A 2.2ps 2-D Gated-Vernier Time-to-Digital Converter with Digital Calibration

Ping Lu, Sr. Member, IEEE, Ying Wu, Student Member, IEEE, and Pietro Andreani, Sr. Member, IEEE

Abstract—This paper presents a 2-dimension (2-D) Vernier time-to-digital converter (TDC) which uses two 3-stage gated-ring-oscillators (GROs) in the X/Y Vernier branches. The already small Vernier quantization noise (~10.6ps) is improved by the 1st-order noise shaping of the GRO. Moreover, since all delay differences between X phases and Y phases can be used (rather than only the diagonal line of the 1-dimension architecture), the intrinsic large latency time of the Vernier architecture is dramatically reduced. The TDC is implemented in a 65nm CMOS process and consumes 2.3mA from 1.0V. The measured total noise integrated over a bandwidth of 1.25 MHz yields an equivalent TDC resolution of 2.2ps, while the average latency time (within 2ns) is less than 1/6 of that in a standard Vernier TDC.

Index Terms—Time to digital converter (TDC), Gated-Ring-Oscillator (GRO), Vernier, 2-dimension (2-D).

I. INTRODUCTION

HIGH-resolution TDCs have obtained more and more popularity due to their use in digital phase-locked loop (DPLL), analog-to-digital converter (ADC), jitter measurement, etc. However, besides the high resolution, a large detection range is also required in many applications. In a fractional divider-based DPLL, for example, the locked state forces a TDC to cover several oscillation periods, which usually amounts from several hundreds of picoseconds to several nanoseconds. Furthermore, while in the PLL lock-capture process, even the whole reference clock period (several tens of ns) should be detected by the TDC. Most open-loop delay-line TDCs (such as flash [1] and Vernier [2, 3]) either suffer from a low resolution or a small range or both, since the delay line cannot be too long. Time-amplification TDCs [4], on the other hand, have the linearity issue by the presence of two-step operation. Cyclic TDCs are good candidates for a large dynamic range, but the resolution of a ring-oscillator TDC cannot exceed an inverter delay, i.e. it is hard to achieve a very high resolution [5]. Noise-shaping cyclic TDCs, however, can give a good performance for both resolution and dynamic range. The basic noise-shaping cyclic TDC is based on a GRO, where the quantization error is accumulated across successive measurements, so that a resolution beyond the inverter delay is obtained [6]. Its generalized version switched-ring-oscillator (SRO) TDC provides the same noise shaping as a GRO TDC and avoids the dead zone (cannot detect linearly the input which is around integer times of GRO period) by using another running frequency instead of freezing the GRO state [7]. The work in [8] combines Vernier and GRO techniques (Gated-Vernier) to further improve the TDC resolution and meanwhile greatly mitigates charge leakage and redistribution, which otherwise increases the noise floor. Like all Vernier TDCs, it achieves a higher resolution at the penalty of a longer latency time.

In this brief, a new 2-dimension Gated-Vernier TDC (2-D GVTDC) with digital calibration of the delay ratio of the two GROs is proposed. This TDC targets applications in divider-based DPLLs, and cover the whole DPLL reference clock period thanks to a new edge-swallowed phase detector (ESPD). The proposed TDC not only improves the resolution by a factor of 5 (from 10.6ps to 2.2ps), but also dramatically reduces the latency time, compared to a standard Vernier TDC.

An overview of the paper is as follows: section II describes the proposed 2-D GVTDC; section III details the circuit implementation; section IV discloses the measurement results; finally, conclusions are drawn in section V.

II. EVOLUTION OF THE 2-D GATED-VERNIER TDC

A. GRO TDC

GRO is a special ring oscillator with its oscillation gated by an enable signal, as shown in Fig. 2. When the “enable” is on, it works like a regular oscillator. When the “enable” is off, the charge of each output node is frozen as if a current gate is

P. Lu is with Department of Electrical and Information Technology, Lund University, Lund, 22363, Sweden. (e-mail: Ping.Lu@eit.lth.se).
Y. Wu is with Department of Microelectronics/DIMES, Delft University of Technology, Delft, 2628CD, Netherlands. (e-mail: Y.Wu-1@tudelft.nl).
P. Andreani is with Department of Electrical and Information Technology, Lund University, Lund, 22363, Sweden. (e-mail: Pietro.Andreani@eit.lth.se).
Adding the Vernier technique to the GRO TDC brings two benefits. One is that the raw resolution (quantization step interval \( \Delta \)) is not limited by (3), the latency time for some output values can be greatly reduced. As an example, let us assume that an input time interval of 38\( \Delta \) is sent in a 3-stage GVTDC with \( k=18 \). According to (1) and (3), the GVTDC can give the result after 38\( \tau_1 \) (i.e., FGRO start catching up with SGRO when \( X=Y=38 \)). However, it is easy to find that the minimum \( X \) to meet (1) is 4 (when \( Y=2 \)), in which case the latency is only 4\( \tau_1 \), without the limitation of (3). To achieve such a quick detection, an extra arbiter must be placed between the 1\( ^{st} \) SGRO phase (the 4\( ^{th} \) phase is just the 1\( ^{st} \) one since (3) can also be rewritten as \( X=\text{mod}(X,N) \)) and the 2\( ^{nd} \) FGRO phase. To get the priority for any input value, arbitration must be done between any two SGRO and FGRO phases, not only between the phases with the same \( p^b \).

By doing this, a Vernier line is extended to a Vernier plane (2-D), where the delay differences between all phase pairs are employed [9].

### 2-D Gated-Vernier TDC

From (1), we can see that if \( X \) and \( Y \) can be chosen more flexibly and not limited by (3), the latency time for some output values can be greatly reduced. As an example, let us assume that an input time interval of 38\( \Delta \) is sent in a 3-stage GVTDC with \( k=18 \). According to (1) and (3), the GVTDC can give the result after 38\( \tau_1 \) (i.e., FGRO start catching up with SGRO when \( X=Y=38 \)). However, it is easy to find that the minimum \( X \) to meet (1) is 4 (when \( Y=2 \)), in which case the latency is only 4\( \tau_1 \), without the limitation of (3). To achieve such a quick detection, an extra arbiter must be placed between the 1\( ^{st} \) SGRO phase (the 4\( ^{th} \) phase is just the 1\( ^{st} \) one since (3) can also be rewritten as \( X=\text{mod}(X,N) \)) and the 2\( ^{nd} \) FGRO phase. To get the priority for any input value, arbitration must be done between any two SGRO and FGRO phases, not only between the phases with the same \( p^b \).

By doing this, a Vernier line is extended to a Vernier plane (2-D), where the delay differences between all phase pairs are employed [9].

### Gated-Vernier TDC (GVTDC)

Adding the Vernier technique to the GRO TDC brings two benefits. One is that the raw resolution (quantization step before shaping) is not limited by the inverter delay, since the inverter delay difference \( \tau_1 - \tau_2 \) is used, rather than inverter delay itself (Fig. 3(a), where \( \tau_1 \) is the slow GRO (SGRO) delay and \( \tau_2 \) is the fast GRO (FGRO) delay). The enable pulse width is not equal to the input interval any more, but determined by the moment that FGRO phases start leading SGRO phases (catch up) which is detected by arbiters. This directly results in the other benefit of a small range for the frozen levels (see the red circles in Fig.3(b)), so that the gating skew due to the charge redistribution is much smaller than with a standard GRO TDC.

To make it easier to understand, only a single phase is illustrated in Fig. 3(b) and Fig. 2(b). In practice, each phase of the ring oscillator is used to provide the raw quantization. Since the absolute delay of each GRO does not set the TDC quantization any more, a relatively large load capacitance can be used in the GRO cell, further reducing the impact of charge redistribution. The large load capacitance also makes the frozen voltage insensitive to the leakage current. Although the GVTDC achieves a higher resolution and a better immunity to the device non-ideality, it suffers from a large latency time, as shown in Fig. 1. The digitized output of an \( N \)-stage GVTDC can be expressed as:

\[
\text{out}_{TDC} = k \times X - (k - 1) \times Y \tag{1}
\]

\[
k = \frac{\tau_1}{\Delta}, \; k - 1 = \frac{\tau_2}{\Delta}, \; \Delta = \tau_1 - \tau_2 \tag{2}
\]

where \( X \) and \( Y \) denote the taps of SGRO and FGRO delay during the enables (enable_S & enable_F) are on, respectively, within the measurement window. Obviously,

\[
X = Y + m \times N; \; (m = 0, 1, 2, ... \tag{3}
\]

since the arbiters only detect the phase relationship between the \( p^b \) SGRO delay cell and the \( p^b \) FGRO delay cell (\( p \leq N \)), where \( m \) denotes the number of times the oscillation has travelled around the GRO (i.e., number of laps). The larger \( N \) is, the longer the latency time will be due to the worse discontinuity of \( X \), as shown in Fig.4.

### 2-D Gated-Vernier TDC

From (1), we can see that if \( X \) and \( Y \) can be chosen more flexibly and not limited by (3), the latency time for some output values can be greatly reduced. As an example, let us assume that an input time interval of 38\( \Delta \) is sent in a 3-stage GVTDC with \( k=18 \). According to (1) and (3), the GVTDC can give the result after 38\( \tau_1 \) (i.e., FGRO start catching up with SGRO when \( X=Y=38 \)). However, it is easy to find that the minimum \( X \) to meet (1) is 4 (when \( Y=2 \)), in which case the latency is only 4\( \tau_1 \), without the limitation of (3). To achieve such a quick detection, an extra arbiter must be placed between the 1\( ^{st} \) SGRO phase (the 4\( ^{th} \) phase is just the 1\( ^{st} \) one since (3) can also be rewritten as \( X=\text{mod}(X,N) \)) and the 2\( ^{nd} \) FGRO phase. To get the priority for any input value, arbitration must be done between any two SGRO and FGRO phases, not only between the phases with the same \( p^b \).

By doing this, a Vernier line is extended to a Vernier plane (2-D), where the delay differences between all phase pairs are employed [9].

Fig.5 gives an example diagram with \( N=3 \) and \( k=18 \). The horizontal lines denote FGRO phases and the vertical lines SGRO phases. A general quantization array in a 2-D TDC is given in Fig.5(b). For a delay-line Vernier TDC, the arbitration track is only at the diagonal of the array of Fig.5(b), i.e. at “1, 2, 3…”; for a 3-stage GVTDC, the track (due to the circulation in Fig.4) includes “1, 2, ..., 54”, “55, 56, ..., 108”, “109, 110, ..., 162”... and for a 11-stage GVTDC, the track includes “1, 2, 198”, “199, 200, ..., 396”, “397, 398, ..., 594”... It is easy to see that the detection track of an \( N \)-stage GVTDC is delay-equivalent to delayed Vernier TDC due to the ring operation, but still cannot...
completely cover the plane of Fig.5(b). In addition, a larger stage \( N \) results in more numbers missing in the plane. A 2-D GVTDC, however, can cover the whole plane without any gap, yielding a minimum latency time. Assuming \( T_{in} \) is the input time interval to the 2-D GVTDC, the normalized latency time \( \tau_1 \) with respect to \( T_{in} \) is derived from (3) and Fig. 5(b):

\[
\tau_1 = \left\lfloor \frac{T_{in}}{k} \right\rfloor + \left\lceil \frac{T_{in}}{k} \right\rceil - \left\lfloor \frac{T_{in}}{k} \right\rfloor \times k
\]

(4)

where \( \lfloor a \rfloor \) denotes the maximum integer that is not larger than \( a \), and \( \lceil a \rceil \) denotes the minimum integer that is not smaller than \( a \). \( \lceil \frac{T_{in}}{k} \rceil \) is just \( \text{out}_{TDC} \). Basically, the latency improvement increases with an increasing \( k \), but not monotonically.

It can be seen from Fig.5 (b) that even negative inputs can be detected directly. However, processing positive and negative inputs together requires a complicated arbiter, which burdens the GRO output buffers and aggravates the influence of the arbiter metastability. We therefore precondition the input by a signed phase detector, and send only positive time interval to the GROs.

Both GVTDC and 2-D GVTDC need to define \( k \) beforehand. Theoretically, \( k \) in a GVTDC is not necessarily an integer, but it is in a 2-D GVTDC. Analog tuning of the delay does not have a good immunity to noise [10]. In this work, a digital calibration is adopted which accurately defines \( k \) within \( \pm 1.25\% \), meanwhile effectively compensating for process and temperature variations. This will be detailed in the next section.

III. CIRCUIT DETAILS

The proposed 2-D GVTDC consists of two 3-stage GROs, arbiter array, phase detector, digital decoder and a digital delay calibration, as shown in Fig.5(a). It should be noted that \( k=18 \) is only for the sake of example. In the practical chip, \( k \) is defined

Fig. 5 2-D GVTDC (a) diagram; (b) quantization array (assume \( k=18 \)).

A. Vernier GROs and Digital Calibration

In the actual implementation, \( k \) is usually not an integer and hard to predict due to GRO delay fluctuations. A simple method to accurately detect the GRO delay is to count its cycles/periods within a known long time. For example, if the number of GRO periods is 1000 (after averaging) during 900ns, the GRO delay is 900ns/1000/3≈300ps \( (N=3 \text{ and only rise edges are considered}) \) with the error of 900ns/3/300ns \( \approx 0.5 \). We can thus calibrate the GRO delay by comparing the number of periods with a series of reference numbers (corresponding different \( k \) and GRO delays) stored in a table (look-up-table). Similar with the coarse-fine tuning approach, these different references cover a wide delay range (coarse), while the fine tuning is implemented by a switched unit capacitor bank. In this work, a metal fringe capacitor is employed for its very small capacitance per unit area. By using 200aF/LSB in the capacitor bank, each GRO achieves a delay step of ~82fs. With an 8-bit input, the bank well covers the delay gap between two adjacent coarse tunings.

Fig.6 shows the diagram of GRO cell with digital calibration scheme. In the calibration mode, GRO gates are always on. Besides the gated inverter core, each GRO cell has an 8-bit switchable unit capacitor load, which is controlled by the calibration block. In Fig.6, look-up-table (LUT) denotes the coarse calibration where 28 addresses store 28 pairs of reference period numbers, covering a delay range of around 80ps~250ps (see table in Fig.6). To relieve the speed bearing of digital counter, the divided-by-2 clock of GRO (rather than GRO output) is used, which allows a counting duration of 960ns to provide a delay error better than ±80fs (if \( C_5 =1000 )\), as calculated before.

Fine calibration starts with a default address, trying to tune
GRO delays to match the reference numbers \((C_S\) and \(C_F\) in Fig.6) in this address. If the 8-bit tuning word of either GRO keeps in the extreme values (“0” or “255”), it indicates that GRO delays are outside of this region. Coarse calibration will therefore jump to the next address to do the same operation till the correct and stable GRO delay ratio is obtained.

After calibration, the phase detector switches the TDC to the normal working mode while the calibration block is disabled, maintaining however the fixed tuning words for the GROs. Meanwhile, the LUT sends the \(k\) value according to the final address to the digital decoder for the calculation of the TDC output.

**B. Edge-Swallowed Phase Detector**

In the 2-D GVTDC, the phase detector is used to generate enables for both GROs. Unlike a regular GRO TDC, the enable pulse width for Vernier GROs is not the same with the input interval (see Fig.3). The gates of the GROs are closed by a reset when FGRO catches up with SGRO, not the slow edge of the interval (see Fig.3). Thus, the phase detector is more complex than a simple D flip-flop. Furthermore, as mentioned above, the phase detector also works for input sign detection. However, the most serious concern is the possibility of edge missing due to the latency time, i.e., when the input interval is very large (close to one reference period), the latency time might mask the next input edge, yielding a wrong detection. To avoid this, an edge swallowed phase detector (ESPD) is proposed, as illustrated in Fig. 7.

In the bottom-right dashed square of Fig.7, waveforms illustrate how the input clock is masked in a traditional phase detector when the input interval is large, and how to eliminate this effect in our work. In this case, \(CLK_S\) leads \(CLK_F\) by a large amount, so the pulse width of \(S1\) should be always wider than the pulse width of \(S2\) and the sign should be “0” (positive). However, FGRO needs some time to catch up SGRO (Vernier operation) after \(F1\) becomes high. In this case, it may happen that the next rise edge of \(CLK_S\) arrives before the input D flip-flops are reset, in such a way that it cannot trigger a valid enable. The skipping of \(CLK_S\) rise edge results in the subsequent wrong phase detection, both in sign and value. However, if the corresponding \(CLK_F\) rise edge can be also skipped (swallowed), the output will skip the edge pair, but it will yield the correct detection for the 3rd edge pair, as shown in Fig.7. Since this case is relevant only during the initial DPLL lock-capturing process, the ensuing TDC non-linearity is not important. Once the DPLL is locked, relatively small phase errors make the TDC always work linearly, without any swallowing. The detailed schematic realization is shown in the left of Fig. 7. During normal operation, the part in the dotted region does not work and RSS/RFF is always low due to the reset by \(CLK_F/CLK_S\) rise edges. However, if e.g. \(CLK_S\) rise edge arrives when \(S1\) is still high (i.e. the system is not reset yet), it immediately converts RSS to high level, resetting the system (see the waveform of Fig.7). The reset state will not be changed until the corresponding \(CLK_F\) rise edge arrives to reset RSS. Thus, the \(CLK_F\) rise edge also cannot trigger the output, as if it is swallowed by the system. The ESPD does not produce any wrong sign at the output, and always yields the approximately correct detected results at least every second period, when the input interval is very close to the clock period. Thus guarantees a DPLL to lock correctly, and will not affect the performance in the locked state.

The sign detection and calibration control are also shown in Fig 7. It can be seen that SGRO always leads FGRO at the start of each measurement, since enable_S is always triggered earlier by the cross-over connection. GRO calibration control is simply realized by using NAND and NOR gates. The detected time difference will be isolated from ESPD output when calib_enable is high, so that enable_S and enable_F are always high to maintain GROs always oscillating.

**C. Digital Decoder**

When the GROs oscillate within their enable windows, the arbiter array compares the phase difference of each phase pair to find the catch-up location in the plane. A digital decoder then translates the coordinate of the location to a TDC output. According to (1), the edge number of two GROs is needed \((k)\) is given by GRO calibration). Using multiple-phase counter is the easiest way to get \(X\) and \(Y\), but it would consume much power [10]. We therefore only count the lap number and add the remainder by comparing the coordinate of starting and ending phase within a measurement window.
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCSII.2016.2548218, IEEE Transactions on Circuits and Systems II: Express Briefs

Higher noise floor. The TDC detection range measurement for GROs. The severe nonlinearity results in more spurs and a step of 2.2ps would be needed. The low spurs in the spectrum with the same OSR, but assuming a white PSD, a flash-TDC equal to \(\approx 141\text{fs}\) integrated noise with an oversampling ratio (OSR) of 20 is quantization noise shaping following the expected first-order window on 16 sequential records before averaging. The terminal is tuned by a sinusoidal voltage. After GRO (Skyworks). The buffer is loaded by the varactor, whose other delay generator by using the low-jitter buffer Si53306-B-GMR (Silicon Laboratories) and the varactor SMV2019-079LF (Skyworks). The buffer is loaded by the varactor, whose other terminal is tuned by a sinusoidal voltage. After GRO calibration, resulting in \(k=15\) and \(\Delta=10.6\text{ps}\), the measured PSD of the TDC output, with sampling frequencies \(F_3=50\text{MHz}\), is shown in Fig. 9(a), where the 32768-point FFT employs a Hann window on 16 sequential records before averaging. The quantization noise shaping following the expected first-order 20dB/decade slope is clearly visible. At \(F_3=50\text{MHz}\), the integrated noise with an oversampling ratio (OSR) of 20 is equal to \(\approx 141\text{fs}_{\text{rms}}\). To achieve the same integrated noise level with the same OSR, but assuming a white PSD, a flash-TDC step of 2.2ps would be needed. The low spurs in the spectrum are most likely caused by the off-chip varactor, given its nonlinear capacitance vs. bias voltage curve. As a comparison, fig.9 (b) shows the measured PSD without digital calibration for GROs. The severe nonlinearity results in more spurs and a higher noise floor. The TDC detection range measurement is also given in Fig.10, which clearly demonstrates the effect of ESPD for a full-period input range.

At 50MHz input frequency, the TDC consumes 2.3mA current from a 1.0V supply. A performance summary and comparison with other state-of-the-art works are given in TABLE I.

V. CONCLUSION

A 2-dimension Gated-Vernier TDC with digital GRO calibration has been presented. The 65nm CMOS prototype achieves an equivalent TDC resolution of 2.2ps with an OSR of 20 and a sampling frequency of 50MHz. Its high resolution and greatly improved latency time for unlimited input range make it very attractive for use in DPLL applications.

TABLE I PERFORMANCE AND COMPARISONS

<table>
<thead>
<tr>
<th>Scheme</th>
<th>GRO</th>
<th>GRO</th>
<th>SRO</th>
<th>MASH (\Delta)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tech. [(\text{nm})]</td>
<td>90</td>
<td>130</td>
<td>90</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td>(F_3) [(\text{MHz})]</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>OSR</td>
<td>20</td>
<td>25</td>
<td>25</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>BW [(\text{MHz})]</td>
<td>1.25</td>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Resolution [(\text{ps})]</td>
<td>2.2</td>
<td>1.4</td>
<td>17</td>
<td>307</td>
<td></td>
</tr>
<tr>
<td>Input Range [(\text{ns})]</td>
<td>20 (input period)*</td>
<td>12</td>
<td>20</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Supply [(\text{V})]</td>
<td>1.0</td>
<td>1.5</td>
<td>1.2</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>Power [(\text{mW})]</td>
<td>2.3</td>
<td>21</td>
<td>2</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>Area [(\text{mm}^2)]</td>
<td>0.068</td>
<td>0.04</td>
<td>0.02</td>
<td>0.11</td>
<td></td>
</tr>
</tbody>
</table>

* The input range can be enlarged when using a lower sampling frequency

is given in Fig.10, which clearly demonstrates the effect of ESPD for a full-period input range.

A 2-D Gated-Vernier TDC with digital GRO calibration has been presented. The 65nm CMOS prototype achieves an equivalent TDC resolution of 2.2ps with an OSR of 20 and a sampling frequency of 50MHz. Its high resolution and greatly improved latency time for unlimited input range make it very attractive for use in DPLL applications.

IV. MEASUREMENTS

The 2-D GVTDC, implemented in 65nm CMOS (Fig.8), occupies an active area of 0.068mm\(^2\). Generally, a sinusoidal delay between two same-frequency clocks is used in TDC measurements. In our measurements, the Tektronix DTG5274 Data Timing Generator always has a relatively large 2.5MHz spur and its harmonics present on the generated sinusoidal delay, making a quantitative analysis of the power spectral density (PSD) impossible. We therefore set up an off-chip delay generator by using the low-jitter buffer Si53306-B-GMR (Silicon Laboratories) and the varactor SMV2019-079LF (Skyworks). The buffer is loaded by the varactor, whose other terminal is tuned by a sinusoidal voltage. After GRO calibration, resulting in \(k=15\) and \(\Delta=10.6\text{ps}\), the measured PSD of the TDC output, with sampling frequencies \(F_3=50\text{MHz}\), is shown in Fig. 9(a), where the 32768-point FFT employs a Hann window on 16 sequential records before averaging. The quantization noise shaping following the expected first-order 20dB/decade slope is clearly visible. At \(F_3=50\text{MHz}\), the integrated noise with an oversampling ratio (OSR) of 20 is equal to \(\approx 141\text{fs}_{\text{rms}}\). To achieve the same integrated noise level with the same OSR, but assuming a white PSD, a flash-TDC step of 2.2ps would be needed. The low spurs in the spectrum are most likely caused by the off-chip varactor, given its nonlinear capacitance vs. bias voltage curve. As a comparison, fig.9 (b) shows the measured PSD without digital calibration for GROs. The severe nonlinearity results in more spurs and a higher noise floor. The TDC detection range measurement

REFERENCES
