively. In contrast, state-of-the-art SPDs (e.g., SNSPDs), which work in the same near- to mid-IR range, have DCRs ranging from <1 to 10 Hz measured at extremely low temperatures ($<$5 K). The DCR data shown in Figure 10 suggest that the SQDSPD is promising to operate at much higher temperatures than the SNSPDs do while maintaining similar DCR performance. The total DCRs in Figure 10 started to rise significantly after the temperature exceeded $\sim$200 K, and it became obvious that the calculated SRH contribution to the total DCR dominated over the direct excitation contribution, which could be explained by the existence of a large number of SRH centers associated with a moderate crystal quality assumed for the InAs and GaSb QDs in the calculation (see the SI for details). For the GaSb-based SPD, the relatively large bandgap helped suppress the DCR contribution from direct excitation. The InAs-based detector showed a total DCR of 3663 Hz at 300 K, which was more than 2 orders of magnitude larger than the DCR ($\sim$13) of the GaSb-based detector. It should be noted that, the operation temperature of the SQDSPD may be well below 300 K, because it needs to satisfy the requirements for Coulomb blockade, as discussed in Working Modes of the Proposed SQDSPD. For example, the operation temperature of SQDSPDs simulated in Simulation of the Potential Change in SQDSPD should be below $\sim$80 K in this respect.

**Readout Scheme.** For all the Modes of the proposed SQDSPD, the voltage change on the sense probe can be used as an output signal. To use the SQDSPD, a circuit needs to be integrated with the device sketched in Figure 1 to read out the voltage change. In our previous work, we suggested a possible readout scheme for picking up this voltage change near the QD blockade, as discussed in Working Modes of the Proposed SQDSPD, which can be as long as microseconds. This long signal retention time helps relax significantly the tight timing requirements imposed on the readout circuit, as compared to the situation in ref 36. Hence, the design of the readout scheme can emphasize more on the optimization of parameters such as the amplification of the voltage signal, rather than be limited by the short readout time. In addition to higher speed readout circuits (like that in ref 36), if layout space and power permit,
slower operational amplifier circuits, like those shown in Figure 11, may also be used. Figure 11a shows an example of a readout scheme that offers an amplification of the voltage signal by a factor of $R_2/R_1$. Since the slew rate of common operational amplifiers is $-\frac{V}{\mu s}$, $V_{in}$'s with a retention time of $\sim \mu s$ can be amplified without distortion at the output end.

For Modes I and II, current spikes appear on the transient drain-source currents may also be used as output signals to indicate single photon events.\footnote{For Modes I and II, current spikes appear on the transient drain-source currents may also be used as output signals to indicate single photon events.} Since it is essentially the current associated with a single charge transport, proper current (or charge) amplification may be needed. In addition, the speed of the readout scheme should be adequately fast to catch the current spikes which have a width of ps–ns. Figure 11b presents a two-stage readout scheme that may fulfill the above requirements. The first stage is essentially a charge amplifier which turns the current spike into a voltage signal ($V_{out}$) with a decay time constant of $R_2C$. Using the elementary charge of 1.6 $\times 10^{-19}$ C and a capacitance of 0.1 pC, $V_{out}$ is of the order of 1 $\mu V$. The second stage serves to amplify this small voltage signal with a gain of $R_3/R_c$.

**CONCLUSION**

In summary, we have proposed a new type of SPD and analyzed six possible working modes of the detector. The proposed SPD leveraged the voltage changes, arising from the capacitive coupling between a single photogenerated charge trapped in the QD and a nearby sense probe, as output signals. Simulation results on working modes II and III indicated a voltage signal of $\sim 7$ mV. Further analyses revealed intrinsic timing jitters of fs to subps and DCRs of negligible values at low temperatures. These results suggest that the proposed SPD holds great potentials for single-photon detection applications where superior timing accuracy is required. Furthermore, the proposed SQDSPD features an ultracompact footprint ($\sim 500$ mm$^2$), which is likely favorable for applications such as integrated quantum optics and communications. On the other side of the coin, this small footprint may bottleneck the detection efficiency of the proposed SQDSPD due to a low coupling efficiency between the light source and the QD in a conventional optical fiber-coupling setup. However, this difficulty may be circumvented by switching to a waveguide–QD coupling scheme that has recently been demonstrated to offer a near-unity coupling efficiency.\footnote{This difficulty may be circumvented by switching to a waveguide–QD coupling scheme that has recently been demonstrated to offer a near-unity coupling efficiency.}

**ASSOCIATED CONTENT**

- Supporting Information: The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsphto1515.
  - Details of Mode IV; Self-consistent calculation scheme
  - Single hole wave function and potential difference for Mode II; Calculation details of timing jitter and DCR (PDF).

**AUTHOR INFORMATION**

**Corresponding Authors**

- E-mail: yangzh08@gmail.com.
- E-mail: jifeng.liu@dartmouth.edu.
- E-mail: syu@uark.edu.

**ORCID**

Yang Zhang: 0000-0002-6208-9392

**Author Contributions**

Y.Z. laid out the scheme of this work, expanding J.L. and S.-Q.Y.'s original plot into all possible working modes. Y.W. and X.W. conducted the simulation. E.R.F. provided input on readout schemes. Y.Z. analyzed the data and wrote the manuscript with the inputs from J.L. and S.-Q.Y. All the authors commented on the manuscript. S.-Q.Y., G.S., and J.L. supervised this work.

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This research was sponsored by the DARPA DETECT program through Army Research Office (ARO) Cooperative Agreement Number W911NF-16-2-0162. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

**REFERENCES**


DOI: 10.1021/acsphotonics.7b01515
ACS Photonics XXX, XXX, XXX–XXX


