JVST B
Journal of Vacuum Science & Technology B

ARTICLE

avs.scitation.org/journal/jvb

Gap distance dependence on field emission at the nanogap between silicon cleavage surfaces

Cite as: J. Vac. Sci. Technol. B **41**, 022805 (2023); doi: 10.1116/6.0002456
Submitted: 26 December 2022 · Accepted: 16 February 2023 ·
Published Online: 16 March 2023

Yuki Akura,[1,a)] Masaki Shimofuri,[1] Amit Banerjee,[1] Jun Hirotani,[1,2] and Toshiyuki Tsuchiya[1]

AFFILIATIONS

[1]Department of Micro Engineering, Graduate School of Engineering, Kyoto University, Kyotodaigaku-Katsura C3, Nishikyo-ku, Kyoto 615-8540, Japan
[2]PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

[a)]**Author to whom correspondence should be addressed:** y_akura@nms.me.kyoto-u.ac.jp

ABSTRACT

In this study, we developed a model of field emission at a vacuum nanogap composed of semiconductor electrodes and conducted numerical calculations to evaluate its field emission characteristics. The current at the silicon nanogap was also measured, and the results were compared with the calculated results. Large-area vacuum nanogaps are expected to be used for highly efficient room-temperature thermionic power generation owing to their good thermal insulation, and their electrical conduction of tunneling current is sufficient for practical use. Because vacuum nanogaps should be parallel and smooth with a large area, we proposed and examined the use of cleavage to fabricate nanogaps in single-crystal materials. A nanogap with a nominal area of 26.5 μm^2 was fabricated in a microelectromechanical device to evaluate its properties, and the model was used for the analysis. The calculated current–voltage characteristics of the semiconductor nanogap of both n-type and p-type silicon were affected by the bandgap and the presence of carriers, i.e., electrons and holes. Moreover, the Fowler–Nordheim plot was confirmed to be valid for semiconductor nanogaps based on the calculated results. We proposed an effective work function that compensates for the discrepancy between semiconductor nanogaps and the Fowler–Nordheim theory and enables a more accurate analysis. Field-emission currents flowing through a 1.1 nm nanogap were measured and matched with the calculated results by adjusting the field enhancement factor and emission area. The model developed and confirmed experimentally in this work will be helpful in examining field emissions at semiconductor nanogaps.

I. INTRODUCTION

In recent years, Internet of Things (IoT) sensors have increasingly been used worldwide, and a major problem limiting their application is the power source for sensors. Considering that many IoT devices are being installed, their power sources should be maintenance-free and standalone. Thermoelectric power generation, which generates electricity from waste heat in the environment, is a strong candidate that can be used to address this problem and has been implemented.[1]

The conversion efficiency η of thermoelectric power generation is expressed as[1]

$$\eta = \left(1 - \frac{T_{\mathrm{cold}}}{T_{\mathrm{hot}}}\right) \frac{\sqrt{1+ZT}-1}{\sqrt{1+ZT}+\frac{T_{\mathrm{cold}}}{T_{\mathrm{hot}}}}, \tag{1}$$

where T_{cold} and T_{hot} are the temperatures of cold and hot reservoirs, respectively, and ZT is a dimensionless thermoelectric figure of merit. ZT is determined using the Seebeck coefficient S, electrical conductivity σ, operating temperature $T = (T_{\mathrm{cold}} + T_{\mathrm{hot}})/2$, and thermal conductivity κ, and is expressed as[1]

$$ZT = \frac{S^2 \sigma T}{\kappa}. \tag{2}$$

A higher ZT results in higher conversion efficiency. ZT has a benchmark value of 1 for practical use.[2] To ensure highly efficient thermoelectric power generation, materials with low thermal conductivity and high electrical conductivity are required. Several studies have attempted to realize such materials in various fields,[1–4] including research on inorganic bulk materials,[5,6] inorganic thin film,[7,8] and organic materials.[9] However, solid materials have

limitations in terms of reducing thermal conductivity owing to heat conduction by lattice vibrations and electron/hole conduction. In other words, a drastic improvement in conversion efficiency is difficult using solid materials.

At a vacuum nanometer-scale gap (nanogap), phonon conduction does not occur, but electrical conduction occurs because of quantum effects; it can be regarded as a material with very low thermal conductivity and sufficiently high electrical conductivity. In addition, a nanogap has a large Seebeck coefficient compared to other thermoelectric materials.[10] Therefore, thermionic emission in a vacuum nanogap can be used to achieve higher efficiency.

Fabrication methods for nanogaps have been investigated to realize nanometer-sized electronic devices,[11] biosensors,[12] etc. Various methods were proposed, including mechanically controllable break junctions,[13] electrochemical deposition,[14] electromigration,[15] and focused ion beam.[16] The nanogap should be parallel and smooth over a large area to achieve low-temperature thermionic power generation. However, nanogaps fabricated using above-mentioned methods are not parallel and smooth, and the area is very small (order of nm^2). We proposed the use of cleavage in single-crystal materials to fabricate a conformal nanogap, which has been implemented using single-crystal silicon. In our previous study,[17,18] a microelectromechanical system (MEMS) device was developed in which fabrication and evaluation of the Si nanogap (cleaving, controlling the gap distance, applying the voltage, and measuring the current) can be performed in the sample chamber of a scanning electron microscope (SEM). The mechanical characteristics of the device were discussed.[18] Field emission characteristics and their gap distance dependence must be evaluated experimentally for the use of nanogaps in thermionic power generation, and a numerical model must be developed for theoretical discussion of measurement data beforehand. While numerical calculation for the theoretical analysis of field emission at a nanogap composed of metal electrodes has been reported,[10] a numerical model at a semiconductor nanogap has not yet been thoroughly examined.

In addition, an analysis method must be discussed. Fowler–Nordheim (FN) plots are generally used in the analysis of field-emission currents.[19–21] Therefore, it could also be useful to analyze the emission current at semiconductor nanogaps. However, an FN plot is based on the approximated theory of field emission from a metal cathode; therefore, it is necessary to verify the validity of applying it to semiconductor nanogaps before analyzing the measurement results obtained using this method.

In this study, we developed a model of field emission at vacuum nanogaps composed of both p-type and n-type semiconductor electrodes and conducted numerical calculations to evaluate their characteristics. The validity of applying an FN plot to semiconductor nanogaps was evaluated using calculation results. Finally, the field-emission current at the silicon nanogap was measured, and the results were compared with those obtained from the calculation.

Section II describes a model of field emission at a metal nanogap that served as the foundation for the developed model, followed by a description of the developed model at a semiconductor nanogap. Section II also includes the theoretical basis for an FN plot. In Sec. III, details of calculations and an experimental procedure are presented. The result of calculations and experiments are shown in Sec. IV. In Sec. V, the calculated results, the validity of an FN plot, and the measurement results are discussed.

II. THEORY

A. Field emission at a metal nanogap

In general, the current density J of field-emitted electrons from a metal electrode is expressed as

$$J = e\int_0^\infty N(E_x)\cdot D(E_x)\,dE_x, \tag{3}$$

where $N(E_x)dE_x$ is the number of electrons incident on a unit area and unit time with the kinetic energy in the x-direction in the range from E_x to $E_x + dE_x$, which is known as the supply function, $D(E_x)$ is the probability that electrons whose kinetic energy in the x-direction (E_x) can penetrate a potential barrier, and e is the electron charge. In this study, the x-direction is perpendicular to the emission surface of the electrode. A schematic diagram of the current is shown in Fig. 1. Electrons tunnel through the potential barrier, as indicated by the blue arrow. $N(E_x)D(E_x)dE_x$ out of $N(E_x)dE_x$ electrons emitted from electrode 1 reach electrode 2.

$N(E_x)$ and $D(E_x)$ are expressed as given in Eqs. (4) and (5), respectively.[10] $N(E_x)$ is calculated by the free-electron theory of metal and $D(E_x)$ is calculated using the Wentzel–Kramers–Brillouin–Jeffreys method,

$$N(E_x) = \frac{mk_BT}{2\pi^2\hbar^3}\ln\left\{1+\exp\left(-\frac{E_x}{k_BT}\right)\right\}, \tag{4}$$

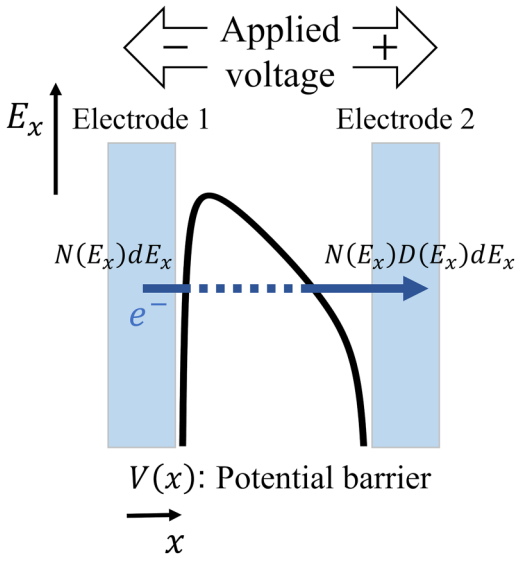

FIG. 1. Schematic of electron transfer at a nanogap. The potential barrier for electrons in electrode 1 is indicated by a black line. The number of electrons that pass through the nanogap with the x-direction kinetic energy E_x in unit time and the unit area is the product of N, D, and dE_x.

$$D(E_x) = \exp\left[-\frac{2}{\hbar}\int_{x_1}^{x_2}\sqrt{2m(V(x)-E_x)}\,dx\right], \tag{5}$$

where m is the effective mass of the electron, k_B is the Boltzmann constant, T is the electrode temperature, $\hbar$ is Dirac's constant, and $V(x)$ is the electron potential at x; x_1 and x_2 are the roots of $V(x) - E_x = 0$ ($x_1 < x_2$).

The potential of the electrons at the nanogap whose gap distance is d_{gap} and where the bias voltage V_{bias} is applied is expressed as given in Eq. (6).[10] The third term in Eq. (6) gives the image charge effect for both electrodes,

$$V(x) = \Phi - \frac{eV_{bias}}{d_{gap}}\cdot x - \frac{e^2}{4\pi\epsilon_0}\left[\frac{1}{4x} + \frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{nd_{gap}}{n^2d_{gap}^2 - x^2} - \frac{1}{nd_{gap}}\right)\right], \tag{6}$$

where Φ is the work function of the electrode and ϵ_0 is the electric constant. In this work, the effects of space charge and exchange-correlation, which were considered by complex quantum calculation in some works in field emission at metal–insulator–metal junctions[22,23] and a metal nanogap,[24] were ignored for simplicity of the calculation.

By substituting Eqs. (4)–(6) into Eq. (3), the current from one side of the metal nanogap electrodes can be calculated. The same calculation was performed on the opposite side. The current at the nanogap is a superposition of these currents.

B. Field emission at a semiconductor nanogap

In semiconductor electrodes, there is a bandgap and two types of carriers: electrons in the conduction band and holes in the valence band. Therefore, the field-emission model for semiconductor electrodes should be modified.

At the semiconductor nanogap, similarly to Eq. (4), the supply functions of electrons in the conduction band $N_{electron}(E_x)$ and those of holes in the valence band $N_{hole}(E_x)$ are expressed as

$$N_{electron}(E_x) = \frac{m_{dse}k_BT}{2\pi^2\hbar^3}\ln\left\{1 + \exp\left(-\frac{E_x - E_F}{k_BT}\right)\right\}, \tag{7}$$

$$N_{hole}(E_x) = \frac{m_{dsh}k_BT}{2\pi^2\hbar^3}\ln\left\{1 + \exp\left(-\frac{E_F - E_x}{k_BT}\right)\right\}, \tag{8}$$

where m_{dse} and m_{dsh} are the density-of-states effective mass of electrons and holes, respectively, and E_F is the Fermi level. The Fermi level for semiconductor electrodes is obtained by solving the following two equations derived from the charge neutrality and mass-action law:[25]

$$N_C\exp\left(-\frac{E_C - E_F}{k_BT}\right) = \frac{n_i^2}{N_C\exp\left(-\frac{E_C - E_F}{k_BT}\right)} + \frac{N_D}{1 + 2\exp\left(\frac{E_F - E_D}{k_BT}\right)}, \tag{9}$$

$$N_V\exp\left(-\frac{E_F - E_V}{k_BT}\right) = \frac{n_i^2}{N_V\exp\left(-\frac{E_F - E_V}{k_BT}\right)} + \frac{N_A}{1 + 4\exp\left(\frac{E_A - E_F}{k_BT}\right)}, \tag{10}$$

where N_C and N_V are the effective densities of states on the conduction band and valence band, respectively; n_i is the intrinsic carrier concentration; E_C and E_V are the energy levels at the bottom of the conduction band and top of the valence band, respectively; N_D and N_A are the donor and acceptor densities, respectively; and E_D and E_A are the donor and acceptor levels, respectively. For silicon doped with phosphorus as the donor, $E_D = E_C - 0.046$ (eV), whereas for silicon doped with boron as the acceptor, $E_A = E_V + 0.044$ (eV).[25] N_C, N_V, and n_i in silicon are expressed in the following equations.[25,26] The expression of n_i was obtained by fitting the experimental data,[26,27]

$$N_C = 2\left(\frac{m_{dse}k_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}}, \tag{11}$$

$$N_V = 2\left(\frac{m_{dsh}k_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}}, \tag{12}$$

$$n_i = 3.87\times10^{16}\times T^{\frac{3}{2}}\times\exp\left(\frac{-0.605\times e}{k_BT}\right)\ [\mathrm{cm}^{-3}]. \tag{13}$$

E_C and E_V are related by the expression $E_C - E_V = E_g$, where E_g is the bandgap. E_g in silicon is expressed as Eq. (14).[28,29] It was obtained by fitting the experimental data with the relation for the temperature dependence of the bandgap proposed by Varshni,[30]

$$E_g = 1.155 - 4.73\times10^{-4}\times\frac{T^2}{635 + T}\ [\mathrm{eV}]. \tag{14}$$

$D(E_x)$ of electrons and holes in the semiconductor is obtained by replacing m in Eq. (5) with m_{dse} and m_{dsh}, respectively. $V(x)$ for the carriers of a semiconductor is also obtained using Eq. (6). The sign of V_{bias} should be changed when $V(x)$ is calculated for holes.

The current emitted from the electrode of the semiconductor nanogap can be obtained by integrating Eq. (3) with Eqs. (7)–(14). When performing this integration, the effect of the presence of the bandgap should be considered: when the carriers are electrons in the conduction band, the interval of integration should be replaced by $E_C < E_x < \infty$, whereas when the carriers are holes in the valence band, the interval of integration should be replaced by $-\infty < E_x < E_V$. The current density at the semiconductor nanogap J_{gap} flowing from electrode 1 to electrode 2 is calculated by subtracting the respective current densities generated by electrons and holes at each electrode. The subtraction is expressed as

$$J_{gap} = (J_{hole,\ 1} - J_{electron,\ 1}) - (J_{hole,\ 2} - J_{electron,\ 2}), \tag{15}$$

where $J_{hole,1}$ and $J_{hole,2}$ are the current densities generated by holes emitted by electrodes 1 and 2, respectively, and $J_{electron,1}$ and

$J_{electron,2}$ are those generated by electrons emitted by electrodes 1 and 2, respectively.

C. Fowler–Nordheim theory

The Fowler–Nordheim theory is an approximated theory of field emission from a metal electrode. In this theory, the following assumptions were made: the electrode temperature is 0 K, the third term of Eq. (6) can be ignored, and electrons near the Fermi level dominate the emission current. By applying these assumptions to Eqs. (3)–(5), the current density J_{FN} can be expressed in the following equations:[31]

$$J_{FN} = \frac{a_{FN}F^2}{\Phi}\exp\left(-\frac{b_{FN}\Phi^{\frac{3}{2}}}{F}\right), \tag{16}$$

$$a_{FN} = \frac{e^3}{16\pi^2\hbar} \approx 2.4696 \times 10^{-25}\ [\mathrm{A\,J\,V^{-2}}], \tag{17}$$

$$b_{FN} = \frac{4\sqrt{2m}}{3e\hbar} \approx 1.0652 \times 10^{38}\ [\mathrm{J^{-3/2}\,V\ m^{-1}}], \tag{18}$$

where F is the electric field calculated as $F = V_{bias}/d_{gap}$.

To apply this theory to evaluate the measured emission current, correction factors should be added. A field enhancement factor (β) is introduced to express the local enhancement of the electric field, where F is replaced by βF. The measured current is the product of the current density and emission area A. Therefore, the current I_{FN} can be expressed as given in the following equation:

$$I_{FN} = \frac{a_{FN}A\beta^2V_{bias}^2}{\Phi d_{gap}^2}\exp\left(-\frac{b_{FN}\Phi^{\frac{3}{2}}d_{gap}}{\beta V_{bias}}\right). \tag{19}$$

Dividing by V_{bias}^2 and taking the logarithm, we obtain

$$\ln\left(\frac{I_{FN}}{V_{bias}^2}\right) = -\frac{b_{FN}\Phi^{\frac{3}{2}}d_{gap}}{\beta}\cdot\frac{1}{V_{bias}} + \ln\left(\frac{a_{FN}A\beta^2}{\Phi d_{gap}^2}\right). \tag{20}$$

The plot of $\ln(I_{FN}/V_{bias}^2)$ against $1/V_{bias}$ yields a linear curve. This plot is known as an FN plot. Because the plot obtained from the measurement result is a straight line, the measured current can be regarded as the emission current. The field enhancement factor and emission area were estimated by fitting the result to a linear curve.[21]

III. METHODS

A. Calculation

The calculations were performed for metal, p-Si, and n-Si electrodes. The work function of the electrode was 4.5 eV, assuming a silicon electrode. The electrode temperature was set to 300 K. The donor and acceptor densities were $N_D = 1.0 \times 10^{18}\ \mathrm{cm^{-3}}$ and $N_A = 0.0\ \mathrm{cm^{-3}}$ for n-type phosphorus-doped silicon and $N_D = 0.0\ \mathrm{cm^{-3}}$ and $N_A = 1.0 \times 10^{18}\ \mathrm{cm^{-3}}$ for p-type boron-doped silicon, respectively. The p-Si values were consistent with those of the device used in the measurement in this study.

B. Experiment

We conducted current measurements to experimentally characterize the field emission at a semiconductor nanogap. The MEMS device shown in Fig. 2 was used in the experiment. The design and fabrication process of the device are described in detail in our previous paper.[18] The device was made from a silicon-on-insulator wafer. The shuttle was suspended by supporting beams. The pin was inserted into the penetration hole and actuated in the direction indicated by the arrow in Fig. 2 using a piezo actuator. The hook made contact by the lateral motion of the pin, and tensile force was applied to the short beam tapered at the middle. As a result, cleavage occurred at the beam, and the nanogap was fabricated. The cleaved surface was the {111} surface of the single-crystal silicon. After cleavage, the shuttle became movable. The nominal size of this nanogap was 26.5 μm^2 (horizontal: 5.3 μm, vertical: 5.0 μm). The device layer of the wafer was doped with boron, and the dopant concentration was $1.0 \times 10^{18}\ \mathrm{cm^{-3}}$.

The measurement circuit is shown in Fig. 3. A bias voltage V_{bias} was applied between the gaps, and the current at the gap was measured. The bias voltage was applied, and the current was measured using a source meter (Keithley 2450, USA). In this experiment, the gap distance was controlled by the pin motion.

The measurement procedure is described as follows. The device was placed in the SEM instrument (JSM-6390, JEOL, Japan). The following steps were performed in the SEM sample chamber.

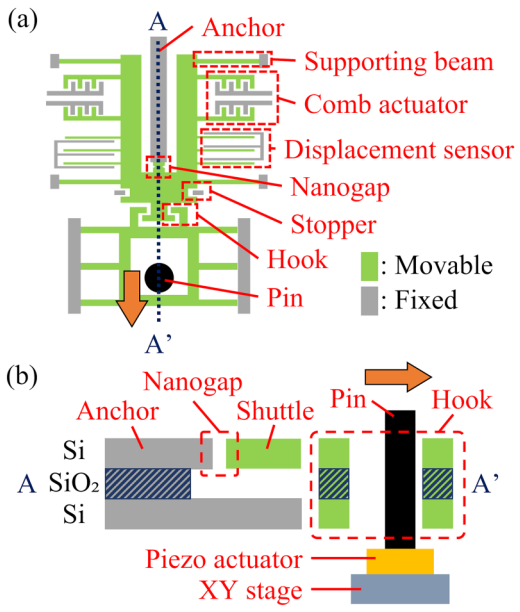

FIG. 2. Schematic of the MEMS device: (a) top view; (b) A–A′ cross-sectional view.

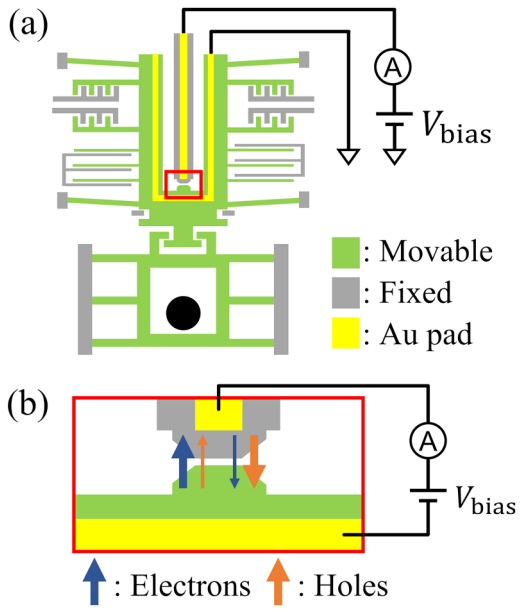

FIG. 3. Circuit used for the current measurement: (a) overall view; (b) enlarged view of the nanogap area.

The pin was actuated using a piezo-actuator, and a nanogap was fabricated. Subsequently, the gap distance was controlled by changing the voltage applied to the piezo actuator V_{pin}. In the experiment, the actuation voltage was decreased from 40.0 to 34.0 V in steps of 0.25 V, thereby decreasing the gap distance. The gap distance was sufficiently large when the actuation voltage was 40.0 V. This was confirmed by current measurement and SEM observation. No significant current was measured when the bias voltage was applied at the actuation voltage of 40.0 V. SEM observation showed the gap distance to be approximately 1 μm. The bias voltage was increased from 0.0 to 10.0 V in steps of 0.2 V, and the current was measured at each bias voltage. The bias voltage sweep was repeated five times at each actuation voltage. The pin was actuated, and the gap distance was made sufficiently large at every interval between bias voltage sweeps to eliminate the effect of previous measurement.

IV. RESULTS

A. Calculated I–V characteristics

The calculated current densities of the metal, p-Si, and n-Si nanogaps with different gap distances are shown in Fig. 4. The current densities depend on the gap distance and electrode material. The figure shows a strong dependence on the gap distance. This is because the gap distance determines the potential barrier, which has an exponential effect on transfer probability, as expressed in Eq. (5). There are large differences in the shape of the curves for the 1, 5 and 10, and 100 nm gap distances. This difference is discussed in Sec. V A with additional calculations. The dependence on the electrode material differs according to the

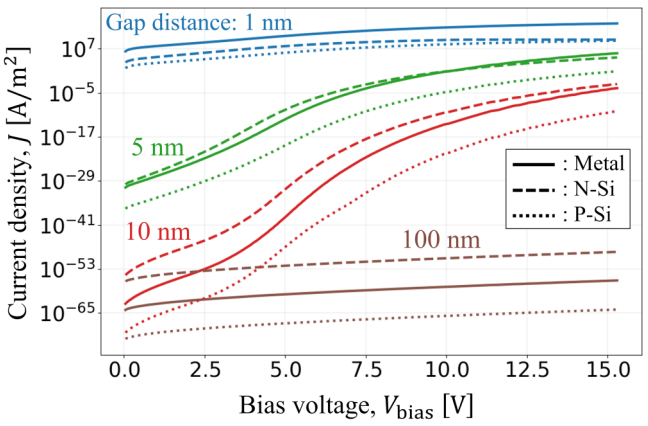

FIG. 4. Calculated current densities at nanogaps as a function of bias voltage.

strength of the electric field, which is the quotient of the bias voltage divided by the gap distance. When the field is low, the currents at the n-Si, metal, and p-Si nanogaps are high in that order because of the difference in Fermi levels of electrodes. When the field is high, the current density at the metal nanogap is higher than that at the n-Si nanogap because of the absence of carriers in the bandgap of the n-Si electrode. When the field is high, the carriers near the Fermi level have a high transfer probability, but they do not exist in the n-Si electrode.

B. Measured I–V characteristics

The measured currents are shown in Fig. 5. The voltage–current plots show significant differences, with the piezo-actuator voltage of 34.75 V as the boundary. No detectable current was measured when the voltage exceeded the threshold because the gap distance was too large. The current was suddenly measured in the fifth voltage sweep at 34.75 V. Pull-in appeared to occur when the detectable current flowed. The nonlinear I–V curves were measured when the actuation voltage was lower than the threshold, and they were not affected by the actuation voltage. This is because the partial contact of the gap derived from the pull-in resulted in a constant gap distance in these measurements.

The currents measured in the fifth sweep at the threshold voltage were smaller than those measured at a lower actuation voltage. At a nanogap, an electrostatic attraction force F_g acts between the gaps, as expressed in Eq. (21),

$$F_g = \frac{1}{2} \cdot \frac{\epsilon_0 A}{d_{gap}^2} V_{bias}^2. \tag{21}$$

In the first pull-in in the fifth sweep at 34.75 V, a large attractive force was exerted on nanogap electrodes, which can cause plastic deformation on surface asperities. This explains why the measured I–V curve at the first pull-in was different from other curves obtained thereafter. The surface became flat, and a larger current flowed.

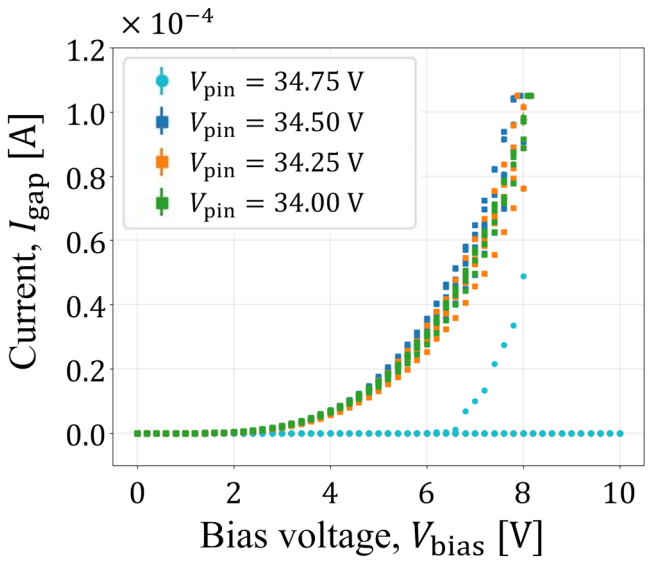

FIG. 5. Measured I–V characteristics. Significant current values were not measured for measurements with an actuation voltage higher than 34.75 V, and they are not shown.

V. DISCUSSION

A. I-V characteristics

The supply function $N(E_x)$, transfer probability $D(E_x)$, and the product of the electron charge, $N(E_x)$ and $D(E_x)$ of 5 and 10 nm metal nanogaps, are shown in Fig. 6. The data shown in Fig. 6 were obtained by calculation conducted for electron emission from one electrode only. The supply functions calculated using Eq. (4) do not depend on the bias voltage and gap distance. As the energy of the electron increases, the supply function decreases linearly when the energy is lower than the Fermi level and decreases exponentially when it is higher.

Figures 6(c)–6(f) show the transfer probabilities and products for bias voltages in the range of 1–15 V at steps of 1 V. The products are the integrands of Eq. (3) and represent the current distributions for the energy of an electron. Figures 6(e) and 6(f) show that the current distributions have two peaks: the Fermi level and the top of the potential barrier. These two peaks arise from behaviors of the supply functions and transfer probabilities. The behavior of the supply functions changes when the energy of an electron approaches the Fermi level. The transfer probabilities have a value of one when the energy of an electron is higher than the top of the potential barrier.

The two peaks in Figs. 6(e) and 6(f) indicate that the electrons of two energy levels dominate the current. When the electric field

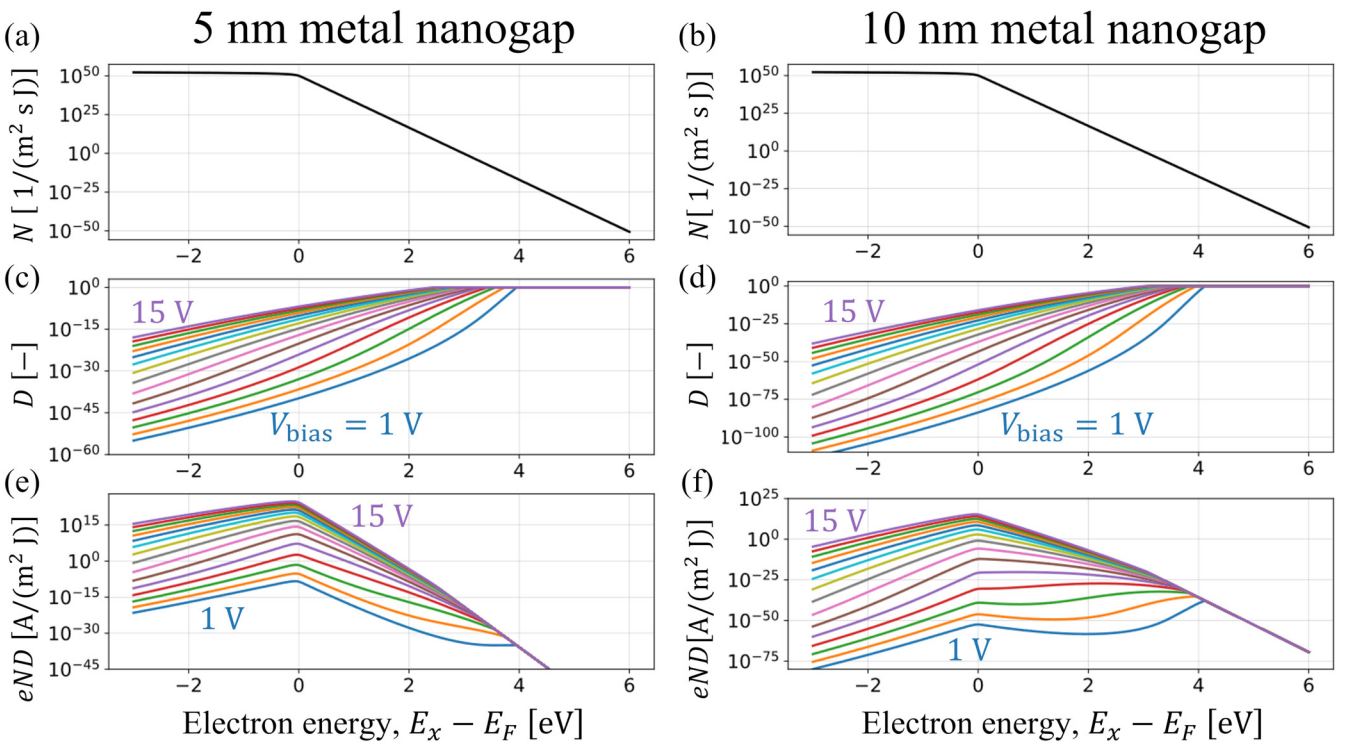

FIG. 6. Supply function $N(E_x)$ at (a) 5 and (b) 10 nm metal nanogaps. Transfer probability $D(E_x)$ for bias voltages in the range of 1–15 V at (c) 5 and (d) 10 nm metal nanogaps. Integrands of Eq. (3) at (e) 5 and (f) 10 nm metal nanogaps. The horizontal axes of all the graphs show the energy of an electron minus the Fermi energy.

is high, electrons near the Fermi level are dominant, whereas when the electric field is low, electrons near the top of the potential barrier are dominant. When the field is high, the transfer probability of the electrons near the Fermi level is sufficiently high such that they are dominant owing to their large population. However, when the field is low, the transfer probability is low, and electrons near the top of the potential barrier are dominant because of the relatively high transfer probability. This analysis was consistent with that of electrode materials described in Sec. IV A.

As shown in Fig. 4, the increase in the current density at 5 and 10 nm nanogaps is higher at bias voltages of approximately 3 and 5 V, respectively, than those at other bias voltages. The transfer probability of electrons near the Fermi level at 5 and 10 nm nanogaps increases rapidly at bias voltages of approximately 3 and 5 V, respectively, as illustrated in Figs. 6(c) and 6(d). These rapid increases in the transfer probability resulted in a rapid increase in the current generated by electrons near the Fermi level. The rapid increases in the current densities in Fig. 4 can be explained by the behavior of electrons near the Fermi level; they indicate the transition of dominant carriers from the carriers near the top of the potential barrier to those near the Fermi level. Thus, transfer probabilities of the carriers near the Fermi level have a large effect on the I–V characteristics of the nanogap, and they are determined by the shape of the potential barrier as shown in Eq. (5). When the gap distance is small, in addition to the image charge effect lowering the potential barrier, the electric field increases rapidly as the applied voltage increases. Therefore, at the 1 nm gap distance, the carriers near the Fermi level have a sufficiently high transfer probability even at low bias voltages; accordingly, the current density shown in Fig. 4 does not have a rapid increase region. As the gap distance increases, the image charge effect decreases. Additionally, a larger gap distance suppresses the growth of the electric field with increasing bias voltage. Because of these two reasons, the bias voltage required for the carriers near the Fermi level to contribute to the current is increased when the gap distance is large. Thus, the bias voltage at the rapid increase region of the 10 nm nanogap is larger than that of the 5 nm one. The current density of the 100 nm nanogap increased slowly because a bias voltage higher than 15 V is required to enable these carriers to transfer over the gap. To obtain a high current density, the carriers near the Fermi level should be allowed to pass through the nanogap.

The current densities generated by electrons and holes at the semiconductor nanogap are shown separately in Fig. 7. The current density generated by holes is $J_{\mathrm{hole},1}-J_{\mathrm{hole},2}$ and that generated by electrons is $J_{\mathrm{electron},2}-J_{\mathrm{electron},1}$. At the n-Si nanogap, the current generated by electrons is always larger than that generated by holes. At the p-Si nanogap, the dominant carrier differs with the strength of the electric field: holes are dominant when the field is high, and electrons are dominant when the field is low. In our calculation model, the transfer probabilities of electrons are higher than those of holes because of their higher energy. Because the number of electrons at the n-Si nanogap is greater than that of holes and electrons have higher transfer probabilities, the electrons dominate the current. At the p-Si nanogap, the number of holes is greater than that of electrons; however, the transfer probabilities of holes are lower than those of electrons. Therefore, when the electric field is low, the transfer probabilities of holes are very low, and the electron current is dominant. The majority carriers as well as minority carriers should be considered in the analysis of currents at p-Si nanogaps.

B. FN plots

Figures 8(a)–8(c) show the FN plots obtained from the calculated currents at each nanogap. The plots are straight lines when the electric field is high, which agrees with the FN theory. As discussed in Sec. IV A, when the electric field is high, electrons near the Fermi level dominate the currents, which is consistent with

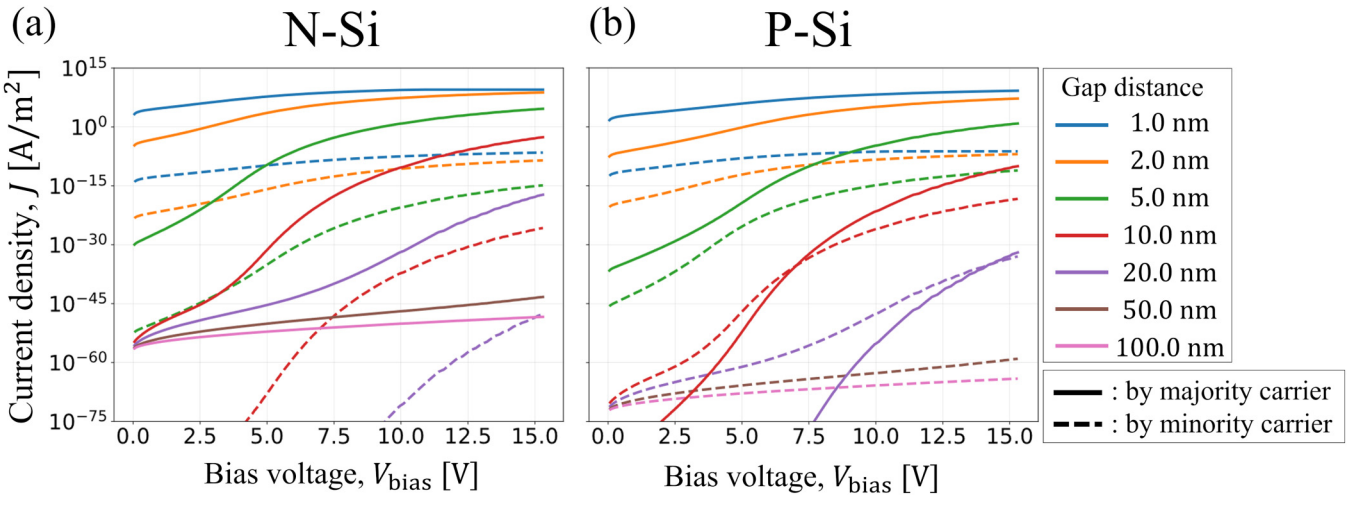

FIG. 7. Calculated current densities by each carrier at (a) n-Si nanogap and (b) p-Si nanogap. The solid line indicates the current by majority carriers (electrons in n-Si, holes in p-Si), and the dashed line indicates that by the minority carriers (holes in n-Si, electrons in p-Si).

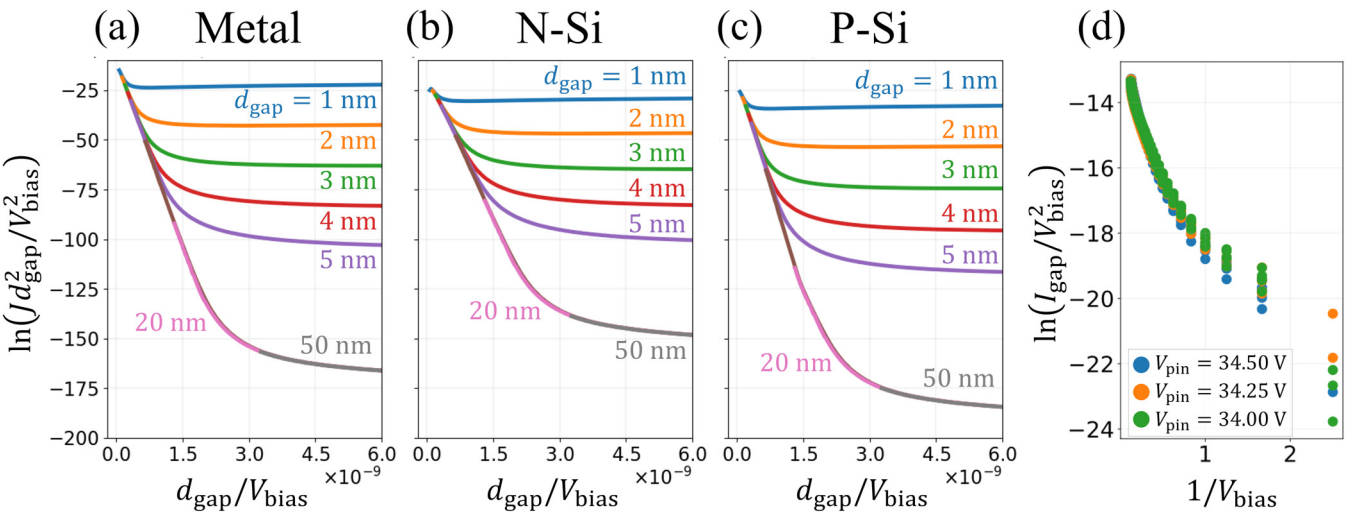

FIG. 8. FN plots obtained from calculated results on (a) metal, (b) n-Si, and (c) p-Si nanogap. (d) FN plots obtained from measurement results.

the third assumption of the FN theory presented in Sec. II C. Therefore, the calculation results show good agreement with the FN theory in the high-field region. This indicates that FN plots can be used to analyze the current in semiconductor nanogaps.

Figure 8(d) shows the FN plots obtained by the measurement. Because the gap distance was constant in measurements, the bias voltage was not converted into the applied electric field. The plots are straight lines when the bias voltage is higher than 2 V, which is similar to the plots obtained from calculated results. The shapes of the plots are also similar to those of metal nanogaps in earlier studies.[32,33] Therefore, the measured currents are emission currents. As noted in Sec. IV B, the gap had partial contact. There are two sources of measured current: emission current through surface oxide layers at the contact point of the gap or that through the vacuum of the noncontact part. The gap surfaces were covered with very thin native oxide layers because the device was cleaved under a vacuum. A previous study[34] reported that an approximately 0.6 nm native oxide layer was grown on p-Si in 100 h of exposure to room air at room temperature; therefore, the thickness of the layers was estimated to be less than 0.6 nm because our experiment was conducted in a vacuum. If the measured current passes through the oxide film at a contact point, the current is considered to be tunneling at a subnanometer oxide layer with a very small area. To calculate the current, the field emission and current crowding effect, which cause additional resistance,[35,36] should be considered. Our measurement results should not differ from the possibility of tunneling in the oxide layer. However, because of the atomic asperity of the surface, the contact area seemed to be substantially smaller than that of the gap surface, and the distance of the noncontact part of the gap seemed to be very small (approximately 1 nm). Because the emission current is proportional to the emission area, a substantial current flow through the noncontact part, and the current through the noncontact part of the gap is significantly larger than that through the contact area. Therefore, it is appropriate that the measured current should be regarded as an emission current through the vacuum of the noncontact part.

The vertical intercepts and slopes of the straight lines in Figs. 8(a)–8(c) seem to vary with the electrode material and gap distance. Because the intercepts and slopes do not depend on the electrode material or the gap distance in the FN theory, the variation of these values decreases the accuracy of the quantitative analysis of the measured I–V characteristics using their FN plots. Compensations are required for a more accurate analysis. In this study, we propose an effective work function as compensation, as shown in Fig. 9. It is defined as the maximum height of the potential barrier from the Fermi level of the electrode when no

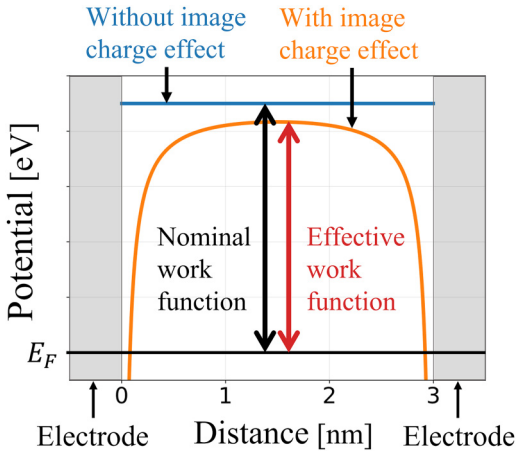

FIG. 9. Definition of effective work function. The lower and upper curves show the potential barrier with and without the image charge effect, respectively.

bias voltage is applied. In nanogaps, the maximum potential barrier is smaller than the nominal work function owing to the image charge effect. With an effective work function, the image charge effect can be considered. Furthermore, because the Fermi level of the electrode is used as a reference for the potential, the effect of the electrode material can also be considered.

The calculated results were analyzed using nominal and effective work functions to evaluate the effectiveness of this compensation. According to Eq. (16), the intercepts of the plots in Figs. 8(a)–8(c) are at $-b_{\mathrm{FN}}\,\Phi^{3/2}$ under the assumptions of the FN theory. This indicates that if the FN theory is valid, the quotients of the intercepts and work functions to the power of 3/2, that is $-b_{\mathrm{FN}}$, are constant, and the value is $-1.0652 \times 10^{38}\ \mathrm{J}^{-3/2}\ \mathrm{V\ m}^{-1}$, which is referred to as the FN constant. In Fig. 10, the FN constants obtained using the nominal work functions are indicated as circles. These values vary with the electrode material and gap distance, and their standard deviation is $2.30 \times 10^{37}\ \mathrm{J}^{-3/2}\ \mathrm{V\ m}^{-1}$. The stars in Fig. 10 show the FN constants obtained using the effective work functions. The dependence of the values on the electrode material and gap distance is smaller. Their standard deviation is $8.45 \times 10^{36}\ \mathrm{J}^{-3/2}\ \mathrm{V\ m}^{-1}$, which is 37% of the value obtained using nominal work functions. This decrease in the standard deviation verifies the effectiveness of applying compensation using effective work functions. The proposed compensation method enables a quantitative analysis using FN plots to include the image potential effect, which is a large deviation of a nanogap from the FN model. The method improves the accuracy of the analysis.

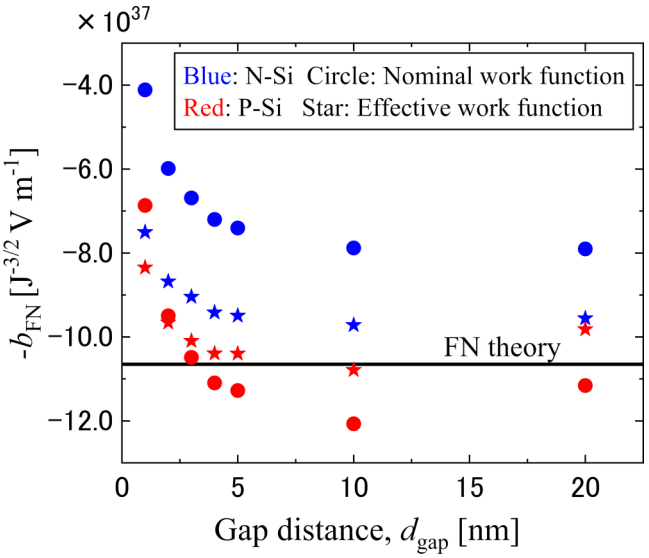

FIG. 10. FN constant calculated via a theoretical analysis. The values are calculated in both n-Si and p-Si. The nominal and effective work functions are used to obtain circles and stars in the plot. The theoretical value of the FN constant is indicated by the black line.

C. Comparison of calculated and measured results

Figure 11 shows calculated and measured results in a semilog plot. The calculated results are scaled with an emission area of $3.42 \times 10^{-14}\ \mathrm{m}^2$ and a field enhancement factor of 2.51. These values were obtained by adjusting them to minimize the sum of the squared error between the measured and calculated results, assuming that the emission area, field enhancement factor, and gap distance do not depend on the bias voltage. These values can depend on the bias voltage and electrostatic attraction force,[37] as is discussed later. The calculated results of the nanogap with a gap distance of approximately 1 nm were fitted to measurement results, and the calculation result for a gap of 1.1 nm yielded the best fit. These fitted parameters indicate that the dominant current flowed through the gap of 1.1 nm with an area of $3.42 \times 10^{-14}\ \mathrm{m}^2$. Because the fitted emission area is smaller than the nominal area by two orders of magnitude, the electric field was enhanced in the part where the dominant current flowed, as indicated by the fitted field enhancement factor.

Figure 11 shows that the measured current is larger than the fitted calculated result when the bias voltage is lower than 5 V. There are two possible causes for this. The first cause is that the conditions are different at low and high bias voltages. The field enhancement factor and emission area are affected by electrostatic attraction. When the bias voltage is low, the attraction force is small, and the gap distance is larger than that at a high bias voltage. Therefore, the current is more concentrated in a small area where the gap distance is locally narrower and the local field enhancement is higher. In other words, when the bias voltage is low, the actual applied voltage is higher than the assumed calculated result shown in Fig. 11, which causes underestimation. The second cause is that the effects of the space charge and exchange-correlation were not considered in calculations. The exchange-correlation effect enhances emission current,[22,23] whereas the space charge effect limits emission current when the bias voltage is high.[22–24,38] These influences are consistent with the results shown in Fig. 11.

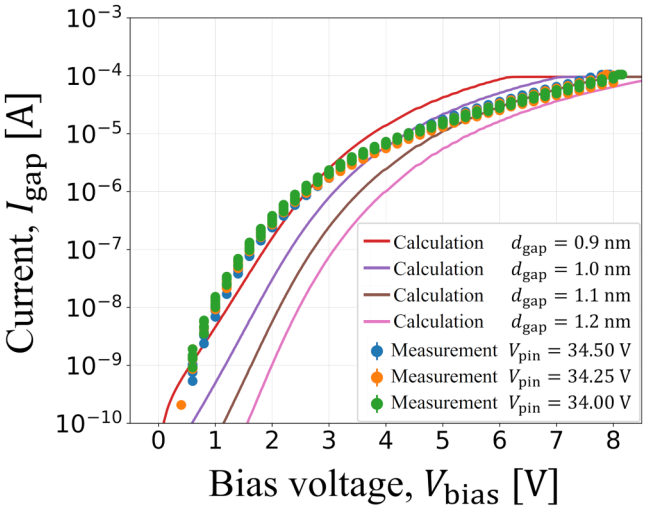

FIG. 11. Comparison of measured and calculated results.

While there are some minor deviations, Fig. 11 indicates a good match between the calculated and measured currents. The results of the measurement cannot be directly compared with data from other papers due to the specificity of the cleaved nanogap in this study, but the match indicates the veracity of our data.

VI. CONCLUSIONS

In this study, we developed a numerical model to evaluate the characteristics of field emission at semiconductor nanogaps and conducted measurements to verify the validity of the model. The model was developed based on the metal nanogap model by incorporating the carrier distribution in semiconductors and was used to determine the I–V characteristics for various gap distances. In terms of carrier distribution, an analysis performed using the models revealed that two types of carriers dominate the current, namely, those near the Fermi level and those near the top of the potential barrier, and there are differences in the behaviors of the minority carriers of n-Si and p-Si. In addition, the calculated results confirmed that the qualitative analysis using the FN plot is possible for semiconductor nanogaps, and we proposed a compensation method for the work function for the quantitative analysis using the FN plot. Furthermore, the I–V characteristics of the p-Si nanogap were measured using a MEMS device. Although the gap had partial contact, we considered that the measured current is the emission current at the vacuum nanogap.

In our model, it was assumed that the band structure is the same on the gap surface and in the bulk; however, there is likely a certain influence from the surface. Additionally, the calculations in this work did not consider the effects of space charge and exchange-correlation. Thus, the above effects should be considered in the future to improve the accuracy of the model. Another problem encountered in this study is the partial contact of gaps in the experiment. Improving the device design would result in more accurate gap spacing control.

ACKNOWLEDGMENTS

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant No. JP21H01261 and Kyoto University Nanotechnology Hub in the "ARIM Project" sponsored by MEXT, Japan.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Yuki Akura: Investigation (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal). **Masaki Shimofuri:** Methodology (equal); Resources (lead); Software (supporting). **Amit Banerjee:** Conceptualization (supporting). **Jun Hirotani:** Conceptualization (supporting); Writing – review & editing (supporting). **Toshiyuki Tsuchiya:** Conceptualization (lead); Funding acquisition (lead); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]M. Haras and T. Skotnicki, Nano Energy **54**, 461 (2018).
[2]W. Liu, Q. Jie, H. S. Kim, and Z. Ren, Acta Mater. **87**, 357 (2015).
[3]M. S. Dresselhaus, G. Chen, M. Y. Tang, R. G. Yang, H. Lee, D. Z. Wang, Z. F. Ren, J.-P. Fleurial, and P. Gogna, Adv. Mater. **19**, 1043 (2007).
[4]Z. Soleimani, S. Zoras, B. Ceranic, S. Shahzad, and Y. Cui, Sustain. Energy Technol. Assess. **37**, 100604 (2020).
[5]T. Caillat, A. Borshchevsky, and J.-P. Fleurial, J. Appl. Phys. **80**, 4442 (1996).
[6]K. Biswas, J. He, I. D. Blum, C. I. Wu, T. P. Hogan, D. N. Seidman, V. P. Dravid, and M. G. Kanatzidis, Nature **489**, 414 (2012).
[7]X. Chen, Z. Zhou, Y. H. Lin, and C. Nan, J. Materiomics **6**, 494 (2020).
[8]S. H. Heo *et al.*, Nat. Commun. **10**, 864 (2019).
[9]C. J. Yao, H. L. Zhang, and Q. Zhang, Polymers **11**, 107 (2019).
[10]Y. Hishinuma, T. H. Geballe, B. Y. Moyzhes, and T. W. Kenny, Appl. Phys. Lett. **78**, 2572 (2001).
[11]T. Li, W. Hu, and D. Zhu, Adv. Mater. **22**, 286 (2010).
[12]T. Adam, T. S. Dhahi, S. C. B. Gopinath, U. Hashim, and M. N. A. Uda, Biotechnol. Appl. Biochem. **69**, 1395 (2022).
[13]B. Liu, M. Tsutsui, and M. Taniguchi, Micromachines **9**, 282 (2018).
[14]J. Sadar, Y. Wang, and Q. Qing, ECS Trans. **77**, 65 (2017).
[15]H. Park, A. K. L. Lim, A. P. Alivisatos, J. Park, and P. L. McEuen, Appl. Phys. Lett. **75**, 301 (1999).
[16]A. Cui, Z. Liu, H. Dong, Y. Wang, Y. Zhen, W. Li, J. Li, C. Gu, and W. Hu, Adv. Mater. **27**, 3002 (2015).
[17]M. Shimofuri, A. Banerjee, Y. Hirai, and T. Tsuchiya, in *IEEE 35th International Conference on Micro Electro Mechanical Systems*, Tokyo, Japan, 9–13 January 2022 (IEEE, New York, NY, 2022), p. 511.
[18]M. Shimofuri, A. Banerjee, J. Hirotani, Y. Hirai, and T. Tsuchiya, J. Microelectromech. Syst. **32**, 67 (2023).
[19]M. Takai *et al.*, Microelectron. Eng. **41–42**, 453 (1998).
[20]J. S. Kang and K. C. Park, J. Vac. Sci. Technol. B **35**, 02C109 (2017).
[21]F. Urban, M. Passacantando, F. Giubileo, L. Iemmo, and A. Di Bartolomeo, Nanomaterials **8**, 151 (2018).
[22]P. Zhang, Sci. Rep. **5**, 9826 (2015).
[23]S. Banerjee and P. Zhang, AIP Adv. **9**, 085302 (2019).
[24]W. S. Koh and L. K. Ang, Nanotechnology **19**, 235402 (2008).
[25]S. M. Sze, Y. Li, and K. K. Ng, *Physics of Semiconductor Devices*, 4th ed. (Wiley, Hoboken, NJ, 2021), pp. 16–23.
[26]M. A. Green, J. Appl. Phys. **67**, 2944 (1990).
[27]F. J. Morin and J. P. Maita, Phys. Rev. **96**, 28 (1954).
[28]G. E. Jellison Jr., and D. H. Lowndes, Appl. Phys. Lett. **41**, 594 (1982).
[29]P. J. Timans, J. Appl. Phys. **74**, 6353 (1993).
[30]Y. P. Varshni, Physica **34**, 149 (1967).
[31]S. D. Liang, *Quantum Tunneling and Field Electron Emission Theories* (World Scientific, Toh Tuck Link, Singapore, 2014), Sec. 12, p. 2.
[32]J. Xu, Q. Wang, Z. Tao, Y. Zhai, C. Guangdian, Z. Qi, and X. Zhang, IEEE Trans. Electron Devices **64**, 2364 (2017).
[33]S. Nirantar, T. Ahmed, G. Ren, P. Gutruf, C. Xu, M. Bhaskaran, S. Walia, and S. Sriram, Nano Lett. **18**, 7478 (2018).
[34]M. Morita, T. Ohmi, E. Hasegawa, M. Kawakami, and M. Ohwada, J. Appl. Phys. **68**, 1272 (1990).
[35]P. Zhang, Y. Y. Lau, and R. S. Timsit, IEEE Trans. Electron Devices **59**, 1936 (2012).
[36]P. Zhang and Y. Y. Lau, J. Appl. Phys. **108**, 044914 (2010).
[37]J. Lin, P. Y. Wong, P. Yang, Y. Y. Lau, W. Tang, and P. Zhang, J. Appl. Phys. **121**, 244301 (2017).
[38]S. Srisonphan, IEEE Electron Device Lett. **42**, 1540 (2021).