Exciton Condensation in Topological Insulators

Harry Fetsch
Physics REU Program, Department of Physics, University of California Davis and Department of Physics, Harvey Mudd College

Topological Insulators are materials of particular interest in condensed matter physics because of their gapless conducting surface states. When illuminated with a laser, an electron-hole pair is created in a topological insulator; under some conditions, these electrons and holes can bond to form quasiparticles called excitons. Because they have integer spin, excitons below some critical temperature $T_c$may undergo Bose-Einstein condensation into the ground state. According to some theoretical predictions, in this state excitons flow through the material without friction, analogously to superfluid helium, or Cooper pairs in a superconductor. Bismuth Selenide (Bi_2Se_3) is a topological insulator that can be grown into long, thin wires and measured with Scanning Photocurrent Microscopy (SPCM). Through photocurrent measurements in antimony-doped Bi2Se3 nanowires, we have observed carriers with a long transport length (greater than 1mm), and high mobility. Based on bias voltage and gating voltage dependence, we conclude that these measurements are explained by an exciton Bose-Einstein condensate, with a critical temperature above 79K. Investigation of these quantum effects on a macroscopic scale may lead to applications in spintronics and high-temperature quantum computing systems.

Background

The background provided here is almost entirely qualitative, and at the level needed to understand the experimental work in this lab. For a deeper and more quantitative explanation of the concepts introduced here, please see the References.

Crystalline solids are frequently modeled as a potential that repeats periodically some finite, but arbitrarily large, number of times. Even in the simplified case of a pure material with non-interacting electrons, solving for the electrons wave function in this model leads to an interesting behavior. If the wave function is assumed to be periodic like the lattice, the energies of the solutions of the Schrodinger equation are quantized, and at certain energies no states are allowed. This has the effect of forming the allowed energy levels into bands of closely-spaced energies separated by large gaps. Semiconductors are a class of materials in which the Fermi level falls within a gap, meaning that the band below this gap, called the valence band, is completely filled, and the band above it, called the conduction band, is empty. For a material to be a semiconductor instead of an insulator, this gap must be small enough that electrons can be thermally excited into the conduction band, allowing them to carry a current. Figure 1 shows the dispersion relation, the graph of allowed energies as a function of the wave vector k (proportional to the momentum), for a typical semiconductor.

Topological Insulators (TIs) are a class of materials which show a different and interesting dispersion relation on their surface. In some materials, the bulk has an insulating band gap, but spin-orbit coupling leads to conducting surface states that cross the band gap (Figure 2). The topology of these materials creates a distinct electronic dispersion relation, giving them their name Topological Insulators [1]. This topological invariance means that, even if there are defects in the crystal, the surface states are protected– unsurprisingly, this is an enormous benefit for experimental investigations where measuring pure samples is nearly impossible. The surface states of a TI are additionally interesting because, as shown in the figure, the dispersion relation in the band gap is linear, and

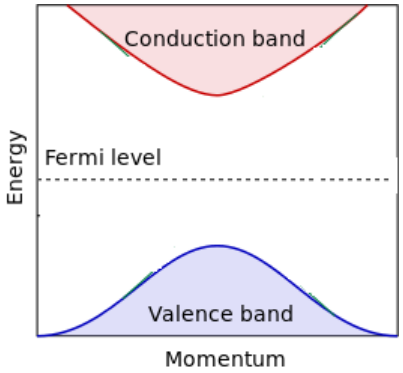

FIG. 1: Simple diagram of the band structure of a semiconductor. The Fermi level lies within the band gap, and at absolute zero the conduction band is completely empty.

fermions in these states experience spin-momentum locking. Potential applications of this fact will be discussed later, but for moment the important implications are that fermions (electrons and holes) can move along the surface of a TI with high velocity and without backscattering. The point where these lines cross within the band gap is called the Dirac point, and is located near the middle of the gap.

The other essential piece of theoretical background is the process of exciton formation. In an ordinary semiconductor, photons incident on the surface can promote electrons from the valence band to the conduction band, provided that the photons energy is greater than the band gap. The vacancy in the valence band created by the electron is called a hole, and within the crystal it can be treated as a positively-charged particle. When the semiconductor is attached to two metal contacts that form a circuit, free carriers (electrons and holes) can travel to the contacts, creating a photocurrent in the circuit. However, under certain conditions, instead of diffusing as free carriers, the Coulomb interactions between electrons and holes can bind them together to form quasiparticles called excitons. A condition for exciton formation is that the energy

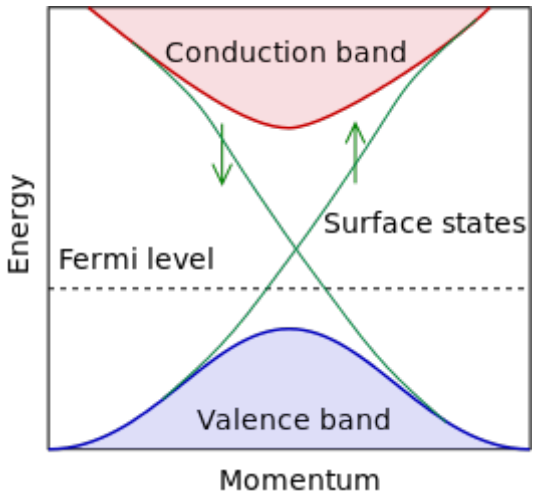

FIG. 2: Band structure for a Topological Insulator

of the Coulomb interactions between electrons and holes is at least on the order of their thermal kinetic energy. This gives an upper temperature limit on exciton formation.

When excitons are generated by photoexcitation in a semiconductor, they diffuse through the sample until they reach the contacts. Here, the internal electric field between the metal contact and semiconducting sample separates the excitons back into electrons and holes. Provided that the contacts are connected to each other through some external circuit, the electrons enter the metal, where they produce a current, and the holes travel back through the sample toward the other contact to equalize charge. Movement of carriers through a semiconductor is governed by the carriers' lifetime τ and mobility μ: $L_D = \sqrt{k_b T/e * \mu\tau}$. The lifetime represents the time constant of the function for the carriers' decay– in the case of excitons decay means recombination of the electron and hole. The mobility represents the speed of the carriers; in conditions where scattering is more likely, mobility is lower. Carrier mobility and lifetime give a characteristic transport length L_D for a semiconductor. When carriers are generated at some point in a sample, for example by photoexcitation, their concentration decays exponentially with distance from that point, with a decay length given by L_D.

Two properties of excitons are relevant to our research. First, because they are made up of an electron tightly bound to a positively-charged hole, excitons are charge-neutral. This is useful in detecting their presence as opposed to free carriers. Secondly, since electrons and holes are both spin-1/2 fermions, excitons are spin-1 or 0, making them bosons. Excitons therefore obey Bose-Einstein statistics, meaning that under certain conditions they can condense into the ground state. In a Bose-Einstein condensate, all bosons involved share the same ground-state wave function. This allows exotic properties quantum properties on a macroscopic scale, such as the ability to flow without friction, as seen in superfluids and superconductors. We have found that excitons created by photoexcitation in Bi_2Se_3 nanowires under certain conditions also undergo condensation, enabling extremely long photocurrent transport length.

Sample Growth and Device Fabrication

The main process used to grow samples was Chemical Vapor Deposition (CVD). A typical growth setup is shown below (Figure 3). A glass tube (22mm diameter) is placed in a tube furnace and attached to a pump at the left end and a controlled argon gas flow at the right. The growth substrate, in this case SiO_2 coated with a thin layer of gold, is placed near the left side of the tube, past the edge of the tube furnace (14cm left of center). For bismuth selenide growths, two precursors were used; each is held in a smaller glass tube (10mm diameter) placed inside of the larger tube. The first precursor was Bi_2Se_3 flakes (approx. 116mg) and antimony powder (approx. 36mg), both vacuum deposition grade (Alfa Aesar). This is placed at the center of the tube furnace. The second precursor consisted of five selenium pellets (approx. 225mg, from Johnson Matthey Inc.), placed to the right of the tube furnace (16cm right of center). The tube is first pumped with a vacuum pump and flushed with argon gas several times. Once the pressure drops to 40mTorr, the pump is turned off, the tube is again filled with argon, and the left side is opened to atmosphere. Flow from the left side passes through a bubbler, preventing air from being pulled back into the tube. A constant argon flow of 150 sccm is then maintained from the right to direct the chemical vapors, and maintain air pressure in the tube during cooling. Next, the tube furnace is heated to 660 over the course of 20 minutes, then held at this temperature for 5 hours and allowed to cool naturally. Because the furnace covers only part of the tube, a temperature gradient is created, with the temperature highest in the center and relatively constant within the furnace, but dropping sharply from the side of the furnace to the ends of the tube. As the temperature increases, the precursors either melt and evaporate, or sublime directly. The vapors generated are carrier to the left by the argon flow. As they reach the dropping temperature gradient at the left of the tube, they deposit on the tube walls and the substrate. The substrates gold coating provides nucleation sites for crystal growth. Depending on temperature and growth time, various types of crystals may be favored; the goal of CVD is usually to grow long, thin nanowires or nanoplates.

After growth, samples have to be transferred to a $Si-SiO_2$ (silicon with a 300nm silicon oxide coating) wafer with gold markings on it, which are needed for future parts of the process. This is done by pressing the wafer onto the substrate once or twice; in some cases, samples can be moved manually between the substrate and new wafer, or within wafer, using a piece of cat hair. The wafer is examined under a microscope to find good samples. Even with theoretically identical conditions, the quality of growths can vary wildly, and it is commonplace to have no useful samples on a wafer. After potentially useful samples are identified, their thicknesses can be measured using an atomic force microscope (AFM). Optical images are then taken of the wafer through a microscope. The pre-patterned gold electrodes markers on the wafer are used

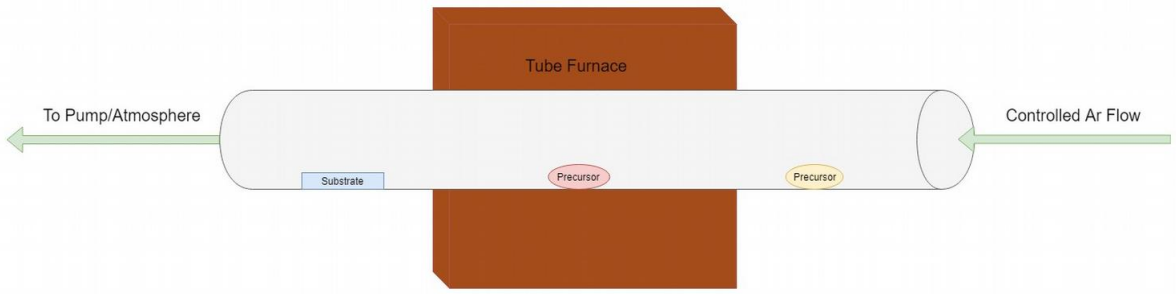

FIG. 3: Schematic of growth setup

to fix coordinates; using these coordinates, a CAD program is used to draw an electrode pattern over the images. Small ($1\mu m$ wide) electrodes are drawn across the samples of interest, and used to connect these samples to larger gold plates on the wafer. For typical samples, four electrodes is the configuration for measurement. In cases where the sample is too small or many samples need to be connected to a limited number of gold plates, two or three electrodes is enough for some measurements.

Next, the wafer is coated with polymers PMMA and MMA (polymethyl methacrylate and metheyl methacrylate). Using the electrode design from the CAD file, an electron beam lithography (EBL) machine draws patterns on the wafer. The electron beam breaks some bonds in the polymer, weakening it in regions where it is hit by the beam. When the wafer is soaked in a solution after EBL, only the weakened regions of the polymer are removed. Thin layers of chromium and gold are then deposited on the wafer, but they only stick to the silicon in areas where it is not covered by polymer. After the rest of the polymer is removed, this leaves gold electrodes, bonded to the silicon with chromium, in only the areas of the wafer targeted by the electrode design (Figure 4). Finally, with electrodes in place connecting the sample to larger gold plates, these plates can be connected to an outside circuit using temporary aluminum wires, and the sample is ready for measurement.

Measurements

For data taking, the wafer is placed on a chip carrier with conductive leads, and the sample to be measured is connected to those leads with aluminum wires. The chip carrier is housed in a cryostat, and five cables within the cryostat are plugged in to the carriers leads; the other ends of these cables are outside and can be connected to external circuits. The cryostat is sealed and pumped to vacuum for measurements. Temperature is controlled by an electric heater and a variable flow of liquid nitrogen or liquid helium. The following measurements are typically taken both at room temperature (300K) and low temperature (77K for liquid nitrogen, 4.2K for liquid helium).

The first type of measurement that is usually taken is the samples I-V curve. One electrode on the sample is connected to a variable voltage source and an electrode on another part of the sample is grounded through an ammeter. Scanning current as a function of voltage gives the resistance of the sample, presuming the I-V curve is linear. Additionally, scanning for each of the six pairs of electrodes on a four-electrode sample allows the contact resistance of the two inner electrodes to be determined, which can be taken into account as a correction on future measurements. Nonlinearity often comes from irregularity in the junction between the electrode and the sample due to residual polymers on the sample, or oxidation of the sample or electrode. This can sometimes be corrected by applying a high voltage across that junction for a short period of time. For a sample with a significantly nonlinear I-V curve, further current and voltage measurements can be difficult or inaccurate, so determining the shape of this curve is vital.

Gate voltage dependence measurements are also taken. Because the wafer is made of conductive silicon with a thin dielectric layer of silicon dioxide above it, voltage applied to the silicon layer (called the gate voltage) acts like a capacitor with the sample. Varying gate voltage changes the carrier concentration, which moves the Fermi level of the sample. Its characteristics can therefore by tuned dynamically to tend toward being n-type or p-type. Because of the antimony doping, all samples are inherently n-type. For some samples, a large enough negative gate voltage can push the carrier concentration to that of an intrinsic semiconductor, and beyond that into the p-type range. To conduct a gate scan, a small bias voltage is applied across the sample and the gate voltage is continuously scanned through up to $\pm$150V. The current through the sample is measured throughout this process, and because current depends on the carrier concentration, this can be used to characterize the sample at each gate voltage.

Treating the gate-sample system as a capacitor, the change in carrier concentration Δn with the application of a gate voltage V_g on a sample of dimensions w, h, l and capacitance C is given by:

$\frac{C*\Delta V_g}{whl} = \Delta n$

Current is a function of the carrier concentration and bias field:

$\Delta I = \Delta n \mu V_b wh/l$

This allows the mobility of free carriers in the sample to be determined from a gate scan:

$\mu = \frac{\Delta I L^2}{C V_b \Delta V_g}$

Photocurrent in the sample is measured using Scanning

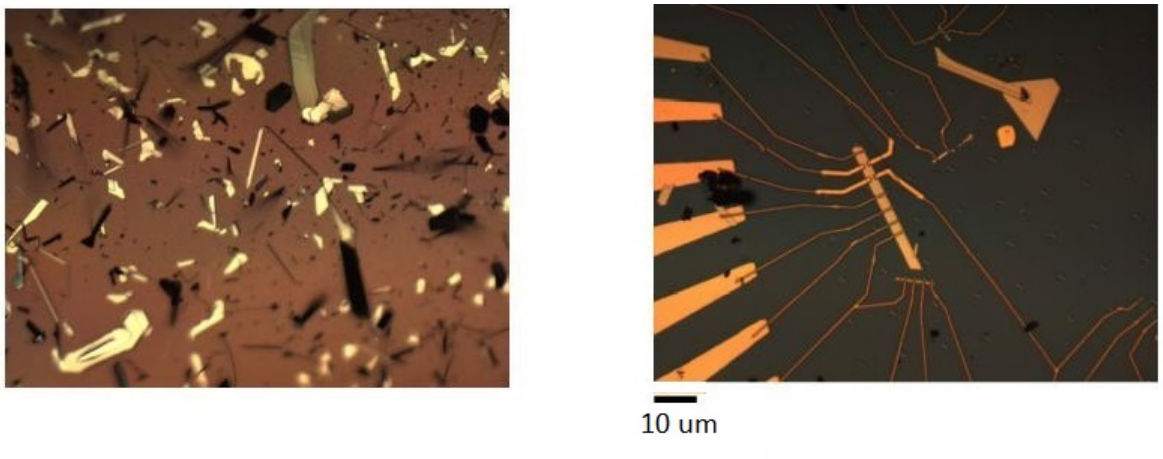

FIG. 4: Optical images of a wafer after sample transfer (left) and a finished electrode pattern (right).The sample is a rare example of one that is connected to six electrodes. Other crystals not suitable for electrode deposition and measurement appear in the upper right of the image. Scale bar: 10μm

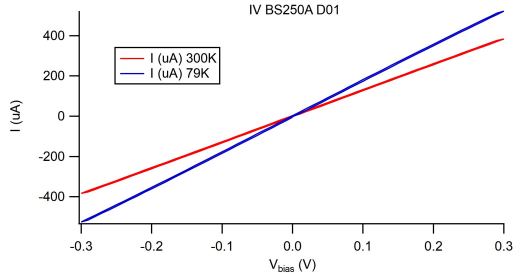

FIG. 5: Typical linear I-V curve for a Bi_2Se_3 sample at 300K and 79K. Conductivity is higher at low temperature because of reduced phonon scattering.

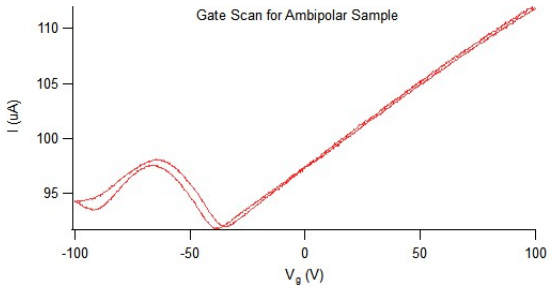

FIG. 6: Gate scan with fixed bias (1V). Current increases linearly at high gate voltage because the sample is n-type. As gate voltage decreases, current reaches its minimum at the ambipolar point, where the sample becomes an intrinsic semiconductor, then increases as it enters the p-type range. Another minimum is seen at lower gate because of differing carrier concentration at the top and bottom surfaces of the sample.

Photocurrent Microscopy (SPCM). A simple diagram of the SPCM setup is below (Figure 7). A 532nm continuous-wave laser is directed at the sample. It passes through an objective lens, usually 10x, immediately above the sample. Computer-controlled scan mirrors in the beam path allow precise control of the laser position. One electrode is grounded, another is grounded through an ammeter, and the rest are left floating. When the laser strikes the sample, a photocurrent is generated between the two grounded electrodes, which can be measured with the ammeter. This photocurrent is measured as a function of the laser position, allowing precise determination of photocurrent decay length. Because photoexcited carriers have to reach the contacts to generate a current, and the magnitude of the current increases with the number of carriers, measured photocurrent is strongest when laser excitation is near the contacts, and decays exponentially away from them. The decay length for this photocurrent is the same as the transport length of the carriers. In cases where the photocurrent is extremely weak or nonexistent, thermoelectric effects at the contacts instead dominate. The laser heats the metal contacts when it passes over them; the thermal gradient across the sample generates a voltage with the opposite polarity of the photocurrent. This effect is usually about 100 times weaker than the photocurrent and therefore negligible except when photocurrent is suppressed or nonexistent.

Carrier diffusion length is given by the SPCM profile, and mobility can be extracted from the gate dependence. In principle, if the type of carrier is known, then carrier lifetime can be determined from these quantities. However, because this research centers on determining the type and behavior of carriers, an independent measurement of lifetime is highly desired. This is accomplished by measuring the decay time of photocurrent after excitation. A 450nm pulsed laser is directed through the same optics as used for SPCM, and fixed at one point on the sample. The pulse duration is 40ns and the laser is driven at some much lower frequency, usually 1MHz. Photocurrent passes through a fast current amplifier (rise time $<$ 1ns) and the amplified signal is detected as a voltage by an os-

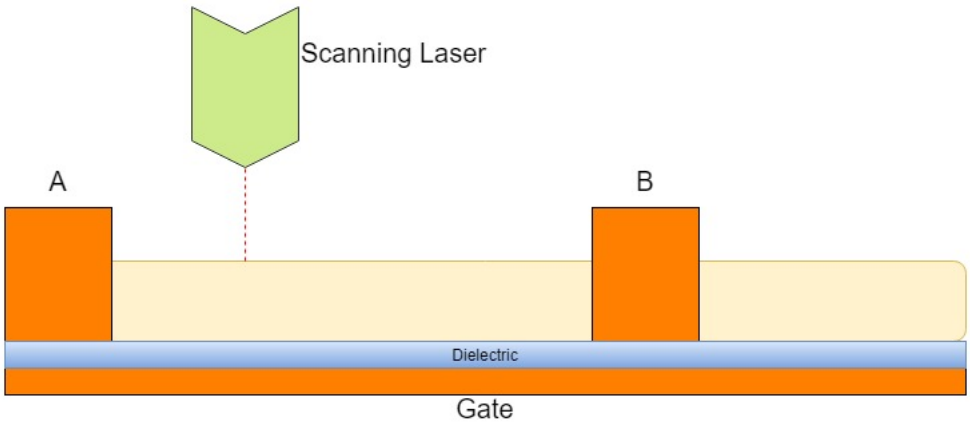

FIG. 7: Schematic of the SPCM measurement apparatus. For reference, electrode separation is typically about 300μm.

cilloscope. The observed signal peaks during laser excitation, then decays exponentially with a time constant of approximately 40ns. It is unknown whether this decay is a result of carrier lifetime or the rise time of the circuit. There is nonnegligible capacitance in the measurement circuit, coming mainly from the coaxial cables used between the cryostat and amplifier, and between the amplifier and oscilloscope. Combined with the input impedance of the amplifier, this may lead to an RC time constant of the circuit that could be on the order of 40ns. Initial decay time measurements of 1s were determined to be a result of these effects. By using a faster amplifier, minimizing the length of coaxial cable in the circuit, and adding termination resistors where possible, this was reduced to 40ns but further investigation of the measurement setup is required to determine whether this value comes from the ideal signal or the circuit. However, the measured decay time gives a 40ns upper limit on the carrier lifetime.

Results

SPCM measurements were taken at various temperatures, gate voltages, bias voltages, and laser powers (controlled with a variable filter). Photocurrent is strongest at low temperatures, and for most samples disappears by 200K. In addition to decreasing intensity, the photocurrent decay length decreases with increasing temperature. This is likely because excitons can recombine more readily at high temperatures, which reduces carrier lifetime, and because phonon scattering in the crystal is higher, which causes current to dissipate more quickly. The photocurrent profile has little dependence on gate voltage within the n-type range; however, as large negative gate voltage pushes the sample to intrinsic or p-type, the intensity and decay length of the photocurrent both decrease. Negative gate lowers the Fermi level of the sample; when the Fermi level is close enough to the valence band, holes at the top of the valence band can interact with photoexcited electrons, leading to faster recombination.

SPCM can be used with varying bias voltage across the sample, which is useful for determining the charge of photocarriers. In addition to the diffusion velocity that all carriers have, applied bias voltage gives free carriers a drift velocity, as electrons move to the higher-potential end of the sample and holes move to the lower-potential end. This should have the effect of shifting the photocurrent profile to the left or right as bias voltage is increased, depending on whether the photocurrent is generated by electrons or holes. On the other hand, if photocarriers are charge-neutral, bias voltage should have no effect on the photocurrent. As shown (Figure 8), applying bias voltages up to $\pm$200mV has no systematic effect on the photocurrent. This provides strong evidence that the charge carriers in our samples are neutral excitons.

The other key characteristic of our samples is the long photocurrent decay length, indicating long carrier transport. One typical photocurrent profile is shown below, along with a graph of intensity vs. position for one horizontal line on the profile (Figure 9). The photocurrent is constant with position for the entire length of the sample, making accurate determination of the decay length impossible. However, based on fitting with a double exponential function, a lower bound can be set at 1mm.

The decay length has to be decreased in order to be measured. This can be done by applying a sufficiently large negative gate voltage. Doing this decreases carrier lifetime or mobility, which decreases the diffusion length. Diffusion length can then be determined from SPCM at low mobility, and extrapolated to find values at higher mobilities.

For other bismuth selenide studies, mobility values have been measured at up to $2m^2V^-1s^-1$ [2]. Even using this as an upper bound for μ, along with a lifetime upper bound of $\tau = 1\mu s$, the decay length would be $L_D < 13\mu m$. Our extreme lower bound on diffusion length is nearly 100 times higher, at $L_D = 1mm$, and the real value may be much higher still. Our measurements therefore cannot be explained by previously-investigated exciton transport mechanisms in TIs. Gate dependence measurements yield a free carrier mobility of $0.1m^2V^-1s^-1$, meaning the measurements can also not be explained by free carrier transport.

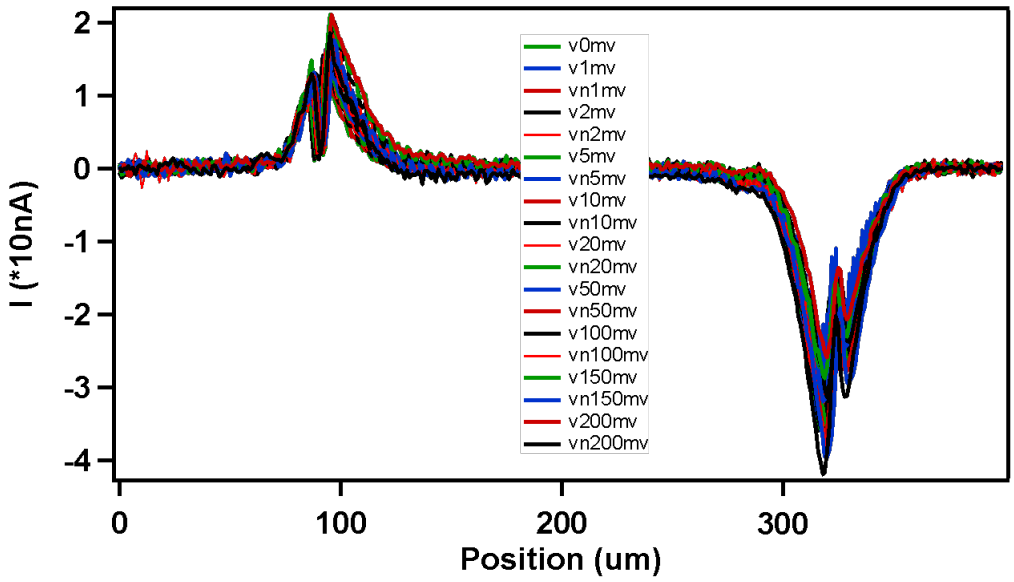

FIG. 8: Horizontal slice of photocurrent profile for a wide range of bias voltages. A gate voltage was applied to reduce the photocurrent from a uniform line to a clear exponential to ensure that bias effects can be seen if they exist.

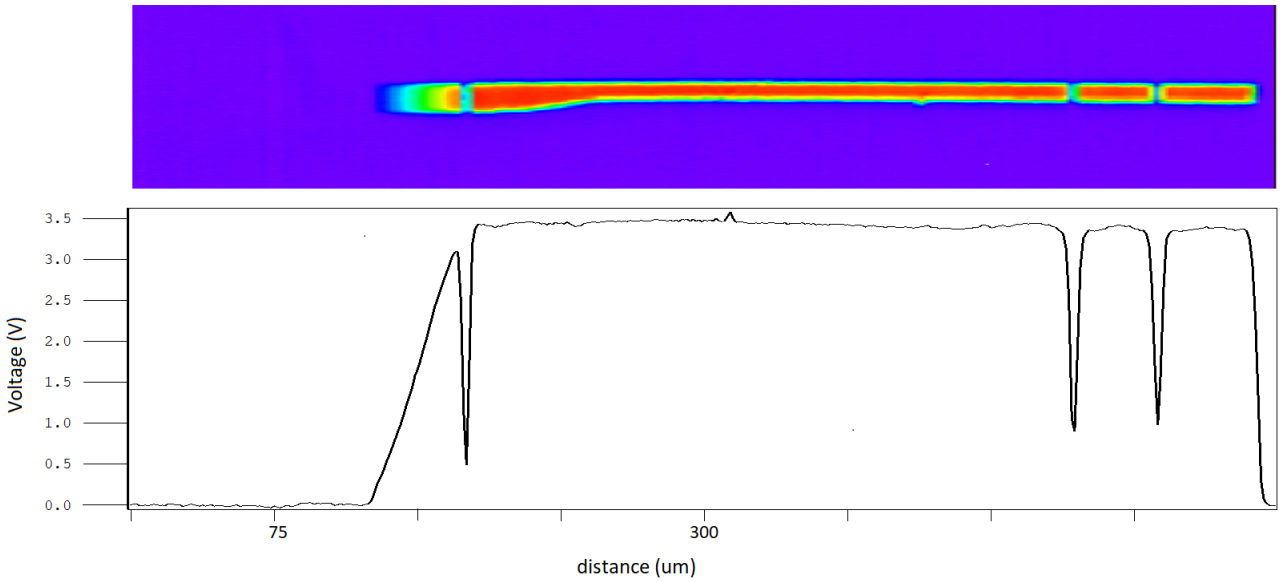

FIG. 9: 2-D image of photocurrent as a function of laser position (top). Red indicates higher current. Photovoltage vs. horizontal position for one slice of the profile (bottom). Signal is plotted as voltage because the ammeter outputs a voltage signal, which is processed by the SPCM program. For these data, photocurrent was measured between the two rightmost electrodes. An electrode is apparent on the far left of the sample. Other than locally blocking the laser signal, it was left floating and had no influence on the photocurrent.

These measurements are best explained instead by a quantum phase that flows analogously to a superfluid and allows highly dissipationless transport– an exciton condensate. Evidence of macroscopic order in exciton quantum states has previously been observed in quantum wells [3], and in bilayer systems[4]. However, these experiments were conducted on smaller-scale systems, and observed exciton behavior using different methods. The fact that experiments on Bi_2Se_3 have not previously uncovered similar evidence of exciton condensation is likely explained by the doping of our samples. Long decay length is only observed in Sb-doped samples. We believe that this doping moves the Fermi level near the Dirac point, supported by the fact that a sufficiently large gate voltage can reduce decay length by moving the Fermi level away

from the Dirac point.

Conclusion

We have observed extremely long photocurrent decay length in antimony-doped bismuth selenide nanowires. This suggests highly dissipationless transport of excitons, providing some of the first experimental evidence of Bose-Einstein condensation in macroscopic exciton systems. Longer samples will need to be grown and studied to pin down the decay length to higher values with less error. Better control of the doping level has the potential to allow further tuning of T_c for condensation in samples and more extensive study of the properties of exciton condensates. However, even the current data hint at exciting applications of this phenomenon. Long dissipationless transport may be useful in various spintronic devices. Additionally, because of excitons' low mass, T_c for exciton condensation has been observed to be relatively high– probably above liquid nitrogen temperature, and certainly beyond the millikelvin range. The observation and potential manipulation of quantum coherence at these temperatures may find applications in high-temperature quantum computing.

[1] M.Z. Hasan and C.L. Kane, "*Colloquium:* Topological Insulators," *Review of Modern Physics* **82**, 3045 (2010).

[2] N.P. Butch et al., "Strong surface scattering in ultrahigh-mobility Bi_2Se_3 topological insulator crystals," *Physical Review B* **81**, 24130 (2010).

[3] L.V. Busov, A.C. Gossard, and D.S. Chemia, "Macroscopically ordered state in an exciton system," *Nature* **418**, 6899 (2002).

[4] J.P Epstein and A.H. MacDonald, "Bose-Einstein Condensation of Excitons in Bilayer Electron Systems," *Nature* **418**, 6899 (2002).

[5] Band gap diagrams taken and modified from $https : //docs.quantumwise.com/tutorials/topological_insulator_bi2se3/t$ Public domain.