Tunable CMOS Delay Gate with Reduced Impact of Fabrication Mismatch on Timing Parameters

Przemyslaw Mroszczyk and Piotr Dudek
School of Electrical & Electronic Engineering
The University of Manchester
Manchester, M13 9PL, United Kingdom
przemyslaw.mroszczyk@postgrad.manchester.ac.uk
p.dudek@manchester.ac.uk

Abstract— **This paper presents the analysis and design of a simple one-stage tunable delay gate with improved matching properties as compared with the commonly used "current starved inverter". The operation of two delay lines employing these structures in a standard 90 nm CMOS technology was verified based on the post layout mismatch Monte Carlo simulations. Accounting for the fabrication mismatch, the delay generated by the proposed "output-split inverter" (OSI) circuit is about 10-30% more accurate as compared to the conventional current starved inverter occupying the same chip area.**

I. Introduction

An inherent shortcoming of any VLSI fabrication process is a certain degree of random variability of all the physical parameters of the manufactured devices, mainly MOS transistors used in any standard CMOS technology. Usually, the nature of these variations can be seen as a global parameter fluctuation, with a long distance correlation, and a local one, randomly affecting the parameters of equally designed and closely laid out devices [1]. Especially local variations, known as fabrication mismatch, are more difficult to mitigate and may significantly degrade the performance of a circuit. A well known method for mismatch optimization is transistor size scaling based on the assumption that the local variations average out as the MOS device area increases [2]. Proper layout drawing techniques (accounting for the proximity effects, the STI stress, the metal coverage, the distance-dependent parameter variability etc.) and some circuit specific design techniques such as offset compensation, bias point optimization, design redundancy and averaging, and post fabrication trimming could be applied to reduce mismatch and increase the fabrication yield [3]. Usually the efficient mismatch optimization of a larger analogue or mixed system is difficult and requires specific simulation tools. In many cases, the correct operation of such a system relies mainly on the correct operation of its critical sub-blocks mostly affecting its total performance, and these circuits should be optimized in the first instance [4].

The functional sub-block considered in this paper is a tunable delay gate used in a number of applications requiring the generation of controlled delay time intervals such as Delay Locked Loops [5], time-to-digital converters [6], silicon pixel readout circuits for particle detection [7]-[9] and asynchronous processor arrays [4], where the mismatch in time intervals generated by different delay gates or delay lines is reported as the dominant factor limiting the precise operation of the whole circuit. The applications found in literature employ a typical structure of a delay gate based on the current starved inverter (CSI) shown in Fig. 1a, where the mismatch reduction is usually achieved by the proper scaling (enlarging) of all the transistors or just the current controlling one (M_1) [5], [8].
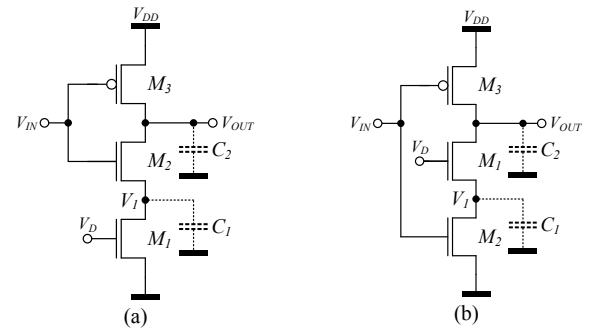

Figure 1. Schematic diagrams of the delay gates: a) the commonly used current starved inverter (CSI), b) the proposed "output-split inverter" (OSI).

In this paper an alternative structure of a delay gate, the "output-split inverter" (OSI) presented in Fig. 1b, is analysed. The only difference is the location of the current controlling transistor M_1 on the drain rather than source side of M_2, which separates or "splits" the output of the inverter. It is shown that this structure has less influence of MOS parameter mismatch on the generated delay variability than the CSI gate when both of them occupy the same area. The proposed OSI structure was also used in the build of a charge pump circuit [5] and a linearly tuned delay element [10]. Nevertheless, to the best of our knowledge, no previous work, except our application-oriented paper [4] postulate the superior matching performance of this circuit. Section II of this paper analyses the operation of both gates, Section III explains the influence of the MOS fabrication mismatch on the delay variability,

Section IV presents and discusses the simulation results and Section V concludes the paper.

II. CIRCUIT ANALYSIS

The operation of both delay gates presented in Fig. 1 is similar. Transistors M_2 and M_3 operate as switches determining the direction of the current flow (as in a static CMOS inverter) and additionally, for the rising edge of V_{IN}, M_2 turns on and M_1 limits the current discharging the output load capacitance depending on V_D. However, the simulation results of these gates reveal substantial differences in their operation (Fig. 2) which are of high importance in terms of the device matching properties (the details of the simulations are presented in Section IV). In the following we provide a simplified analysis presenting only the first-order behaviour addressing the major differences between circuits from Fig. 1.
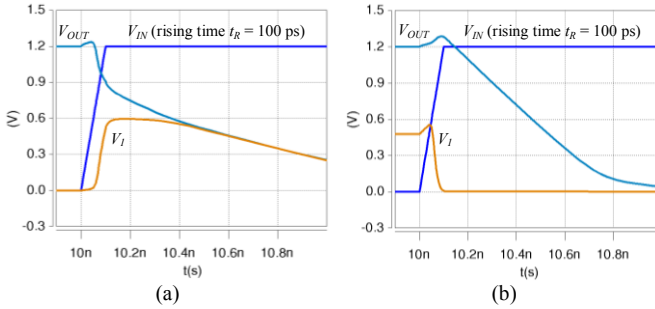

Figure 2. Transient analysis of the circuits from Fig. 1 for the rising edge of the input volage V_{IN} (W_1/L_1 = 1μm/1μm, $W_{2,3}/L_{2,3}$ = 1μm/90nm, C_2 = 1 fF, C_1 = 0 - only inherent source/drain MOS capacitances assumed, V_{DD} = 1.2 V, t_R = 100 ps): a) CSI gate for V_D = 450 mV, b) OSI gate for V_D = 500 mV.

A. Current Starved Inverter (CSI)

A simplified analysis of the CSI circuit showing the transitions of voltages V_{IN}, V_{OUT} and V_1 is presented in Fig. 3. There are three phases on the timeline: the initial phase ($t<t_1$), the switching phase ($t_1<t<t_2$) and the discharge phase ($t>t_2$).
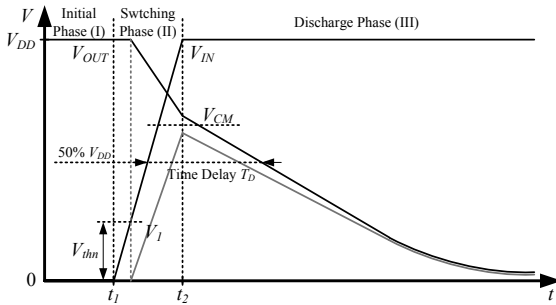

Figure 3. The behaviour of the CSI circuit in a transient state.

In the initial phase, the capacitance C_1 is discharged to zero through M_1 (for $V_D > 0$) and the capacitance C_2 is charged to V_{DD} through M_3 (V_{IN} = 0). In the switching phase, the rising edge of V_{IN} simultaneously turns M_2 on and turns M_3 off. Due to the charge sharing between C_1 and C_2, and the drain current of M_3, the capacitance C_2 gradually discharges and C_1 charges to a higher voltage. As a result V_{OUT} and V_1 converge closer to the common level $V_{CM} \approx V_{DD} \times C_2/(C_1+C_2)$ denoting the starting point for the discharge phase where M_2 (for $V_{IN} = V_{DD}$) keeps

voltages V_{OUT} and V_1 almost equal. The discharge rate of C_1 and C_2 depends mostly on the drain current of M_1 controlled by V_D. The proposed analysis does not consider the variable signal slopes and the charge injection effects of M_2 and M_3. These issues are covered in Section IV when simulating delay lines.

B. Output-Split Inverter (OSI)

The analysis of the OSI circuit showing the transitions of voltages V_{IN}, V_{OUT} and V_1 is presented in Fig. 4.
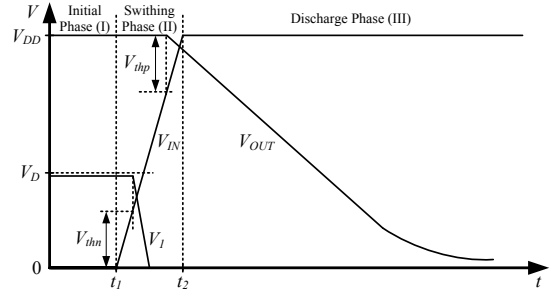

Figure 4. The behaviour of the OSI circuit in a transient state.

In the initial phase the capacitance C_2 remains charged to V_{DD} through M_3 and the capacitance C_1 remains charged closely to V_D through M_1 operating in the weak inversion and conducting the off leakage current of M_2 (V_{IN} = 0). In the switching phase the rising edge of V_{IN} turns M_2 on (when $V_{IN} > V_{thn}$), which quickly discharges C_1 to zero and, after that, turns M_3 off (when $V_{IN} > V_{DD} - |V_{thp}|$) and C_2 starts discharging with a rate dependent mainly on the drain current of M_1 controlled by V_D. The further operation of this gate is practically the same as in the case of the CSI one.

III. MISMATCH ANALYSIS

In literature the mismatch analysis and optimization of the timing parameters of the current starved inverter is usually done assuming that the transistors M_2 and M_3 (Fig. 1a) work as ideal switches and the precision of the generated time delay T_D depends mainly on the current limiting transistor M_1 [5]. More elaborate analysis of this circuit, accounting for the inter- and intra- die parameter fluctuations (also in switching transistors), was presented in [8]. The majority of works conclude that the precision of the time T_D could easily be improved by enlarging the current limiting transistor but the corresponding precision-area trade off should also be considered. In this section we focus on the influence of the MOS parameter variability on the signal transitions in the switching phase and show its impact on the starting point of the discharge phase in both circuits.

A. Mismatch in the CSI circuit

The analysis of the CSI circuit showing the influence of the MOS parameter fluctuations on the variability of the V_{CM} voltage and the generated delay time T_D is presented in Fig. 5. The common level $V_{CM} \approx V_{DD} \times C_2/(C_1+C_2)$ will also depend on slightly different parameters of all the MOS transistors resulting from the fabrication mismatch. For example, due to the random variation of the threshold voltage V_{thn} of M_1 this transistor may be slightly "stronger" (for lower values of V_{thn})

or slightly "weaker" (for higher values of V_{thn}) than a regular one. For the "stronger" M_1 the V_{CM} voltage will be lower and the corresponding discharge slope will be steeper, whereas for the "weaker" M_1 the V_{CM} voltage will reach a higher value and the discharge phase will take a longer time. As a result it can be observed that the variability of V_{CM} voltage (ΔV_{CM}) affects the variability ΔT_D of the generated time delay T_D.
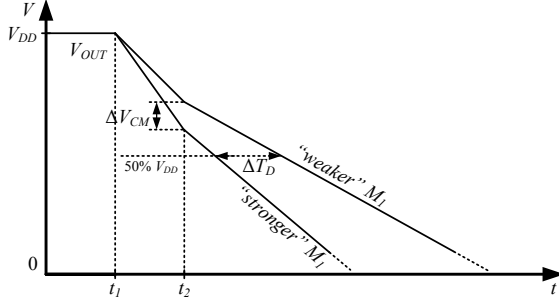

Figure 5. The transient state of the CSI circuit showing the influence of the MOS parameter variability on the V_{CM} voltage and the generated delay T_D.

B. Mismatch in the OSI circuit

The analysis of the OSI circuit showing the influence of the MOS parameter fluctuations on the variability of the generated time delay T_D is presented in Fig. 6. Due to the current limiting transistor M_1 "splitting" the output of the inverting stage, the rising edge of V_{IN} may not force an immediate transition of V_{OUT}. While the capacitance C_1 quickly discharges to zero (Fig. 4), M_3 still pulls the output node up to V_{DD} postponing the discharge phase roughly until $V_{IN} > V_{DD} - |V_{thp}|$. For the slightly "stronger" transistor M_3 with lower threshold voltage $|V_{thp}|$ the discharge phase may begin later than for the slightly "weaker" one. Additionally, the slightly "stronger" transistor M_1 may discharge C_2 faster also increasing the variability of the generated delay time. The influence of the first factor (ΔV_{thp}) is usually suppressed by a sharp slope of V_{IN}, whereas in case of the second one the discharge phase always begins for the fully charged C_2. Because of this the starting point of the discharge phase is more stable and the discharge time of C_2 is longer for the same current of M_1 (C_2 charged to $V_{DD} > V_{CM}$) as compared to the CSI structure, which makes the generated delay time of the OSI circuit less prone to fabrication mismatch.
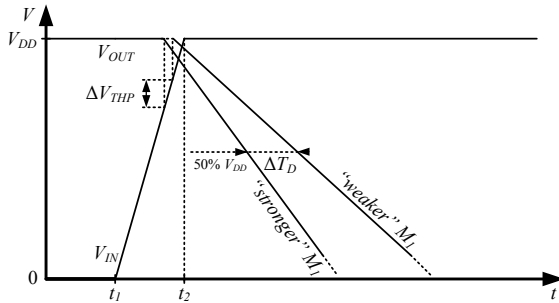

Figure 6. The transient state of the OSI circuit showing the influence of the MOS parameter variability on the generated delay time T_D.

IV. SIMULATION RESULTS AND DISCUSSION

The operation of both circuits (Fig. 1) was verified using HSPICE and mismatch Monte Carlo MOS transistor models from a standard 90 nm CMOS technology. We assumed fixed load $C_2 = 1$ fF corresponding to a typical load capacitance in this technology and considered only the size-dependent source /drain junction capacitances of MOS transistors ($C_1 = 0$). The rising time of the input signal $t_R = 100$ ps (not critical in terms of the qualitative analysis) and $V_{DD} = 1.2$ V.
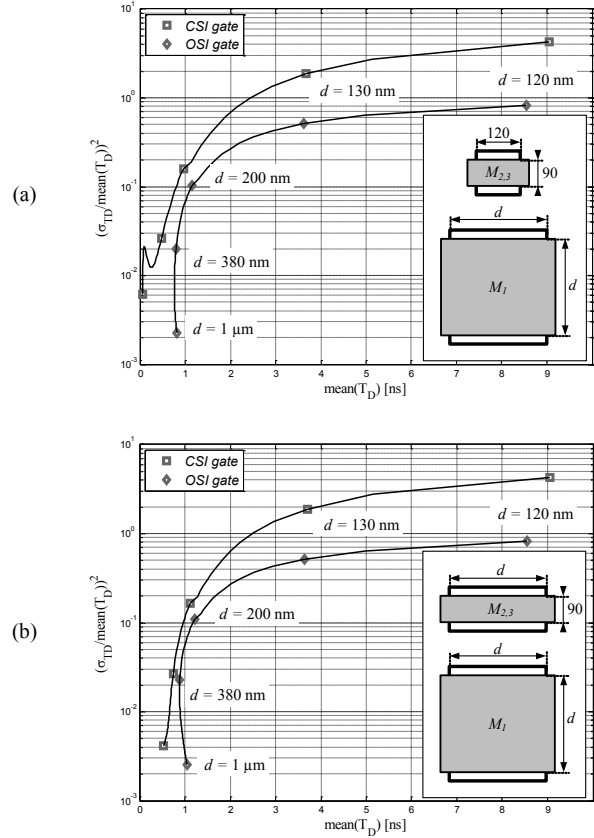

Figure 7. The normalised delay variance in terms of the mean T_D of the CSI and OSI circuits for two different scaling approaches where: a) only the current limiting transistor M_1 is scaled, b) all the transistors are scaled.

The statistical parameters of both structures using two different approaches to transistor scaling, where only the current limiting transistor M_1 is scaled ($W_1 = L_1 = d$, $W_{2,3} = 120$ nm, $L_{2,3} = 90$ nm) and when widths of the switching transistors $M_{2,3}$ are also scaled simultaneously with M_1 ($W_{1,2,3} = L_1 = d$, $L_{2,3} = 90$ nm), where d is swept between 120 nm and 1 μm, were obtained from 500 Monte Carlo runs for each case. The corresponding results showing the normalized delay variance [9] in terms of the mean value of the delay time for the CSI and the OSI gates (for V_D fixed to 420 mV) are presented in Fig 7. In both scaling schemes, the gate of M_1 is square in order to reduce the variability of its width and length and increase its active area and hence reduce the parameter mismatch [2]. It can be observed that the OSI structure exhibits less delay time variability than the CSI variant for the same area (e.g. for $W_1/L_1 = 1$μm/1μm in Fig. 7b the accuracy of the OSI circuit is about 20% higher than for the CSI one).

In order to compare the operation of both gates accounting for the variability in the switching levels, signal slopes and load capacitances, models of two delay lines consisting of 16 serially connected symmetric (with two complementary current limiting transistors) CSI and OSI gates were designed for $W_{1,4}/L_{1,4} = 1\mu m/0.5\mu m$ and $W_{2,3}/L_{2,3} = 1\mu m/90nm$ (Fig. 8).
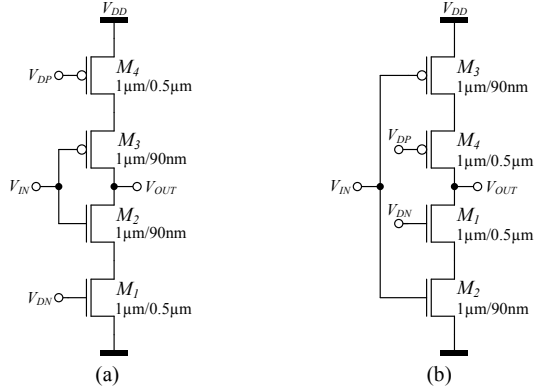

Figure 8. Schematic diagrams of the delay gates with two complementary current limitting transistors M_1 and M_4: a) CSI variant, b) OSI variant.

The layouts of both delay lines designed in a standard 90 nm CMOS technology are presented in Fig. 9. Fig. 10 shows the schematic diagram of the test circuit used for the simulations. The two additional inverters (with transistors of $W/L = 1\mu m/90nm$) on the input and the output of the delay line were added to assure more realistic models of the input signal source and the output load.
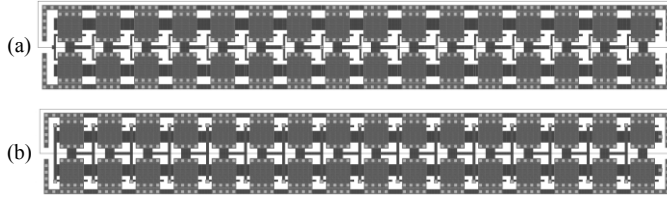

Figure 9. Layouts of the delay lines consisting of 16 serially connected a) CSI gates and b) 16 OSI gates (size of one line: 3.7 μm × 27 μm).
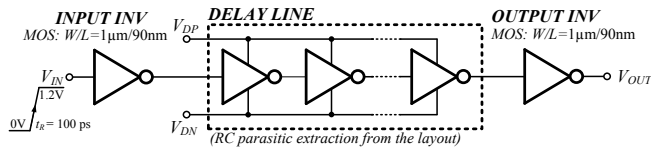

Figure 10. Schematic diagram of the test circuit used in the simulations.

The results of the Monte Carlo mismatch simulations of the test circuits with the post layout RC extracted models of the delay lines (also accounting for the mismatch in MOS transistors in the additional inverters) are presented in Fig. 11. The figure shows the normalized variance of the generated delay time for different bias voltages ($V_{DN} = 400...900$ mV and $V_{DP} = V_{DD} - V_{DN} = 800...300$ mV) in terms of the mean value of T_D. The simulation results (verified across all PVT corners) show that the use of the OSI gates in the design of delay lines improves the precision of the generated delay time T_D (compared to the CSI variant of the same size) not requiring any additional transistor scaling or area increase.
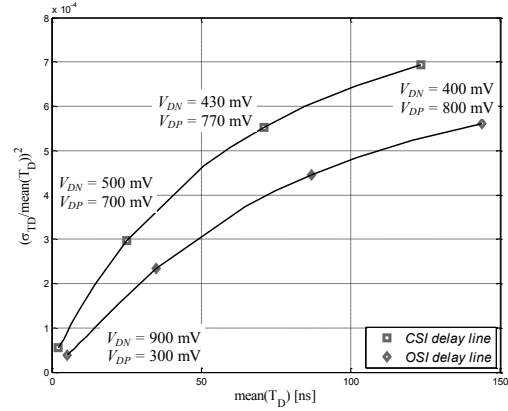

Figure 11. The normalized delay variance in terms of the mean T_D generated by the CSI and OSI delay lines for $V_{DN} = 400...900$ mV and $V_{DP} = V_{DD} - V_{DN}$.

V. CONCLUSIONS

This paper presents the design and the analysis of the tunable delay gate based on the "output-split inverter" (OSI). The proposed structure operates with a higher precision than the commonly used current starved inverter (CSI) circuit in the presence of fabrication mismatch. The operation and timing parameters of both structures were verified based on the mismatch Monte Carlo simulations of the post layout models of delay lines designed in a standard 90 nm CMOS technology. In typical applications the proposed OSI structure may show about 10-30% higher accuracy than its CSI counterpart when occupying the same chip area.

REFERENCES

[1] Pelgrom M., Duinmaijer A. C., Welbers A., "Matching Properties of MOS Transistors", IEEE Journal of Solid-State Circuits, vol. 24, no. 5, pp. 1433-1439, Oct. 1989.

[2] Lovett S., Welten M., Mathewson A., Mason B., "Optimising MOS Transistor Mismatch", IEEE Journal of Solid-State Circuits, vol. 33, no. 1, pp. 147-150, Jan. 1998.

[3] Kinget P., "Device Mismatch and Tradeoffs in the Design of Analog Circuits", IEEE Journal of Solid-State Circuits, vol. 40, no. 6, pp. 1212-1224, Jun. 2005.

[4] Mroszczyk P., Dudek P., "Trigger-Wave Collision Detecting Asynchronous Cellular Logic Array for Fast Image Skeletonization", ISCAS 2012, May 2012.

[5] Christiansen J., "An Integrated CMOS 0.15ns Digital Timing Generator for TDC's and Clock Distributed Systems", IEEE Trans. on Nuclear Science, vol. 42, no. 4, pp. 753-757, Aug. 1995.

[6] Dudek P., Szczepanski S., Hatfield J., "A High-Resolution CMOS Time-to-Digital Converter Utilizing a Vernier Delay Line", IEEE Trans. on Solid-State Circuits, vol. 35, no. 2, pp. 240-247, Feb. 2000.

[7] Heijne E., et al., "LHC1: A semiconductor pixel detector readout chip with internal, tunable delay providing a binary pattern of selected events", Nuclear Instruments and Methods in Physics Research, Section A 383, pp. 55-63, Elsevier, 1996.

[8] Cantatore E., et al., "Statistical analysis and optimisation of delay line chains for pixel readout electronics", Nuclear Instruments and Methods in Physics Research, Section A 395, pp. 318-323, Elsevier, 1997.

[9] Bolt M., et al., "Matching Properties of MOS Transistors and Delay Line Chains with Self-Aligned Source/Drain Contacts", Proc. of the IEEE Int. Conference on Microelectronic Test Structures, Mar. 1996.

[10] Jovanovic G., Stojcev M., "Linear Current Starved Delay Element", XL Int. Scientific Conference on Information Communication and Energy Systems and Technologies (ICEST), Jun. 2005.