Ultra Low Power Circuits

Lab #3 - Digital Report

November 7th 2022

Bologna Master Degree in Electrical and Computer Engineering (MEEC)

Group 12: Duarte Miguel de Aguiar Pinto e Morais Marques (96523) | Jun Zhou (96261)

1) Determining input capacitance $\mathrm{C_i}$

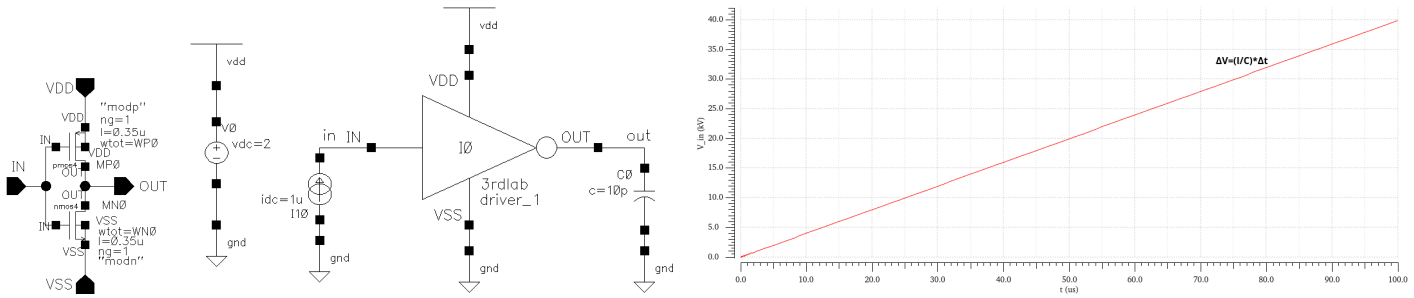

Figure 1: On the left, schematic of a single inverter, used in the test bench shown in the middle, in which widths $\mathrm{W_n} = 0.4\mu\mathrm{m}$ and $\mathrm{W_p} = 1.172\mu\mathrm{m}$ for the nMOS and pMOS transistors (respectively) were selected. By running a transient analysis for $\Delta\mathrm{t} = 100\mu\mathrm{s}$, the plot of the input voltage (measured in `in`), shown on the right, was obtained. Having measured $\Delta\mathrm{V} = 39.909\mathrm{kV}$, an input capacitance of $\mathrm{C_i} \approx 2.5\mathrm{fF}$ was obtained.

2) Single gate transistors

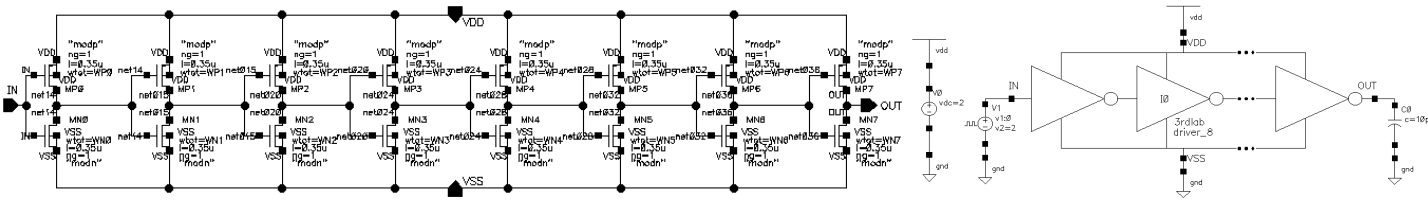

Figure 2: On the left, schematic of the digital output driver with $\mathrm{N} = 8$ stages. On the right, cellview in which the previously mentioned driver was included and from which the plots in Figure 3 were obtained. For a different number of stages N, similar schematics and respective test benches were designed in Virtuoso. In these cases, all transistors have a single gate, channel length $\mathrm{L} = 0.35\mu\mathrm{m}$ and varying channel widths.

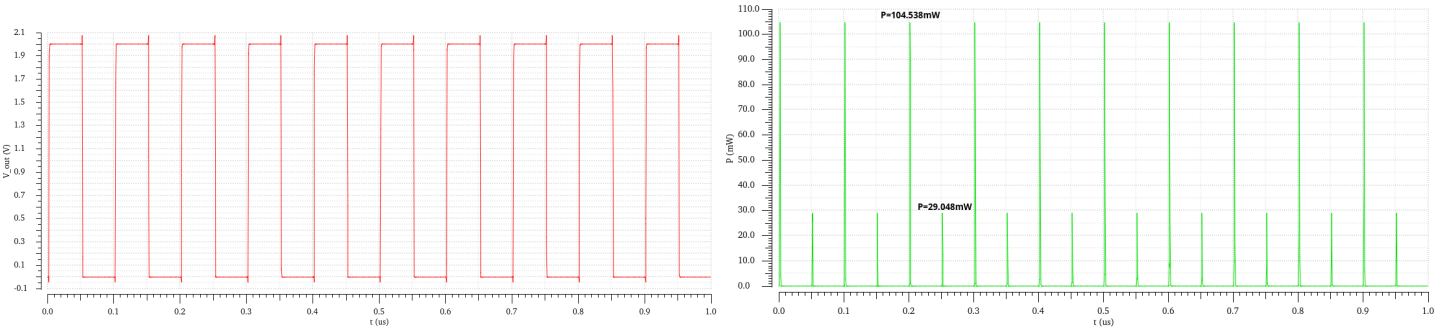

Figure 3: Plots obtained with the test bench shown in Figure 2, with a transient analysis ran from $\mathrm{t} = 0$ to $\mathrm{t} = 1\mu\mathrm{s}$ and for $\mathrm{W_n} \neq \mathrm{W_p}$. On the left, output voltage measured at the final stage of the digital output driver. On the right, power consumed by the driver, given by $\mathrm{P} = \mathrm{I} \times \mathrm{V_{DD}}$ (with $\mathrm{V_{DD}} = 2\mathrm{V}$), in which I is the current measured at the input pin `VDD`. Similar plots were obtained with $\mathrm{N} = 8$, $\mathrm{N} = 7$ and $\mathrm{N} = 6$ stages, both when $\mathrm{W_n} \neq \mathrm{W_p}$ and $\mathrm{W_n} = \mathrm{W_p}$ were considered.

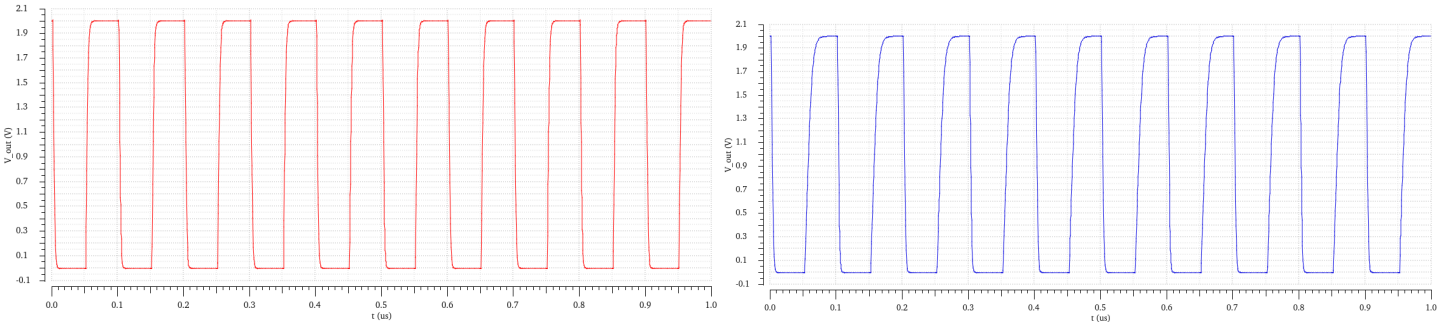

Figure 4: Plots obtained with a test bench similar to the one shown in Figure 2, but for $\mathrm{N} = 5$ stages, in which a distortion in the output voltage was more apparent. Once again, a transient analysis was run from $\mathrm{t} = 0$ to $\mathrm{t} = 1\mu\mathrm{s}$. On the left, plot of the output voltage obtained for $\mathrm{W_n} \neq \mathrm{W_p}$; on the right, plot obtained for $\mathrm{W_n} = \mathrm{W_p}$.

3) Variable number of gates

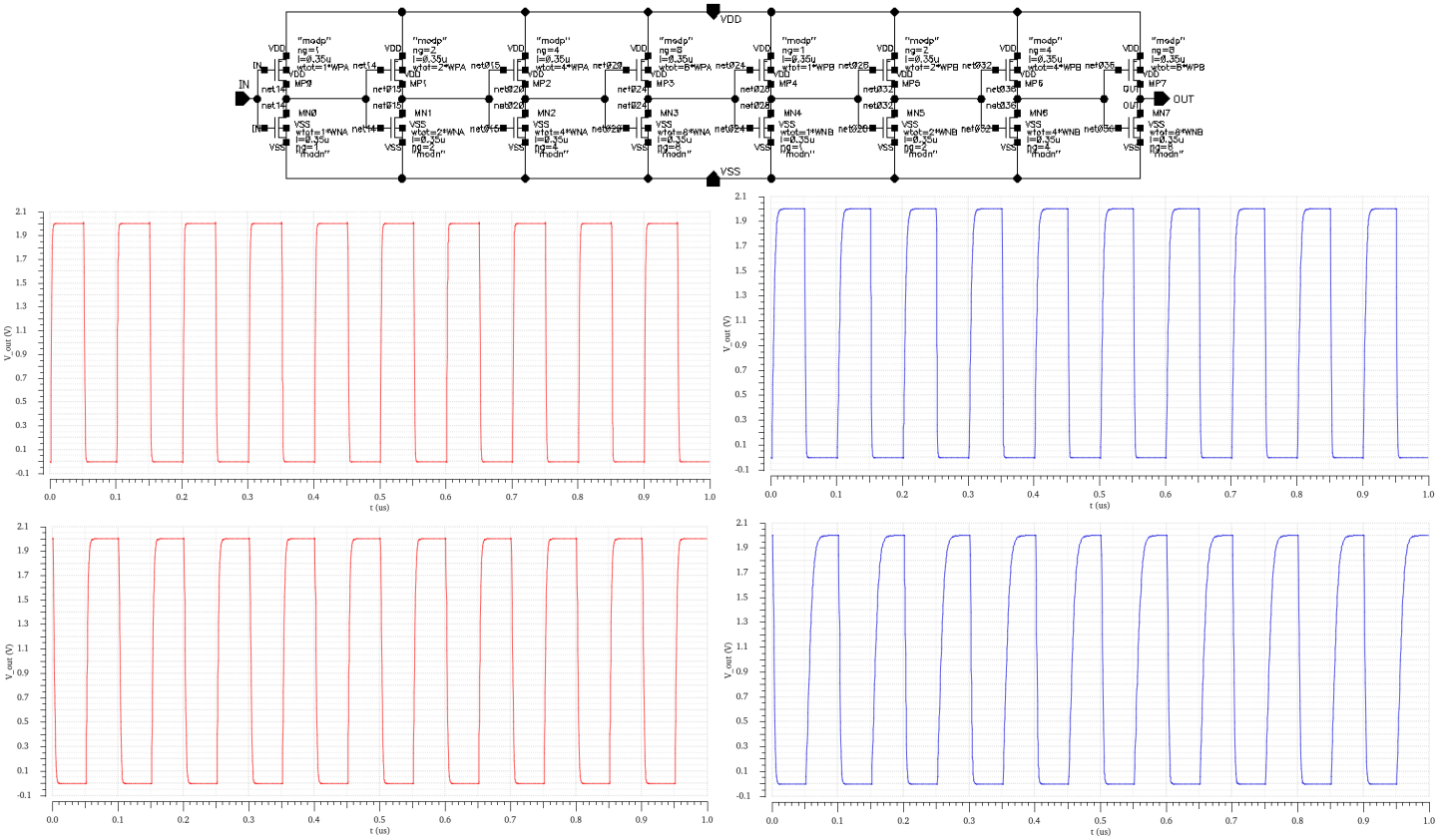

Figure 5: On the top, schematic of the digital output driver with $\mathrm{N} = 8$ stages and transistors with a variable numbers of gates. Due to the limit of 24 gates for each transistor (imposed by the software), the widths of the transistors were changed at the fifth stage. Similar schematics were used for different number of stages N. In the middle, plots of the output voltage for $\mathrm{N} = 8$ stages, considering $\mathrm{W_n} \neq \mathrm{W_p}$ (on the left) and $\mathrm{W_n} = \mathrm{W_p}$ (on the right). On the bottom, plots of the output voltage for $\mathrm{N} = 7$ stages, considering $\mathrm{W_n} \neq \mathrm{W_p}$ (on the left) and $\mathrm{W_n} = \mathrm{W_p}$ (on the right), in which distortion is already apparent.

4) Alternative solution

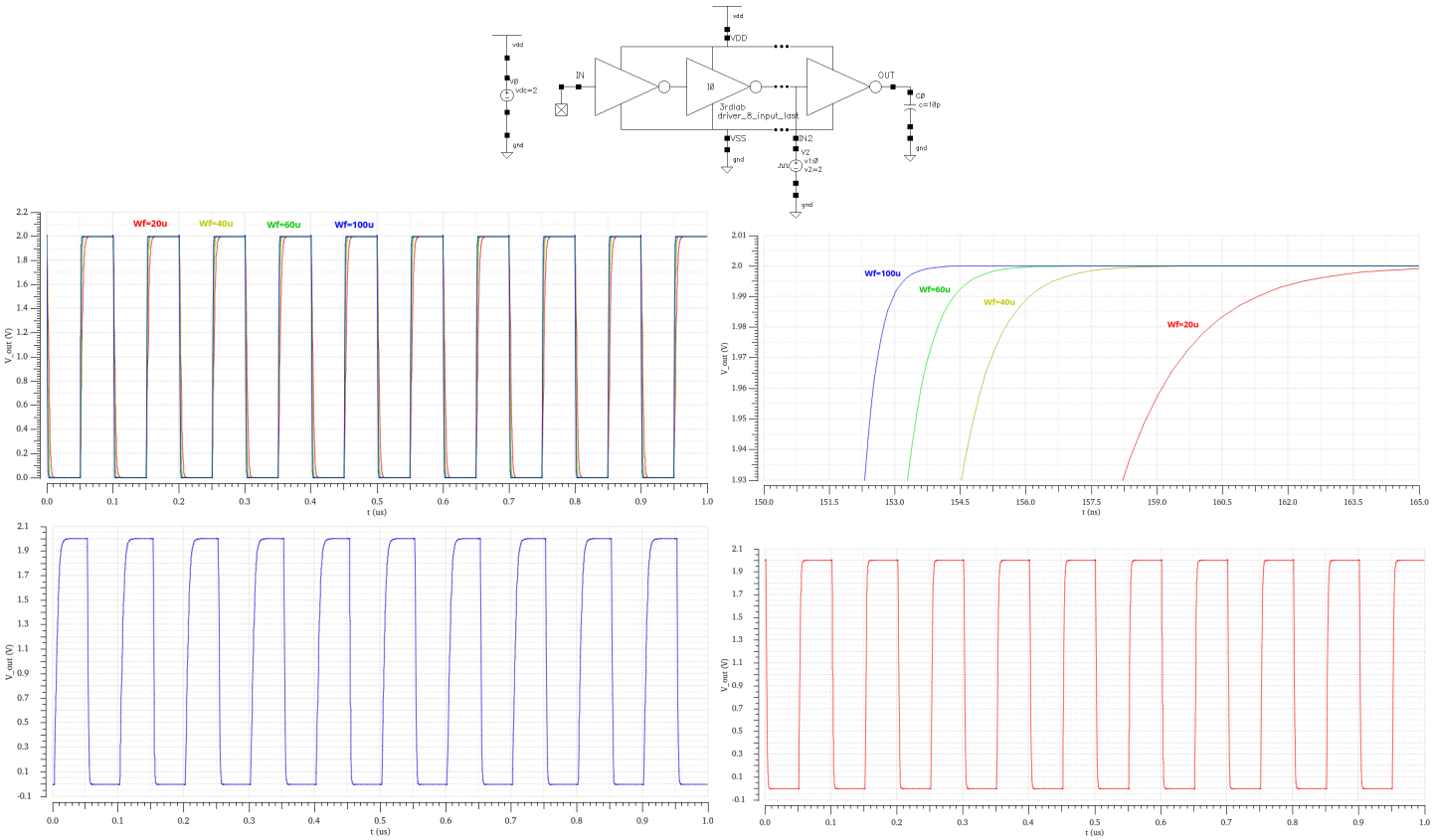

Figure 6: On the top, test bench used to obtain the minimum channel widths of the last stage. In the middle: on the left, plot of the output voltage obtained for different values of the width $\mathrm{W_f}$ for the nMOS transistor in the final output stage; on the right, zoomed in detail at the transition from Low to High. On the bottom: on the left, plot of the output voltage for $\mathrm{N} = 8$ stages and $\mathrm{W_n} = \mathrm{W_p}$ (with $\mathrm{W_f} = 40\mu\mathrm{m}$), in which a clear distortion is apparent; on the right, plot of the output voltage for $\mathrm{N} = 5$ stages and $\mathrm{W_n} \neq \mathrm{W_p}$ (with $\mathrm{W_f} = 40\mu\mathrm{m}$), in which distortion is not significant.

1) Determining input capacitance $\mathrm{C_i}$

For this digital output driver, an estimate of the optimal number of stages is given by the equation shown below, where $\mathrm{C_i}$ and $\mathrm{C_L}$ are the input and output capacitances, respectively.

$$\mathrm{N} = \ln\left(\frac{\mathrm{C_L}}{\mathrm{C_i}}\right) \text{ (nearest integer)}$$

In this case, the output capacitance is necessarily defined as $\mathrm{C_L} = 10\mathrm{pF}$. In order to determine the value of C_i, the test bench shown in Figure 1 was used. The value of the input capacitance will depend on the first inverter (i.e., the first stage). For this purpose, the minimum channel width of $\mathrm{W_n} = 0.4\mu\mathrm{m}$ was used for the nMOS transistor. On the other hand, a width of $\mathrm{W_p} = 1.172\mu\mathrm{m}$ was used for the pMOS transistor, in order to compensate for the lower mobility of holes, when compared to the mobility of electrons ($\mu_\mathrm{p} < \mu_\mathrm{n}$). According to the datasheet, $\mu_\mathrm{n}\mathrm{C_{ox}} = 170\mu\mathrm{A/V^2}$ and $\mu_\mathrm{p}\mathrm{C_{ox}} = 58\mu\mathrm{A/V^2}$, thus $\mathrm{W_p/W_n} = 170/58 \approx 2.931$. Additionally, a length of $\mathrm{L} = 0.35\mu\mathrm{m}$ was used for both transistors.

Using an input DC current of $1\mu\mathrm{A}$, the voltage in the input node was simulated in a transient analysis during $\Delta\mathrm{t} = 100\mu\mathrm{s}$, having an initial condition of $0\mathrm{V}$ been forced at the input. A linear response was obtained, as seen in Figure 1, having the voltage changed $\Delta\mathrm{V} = 39.909\mathrm{kV}$. With equation $\mathrm{I} = \mathrm{C_i}\frac{\Delta\mathrm{V}}{\Delta\mathrm{t}}$, the value $\mathbf{C_i \approx 2.5pF}$ was obtained. This led to a number of stages $\mathbf{N = 8}$.

Finally, an estimate $\mathbf{K = 2.819}$ for the ratio between the channel widths of successive stages was obtained from the following equation:

$$\mathrm{K} = \left(\frac{\mathrm{C_L}}{\mathrm{C_i}}\right)^{\frac{1}{\mathrm{N}}}$$

2) Single gate transistors

Considering the previously obtained estimates $\mathrm{N} = 8$ and $\mathrm{K} = 2.819$, the schematic shown in Figure 2 was designed in Cadence Virtuoso. Using it, the test bench shown next to it was used to simulate the digital output driver. For this purpose, an input square wave of frequency $\mathrm{f} = 10\mathrm{MHz}$ varying between 0V and 2V and an output capacitor of $\mathrm{C_L} = 10\mathrm{pF}$ were included, having a supply voltage of $\mathrm{V_{DD}} = 2\mathrm{V}$ been used. In a transient analysis performed until $\mathrm{t} = 1\mu\mathrm{s}$, the plot of the voltage in the output node was obtained. Moreover, by measuring the current at `VDD` and multiplying it by $\mathrm{V_{DD}} = 2\mathrm{V}$, the plot of the power consumption was obtained. Both these graphs are shown in Figure 3. It can be seen that the input square wave was successfully transposed to the output, with only small spikes in voltage being apparent in some transitions between different states. As for the power consumption, the largest values occur in the transitions from Low to High (i.e., from 0V to 2V) - these values will be designated as '$\mathbf{P_M}$'. Other smaller spikes in power consumption are seen in the transitions from High to Low - these smaller spikes in power will be designated as '$\mathbf{P_m}$' throughout this report.

In order to attempt the minimization of power consumption and area, similar schematics to the one shown in Figure 2 were designed, but with different number of stages (N). Moreover, for each N, the same circuits were simulated in two different ways. Firstly, the widths $\mathrm{W_P}$ of the pMOS transistors were obtained by $\mathrm{W_p} = 2.931 \times \mathrm{W_n}$ for each inverter, as mentioned before. Secondly, the difference in mobilities $\mu_p < \mu_n$ was ignored, having the same widths been used for nMOS and pMOS transistors (i.e., $\mathrm{W_p} = \mathrm{W_n}$); by doing this, even though an increase in the delay time is to be expected, the area and power consumption can be decreased. In Figure 4, the results obtained for $\mathrm{N} = 5$ stages are shown; since a considerable distortion in the signal becomes more apparent in this case, the number of stages was not further decreased.

For each case, different parameters were calculated and are shown in Table 1. These include $\mathrm{P_M}$ and $\mathrm{P_m}$ (mentioned previously), as well as a value for the total **area** of the design, which can be estimated by summing all $\mathrm{W} \times \mathrm{L}$ products. Moreover, the rise and fall times of the output voltage ($\mathbf{t_{TLH}}$ and $\mathbf{t_{THL}}$), as well as the delay times from High to Low ($\mathbf{t_{PHL}}$) and from Low to High ($\mathbf{t_{PLH}}$), are shown. In turn, the delay time $\mathbf{t_p}$ is given by the following equation:

$$\mathrm{t_p} = \frac{1}{2}(\mathrm{t_{PHL}} + \mathrm{t_{PLH}})$$

3) Variable number of gates

With the purpose of minimizing area, the same results were obtained with a different solution. Instead of using transistors with a single gate and different widths (as mentioned for section 2), the same width $\mathrm{W_n} = 0.4\mu\mathrm{m}$ was used, as well as $\mathrm{W_p} = 1.172\mu\mathrm{m}$ (if $\mathrm{W_n} \neq \mathrm{Wp}$ is being considered) or $\mathrm{W_p} = 0.4\mu\mathrm{m}$ (if the difference in mobilities is not taken into account). By maintaining the width and changing the number of gates, a doubling in the number of gates is approximately equivalent to doubling the length L of the transistor, a much lower value than the widths considered before, thus leading to lower total areas.

In this case, a significant distortion was obtained for $\mathrm{N} = 7$, thus no other number of stages was tested, as implied by Figure 5. Additionally, it is worth mentioning that the software Virtuoso only allows a number of gates ≤ 24, thus the widths $\mathrm{W_n}$ and $\mathrm{W_p}$ were changed accordingly in the fifth stage of each schematic (as shown in Figure 5), reverting back to 1 gate per transistor and increasing it once again up until the last stage.

4) Alternative solution

Finally, the test bench shown in Figure 6 was designed. In this case, the square wave input signal was placed at the input of the final stage (having an instance `noConn` of the library `simple` been placed at the input of the initial stage). By running a transient analysis simulation in $\Delta\mathrm{t} = 1\mu\mathrm{s}$, the voltage in the output node was plotted for different values of the width $\mathrm{W_f}$ for the nMOS transistor in the final stage, in order to obtain an estimate of the minimum value for each there is still no significant distortion in the signal - some of these plots are shown in Figure 6. From these simulations, it was clear that, for $\mathrm{W_f} < 40\mu\mathrm{m}$, a more significant distortion is apparent.

The same parameters were obtained for these simulations, as shown in Table 1. However, in this case, a distortion was immediately apparent for $\mathrm{N} = 8$ stages when widths $\mathrm{W_p} = \mathrm{W_n}$ were used, as shown in Figure 6, thus the different parameters were not determined for situations in which $\mathrm{W_p} = \mathrm{W_n}$. On the other hand, almost no imperfections were apparent even for $\mathrm{N} = 5$ when $\mathrm{W_p} \neq \mathrm{W_n}$. For each schematic (i.e., for each number of stages N), the ratio K between channel widths of the nMOS transistors was determined so that $0.4\mu\mathrm{m} \times \mathrm{K^{N-1}} = 40\mu\mathrm{m}$.

	Number of stages	$\mathrm{P_M}$ (mW)	$\mathrm{P_m}$ (mW)	Area ($\times 10^{-12}\mathrm{m}^2$)	$\mathrm{t_{THL}}$ (ns)	$\mathrm{t_{TLH}}$ (ns)	$\mathrm{t_{PHL}}$ (ns)	$\mathrm{t_{PLH}}$ (ns)	$\mathrm{t_P}$ (ns)
Single gate transistors	N=8 ($\mathrm{W_n} \neq \mathrm{W_p}$)	104.5	29.0	1206.3	0.4	0.5	27.6	27.5	27.6
	N=8 ($\mathrm{W_n} = \mathrm{W_p}$)	59.5	17.4	613.7	0.4	0.9	28.5	27.9	28.2
	N=7 ($\mathrm{W_n} \neq \mathrm{W_p}$)	58.2	12.9	427.7	0.6	0.9	27.6	27.5	27.5
	N=7 ($\mathrm{W_n} = \mathrm{W_p}$)	26.5	6.7	217.6	0.6	1.8	28.9	27.7	28.3
	N=6 ($\mathrm{W_n} \neq \mathrm{W_p}$)	26.4	5.2	151.5	1.4	1.8	27.8	27.8	27.8
	N=6 ($\mathrm{W_n} = \mathrm{W_p}$)	10.4	2.4	77.1	1.4	4.6	28.7	29.3	29.0
	N=5 ($\mathrm{W_n} \neq \mathrm{W_p}$)	10.5	1.9	53.6	3.7	4.4	29.0	28.7	28.9
	N=5 ($\mathrm{W_n} = \mathrm{W_p}$)	3.8	0.8	27.2	3.7	12.3	33.5	29.0	31.3
Variable number of gates	N=8 ($\mathrm{W_n} \neq \mathrm{W_p}$)	21.9	5.6	140.3	1.8	2.1	27.6	27.6	27.6
	N=8 ($\mathrm{W_n} = \mathrm{W_p}$)	7.9	2.9	71.4	1.8	6.1	27.8	29.4	28.6
	N=7 ($\mathrm{W_n} \neq \mathrm{W_p}$)	11.1	2.8	69.9	3.6	4.3	28.5	28.4	28.4
	N=7 ($\mathrm{W_n} = \mathrm{W_p}$)	4.0	1.4	35.6	3.6	12.1	28.4	32.4	30.4
Alternative solution	N=8 ($\mathrm{W_n} \neq \mathrm{W_p}$)	16.1	4.7	113.5	2.4	2.9	28.6	28.5	28.5
	N=7 ($\mathrm{W_n} \neq \mathrm{W_p}$)	16.1	4.1	102.2	2.4	2.9	28.5	28.3	28.4
	N=6 ($\mathrm{W_n} \neq \mathrm{W_p}$)	16.0	3.5	91.0	2.4	2.9	28.3	28.3	28.3
	N=5 ($\mathrm{W_n} \neq \mathrm{W_p}$)	15.9	2.6	80.2	2.4	2.9	28.3	28.1	28.2

Table 1: Parameters relative to the different configurations used for the digital output driver design and simulations. The values of $\mathrm{P_M}$ and $\mathrm{P_m}$ correspond to the power at the higher and lower peaks in the plots for the power consumption (as shown in Figure 3), respectively. When a single gate is used in all transistors, the area can be estimated by summing the products $\mathrm{W} \times \mathrm{L}$ of all transistors. On the other hand, when a variable number of gates is considered, the respective length is multiplied by the respective number of gates (while the width remains constant). The values $\mathrm{t_{THL}}$ and $\mathrm{t_{TLH}}$ correspond to fall and rise times (respectively) in the output signal, whereas $\mathrm{t_{PHL}}$ and $\mathrm{t_{PLH}}$ are the delay times from High to Low and from Low to High, respectively. The delay time in the final column is given by $\mathrm{t_P} = \frac{1}{2}(\mathrm{t_{PHL}} + \mathrm{t_{PLH}})$.

Discussion of results

a) Same widths for nMOS and pMOS transistors in each stage ($\mathrm{W_n} = \mathrm{W_p}$)

From the data collected in Table 1, it can be concluded that the use of equal widths for the nMOS and pMOS transistors in each stage leads to a significant decrease in power consumption. With single gate transistors, the decrease in $\mathrm{P_M}$ varied between 54.5% (for N=7) and 75.6% (for N=8), whereas the decrease in $\mathrm{P_m}$ varied between 40% (for N=8) and 57.9% (for N=5). On the other hand, with a variable number of gates, the decreases in $\mathrm{P_M}$ and $\mathrm{P_m}$ were, for N=8, 63.9% and 48.2%, whereas N=7 led to 64.0% and 50% decreases (respectively). It can be noticed that the decreases in $\mathrm{P_M}$ relatively close to 50%, but higher (from 54.5% to 75.6%). By considering $\mathrm{W_n} = \mathrm{W_p}$ (instead of $\mathrm{W_n} \neq \mathrm{W_p}$), the widths of the pMOS transistors are effectively decreased by a factor of 2.931, which leads to an increase in the resistance by the same factor. This higher resistance corresponds to lower currents (due to Ohm's Law), which means lower power consumption. On the other hand, it is worth noting that the decreases in $\mathrm{P_m}$ are always less significant; $\mathrm{P_m}$ corresponds to transitions from High to Low - thus, the pMOS transistors turn OFF and the nMOS transistors turn ON. Since $\mathrm{W_n}$ does not change (it is $\mathrm{W_p}$ that changes), the decrease in current values for these transitions are not as significant.

Another positive aspect which comes from changing $\mathrm{W_p}$ is the decrease in area. This decrease is very close to 50% in all situations. For N=8 and single gate transistors, its decrease is 47.6%, while it is 49.3% for N=5; in all other configurations, its value is 49.1%. These decreases ($\lesssim 50\%$) are once again due to the factor 2.931 used to obtain $\mathrm{W_p}$ in the initial configurations.

On the other hand, the decrease in power consumption and area leads to an **increase in** rise, fall and **delay times**. When the width of the pMOS transistors is changed, the only parameter which does not suffer any change is $\mathrm{t_{THL}}$, the fall time in the output signal. This is because, in High to Low transitions, the pMOS turn OFF and the nMOS turn ON, which makes it so that the change in $\mathrm{W_p}$ is not as significant (the widths of the nMOS transistors, $\mathrm{W_n}$, are not altered). Contrarily, the increases in $\mathrm{t_{TLH}}$ are much more significant since, in Low to High transitions, the nMOS turn OFF and the pMOS turn ON. In this case, the increases varied between 80% (N=8, single gate) and 190.5% (N=8, variable number of gates). Once again, this is due to the fact that lower $\mathrm{W_p}$ lead to lower currents, which, in turn, increase the rise times in a similar factor. As for the delay times, an increase was always observed, even though not as significant (percentage wise) - these varied between 2.2% and 8.3%.

For all configurations used in this laboratory assignment, the highest delay time occurred with N=5 and single gate transistors ($\mathrm{t_P} = 31.3\mathrm{nS}$), which confirms the negative effect of area and power minimization on the increase in delay times.

b) Changing the number of stages (N)

In this analysis, only the situations in which $\mathrm{W_n} \neq \mathrm{W_p}$ will be taken into account. As expected, as the number of stages is decreased, the power consumption also decreases, since there are less transistors to power up (i.e., lower currents are at play). In the case of single gate transistors and with a variable number of gates, the decrease in $\mathrm{P_M}$ varied between 60.2% (N=6→N=5, single gate) and 44.3% (N=8→N=7, single gate). As for $\mathrm{P_m}$, the changes were also quite significant, varying between 50.2% and 64.6%. The values obtained for single gate transistors are all close to 50% and this is due to the fact that $\mathrm{K} = 2.819$, as determined in section 1. Because of this, as one stage is removed, the total area decreases a little above 50% ($\approx 65\%$ for single gate transistors). With a variable number of gates, each stage had two times more gates than the previous, which led to an area decrease of 50.2% (much closer to 50%) and similar decreases of 49.3% and 50.0% for $\mathrm{P_M}$ and $\mathrm{P_m}$, respectively.

However, it is worth noting that $\mathrm{P_M}$ **did not** change significantly in the alternative solution results. This should be because, in this case, the last stage transistors had a fixed length (determined in section 4) and the widths $\mathrm{W_n}$ and $\mathrm{W_p}$ in the other stages were determined with a varying K ratio in order to have $0.4\mu\mathrm{m}$ in the initial nMOS transistor. This led to a decrease in area much less significant than in the other cases (between 10.0% and 11.9%), thus lower changes in all parameters shown in Table 1.

Once again, the minimization of area and power consumption leads to an increase in the rise, fall and delay times. However, as explained above, this does not happen in the alternative solution (in fact, in some situations, these values decrease by a small amount). It is worth mentioning, however, that the delay time actually decreased by 0.1ns in a single gate transistor configuration when the number of stages decreased from N=8 and N=7, which perhaps indicates a significant advantage of this configuration over the others (in which the minimization of power and area have the disadvantage of increasing all time related parameters).

c) Using transistors with a variable number of gates

Because of different aspects, the results shown in Table 1 could lead to the conclusion that using a variable number of gates in the transistors might be the optimal solution for the digital output driver design.

Firstly, it is worth noting that, even though the number of stages only varied between N=8 and N=7, the total area estimated for the output driver is, apart from N=8 with $\mathrm{W_n = W_p}$, lower than all values for the alternative solution and lower than all but two configurations with single gate transistors. This happens because, in this case, it is not the width of the transistors that varies. Increasing the number of gates by a factor of n can be considered equivalent to multiplying the channel length by n - being $\mathrm{L} = 0.35\mu\mathrm{m}$ a much lower value than most channel widths used in other configurations. Additionally, each successive stage had double of the number of gates from the previous, which is a lower number than $\mathrm{K} = 2.819$ considered previously. Moreover, even for the alternative solution, mostly larger ratios were considered. Since in those configurations K was determined in order to go from $\mathrm{W_n} = 0.4\mu\mathrm{m}$ (at the first stage) to $\mathrm{W_n} = 40\mu\mathrm{m}$ (in the last stage), the correspondent ratios are $\mathrm{K} = \sqrt[7,6,5,4]{100} \approx 1.9, 2.2, 2.5, 3.2$ (for N=8, 7, 6 and 5, respectively).

Associated with the (generally) lower areas, the configurations with a variable number of gates also lead to incredibly low power consumption. In fact, the values of $\mathrm{P_M}$ for $\mathrm{W_n = W_p}$ were lower than all others, apart from $\mathrm{N} = 5(\mathrm{W_n = W_p})$ (single gate). The power $\mathrm{P_m}$ showed similarly low values when compared to the other configurations.

The two most significant disadvantages of using a variable number of gates are the higher rise, fall and delay times, as well as the fact that only a minimum N=7 can be used - below that, significant distortion was apparent in the output signal, as transmitted by the plots in Figure 5. Considering that the delay times $\mathrm{t_P}$ were maintained within the range 27.6ns-30.4ns, with very low increases in $\mathrm{t_{THL}}$ and $\mathrm{t_{TLH}}$ (apart from $\mathrm{t_{TLH}} = 12.1\mathrm{ns}$ obtained for $\mathrm{N} = 7$) with $\mathrm{W_n = W_p}$), it may be reasonable to indicate that the significant minimization in power consumption and area would suggest that configurations with a variable number of gates (in this case, multiplied by a factor of 2 from one stage to another) should provide a better option for the digital output driver design.