On the Design of a Linear Delay Element for the Triggering Module at CERN LHC

Jordan Lee Gauci*, Edward Gatt*, Owen Casha*, Giacinto De Cataldo†, Ivan Grech* and Joseph Micallef*
*Department of Microelectronics and Nanoelectronics, University of Malta
†Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Italy
*E-mail: jordan-lee.gauci.10@um.edu.mt

***Abstract*—This paper presents an analytical model of a linear delay element circuit to be employed in the triggering module for the High Momentum Particle Identification Detector (HMPID) at the CERN Large Hadron Collider (LHC). The aim of the analytical model is to facilitate the design of the linear delay element circuit, while maximizing its linearity and delay range. The analytical model avoids the need of time consuming parametric sweeps on the aspect ratios of the various transistors of the delay element in order to optimize it. In addition, the analytical model can be used to predict the variation of the delay with the input tuning voltage. The proposed analytical model is verified via the simulation of the delay element circuit using the $0.18\,\mu\mathrm{m}$ X-FAB technology.**

***Index Terms*—Delay Lines, Delay Range, Linearity, Jitter, Modelling**

I. Introduction

Precise delay generation is an active research area due to the employment of delay generators in high-energy physics, time-of-flight experiments, time-to-digital converters and time interval measurement circuits. The core element of the delay generator is the delay line, which is a device capable of delaying the input signal by a predefined value. This delay can either be fixed or variable. In the case of variable delay lines, the delay can be tuned by both an analog or a digital mechanism. Delay lines have four important characteristics which determine their performance [1]–[3]:

- **Delay Step**: The finest incremental time delay step that can be produced by the delay line.
- **Delay Range**: The maximum time by which a signal can be delayed.
- **Linearity**: The ability to achieve equal and uniform delay steps.
- **Jitter**: Variation in the delay of the output signal due to noise, which has a direct effect on the smallest delay step that may be generated.

Some analog delay lines have fixed delays, such as those employing transmission gate delay elements and inverter based delay elements. In these cases, the generated delay mainly depends on the dimensions of such devices. However, the delay may be tuned either by modulating the power supply voltage or through the use of additional circuitry such as a current-starved inverter, where by tuning the quiescent current, the delay is increased or decreased accordingly. Although this type of delay element has a non-linear transfer characteristic, techniques were proposed to linearize it [4]. The work in [5], proposes to add diode connected transistors to the inverter in order to generate very linear delays. Thyristor-based delay elements can also be used, where long delay ranges can be achieved at the expense of area and higher power consumption [6]. This paper presents an analytical model of the delay element proposed in [7]. The aim of this model is to facilitate the design of the delay element circuit, while maximizing its linearity and delay range. Compared to the design methodology employed in [7], this analytical model avoids the need of time consuming parametric sweeps on the aspect ratios of the various transistors of the delay element in order to optimize it. The analytical model can also be used to predict the variation of the delay with the input tuning voltage.

II. Background

This work will aid the design of the trigerring module (better known as the Fan-In/Fan-Out module) to be used by the High Momentum Particle Identification Detector (HMPID) at the CERN Large Hadron Collider (LHC). HMPID is a triggered detector, where the signals on the detector pads are read on receiving a trigger signal from the Central Trigger Processor [3]. Currently, during Run1 and Run2, the HMPID receives the Level 0 (L0) trigger signal, after approximately $1.2\,\mu\mathrm{s}$, which corresponds exactly to the peaking time of the signal on the pads. A digital delay generator is used to coarsely adjust the delay on the trigger signal in steps of $25\,\mathrm{ns}$. Because of the upgrade that will occur during the second long shutdown (2019-2021), the HMPID will use the Level Minus (LM) trigger instead of L0. This signal arrives at approximately $700\,\mathrm{ns}$ after a collision. This means that a new delay generator is required to fine-tune the delay on the trigger. In [3], an improved digital delay generator was proposed as a preliminary solution. However, results have shown the issue of the creation of a random offset in the generated delay profile which is not deterministic [3]. An analogue delay generator is therefore being considered to overcome this issue, and an analytical model to facilitate the design of the chosen delay element is being proposed.

III. Delay Element Circuit

The current-starved inverter architecture suffers from a non-linear relationship between the delay and the tuning voltage while having an input range limited to the from V_t to V_{DD}, where V_t is the threshold voltage and V_{DD} is the supply

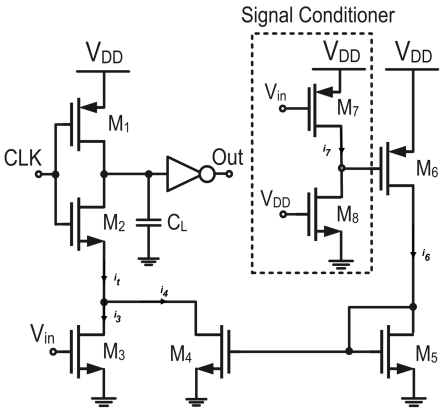

Fig. 1. Linear Delay Element Circuit [7].

voltage. The delay time (T_d) between the input and the output signal of a current-starved inverter can be modelled by [7, 8]:

$$T_d \propto \frac{C_L}{I_{cp}} V_{DD} \tag{1}$$

where C_L represents the capacitive load of the inverter and I_{cp} is the charging/discharging current through the capacitive load. Eq. 1 shows that the delay may either be varied through the control of C_L, I_{cp}, or V_{DD}. Eq. 1 further illustrates how the delay is non-linear, as it is inversely proportional to the current. Furthermore, the constant of proportionality is equal to $\ln(2)$ [8].

To overcome the linearity problem, the work in [7] presents a low power linear delay element circuit, shown in Fig. 1. The circuit is based on a current-starved inverter architecture and can obtain both a linear delay and rail-to-rail operation. This is achieved through the addition of transistors M4-M8. If the tuning voltage, V_{in}, is applied directly through transistor M_6, the delay response of the circuit would be highly non-linear and non-monotonic. For this reason, an inverting common-source amplifier is added consisting of M_7 and M_8 and this helps in achieving a monotonic and quasi-linear relationship in the delay response of the circuit.

The current, i_2, through transistor M_2 determines the delay time of the circuit, and it consists of the currents i_3 and i_4. When $V_{in} < V_{tn}$, M_3 works in the sub-threshold region, while M_6 works in saturation. The delay in this case is therefore primarily dependent on the current through M_6. For values of V_{in} in the region $V_{tn} < V_{in} < V_{DD} - V_{tp}$, both M3 and M6 are on and a linear delay-voltage characteristic may be achieved. When $V_{in} > V_{DD} - V_{tp}$, M_7 is switched off and thus the current through M_6 would be saturated. In this case, the delay would depend on the M_3. Transistor sizing optimization was performed by means of a parameteric sweep, in order to obtain the most linear voltage-delay characteristic, particularly when both M_3 and M_6 are turned on [7].

In particular, the sizing of M_3 and M_4, were chosen to be as small as possible such that the maximum voltage-to-delay gain is achieved. The size of M_5 and M_6 was minimized to reduce the current. The sizing of the transistors of the inverting common-source amplifier, M_7 and M_8, was chosen specifically to maximize the gain without affecting the linearity. As such, M_7 is chosen as large and M_8 as small. However, if M_8 is chosen too small, there would be a large voltage swing on the gate of M_6 and this will have a negative effect on linearity [7].

IV. Analytical Model

An analytical model for the current of the delay element shown in Fig. 1 is presented in this section. This model is valid for the range $V_{tn} \leq V_{in} \leq V_{dd} - V_{tp}$. Two versions of the model are presented in this section: a first-order model and a second-order model. In addition, an approximation method is used in order to linearize the delay transfer characteristic of the circuit. Let i_3 be the current through M3, i_4 be the current through M4, and so on. The total current, i_t, is equal to:

$$i_t = i_3 + i_4 \tag{2}$$

$$= \frac{K'_n}{2}\frac{W_3}{L_3}(V_{in} - V_{tn})^2 + \frac{\frac{W_4}{L_4}}{\frac{W_5}{L_5}} i_5 \tag{3}$$

The current $i_6 = i_5$ is equal to

$$i_6 = \frac{K'_p}{2}\frac{W_6}{L_6}(V_{DD} - V_x - V_{tp})^2 \tag{4}$$

where

$$V_x = i_7 r_{ds8} = \frac{K'_p}{2}\frac{W_7}{L_7}(V_{DD} - V_{in} - V_{tp})^2 r_{ds8} \tag{5}$$

and r_{ds8} is the drain to source resistance for M_8. To simplify some terms,

$$\mu_1 = \frac{K'_n}{2}\frac{W_3}{L_3} \tag{6}$$

$$\mu_2 = \frac{\frac{W_4}{L_4}}{\frac{W_5}{L_5}}\frac{K'_p}{2}\frac{W_6}{L_6} \tag{7}$$

This means that i_t is equivalent to:

$$i_t = \mu_1(V_{in} - V_{tn})^2 + \mu_2(V_{DD} - V_x - V_{tp})^2 \tag{8}$$

Eq. 8 may be expanded to a fourth order polynomial and therefore the current equation may be expressed as:

$$i_t = \alpha_4 V_{in}^4 + \alpha_3 V_{in}^3 + \alpha_2 V_{in}^2 + \alpha_1 V_{in} + \alpha_0 \tag{9}$$

where $\alpha_4, \alpha_3, \alpha_2, \alpha_1, \alpha_0$ are functions of the transistor aspect ratios $\frac{W}{L}$, the process parameters K'_p and K'_n, the supply voltage and the threshold voltages of the NMOS and PMOS transistors, V_{tn} and V_{tp}, respectively. This implies that the delay would be equal to

$$T_d \propto \frac{1}{i_t} = \frac{1}{\alpha_4 V_{in}^4 + \alpha_3 V_{in}^3 + \alpha_2 V_{in}^2 + \alpha_1 V_{in} + \alpha_0} \tag{10}$$

The above equation assumes that r_{ds8} has a constant value. In reality, this is not true since it depends on the current i_8 and the drain-source voltage $V_{ds8} = V_x$. Thus a better model of V_x should be taken. To achieve this, we know that the current through M_7 is equal to the current through M_8. Thus,

$$i_7 = \frac{K'_p}{2}\frac{W_7}{L_7}(V_{dd} - V_{in} - V_{tp})^2 \quad (11)$$

$$= K'_n\frac{W_8}{L_8}((V_{dd} - V_{tn})V_x - \frac{V_x^2}{2}) \quad (12)$$

From Eq. 11 and Eq. 12, V_x, can be found (Eq. 13).

$$V_x = V_N - \sqrt{(V_N)^2 - 2\frac{K_p}{K_n}(V_{dd} - V_{in} - V_{tp})^2} \quad (13)$$

where $V_N = V_{dd} - V_{tn}$, $K_p = K'_p\frac{W_7}{L_7}$ and $K_n = K'_n\frac{W_8}{L_8}$. Eq. 10 shows that the time delay that can be achieved is a highly non-linear function. However linearity may be increased through the choice of values of $\alpha_4, \alpha_3, \alpha_2, \alpha_1$, and α_0. The Lagrange Polynomial approximation was used to transform Eq. 10 into a second order polynomial equation. The second order equation is a good approximation to the equation as it will provide two degrees of freedom, in terms of the coefficients of V_{in}, and V_{in}^2. The former would enable to control the range, while the latter would permit to control the linearity. Thus, Eq. 10 may be rewritten in the form of:

$$T_d \approx AV_{in}^2 + BV_{in} + C \quad (14)$$

where the coefficients A, B, and C can be directly related to the parameters of the transistors. A closer examination of the above coefficients shows that the coefficient A, can be minimized through an increase in the dimensions of transistor M_5, a decrease in the sizing of M_3, and an increase in r_{ds8} (decrease in $\frac{W_8}{L_8}$). This however will have an effect on the second coefficient, B. When the aforementioned parameters are changed, the gradient (and therefore the range) changes significantly. This is due to the fact that the denominator of the three coefficients is the same, and it is a function of the transistor ratios $M_3 - M_8$. This therefore implies that in this architecture there is a trade-off between delay range, resolution, and linearity.

V. Model Verification

The circuit illustrated in Fig. 1 was implemented and simulated in Cadence using the 0.18 μm X-FAB technology, with the transistor aspect ratios, as implemented in [7] (refer to Table I). A scaling factor was added to the dimensions such that any channel-length modulation effects are minimized. In addition, each individual transistor was characterized to find the process parameters K'_n and K'_p together with the transistors' respective threshold voltages.

An input pulse was applied to the CLK terminal while varying the tuning voltage, V_{in}. The delay was calculated by measuring the time difference between the input and output waveforms, when the voltage reaches 50% of the final value. A MATLAB script was written in order to test the analytical model. The script contains the equations for the currents in each branch of the circuit. Through a DC analysis, with an input voltage of 1 V, it was found that there is a good relationship between the values, thus showing that the

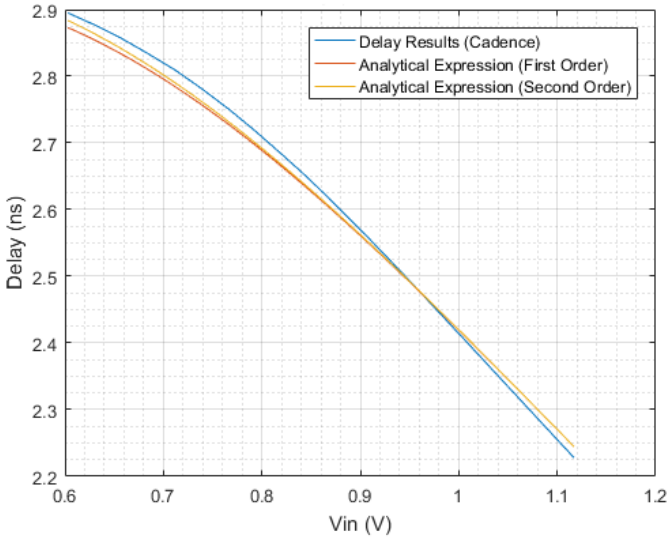

Fig. 2. Comparison of the simulated results with those obtained from the first order and second order analytical models.

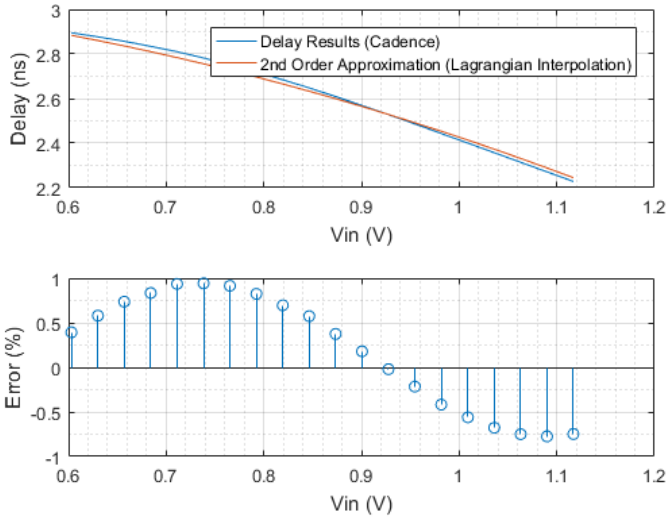

Fig. 3. Comparison of the simulated delay variation with that obtained from a second order Lagrangian polynomial.

TABLE I
Transistor Aspect Ratios.

Transistor Name	Width (μm)	Length (μm)
M1	6.9	0.18
M2	3.5	0.18
M3	0.225	0.25
M4	0.69	0.18
M5	2.08	0.18
M6	2.21	0.18
M7	3.18	0.18
M8	1.94	0.18

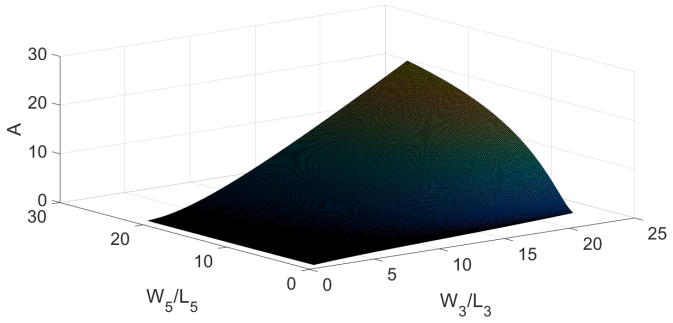

Fig. 4. Variation of the quadratic coefficient A with the aspect ratio of transistors M_3 and M_5.

equations describing the current are correct. A voltage sweep on the tuning voltage, in the range V_{tn} to $V_{DD} - V_{tp}$ was then performed and the delay results are illustrated in Fig. 2, where the delay is shown in blue. The first order analysis is illustrated in orange, and the improved analysis is illustrated in yellow. It can be seen that there is a good relationship between the analytical results and the simulated delay. There is, however, a discrepancy between the analytical results and the obtained result. This can be attributed to parasitic effects.

The Lagrangian interpolation method was used to approximate the delay curve by a second-order polynomial. The results are shown in Fig. 3, where the simulated delay results from Cadence are plotted in blue, and the approximation is plotted in orange. The error is also plotted in the subplot, where the error illustrated reaches a maximum of 0.94%. This therefore shows that the proposed model can be used to effectively approximate the delay transfer charactertistic with a second-order polynomial. To this end, different values for the transistors aspect ratios can be modified to improve linearity and range. As stated in Section IV, the coefficient of the term in V_{in}^2, can be minimized through an increase in r_{ds8}, an increase in the sizing of $M3$, or a decrease in the sizing of M_5. The aspect ratio of M_8 was decreased to $700\,\mathrm{nm}/180\,\mathrm{nm}$, which yields an r_{ds} value of around $1.7\,\mathrm{k}\Omega$. A sweep on the dimensions of M_3 and M_5 was then performed with the coefficient in V_{in}^2 worked out on each iteration. The results are presented in Figure 4. From this plot, the values of the aspect ratios of M_3 and M_5, that yields the lowest value of the coefficient of V_{in}^2 can be selected. These aspect ratios were calculated to be 0.9 and 20.8, respectively.

Transient simulations were performed in order to obtain the variation of the delay with the input tuning voltage for different supply voltages. The results are reported in Fig. 5 and are compared to those estimated using the Lagrangian Interpolation analytical model. There is good agreement between the plots, which shows that the Lagrangian Interpolation for Eq. 14 can be successfully used to set the transistor aspect ratios in order to obtain the most linear response.

VI. Conclusion

This paper presented an analytical model for the current-starved inverter proposed in [7]. A model was derived from

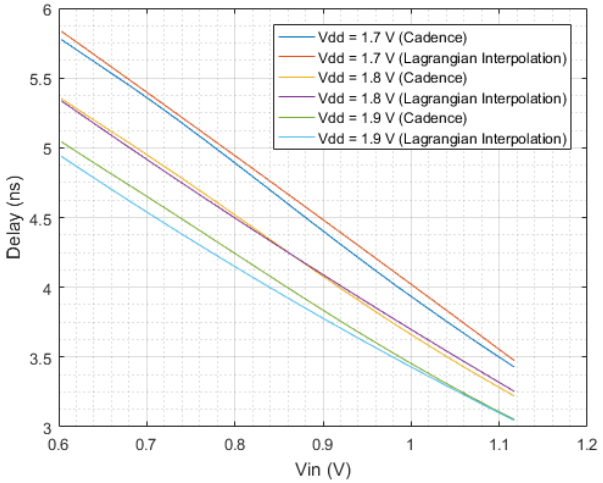

Fig. 5. Plot of the delay versus the input tuning voltage for the optimised linear delay element circuit. The plot presents both the simulation results and the second order approximation results obtained by means of the analytical model, for different values of V_{dd}.

first principles, where the delay was expressed in terms of the current and the input tuning voltage. This resulted in an inverse polynomial equation. The equation was approximated using the Lagrangian interpolation method and by means of this approximation, the transistor aspect ratios were found in order to achieve a linear delay transfer characteristic. The design was validated via simulation using the X-FAB $0.18\,\mu\mathrm{m}$ technology.

Acknowledgements

The research work disclosed in this publication is funded by the ENDEAVOUR Scholarship Scheme (Malta). The scholarship may be part-financed by the European Union - European Social Fund (ESF) under Operational Programme II - Cohesion Policy 2014-2020, *"Investing in human capital to create more opportunities and promote the well being of society."*

References

[1] B. Abdulrazzaq, I. A. Halin, S. Kawahito, R. M. Sidek, S. Shafie, and N. A. Yunus, "A Review on High-Resolution CMOS Delay Lines: Towards sub-picosecond Jitter Performance," *SpringerPlus*, vol. 5, no. 1, pp. 1–32, 2016.

[2] M. Maymandi-Nejad and M. Sachdev, "A Monotonic Digitally Controlled Delay Element," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 11, pp. 2212–2219, 2005.

[3] J. L. Gauci, E. Gatt, G. De Cataldo, O. Casha, and I. Grech, "An analytical model of the delay generator for the triggering of particle detectors at cern lhc," in *CAS (NGCAS), 2017 New Generation of*. IEEE, 2017, pp. 69–72.

[4] G. Jovanović and M. Stojčev, "Current starved delay element with symmetric load," *International journal of electronics*, vol. 93, no. 03, pp. 167–175, 2006.

[5] B. Markovic, S. Tisa, F. A. Villa, A. Tosi, and F. Zappa, "A high-linearity, 17 ps precision time-to-digital converter based on a single-stage vernier delay loop fine interpolation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 3, pp. 557–569, 2013.

[6] N. R. Mahapatra, S. V. Garimella, and A. Tareen, "An empirical and analytical comparison of delay elements and a new delay element design," in *VLSI, 2000. Proceedings. IEEE Computer Society Workshop on*. IEEE, 2000, pp. 81–86.

[7] H. Rivandi, S. Ebrahimi, and M. Saberi, "A low-power rail-to-rail input-range linear delay element circuit," *AEU-International Journal of Electronics and Communications*, vol. 79, pp. 26–32, 2017.

[8] E. Zafarkhah, M. Maymandi-Nejad, and M. Zare, "Improved accuracy equation for propagation delay of a cmos inverter in a single ended ring oscillator," *AEU-International Journal of Electronics and Communications*, vol. 71, pp. 110–117, 2017.