



Cite this: *Nanoscale*, 2015, 7, 8076

Scaling of graphene integrated circuits†

Massimiliano Bianchi,^a Erica Guerriero,^a Marco Fiocco,^a Ruggero Alberti,^a Laura Polloni,^a Ashkan Behnam,^b Enrique A. Carrion,^b Eric Pop^c and Roman Sordan*^a

The influence of transistor size reduction (scaling) on the speed of realistic multi-stage integrated circuits (ICs) represents the main performance metric of a given transistor technology. Despite extensive interest in graphene electronics, scaling efforts have so far focused on individual transistors rather than multi-stage ICs. Here we study the scaling of graphene ICs based on transistors from 3.3 to 0.5 μm gate lengths and with different channel widths, access lengths, and lead thicknesses. The shortest gate delay of 31 ps per stage was obtained in sub-micron graphene ROs oscillating at 4.3 GHz, which is the highest oscillation frequency obtained in any strictly low-dimensional material to date. We also derived the fundamental Johnson limit, showing that scaled graphene ICs could be used at high frequencies in applications with small voltage swing.

Received 16th February 2015,
Accepted 27th March 2015

DOI: 10.1039/c5nr01126d

www.rsc.org/nanoscale

1. Introduction

Scaling of electronic devices has been the dominant trend in modern electronics over the last 50 years as smaller transistors operate faster.¹ Both digital and analog circuits benefit from scaling because a reduction of transistor feature sizes allows for higher data and analog signal bandwidths.² However, scaling also has a detrimental influence on the properties of field effect transistors (FETs), manifested by weakening of the gate control over the channel at the expense of an increased drain control, resulting in short-channel effects.¹ The most promising way of keeping the channel under strong control of the gate is to also scale the devices in the direction perpendicular to the channel, which has mostly been done in the past by reducing the gate oxide thickness. However, due to aggressive scaling, the oxide thickness has been reduced to critically low levels (just a few atomic layers, limited by gate oxide tunneling). Therefore the remaining viable path to continue transistor scaling is to reduce the thickness of the channel itself, typically to $<1/4$ of the channel length.³ This approach has resulted in the development of silicon-on-insulator technology⁴ and a recent commercial push to Si Fin-FETs at the 22 nm technology node.^{5,6}

Two-dimensional (2D) materials such as graphene, MoS_2 , or WSe_2 offer the realization of FETs at few-nanometer gate lengths due to their sub-nanometer, atomically thin channels. What sets graphene apart from other 2D materials and other conventional semiconductors are its large and equal^{7,8} electron and hole mobilities^{9–11} due to its symmetric band structure.¹² Different figures of merit have been used in the past to evaluate to what extent the high mobility of graphene, despite its zero band gap, can be utilized in FETs and how such FETs scale with the gate length L . The most commonly used performance metrics have been the transistor cutoff frequency f_T ¹³ and the maximum frequency of oscillation f_{max} .¹⁴ However, despite the importance of f_T and f_{max} in the characterization of individual high-frequency transistors, they do not reflect the actual speeds at which the realistic ICs operate.¹⁵ ICs run at much lower frequencies because both f_T and f_{max} are measured under highly-idealized conditions in simple single-transistor configurations.† In realistic electronic circuits with several transistor stages connected together, many other factors, *e.g.*, loading by the following stage, interconnects, gate resistances, and mismatch between different stages, limit the highest operating frequency. Moreover, both f_T and f_{max} are valid only in a small-signal regime, while most of the electronic circuits (especially digital) operate in a large-signal regime. The typical figure of merit in a large-signal regime is the intrinsic gate delay CV/I ¹⁶ which has been widely used to demonstrate the progress of Moore's law¹⁷ over the years. The scaling of the Si complementary metal–oxide–semiconductor (CMOS) logic was closely followed by the reduction of the intrinsic gate delay, reaching just below 1 ps at the 22 nm node.¹⁸ However, CV/I delay cannot directly be measured and is also a single-transistor metric which, similarly to f_T and f_{max} ,

^aL-NESS, Department of Physics, Politecnico di Milano, Polo di Como, Via Anzani 42, 22100 Como, Italy. E-mail: roman.sordan@polimi.it

^bElectrical & Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

^cElectrical Engineering, Stanford University, Stanford, CA 94305, USA

†Electronic supplementary information (ESI) available: Discussions on the cutoff frequency f_T , the maximum frequency of oscillation f_{max} , and the intrinsic gate delay CV/I . See DOI: 10.1039/c5nr01126d

does not reflect realistic transistor delays in ICs. For this reason the International Technology Roadmap for Semiconductors (ITRS)¹⁹ decided to use fan-out-of-one (FO1) inverter gate delay τ in ring oscillators (ROs) as the main speed metric of scaled ICs since 2009.²⁰ In the Si CMOS technology this delay is currently $\tau = 2.2$ ps at the 22 nm node.^{18,20}

Here, we demonstrate the scaling of graphene sub-micron ICs which has not been done so far. Only the scaling of single-transistor frequencies f_T and f_{\max} has been demonstrated by collecting data from different sources,^{21,22} and no scaling data are available for CV/I delay. Non-systematic studies on the scaling of graphene ICs are available only for three different gate lengths above 1 μm .²³ We demonstrate the scaling by fabricating and characterizing graphene ROs of nine different gate lengths, from 0.5 to 3.3 μm , while varying the channel width, access length, and lead thickness at the same time. The shortest FO1 gate delay was $\tau = 31$ ps at the gate length $L = 0.9$ μm , which represents the shortest FO1 gate delay demonstrated to date in any low-dimensional material. By comparison, the shortest reported FO1 delays in low-dimensional materials are 104 ps in graphene,²³ 1.9 ns in carbon nanotubes,²⁴ and 62.5 ns in bilayer MoS₂,²⁵ while the shortest delay in conventional thin-film materials is 100 ps, as obtained in 40 nm thick polycrystalline Si thin-films.²⁶ The shortest demonstrated delay in graphene allowed the realization of the fastest low-dimensional oscillators running at 4.3 GHz, exceeding the oscillation frequency of previously realized graphene ROs by more than a factor of three.²³ The oscillations in the investigated ROs were possible because a high voltage gain ($A_v \sim -5$) was obtained in the fabricated sub-micron graphene inverters. High voltage gain resulted in a static noise margin (NM) equal to 17% of the output voltage swing, similar to other transistor technologies. We also derived the fundamental Johnson limit²⁷ for graphene, demonstrating a trade-off between the highest operating frequency and output device power in scaled graphene FETs. Finally, our work highlights the need for further reduction of contact resistance in graphene FETs in order to continue the observed scaling trend of graphene circuits and further improve their noise margin.

2. Experimental

Graphene monolayers were grown by chemical vapor deposition (CVD) on Cu with a CH₄ precursor and transferred to SiO₂ (300 nm)/Si substrates. Conductive channels of widths $W = 5$ μm and $W = 10$ μm were defined by patterning large-area CVD graphene by electron-beam (e-beam) lithography and reactive-ion etching. The source, drain, and gate contacts were patterned by e-beam lithography and deposited in an e-beam evaporator. Top gates composed of Al/Ti/Au (32/1/9 nm) were fabricated first by direct evaporation of Al on graphene, which upon exposure to air formed a very thin (~ 4 nm) native AlO_x gate insulator at the interface with graphene.²⁸ The gates were terminated by Ti/Au during the same evaporation step in order to form Ohmic contacts with the source/drain terminals of the

following inverter stage. Source and drain contacts were made in the following step and consisted of Au (75 nm or 100 nm), without the use of any adhesion layer. Two different contact lengths were used (2.1 μm and 0.7 μm) which did not influence the operation of ROs as long as the lead resistance was the same. At even shorter contact lengths the contact resistance will increase. However, this could be compensated for by increasing the injection of charge carriers through graphene edges (*e.g.*, by etching graphene below the metal contacts).^{29,30}

As fabricated, both FETs in an inverter were identical. Complementary operation was obtained between the Dirac points of the two FETs³¹ which split after the supply voltage $V_{\text{DD}} > 0$ is applied.²³ The inverters exhibited the highest voltage gain at the dc operating point that lies approximately halfway between the Dirac points of the two FETs. The dc mismatch between the input and output voltage of an inverter at the highest gain point was equal to the Dirac voltage V_0 of the unbiased FETs.³² In ROs with relatively large Dirac voltages ($V_0 > 0.1$ V) it was therefore necessary to apply a positive back-gate voltage to shift the Dirac point back to zero and eliminate the in/out mismatch. All measurements were performed at room temperature. ROs which required low (or zero) back-gate voltages were operated in air while ROs which required larger back-gate voltages ($V_{\text{BG}} > 50$ V) were operated in air under N₂ flow (to stabilize the position of the Dirac points at zero).

The exact dimensions of the graphene FETs were characterized using an SEM. We found that the previously reported ROs²³ have gates $\sim 10\%$ longer than those reported (due to the e-beam proximity effect), *i.e.*, the shortest gate length previously demonstrated in graphene ROs was 1.1 μm instead of 1 μm . Here, we additionally investigated five shorter gate lengths: 1, 0.9, 0.8, 0.7 and 0.5 μm .

The voltage signals were measured using an Agilent Infiniium DSOX91304A (bandwidth 13 GHz) digital storage oscilloscope while the parasitic capacitive load of the ROs was minimized by connecting the output of the ROs to the oscilloscope through low-capacitance (< 0.08 pF) active probes.

3. Results and discussion

The fabricated graphene ROs were composed of three inverters (each with two FETs) cascaded in a loop (Fig. 1). The loop makes a RO unstable and therefore induces oscillation if all inverters are identical and exhibit signal matching and voltage gain $|A_v| > 1/\cos(\pi/n)$, where $n \geq 3$ is the odd number of inverters in the loop³³ (here $n = 3$ and $|A_v| > 2$). Since each inverter stage in the loop is both driven and loaded by another stage, the frequency $f_{o,\text{FO1}} = 1/(2n\tau)$ at which ROs oscillate represents the highest operating frequency under realistic conditions (characterized by the absence of driving current/voltage sources, zero or infinite loads, matched impedances, or external unilateralization feedbacks). The fourth inverter was added to the RO to isolate it from the measurement equipment. This reduces the oscillation frequency to $f_o = 0.82f_{o,\text{FO1}}$ as the fourth inverter doubles the load of the third inverter.²³ Measuring f_o

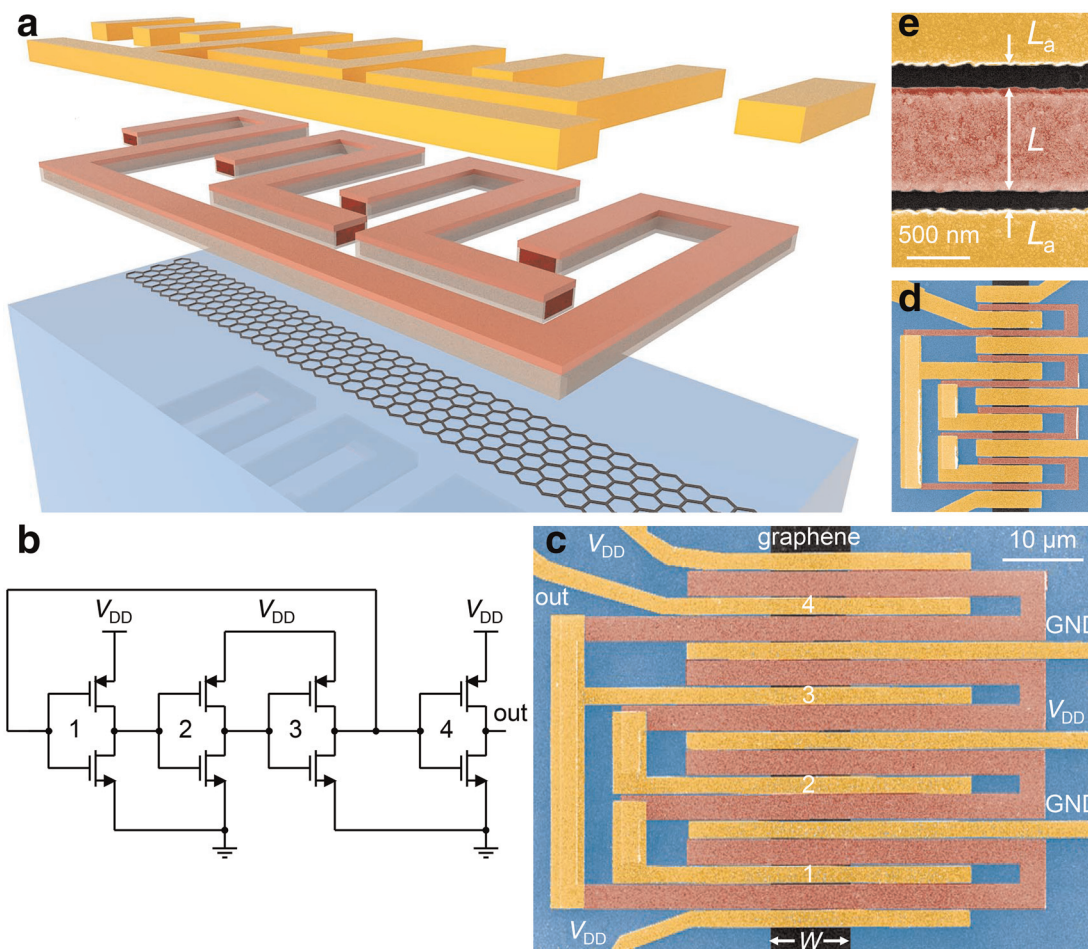


Fig. 1 Integrated graphene ring oscillators (ROs). (a) Exploded view of a buffered three-stage graphene RO. Monolayer graphene channel (black) on an insulating substrate (blue) was defined by reactive-ion etching of large-area graphene in the first step (bottom). The $\text{AlO}_x/\text{Al}/\text{Au}$ gate stack (gray/ruby/red) was patterned in the second step (middle). The Au source/drain contacts (yellow) were patterned in the third step (top). The drains of the previous inverter stage and Au-terminated gates of the next stage were overlapped in order to form Ohmic contacts between the cascaded stages. (b) Circuit diagram of a buffered three-stage RO. The RO is composed of three inverters (1–3) cascaded in a loop with the fourth inverter (4) acting as a buffer. (c) Scanning electron microscopy (SEM) image of a RO with the largest investigated dimensions ($L = 3.3 \mu\text{m}$ and $W = 10 \mu\text{m}$). The false colors correspond to the colors used in (a) and the numbers mark the inverters shown in (b). (d) SEM image of a RO with the smallest investigated dimensions ($L = 0.8 \mu\text{m}$ and $W = 5 \mu\text{m}$) at which it was possible to obtain oscillation. To demonstrate scaling, this image is shown in the same scale as the image in (c). (e) SEM image of an FET in a RO showing a gate of length L and access parts of the channel (the ungated parts of the channel between the gate and source/drain) of lengths L_a .

it is possible to determine $f_{o,FO1}$ and therefore the FO1 gate delay τ . The graphene ROs were laterally scaled by fabricating them from inverters with nine different gate lengths, in the range $0.5 \mu\text{m} \leq L \leq 3.3 \mu\text{m}$. The other two lateral dimensions, channel width W ($10 \mu\text{m}$ and $5 \mu\text{m}$) and access lengths L_a ($0.5 \mu\text{m}$ and $0.25 \mu\text{m}$), were also scaled in order to investigate the influence of parasitics on the oscillation frequency. Similarly to the state-of-the-art Si CMOS circuits, only lateral dimensions were scaled, while the equivalent gate oxide thickness (EOT) was kept constant, $\text{EOT} = \epsilon/C_{\text{ox}} = 2.5 \text{ nm}$, where ϵ is the permittivity of SiO_2 , and $C_{\text{ox}} = 1.4 \mu\text{F cm}^{-2}$ is the capacitance of the top-gate stack in which a thin ($\sim 4 \text{ nm}$) AlO_x layer was used as a gate insulator (see the Experimental section).²³

The measured oscillation frequencies and FO1 gate delays of 65 fabricated ROs are shown in Fig. 2. For a constant channel width W and access length L_a the oscillation frequency was found to scale as $f_o \propto L^{-1}$ (Fig. 2a) in an identical manner to conventional semiconductor ROs.³⁴ Ideally, a square dependence L^{-2} is expected because both channel resistance and geometric gate capacitance scale down with L and therefore the corresponding time constant scales as L^2 . However, the channel resistance is only one part of the total resistance, which contains several additional components that do not scale with L , such as contact resistance and resistance of interconnects and leads. Moreover, the gate resistance increases with decreasing L . Similarly, the capacitance which influences the gate delay also includes parasitic capacitances that do not

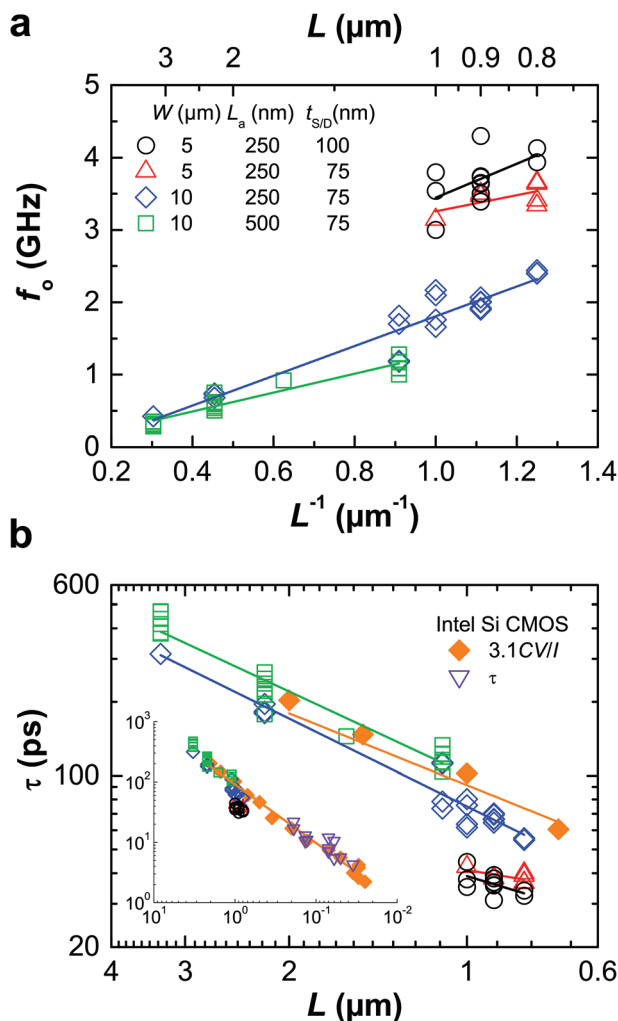


Fig. 2 Scaling of integrated graphene ROs. (a) Oscillation frequency f_o of 65 buffered three-stage ROs as a function of the gate length L . Four different RO layouts (coded in different colors/symbols) were used in order to demonstrate the influence of the channel width W , access length L_a , and thickness t_{SD} of the source/drain metal (*i.e.*, lead resistance) on scaling. (b) FO1 gate delay τ as a function of the gate length L . The delay was calculated from the measured f_o as $\tau = 0.82/(6f_o)$. The FO1 gate delays of Si CMOS ROs made by Intel are plotted as a reference although graphene and Si CMOS circuits cannot directly be compared, see the main text. The delays for all nodes (filled orange rhombuses)¹⁸ were calculated from the CV/I delays provided by Intel as $\tau = 3.1\text{CV}/I$, where the proportionality constant of 3.1 was obtained by fitting CV/I delays with actual FO1 gate delays (open purple triangles) of Si CMOS ROs which Intel provided for some of the nodes.^{34,36,37} Inset: the same plot in the full scale ($25 \text{ nm} \leq L \leq 3.3 \mu\text{m}$).

scale with L , such as gate fringe capacitances and capacitances between the source/drain contacts and the substrate. As a consequence, the gate delay scales slower, typically as $\tau \propto L$ (Fig. 2b) and therefore $f_o \propto L^{-1}$.

At smaller access lengths L_a , the total channel resistance is smaller and therefore the gate delay is smaller, as evidenced by comparing the green/square and blue/rhombus plots in Fig. 2. At smaller channel widths W the gate delays were also

found to be smaller, as evidenced by the blue/rhombus and red/triangle plots in Fig. 2. Although this seems counterintuitive (because channel and contact resistances scale as W^{-1} and geometric and fringe gate capacitances as W) at smaller channel widths the gate resistance is smaller which reduces the gate delay. In addition, there are several other components in the total FET resistance which do not depend on W , *e.g.*, resistance of the leads and interconnects, which also lead to a smaller gate delay at smaller W (due to the reduction of gate capacitances). Finally, the larger thickness of the source/drain leads (t_{SD} up to 100 nm) results in smaller source/drain resistances and therefore smaller delays, as evidenced by comparing the red/triangle and black/circle plots in Fig. 2.

In contrast to oscillation frequency f_o which depends on the number of inverter stages in a RO, the FO1 gate delay is independent of the number of inverters and mostly depends on the channel mobility in otherwise identical ROs.³⁵ For this reason, FO1 gate delay is used to compare different IC technologies and represents the main IC speed metric used by the ITRS.²⁰ Fig. 2b shows FO1 gate delays τ in the fabricated graphene ROs and Si CMOS ROs.^{34,36,37} Although most of the fabricated graphene ROs scale faster than Si CMOS ROs (due to the larger channel mobility of graphene FETs), these two technologies cannot directly be compared because the leakage drain current in graphene ROs is ~ 3 orders of magnitude larger than in Si ROs.²³ Si CMOS is designed for low-power applications which impose the lower limit for the threshold voltage V_{th} of Si MOSFETs. Without this power constraint it would be possible to further reduce V_{th} and therefore reduce the MOSFET on-state resistance $\propto (V_{\text{DD}} - V_{\text{th}})^{-1}$, resulting in shorter gate delays.³⁸ In addition, the scaling trend of Si CMOS ROs has been preserved down to the 22 nm node (physical gate length $L \sim 34 \text{ nm}$),⁵ while in the graphene case no oscillations were observed for physical gate lengths below $L = 0.8 \mu\text{m}$.

The absence of oscillation in the fabricated graphene ROs with $L < 0.8 \mu\text{m}$ can be explained by a loss of voltage gain at short gate lengths. Fig. 3 shows the dc voltage gain in the best performing graphene inverters with $L = 0.8 \mu\text{m}$. The obtained highest gain $A_v \sim -5$ is large enough to satisfy the oscillation criterion and compensate for small in/out signal mismatches at the highest gain points of the inverters in a RO.³² However, at the shorter investigated gate lengths ($0.7 \mu\text{m}$ and $0.5 \mu\text{m}$) the voltage gain of the inverters was reduced to $A_v \sim -2$. This reduction is caused by contact and access resistances which do not scale with gate length and cannot be gated. At short gate lengths, these constant series resistances (here, in the range from $300 \Omega \mu\text{m}$ to $1 \text{ k}\Omega \mu\text{m}$, normalized by the FET width) become comparable to the channel resistance and therefore suppress the voltage gain. Even though the gain of 2 is just enough to support oscillation in three-stage ROs, it is unlikely that all three inverter stages can exhibit such a gain due to fabrication-induced variability. This leads to the absence of oscillation if the voltage gain drops below 2 in just one stage. The constant series resistances also suppress the decrease of the total channel resistance as the

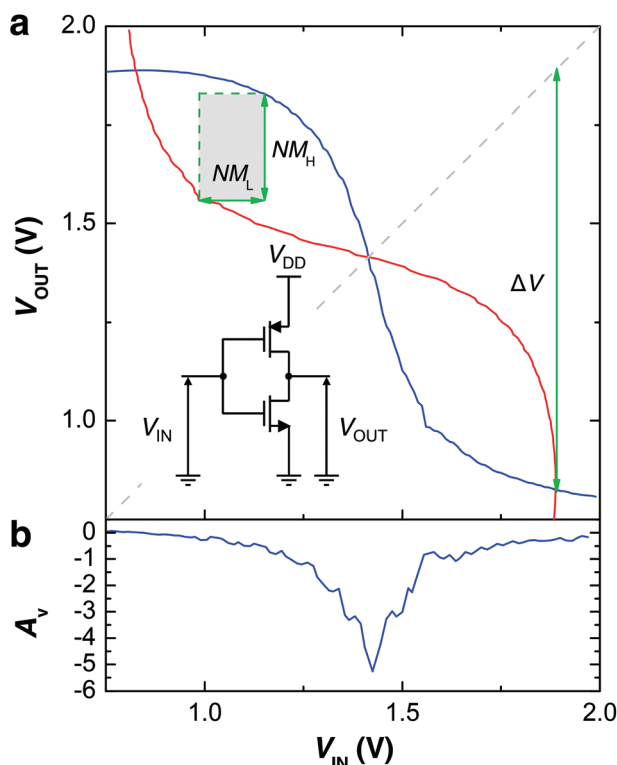


Fig. 3 Voltage gain and NMs in short-gated graphene inverters at $V_{DD} = 2.5$ V. (a) Static voltage transfer characteristic (blue) and its mirrored reflection (red) of an inverter with $L = 0.8$ μm , $W = 5$ μm , and $L_a = 250$ nm. The output voltage swing $\Delta V = 1$ V. The low ($NM_L = 0.17$ V) and high ($NM_H = 0.27$ V) static NMs are determined by the rectangle of the largest area that can be drawn inside the loops. The actual static NM is the smaller of the two, *i.e.*, $NM = NM_L$. Inset: schematic of an inverter. (b) Low-frequency voltage gain obtained as the first derivative of the static transfer curve shown in (a).

graphene FETs are reduced in size, thereby slowing the scaling trend for $L < 1$ μm (Fig. 2). These findings demonstrate the need for alternative contacts if the scaling of graphene ICs is to be continued.

The static voltage transfer characteristics shown in Fig. 3 can also be used to determine the static NM of the fabricated graphene ICs. The static NM based on a maximum product criterion is obtained by maximizing the area of a rectangle inside the inverter transfer characteristic loop³⁹ as shown in Fig. 3. For the present inverters with $L = 0.8$ μm this gives $NM = 0.17$ V at the output voltage swing $\Delta V = 1$ V, *i.e.*, 17% of the swing. At longer gate lengths the voltage gain increases and the static NM reaches $\sim 35\%$ of the voltage swing. Similar NMs are obtained in conventional semiconductor circuits. High-speed InP heterojunction bipolar transistor (HBT) emitter-coupled logic (ECL) gates, which are used in wide bandwidth (>100 GHz) digital and mixed-signal ICs,⁴⁰ have $NM = 1.8V_T$ at $\Delta V \sim 10V_T$, *i.e.*, 18% ΔV , where V_T is the thermal voltage (25 mV at room temperature).^{40–42} The NM of Si CMOS logic gates ranges from $\sim 25\%$ ΔV in the past logic gates (*e.g.*, at $L = 0.8$ μm and

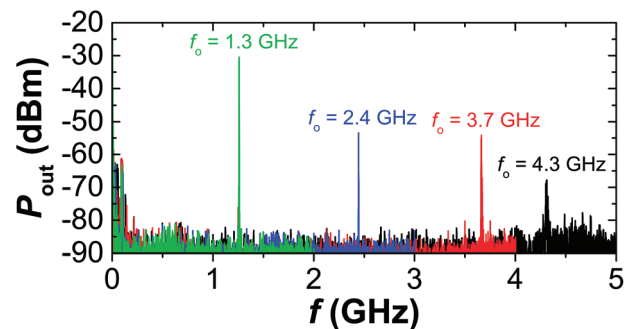


Fig. 4 The power spectra of the output signals with the highest oscillation frequency for each of the investigated RO layouts (coded in the same colors as in Fig. 2). The finite capacitance of the output pad results in signal attenuation which is stronger at higher frequencies, as described in the main text. In the fastest two RO layouts (red and black curves) the surface area of the output pad was reduced from $(150$ $\mu\text{m})^2$ to $(80$ $\mu\text{m})^2$ because of which the attenuation of the signal in red is almost the same as that for the signal in blue. A pure sine wave with a voltage swing $\Delta V = 1$ V (as in Fig. 3) has a power level of 4 dBm.

$\Delta V = 5$ V)^{42,43} to 15% ΔV in the aggressively scaled state-of-the-art logic gates ($L = 10$ nm and $\Delta V = 0.55$ V).⁴⁴ However, in absolute terms, Si CMOS still has the highest NM because it has the highest output voltage swing which is almost equal to the voltage supply, $\Delta V \approx V_{DD}$. The small absolute swing in ECL gates is compensated by the fact that these logic gates generate very low levels of ground and supply noise.^{42,43} The extent of noise generation in graphene logic gates at high frequencies is yet to be investigated.

The actual voltage swing ΔV measured at the output of the fabricated ROs is much smaller than the swing obtained from the static voltage transfer characteristics (Fig. 3), as evidenced by the low power levels of the output signals shown in Fig. 4. The main reason for this discrepancy is the low-pass filtering of the fourth inverter which is loaded by a fixed parasitic capacitance between the output on-chip pad and the back gate ($C_{BG,ox} = 11.5$ nF cm^{-2}). This capacitance does not scale with gate length L and, when multiplied by the total FET channel resistance including the contacts (which scales slower than L), results in a time constant τ_{out} which scales slower than L . Because of this, the oscillation frequency (which scales as $f_o \propto L^{-1}$) increases faster than the output bandwidth $(2\pi\tau_{out})^{-1}$ as the gate length is reduced and therefore the suppression of the output signal is stronger at higher oscillation frequencies.

The voltage gain in the fabricated graphene inverters can be increased by increasing the voltage supply above that used here ($V_{DD} = 2.5$ V),⁴⁵ which could be used to restore oscillation in short-gated ROs and also increase the output voltage swing. However, scaling of electronic circuits is usually accompanied by the decrease of the voltage supply in order to keep the electric field in scaled devices below the breakdown field of the material. High electric fields arise both in the gate dielectric and channel of scaled FETs. Since the investigated FETs are

not scaled in the vertical direction, gate dielectric breakdown limits the voltage supply to a fixed value, *i.e.*, $V_{DD} < 2BV_{ox} = 5$ V, where $BV_{ox} = 2.5$ V is the breakdown voltage of the used gate oxide, here AlO_x (see the Experimental section). On the other hand, the maximum voltage drop V_{max} along the channel ($V_{DS} < V_{max}$), which imposes the limit $V_{DD} < 2V_{max}$, depends on channel dimensions as $V_{max} = RWJ_{max}$, where R is the total channel resistance and $J_{max} = 1.2$ mA μm^{-1} is the maximum current density in graphene.⁴⁶ Although it may seem from this expression that the contact, access, and other series resistances are beneficial in increasing R and therefore V_{max} , they should be excluded from the consideration because they also deteriorate transistor properties. Hence, the lower limit for V_{max} is obtained by assuming that the channel resistance is equal only to the gated part of graphene, *i.e.*, $R = R_{sh}L/W$, where R_{sh} is the sheet resistance of graphene in the fabricated top-gated FETs (here $R_{sh} \sim 4$ k Ω sq⁻¹ at the Dirac point). From this, $V_{max} = R_{sh}LJ_{max} = 4.8L$ V μm^{-1} which leads to $V_{max} = 3.8$ V for $L = 0.8$ μm . Therefore, the gate insulator breakdown limits the voltage supply to 5 V which is above the supplies used here. However, the highest voltage supply used in the fabricated ROs was 3.5 V because we found that at higher voltage supplies transistor properties were unstable due to intense power dissipation (Joule heating).⁴⁷

One of the consequences of scaling is that at the shorter gate lengths L , the maximum voltage drop V_{max} on a transistor decreases while the cutoff frequency f_T at the same time increases. In HBTs, a similar consideration leads to a trade-off between f_T and V_{max} given by the Johnson limit $f_TV_{max} = \nu_{sat}E_B/(2\pi)$, where ν_{sat} is the saturation velocity of charge carriers and E_B is the dielectric strength of the transistor material.²⁷ This limit, although suggested to be valid also for graphene transistors,⁴⁸ is not directly applicable to graphene FETs. The Johnson limit was originally derived by considering the dielectric breakdown of the reverse biased collector p-n junction in bipolar junction transistors.²⁷ Since graphene FETs do not have any p-n junctions, the correct limit can be obtained by multiplying the intrinsic cutoff frequency¹³ $f_T = \nu_{sat}/(2\pi L)$ with $V_{max} = R_{sh}LJ_{max}$. This gives the valid limit for graphene FETs $f_TV_{max} = \nu_{sat}R_{sh}J_{max}/(2\pi) \sim 229$ GHzV, for $\nu_{sat} \sim 3 \times 10^7$ cm s⁻¹,¹¹ which is comparable to that of Si, in which $f_TV_{max} \sim 200$ GHzV.²⁷ However, the numerical value for the limit in graphene is given for illustration purposes, because it is obtained using relatively large sheet resistance in the fabricated top-gated graphene FETs at the Dirac point ($R_{sh} \sim 4$ k Ω sq⁻¹). At smaller sheet resistances, higher f_T with respect to that of Si FETs could be reached only at the expense of a lower V_{max} , *i.e.*, smaller signal amplitude and lower power. Alternatively, this shows that even though the intrinsic cutoff frequency is independent of carrier depletion, it is still smaller in graphene FETs than in Si FETs at the same signal amplitude. This observation is not necessarily a drawback, but serves to guide graphene technology towards low-voltage oscillators, which is not unexpected as it is well-known that the optimal operating voltage of a semiconductor technology tends to scale with the material band gap.

4. Conclusions

We have demonstrated the scaling of FO1 gate delay of graphene ROs, which represents the main IC speed metric adopted by the ITRS. The FO1 gate delay of the fabricated graphene ROs is shorter than that of any strictly low-dimensional materials (*e.g.*, one-dimensional nanotubes, 2D graphene, and MoS₂) to date. The shortest obtained FO1 gate delay was 31 ps, which corresponds to the highest oscillation frequency of 4.3 GHz measured in low-dimensional oscillators. The fabricated graphene ROs also have a static noise margin comparable to that of Si ROs, as a fraction of the output voltage swing. The obtained results stem from the relatively large voltage gain ($A_v \sim -5$) exhibited by sub-micron graphene inverters from which the ROs were composed. The derived Johnson limit for graphene and measured scaling data indicate that graphene ICs would not be able to deliver the same power at high frequencies as the Si circuits, but could outperform them in terms of oscillation frequency at the same gate length. The present graphene ICs could find applications in simple graphene mixed-signal circuits (*e.g.*, for baseband processing) in which fast operation is favored over power dissipation. Our results also emphasize the need for further reduction of contact resistance and device variability in graphene ICs in order to preserve the demonstrated scaling trend.

Acknowledgements

We thank M. Leone for technical support. This research was supported in part by Fondazione Cariplo (grant no. 2011-0373), EU FP7 Graphene Flagship (grant no. 604391), the PRIN project GRAF, and by the United States National Science Foundation (NSF) and Air Force Office of Scientific Research (AFOSR).

References

- 1 F. Schwierz, H. Wong and J. J. Liou, *Nanometer CMOS*, Pan Stanford Publishing, Singapore, 2010.
- 2 A. M. Niknejad and H. Hashemi, *mm-Wave Silicon Technology 60 GHz and Beyond*, Springer, New York, 2008.
- 3 *Handbook of Semiconductor Manufacturing Technology*, ed. R. Doering and Y. Nishi, CRC Press, Boca Raton, 2008.
- 4 J.-P. Colinge, *Silicon-on-Insulator Technology: Materials to VLSI*, Springer, New York, 2004.
- 5 C. Auth, C. Allen, A. Blattner, D. Bergstrom, M. Brazier, M. Bost, M. Buehler, V. Chikarmane, T. Ghani, T. Glassman, R. Grover, W. Han, D. Hanken, M. Hattendorf, P. Hentges, R. Heussner, J. Hicks, D. Ingerly, P. Jain, S. Jaloviar, R. James, D. Jones, J. Jopling, S. Joshi, C. Kenyon, H. Liu, R. McFadden, B. McIntyre, J. Neirynek, C. Parker, L. Pipes, I. Post, S. Pradhan,

- M. Prince, S. Ramey, T. Reynolds, J. Roesler, J. Sandford, J. Seiple, P. Smith, C. Thomas, D. Towner, T. Troeger, C. Weber, P. Yashar, K. Zawadzki and K. Mistry, *Symp. VLSI Technol.*, 2012, 131–132.
- 6 C.-H. Jan, U. Bhattacharya, R. Brain, S.-J. Choi, G. Curello, G. Gupta, W. Hafez, M. Jang, M. Kang, K. Komeyli, T. Leo, N. Nidhi, L. Pan, J. Park, K. Phoa, A. Rahman, C. Staus, H. Tashiro, C. Tsai, P. Vandervoorn, L. Yang, J.-Y. Yeh and P. Bai, *Int. Electron Devices Meet.*, 2012, 3.1.1–3.1.4, San Francisco, CA, USA.
- 7 D. S. Novikov, *Appl. Phys. Lett.*, 2007, **91**, 102102.
- 8 J.-H. Chen, C. Jang, S. Adam, M. S. Fuhrer, E. D. Williams and M. Ishigami, *Nat. Phys.*, 2008, **4**, 377–381.
- 9 X. Hong, A. Posadas, K. Zou, C. H. Ahn and J. Zhu, *Phys. Rev. Lett.*, 2009, **102**, 136808.
- 10 C. R. Dean, A. F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. L. Shepard and J. Hone, *Nat. Nanotechnol.*, 2010, **5**, 722–726.
- 11 V. E. Dorgan, M.-H. Bae and E. Pop, *Appl. Phys. Lett.*, 2010, **97**, 082112.
- 12 P. R. Wallace, *Phys. Rev.*, 1947, **71**, 622–634.
- 13 S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, Wiley-Interscience, Hoboken, 2007.
- 14 S. Mason, *Trans. IRE Prof. Group Circuit Theory*, 1954, **CT-1**, 20–25.
- 15 G. Taylor and J. Simmons, *Solid-State Electron.*, 1986, **29**, 941–946.
- 16 R. Chau, S. Datta, M. Doczy, B. Doyle, B. Jin, J. Kavalieros, A. Majumdar, M. Metz and M. Radosavljevic, *IEEE Trans. Nanotechnol.*, 2005, **4**, 153–158.
- 17 G. Moore, *Proc. IEEE*, 1998, **86**, 82–85.
- 18 K. Kuhn, U. Avci, A. Cappellani, M. Giles, M. Haverty, S. Kim, R. Kotlyar, S. Manipatruni, D. Nikonov, C. Pawashe, M. Radosavljevic, R. Rios, S. Shankar, R. Vedula, R. Chau and I. Young, *Int. Electron Devices Meet.*, 2012, 8.1.1–8.1.4.
- 19 The international technology roadmap for semiconductors, <http://www.itrs.net>.
- 20 The international technology roadmap for semiconductors: Process Integration, Devices, and Structures, 2011, <http://www.itrs.net/Links/2011ITRS/2011Chapters/2011PIDS.pdf>.
- 21 Y. Wu, K. A. Jenkins, A. Valdes-Garcia, D. B. Farmer, Y. Zhu, A. A. Bol, C. Dimitrakopoulos, W. Zhu, F. Xia, P. Avouris and Y.-M. Lin, *Nano Lett.*, 2012, **12**, 3062–3067.
- 22 F. Schwierz, *Proc. IEEE*, 2013, **101**, 1567–1584.
- 23 E. Guerriero, L. Polloni, M. Bianchi, A. Behnam, E. Carrion, L. G. Rizzi, E. Pop and R. Sordan, *ACS Nano*, 2013, **7**, 5588–5594.
- 24 Z. Chen, J. Appenzeller, Y.-M. Lin, J. Sippel-Oakley, A. G. Rinzler, J. Tang, S. J. Wind, P. M. Solomon and P. Avouris, *Science*, 2006, **311**, 1735.
- 25 H. Wang, L. Yu, Y.-H. Lee, Y. Shi, A. Hsu, M. L. Chin, L.-J. Li, M. Dubey, J. Kong and T. Palacios, *Nano Lett.*, 2012, **12**, 4674–4680.
- 26 S. D. Brotherton, C. Glasse, C. Glaister, P. Green, F. Rohlfing and J. R. Ayres, *Appl. Phys. Lett.*, 2004, **84**, 293–295.
- 27 E. Johnson, *RCA Rev.*, 1965, **26**, 163–177.
- 28 S.-L. Li, H. Miyazaki, A. Kumatani, A. Kanda and K. Tsukagoshi, *Nano Lett.*, 2010, **10**, 2357–2362.
- 29 J. T. Smith, A. D. Franklin, D. B. Farmer and C. D. Dimitrakopoulos, *ACS Nano*, 2013, **7**, 3661–3667.
- 30 W. S. Leong, H. Gong and J. T. L. Thong, *ACS Nano*, 2014, **8**, 994–1001.
- 31 F. Traversi, V. Russo and R. Sordan, *Appl. Phys. Lett.*, 2009, **94**, 223312.
- 32 L. G. Rizzi, M. Bianchi, A. Behnam, E. Carrion, E. Guerriero, L. Polloni, E. Pop and R. Sordan, *Nano Lett.*, 2012, **12**, 3948–3953.
- 33 B. Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill, Boston, 2001.
- 34 S. Tyagi, C. Auth, P. Bai, G. Curello, H. Deshpande, S. Gannavaram, O. Golonzka, R. Heussner, R. James, C. Kenyon, S. H. Lee, N. Lindert, M. Liu, R. Nagisetty, S. Natarajan, C. Parker, J. Sebastian, B. Sell, S. Sivakumar, A. St Amour and K. Tone, *Int. Electron Devices Meet. Tech. Dig.*, 2005, 245–247.
- 35 A. Nathan, A. Ahnood, M. T. Cole, S. Lee, Y. Suzuki, P. Hiralal, F. Bonaccorso, T. Hasan, L. Garcia-Gancedo, A. Dyadyusha, S. Haque, P. Andrew, S. Hofmann, J. Moultrie, D. Chu, A. Flewitt, A. Ferrari, M. Kelly, J. Robertson, G. Amaratunga and W. I. Milne, *Proc. IEEE*, 2012, **100**, 1486–1517.
- 36 S. Yang, S. Ahmed, B. Arcot, R. Arghavani, P. Bai, S. Chambers, P. Charvat, R. Cotner, R. Gasser, T. Ghani, M. Hussein, C. Jan, C. Kardas, J. Maiz, P. McGregor, B. McIntyre, P. Nguyen, P. Packan, I. Post, S. Sivakumar, J. Steigerwald, M. Taylor, B. Tufts, S. Tyagi and M. Bohr, *Int. Electron Devices Meet. Tech. Dig.*, 1998, 197–200.
- 37 S. Tyagi, M. Alavi, R. Bigwood, T. Bramblett, J. Brandenburg, W. Chen, B. Crew, M. Hussein, P. Jacob, C. Kenyon, C. Lo, B. McIntyre, Z. Ma, P. Moon, P. Nguyen, L. Rumaner, R. Schweinfurth, S. Sivakumar, M. Stettler, S. Thompson, B. Tufts, J. Xu, S. Yang and M. Bohr, *Int. Electron Devices Meet. Tech. Dig.*, 2000, 567–570.
- 38 J. P. Uyemura, *CMOS Logic Circuit Design*, Springer, New York, 2002.
- 39 J. R. Hauser, *IEEE Trans. Educ.*, 1993, **36**, 363–368.
- 40 M. J. W. Rodwell, M. Le and B. Brar, *Proc. IEEE*, 2008, **96**, 271–286.
- 41 R. C. Jaeger and T. N. Blalock, *Microelectronic Circuit Design*, McGraw-Hill, New York, 2011.
- 42 M. L. Minges, *Electronic Materials Handbook: Packaging, Volume I*, ASM International, Materials Park, 1989.
- 43 J. F. Wakerly, *Digital Design Principles and Practices*, Pearson Education, Upper Saddle River, 2006.
- 44 M.-C. Chen, C.-H. Lin, Y.-F. Hou, Y.-J. Chen, C.-Y. Lin, F.-K. Hsueh, H.-L. Liu, C.-T. Liu, B.-W. Wang, H.-C. Chen, C.-C. Chen, S.-H. Chen, C.-T. Wu, T.-Y. Lai, M.-Y. Lee, B.-W. Wu, C.-S. Wu, I. Yang, Y.-P. Hsieh, C. Ho, T. Wang,

- A. Sachid, C. Hu and F.-L. Yang, *Symp. VLSI Technol.*, 2013, T218–T219.
- 45 E. Guerriero, L. Polloni, L. G. Rizzi, M. Bianchi, G. Mondello and R. Sordan, *Small*, 2012, **8**, 357–361.
- 46 A. D. Liao, J. Z. Wu, X. Wang, K. Tahy, D. Jena, H. Dai and E. Pop, *Phys. Rev. Lett.*, 2011, **106**, 256801.
- 47 S. Islam, Z. Li, V. Dorgan, M.-H. Bae and E. Pop, *IEEE Electron Device Lett.*, 2013, **34**, 166–168.
- 48 M. J. Kelly, *Semicond. Sci. Technol.*, 2013, **28**, 122001.

Scaling of Sub-Micron Graphene Integrated Circuits

Electronic Supplementary Information

Massimiliano Bianchi,^a Erica Guerriero,^a Marco Fiocco,^a Ruggero Alberti,^a Laura Polloni,^a Ashkan Behnam,^b Enrique A. Carrion,^b Eric Pop,^c and Roman Sordan^{*a}

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXX 20XX

First published on the web Xth XXXXXXXXXX 200X

DOI: 10.1039/b000000x

Cutoff frequency. The most commonly used transistor performance metric is the cutoff frequency f_T ; the intrinsic f_T can be very large in graphene FETs ($f_T \sim 400$ GHz)^{1,2} due to large mobility in graphene, and compares favorably to that of conventional FETs.³ The cutoff frequency was also found to scale as $f_T \propto L^{-1}$ similar to conventional semiconductor FETs.^{1,3} However, the cutoff frequency is a measure of the internal transistor transit time rather than transistor functionality at high-frequencies. This is because the cutoff frequency is obtained from the small-signal current gain measured in a circuit in which the gate is driven by a current source while the output (between the source and drain) is shorted.⁴ Use of an input current source makes the cutoff frequency independent of the gate resistance which is one of the main limiting factors affecting high-frequency response of FETs. Similarly, the short circuit between source and drain shorts the output conductance (in the absence of the parasitic source/drain series resistances) and therefore makes the cutoff frequency independent of the extent of the drain current saturation (shorted output also does not have any practical relevance).

Maximum frequency of oscillation. A more realistic figure of merit is the maximum frequency of oscillation f_{max} which represents the frequency at which the Mason's unilateral power gain⁵ falls to unity. Since the power gain is obtained from a circuit in which the impedance of the input voltage source and the load impedance are finite,⁴ both the gate resistance and output conductance influence the power gain and therefore f_{max} . For this reason graphene FETs exhibit much smaller f_{max} than other types of transistors, typically ~ 50 GHz.⁶ By comparison, $f_{max} > 1$ THz has been obtained both in InP high electron mobility transistors⁷ and heterojunction bipolar transistors.⁸ In addition, f_{max} of graphene FETs was found to scale only weakly with the inverse gate length.³ The maximum frequency of oscillation represents an impor-

tant benchmark because it is an invariant quantity which allows comparison of different transistor technologies and accounts for real signal amplification.⁵ However, it suffers from the requirements for device unilateralization and impedance matching which cannot be performed in realistic circuits.

Intrinsic gate delay. The typical figure of merit in a large-signal regime is the intrinsic gate delay CV/I which represents the time required to increase the voltage of the gate capacitor C_G by the large-signal output voltage swing ΔV when charged by a constant drain current I_D .⁹

References

- 1 Y. Wu, K. A. Jenkins, A. Valdes-Garcia, D. B. Farmer, Y. Zhu, A. A. Bol, C. Dimitrakopoulos, W. Zhu, F. Xia, P. Avouris and Y.-M. Lin, *Nano Lett.*, 2012, **12**, 3062–3067.
- 2 R. Cheng, J. Bai, L. Liao, H. Zhou, Y. Chen, L. Liu, Y.-C. Lin, S. Jiang, Y. Huang and X. Duan, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 11588–11592.
- 3 F. Schwierz, *Proc. IEEE*, 2013, **101**, 1567–1584.
- 4 F. Schwierz, H. Wong and J. J. Liou, *Nanometer CMOS*, Pan Stanford Publishing, Singapore, 2010.
- 5 S. Mason, *T. IRE Prof. Group Circ. Theory*, 1954, **CT-1**, 20–25.
- 6 Z. Guo, R. Dong, P. S. Chakraborty, N. Lourenco, J. Palmer, Y. Hu, M. Ruan, J. Hankinson, J. Kunc, J. D. Cressler, C. Berger and W. A. de Heer, *Nano Lett.*, 2013, **13**, 942–947.
- 7 R. Lai, X. Mei, W. Deal, W. Yoshida, Y. M. Kim, P. H. Liu, J. Lee, J. Uyeda, V. Radisic, M. Lange, T. Gaier, L. Samoska and A. Fung, *Int. El. Devices Meet.*, 2007, pp. 609–611.
- 8 J. Hacker, M. Urteaga, M. Seo, A. Skalare and R. Lin, *Proc. IEEE MTT-S Int. Microwave Symp.*, 2013, pp. 1–3.
- 9 R. Chau, S. Datta, M. Doczy, B. Doyle, B. Jin, J. Kavalieros, A. Majumdar, M. Metz and M. Radosavljevic, *IEEE T. Nanotechnol.*, 2005, **4**, 153–158.

^a L-NESS, Department of Physics, Politecnico di Milano, Polo di Como, Via Anzani 42, 22100 Como, Italy. E-mail: roman.sordan@polimi.it

^b Electrical & Computer Engineering, University Illinois Urbana-Champaign, Urbana IL 61801, USA.

^c Electrical Engineering, Stanford University, Stanford, CA 94305, USA.