A sub-1V bandgap voltage reference (BGVR) in 0.16µm CMOS is presented that circumvents the trade-off between area and power consumption in conventional sub-1V BGVR topologies. This circuit can be used in systems to generate reference voltages locally, eliminating the conventional relatively large BGVR and its global reference voltage distribution interconnect. The active area is 0.0025mm²; at 295K, \( V_{\text{REF}} = 0.94 \text{V} \) with \( \sigma = 0.8\% \) at a minimum supply voltage of 1.1V, with 1.4µA supply current.

**Text:**

Today’s ICs usually employ one bandgap voltage reference (BGVR) circuit to generate a well defined voltage that is reused at many places in that IC. The classical BGVR generates a reference voltage that is slightly larger than the material bandgap: a little above 1200mV in silicon. For deep sub-micron technologies the supply voltage is about the same as the material bandgap which prevents using the classical bandgap structure. As a solution a number of BGVR topologies that create a sub-1V are invented; most of them are based on the structure introduced by Banba [1], some are using resistive voltage division [2] or voltage averaging [3]. For low-
power operation high-ohmic resistors (occupying a large area!) must be used in all these techniques, leading to an immediate trade-off between power consumption and chip-area. This trade-off prevents the local generation of reference voltages where they are required: either the power penalty or the area penalty would be too significant. Alternative topologies that do not require high-ohmic resistors typically are not-BGVR-based circuits relying on threshold voltages and hence require trimming to achieve low spread.

The current manuscript presents a sub-1V BGVR that circumvents the power-area trade off of conventional sub-1V BGVR, aiming at local reference voltage generation wherever a reference voltage is needed in a die. To break the power-area trade-off no resistive averaging nor subdivision can be used to get sub-1V operation, while a drive to minimum area results in avoiding any topology using opamps, high ohmic resistors or multiple diodes. The basic topology of the BGVR is shown in Fig. 1 and is known [4]. For this topology and assuming that the PMOS transistors MP1 and MP2 are in weak-inversion with an ideal exponential behavior, the generated reference voltage ($V_{REF}$) is known to be

$$V_{REF} = \frac{kT}{q} \cdot N \cdot \ln(A) + \frac{kT}{q} \cdot \ln\left(\frac{I_{C}(T)}{I_{C,0} \cdot T^\eta}\right) + V_{gap,0} \approx 1.2V,$$

with a curvature that amounts to about 4mV over a 160K temperature range. In this relation the factor $N$ equals the ratio between resistors $R_2$ and $R_1$, the factor $A$ equals the ratio of the current factors of transistors MP1 and MP2, $I_C(T)$ is the diode current as a function of temperature $T$, $\eta$ is a technology and topology dependent factor typically around 4 while the other parameters are technology dependent or just plain constants. For actual, non-ideal, weak-inversion operated MOS transistors the expression for $V_{REF}$ is more complex, which is not of main relevance for the current manuscript.

To get a sub-1V BGVR without any subdivision, we introduce an almost temperature independent offset between MP1 and MP2, realized using the dependency between threshold
voltage $V_T$ and transistor length $L$. A typical $V_T(L)$ relation for three temperatures is given in Fig. 2. On the x-axis in this figure is the transistor length, starting at minimum length. With increasing length the $V_T$ first increases and then decreases, due to short channel effects [5]. For the current circuit we selected transistor lengths $L_1$ and $L_2$ for transistors MP1 and MP2 respectively, yielding an almost temperature independent offset voltage $\Delta V_T$. The result is that now

$$V_{REF} = N \cdot \left[ \frac{kT}{q} \cdot \ln(A) - \Delta V_T \right] + \frac{kT}{q} \cdot \ln \left( \frac{I_c(T)}{I_{c0}T^q} \right) + V_{gap,0} \approx 1.2V - N \cdot \Delta V_T. \quad (1)$$

An associated effect is that the BGVR does not operate at temperatures below $T_{min} = \frac{q \Delta V_T}{k \cdot \ln(A)}$, whereas a conventional BGVR has a lower temperature limit that is set by some other limitations of the circuitry in the BGVR or is determined just by the application; a typical lower useful temperature limit is -25°C or -40°C; reaching -273°C (0 K) is not often necessary. To put it the other way round, if the BGVR of Fig. 1 must operate down to a temperature $T_{min}$, the minimum $V_{REF} \approx 1.2V - (T_{min} - 273.15) \cdot \frac{2mV}{K}$. The 2mV/K in this relation corresponds to the temperature dependency of a PN-junction at constant or weakly temperature dependent current.

The circuit we implemented equals the principle circuit of Fig. 1, with regulated cascodes, see Fig. 3. For the design, a minimum usable temperature of -40°C was defined, with a design margin of 60°C, which yields effectively $T_{min} = -100°C$ and results in $V_{REF} \approx 900mV$. In our design we did some extra optimization in the length ratios and width ratios of the transistors to get maximum curvature correction. The designed circuit measures 50µm * 50µm in standard 0.16µm bulk CMOS technology. Measurements were done on a Süss Microtec PM200 wafer prober with cooling, using a Keithley 4200 analyzer.
Fig. 4 shows the generated $V_{\text{REF}}$ and supply current $I_{\text{DD}}$ as a function of the supply voltage $V_{\text{DD}}$, for -45°C, 25°C and 145°C for one sample. Note that the $V_{\text{REF}}$ curves are almost overlapping; only the curve for 145°C shows a slight deviation from the others. As expected, the $I_{\text{DD}}$ curves are very much different, approximately corresponding to an offset proportional-to-temperature current. Fig. 5 shows $V_{\text{REF}}$ and $I_{\text{DD}}$ as a function of temperature for one sample (with a near average $V_{\text{REF}}$) at $V_{\text{DD}}=1.2V$; note that the $V_{\text{REF}}$-axis is zoomed in to be able to visualize any variation in the reference voltage. Also visible in the figure is the small curvature in the $I_{\text{DD}}$- curve which was designed for maximum curvature correction. Extrapolation of this non-linear curve shows a $T_{\text{min}} \approx -120^\circ\text{C}$ which nicely complies with the measured $V_{\text{REF}} \approx 940mV$. As for any BGVR, sufficiently low spread is essential; the upper part of Fig. 5 shows the measured $V_{\text{REF}}$ for 60 (untrimmed and unselected) samples at $V_{\text{DD}}=1.2V$ and at 23°C; the sample used for the lower part in Fig. 5 is sample 33. From this the average $V_{\text{REF}}$ equals 944mV with a 1 sigma spread figure $\sigma_{V\text{REF}}=0.76\%$. A micrograph of the chip is shown in Fig. 7, in which the sub-1V BGVR is inside the black square.

The comparison chart in Fig. 6 shows properties of a number of published sub-1V CMOS voltage reference circuits that include measured spread data. Standard bandgap voltage references producing the conventional 1.2V and BiCMOS circuits were not taken into account. The data in Fig. 6 shows that the presented sub-1V BGVR has good untrimmed performance (reference voltage spread $\sigma_{V\text{REF}}$ and temperature coefficient $TC$) at simultaneously a low power consumption and a low area consumption.

In summary, an ultra-compact sub-1V BGVR that circumvents the usual power-area trade off is presented, operating at 1.4μA at room temperature with supply voltages down to 1.1V,
generating $V_{\text{REF}}=0.94\text{V}$ with $\sigma=0.76\%$, with an active area of 50$\mu$m * 50$\mu$m in 0.16$\mu$m CMOS technology. These specifications enable local generation of reference voltages wherever needed.

Acknowledgements: We thank NXP for chip fabrication, and G. van der Weide, G. Wienk, H. de Vries and S. Smits for assistance, and M. Pelgrom for challenging us.

References:


Figure captions:

Figure 1: Principle schematic of the area efficient low-power sub-1V BGVR circuit: a conventional CTAT voltage ($V_D$) is added to a down-shifted PTAT voltage ($V_{R2}$).

Figure 2: Typical threshold voltage ($V_T$) dependency on transistor length in modern CMOS technologies for 3 temperatures: using a balanced pair using lengths $L_1$ and $L_2$ creates an almost constant offset voltage $\Delta V_T$.

Figure 3: Actual circuit implementation: the circuit of Fig. 2 with regulated cascades.

Figure 4: Measured output voltage $V_{REF}$ and supply current $I_{DD}$ of the BGVR as a function of supply voltage $V_{DD}$, at $T$= -45°C, $T$=25°C and $T$=145°C.

Figure 5: Upper: measured $V_{REF}$ on 60 samples at a supply voltage $V_{DD}$=1.2V at 23°C: average $V_{REF}$=944mV with $\sigma_{VREF}$=0.8%; lower: measured $V_{REF}$ and $I_{DD}$ as a function of temperature at $V_{DD}$ =1.2V for a near average sample (sample 33 in upper part)

Figure 6: Comparison of reported sub-1V voltage references in CMOS.

Figure 7: Micrograph of the chip fabricated in a standard 160nm bulk CMOS; the sub-1V bandgap voltage reference circuit is inside the black square.
Figure 1: Principle schematic of the area efficient low-power sub-1V BGVR circuit: a conventional CTAT voltage ($V_D$) is added to a down-shifted PTAT voltage ($V_{R2}$).
Figure 2: Typical threshold voltage ($V_T$) dependency on transistor length in modern CMOS technologies for 3 temperatures: using a balanced pair using lengths $L_1$ and $L_2$ creates an almost constant offset voltage $\Delta V_T$. 
Figure 3: Actual circuit implementation: the circuit of Fig. 2 with regulated cascades.
Figure 4: Measured output voltage $V_{REF}$ and supply current $I_{DD}$ of the BGVR as a function of supply voltage $V_{DD}$, at $T= -45^\circ C$, $T= 25^\circ C$ and $T=145^\circ C$. 

$V_{REF}$ [V]  
$I_{DD}$ [$\mu$A]  
$V_{DD}$ [V]
Figure 5: Upper: measured $V_{\text{REF}}$ on 60 samples at a supply voltage $V_{\text{DD}}=1.2\text{V}$ at 23°C:

average $V_{\text{REF}}=944\text{mV}$ with $\sigma_{\text{VREF}}=0.8\%$; lower: measured $V_{\text{REF}}$ and $I_{\text{DD}}$ as a function of temperature at $V_{\text{DD}}=1.2\text{V}$ for a near average sample (sample 33 in upper part)
<table>
<thead>
<tr>
<th></th>
<th>[2]</th>
<th>[6]</th>
<th>[7]</th>
<th>[8]</th>
<th>[9]</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology (all CMOS)</td>
<td>0.35μm</td>
<td>0.35μm</td>
<td>0.5μm</td>
<td>0.6μm</td>
<td>0.35μm</td>
<td>0.16μm</td>
</tr>
<tr>
<td>V&lt;sub&gt;DD&lt;/sub&gt;</td>
<td>&gt;1.4V</td>
<td>&gt;1V</td>
<td>&gt;0.95V</td>
<td>&gt;0.98V</td>
<td>&gt;0.85V</td>
<td>&gt;1.1V</td>
</tr>
<tr>
<td>V&lt;sub&gt;REF&lt;/sub&gt;</td>
<td>858mV</td>
<td>190.1mV</td>
<td>628mV</td>
<td>603mV</td>
<td>650mV</td>
<td>944mV</td>
</tr>
<tr>
<td>σ&lt;sub&gt;V&lt;sub&gt;REF&lt;/sub&gt;&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trimmed</td>
<td>~0.9%</td>
<td>0.26%</td>
<td>~0.5%</td>
<td>~1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>untrimmed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2%</td>
<td>0.8%</td>
</tr>
<tr>
<td>min temp</td>
<td>max temp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-20°C</td>
<td>120°C</td>
<td>-20°C</td>
<td>120°C</td>
<td>-40°C</td>
<td>120°C</td>
<td>0°C</td>
</tr>
<tr>
<td>TC</td>
<td>trimmed</td>
<td>12.4ppm/°C</td>
<td>16.9ppm/°C</td>
<td>17ppm/°C</td>
<td>&lt;25ppm/°C</td>
<td></td>
</tr>
<tr>
<td>untrimmed</td>
<td></td>
<td></td>
<td>670ppm/°C</td>
<td></td>
<td>60 ppm/°C</td>
<td>30ppm/°C</td>
</tr>
<tr>
<td>I&lt;sub&gt;DD&lt;/sub&gt;@~300K</td>
<td>116μA</td>
<td>0.25 μA - 0.56μA</td>
<td>&lt;10μA</td>
<td>&lt;18μA</td>
<td>1.2μA</td>
<td>1.4μA</td>
</tr>
<tr>
<td>Area</td>
<td>1.2mm²</td>
<td>0.049 mm²</td>
<td>1.1mm²</td>
<td>0.24mm²</td>
<td>0.063 mm²</td>
<td>0.0025mm²</td>
</tr>
</tbody>
</table>

Figure 6: Comparison of reported sub-1V voltage references in CMOS.
Figure 7: Micrograph of the chip fabricated in a standard 160nm bulk CMOS; the sub-1V bandgap voltage reference circuit is inside the black square.