Optimum Number of Transistors in Stacked CMOS Millimeter-Wave Power Amplifiers

Mohammad Hassan Montaseri, Janne Aikio, Timo Rahkonen, Aarno Pärssinen
Centre for Wireless Communications and Circuits and Systems Research Group
Faculty of Information Technology and Electrical Engineering (ITEE)
University of Oulu
Oulu, Finland
{Firstname.Lastname@oulu.fi}

Abstract—This paper proposes how to define the optimum number of stacked transistors in a multi-stacked CMOS power amplifier (PA) topology, based on several physical as well as circuit design aspects. Starting with a systematic concept, the analysis then goes through the relevance of transistor transconductance, aspect ratio, parasitics, operating frequency, and number of transistor stages in a pentagonal trade-off concept. While this is done based on theoretical circuit analysis, the results, then, were evaluated using simulations based on 45nm CMOS technology.

Keywords—CMOS; millimeter-wave integrated circuits; millimeter-wave power amplifiers; multi-stacked topology;

I. INTRODUCTION

With the growth of millimeter-wave spectrum applications the need for fully integrated efficient millimeter-wave power amplifiers (PA) is inevitable. Higher integration capability and lower fabrication cost offered by CMOS technology, has made PA integration a highly desirable feature and challenging target for research. The main drawback of scaled CMOS transistors, however, is their low junction breakdown levels, making them inappropriate for power amplification applications if used in a single transistor topology. Stacking CMOS transistors in series connection is an approach to prevent the foregone problem [1]–[4] in a way that with N transistors, the overall structure can tolerate N times the drain-source breakdown voltage of a single device without violating any transistor junction. This is possible in IC processes, like SOI, where bulk can be isolated properly from the substrate and connected directly to source. Nevertheless, the presence of transistor parasitics dictate phase rotation along each and every stage in stacked transistor topologies; if not tolerable, additional matching networks are necessary so as to both match each stage to its successor and to phase align each single amplifier stage in order for the signals to be efficiently summed up to the output node [5]–[7]. Accordingly, the number of stacked transistors remains limited regardless of either approach, the definition of which is the aim of this paper.

II. STACKED CMOS POWER AMPLIFIERS

As shown in Fig. 1, a stacked-transistor PA is composed of series connection of several transistors on top of each other, where $V_n$ is the drain-source amplitude of each transistor. As mentioned earlier, if superimposed efficiently, this results in output voltage swing increase across the overall device to an order of $N \times V_n$ whilst none of the junctions experiences any stress. On the other hand, the design requirements mandate each stacked transistor to provide a resistive load of $R_{opt}$, in a way that the series connection of all stacked transistors is equal to the load to be driven, i.e. $R_L$. This is done through dimensioning capacitances $C_2$ to $C_N$ using either approaches proposed in [4], [5], [7], or [8].

Due to the presence of parasitics, the angles of drain-source voltage vectors tend to gradually rotate per stacked transistor along the device. For example, in case of uniform phase rotation, summing all drain-source voltages yields the amplitude of the output of interest as

$$|V_{out}| = \left| \sum_{n=1}^{N} V_n e^{i\theta} \right| = V_n \times \frac{\sin(N\theta/2)}{\sin(\theta/2)} \tag{1}$$

From (1) the dependence of maximum output amplitude on the phase rotation posed by each stage can be demonstrated. Indeed, inter-stage phase misalignment results in radio frequency (RF) signal amplitude to destructively depart from its ideal amplitude (see Fig. 2) and hence output power and efficiency reduction.

It can be seen from Fig. 2 that the maximum swing degrades from ideal case with phase rotation per stage. In addition, due to excess phase rotation, after a certain number of transistor stages, the maximum amplitude starts to saturate reaching an absolute maximum and decreases afterwards. It
should be noted that this saturation is due to the fact that excess phase rotation gradually reduces amplitude of the previous stages.

A closer look at (1) in polar plot reveals that the total phase rotation should be bound in practice to about 120° instead of 180°. This is because after this amount of phase shift the increase in signal amplitude is rather very low. For example, consider a phase rotation of 10° per stacked transistor; the first 12 stages reach a level of 10×V_m while the next 6 stages will only add approximately 2×V_m more which leads to an approximately 64% energy loss compared to optimum case. By the same token, for θ = 15° and θ = 20° the optimum numbers of stages are 8 and 6, respectively; while the remaining stages, i.e. from 120° up to 180°, will only add approximately 1–1.3 times V_m which yields drastic efficiency degradation and must be avoided. Thus, 120° can be regarded as a sweet spot of the stacked transistor PAs design.

Since phase rotation stems from the presence of parasitics, it can be defined from the input impedance of each stage using the MOS transistor model at high frequencies as shown in Fig. 4.

Analytical model gives the input impedance of the n^th stage to be

\[
Z_{in,n} = \frac{C_{gd} + C_{gd}(1 + g_m(n + 1)R_{opt}) + C_n + n(n + 1)R_{opt}C_{n+1}C_{gd}}{(g_m + sC_{gd})(C_{gd} + C_n + s(n + 1)R_{opt}C_{n+1}C_{gd})}, \tag{2}
\]

where \(R_{opt}\) is the optimum load of each single transistor and \(C_n\) is the designed gate capacitance of common gate stages. Apparently (2) reveals phase variation as a direct consequence of the presence of frequency dependent terms in the input impedance. Thus the amount of phase variation of each stage can be calculated as

\[
\theta_n = \tan^{-1}(nR_{opt}/C_{gd}) + \tan^{-1}(C_{gd}/(g_m)), \tag{3}
\]

and the total phase variation can be defined to be

\[
\theta_{tot} = \sum_{n=1}^{N} \left( \tan^{-1}(nR_{opt}/C_{gd}) + \tan^{-1}(C_{gd}/(g_m)) \right). \tag{4}
\]

Substituting the method explained in [5] for dimensioning \(C_n\) and using “Euler-Maclaurin” summation method, on the condition that \(nR_{opt}/C_{gd} \ll C_n\) and \(C_{gd}/g_m\) are less than 1, (4) can be approximated as

\[
\theta_{tot} \approx \frac{(N + 1)nR_{opt}C_{gd} + C_{gd}(1 + g_mR_L))}{2(C_{gd} + g_mR_LC_{gd})} + \frac{NnC_{gd}}{g_m}. \tag{5}
\]

Regarding 120° as the total phase rotation threshold for the design, (4) can be solved for the optimum number of stacked transistors as follows

\[
\theta_{tot} \approx \frac{(N + 1)nR_{opt}C_{gd} + C_{gd}(1 + g_mR_L))}{2(C_{gd} + g_mR_LC_{gd})} + \frac{NnC_{gd}}{g_m} = \frac{2\pi}{3}. \tag{6}
\]

Equation (6) predicts the optimum number of stacked transistors for which the input impedance of \(n+1^\text{th}\) stage, i.e. the load impedance of \(n^\text{th}\) stage, starts to destroy the drain node voltage hence overall device output power and efficiency. It should be noted that (5) reveals several
interesting design trade-offs as well. Accordingly, it can be
deduced that any increase in phase yields reduction in the
number of transistor stages. More importantly, phase varies in
direct proportion to the operating frequency, transistor
parasitics, and the load line, while indirectly to transconductance of the transistors. The transconductance,
however, is defined both by transistor biasing and aspect ratio.
This gives a trade-off between the mentioned design
parameters shown in Fig. 5.

\[
\begin{align*}
\text{Power/Efficiency} & \quad \text{Device Dimensions (Cgs, Cgd, W/L)} \\
\text{Operating Frequency (ω/θ)} & \quad \text{Biasing (gm)} \\
\text{Number of stacks (Nmax)} & \quad \text{Number of stacks (Nmax)}
\end{align*}
\]

Fig. 5. Multi-stacked transistors PAs design trade-offs.

III. SIMULATION RESULTS

In order to evaluate the theoretical analysis presented in
previous section, a multi-stacked-transistor PA (Fig. 6) was
designed to supply 100mW into a load of \( R_L = 50\Omega \) with a
25dB voltage gain at 28GHz operating frequency, based on
45nm CMOS technology. \( R_1 - R_{N+1} \) are designed to be much
higher compared to the \( C_n \) reactances at the desired
frequencies.

Based on the requirements, transistor biasing and
dimensions were designed such that their transconductance \( g_m \),
gate-source capacitance \( C_{gs} \), and gate-drain capacitance \( C_{gd} \)
were 350mS, 200fF, and 45fF, respectively. Substituting the
given parameters in (6) and solving for the number of stacked
transistors \( N \) gives 10.5. Since this is the number of transistors
for which a 120° phase rotation takes place, it is recommended
to round 10.5 down to 10 transistors for the structure, i.e. an
approximately average of 12° of phase rotation per stage. Fig.
7 depicts the simulated phase rotation of each single stage as
well as cumulative phase rotation of each added stage
compared to the proposed theoretical formulae (3) and (5),
respectively.

From (6) it can be seen that while the analysis predicts the
optimal number of stacked transistors to be ten, simulation
results reveal that it is roughly eight. This indicates that the
average phase rotation per stage is somewhat more than what
was calculated, i.e. almost 16° per stage. Thus, the proposed
analytical calculations comply fairly well with the simulation
results. However, their difference stems first from the fact
that some approximation has been made in calculating (5), and
secondly the parasitics (capacitances, resistances, etc.) are
nonlinear in their nature; since they tend to vary with signal
level, AM - PM conversion distortion is inevitable.

Fig. 8 shows a plot of maximum efficiency, \( \eta_{\text{max}} \), as a
function of the number of the stacked transistors. As
mentioned earlier, due to phase misalignment of the stages, the
output voltage amplitude deviates destructively from its ideal
value. Accordingly, the drain efficiency (DE) starts to
degrad. 

\[
\begin{align*}
\text{Operating Frequency} & \quad \text{Biasing (gm)} \\
\text{Number of stacks (Nmax)} & \quad \text{Number of stacks (Nmax)}
\end{align*}
\]

Fig. 6. The designed multi-stacked CMOS schematic circuit.

Fig. 7. Single-stage/cumulative phase rotation posed by each stacked
transistor.

The rate at which the designed multi-stacked transistor PA
loses efficiency per added stage is also plotted in Fig. 8.
Firstly, it can be seen that the loss curvature is an exponent of
the phase rotation per stacked transistor. Secondly, the
steepness of the curvature increases with phase misalignment.
In case of excess phase rotation, due to additional stages, the
structure may produce RF signals which

transistor will be totally useless if used. This, in addition to previous explanations, supports the analytical calculations, which were proposed in previous section. In this example, eight transistor stacks can be considered if maximum power delivering capability is sought. However, that already reduces drain efficiency more than 10% compared to optimum class-A performance. In case of maximum efficiency, only stack of 3...4 transistors should be used.

**Fig. 8.** Drain efficiency (DE) of the overall PA device per number of stacked transistors and efficiency degradation rate of each stacked transistor.

**IV. CONCLUSION**

In this paper, a theoretical method was developed to define the maximum/optimum number of stacked transistors in multi-stacked transistor PAs using several theoretical formulae. The study started from a systematic viewpoint up to transistor level physical parameters. It was shown how the operating frequency, and transistor aspect ratio, through gate-source and gate-drain capacitances, and transconductance through biasing, define the maximum as well as the optimum number of stacked transistors. It was justified that any increase in either or both operating frequency and parasitic capacitances will translate into phase rotation, which in turn reduces the maximum allowed number of transistor stages. Conversely, increasing transconductance of transistors results in phase reduction while allowing some room for either more device stacking or higher operating frequency. However, it should be noted that this is in a compromise with the transistor dimensioning.

Based on the presented analysis, a CMOS PA was designed using 45nm CMOS technology. An optimum number of 10 stacked transistors was calculated analytically at 28GHz operating frequency. According to simulations 20dBm saturated output power with drain efficiency of 32% could be reached. This accounted for 10% discrepancy between simulation and analysis. However, the simulation results were in good agreement with the analytical calculations which proves the validity of the proposed method.

**ACKNOWLEDGMENT**

The authors would like to express their gratitude towards InfoTech Oulu Doctoral Program.

**REFERENCES**


