The injection of holes from silicon through silicon oxide (SiO₂) in a tantalum nitride—aluminum oxide—silicon nitride—silicon oxide (TANOS) structure has been studied experimentally. Using the high-permittivity Al₂O₃ insulator as a blocking one suppresses the parasitic injection of electrons from the conducting TaN contact. This allows the injection of holes from the substrate into nitride to be studied up to comparatively high electric fields. The experimental data are not described by the standard Fowler—Nordheim law with reasonable physical parameters. At the same time, these data are in good agreement with the model of trap-assisted tunneling hole injection in SiO₂. The developed theory shows that the traps in a narrow energy band make a major contribution to this process, i.e., this injection is resonant in nature.

PACS numbers: 72.20.-i, 72.20.Jv

DOI: 10.1134/S1063776109110089

1. INTRODUCTION

Amorphous silicon oxide (SiO₂) and nitride (Si₃N₄) are the two key insulators in silicon devices. The electronic structure and charge transport in these materials are difficult to study theoretically, because band calculations are inapplicable to these (amorphous) materials. This leads to problems with the determination of the band gap, the energies of the highest occupied (HOMO) and lowest unoccupied (LUMO) orbitals but also bonding Si₃N₄. Effective mass calculations of the electronic structure, both semiempirical [6] and nonempirical [7] ones, also suggest the presence of bonding Si₃N₄, p–O₂p orbitals forming a fairly wide band. “Light” holes with an effective mass $m^* \approx m_0$ must correspond to the wide band of bonding orbitals.

The height of the barrier for the injection of electrons at the Si/SiO₂ interface has been determined reliably, 3.1 eV [8]. According to various experimental data, the band gap $E_g$ for SiO₂ lies within the range 5.0–10.6 eV. $E_g = 10.6$ eV is given, for example, in Mott’s well-known book [9]. Thus, at a silicon band gap of 1.12 eV, the height of the barrier for holes at the Si/SiO₂ interface is 0.8–6.4 eV. The holes at the Si/SiO₂ interface are injected by the tunneling mechanism [10, 11]. Recently, however, the injection current density in an electric field of 10⁷ V cm⁻¹ lies within the range $10^{30}–10^{33}$ A cm⁻². Clearly, the height of the hole barrier at the Si/SiO₂ interface gives a great uncertainty in the estimate of the hole injection current. Thus, for example, for this uncertainty in the barrier height at an effective hole mass $m^* \approx 0.5m_0$, the injection current density in an electric field of 10⁷ V cm⁻¹ lies within the range $10^{7}–10^{10}$ A cm⁻². Clearly, the height of the hole barrier at the Si/SiO₂ interface and the hole effective mass in SiO₂ require a refinement.

The electrons and holes are believed to be injected from silicon into SiO₂ by the Fowler—Nordheim tunneling mechanism [10, 11]. Recently, however, the injection of electrons from silicon into SiO₂ has been shown to be described more accurately by the trap-assisted tunneling (TAT) mechanism [12–17]. As regards the injection of holes, there is no understanding of the injection mechanism here.

Amorphous silicon nitride has the property of localizing (trapping) the electrons and holes injected into it with a great confinement time (about 10 years) in a localized state at 85°C [18]. This localization (memory) effect is used in developing flash memory. Thus, for example, the development of 32-Gbyte flash...
memory based on the localization effect in silicon nitride has been reported [19]. Information is written in such a memory element (a storing field-effect transistor) by the injection of electrons from silicon into silicon nitride through an insulating SiO₂ layer followed by the trapping of electrons and is erased by the injection of holes from silicon through SiO₂ followed by the trapping of holes in silicon nitride. Thus, refining the mechanism of hole injection from silicon into SiO₂ and refining the barrier heights and the hole effective mass are of considerable interest not only in studying the electronic structure of the Si/SiO₂ interface and charge transport but also in developing modern devices with flash memory.

The goal of this paper is to experimentally and theoretically study the mechanism of hole injection from silicon through silicon oxide. We have found for the first time that the holes in strong electric fields are injected through trap-assisted tunneling in SiO₂. Previously, this hole injection mechanism has not been observed experimentally and has not been described theoretically. We derive relations between the height of the hole barrier at the Si/SiO₂ interface and the hole effective mass \( m^* \). The hole trap energy and the trap density in SiO₂ are estimated.

2. THE EXPERIMENT

We experimentally studied metal–insulator–semiconductor (MIS) structures with a three-layer insulator: TaN–Al₂O₃–Si₃N₄–SiO₂–Si (TANOS). The substrate was (100)-oriented \( p \)-type silicon with a resistivity of 20 Ohm cm. After standard purification, 3.5-nm-thick nitrided thermal oxide was grown on silicon. A 7.0-nm-thick layer of amorphous silicon nitride (Si₃N₄) was deposited on silicon oxide in a low-pressure reactor from a mixture of dichlorosilane (SiH₂Cl₂) and ammonia (NH₃) at a temperature of 750°C. A 16-nm-thick layer of aluminum oxide was deposited on the layer of silicon nitride from a mixture of trimethylaluminum ((AlCH₃)₃) and water vapors (H₂O) by the atomic layer deposition method. A 25-nm-thick TaN layer was used as the conducting electrode. The thicknesses of the dielectric layers were measured on control silicon wafers by a single-wavelength ellipsometer at the helium–neon laser wavelength \( \lambda = 6328 \) Å.

In the experiment, the flat-band potential \( V_{fb} \) for the MIS structure was measured as a function of the amplitude and duration of the applied voltage pulse. The flat-band potential is the voltage that must be applied to the MIS structure for the bands in silicon to become flat. To a first approximation, the flat-band potential is proportional to the charge accumulated in the insulator. In our numerical model, \( V_{fb} \) was calculated exactly. The flat-band potential was experimentally determined by measuring the capacitance–voltage (C–V) characteristic at a frequency of 100 kHz.

The energy diagram for the TANOS structure is shown in Fig. 1a. Figure 1 presents the silicon oxide band gap \( E_g = 8.0 \) eV as the most reliable value, in our view, for thermal silicon oxide [8]. For a positive potential at the (TaN) gate, the electrons are injected from silicon through silicon oxide and are trapped in silicon nitride (information writing), Fig. 1b. For a negative potential at TaN, the holes are injected from silicon through silicon oxide and are trapped in silicon nitride (information erasing), Fig. 1c. Here, we studied the injection of holes from silicon into SiO₂ at a negative potential of various amplitudes and recorded the time dependence of the flat-band potential \( V_{fb} \) for the structure. A fresh specimen was chosen for each voltage without any preinjected electrons. Thus, we eliminated the change in \( V_{fb} \) due to the ionization of electrons from silicon nitride and recorded only the injection of holes from the semiconductor substrate into silicon nitride through the layer of silicon oxide.

We compared the experimental results with the results of our simulations using a model in which we calculated the injection of holes from silicon followed by their trapping in silicon nitride based on the Shock-
The TANOS + capacitor system was chosen in such a way that the same voltage as that in the preceding experiment was applied to TANOS. In accordance with the assumption made, the experimental curves $V_{fb}(t)$ were expected to be shifted leftward, toward shorter times. However, the positions of the onset of $V_{fb}$ growth virtually coincided in the two experiments. This indicates that the entire charge injected into TANOS is trapped here and there is no through current.

3. ANALYSIS OF THE EXPERIMENTAL DATA

The subsequent analysis of the experimental data is based on the following. Assuming that in the initial period of charge injection into TANOS, at least until $V_{fb}$ has shifted greatly from its initial value, the current is constant and is described by an FN–like law

$$J_{FN} = J_0 \exp \left( -\frac{4}{3} \frac{\Phi^2}{eFh} \right).$$

Here, $\Phi$ is the height of the hole barrier at the Si/SiO$_2$ interface, $m^*$ is the hole effective tunneling mass in SiO$_2$, $F$ is the electric field strength in SiO$_2$, and $J_0$ is the pre-exponential factor that for a metal–insulator contact is given by the expression

$$J_0 = e^2 F^2 / 8 \pi h \Phi m^*.$$ 

The typical value of this factor for FN injection from a metal and a semiconductor is

$$J_0 \approx 10^7 \text{ A cm}^{-2}.$$ 

Let us specify some shift in $V_{fb}$ (in Fig. 2, this shift is shown in the form of a horizontal dashed line, $V_{fb} = -1$ V). Some charge accumulated in silicon nitride corresponds to this shift. Using Eq. (1), we will find that the time $\tau$ in which this charge is injected satisfies the equation

$$\ln \tau = \frac{4}{3} \frac{\Phi^2}{eFh} \frac{d_{eff}}{d_{AO}} + \ln \frac{CV_{fb}}{J_0}.$$ 

Here, $U$ is the amplitude of the applied voltage, $d_{eff}$ is the TANOS effective thickness recalculated to the dielectric constant $\varepsilon_0$ for SiO$_2$.

$$C = \frac{\varepsilon_0}{d_{N}/\varepsilon_N + d_{AO}/\varepsilon_{AO}}$$

is the capacitance of the TANOS layer that consists of silicon nitride of thickness $d_{N}$ with permittivity $\varepsilon_N$ and aluminum oxide of thickness $d_{AO}$ with permittivity $\varepsilon_{AO}$.

The experimentally measured values of $\tau$ are presented in Fig. 3. We see that these data points fall nicely on a straight line in $(\ln \tau, 1/U)$ coordinates. By fitting based on Eq. (4), we can determine the parameters

$$J_0 \approx 5 \times 10^3 \text{ A cm}^{-2}.$$
Thus, $\Phi$ and $m^*$ cannot be determined independently.

If we take the universally accepted value of 3.8 eV as $\Phi$, then we will obtain $m^* \approx 0.2 m_0$ for the hole effective mass, which is considerably lower than that reported in other works on determining the hole effective mass in SiO$_2$ [23]. If, however, we choose a typical hole effective mass $m^* = 0.5 m_0$, then we will obtain $\Phi = 2.85$ eV for the barrier, which is lower than the universally accepted value of 3.8 eV by 1 eV. The derived pre-exponential factor $J_0$, which turned out to be lower than its typical value for FN injection by three orders of magnitude, causes even greater bewilderment.

Thus, we see that, on the one hand, the FN-like dependence (1) describes well the experimental data, but, on the other hand, the physical parameters turn out to be lower than its typical value for FN injection by three orders of magnitude, causes even greater bewilderment.

Thus, the FN-like dependence (1) describes well the experimental data, but, on the other hand, the physical parameters turn out to be lower than the expected ones. This contradiction can be resolved if the holes are assumed to be injected not by the FN mechanism (hole tunneling from the silicon valence band through the silicon oxide band gap into the silicon nitride valence band) but by trap-assisted tunneling in the bulk of SiO$_2$. In other words, the holes are injected stepwise: the holes are tunneled from the silicon valence band into traps in the bulk of SiO$_2$ and from there into the valence band (Fig. 4).

### 4. THE THEORY OF TAT HOLES THROUGH SiO$_2$

To derive the equation for the current of TAT holes, let us write the equation for the filling of traps in SiO$_2$:

$$\frac{\partial}{\partial t} p(z) = \text{Inj}_s(z) [ P(z) - p(z)] - [\text{Ion}_s(z) + \text{Ion}_n(z)] p(z).$$

(8)

Here, $P(z)$ is the volume distribution of the hole trap density in SiO$_2$, $p(z)$ is the density of the filled traps, $\text{Inj}_s$ is the rate of filling of the traps with holes from the substrate, $\text{Ion}_s$ and $\text{Ion}_n$ are the trap ionization rates into the substrate and the silicon nitride valence band, respectively. A general expression for the ionization of a trap in an insulator into the free volume of a semiconductor was derived in [24]. For a wide range of cases, the trap ionization rate can be represented as

$$\text{Ion}^0_{s,n} = \frac{V_{\text{out}}}{2z} T_S,$$

(9)

where $V_{\text{out}}$ is the hole velocity in the substrate valence band for an energy equal to the energy of the trapped hole, $z$ is the distance from the trap in SiO$_2$ to the substrate, and $T_S$ is the tunneling factor. The latter is calculated in the WKB approximation and for a trapezoidal barrier is

$$T_S = \exp \left\{ \frac{4(2m^*)^{1/2}}{3} \left[ (W + eFz)^{1/2} - W^{1/2} \right] \right\},$$

(10)

However, not all of the states in the substrate are free and, hence, the following expression should be used for the ionization rate:

$$\text{Ion}^0_{s,n} = \frac{V_{\text{out}}}{2z} T_S \left\{ 1 - \left[ 1 + \exp \left( \frac{E_F - E_t}{kT} \right) \right]^{-1} \right\},$$

(11)

where $E_F$ is the Fermi energy for the substrate and $E_t$ is the position of the trap energy level. The energy here is counted as for electrons: it increases from the valence band to the conduction band.

The injection coefficient $\text{Inj}_s$ can be found from the following considerations. Assume that the holes from...
traps in SiO₂ can be ionized only back into the substrate. At thermodynamic equilibrium, it then follows from Eq. (8) that
\[ p_t = \frac{\text{Inj}_S}{\text{Inj}_S + \text{Inj}_T} P_t. \]  
(12)

On the other hand, the equilibrium filling of traps must correspond to Fermi statistics with the substrate Fermi energy:
\[ p_t = P \left[ 1 + \exp \left( \frac{E_F - E}{kT} \right) \right]^{-1}. \]  
(13)

Comparing Eqs. (12) and (13), given (11), we obtain
\[ \text{Inj}_S = \text{Inj}_S \exp \left( \frac{E_F - E}{kT} \right) \]
\[ = \frac{V_{\text{out}} T_S}{2 \pi} \left[ 1 + \exp \left( \frac{E_F - E}{kT} \right) \right]^{-1}. \]  
(14)

For the trap ionization rate into the silicon nitride valence band, we will use the expression
\[ \text{Ion}_N = \frac{W}{2 \pi \hbar} T_N, \]  
(15)

where \( W \) is the depth of the level in the trap (see Fig. 4), \( T_N \) is the tunneling factor for the hole release from the trap into the silicon nitride valence band. For a trapezoidal barrier, we have
\[ T_N(z) = \exp \left\{ - \frac{4}{3} \right\} \]
\[ \times \left\{ 1 - \frac{(2m^*)^{1/2}}{eFh} \left( \frac{W^{3/2} - (W - eF(D - z))^{3/2}}{eFh} \right) \right\}, \]  
(16)

where \( D \) is the SiO₂ thickness. For a triangular barrier, Eq. (16) should be used without the last term in square brackets.

The trap-assisted injection current is
\[ J_{\text{TAT}} = e \int \text{Ion}_N(z) p_t(z) dz. \]  
(17)

From the stationary equation (8), we will obtain
\[ J_{\text{TAT}} = e \int_0^D \frac{\nu_S \nu_N T_S(z) T_N(z) \nu_N T_N(z)}{\nu_S T_S(z) + \nu_N T_N(z)} \]
\[ \times P_s(z) \left[ 1 + \exp \left( \frac{E_F - E_s(z)}{kT} \right) \right]^{-1} dz, \]  
(18)

\[ \nu_S = \frac{V_{\text{out}} T_S}{2 \pi}, \quad \nu_N = \frac{W}{2 \pi \hbar}. \]

Here, \( E_s(z) \) is the local position of the trap energy level,
\[ E_s(z) = E_s + \Delta - \Phi + W + eFz, \]  
(19)

\( E_s \) is the energy of the top of the substrate valence band and \( \Delta \) is the shift of the flat bands at the Si/SiO₂ interface (see Fig. 4). The hole velocity in the substrate is \( v_{\text{out}} = 0 \) if the trap energy is projected onto the semiconductor band gap \( (E_s > E_F + \Delta) \) and
\[ V_{\text{out}} = \left[ \frac{2(E_s + \Delta - E_F)}{m^*_{\text{Si}}} \right]^{1/2} \]  
(20)
in the opposite case. Here, \( m^*_{\text{Si}} \) is the hole effective mass in the silicon valence band.

Expression (18) for the injection current can be generalized for the distribution of traps in energy \( W \). For this purpose, we will introduce the density of traps in space and energy, \( \Pi(z, W) \). Instead of Eq. (18), we should then use
\[ J_{\text{TAT}} = e \int_0^D \frac{\nu_S \nu_N T_S(z) T_N(z) \nu_N T_N(z)}{\nu_S T_S(z) + \nu_N T_N(z)} \]
\[ \times \Pi(z, W) \left[ 1 + \exp \left( \frac{E_F - E_s(z)}{kT} \right) \right]^{-1} dW, \]  
(21)

where
\[ G(E_s, W) = \frac{\nu_S \nu_N T_S(z) T_N(z) \nu_N T_N(z)}{\nu_S T_S(z) + \nu_N T_N(z)} \]  
(22)
is the TAT rate per trap. For an arbitrary distribution of hole traps in SiO₂ volume and in energy, the current can be calculated using Eq. (21) only numerically. However, in some cases, this expression can be simplified. The point is that the function \( G(E_s, W) \) has a sharp maximum at
\[ v_S T_S(z) = v_N T_N(z), \]  
(23)
when the tunneling rates from the substrate into the trap and from the trap into the silicon nitride valence band are equal. If we assume, for simplicity, that the field in oxide or its thickness are so large that the trap is ionized into silicon nitride through a triangular barrier, then the maximum of the function \( G(E_s, W) \) is reached at the trap energy
\[ W_m = 2^{-2/3} \Phi, \quad \Phi = \Phi + E_s + eF - \Delta. \]  
(24)

Near this maximum, the dependence of function (22) on the trap energy can be approximated by the expression
\[ G(E_s, W) = v_S \left[ 2 \cosh \left( \frac{W - W_m}{\delta W} \right) \right]^{-1} \]
\[ \times \exp \left( \frac{2(2m^*)^{1/2} \delta W}{3 eF} \right), \]  
(25)

where
\[ \delta W = \frac{eFh}{2(2m^*)^{1/2} W_m}. \]

The quantity \( \pi \delta W \) characterizes the width of this maximum. For example, for \( m^* = 0.5 m_0, W_m = 2 \mathrm{eV}, \) and \( F = 10^7 \mathrm{V \ cm}^{-1}, \) we have \( \pi \delta W = 0.3 \mathrm{eV}, \) i.e., the maxi-
For a highly doped $p$-type semiconductor, the Fermi energy is close to the energy of the top of the substrate semiconductor valence band and the bending of the bands in enrichment is insignificant. Under these conditions, instead of Eq. (28), we can use

$$J_{\text{TAT}} = \frac{\pi e \delta W^2}{2 e^2 F}$$

In addition, we will consider the case where

$$E_v + \Delta > E_F.$$  \hspace{1cm} (27)

Equation (29) was derived by assuming a triangular barrier for the tunneling of holes from the substrate into the silicon nitride valence band. However, it can be easily generalized to the case of a trapezoidal barrier:

$$J_{\text{TAT}} = \frac{\pi e \delta W^2}{2 \sqrt{2 m^* \Phi}}$$

Equation (30) shows that for a trapezoidal barrier, traps with a larger depth than that for a triangular barrier are involved in TAT.
5. DISCUSSION

Note once again that the exponent in Eqs. (29) and (30) is smaller than that for FN injection (1) by a factor of 2. If the assumption about a fairly smooth energy distribution of traps in SiO₂ is valid, then it follows from Eq. (29) that traps with an energy near

\[ W \approx 0.6\Phi \]  

make a major contribution to TAT for a triangular barrier. If we interpret the experimental results presented in Figs. 2 and 3 based on Eq. (29) (see Fig. 6), then we will obtain

\[ \Phi(m^*/m_0)^{1/3} = 3.6 \text{ eV}, \]  

\[ \frac{\pi e^2}{\hbar^2} \delta W T 1 \left(2^{2/3} \Phi \right) = 5 \times 10^3 \text{ A cm}^{-2}. \]

(32)

(33)

It follows from Eq. (32) that the hole effective mass in SiO₂ is \( m^* \approx 0.85 m_0 \). If we take the universally accepted hole barrier \( \Phi = 3.8 \text{ eV} \), then the barrier height for this hole mass will be \( \Phi = 4.5 \text{ eV} \).

The volume density of the traps in SiO₂ that are resonantly involved in TAT can be estimated from Eq. (33):

\[ P_v = \pi \delta W T 1 \left(2^{2/3} \Phi \right) = 5 \times 10^{16} \text{ cm}^{-3}. \]  

(34)

As we see, a moderately large number of traps is required to ensure effective TAT, with these traps actively involved in TAT being located in a narrow spatial layer,

\[ \delta z = \pi \delta W / eF. \]  

(35)

The center of this layer lies at a distance

\[ z = (\Phi - W_v)/eF = 0.37 \Phi / eF, \]  

(36)

i.e., this layer narrows and approaches the Si/SiO₂ interface as the electric field grows. For example, for the TANOS geometry used in our experiment and at \( U = -20 \text{ V} \), we have the field in oxide \( F = 1.4 \times 10^7 \text{ V cm}^{-1} \) and \( \delta z = 2.2 \times 10^{-3} \text{ cm}, z = 10^{-2} \text{ cm} \). If the traps are assumed to be distributed in energy uniformly, then the total trap density in SiO₂ with a band gap of 8 eV will be \( 1.3 \times 10^{18} \text{ cm}^{-3} \).

6. CONCLUSIONS

Analysis of our experiments on the accumulation of holes in a TANOS structure has shown that the holes are injected into silicon nitride through trap-assisted tunneling in SiO₂. The hole effective tunneling mass in SiO₂ and the height of the hole barrier at the Si/SiO₂ interface cannot be determined independently from analysis of the experimental data. Only the factor \( \Phi(m^*/m_0)^{1/3} \) combining both parameters can be found. If we take the barrier height \( \Phi = 3.8 \text{ eV} \) as the most reliably measured one, then the hole tunneling mass in SiO₂ is \( m^* \approx 0.85 m_0 \). Assuming that the distribution of traps in SiO₂ layer volume and in energy within the SiO₂ band gap is uniform, the density of the traps that provide the observed injection current is \( 1.3 \times 10^{18} \text{ cm}^{-3} \).

Note that the detected TAT effect through hole traps in tunneling silicon oxide speeds up the injection of holes in the TANOS elements of flash memory, i.e., the trap-assisted hole injection in tunneling oxide increases the speed of such memory elements in erasing mode. On the other hand, traps in tunneling oxide can lead to the undesirable draining of a positive charge (holes) through the same TAT mechanism. Studying this phenomenon is beyond the scope of this paper. The hole effective tunneling mass in SiO₂ determined here from our experimental data lies within the range \( m^* = (0.5–0.85) m_0 \). These values agree with \( m^* = 0.45 m_0 \) [20] and are an order of magnitude lower than the hole masses obtained in SiO₂ band calculations [1–4]. It seems possible that the detected light holes in SiO₂ correspond to the bonding Si3p, p–O2p orbitals. Quantum-mechanical SiO₂ band calculations must be performed to confirm this conclusion.

ACKNOWLEDGMENTS

This work was supported by the Siberian branch of the Russian Academy of Sciences (project no. 70) and the Ministry of Science and Technology of the Republic of Korea (National Program of Tera-Level Nanodevice).

REFERENCES

8. V. A. Gritsenko, Composition and Electronic Structure of Amorphous Dielectrics in the Silicon MIS Structures (Nauka, Novosibirsk, 1993) [in Russian].

Translated by V. Astakhov