Electronic and thermal effects in the insulator-metal phase transition in VO$_2$ nano-gap junctions

Arash Joushaghani,¹ Junho Jeong,¹ Suzanne Paradis,² David Alain,² J. Stewart Aitchison,¹ and Joyce K. S. Poon¹

¹) Department of Electrical and Computer Engineering, University of Toronto, 10 King’s College Road, Toronto, Ontario, M5S 3G4, Canada
²) Defence Research and Development Canada - Valcartier, 2459 Pie-XI Blvd. North, Quebec, Quebec G3J 1X5, Canada

(Dated: 27 November 2014)

By controlling the thermal transport of VO$_2$ nano-gap junctions using device geometry, contact material, and applied voltage waveforms, the electronically induced insulator-metal phase transition is investigated in the adiabatic heating and transient carrier injection regimes. With a gradual ramping of an applied voltage on a microsecond time scale, the transition electric field threshold can be directly reduced by the Joule heating. With an abrupt applied voltage, the transition threshold is initiated by carriers injected within the first tens of nanoseconds, but the complete insulator-metal phase transition is limited by thermal redistribution times to hundreds of nanoseconds.

The dynamics of the insulator-metal phase transition (IMT) in the correlated electron material vanadium dioxide (VO$_2$)$^{1,2}$ have attracted much attention$^{3,4}$ due to potential application of VO$_2$ in transistors$^{5,6}$ and opto-electronic$^{7,8}$ switches. A reversible IMT through a structural reconfiguration occurs in VO$_2$ when the local temperature rises beyond the insulator-metal transition temperature, $T_{I\rightarrow M} = 340$ K. This process takes place on the time scale of $10^{-7}$-$10^{-5}$s for nano-scale devices$^{9}$. Alternatively, the IMT due to an electronic reconfiguration takes place on time scales of $10^{-12}$-$10^{-8}$ s due to a perturbation in the electronic structure$^{11}$, accumulated surface charge$^{12}$, external strain$^{13}$, applied electric fields or voltages$^{14-18}$, and currents$^{19}$. An electronically induced IMT produced by a current or applied voltages is typically accompanied by a local temperature increase in the VO$_2$ from Joule heating$^{20,21}$. The temperature rise modifies the temporal dynamics of the phase transition by inducing a secondary structural phase transition$^{21}$ when the temperature exceeds $T_{I\rightarrow M}$. Although thermal effects can be reduced by limiting the current$^{22}$ or increasing the heat dissipation rate of the devices$^{23}$, the role of Joule heating during an electronically induced IMT remains unclear$^{5,24}$.

In this Letter, we report the effects of heating on the electronically induced IMT in VO$_2$ nano-gap junctions. The thermal transport was controlled by the applied voltage waveform, device geometry, and metal contact. We studied the IMT in the adiabatic and transient regimes. The dependence of the IMT threshold voltage on the duration of the applied voltage pulses and device parameters can be explained by thermal accumulation and dissipation rates. The IMT could not occur in the absence of injected carriers, so carrier effects are critical in the electronically induced phase transition of VO$_2$.

The schematic and scanning electron micrograph (SEM) of the typical device are presented in Figs. 1(a)-1(b) respectively. A VO$_2$ film with a thickness of $h = 100$ nm was deposited using radio-frequency magnetron sputtering of a vanadium target on 2 $\mu$m thermally grown layer of silica (SiO$_2$) on a silicon (Si) substrate. The VO$_2$ junctions were formed by a combination of electron beam (e-beam) lithography, metal evaporation, and lift-off to produce two metal contacts separated by a gap of length $L$. These contacts had a width of $W = 10 \mu$m and a thickness of $t = 100$ nm. The contacts were either gold (Au) or palladium (Pd).

The phase transition dynamics are limited by the $RC$ time constant of the circuit, $\tau_{RC}$, the thermal extraction time of the junction, $\tau_{T}$, and the forward electronic phase transitions time constant, $\tau_E$. An independent measurement of $\tau_T$ and $\tau_E$ of the junctions is difficult, since the electronic phase transition is typically accompanied with a temperature increase. As the temperature increases, the average electric field at the onset of the phase transition, $E_{I\rightarrow M}$, decreases$^{25}$ according to $E_{I\rightarrow M} \propto (1 - (T/T_c))^2$. However, by independently varying $\tau_T$, we can evaluate the roles of the thermal and electronic effects, as well as their interaction in the phase transition.

The heat transfer channels in the VO$_2$ junction are illustrated in Fig. 1(c). The total heat dissipation rate of the junction is the sum of the effective heat extraction rates of the surrounding media, $\gamma_x$, where $x$ is either air, VO$_2$, SiO$_2$ or the metal contact. Since the thermal conductivity of the contact metals is much larger than that of the VO$_2$, air, or SiO$_2$, the
The total heat dissipation rate of the junction can be approximated as $\gamma_{\text{Metal}}$. The heat generation rate in the junction is the total dissipated electrical power, $P(t)$. As a result, the simplified heat transfer equation for the VO$_2$ junctions can be written as:

$$
\rho D C_p \frac{\partial T}{\partial t} = \kappa \nabla^2 T - \gamma_{\text{Metal}} T + \frac{P(t)}{V_{\text{eff}}},
$$

where $V_{\text{eff}}$ is the effective voltage of the junction, and $\rho_D$, $C_p$, and $\kappa$ are the density, the specific heat capacity and the thermal conductivity of VO$_2$ respectively. This equation shows that $\gamma_{\text{Metal}}$ and $P(t)$ changes the thermal transport rate of the system and allows for the controlled study of $\tau_p$ and $\tau_E$. In the experiments to follow, the contacts were either Au or Pd, and we controlled $P(t)$ by the applied voltage waveform. The thermal conductivities of Au and Pd are about 315 W/(m·K) and 70 W/(m·K), respectively.

To measure the time constants of the VO$_2$ devices, an arbitrary function generator was used to apply a voltage, $V_{\text{App}}$, across the device contacts and a series resistor $R_L = 550$ Ω. The simplified circuit model of the device and the experiment is shown in Fig. 1(d). The voltage drop across the gap, $V_{\text{Gap}}$, and $V_{\text{App}}$ were monitored using an oscilloscope. The measured $\tau_{RC}$ for the setup was $\approx 40 \pm 10$ ns and remained constant within the experimental error range for devices with different values of gap lengths and contact metals.

We investigated the adiabatic Joule heating regime by applying a periodic train of triangular voltage pulses, which gradually increased $I$ and $V_{\text{Gap}}$ and allowed sufficient time for Joule heating to occur. The rise time of the pulses, $t_p$, was at least 5 μs. The period was 2 s to prevent cumulative heating from successive pulses. Figure 2(a) shows an example of the measured $V_{\text{Gap}}$ and $V_{\text{App}}$.

To extract the electric field for the phase transition, we measured $V_{I\rightarrow M}$ for devices with different $L$ and took the slopes of to be the “average” transition electric field, $E_{I\rightarrow M}$. The voltage offset, $V_0$, taken to be the value of $V_{I\rightarrow M}$ as $L \to 0$, is due to the finite contact resistance and work function mismatch between VO$_2$ and the contact material. As an example, Figure 2(b) shows the measured $V_{I\rightarrow M}$ as a function of $L$ for devices with Pd contacts and for $t_p$ between 5 μs and 1 s. For all device lengths, shorter voltage pulses required higher values of $V_{I\rightarrow M}$, which meant more energy had to be deposited to the device to initiate the transition. This could be caused by a reduction of thermal accumulation (which would have decreased $V_{I\rightarrow M}$) and/or a limitation due to $\tau_E$.

To investigate thermal influence on $E_{I\rightarrow M}$, we extracted $E_{I\rightarrow M}$ as a function of $t_p$ for devices with Au or Pd metal contacts in Fig. 2(c). The solid lines are fits to guide the eye. For both types of samples, $E_{I\rightarrow M}$ was constant when $t_p \geq 1$ ms. As $t_p$ decreased, $E_{I\rightarrow M}$ increased and saturated at around 55 kV/cm for the Au samples when $t_p < 10^{-5}$s. $E_{I\rightarrow M}$ did not saturate for the Pd samples for the measured values of $t_p$. We define a characteristic pulse width, $t_p$, as the value of $t_p$ that resulted in a 50% increase in $E_{I\rightarrow M}$ from its value at $t_p = 1$ s. As shown in Fig. 2(c), $t_p^{10\%} \approx 3 \times 10^{-5}$s for the Au samples and $t_p^{10\%} \approx 5 \times 10^{-6}$s for the Pd samples.

The dependence of $E_{I\rightarrow M}$ on $t_p$ and the contact metal can be explained by thermal transport. When $t_p \gg t_r$, Joule heating raised the temperature in the VO$_2$ junction, which reduced $E_{I\rightarrow M}$, As $t_p$ decreased, heating had a lesser contribution and consequently $E_{I\rightarrow M}$ increased. Since Au has a higher thermal conductivity than Pd (which would reduce the temperature of the VO$_2$ in the Au samples), $E_{I\rightarrow M}$ for Au was larger than in the Pd samples at all values of $t_p$. $E_{I\rightarrow M}$ saturated when $t_p \ll t_r$, which was approximately the value of $E_{I\rightarrow M}$ when no thermal effects were present. In the Au samples, the saturation of $E_{I\rightarrow M}$ to about 55 kV/cm when $t_p \ll t_r$ occurred when almost no thermal effects were present. This field value agrees well with measurements of vertical VO$_2$ junctions, in which heat accumulation is minimized owing to large device cross-sections.

The temperature change in the VO$_2$ due to the applied voltage further affects the carrier density in the devices. Fig. 2(d) shows the offset voltage, $V_0$, vs. $t_p$. Both the carrier density and the work function of VO$_2$ increase across the phase transition with a similar temperature dependence as $E_{I\rightarrow M}$. However, while a higher carrier density reduces $V_0$, a higher work function increases $V_0$. The carrier density effects tend to dominate over the work function changes, because the carrier density can change by about 3 orders of magnitude across the IMT, while the work function changes by less than 3%.

The measured resistivity change of the junctions was slightly larger than 2 orders of magnitude. Therefore, the increase in $V_0$ when $t_p < t_r$, in a similar way as $E_{I\rightarrow M}$ in Fig. 2(c), suggests the carrier density decreased as the voltage pulses became shorter and the role of heating was reduced.

To study the IMT without thermal effects, we applied square-wave pulses with rise time $t_p \ll t_r$, which was limited to 40 ns due to the experimental setup. The pulses had a duration of 1 μs and a period of 2 s. The amplitude of $V_{\text{App}}$,
\( v_{\text{App}} \) was varied and the time profile of \( V_{\text{Gap}} \) was measured. Figure 3(a) shows an example of the measured \( V_{\text{App}} \) and \( V_{\text{Gap}} \), and \( v_{\text{App}} \) is indicated. As before, the drop in \( V_{\text{Gap}} \) indicated the start of the phase transition. The value of \( E_{I\rightarrow M} \) was determined using the smallest value of \( v_{\text{App}} \) that triggered the IMT for each \( L \) and was 52 ± 5 kV/cm for Pd and 55 ± 8 kV/cm for Au devices. These values agreed well with the saturation values of \( E_{I\rightarrow M} \) in Fig. 2(c).

Figure 3(b) shows the measured values of \( R_{\text{gap}} \) as \( v_{\text{App}} \) increased for a device with \( L = 800 \) nm and Pd contacts. \( R_{\text{gap}} \) was constant when \( V_{\text{Gap}} < V_{I\rightarrow M} \). Increasing \( v_{\text{App}} \) such that \( V_{\text{Gap}} \geq V_{I\rightarrow M} \) resulted in a rapid drop in \( R_{\text{gap}} \) at a time of \( \tau_1 \) after the rise of the pulse. \( R_{\text{gap}} \) continued to decrease and dropped below 10 dB of its initial value at \( \tau_2 \) after the rise of the pulse. Both \( \tau_1 \) and \( \tau_2 \) decreased with increasing \( v_{\text{App}} \). Figures 3(c)-3(d) show the measured \( \tau_1 \) and \( \tau_2 \) as a function of \( v_{\text{App}} \) for devices with two different values of \( L \) and contact metals. As \( v_{\text{App}} \) increased, \( \tau_1 \) for the \( L = 500 \) nm Pd and Au devices and \( L = 800 \) nm Pd devices decreased and saturated at approximately \( \tau_{\text{RC}} \). This suggests \( \tau_1 \) is due to the combined effects of \( \tau_{\text{RC}} \) and \( \tau_{\text{E}} \) with \( \tau_{\text{E}} \leq \tau_{\text{RC}} \). \( \tau_1 \) did not saturate for the \( L = 800 \) nm Au device for the voltages tested. \( \tau_2 \) also decreased with increasing \( v_{\text{App}} \) at a lower rate than \( \tau_1 \), consistent with \( \tau_2 \) being caused by Joule heating of the VO\(_2\). As expected, longer channels required more time for the heat to redistribute. For the same device length and the same voltage difference from the phase transition threshold, \( \tau_2 \) for Au was higher than that for Pd by up to a factor of about 2, because the local temperature of the Au samples would be lower owing to the higher thermal conductivity of Au. As \( v_{\text{App}} \), or equivalently, the applied electrical power, increased, \( \tau_2 \) decreased because of the faster rate of heating as described by Eq. 1.

These measurements confirm that the phase transition is initiated electronically and occurs on a timescale limited by the \( RC \) time constant of the measurement setup, and is followed by a thermal transition that is limited by the thermal dissipation of the junction. To suppress the thermal effects, it is essential to reduce \( I \) by electrically insulating the VO\(_2\) from the metallic contacts. One approach is to use a thin high-k dielectric between the VO\(_2\) and the metallic contacts. Figure 4(a) shows the schematic of the device. A 7 nm thick layer of zirconium dioxide (ZrO\(_2\)) was deposited using atomic layer deposition and patterned using a combination of electron beam lithography and lift-off. The metal contacts were formed as before using Pd and had the width of \( W = 10 \) \( \mu \)m and thickness of 100 nm. Figure 4(b) is an X-ray enhanced SEM image of the device near the gap showing that the thin dielectric layer that covers the VO\(_2\) surface.

To observe the phase transition, we measured \( V_{\text{Gap}} \) and the current, \( I \), that leaked through the ZrO\(_2\) and VO\(_2\) layers. Figure 4(c) shows the measured \( IV \) characteristics of two junctions with (green) and without (black) the ZrO\(_2\) layer. Both devices had the same \( L = 500 \) nm. The current, \( I \), initially increased with increasing \( V_{\text{Gap}} \) and the phase transition resulted in a drop in \( V_{\text{Gap}} \). As expected, the ZrO\(_2\) layer reduced \( I \) by an order of magnitude at any value of \( V_{\text{Gap}} \). However, the ZrO\(_2\) layer also increased \( V_{I\rightarrow M} \) by a factor of 2.5, which resulted in a value of \( I \) that was similar to devices without ZrO\(_2\). Figure 4(d) shows the measured \( V_{I\rightarrow M} \) vs. \( L \) for devices with (green) and without (black) the ZrO\(_2\) layer. \( V_{I\rightarrow M} \) did not change with \( L \) for devices with the ZrO\(_2\) layer, indicating that the phase transition occurred due to the electric field breakdown of the ZrO\(_2\) that allowed sufficient current to pass through the device before the phase transition occurred.

As another approach to limit the current, we fabricated VO\(_2\) nano- and micro-wires placed between a capacitor with an air gap. The VO\(_2\) wires were fabricated using e-beam lithography and subsequent reactive ion etching. The metal contacts
were defined using e-beam lithography, thermal evaporation of 400 nm of Au, and a subsequent lift-off step. Figure 5(a) shows the SEM images of the fabricated device and the magnified images, showing the VO₂ nanowire at the gap and on the metal contacts. Four metal pads contacted the VO₂ wire to accurately measure the wire resistivity, ρ\textsubscript{Wire}. A voltage of V\textsubscript{Gap} was applied across the capacitor, which established an electric field in a section of the VO₂ between the electrodes.

Figure 5(b) shows the normalized resistivity of the wire \(\rho\textsubscript{Wire}(T)/\rho\textsubscript{Wire}(T = 295 K)\) as a function of T for a wire with a width of 500 nm and a thickness of 100 nm. When \(T > 335\) K, the VO₂ underwent a phase transition and the resistivity of the wires dropped. Figure 5(c) shows the normalized resistivity as V\textsubscript{Gap} increased. The resistivity did not change with V\textsubscript{Gap} up to the breakdown voltage of the gap, which was about 20 V. Below the breakdown voltage, the leakage current remained less than 10 pA. The measurements were repeated at different temperatures to decrease the required \(E_{1\rightarrow M}\). Even when \(T \approx T_{1\rightarrow M}\), the normalized resistivity was unaffected by V\textsubscript{Gap}. As previously mentioned, the required electric field for the phase transition of VO₂ scales as \(1/T\) and is \(< 5 \text{ kV/cm}\) close to \(T_{1\rightarrow M}\). According to finite element simulations, the average electric field in these devices was \(> 10 \text{ kV/cm}\). However, these electric field values did not result in any detectable changes in \(\rho\textsubscript{Wire}\) even at elevated temperatures. If an external electric field could induce a phase transition in these films, its threshold would have been \(\gg 10 \text{ kV/cm}\), a value which was not achieved in this experiment. Consequently, we conclude that the most plausible cause of the phase transition in these films was carrier injection.

In summary, we have shown the influence of Joule heating on the threshold electric field, carrier densities, and phase transition time scales of the IMT in micro- and nano-scale two terminal VO₂ junctions. The phase transition was initiated by carrier injection (not electric field). The measured thermal dissipation time was \(10^{-7} \text{ s}\), while the measured electronic transition time was limited by the \(RC\) time constant of the setup to 40 ns. The thermal effects can be suppressed using high frequency circuits for current regulation after the phase transition. These circuits are typically used for current quenching in avalanche photodiodes and can be potentially adapted to VO₂ devices.