Highly Sensitive Bulk Silicon Chemical Sensors with Sub-5 nanometer Thin Charge Inversion Layers

Hossain M. Fahad,1,2,3† Niharika Gupta,1† Rui Han,1 Sujay B. Desai,1,2,3 Ali Javey1,2,3*  

1Electrical Engineering & Computer Sciences, University of California, Berkeley, CA 94720, USA  
2Berkeley Sensor and Actuator Center, University of California, Berkeley, CA 94720, USA  
3Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA  

*Correspondence should be addressed to A.J. (ajavey@berkeley.edu)  
†These authors contributed equally to this work.

Abstract: There is an increasing demand for mass-producible, low-power gas sensors in a wide variety of industrial and consumer applications. Here, we report chemical sensitive field-effect-transistors (CS-FETs) based on bulk silicon wafers, wherein an electrostatically confined sub-5 nm thin charge inversion layer is modulated by chemical exposure to achieve a high sensitivity gas sensing platform. Using hydrogen sensing as a ‘litmus’ test, we demonstrate large sensor responses (> 1000%) to 0.5% H₂ gas, with fast response (< 60 sec) and recovery times (< 120 sec) at room temperature.
and low power (<50µW). Based on these performance metrics as well as standardized benchmarking, we show that bulk silicon CS-FETs offer similar or better sensing performance compared to emerging nanostructured semiconductors while providing a highly scalable and manufacturable platform.

**Keywords**: CS-FET, electrostatic confinement, CMOS gas sensors, tunable sensitivity, low power, charge inversion layer.

In recent years, the micro-hotplate based resistive ceramic sensor has been the dominant commercial technology for miniaturized gas sensing applications. These sensors are made of thick (hundreds of nanometers) films of transition metal oxides, for example ZnO, SnO$_x$ and InO$_x$, that get oxidized or reduced by a target gas at high temperatures.$^{1,2}$ Consequently, this technology suffers from high power consumption requirements ($>>$1 mW). Furthermore, such ceramic films need to be electrically conductive, thereby limiting the choice of metal oxides which can be used for sensing, and contributing to poor selectivity against interfering gases. Despite these drawbacks, major manufacturers continue to develop portable gas sensors based on this technology. Resistive sensors based on metallic nanowires (Pd, Pt) have also shown promise for low-power hydrogen gas sensing.$^{3,4}$ However, their applicability to detect other gases remains mostly undetermined.
Another promising class of sensors is based on functionalized field-effect-transistors (FETs).\textsuperscript{5-9} Chemical sensitive FET (CS-FET) sensors based on low dimensional nano-materials such as carbon nanotubes, silicon nanowires, graphene and transition metal dichalcogenides have shown high sensitivity in detecting a wide variety of gases at room temperature.\textsuperscript{10-24} This is primarily due to i) large surface area to volume ratio and ii) confinement of charge transport in one or two dimensions. Among these materials, pristine single-crystalline silicon is comparatively inert and can respond to specific gases only upon functionalization with appropriate chemical sensitive layers. We recently demonstrated this selectivity advantage using ultra-thin-body (3.5 nm) silicon CS-FETs integrated with different chemical sensing layers (~ 5 nm thin metal alloys) sensitive to specific gases.\textsuperscript{25} While such nanoscale silicon can provide high sensitivity, thickness uniformity control across wafers can lead to process, cost and yield complexities in large-scale manufacturing. On the other hand, conventional bulk silicon transistors do not have a physically thin channel resulting in less susceptibility to such complexities, and can be manufactured very economically.

In this work, we demonstrate bulk silicon CS-FETs as a highly sensitive low power gas sensing platform. The well-established concept of few nanometer thin inversion layers in conventional MOSFETs is adopted here using proper device architecture and operating voltage conditions. Through the electrostatic confinement of the inversion layer, the capacitive coupling between the sensing layer and channel is maximized, enabling high
detection sensitivity. To evaluate this platform, hydrogen gas sensing is used as the test application. Monitoring hydrogen leaks is becoming increasingly important in several applications, requiring stable sensors that can detect below the lower explosion limit of 4% (v/v in air) at low power, low cost and in a very small form factor.\textsuperscript{26}

**Device description**

Conceptually, bulk silicon CS-FETs are similar to conventional enhancement-mode silicon transistors with the exception of the electrically active gate stack that is replaced by a large surface area, ultra-thin chemical sensing layer as depicted in Fig. 1. This sensing layer is electrically floating and can be engineered to be sensitive to a target gas, where interactions can lead to reversible changes in work-function and/or morphology. For the purpose of this work, ultra-thin sensing layers composed of Ni (0.3 nm) and Pd (1 nm) is used for H\textsubscript{2} gas sensing, where H\textsubscript{2} readily dissociates over Pd at room temperature into atomic hydrogen leading to the formation of PdH\textsubscript{x}.\textsuperscript{27} CS-FETs are configured as n-type transistors with light p-body doping (~ 8e14 boron atoms cm\textsuperscript{-3}) and the electrically floating sensing layer is capacitively coupled to the silicon channel via the native oxide (effective oxide thickness, EOT, of 2.5 to 3 nm). The sensitivities of these sensors are dependent on the threshold voltage of the transistors.

Under equilibrium and ambient conditions, the CS-FET threshold voltage ($V_t$) is determined by the body doping and the effective work-
function (EWF) of the sensing layer, which for ultra-thin Ni-Pd is expected to be lower (~ 4.2 eV) than bulk values (~ 5.11 eV) due to work-function dependence on metal thickness. If the $V_t$ is sufficiently low, an inversion layer of electrons (transistor channel) is created at the Si/SiO$_2$ interface. The total electron density and thickness of the inversion layer (which is directly dependent on $V_t$) can then be controlled by applying a reverse-bias to the silicon substrate (also called body), as depicted in Fig. 1. In conventional transistors, this mechanism of $V_t$-control is called the “body-effect”, where an appropriate body-voltage ($V_{SUB}$) effectively controls the p-n junction formed between the p-body and the n-inversion-layer. From a sensors perspective, this provides a highly tunable operation for CS-FETs, where the device can be tuned to the optimal performance by using $V_{SUB}$. Eq. 1 describes the relation between applied body bias and threshold voltage:

$$\Delta V_t = \sqrt{\frac{2 \varepsilon_s q N_{SUB}}{C_{ox}} V_s}$$

where, $\varepsilon_s$ is the dielectric permittivity of silicon, $q$ is the electron charge, $N_{SUB}$ is the body doping, $C_{ox}$ is the capacitance of the native oxide, $V_s$ is the source voltage (ground) and $V_{SUB}$ is the applied body bias. 

RESULTS AND DISCUSSION

Device modeling and simulation

A physical understanding of the sensing mechanism in bulk silicon CS-FET gas sensors is presented in Fig. 2 using TCAD (Synopsys v.2016) device
modeling and simulations. Table 1 lists the parameters used in simulating the n-type transistors where the sensing layer work-function is set at 4.2 eV. Details on device modeling and simulation are in the methods section.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate length ($L_g$)</td>
<td>3 µm</td>
</tr>
<tr>
<td>Effective oxide thickness (EOT)</td>
<td>3 nm</td>
</tr>
<tr>
<td>Source/drain doping</td>
<td>1e20 cm⁻³, Phosphorus</td>
</tr>
<tr>
<td>Body doping ($N_{sub}$)</td>
<td>8e14 cm⁻³, Boron</td>
</tr>
</tbody>
</table>

Table 1. CS-FET device parameters.

As it can be seen in Fig. 2a, at $V_{SUB} = 0V$, a simulated peak electron density of 2e16 cm⁻³ is observed at an inversion layer depth of 3.3 nm with a total inversion layer thickness ($T_{inv}$) of 17.4 nm. $T_{inv}$ is extracted as the location of the charge centroid in Fig. 2a, and the electron density distribution is extracted across the mid-point of the CS-FET silicon channel. Applying $V_{SUB} =-4V$, lowers the peak electron density to 0.65e16 cm⁻³ at an inversion layer depth of 2.7 nm and a $T_{inv}$ of 6.3 nm. With higher reverse body bias, the inversion layer is not only thinner but is also pushed closer to the interface of the native oxide and silicon, leading to improved electrostatic control in the channel by the sensing layer. As will be seen in the subsequent sections, this drastically improves sensor response and sensitivity.
Next, we explore various device parameters that can be varied to optimize the charge inversion layer. $T_{\text{inv}}$ is inversely dependent on body doping ($N_{\text{SUB}}$) as shown in Fig. 2b. As $N_{\text{SUB}}$ is increased from 8e14 cm$^{-3}$ to 8e16 cm$^{-3}$ of boron atoms, $T_{\text{inv}}$ (at $V_{\text{SUB}} = -4V$, EOT = 3 nm) is reduced from 6.3 nm to 2 nm. Body doping and charge density in the inversion layer are also inversely related to each other. As indicated in Fig. 2c, increasing $N_{\text{SUB}}$ from 8e14 cm$^{-3}$ to 8e16 cm$^{-3}$ of Boron, leads to a decrease in peak electron density ($n_{\text{electron}}$) from 0.65e16 cm$^{-3}$ to 0.4e11 cm$^{-3}$ ($V_{\text{SUB}} = -4V$, EOT = 3 nm). The effect of oxide thickness on $T_{\text{inv}}$ is minimal as seen in Fig. 2b, where only at low $V_{\text{SUB}}$, EOT reduction leads to minimal decrease in $T_{\text{inv}}$. Reducing the EOT from 5 nm to 1 nm leads to an increase in peak electron density from 0.5e9 cm$^{-3}$ to 1e14 cm$^{-3}$ ($V_{\text{SUB}} = -4V$, $N_{\text{SUB}} = 8e16$ cm$^{-3}$). Based on this discussion, body doping and effective oxide thickness are key optimization knobs to consider in designing a sensitive bulk CS-FET gas sensor. However, it is important to note that $T_{\text{inv}}$, $n_{\text{electron}}$, EOT and $N_{\text{SUB}}$ are intricately linked to each other and that the above discussion provides highly generalized guidelines for CS-FET sensor design based on simulation results.

**Experimental validation of correlation between sensitivity and inversion layer thickness**

Bulk silicon CS-FETs are fabricated using a fully CMOS-compatible, gate-last processing scheme (see methods section and Supplementary Information S1), where the Ni-Pd sensing layer is deposited in the penultimate process step. Following this, the sensor is annealed in $N_2$ at 150°C for one hour. Fig.
3a shows the experimentally measured room-temperature sensor response of a Ni-Pd CS-FET to different concentrations of hydrogen ranging from 0.05% to 0.5% (diluted in dry air) in steps of 0.05%, at different body biases. Details on the measurement apparatus can be found in the methods section. With increasing reverse body bias from 0V to -2V, % sensor response $((I_{\text{peak}} - I_{\text{baseline}})/I_{\text{baseline}})$ to 0.5% H$_2$ concentration increases from 291 % to 1383 %, as indicated in Fig. 3b. Furthermore, sensor linearity is also drastically improved, where the sensitivity increases from 0.04%/ppm to 0.27%/ppm upon changing body biases from 0V to -2V. Here, sensitivity is defined as the slope of % sensor response $((I_{\text{peak}} - I_{\text{baseline}})/I_{\text{baseline}})\) per ppm of hydrogen gas. It is to be noted that variations in processing conditions of the sensing layer, for example, annealing in forming gas instead of N$_2$, can result in high sensor responses (~15000%) (shown in Supplementary Information S2). This is due to formation of large size Pd-clusters providing increased surface area for H$_2$ interaction. However, this negatively impacts sensor response time due to longer H$_2$ diffusion paths. It is also important to note that bare silicon CS-FETs without any sensing layers do not show any response to hydrogen (see supplementary information S3). Fig. 3c shows the sensor response time ($t_{90}$) vs. hydrogen concentration, with minimum and maximum $t_{90}$ ~ 36s (for 0.5% H$_2$) and 196s (for 0.05% H$_2$) respectively. $t_{90}$ is defined as the time taken for the sensor to reach 90% of its peak response value from the baseline current. Varying the body bias appears to have no significant effect on sensor response time. This is expected, as response times are dependent on
the rate at which hydrogen diffuses and adsorbs on the Ni-Pd sensing layer.\textsuperscript{30} Fig. 3d depicts the room temperature recovery times ($t_{10}$) for different hydrogen concentrations, with minimum and maximum $t_{10}$ of $\sim 62$ s (for 0.5% H$_2$) and 679 s (for 0.05% H$_2$) respectively. $t_{10}$ is defined as the time taken for the sensor to recover to 10% of its baseline current from the peak value. Varying the body biases does not change the rate of the desorption reaction but has a dramatic effect on sensor $t_{10}$, with larger reverse biases enabling shorter $t_{10}$. This is primarily due to the different I-V characteristics demonstrated by different reverse biases leading to varied current vs concentration relationships (see supplementary information S4). Additionally, $t_{10}$ can be further reduced by using integrated micro-heaters, which we have demonstrated in the past. It is important to note that silicon CS-FETs will have a temperature dependence, requiring appropriate compensation for harsh environment operation. Results pertaining to this will be described in a future work. In all of the above measurements, the total power consumption of the hydrogen sensors is below 50 $\mu$W, re-affirming bulk silicon CS-FETs as a low power gas sensing platform.

Fig. 4 compares the experimental data to simulation results, where simulated sensor responses are obtained at different body biases for a constant -0.1 eV work-function change in the sensing layer (analogous to a simulated gas exposure of 0.5% H$_2$). As shown, the trend of increasing sensor responses with higher reverse body bias is consistent in both theory and experiment. However, the simulated work-function change does not
capture the interaction between hydrogen and the sensing layer, which may explain the discrepancy in the trends.

**Sensor hysteresis, ambient drift and long-term stability**

Several experiments were performed to gauge sensor hysteresis and long-term stability. Exposing the sensor to cycles of low (0.05%), medium (0.25%) and high concentrations (0.5%) of hydrogen indicates minimal hysteresis in sensor performance as indicated by Fig. 5a. Fig. 5b captures the baseline drift of two sensors ($V_{DS} = 3\text{V}, V_{SUB} = 0\text{V}$) over a period of 5.7 days, where the sensors were measured in ambient air without any gas flow and uncontrolled room humidity. The maximum variation from mean baseline current in both sensors is approximately 10%, indicating stable sensor baselines. Additionally, bare silicon CS-FETs without any sensing layers also exhibit similar stability as depicted in supplementary information S5. Fig. 5c shows the variation in peak sensor response current to a fixed hydrogen concentration for nearly a week, where a sensor is exposed to 0.5% H$_2$ (for 10 minutes) once per day. This measurement was done at room temperature with the relative humidity left uncontrolled and varying between 20%-40%. Based on these results, the CS-FET platform exhibits minimal sensor performance degradation and long-term stability. With respect to sensor selectivity, we previously demonstrated the Ni-Pd sensing layer to be selective to H$_2$S (hydrogen sulfide) and NO$_2$ (nitrogen dioxide) via multiplexed gas sensing gas experiments. Detailed selectivity results of the
bulk silicon CS-FETs in contextually defined applications will be described in a future work.

**Benchmark comparison with emerging low dimensional materials**

Finally, the performance of bulk silicon CS-FETs was benchmarked against emerging materials like carbon nanotubes, MoS$_2$ and graphene for hydrogen gas sensing at the same concentration level (0.5%). Sensitive and fast detection at this concentration is important from a safety perspective, as it is below the lower explosion limit of 4%. This benchmark cites research works that have used both functionalized and non-functionalized materials. As indicated in Fig. 6, bulk silicon outperforms these low dimensional materials in terms of normalized sensor response (%). The results suggest that electrostatic charge confinement can be an effective route towards achieving high sensitivity with potential advantages over structural charge confinement.

**CONCLUSION**

To summarize, we have demonstrated chemical sensitive field effect transistors on bulk silicon, with an electrically floating ultra-thin Ni$_{0.3}$Pd$_{1}$ sensing layer for H$_2$ gas sensing. Through device modeling and simulations, we have shown that by applying different $V_{\text{SUB}}$, the sensitivity of the CS-FET can be tuned electrically. We have corroborated this by measuring the H$_2$ sensor response of fabricated Ni-Pd CS-FETs which results in improved sensor linearity and recovery times. Moreover, this platform exhibits minimal
sensor hysteresis and long-term drift. The results presented in this work build a compelling case for bulk silicon CS-FETs from both performance and manufacturability perspectives. This platform provides opportunities in a wide variety of applications such as industrial safety, environmental air quality monitoring, wireless sensor networks and consumer electronics.

METHODS

CS-FET device modeling and simulation

CS-FET device simulations in Fig. 2 and Fig. 4 were carried using Synopsys TCAD (Version M-2016.12). Carrier transport in devices are handled by self-consistently solving the Poisson`s continuity equation with the drift-diffusion model. The Philips unified model is used for calculating mobility in the devices. Quantum confinement effects associated with nanoscale devices are taken into consideration using the density-gradient based quantization model. The Slotboom and Graaff bandgap narrowing model is incorporated throughout the device. In addition to this, doping dependent Shockley-Reed-Hall recombination model is utilized in conjunction with Hurkx band-band tunneling model.

CS-FET fabrication process

CS-FET gas sensors were fabricated on prime grade silicon <100> wafers with sheet resistivity in the range of 10-20 ohm.cm. A schematic representing the fabrication process is depicted in supplementary information S1. Before processing, all wafers were cleaned in a standard
Piranha (1:4, hydrogen peroxide: sulfuric acid) bath at 120°C and native oxide removed using a 10s dip in 1:10 hydrofluoric acid. First, a 350 nm silicon dioxide was thermally grown on the silicon wafers for device isolation, using a three step dry (5 min)-wet (55 min)-dry (5 min) oxidation process at 1000°C, atmospheric pressure for 55 min. Oxide thickness was verified using fixed angle (70°) ellipsometry. Next, source and drain doping regions in silicon were defined using standard i-line photolithography process (Fujifilm, photoresist: OiR 906-12, developer: OPD-4262) and wet etching the isolation oxide (in 5:1 buffered hydrofluoric acid for 5 minutes). Following this, approximately 300 nm thick phospho-silicate glass (PSG) layer was deposited at 450°C using low pressure chemical vapor deposition (LPCVD). To complete the formation of n++ doped regions, phosphorus drive-in and activation was performed in the silicon source and drain by rapid thermal annealing (RTA) at 1050°C for 30s in N₂. The PSG layer was then removed in a 1:10 hydrofluoric acid bath for 1 minute. This process step involves over-etching that can lead to some loss in field oxide from the original 350 nm, but is inconsequential to the overall device isolation. The “gate” or sensing layer region was patterned next and etched in 5:1 buffered hydrofluoric acid for 4 minutes. To define source and drain contacts, a separate source-drain metallization mask was used, which underlaps the doped source and drain regions by 11 µm. After this, 50 nm of nickel was then deposited in the source and drain contact regions, using thermal evaporation and lift-off in acetone. To achieve ohmic source and drain contacts, nickel silicidation (NiSi) was performed in forming
gas using an RTA at 420°C for 5 minutes. Following this, an ultra-thin Ni-Pd sensing layer was deposited by sequentially evaporating 1 nm Pd (using e-beam) and then 0.3 nm Ni (using thermal), without any vacuum break. Finally, the sensing layer was annealed in N₂ at 150°C for 1-hour post-deposition, which completed the sensor fabrication process.

**Sensor measurement apparatus**

All gas sensing experiments described in this paper were done in a walk-in fume hood. CS-FET device chips were wire bonded to a 28 pin J-bend leaded chip carrier. A small volume (~ 0.83 cm⁻³) 3D printed housing (made of Polyactic acid) consisting of a ¼” gas inlet was used to cover the chip carrier. Pure dry air was used as diluent gas and was procured from Praxair Technology Inc. For H₂ sensing experiments, 5% H₂ in N₂ (Praxair) was used as source. Ultra-high purity H₂ (Praxair) was used for the experiment in supplementary information S2. House compressed dry air was used for week long extended measurements (in Fig. 5c). Typical gas flow rates were from 1-100 sccm and diluent (air) flow rate was approximately 1000 sccm. Ambient temperature and humidity was monitored by commercial sensors purchased from Sensirion AG (Model SHT2x and SHT3x). Gas delivery was controlled by mass flow controllers (Alicat Scientific Inc.). CS-FET sensors were biased using a Keithley 428 current preamplifier and the current signals were acquired using a LabVIEW controlled data acquisition unit (National Instruments, NI USB-6259). A Keysight 4155C semiconductor parameter analyzer was used for extended length ambient drift measurements in Fig.
5b. Measurements in supplementary information S4 were carried out using a different electronic readout setup and data acquisition board (National Instruments, NI-USB 6218).

**ASSOCIATED CONTENT**

**Supporting Information**

The supporting information is available free of charge on the ACS Publications website.

Bulk CS-FET fabrication process; $\text{H}_2$ sensor response characteristics of bulk CS-FETs with forming gas annealed Ni-Pd sensing layer; $\text{H}_2$ sensor response characteristics of bare (un-functionalized) bulk CS-FETs; Analysis of $\text{H}_2$ surface concentration, desorption rates and correlation to sensor recovery at different $V_{\text{SUB}}$; Ambient drift characteristics of bare (un-functionalized) bulk CS-FETs; References.

**Conflict of Interest**

H.M.F and A.J. declare competing financial interests in equity on shares of Serinus Labs, Inc. An invention disclosure on this work has been filed with the University of California, Berkeley.

**AUTHOR INFORMATION**

**Corresponding Author**

*E-mail: Ali Javey (ajavey@berkeley.edu)*
ORCID

Hossain M. Fahad: 0000-0002-6758-5432

Ali Javey: 0000-0001-7214-7931

Author Contributions

H.M.F and N. G. contributed equally to this work. H.M.F led the project and fabricated the CS-FET sensors as well as device modeling and simulation. N.G. and H.M.F. carried out the measurements and analysis. N.G. and R.H. contributed to measurement setup and programming. S.B.D contributed to device analysis. A.J. supervised the project. All authors discussed the results and wrote the paper.

ACKNOWLEDGEMENTS

A. J. acknowledges the Bakar Fellows Program in funding this research work. H.M.F. and N.G. thank H. Kim, D.-H. Lien, A.B. Sachid and M. Amani for fruitful discussions.

REFERENCES


Fig. 1. Cross-sectional schematic of a bulk silicon CS-FET with electrostatic confinement of the charge inversion layer for achieving high sensitivity.
**Fig. 2.** Characteristic of simulated silicon CS-FET devices depicting (a) inversion layer profiles, (b) extracted inversion layer thickness (sensing layer work-function ($\Phi_m$) is set to 4.2 eV) and (c) peak electron density at different body bias. (Note: Inversion layer profiles are extracted across the channel mid-point of the simulated device).
Fig. 3. (a) Experimentally measured current of a Ni-Pd CS-FET in response to different H₂ concentrations and at different body biases (V_DS = 3V, R.H. < 10 %). (b) Extracted sensor response (ΔI/I₀) vs. H₂ concentrations, (c) t₉₀ vs concentration and (d) t₁₀ vs concentration from panel (a) at different V_SUB.
**Fig. 4.** Theory vs. experiment comparing sensor response. (In experiment: \( V_{DS} = 3V \), R.H. < 10 \%, for theory sensing layer work-function \( \Phi_m = 4.2 \text{ eV} \)).
Fig. 5. (a) H\(_2\) concentration cycling for gauging sensor hysteresis of two Ni-Pd CS-FET sensors (\(V_{DS} = 3\)V, \(V_{SUB} = -1.5\)V, R.H. < 10 %). (b) Still ambient drift characteristics of two Ni-Pd CS-FETs with uncontrolled relative humidity (\(V_{DS} = 3\)V, \(V_{SUB} = 0\)V). (c) Sensor baseline current, peak current and response characteristics to multiple H\(_2\) exposure pulses of 0.5% over 6 days (\(V_{DS} = 3\)V).
Fig. 6. Performance benchmark at 0.5% H\textsubscript{2} concentration comparing different H\textsubscript{2} sensors reported in literature based on emerging nanomaterials and bulk silicon (this work).
TOC Figure: