Transport mechanisms of electrons and holes in dielectric films

K A Nasyrov, V A Gritsenko

DOI: 10.3367/UFNe.0183.201310h.1099

Contents

1.	Introduction: the importance of understanding electron and hole transport mechanisms	
	in silicon device dielectrics	999
2.	Two-band conduction in dielectrics	1002
3.	Contact-limited mechanism of electron and hole injection	1002
	3.1 Fowler-Nordheim tunneling injection at a contact; 3.2 Thermally assisted contact tunneling; 3.3 Thermoionic	
	contact emission: the Schottky effect; 3.4 Trap mediated injection in silicon nitride	
4.	Trap ionization models	1005
	4.1 Trap ionization by the Frenkel mechanism; 4.2 Multiphonon ionization of traps; 4.3 Charge transport by	
	trap-to-trap tunneling	
5.	Space charge in a dielectric and the Shockley–Reed–Hall equations	1007
6.	Two band conduction and trap parameters in Si_3N_4	1007
7.	Conduction in silicon-enriched silicon nitride	1009
8.	Monopolar electron conduction in Al ₂ O ₃	1010
9.	Conclusion	1011
	References	1011

<u>Abstract.</u> Electron and hole transport mechanisms in amorphous silicon oxide, silicon nitride, and aluminum oxide, dielectric materials of high relevance to silicon device technology, are reviewed. It is established that the widely accepted Frenkel model provides a formal description of transport in trap-containing insulators, but nonphysical model parameters must be introduced in order to obtain quantitative agreement. It is shown that the multiphonon model of trap ionization consistently describes charge transport in insulators with traps.

1. Introduction: the importance of understanding electron and hole transport mechanisms in silicon device dielectrics

Dielectric films in silicon devices have a thickness in the range 1–300 nm. The electric field in a dielectric is in the range $3 \times 10^4 - 10^7$ V cm⁻¹ for a voltage of 1 V. Although the most frequently used dielectrics are amorphous films, polycrystalline films have also found applications in recent years. The

K A Nasyrov Institute of Automation and Electrometry, Siberian Branch of Russian Academy of Sciences, prosp. Akademika Koptyuga 1, 630090 Novosibirsk, Russian Federation Fax +7 (383) 333 33 63. E-mail: nasyrov@iae.nsk.su V A Gritsenko Rzhanov Institute of Semiconductor Physics, Siberian Branch of Russian Academy of Sciences, prosp. Lavrent'eva 13, 630090 Novosibirsk, Russian Federation Tel. +7 (383) 330 88 91. Fax +7 (383) 330 27 71 E-mail: grits@isp.nsc.ru

Received 4 March 2013 Uspekhi Fizicheskikh Nauk **183** (10) 1099–1114 (2013) DOI: 10.3367/UFNr.0183.201310h.1099 Translated by E G Strel'chenko; edited by A M Semikhatov most important silicon device dielectrics are the amorphous silicon compounds SiO₂ (dioxide), Si₃N₄ (nitride), and SiO_xN_y (oxynitride) [1]. Because electrons scatter strongly in amorphous materials lacking a long-range order, amorphous films have high breakdown fields, as is exemplified by the value 1.1×10^7 V cm⁻¹ for SiO₂. This is too much for a crystalline semiconductor to withstand because, due to the large mean free path, the avalanche multiplication of carriers results in breakdown fields $\approx 10^5$ V cm⁻¹.

Dielectrics placed in a strong electric field profusely exhibit nonlinear phenomena that are difficult to observe in crystalline semiconductors, such as tunnel injection of carriers in contacts, Frenkel trap ionization, and multiphonon trap ionization.

The following applications of dielectrics crucially require understanding transport mechanisms for the development of reliable silicon devices. Currently (as of 2013), silicon field effect transistors based on metal-dielectric-semiconductor (MDS) structures routinely use 63 m as a characteristic transistor size (the length of the transistor channel is about 70% of this number), the subgate silicon oxide being 1.2 nm thick. At the gate voltage ≈ 1 V, the field in the dielectric is close to 10^7 V cm⁻¹, and the tunnel current reaches 1 A cm⁻². A high leakage current leads to the injection of electrons from the inversion layer to the gate, thus decreasing the slope of the transistor I-V curve and leading to additional power dispersion (and hence to the overheating of the microcircuit). A further decrease in the transistor design rule to 45, 32, 22, and 16 nm requires a corresponding increase in the capacity of the subgate dielectric of the MDS transistor.

For four decades, MDS transistors have used thermal silicon oxide SiO_2 on silicon as a subgate dielectric. The capacity of subgate dielectrics was increased by decreasing the SiO_2 thickness. As noted above, increasing the subgate

capacity further by decreasing the SiO₂ thickness is impossible because of the prohibitively large tunneling current. The most general approach to solving this problem is to replace SiO₂ by so-called alternative dielectrics, or dielectrics with high dielectric constants (high-*k* dielectrics). We note that the dielectric constant of silicon dioxide is $\varepsilon = 3.9$. The current high-*k* dielectric candidates are Al₂O₃ ($\varepsilon \approx 10$), Hf_xSi_yO ($\varepsilon \approx 15$), Hf_xAl_yO_z ($\varepsilon \approx 15$), HfO_xN_y ($\varepsilon \approx 15$), HfO₂ ($\varepsilon \approx 25$), ZrO₂ ($\varepsilon \approx 40$), and TiO₂ ($\varepsilon \approx 80$) [2–5]. Using high-*k* dielectrics can increase the physical thickness while keeping the capacity, and can therefore suppress the tunneling current in an MDS transistor.

Similarly to silicon nitride Si₃N₄, high-*k* dielectrics generally have a high density ($\approx 10^{19}$ cm⁻³) of electron and/or hole traps. The injection of electrons and holes and their subsequent localization on (capture by) traps leads to the accumulation of charge in high-*k* dielectrics, shifting the transistor threshold voltage, and degrading the device performance. We note the extremely tight stability requirements for the MDS transistor threshold voltage. For example, even after 10 years of the continuous operation of a microprocessor, the accumulation of electrons (holes) on traps should not exceed the threshold by more than 10 mV. Therefore, predicting the stability and reliability of a high-*k* MDS transistor requires a detailed understanding of the mechanisms of charge carrier injection at the contacts and the capture and trap ionization process.

Another important application of dielectrics is as insulators in RAM storage capacitors. For charge accumulation on the capacitor plates not to cause any loss of information during the ≈ 1 ms reprogramming cycle, the dielectric leakage current at a voltage of 1 V should not exceed 10^{-7} A cm⁻² [3, 4]. The first material used as an insulator in a RAM storage capacitor was SiO₂, followed by silicon nitride Si₃N₄ ($\varepsilon = 7$) as an element in oxide–nitride–oxide (ONO) structures (oxide here meaning SiO₂). Currently, Al₂O₃, Ta₂O₅, HfO₂, ZrO₂, TiO₂, and SrTiO₃ ($\epsilon \approx 1000$) are being studied for the role of storage capacitor insulator. Using insulators in a storage capacitor requires the development of technology for synthesizing low-conductivity dielectrics, which implies the need for understanding the conduction mechanism and the dependence of the conductivity on the concentration, energy, and trap capture cross section (understanding this allows directionally optimizing, i.e., decreasing, the conductivity of the RAM capacitor dielectric).

In addition to the RAM storage capacitor, passive capacitors are involved in the design of silicon integrated circuits, which typically use SiO_2 and Si_3N_4 as an insulator. Increasing the information capacity of a silicon device decreases the capacitor area. Currently, high-*k* dielectrics (Al₂O₃, Ta₂O₅, HfO₂, ZrO₂, TiO₂, and SrTiO₃) are considered as candidates for such capacitors.

Flash memory design is yet another major application of dielectrics [5–10]. A flash memory cell is an MDS transistor whose threshold voltage is variable due to the accumulation of electrons and/or holes in the storage medium. The threshold voltage retention time is 10 years at 85°C. The storage medium is insulated from the semiconductor channel by the tunneling dielectric through which electrons and holes are injected (read–erase system). The gate and the storage medium are also insulated from each other by a dielectric. Flash memory storage media include polysilicon 'floating' gates, silicon nitride with a high concentration of electron and hole traps, and semiconductor (Si, Ge, SiGe) or metallic

(Ni, Au, Pt) nanoclusters. The materials used as flash memory tunnel dielectrics are thermal silicon dioxide (SiO₂) and silicon oxynitride (SiO_xN_y).

Figure 1 presents the energy diagram of TaN-Al₂O₃-Si₃N₄-SiO₂-Si (TANOS) flash memory for different states [9, 10]. In the data recording mode, the conducting TaN-based gate (TaN has metallic properties) is subjected to a positive potential with respect to the silicon substrate (Figs 1a, b). Silicon's electrons perform Fowler-Nordheim tunneling through the tunneling oxide into the silicon nitride, to be captured there by deep electron traps ($W_{\rm t} \approx 1.5 \, {\rm eV}$). The essential requirement here is to block the parasitic injection of holes from the gate to the silicon nitride, for which there are two reasons: the high barrier to hole injection and the fact that due to the large dielectric constant in the blocking dielectric (Al₂O₃), the electric field in it is low. In the storage mode, the negative electron charge in short-circuited silicon nitride induces an enrichment layer in the p-silicon substrate (Fig. 1c), and therefore there is no conducting channel in the transistor (logic 0). The logic 0 hold time is limited by the ionization of electron traps in silicon nitride, a process which (similarly to hole traps) occurs by either the Frenkel mechanism or the multiphonon mechanism. In the erase mode, the tunneling mechanism results, for a negative potential on the silicon gate, in holes being injected through the silicon oxide into the silicon nitride, where they are captured by hole traps (Fig. 1d). (Importantly, the parasitic injection of electrons into the storage medium should be blocked.) In the data store mode, an inversion layer (conducting channel) is induced in the silicon substrate (logic 1, Fig. 1e). The data storage time (for logic 1) is determined by the ionization of hole traps in silicon nitride. Parasitic electron injection from the silicon and from the gate is also possible (Fig. 1e).

In the write–erase mode, knowing the carrier injection mechanisms from the semiconductor and the conducting gate (here, the conducting tantalum nitride TaN) to the dielectric (here, SiO₂ and Al₂O₃) is essential. We note that the flash memory data (logic 0 and logic 1) retention time should be 10 years at 85°C. To predict the retention time, a study is made of the accelerated spread of electrons and holes in a TANOS flash at increased temperatures (100–300°C). Clearly, a reliable determination of the retention time requires a detailed understanding of the ionization mechanisms of electron and hole traps in a storage medium (here, silicon nitride).

In the store mode, the electron and hole charges localized on traps in the silicon nitride can be compensated by carriers parasitically injected from the silicon and the gate. Predicting this phenomenon requires knowing the electron and hole injection mechanisms in dielectrics in weak electric fields.

Blocking layers used in TANOS flash memories are those of amorphous aluminum oxide Al_2O_3 , a material which, as shown by experience, contains traps. In the logic 0 and logic 1 storage modes, a charge flow can occur via Al_2O_3 traps from traps in the silicon nitride to the gate, resulting in limiting the data retention time. In this connection, the conduction mechanism in aluminum oxide is important to know.

In a yet another flash memory type that is currently dominating the market, the memory medium is a so-called 'floating' polysilicon gate, separated from the channel by a thermal SiO_2 layer and from the conducting silicon gate by an ONO structure. Cycling write–erase processes involve the



Figure 1. Energy diagram of a $TaN - Al_2O_3 - Si_3N_4 - SiO_2 - Si$ (TANOS) flash memory with a silicon nitride storage medium: (a) without applied voltage, (b) positive potential on metal, electron injection into silicon nitride, (c) closed-circuit electron-accumulating state, information storage, logic 0, (d) negative potential on metal, hole injection into silicon nitride, (e) closed-circuit hole-accumulating state, information storage, logic 1.

formation of electron traps in the tunneling oxide of such a flash memory. When in the store mode, the electrons residing in the floating gate can flow to the silicon via traps in the tunneling oxide by the thermally assisted tunneling mechanism [8–12], decreasing the charge accumulated in the floating gate. The thermally assisted tunneling limits the number of write–erase cycles in the floating gate flash memory structure. The number of flash memory reprogramming cycles lies in the range $10^4 - 10^6$. The limited number of write–erase cycles is a key disadvantage of flash memory versus magnetic (hard disc) memory. The mechanism of thermally assisted tunneling is currently the subject of intense research [11, 12].

A silicon integrated circuit has its speed determined by the signal delay in the circuit wire layout. Signal transfer between instruments occurs by means of a multilevel metal layout, the number of metallization levels reaching eight in current microprocessors. The delay τ of a signal passing through a metal line is determined by the line resistance R times the parasitic capacitance C between neighboring lines, $\tau = RC$. The speed is increased by decreasing both the resistance of the conducting lines and the parasitic capacitance between neighboring lines. Whereas early integrated circuits used aluminum lines, lower-resistivity copper is currently employed. The microcircuit insulator used early on was pyrolitic aluminum dioxide SiO₂ ($\varepsilon \approx 3.9$). Increasing the interline parasitic capacitance requires an increase in the dielectric constant of the insulator ('low-k' insulators). The low-k insulator currently in use is $SiO_xC_y - SiO_xF_y$ [13]. The leakage current of a low-k dielectric should not exceed a specified value.

In recent years, much work has been done toward developing the next generation of high-speed (resistive) memory technology [14–23]. Resistive memory (also known as a memristor) consists of a metal-dielectric-metal-based capacitor whose dielectric has its resistance varied and memorized as the applied voltage is varied. A key problem in the physics of the device is the charge transport mechanism in high- and low-resistive states. Resistive memory dielectrics used include stoichiometric SiO_x, GeO_x, TiO_x, HfO_x, TaO_x, ZrO_x, and AlO_x.

The conductivity of a dielectric should, by definition, be zero. However, in 1917, Pool observed an increase in the conductivity of mica with an electric field [24]. In 1938, Frenkel proposed a theoretical explanation of the exponential electric field dependence of the current in a dielectric with traps [25, 26]. Currently, Frenkel's model is ubiquitously used for explaining charge transport in a wide class of dielectrics, including nonstoichiometric oxide SiO_x (x < 2), silicon nitride SiN_x ($x \le 4/3$) [27], high-dielectric-constant dielectrics [28–31], and ferroelectrics [32]. However, as an explanation of trap ionization in semiconductors, the Frenkel effect can only be used for weak electric fields ($< 10^3$ V cm⁻¹) [33]. For strong electric fields, the multiphonon mechanism should be used [34, 35].

In two extreme cases, the conductivity of a dielectric can be limited either by contact carrier injection (Fig. 2a) or by trapmediated transport in a bulk dielectric. This classification is often vague, depending on how strongly the electric field on the contact is screened by the space charge accumulated in the traps (Fig. 2b).

Early work assumed that the conductivity of a dielectric is due to electrons, whether monopolar or single band in nature. Figure 3a presents the energy diagram of an MDS structure in the absence of an applied voltage, and Figure 3b shows an MDS structure for a monopolar (electron) conduction model. More detailed studies have shown that similarly to crystalline dielectrics, holes can contribute to the conductivity of a dielectric. In general, dielectrics exhibit bipolar, two-band



Figure 2. Electron injection at a contact and electron transport in a bulk dielectric for a dielectric (a) with and (b) without the electron space charge.



Figure 3. Energy diagram of an MDS structure in the absence of an applied voltage: (a) monopolar (electron), (b) bipolar, and (c) dielectric conduction model.

conduction (Fig. 3c) with electrons and holes respectively injected from the negatively biased contact (cathode) and from the positively biased contact (anode), and possibly with electron-hole recombination in trap centers near contacts.

2. Two-band conduction in dielectrics

As is known, conduction in crystalline semiconductors usually occurs via free electrons and/or holes that arise from the ionization of shallow donors and acceptors or results from band-to-band excitation due to the small width of the gap. At low temperatures ($T \leq 20$ K), when carriers are frozen out to donors and acceptors, the hopping mechanism sets in.

With their wide band gap (3–8 eV), dielectrics have a negligibly low concentration of intrinsic carriers. With the exception of thermal SiO₂ oxide on silicon, all the remaining dielectrics (the oxides SiO_x, Al₂O₃, Ta₂O₅, HfO₂, ZrO₂, TiO₂; the nitrides Si₃N₄, Ge₃N₄, BN) have a high density of traps, $\sim 10^{18} - 10^{20}$ cm⁻³. Amorphous silicon nitride Si₃N₄, a typical high-trap-density material, is currently a model

dielectric to study electron and hole transport mechanisms and the nature of electron and hole localized states (traps). The lifetime of localized (trapped) electrons and holes in silicon nitride is 10 years at 400 K (the so-called memory effect, which in Si_3N_4 is used in flash memory devices).

The energy of electron and hole traps in dielectrics lies in the range 0.5–2.0 eV. Also, the experimental evidence is that when in the original state, traps in a dielectric are empty of electrons and holes. For these reasons, and unlike donors and acceptors in semiconductors at relatively high temperatures (sufficient to ionize donors and acceptors), traps do not supply free electrons and holes to the conduction band. As shown by experiment, conduction in dielectrics is due to the nonequilibrium electrons and holes injected from a contact.

In semiconductors, the sign of charge carriers (electrons and holes) is determined from the Hall effect (or thermo-emf) measurements. For dielectrics, with their negligibly low concentration of free carriers (see above), this is not possible. Silicon nitride is the first material for which methods for separating the electron and hole current components were developed. Three such methods are currently known.

The first method is based on the accumulation in traps of electrons or holes injected from silicon or metal [36–40], while the other two rely on determining the sign of charge carriers at the silicon–dielectric interface, respectively, by measuring the $\lg J - V$ and C - V curves in the regime of nonequilibrium depletion of silicon [41–44] or by separating the electron and hole current components in an MDS transistor [45].

As already noted, silicon nitride exhibits bipolar twoband conductivity, with electrons and holes respectively injected from the negatively and positively biased contacts. In general, for either potential polarity, bipolar injection of electrons from the cathode and of holes from the anode occurs (Fig. 3c). The electron and hole contributions to the injection current can vary strongly depending on the following factors:

- the height of the contact-dielectric barrier,
- the thickness of the tunneling oxide,
- the pulse duration,
- the polarity of the potential,
- the thickness of the silicon nitride.

3. Contact-limited mechanism of electron and hole injection

As noted in the Introduction, two mechanisms can be conventionally identified for the conductivity of thin dielectric films: carrier injection from a contact into the dielectric and charge transfer in the bulk dielectric. In the former case, three injection mechanisms are usually considered: the Fowler–Nordheim mechanism, thermally assisted tunneling, and Schottky thermoelectron emission (Fig. 4)

3.1 Fowler-Nordheim tunneling injection at a contact

The Fowler–Nordheim mechanism (autoelectron emission) (Fig. 4, arrow 1) dominates in high electric fields and at low temperatures, at which charge carriers can tunnel from the conduction band of a (metal or semiconductor) contact into the conduction band of the dielectric. Although tunneling is of a quantum mechanical (and hence temperature-independent) nature, the number of free carriers at the Fermi level depends on the temperature, leading to a weak temperature dependence in the electric current [46–50].



Figure 4. Electron injection mechanisms at a metal-semiconductor contact: *1*, Fowler–Nordheim tunneling; *2*, thermally assisted tunneling; *3*, over-barrier emission; *4*, Schottky effect. The hatched region is a metal.

For the electron injection from a metal, the current is given by [48]

$$J = AF^{2} \exp\left(-\frac{B\sqrt{m^{*}}\Phi_{0}^{3/2}}{F}\right),$$

$$A = \frac{e^{3}}{8\pi h \Phi_{0}}, \quad B = \frac{8\pi\sqrt{2}}{3he}.$$
(1)

where J is the current density, F is the electric field strength, Φ_0 is the height of the triangle potential barrier at the electrode-dielectric interface, and m^* is the effective tunneling mass.

Because the Fowler–Nordheim mechanism leads to a weak temperature dependence of the current, it manifests itself at high electric fields such that the barrier for tunneling becomes sufficiently narrow. The Fowler–Nordheim effect is observed at sufficiently large contact–dielectric barrier heights.

It has been established that the Fowler-Nordheim effect limits the electron injection in $Si - SiO_2 - Me$ structures [48]. Represented in Fig. 5 are the V-I curves of such structures with different metal contacts. The height of the $Si - SiO_2 - Me$ barrier for electrons depends on the work function of the metal, and the leakage current depends exponentially on the work function of the metal. The V-I curves are straightened in the coordinates that correspond to expression (1) (Fig. 6) [48]. In accordance with theoretical predictions, the tunneling current in silicon oxide depends weakly on the temperature. The effective tunneling electron mass in silicon oxide estimated from the slope of the V–I curve in Fig. 6 is $m_e^* \approx$ $0.4m_0$, where m_0 is the free electron mass. In [49], electron and hole injection from silicon to silicon oxide were studied in the case where the upper contact material was taken to be corona discharge plasma rather than a metal. It was found that for a positive potential on the silicon, hole injection from the silicon to the silicon oxide occurs.

The Fowler–Nordheim effect was observed at liquid nitrogen and room temperatures for electron injection from aluminum to silicon oxide (Fig. 7) [39]. The effective tunneling electron mass in silicon nitride determined from the slope of the *V*–*I* curve is $m_e^* \approx 0.4m_0$. Observations at liquid nitrogen temperatures were made for hole tunneling injection from gold to silicon nitride (see Fig. 7) [39]. The effective tunneling hole mass in silicon nitride as determined from the slope of the *V*–*I* curve is $m_h^* \approx 0.35m_0$.



Figure 5. V-I characteristics of Si-SiO₂-Me structures for electron injection from metals with different work functions.



Figure 6. V-I characteristics of Si-SiO₂-Me structures in Fowler-Nordheim coordinates.

3.2 Thermally assisted contact tunneling

The process of thermally assisted tunneling in a contact (Fig. 4, arrow 2) can be divided somewhat loosely into two stages. The first is the phonon-absorption-assisted excitation of charge carriers to a certain energy level, and at the second stage, the carriers tunnel through a triangle barrier. For the carrier injection from a metal, the expression for the current has the form [47, 51]

$$J = CF \exp\left\{-\frac{\Phi_0 - (1/6)\left[(heF)/(4\pi kT\sqrt{m^*})\right]^2}{kT}\right\}, (2)$$

1003

where
$$C = (2\pi m^* kT)^{1/2} (e/h)^2$$
.



Figure 7. Fowler–Nordheim effect for (1) electron injection from aluminum into silicon nitride and (2) hole injection from gold into silicon nitride.

Thermally assisted tunneling dominates at medium fields and medium temperatures (when the thermal energy is insufficient to overcome the contact barrier by the Schottky mechanism) and has the following features: the current density depends exponentially on the square of the field, the activation energy increases with the temperature, and the current strongly (exponentially) depends on the metal work function.

The effect was predicted for the vacuum–metal system [47, 52] and analyzed theoretically (for electrons) for a metal– dielectric interface [51]. Experimentally, thermally assisted injection of holes was observed at the gold–silicon-nitride interface in Ref. [39]. Figure 8 presents *V*–*I* curves for hole injection from gold into silicon nitride in coordinates corresponding to the mechanism of thermally assisted tunneling (Fig. 8a) and the Schottky effect (Fig. 8b). In case (a), the *V*–*I* curves are straightened at temperatures of 523 K and 593 K, and in case (b) are not straightened at all.

3.3 Thermoionic contact emission: the Schottky effect

Unlike the Fowler–Nordheim emission, the thermoelectron Schottky emission (Fig. 4, arrow 3) dominates at increased temperatures at which thermal energy is sufficient for overcoming the junction barrier, and charge carriers can be excited from the contact Fermi level to the dielectric conduction band. The strong field dependence in the Schottky effect occurs because of the lowering of the barrier due to imaging forces.



Figure 8. *V–I* characteristics for hole injection from gold to silicon nitride in coordinates corresponding to (a) thermally assisted tunneling coordinates and (b) the Schottky effect. Curve 5: theoretical dependence of the Schottky effect at 523 K. Curves: *I*, T = 593K; *2*, *4*, T = 523 K; *3*, T = 77 K.

The expression for the current for the metal-to-dielectric injection has the form [47, 52]

$$J = AT^2 \exp\left(-\frac{\Phi_0 - \beta_{\rm Sh} F^{1/2}}{kT}\right),\tag{3}$$

where $A = (4\pi m^* k^2 e)/(m_e h^2) = 120m^*/m_e$ [A cm⁻² deg⁻²] is the Richardson–Dushman constant, $\beta_{\text{Sh}} = [e^3/(4\pi \varepsilon_{\infty}\varepsilon_0)]^{1/2}$ is the Schottky constant, and ε_{∞} is the high-frequency dielectric constant.

A characteristic feature of the Schottky emission is that the current depends exponentially on both the temperature and the electric field. The V-I curves plotted in the coordinates $\lg J = f(\sqrt{F})$ are straight lines. The Schottky effect is characterized by an exponential variation of the current with the work function of the contact (i.e., the contact-dielectric barrier height) and dominates at high temperatures and low values of the potential barriers at the MDS contact.

3.4 Trap mediated injection in silicon nitride

If the silicon oxide layer at the silicon interface is thick (> 5 nm), then charge carriers can be injected through traps contained in the silicon oxide (a charge first tunnels from the silicon to a trap in the bulk silicon oxide and then from the trap to the free band of the silicon nitride). Due to the lower subbarrier factor, this process, occurring in stages, can be more advantageous, for example, than the Fowler–Nordheim injection through a triangle barrier. Theoretical and experimental studies of thermally assisted tunneling for electron injection are cited in Ref. [53]. We note, however, that this injection mechanism also applies to holes. The expression [54]

$$J_{\text{TAT}} = \frac{e^2 F \hbar P_{\text{t}}}{2^{1/3} m^* 16} \exp\left(-\frac{2}{3} \frac{(2m^*)^{1/2} \Phi^{3/2}}{eF \hbar}\right)$$
(4)

can be used to describe the thermally assisted injection of electrons and holes. The notation here is as follows: Φ is the barrier height (for a hole or an electron) and P_t is the density of the energy and volume distributions of traps (assumed to be uniform in the derivation). It is to be noted that the underbarrier factor in the exponent turns is two times less than the one for the Fowler–Nordheim law. For this

mechanism to be efficient, the traps must have the density greater than 10^{18} cm⁻³ for the given value (8 eV) of the silicon oxide gap width.

4. Trap ionization models

If traps have a high concentration in a dielectric, the volume charge on them limits injection from the contact, and the conductivity is determined by the trap ionization rate. There are currently two trap ionization mechanisms dominating the literature on trap modes: the Frenkel effect and the multiphonon ionization. In semiconductors containing shallow traps, hopping conductivity is observed at low temperatures, when the carriers are localized on traps [55].

4.1 Trap ionization by the Frenkel mechanism

The Frenkel effect is the reduction in the Coulomb potential of an isolated trap due to an applied electrical field (Fig. 9). It is assumed that there is a trap (localized state) with a captured electron in the band gap of the dielectric and that this electron interacts via the Coulomb law with a positively charged defect [25, 26]. The electric field lowers the potential barrier, increasing the probability of an electron to undergo a transition from a localized state to the conduction band. Three trap ionization mechanisms are possible:

1. Over-barrier ionization dominating at sufficiently high temperatures and relatively low fields, for which the barrier width is important and the barrier tunneling probability is low (Fig. 9, case 1). It is this case that was treated in the original work of Frenkel [25, 26].

2. Thermally assisted tunneling (Fig. 9, case 2), the intermediate case between tunneling and over-barrier ionization.

3. Tunneling through a potential barrier (Fig. 9, case 3) dominates in high fields (for which the potential barrier is narrow enough to allow tunneling) and at low temperatures.

The thermal ionization of traps, which is exactly what is called the Frenkel effect, yields the following expression for the trap ionization rate:

$$P = v \exp\left(-\frac{W - \beta_{\rm Fr}\sqrt{F}}{kT}\right),\tag{5}$$



Figure 9. Frenkel effect: electric-field-assisted lowering of the trap Coulomb potential. Arrows indicate possible ionization mechanisms: I, thermal over-barrier emission; 2, thermally assisted tunneling; 3, tunneling ionization. q is the electron charge and r is the spatial coordinate.

where v is the frequency factor and $\beta_{\rm Fr} = (e^3/(\pi \epsilon_{\infty} \epsilon_0))^{1/2}$ is the Frenkel constant.

For thermally assisted tunneling, the trap ionization probability per unit time takes the form [56]

$$P = v \exp\left(-\frac{W - \beta\sqrt{F}}{kT}\right) + \frac{v}{kT}$$

$$\times \int_{0}^{W - \beta\sqrt{F}} \exp\left(-\frac{\varepsilon}{kT} - \frac{2}{h}\int_{x_{1}}^{x_{2}} \sqrt{2m^{*}(eV(x) - \varepsilon)} \,\mathrm{d}x\right) \mathrm{d}\varepsilon,$$
(6)
$$V(x) = W - \frac{e}{4\pi\varepsilon_{\infty}\varepsilon_{0x}} - Fx,$$

where ε is the energy of the excited level, and x_1 and x_2 are the classical turning points

$$x_{1,2} = \frac{1}{2} \frac{W - \varepsilon}{eF} \left(1 \mp \left(1 - \frac{eF}{\pi \varepsilon_{\infty} \varepsilon_0 (W - \varepsilon)^2} \right)^{1/2} \right).$$
(7)

Tunneling through the Coulomb potential barrier is observed at low temperatures in strong electric fields.

A high concentration of traps results in lowering the barrier due to an overlap of the Coulomb potentials of neighboring traps (Fig. 10) [57]. At high temperatures, the expression for the V-I characteristic in the one-dimensional model has the form

$$J = veN^{2/3} \exp\left(-\frac{W}{kT}\right) \sinh\left(\frac{eF}{2N^{1/3}kT}\right).$$
 (8)

In strong fields, the logarithm of the current is proportional to the electric field strength (Pool's law).

4.2 Multiphonon ionization of traps

This model assumes that a free charge can be captured by (generally neutral) lattice defects.

As noted above, traps in dielectrics are deep, and their captured electrons have small localization regions. Such a concentrated charge deforms the nearest vicinity of the trap, which in turn affects the energy level of the captured charge. It is thus assumed that the energy of the trap and the dielectric lattice are strongly related (the so-called phonon-coupled trap).

Because the binding energy of electrons in deep centers greatly exceeds the mean phonon energy, only multiphonon

Wa/2 a/2

Figure 10. Potential barrier reduction due to an overlap of the Coulomb potentials of neighboring traps (Hill model).

processes can provide thermal ionization. The theory of multiphonon ionization uses configuration diagrams to describe electron transitions from a localized to a delocalized state. Vibrational motions of the defect and those of the nearest surroundings of the trap are described by a change in the coordinate Q. Clearly, this one-dimensional description is approximate if the trap has many nearest neighbors. Figure 11a shows the configuration diagram of a defect (trap) for an empty trap and for the trap with a captured electron. The potential curve $U_f(Q)$ corresponds to the trap with a captured electron, and $U_0(Q)$ corresponds to the empty trap. The tilted line is the dependence of the energy of the captured electron $\varepsilon(Q)$ on the configuration coordinate. When $\varepsilon(Q) < E_{\rm c}$ ($E_{\rm c}$ is the dielectric conduction band edge), the electron is considered to be captured; ω is the defect vibration frequency and $W_{\rm ph} = \hbar \omega$ is the local phonon energy.

As can be seen from Fig. 11, there are two ways in which a trap can be ionized. In the first case, the transition from the ground to an excited state occurs due to a vertical (optical) transition; such a transition occurs at a fixed value of Q and requires the optical trap ionization energy W_{opt} . In an optical transition, the electron becomes free, and the nearest vicinity of the defect remains deformed at the first instant. This deformation is then carried by phonons throughout the dielectric. In the second case, an electron is excited by absorbing phonons with a change in the configuration coordinate (to ensure that $\varepsilon(Q) > E_c$). The required energy



Figure 11. (a) Configuration diagram of a trap in the theory of multiphonon ionization. Potentials U_f and U_0 respectively correspond to a trap with a captured electron and an ionized trap. (b) Oscillator model illustrating the multiphonon trap ionization, equilibrium state. (b) Perturbed state, transition to the continuous spectrum.

is the thermal ionization energy $W_{\rm T}$; $W_{\rm opt}$ always exceeds $W_{\rm T}$. A trap with a captured electron has the minimum energy at $Q_0 = [2(W_{\rm opt} - W_{\rm T})]^{1/2}$.

In the simplest case, a deep trap has one bound state. The position of the local level is determined by the defect-induced potential and depends essentially on the distance between the defect and the neighboring atoms. Thus, the vibrations of the defect modulate the position of the level of the localized electron (Fig. 11b). In the case of strong thermal vibrations, the level can move to the continuous spectrum, leading to the ionization of the trap [58].

A trap can also be ionized by a DC electric field, resulting in a strongly temperature-dependent ionization rate.

Treating defect vibrations in the semiclassical approximation, the ionization probability per unit time is given by [35]

$$P = \frac{eF}{2\sqrt{2m^*W_{\text{opt}}}} \exp\left(-\frac{4}{3}\frac{\sqrt{2m}}{\hbar eF}W_{\text{opt}}^{3/2} + 4\frac{W_{\text{opt}} - W_{\text{T}}}{W_{\text{opt}}}\frac{m^*\omega}{\hbar}\frac{W_{\text{opt}}^2}{e^2F^2}\coth\frac{W_{\text{ph}}}{2T}\right).$$
(9)

The corresponding quantum mechanical result is [24]

$$P = \sum_{n} \exp\left(\frac{nW_{\rm ph}}{2kT} - \frac{W_{\rm opt} - W_{\rm T}}{W_{\rm ph}} \coth\frac{W_{\rm ph}}{2kT}\right)$$
$$\times I_n\left(\frac{W_{\rm opt} - W_{\rm T}}{W_{\rm ph}\sinh\left(W_{\rm ph}/2kT\right)}\right) P_i(W_{\rm T} + nW_{\rm ph}), \qquad (10)$$

$$P_i(W) = \frac{eF}{2\sqrt{2m^*W}} \exp\left(-\frac{4}{3}\frac{\sqrt{2m}}{\hbar eF}W^{3/2}\right).$$
 (11)

Here, I_n is the modified Bessel function and $P_i(W)$ is the tunneling probability through a triangle barrier of height W.

4.3 Charge transport by trap-to-trap tunneling

The Shockley–Reed–Hall transport theory assumes that a charge (which we specify to be an electron) is injected from a contact into the dielectric conduction band and can be captured by a trap in the bulk dielectric. This captured electron can then escape into the conduction band and be again captured by another empty trap lower down in the field. A transitional trap filling process occurs and continues until the filling rate of empty traps becomes equal to the ionization rate of the filled ones. Following this, a steady-state charge transport process sets in.

But when the trap concentration is high and the distances between them are small, there is a possibility for the electron to tunnel between traps without being excited to the conduction band.

Figure 12 shows a diagram of electron tunneling from one phonon-coupled trap to another through a distance a in the presence of an electric field F. Solid and dashed lines respectively represent the original (prior to tunneling) and final (post-tunneling) states of the system. In an external field, the electron levels in neighboring traps are not coincident, and it therefore follows that inelastic processes — the emission and absorption of lattice phonons — should be involved in the transition to compensate the energy difference. The phonon-coupled trap model easily takes this into account. Within this model, the tunneling transition rate between traps



Figure 12. Diagram of electron tunneling from one phonon-coupled trap to another over a distance *a* in the electric field *F*. Solid lines: initial state of the system before tunneling. Dashed lines: final state of the system after tunneling. Horizontal dashed line: the tunneling transition of an electron from one trap to another.

is given by [59]

$$P = \frac{\sqrt{\pi} \hbar W_{\rm T}}{m^* a^2 Q_0 \sqrt{kT}} \exp\left(-\frac{W_{\rm opt} - W_{\rm T} + eFa}{2kT}\right)$$
$$\times \exp\left(-\frac{2a\sqrt{2m^*W_{\rm T}}}{\hbar}\right). \tag{12}$$

The numerical transport model uses the following equation to calculate the trap-filling rate in the bulk of the dielectric:

$$\frac{\partial}{\partial t} n_{i} = P_{i-1,i} n_{i-1} \left(1 - \frac{n_{i}}{N_{t}} \right) - P_{i,i-1} n_{i} \left(1 - \frac{n_{i-1}}{N_{t}} \right) + P_{i+1,i} n_{i+1} \left(1 - \frac{n_{i}}{N_{t}} \right) - P_{i,i+1} n_{i} \left(1 - \frac{n_{i+1}}{N_{t}} \right). \quad (13)$$

Here, N_t is the total trap density and n_i is the density of electron-filled traps on site *i* of the model lattice. The distance between sites is equal to the mean distance between traps, $a = N_t^{-1/3}$, and $P_{i-1,i}$ is the tunneling rate from site i - 1 to site *i*; for the inverse process, the tunneling rate is $P_{i,i-1}$.

The filling of the very first trap occurs from the contact and is described by

$$\frac{\partial}{\partial t} n_1 = v_{\rm inj} (N_{\rm t} - n_1) - v_{\rm ion} n_1 + P_{2,1} n_2 \left(1 - \frac{n_1}{N_{\rm t}} \right) - P_{1,2} n_1 \left(1 - \frac{n_2}{N_{\rm t}} \right), \qquad (14)$$

where v_{inj} is the rate of trap filling by electrons from the contact and v_{ion} is the inverse ionization of the trap into the contact:

$$\begin{aligned} v_{\rm inj} &= \exp\left(-\frac{\Phi - W_{\rm T}}{kT}\right) v_{\rm ion} \,,\\ v_{\rm ion} &= \int \frac{V_{\rm out}}{2z\sqrt{2\pi kT}} \\ &\times \exp\left[-\frac{\left(Q - Q_0\right)^2}{2kT} - \frac{4}{3} \frac{\left(2m^*\right)^{1/2} \left[\left(-\varepsilon\right)^{3/2} - \left(-eFz - \varepsilon\right)^{3/2}\right]}{eF\hbar}\right] \\ &\times \left[1 + \exp\left(-\frac{\Phi + \varepsilon + eFz}{kT}\right)\right]^{-1} dQ, \end{aligned}$$
(15)

 $\varepsilon = -Q_0(Q-Q_0) - W_{\text{opt}}$.

Here, Φ is the energy difference between the conduction band bottom of the dielectric and the Fermi level of the contact, z is the trap-contact distance, and V_{out} is the velocity of the escaped electron in the contact.

The boundary condition assumed for the opposite end of the dielectric is that the trap is empty near the contact.

5. Space charge in a dielectric and the Shockley–Reed–Hall equations

The experimental evidence is that in the unperturbed state, electron and hole traps in a dielectric are neutral and unfilled. As electrons and holes are accumulated on traps, a space charge forms in the dielectric. Charge transport, the capture of electrons and holes injected from the contact to the traps, and the ionization of traps are described by the Shockley– Reed–Hall equations

$$\frac{\partial n(x,t)}{\partial t} = \frac{1}{e} \frac{\partial J(x,t)}{\partial x} - \sigma v n(x,t) \left(N_{t}^{e} - n_{t}(x,t) \right) + n_{t}(x,t) P(x,t) - \sigma_{r} v n(x,t) p_{t}(x,t) , \qquad (16)$$

$$\frac{\partial n_{t}(x,t)}{\partial t} = \sigma v n(x,t) \left(N_{t}^{e} - n_{t}(x,t) \right) - n_{t}(x,t) P(x,t) - \sigma_{r} v p(x,t) n_{t}(x,t) , \qquad (17)$$

$$\frac{\partial p(x,t)}{\partial t} = \frac{1}{e} \frac{\partial J_{p}(x,t)}{\partial x} - \sigma v p(x,t) \left(N_{t}^{h} - p_{t}(x,t) \right) + p_{t}(x,t) P(x,t) - \sigma_{r} v p(x,t) n_{t}(x,t) , \qquad (18)$$

$$\frac{\partial p_{t}}{\partial t} = \sigma v p(x,t) \left(N_{t}^{h} - p_{t}(x,t) \right) - p_{t}(x,t) P(x,t) - \sigma_{T} v p(x,t) n_{t}(x,t), \qquad (19)$$

where *n* (n_t) is the concentration of free (trapped) electrons, *p* (p_t) is the concentration of free (trapped) holes, N_t^e (N_t^h) is the concentration of electron (hole) traps, F(x, t) is the local electric field, *e* is the electron charge, σ is the trap capture cross section, σ_r is the recombination cross section between the free and trapped carriers of the opposite sign, *v* is the drift velocity, *e* is the low-frequency dielectric constant, and *P* is the probability of the ionization of filled traps per unit time. The electron drift velocity is related to the current density as J = env.

The corresponding laws of injection (Fowler–Nordheim injection, the Schottky effect, thermally assisted tunneling, etc.) are used as boundary conditions, and the Poisson law

$$\frac{\partial F}{\partial x} = -e \, \frac{n_{\rm t}(x,t) - p_{\rm t}(x,t)}{\varepsilon \varepsilon_0} \tag{20}$$

is used to describe the space charge of trapped electrons and holes.

6. Two-band conduction and trap parameters in Si₃N₄

Amorphous semiconductors and amorphous dielectrics share the ability to localize electrons and holes. Amorphous silicon nitride Si_3N_4 can localize electrons and holes for long periods of time (a holding time of 10 years at 400 K [6]). The memory effect in Si₃N₄ is widely used in energy-autonomous memory microcircuits. The mechanism governing charge transport in silicon–oxide–nitride–oxide–silicon (SONOS) memory devices is the ionization of deep centers in Si₃N₄. The dominant current view is that the ionization mechanism in Si₃N₄ is due to the Frenkel effect.

In this section, experimental results on charge transport in Si₃N₄ are presented for a wide range of electric fields and temperatures and are quantitatively compared with the predictions of the Frenkel model and with those of the quantum model of the multiphonon trap ionization mechanism. In Ref. [56], a quantitative comparison is made between the experimental results on the conductivity of metal-nitrideoxide-semiconductor (MNOS) structures and the theory of multiphonon ionization for a monopolar conduction model, where only the electron injection from silicon is considered and the hole injection from a metal is neglected. Here, experimental results are compared with a more general bipolar model, which allows the electron injection from silicon and the hole injection from the metal and in which free electrons recombine with localized holes and free holes recombine with localized electrons (Fig. 13). For a MNOS structure, the temperature dependence of the current (Fig. 14) was measured for different voltages, and V-I characteristics were measured for different temperatures (Fig. 15). In all measurements, the voltage on the metal was positive, a polarity at which both electron injection from the silicon substrate and hole injection from aluminum occur (Fig. 13b).

For comparison with experiment, a one-dimensional bipolar model of conduction in Si₃N₄ [56] was used. Charge transport was described using Shockley–Reed–Hall equations (16)–(19) and Poisson equation (20), which takes the nonuniform distribution of the electric field in silicon nitride into account. The capture and recombination cross sections in silicon nitride were taken to have the values $\sigma_t^e = \sigma_t^h = \sigma_r = 5 \times 10^{-13} \text{ cm}^2$ [6, 36, 54, 56]. The trap ionization probability *P* in Si₃N₄ is estimated using either the Frenkel model in (5)–(7) or the multiphonon ionization model in (10).

For modeling purposes, electrons and holes were taken to have the same effective mass $0.5m_0$. The electron and hole injection from the silicon substrate and the aluminum electrode were respectively calculated for the Fowler–Nordheim mechanism.

The temperature dependence of the current was measured in the temperature range 77–410 K for voltages of 44, 37, and 30 V. Referring to Fig. 14, which plots the results in the Arrhenius coordinates lg $J-T^{-1}$, the current depends weakly on temperature for T < 200 K, suggesting tunneling as the trap ionization mechanism.

Using first the Frenkel model with thermally assisted tunneling to describe the experimental data, the best fit is obtained for electron and hole traps having the same energies $(W = W^{h} = W^{e} = 1.2 \text{ eV})$, the same effective masses $(m^{e*} = m^{h*} = m^{*})$, and the same concentrations $(N_{t}^{e} = N_{t}^{h} = 7 \times 10^{19} \text{ cm}^{-3})$ (dashed lines in Fig. 14). Complete agreement was obtained only by using an unphysically small frequency factor, $v = 6 \times 10^{6} \text{ s}^{-1}$ (compared to $v = W/h \approx 4 \times 10^{14} \text{ s}^{-1}$ originally estimated by Frenkel[25]). Earlier work [60–63] also used a lowered frequency factor, $v = 10^{6} - 10^{9} \text{ s}^{-1}$, to obtain agreement with experiment. The low-temperature range is dominated by the tunneling effect, and of key importance is the tunneling effective mass. Maximizing the agreement between experiment and calculations required using different effective masses for different voltages: $m^{*} = 4.8m_{0}$ for



Figure 13. Energy diagram of an MNOS structure (a) in the flat band regime and (b) for a positive potential on the metal.



Figure 14. Temperature dependence of the current in an MNOS structure for different positive potentials on the metal. Dots: experiment, solid lines: multiphonon model calculations. Model parameters: $m_e^* = m_h^* = 0.5m_0$, $W_{opt} = 2.8$ eV, $W_T = 1.4$ eV, $W_{ph} = 60$ meV, $\sigma_t^e = \sigma_t^h = \sigma_r = 5 \times 10^{-13}$ cm². Dashed lines: Frenkel model for equal trap energies, $W = W^h = W^e = 1.2$ eV, equal effective masses $m^{e*} = m^{h*} = m^*$ and equal concentrations $N_t^e = N_t^h = 7 \times 10^{19}$ cm⁻³ for electron and hole traps.



Figure 15. *V*–*I* characteristics of Si_3N_4 measured at different temperatures (dots) for a positive potential on Al and calculated (solid lines) from the multiphonon ionization theory. The parameters are the same as in Fig. 14.

V = 44 V, $m^* = 3.8m_0$ for V = 37 V, and $m^* = 2.0m_0$ for V = 30 V. The effective mass values used in the calculations are an order of magnitude greater than the electron and hole tunneling effective masses $0.3-0.6m_0$ found experimentally for Si₃N₄ [39]. We therefore conclude that the Frenkel model does describe charge transport in Si₃N₄ formally, but requires an unphysically small frequency factor and an anomalously large tunneling effective mass for the quantitative agreement.

At the next stage, the experimental and calculated results were compared using the multiphonon ionization theory in Ref. [34] (solid lines in Fig. 14). It is established that the agreement is closest when the parameters of the electron and holes traps are taken to be the same: $W_T^e = W_T^h = 1.4 \text{ eV}, \quad W_{opt}^e = W_{opt}^h = 2.8 \text{ eV}, \quad N_t^e = N_t^h = 7 \times 10^{19} \text{ cm}^{-3}, W_{ph} = 0.06 \text{ eV}.$ Calculations using the multiphonon theory were performed to obtain *V*–*I* characteristics for T = 77, 300, and 410 K (Fig. 15, solid lines) and temperature dependences of the current for V = 44, 37, and 30 V (Fig. 14, solid lines). Overall, rather good agreement is observed between the modeled and measured results. The small discrepancy observed for voltages below 35 V may be due to the slow relaxation of the current in silicon nitride, a phenomenon whose nature is not yet understood [64].

Therefore, the theory of multiphonon trap ionization quantitatively describes experimental charge transport in silicon nitride in a wide range of electric fields and temperatures.

The results suggest that the theory of multiphonon ionization fits the data best if the concentration, the capture cross section, and the optical and thermal energy are taken to be the same for electron and hole traps. The idea of equal parameters was first introduced in Ref. [38] to explain experimental results on the draining of charge at different 'contracting' potentials and was also hypothesized in Refs [65, 66] (see also Ref. [67]) based on luminescence experiments and quantum mechanical modeling.

The calculations predict the optical energy to be twice the thermal energy, $W_{opt} = 2W_T$, suggesting that the activation energy for the capture of an electron or a hole is near zero.

This fact is consistent with the observed weak temperature dependence of the capture cross section for electron and holes in Si_3N_4 [36].

The calculated phonon energy $W_{\rm ph} = 60$ eV is equal to the phonon energy in amorphous silicon obtained from Raman scattering spectra, identifying amorphous silicon nanoclusters in silicon nitride as likely candidates for electron and hole traps.

Quantum size effects in amorphous silicon dots in silicon nitride were observed in optical absorption and photoluminescence experiments [68, 69]. Amorphous Si nanoclusters can capture electrons and holes in silicon nitride [70]. The large difference between the thermal and optical ionization energies predicted by the multiphonon theory is possibly due to the strong deformation of the Si₃N₄ lattice. The capture of electrons and holes in Si₃N₄ probably occurs on the minimal silicon cluster, the Si–Si bond. This model assumes that the Si–Si bond or, alternatively, a cluster of a few silicon atoms, is a deep center for electrons and holes and that it also makes an ionization center. The quantum mechanical modeling of the Si–Si bond in Si₃N₄ supports this hypothesis qualitatively [66].

7. Conduction in silicon-enriched silicon nitride

The theory of transport by tunneling between traps was tested by comparing its predictions with experimental results on the conductivity of silicon-enriched amorphous films of SiN_x grown at various SiH₄/NH₃ ratios [71]. The experiment measured the dependence of the current on the voltage applied to a dielectric with a thickness of ≈ 100 nm [16]. It is known that as the SiH₄/NH₃ ratio increases during the synthesis of Si₃N₄, and the parameter x decreases, signifying an increase in the concentration of stoichiometrically excessive silicon. Increasing the SiH₄/NH₃ ratio increases the refractive index n (see the caption to Fig. 17 below for its value).

In Fig. 16, the measured room-temperature V-I characteristics of stoichiometric (1) and silicon-enriched (2)silicon nitride are presented in coordinates corresponding to the Frenkel law and in coordinates corresponding to Pool's law, $\lg J \sim F$. Worthy of note, the current in siliconenriched nitride is much larger than in its stoichiometric counterpart. Furthermore, the former compound conducts a noticeable current even at low voltages, whereas in the latter this is only possible after a certain voltage threshold is reached—a striking contrast, which we believe is due to the high concentration of traps in the enriched compound. This is consistent with the earlier hypothesis [65] that the nature of traps in silicon nitride relates to silicon clusters and, in particular, the Si-Si bond can serve as a trap in the compound, and a local phonon is associated with this bond.

To obtain the voltage dependence of the current, we performed numerical calculations that allowed the ionization of neutral traps, but not tunneling between them. The calculations showed good agreement with experimental data for near stoichiometric nitride (curves 1, 2 in Fig. 17), but under no conditions could they reproduce the results obtained for silicon-enriched nitride. This could only be satisfactorily done by introducing trap-to-trap tunneling (curves 3, 4 in Fig. 17). To achieve good agreement between theory and experiment, a high concentration (> 10^{21} cm⁻³) of traps is needed.



Figure 16. (a) *V*–*I* characteristics of silicon-enriched silicon nitride in coordinates corresponding to the Frenkel law: SiH₄/NH₃ = 0.02 (*I*); 0.1 (*2*); 0.2 (*3*); 0.33 (*4*); 0.5 (*5*). (b) The same in coordinates corresponding to the Pool law: SiH₄/NH₃ = 0.02 (*I*'); 0.1 (*2*'); 0.33 (*3*').



Figure 17. Experimental (dots) and calculated (solid lines) *V*–*I* characteristics of Si₃N₄ for different trap concentrations. The trap parameters used in the calculations are (*I*) SiH₄/NH₃ = 0.02, *n*=1.96; other parameters: $N_t = 10^{19}$ cm⁻³, $W_T = 1.7 \text{ eV}$, $W_{opt} = 3.5 \text{ eV}$, $\Phi = 2.3 \text{ eV}$, $m^* = 0.5m_0$; (*2*) SiH₄/NH₃ = 0.1, *n* = 2.05, $N_t = 2 \times 10^{19}$ cm⁻³, $W_T = 1.5 \text{ eV}$, $W_{opt} = 3 \text{ eV}$, $\Phi = 1.8 \text{ eV}$, $m^* = 0.5m_0$; (*3*) SiH₄/NH₃ = 0.2, n = 2.2, $N_t = 3 \times 10^{21}$ cm⁻³, $W_T = 1.55$ eV, $W_{opt} = 3 \text{ eV}$, $\Phi = 2.1 \text{ eV}$, $m^* = 0.5m_0$; (*4*) SiH₄/NH₃ = 0.5, *n* = 2.35, $N_t = 3 \times 10^{22}$ cm⁻³, $W_T = 1.7 \text{ eV}$, $W_{opt} = 3 \text{ eV}$, $\Phi = 2.1 \text{ eV}$, $m^* = 0.5m_0$. Here, Φ is the barrier height for electron injection from the silicon substrate into nitride (the energy difference between conduction bands in silicon nitride and silicon). Calculations are done by varying N_t , W_T , W_{opt} , and Φ while keeping the effective electron mass fixed.

8. Monopolar electron conduction in Al₂O₃

In amorphous aluminum oxide, the band gap is $E_g \approx 6.2$ eV and the barrier height for electrons at the Si/Al₂O₃ interface is $\phi^e \approx 2.0$ eV [5]. Al₂O₃ films have a relatively low concentration of traps, and hence their leakage current is lower than in HfO₂. Using Al₂O₃ as a blocking layer in silicon nitride quantum dot flash memories was proposed in [6, 9, 10]. Conduction in Al₂O₃ has been interpreted by many in terms of the Pool–Frenkel model [72,73].

The potential barrier for holes at the Si/Al_2O_3 interface equals 3.1 eV, which is much larger than the barrier for electrons at the Si/Al₂O₃ and Al/Al₂O₃ interfaces. Therefore, if the Al electrode is at a negative potential, the dominant mechanism is the electron injection from the upper electrode (i.e., from aluminum) [74]. For this reason, monopolar conduction in Al₂O₃ due to electrons is considered. The injection current at the Al/Al₂O₃ interface was calculated using the Fowler–Nordheim mechanism, with the electron effective mass taken to be $0.45m_0$ [75–77].

The experimental and predicted results (the Frenkel model with thermally assisted tunneling) on the temperature dependence of the current in and V-I characteristics of Al₂O₃ are compared in Figs 18 and 19. The calculations were performed for the capture cross section $\sigma^{e} = 5 \times 10^{-13} \text{ cm}^{2}$ and the electron trap concentration $N_{\rm t} = 5 \times 10^{19} {\rm ~cm^{-3}}$. At high temperatures in weak electric fields, the conductivity depends on the temperature exponentially. The trap energy is estimated as W = 1.4 eV from the slope of the lg $J - T^{-1}$ curve. The best fitting value of the frequency factor was $v = 10^9 \text{ s}^{-1}$. This value is also anomalously small. At low temperatures, the current depends weakly on temperature and is determined by the tunneling mechanism. The best fitting tunneling electron effective mass is $m^* = 3.5m_0$. We note the necessity of using different electron effective masses at the boundary and in the bulk of Al_2O_3 ($m^* = 3.5m_0$ in the bulk and $m^* = 0.45m_0$ at the Si/Al₂O₃ interface). Using the bulk value $m^* = 0.45m_0$ at the boundary results in a contactlimited conductivity.

In Fig. 20, the experimental temperature dependences of the V-I characteristics (dots) are compared with those calculated for the multiphonon ionization in Al₂O₃ (solid lines) [34] using a cross section of 5×10^{-15} cm² and a concentration of 2×10^{20} cm⁻³ for neutral traps. The electron effective mass and the magnitude of W_{opt} were estimated under low-temperature, high-field conditions to give the electron effective mass $0.4m_e$, $W_{opt} = 3.0$ eV, $W_T = 1.5$ eV, and $W_{ph} = 0.05$ eV as the best fitting values for liquid nitrogen temperature.

A detailed study of optical transition on a hydrogen vacancy in amorphous aluminum oxide yielded the lumines-



Figure 18. Temperature dependences of the current in Al₂O₃ measured for different negative potentials (dots) on Al and calculated using the Frenkel model including thermally assisted tunneling (solid lines). Trap parameters: W = 1.4 eV, $m^* = 3.5m_{\text{e}}$, $v = 1 \times 10^9 \text{ s}^{-1}$, $\sigma = 5 \times 10^{-13} \text{ cm}^2$, $N_{\text{t}} = 5 \times 10^{19} \text{ cm}^{-3}$.



Figure 19. V–I characteristics of Al₂O₃ measured (dots) at different temperatures and calculated (solid lines) using the Frenkel model including thermally assisted tunneling. Trap parameters: W = 1.4 eV, $m^* = 3.5m_{\rm e}$, $v = 1 \times 10^9$ s⁻¹, $\sigma = 5 \times 10^{-13}$ cm², $N_{\rm t} = 5 \times 10^{19}$ cm⁻³.

cence energy 3.0 eV and the excitation energy 6.0 eV [78–80]. The half of the luminescence Stokes shift is 1.5 eV, which equals the thermal energy of the trap. This identifies oxygen vacancies as traps for electrons in amorphous aluminum oxide.

In amorphous aluminum oxide, the ratio $W_{\text{opt}}/W_{\text{T}} = 2$ corresponds to the zero activation energy for trapping. Previously, the same value of this ratio was obtained for Si₃N₄ [61].

Thus, just as for silicon nitride, the Frenkel model quantitatively (formally) describes charge transport in a wide range of electric fields and temperatures, but requires an anomalously large effective mass and an anomalously small frequency factor to fit the data. At the same time, the multiphonon conduction model is capable of quantitatively (consistently) describing the transport of electrons in amorphous aluminum oxide.



Figure 20. *V*–*I* characteristics of Al₂O₃ measured at different temperatures for negative potential on Al (dots) and calculated using the theory of multiphonon ionization (solid lines). Trap parameters: $W_{\rm T} = 1.5$ eV, $W_{\rm opt} = 3.0$ eV, $W_{\rm ph} = 0.05$ eV, $m^* = 0.4m_{\rm e}$, $N_{\rm t} = 2 \times 10^{20}$ cm⁻³, $\sigma = 5 \times 10^{-15}$ cm².

9. Conclusion

With the exception of thermal oxide on silicon, the majority of dielectrics have their conduction limited by the ionization of traps. The widely accepted Frenkel model based on the ionization of charged Coulomb traps describes experiment quantitatively (formally) over a wide range of fields and temperatures. However, quantitative agreement with experiment requires an unphysically large frequency factor and an anomalously large effective mass, an order of magnitude larger than in experiment. Experimental V-I dependences are described quantitatively by the theory of multiphonon ionization of neutral traps in a wide temperature range. Depending on the barrier height for electrons and holes, a dielectric can be a monopolar conductor, as exemplified by amorphous Al₂O₃, for which hole barriers are much higher than those for electrons. Al₂O₃ exhibits monopolar electron conduction, and in amorphous silicon nitride, with its nearly equal-height barriers for electrons and holes, two-band bipolar conduction occurs, with electrons (holes) injected from the negatively (positively) biased electrode. In the bulk of the dielectric, two recombination processes occur: between free electrons and localized holes, and between free holes and localized electrons. Experiments on transport are quantitatively explained by assuming that traps for electrons and holes are neutral.

As shown by experiment, concentrations, energies, and cross sections are the same for electron and hole traps in silicon nitride. The thermal activation energy is twice the optical ionization energy, suggesting that there are no barriers for the capture of electrons and holes onto traps. This does not contradict the weak temperature dependence of the cross section for the capture of electrons and holes in silicon nitride.

References

- Gritsenko V A Phys. Usp. 51 699 (2008) [Usp. Fiz. Nauk 178 727 (2008)]
- 2. Kingon A I, Maria J-P, Streiffer S K Nature 406 1032 (2000)
- 3. Wilk G D, Wallace R M, Anthony J M J. Appl. Phys. 89 5243 (2001)
- 4. Robertson J Eur. Phys. J. Appl. Phys. 28 265 (2004)

- Perevalov T V, Gritsenko V A Phys. Usp. 53 561 (2010) [Usp. Fiz. Nauk 180 587 (2010)]
- Roizin Y, Gritsenko V, in *Dielectric Films for Advanced Microelec*tronics (Eds M Baklanov, M Green, K Maex) (Chichester: John Wiley & Sons, 2007) p. 251
- 7. Nasyrov K A et al. *JETP* **102** 810 (2006) [*Zh. Eksp. Teor. Fiz.* **129** 926 (2006)]
- Gritsenko V A et al. *Dielektriki v Nanoelektronike* (Dielectrics in Nanoelectronics) (Exec. Ed. A L Aseev) (Novosibirsk: Izd. SO RAN, 2010)
- 9. Lee C-H, Park K-C, Kim K Appl. Phys. Lett. 87 073510 (2005)
- 10. Nasyrov K A, Shaimeev S S, Gritsenko V A *JETP* **109** 786 (2009) [*Zh. Eksp. Teor. Fiz.* **136** 910 (2009)]
- 11. Goguenheim D et al. J. Non-Cryst. Solids 245 41 (1999)
- 12. Endoh T et al. J. Appl. Phys. 86 2095 (1999)
- 13. Maex K et al. J. Appl. Phys. 93 8793 (2003)
- 14. Waser R, Aono M Nature Mater. 6 833 (2007)
- 15. Strukov D B et al. *Nature* **453** 80 (2008)
- 16. Yang J J et al. Nature Nanotechnol. 3 429 (2008)
- 17. Lee M-J et al. Nano Lett. 9 1476 (2009)
- 18. Lee M-J et al. Adv. Functional Mater. **19** 1587 (2009)
- 19. Borghetti J et al. Nature 464 873 (2010)
- 20. Goux L et al. Appl. Phys. Lett. 97 243509 (2010)
- 21. Lee M-J et al. Nature Mater. 10 625 (2011)
- 22. Shaposhnikov A V et al. *Appl. Phys. Lett.* **100** 243506 (2012)
- 23. Wang Z et al. Phys. Rev. B 85 195322 (2012)
- 24. Pool H H Philos. Mag. 34 195 (1917)
- 25. Frenkel' Ya I Zh. Eksp. Teor. Fiz. 54 647 (1938)
- 26. Frenkel J Phys. Rev. 54 647 (1938)
- 27. Sze S M Physics of Semiconductor Devices (New York: Wiley, 1981)
- 28. Yeh C-C et al. Appl. Phys. Lett. 91 113521 (2007)
- 29. Zhu W J et al. IEEE Electron Dev. Lett. 23 97 (2002)
- 30. Xu Z et al. *Appl. Phys. Lett.* **80** 1975 (2002)
- 31. Chim W K et al. J. Appl. Phys. 93 4788 (2003)
- 32. Laha A, Krupanidhi S B J. Appl. Phys. 92 415 (2002)
- 33. Ganichev S D et al. *Phys. Rev. B* **61** 10361 (2000)
- 34. Makram-Ebeid S, Lannoo M Phys. Rev. B 25 6406 (1982)
- Abakumov V N, Perel V I, Yassievich I N Nonradiative Recombination in Semiconductors (Amsterdam: North-Holland, 1991); Bezyzluchatel'naya Rekombinatsiya v Poluprovodnikakh (St. Petersburg: Izd. PIYaF, 1997)
- 36. Hampton F L, Cricchi J R Appl. Phys. Lett. 35 802 (1979)
- 37. Gritsenko V A, Meerson E E Mikroelektronika 12 580 (1983)
- 38. Gritsenko V A et al. *Mikroelektronika* **16** 42 (1987)
- 39. Gritsenko V A, Meerson E E, Morokov Yu N *Phys. Rev. B* 57 R2081 (1998)
- 40. Nasyrov K A et al. *JETP Lett*. **77** 385 (2003) [*Pis'ma Zh. Eksp. Teor. Fiz.* **77** 455 (2003)]
- 41. Hielscher F H, Preier H M Solid-State Electron. 12 527 (1969)
- 42. Quast W, Schiek B Electron. Lett. 5 485 (1969)
- Ginovker A S, Gritsenko V A, Sinitsa S P Mikroelektronika 2 283 (1973)
- 44. Gritsenko V A, Meerson E E Phys. Status Solidi A 62 K131 (1980)
- Ginovker A S, Gritsenko V A, Sinitsa S P Phys. Status Solidi A 26 489 (1974)
- 46. Fowler R H, Nordheim L Proc. R. Soc. Lond. A 119 173 (1928)
- 47. Murphy E L, Good R H (Jr.) Phys. Rev. 102 1464 (1956)
- 48. Lenzlinger M, Snow E H J. Appl. Phys. 40 278 (1969)
- 49. Weinberg Z A, Johnson W C, Lampert M A J. Appl. Phys. 47 248 (1976)
- 50. Petrin A B JETP 109 314 (2009) [Zh. Eksp. Teor. Fiz. 136 369 (2009)]
- 51. Roberts G G, Polanco J I Phys. Status Solidi A 1 409 (1970)
- 52. Jensen K L J. Appl. Phys. **102** 024911 (2007)
- 53. Houssa M et al. J. Appl. Phys. 87 8615 (2000)
- 54. Nasyrov K A et al. J. Appl. Phys. 105 123709 (2009)
- Shklovskii B I, Efros A L Electronic Properties of Doped Semiconductors (Berlin: Springer-Verlag, 1984) [Translated from Russian: Elektronnye Svoistva Legirovannykh Poluprovodnikov (Moscow: Nauka, 1979)]
- 56. Vishnyakov A V et al. Solid-State Electron. 53 251 (2009)
- 57. Hill R M Philos. Mag. 23 59 (1971)
- Ganichev S D, Prettl W, Yassievich I N Phys. Solid State 39 1703 (1997) [Fiz. Tverd. Tela 39 1905 (1997)]

- 59. Nasyrov K A, Gritsenko V A J. Appl. Phys. 109 093705 (2011)
- 60. Nasyrov K A et al. IEEE Electron Dev. Lett. 23 336 (2002)
- 61. Nasyrov K A et al. J. Appl. Phys. 96 4293 (2004)
- Gadiyak G B, Obrekht M A, Sinitsa S P Mikroelektronika 14 512 (1985)
- 63. Svensson C M J. Appl. Phys. 48 329 (1977)
- 64. Gritsenko V A, Meerson E E, Sinitsa S P *Phys. Status Solidi A* **48** 31 (1978)
- Gritsenko V A "Electronic structure and optical properties of silicon nitride", in *Silicon Nitride in Electronics* (Amsterdam: Elsevier, 1988) p. 138 [Translated from Russian: in *Nitrid Kremniya v Elektronike* (Exec. Ed. A V Rzhanov) (Novosibirsk: Nauka, 1982)]
- 66. Gritsenko V A et al. *Microelectron. Eng.* **86** 1866 (2009)
- Gritsenko V A Phys. Usp. 55 498 (2012) [Usp. Fiz. Nauk 182 531 (2012)]
- 68. Park N-M et al. Phys. Rev. Lett. 86 1355 (2001)
- 69. Park N-M, Kim T-S, Park S-J Appl. Phys. Lett. 78 2575 (2001)
- Gritsenko V A et al. JETP 98 760 (2004) [Zh. Eksp. Teor. Fiz. 125 868 (2004)]
- 71. Voskoboinikov V V et al. *Mikroelektronika* **5** 369 (1976)
- 72. Kolodzey J et al. IEEE Trans. Electron Dev. 47 121 (2000)
- 73. Specht M et al. Appl. Phys. Lett. 84 3076 (2004)
- 74. Ohta K, Hamano K Jpn. J. Appl. Phys. 11 546 (1972)
- Novikov Yu N, Gritsenko V A, Nasyrov K A JETP Lett. 89 506 (2009) [Pis'ma Zh. Eksp. Teor. Fiz. 89 599 (2009)]
- 76. Novikov Yu N et al. Microelectron. Reliability 50 207 (2010)
- Novikov Yu N, Gritsenko V A, Nasyrov K A Appl. Phys. Lett. 94 222904 (2009)
- 78. Perevalov T V et al. J. Appl. Phys. 108 013501 (2010)
- Pustovarov V A et al. *JETP* 111 989 (2010) [*Zh. Eksp. Teor. Fiz.* 138 1119 (2010)]
- 80. Pustovarov V A et al. Thin Solid Films 519 6319 (2011)