Solid-State Electronics 52 (2008) 1810–1814

Contents lists available at ScienceDirect

Solid-State Electronics

journal homepage: www.elsevier.com/locate/sse

A quantum mechanical mobility model for scaled NMOS transistors with ultra-thin high-*K* dielectrics and metal gate electrodes

Yanli Zhang *, Marvin H. White

Sherman Fairchild Center, Electrical and Computer Engineering Department, Lehigh University, 16A Memorial Drive East, Bethlehem, PA 18015, United States

ARTICLE INFO

Article history:
Received 9 May 2008
Received in revised form 22 July 2008
Accepted 11 August 2008
Available online 17 October 2008

The review of this paper was arranged by Prof. A. Zaslavsky

Keywords:
Quantum mechanical
Electron mobility
Inversion
High-*K*
Metal gate

ABSTRACT

A quantum mechanical model of electron mobility for scaled NMOS transistors with ultra-thin SiO_2/HfO_2 dielectrics (effective oxide thickness is less than 1 nm) and metal gate electrode is presented in this paper. The inversion layer carrier density is calculated quantum mechanically due to the consideration of high transverse electric field created in the transistor channel. The mobility model includes: (1) Coulomb scattering effect arising from the scattering centers at the semiconductor–dielectric interface, fixed charges in the high-*K* film and bulk impurities, and (2) surface roughness effect associated with the semiconductor–dielectric interface. The model predicts the electron mobility in MOS transistors will increase with continuous dielectric layer scaling and a fixed volume trap density assumption in high-*K* film. The Coulomb scattering mobility dependence on the interface trap density, fixed charges in the high-*K* film, interfacial oxide layer thickness and high-*K* film thickness is demonstrated in the paper.

1. Introduction

CMOS logic and microprocessor technology has steadily and reliably moved along the international technology roadmap for semiconductor (ITRS) following Moore's Law in device performance, chip density, energy efficiency and manufacturing cost for industrial electronic products over the past four decades [1]. At and beyond the 45 nm ITRS technology node, high-*K* materials have been introduced to replace the traditional SiO_2 and SiON as the gate dielectric layer to achieve a reduced equivalent electrical oxide thickness with a larger physical thickness to reduce the gate leakage and improve the device performance. Increased channel doping is concomitant with the scaling of the dielectric layer thickness to improve device performance and reliability, while suppressing short-channel effects. Ultra-thin gate oxide and high channel doping create large transverse electric fields, which cause significant quantization of the carriers in the potential well at the semiconductor–dielectric interface and must be considered for scaled NMOS devices to understand electron transport and mobility within the surface inversion layers.

In this paper, we describe a 2-D quantum mechanical mobility model with consideration to surface roughness and Coulomb scattering in scaled NMOS transistors with ultra-thin SiO_2/HfO_2 gate dielectrics and metal gate electrodes. This model can be used to characterize the experimental *I–V* (current–voltage) behavior of a scaled NMOS transistor with an ultra-thin dielectric layer.

2. Device fabrication

The scaled NMOS transistor that we are going to characterize and model is fabricated on a highly-doped P-type silicon substrate (1.2E18 cm^{-3}). An interfacial oxide layer is employed between the semiconductor and the high-*K* film with a thickness of 0.5 nm to provide a high quality semiconductor–dielectric interface with low interface trap density. An ALD (atomic layer deposition) HfO_2 dielectric of 1.6 nm is deposited on the top of the interface oxide layer. The gate electrode is 10 nm TiN which is covered with polysilicon as shown in Fig. 1. The process temperature is controlled under 520 °C after dielectric layer deposition in order to reduce the gate leakage current.

3. Electron mobility

The total electron mobility in a NMOS transistor at low drain bias can be expressed as

$$\frac{1}{\mu_{TOT}} = \frac{1}{\mu_0} + \frac{1}{\mu_{Ph}} + \frac{1}{\mu_{SR}} + \frac{1}{\mu_c} \quad (1)$$

* Corresponding author. Present address: IBM, Essex Junction, VT, United States. Tel.: +1 802 769 4262.
E-mail address: yaz204@lehigh.edu (Y. Zhang).

0038-1101/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.sse.2008.08.003

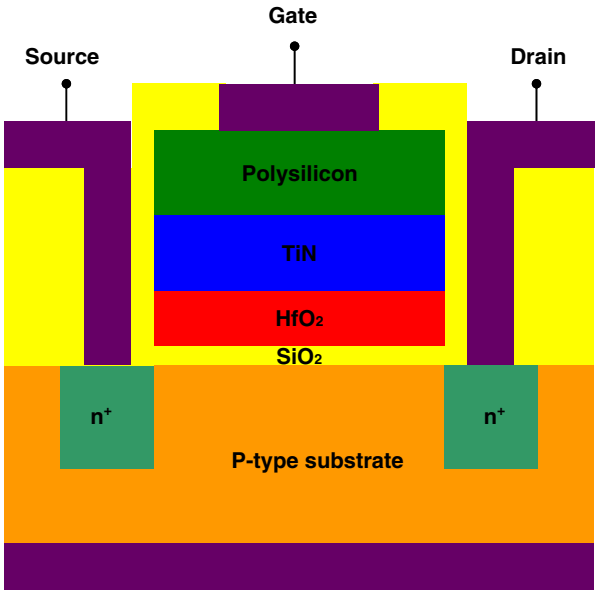

Fig. 1. Cross section of a scaled NMOS transistor with 0.5 nm SiO_2/1.6 nm HfO_2 as the gate dielectric and 10 nm TiN as the gate electrode.

with semiconductor bulk mobility, μ_0, which accounts for the Coulomb scattering and phonon scattering from the semiconductor bulk impurities and lattice vibrations [2]; remote phonon scattering mobility, μ_{ph}, which is related to the intrinsic properties of the high-*K* material; surface roughness mobility, μ_{SR}, which occurs at the Si–SiO_2 interface; and remote Coulomb scattering mobility, μ_c, which includes the scattering from the interface traps and the fixed charges in the high-*K* film. Experimental devices with effective oxide thicknesses less than 1 nm have very high transverse surface electric fields (>1 MV/cm) – even in the subthreshold region of MOS device operation. Film thicknesses and high electric fields and doping densities were not anticipated over a decade. Yamakawa et al. studied semiconductor bulk phonon scattering and surface roughness effects for a device with a substrate doping density of 5E16 cm^{-3} [3]. The bulk phonon scattering (acoustic and inter-valley) contributes to the mobility temperature dependence, which is incorporated into the surface roughness mobility term through the semiconductor bulk mobility, μ_0, in our model (see Section 3.2).

In high-*K* and metal gate devices, remote phonon scattering can be neglected due to metal gate screening [4,5] and the electron mobility expression can be written as

$$\frac{1}{\mu_{TOT}} = \frac{1}{\mu_0} + \frac{1}{\mu_{SR}} + \frac{1}{\mu_c} \tag{2}$$

At low drain bias, the drain current for a NMOS transistor is

$$I_D = \mu_{TOT}\left(\frac{W}{L}\right) Q_n V_D \tag{3}$$

where W and L are the width and length of the transistor respectively, V_D is the drain voltage applied to the device and Q_n is the inversion layer charge density.

3.1. Quantum mechanical analysis for inversion layer charge density

The inversion layer charge density can be obtained with the classical method [6] by treating the energy to be continuous at the surface. However, with continuous scaling of CMOS devices, quantum mechanical effects within the silicon substrate become non-negligible, which leads to significant quantization of the carrier energy and a redistribution of the carriers at the semiconductor and gate dielectric interface (Fig. 2). We calculate the inversion layer charge density with a quantum mechanical approach [7], which entails a variational approach to model electron transport by considering the first three sub-bands, as more than 90% of the surface charge is located in these three bands and this value increases to more than 98% for inversion layer charge densities above 1E12 cm^{-2} [8].

3.2. Quantum mechanical model for surface roughness mobility

Surface roughness scattering is a short-range scattering process arising from interface disorder, which limits the mobility of two dimensional electrons at the Si–SiO_2 interface. If we assume a Gaussian form to describe the positional autocorrelation function for the interface roughness with a correlation length, l_c, and a r.m.s. surface roughness height, Δ, for the random displacement of the interface, then the relaxation time for the surface roughness mobility can be written as [9]

$$\tau_{SR}(E) = \frac{l_c}{\Delta^2} \frac{2^{3/2} m_d^{1/2} E^{3/2}}{q^2 \xi_{eff}^2 \pi^{1/2}} \tag{4}$$

where q is the electron charge, m_d is the density of states mass, E is the energy level, $\xi_{eff} = \frac{|Q_s - Q_n/2|}{\varepsilon_s}$ is the effective electric field, ε_s is the semiconductor permittivity, Q_s is the semiconductor charge density and Q_n is the inversion layer charge density which can be obtained with a quantum mechanical analysis by considering the first three sub-bands. As a function of energy, the surface roughness mobility may be expressed as

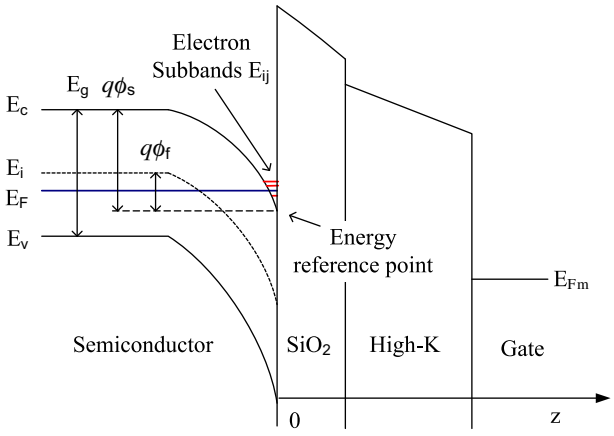

Fig. 2. Energy band diagram for SiO_2/high-*K* structure in the strong inversion region.