<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>RFCMOS unit width optimization technique (Published version)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Tong, Ah Fatt; Lim, Wei Meng; Sia, Choon Beng; Yeo, Kiat Seng; Teng, Zee Long; Ng, Pei Fern</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Tong, A. F., Lim, W. M., Sia, C. B., Yeo, K. S., Teng, Z. L., &amp; Ng, P. F. (2007). RFCMOS unit width optimization technique. IEEE Transactions on Microwave Theory and Techniques, 55(9), 1844-1853.</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2007</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/4657">http://hdl.handle.net/10220/4657</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2007 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE. This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder. <a href="http://www.ieee.org/portal/site">http://www.ieee.org/portal/site</a>.</td>
</tr>
</tbody>
</table>
RFCMOS Unit Width Optimization Technique

Ah Fatt Tong, Wei Meng Lim, Choon Beng Sia, Kiat Seng Yeo, Zee Long Teng, and Pei Fern Ng

Abstract—In this paper, we demonstrate a unit width ($W_f$) optimization technique based on their unity short-circuit current gain frequency ($f_T$), unilateral power gain frequency ($f_{MAX}$), and high-frequency (HF) noise for RFCMOS transistors. Our results show that the trend for the above figures-of-merit (FOMs) with respect to the $W_f$ change is different; hence, some tradeoff is required to obtain the optimum $W_f$ value. During the HF noise analysis, a new FOM is proposed to study the $W_f$ effect on the HF noise performance. In our experiment, the flicker noise of the transistor is also measured and the result shows that the change in $W_f$ does not affect the noise spectral density at the low-frequency range. This technique enables RF engineers to optimize the transistor’s layout and helps to select the optimum $W_f$ for transistors used in specific circuit design such as the low-noise amplifier, voltage-controlled oscillator, and mixer. Furthermore, by using layout optimized transistors in the RF circuit, the optimal circuit’s performance can be easily achieved and, thus, greatly reduced the circuit development time. In the aspect of RF device modeling, by knowing the optimum $W_f$ for a particular process or technology, the number of transistors to model is reduced and, hence, greatly shortens the RF modeling development time for existing and future technologies.

Index Terms—Flicker noise, high-frequency (HF) noise, layout, optimization, RF, RFCMOS, unilateral power gain frequency, unity short-circuit current gain frequency.

I. INTRODUCTION

As the CMOS processing technology continues to advance, its RF properties such as unity short-circuit current gain frequency ($f_T$), unilateral power gain frequency ($f_{MAX}$), and the noise figure relentlessly improves [1]–[7]. It has been reported that the $f_T$ of a 0.13-$\mu$m gate-length MOS transistor can reach the 80-GHz region [1]. The down scaling of channel length has allowed higher integration density and the possibility of integrating digital, analog, and RF circuits into a single chip [8], [9], making CMOS technology a cost-effective solution for fabricating RFCIC.

One of the most commonly used components in RF circuit design is the RFCMOS transistor. In order to optimize the transistor’s RF characteristics, its layout design methodology has been extensively studied and researched [10]–[14]. Although different foundries, integrated device manufacturers (IDMs), and other fab-less semiconductor companies use various transistor layout designs, the typical RF transistors are designed with a multifingered and double-contacted gate configuration. The layout design of the transistor is normally symmetry with a special guard ring added to isolate any interference from other components in the circuit.

In multifingered transistor design, the size of the transistor is mainly controlled by the finger number ($N_f$), unit width ($W_f$), and channel length ($L_g$) of the transistor. For most RF circuit design, transistors with the smallest gate length are always used due to their fast response and high drain current. Therefore, the RF circuit designer will need to select transistors based on either $N_f$ or $W_f$. By optimizing per finger unit width with respect to $f_T$, $f_{MAX}$, minimum noise figure ($NF_{min}$) and flicker noise, the best $W_f$ of the transistor can be selected to use in a specific circuit application such as a low-noise amplifier (LNA), voltage-controlled oscillator (VCO), and mixer. The ability to select and use the optimized layout for the transistor can ensure that the designed RF circuit produces the best performance at the first design cycle, and this greatly reduces the RF circuit development time. This methodology, when applied to a certain process or technology, can help the IDM and foundry select the optimal $W_f$ for transistors meant for model development, hence, saving a large amount of SPICE model development time.

The following figures-of-merit (FOMs) $f_T$, $f_{MAX}$, $NF_{min}$, and flicker noise spectral density have been commonly used in [7], [10], [15], and [16] to characterize the performance for the RFCMOS transistor, but these FOMs are normally presented with respect to the change in technology or channel length. In [11], only the $f_T$ optimization with $W_f$ is shown, while in [2] and [14], $f_T$, $f_{MAX}$ and $NF_{min}$ optimization are presented using only large total width transistors. Due to the transistor’s parasitic changes with respect to $W_f$, their effects toward each individual FOM are different. Hence, it is important to study these four FOMs with respect to $W_f$ on different transistor sizes so as to obtain the optimized width per finger based on either one or all four FOMs. In this paper, a study on the four FOMs with respect to the transistor’s $W_f$ is done for small (48 $\mu$m), medium (120 $\mu$m), and large (240 $\mu$m) total width transistors and a new FOM is proposed for the study of the transistor’s HF noise. All of the transistor’s $f_T$, $f_{MAX}$, $NF_{min}$ and flicker noise spectral density are measured at the maximum transconductance ($g_m$) operating point of the transistor. The test structures consist of transistors with a different total width of 48, 120, and 240 $\mu$m with four different $W_f$ variations.
of 4, 8, 12, and 24 μm. The measured results show that the optimized $W_f$ for $f_T$, $f_{\text{MAX}}$ and $N_{\text{fin}}$ do not coincide at the same point and, hence, some tradeoff is required when selecting the transistor’s $W_f$ for different circuit applications. In Section II, the $W_f$ effect on $f_T$ and $f_{\text{MAX}}$ of the transistor is presented. Section III shows the measured high-frequency (HF) noise data versus the $W_f$ effect and a new FOM is proposed for the analysis of the transistor’s noise. Section IV shows the $W_f$ variation on the flicker noise characteristics for the RF transistor. Finally, the selection criteria for transistors to be used in RF circuits are discussed in Section V.

II. UNIT WIDTH OPTIMIZATION ON $f_T$ AND $f_{\text{MAX}}$

A. $f_T$ Definition and Extraction

$f_T$ is defined as the unity current gain frequency at which the short-circuit current gain of the transistor becomes unity, which is shown in (1) as follows:

$$\omega_T = 2 \cdot \pi \cdot f_T = \frac{g_m}{C_g} \quad (1)$$

$$C_g = C_{gs} + C_{gb} + C_{gd} \quad (2)$$

$$H_{21} = \frac{I_{\text{out}}}{I_{\text{in}}} \quad (3)$$

The short-circuit current gain (3) is used for the extraction of $f_T$. It can be easily obtained by performing a two-port conversion into $H$-parameters from the measured deembedded $S$-parameters of the transistor. In this paper, the extraction of $f_T$ is based on the extrapolation of the $H_{21}$ curves at 10.25 GHz where it has a slope of $-20$ dB/decade. The extrapolated line will cross the frequency axis at 0 dB for $H_{21}$ and the $X$-intercept is the estimated $f_T$ for the transistor. Note that $C_g$ is the overall capacitances looking into the gate terminal, which includes the gate-to-source ($C_{gs}$), gate-to-body ($C_{gb}$), and gate-to-drain ($C_{gd}$) overlap capacitances.

Fig. 1 shows an example of the measured $H_{21}$ (in decibels) versus frequency plot at $V_{gs} = 1.05$ V and $V_{ds} = 1.8$ V. By extrapolating at 10.25 GHz, the $X$-interception is the extracted $f_T$ of the transistor.

B. $f_{\text{MAX}}$ Definition and Extraction

$f_{\text{MAX}}$ is defined as the frequency at which the ratio of the load power to input power becomes unity, and from [4], it is derived to be shown in (4) as follows:

$$f_{\text{MAX}} \approx \sqrt{\frac{f_T}{8\pi R_g C_{gd}}} \quad (4)$$

$$\text{GU} = \frac{0.5 \cdot \left| \frac{S_{21}}{S_{12}} \right|^2}{k \cdot \left| \frac{S_{21}}{S_{12}} \right| - \text{real} \left( \frac{S_{21}}{S_{12}} \right)} \quad (5)$$

$$k = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |S_{11} + S_{22} - S_{12} \cdot S_{21}|^2}{2 \cdot |S_{12}||S_{21}|} \quad (6)$$

The extraction of $f_{\text{MAX}}$ is done using the unilateral power gain (GU), as shown in (5). This power gain can be obtained when the input of the transistor is conjugate matched to the input signal source, the load is also conjugate matched with the transistor output impedance, and an appropriate network is used to cancel the effect of feedback from the output to the input [17]. As frequency increases, GU will decrease and when it reaches unity, the frequency is the maximum transistor operating frequency. At low frequency, the measured GU is normally very unstable and their corresponding slope is not $-20$ dB/decade. Hence, the $S$-parameters measurement is performed up to maximum equipment capability of 50 GHz and uses the data from the higher frequency region to extrapolate and obtain the $f_{\text{MAX}}$ value.

Fig. 2 shows an example of the measured GU (in decibels) versus frequency plot at $V_{gs} = 1.05$ V and $V_{ds} = 1.8$ V. By extrapolating at 40.85 GHz, the $X$-interception is the extracted $f_{\text{MAX}}$ of the transistor.

C. Experimental Results and Discussion

After reviewing the $f_T$ and $f_{\text{MAX}}$ definition and extraction, $S$-parameters measurement is performed on the designed set
TABLE I
UNIT WIDTH OPTIMIZATION TEST STRUCTURES FABRICATED IN 0.18-μm CMOS TECHNOLOGY WITH CHANNEL LENGTH OF 0.18 μm

<table>
<thead>
<tr>
<th>Total Width (μm)</th>
<th>Unit Width (μm)</th>
<th>Finger Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>48</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>48</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>48</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>48</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td>120</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>120</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>120</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>120</td>
<td>24</td>
<td>5</td>
</tr>
<tr>
<td>240</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>240</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td>240</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>240</td>
<td>24</td>
<td>10</td>
</tr>
</tbody>
</table>

Fig. 3. Designed test structure layout for a 15-finger thin gate nMOS transistor.

of thin gate nMOS transistors fabricated in a 0.18-μm CMOS technology with a channel length of 0.18 μm and varying $W_f$ of 4, 8, 12, and 24 μm and total width of 48, 120, and 240 μm, as shown in Table I. Deembedding of their corresponding OPEN and SHORT structures are performed to obtain the true transistor RF performance [18].

An example of the test structure layout from Table I is shown in Fig. 3. The transistor is designed with a multifingered and double-contacted gate configuration with a guard ring included for isolation of the interferences from external components when used in the circuit.

Fig. 4 shows the extracted $f_T$ and $f_{MAX}$ values for a total width of 48-, 120-, and 240-μm test structures, as shown in Table I. The extraction is performed at a maximum $g_m$ biasing condition, whereby $V_{ds} = 1.05$ V and $V_{bg} = 1.8$ V. From the plots, the extracted $f_T$ shows a similar trend for all of the three total width transistors. Fig. 5(a) and (b) shows the extracted maximum $g_m$ and $C_g$ values and the $R_d$ and $C_{gd}$ values versus unit width for the three total widths transistors.

It is well known that $g_m$ and $C_g$ from (1) are directly proportional to the transistor’s width and, hence, their ratio will result in $f_T$ to be independent of the transistor’s width, but from Fig. 4, it shows that the extracted $f_T$ changes with the transistor’s $W_f$. For smaller $W_f$ of 4- and 8-μm transistors, their extracted $f_T$ is smaller than the transistors with $W_f$ of 12 and 24 μm. Such behavior of $f_T$ versus $W_f$ is due to the parasitic capacitances.
that exist in the different $W_f$ layout transistor. For the same total width, smaller $W_f$ transistors will have more fingers and, hence, the amount of overlap extrinsic parasitic capacitances will be larger when compared to large $W_f$ transistors that have a smaller finger number. In Fig. 5(a), the overall gate capacitance $C_g$, which includes all the parasitic overlap capacitances, is increasing with decreasing $W_f$ for all three total width transistors. This clearly indicates that transistor with small $W_f$ will have more fingers and cause higher overlap parasitic capacitances. As for the extracted $g_{m}$ in Fig. 5(a), it increases slightly with decreasing $W_f$ at a constant total width. Due to the self-heating effect, the increasing trend of $g_{m}$ with decreasing $W_f$ is more obvious in a large total width transistor. This is because a large total width transistor with a small $W_f$ has a higher number of source and drain diffusions as compared to the same total width transistor with larger $W_f$. Hence, with more source and drain diffusions, the drain current is distributed more evenly. This means that less current will flow through per diffusion. Thus, the heat generated at each diffusion region and its diffusion resistance change due to temperature will be smaller. Therefore, a transistor with a smaller $W_f$ will have a higher $g_{m}$, due to less self-heating. Furthermore, it can be observed that the changes in the extracted $g_{m}$ for a fixed total width is small so only the parasitic overlap capacitances in $C_g$ affect the $f_T$ versus $W_f$ plot in Fig. 4. The extrinsic overlap parasitic capacitances for $C_{gs}$, $C_{gd}$, and $C_{gds}$ are shown in Fig. 3. Therefore, to optimize $W_f$ based on $f_T$, the width per finger cannot be chosen to be too small. It is observed from Fig. 4 that $f_T$ optimization for $W_f$ happens at approximately 12 $\mu$m for all three sets of the total width transistors. Hence, increasing $W_f$ above 12 $\mu$m will not significantly improve the $f_T$ of the transistor.

From (4), it can be deduced that $f_{\text{MAX}}$ can be lower or higher than $f_T$ based on the layout of the transistor. In Fig. 4, for the three total widths, the extracted $f_{\text{MAX}}$ is observed to be increasing with decreasing $W_f$. This behavior of $f_{\text{MAX}}$ is mainly caused by the change in $R_g$, $f_T$, and $C_{gds}$ with respect to the change in $W_f$. For the same total width with $W_f$ decreasing, the finger number of the transistor is increasing and since $R_g$ is proportional to ($W_f/N_f$), $R_g$ will decrease and cause $f_{\text{MAX}}$ to increase. At the same time, when the finger number increases, the overlap parasitic capacitances (gate-to-drain, gate-to-source, and gate-to-body) will increase, and from (1) and (2), the transistor’s $f_T$ will drop and $C_{gds}$ (intrinsic and extrinsic capacitances) will increase, and from (4), these two parameters will cause $f_{\text{MAX}}$ to decrease. Hence, based on the above analysis, depending on the transistors’ total width and $W_f$, certain parasitic effects ($R_g$ and extrinsic overlap capacitances) will be more dominant than the other; hence, influencing the final $f_{\text{MAX}}$ trend when $W_f$ and total width change. Fig. 5(b) shows the trend of $R_g$ and $C_{gds}$ versus unit width for the three total widths. In Fig. 4(a) and (b), the trend of $f_{\text{MAX}}$ versus $W_f$ is dominated mainly by the effect from $R_g$, while for Fig. 4(c), the extrinsic overlap parasitic capacitances dominant over $R_g$ and cause the trend of $f_{\text{MAX}}$ versus $W_f$ to change. In addition to that, the extracted $f_{\text{MAX}}$ value is decreasing with increasing total width size at the same $W_f$; this can be explained by the increase in the overlap parasitic capacitances in $C_g$ and $C_{gds}$, as shown in Fig. 5(a) and (b) for a larger total width size transistor that has a larger finger number at a fixed $W_f$. Therefore, based on Fig. 4, it is observed that to optimize the transistor layout using $f_{\text{MAX}}$, the transistor’s $W_f$ should be chosen to be small.

This study of the $W_f$ effect on $f_T$ and $f_{\text{MAX}}$ is important and it shows that the wrong selection of $W_f$ for the transistor layout will cause either $f_T$ or $f_{\text{MAX}}$ to suffer. Since the trend of $f_T$ and $f_{\text{MAX}}$ versus $W_f$ is different, some tradeoff is needed when selecting the best $W_f$ value for the transistor to be used in a specific application. Although both $f_T$ and $f_{\text{MAX}}$ are commonly used to compare the performance of the transistor, in the circuit design such as a VCO and transistor gain stages, the power amplification capability is of more importance and, hence, the $f_{\text{MAX}}$ optimization for the transistor $W_f$ will be more crucial in these applications.

III. UNIT WIDTH OPTIMIZATION ON HF NOISE

The HF noise measurements are performed on-wafer using the ATN NP5 Microwave Noise Parameter System. All three sets of total width transistors are measured at the maximum $g_{m}$ condition and the frequency range is from 2 to 26.5 GHz.
All measured noise parameters are deembedded with their corresponding OPEN structures using the deembedding feature in the NP5 system. This embedding procedure can be found in [19], whereby the pad and interconnect parasitic capacitances are deembedded.

A. HF Noise Definition and Theory

From the classical RF noise model from Van Der Ziel [20, 21], the HF noise of a transistor can be described by two correlated current noise sources, the drain noise current ($i_{d1}$) and the induced gate noise current ($i_{g1}$). The induced gate noise is caused by the capacitive coupling from the channel noise current. The mean-square representation for drain current noise and gate noise are shown in (7) and (8), respectively. Note that $c$ is the correlation factor between the drain and gate noise sources

$$\overline{\gamma^2_{d1}} = 4kT\gamma g_{d0}\Delta f$$  \hspace{1cm} (7)$$  
$$\overline{\gamma^2_{g1}} = 4kT\delta \left( \frac{\omega^2 C_{gs}}{g_{d0}} \right) \Delta f$$  \hspace{1cm} (8)$$  
$$c \equiv \frac{\overline{i_{g1}^* i_{d1}}}{\overline{\gamma^2_{g1}} \overline{\gamma^2_{d1}}}.$$  \hspace{1cm} (9)$$

In (7), the parameter $g_{d0}$ is the drain–source conductance at zero $V_{DS}$, and $\gamma$ is the noise factor and has a value of unity at zero $V_{DS}$ and, in long channel devices, it decreases to a value of 2/3 in the saturation region. Note that for short channel devices operating in the saturation region, $\gamma$ can be considerably higher than the long channel value. In (8), the parameter $\delta$ is the gate noise coefficient and it is given a value of 4/3 in [20].

In [22] and [23], the approximated expression for minimum noise factor ($F_{\text{min}}$) is derived as shown in (10). In Section III-B, we will relate this expression to our measured $NF_{\text{min}}$. It is noted that $NF_{\text{min}}$ is equivalent to $F_{\text{min}}$ in decibels as follows:

$$F_{\text{min}} \approx 1 + \frac{f}{f_{\text{MAX}}} \sqrt{\frac{P + R - 2C}{P}}$$

$$P = \frac{2\pi}{4kTg_{m}\Delta f} \overline{\gamma^2_{d1}}$$  \hspace{1cm} (11)$$  
$$R = \frac{\omega^2 C_{gs}}{g_{m}} \overline{\gamma^2_{g1}}$$  \hspace{1cm} (12)$$  
$$C = \text{Im} \left( \frac{i_{g1}^* i_{d1}^*}{\overline{\gamma^2_{g1}} \overline{\gamma^2_{d1}}} \right)$$  \hspace{1cm} (13)$$  
$$f_{c} = \frac{g_{m}}{2\pi C_{gs}}.$$  \hspace{1cm} (14)$$

The $PRC$ noise parameters used in (10) are related to the drain and gate noise current sources, as shown in (11)–(13).

B. Experimental Results and Discussion

Fig. 6 shows the transistor’s $NF_{\text{min}}$ versus frequency plot for different $W_f$ with a total width of 48, 120, and 240 $\mu$m. All three plots show that transistors with $W_f$ of 24 $\mu$m exhibit the highest $NF_{\text{min}}$, while the $NF_{\text{min}}$ characteristics for transistors with $W_f$ of 4, 8, and 12 $\mu$m overlap each other and cannot be differentiated clearly. These plots show an important observation that is by optimizing the $W_f$ of 24-$\mu$m transistors to a smaller $W_f$ of...
4, 8, or 12 μm at the same total width, the NF_{min} response can be greatly improved. Furthermore, for \( W_f \) less than 12 μm, the improvement in NF_{min} is not significant.

Note that the extracted NF_{min} plots for the three total widths transistor in Fig. 6 show very close results. The geometry dependence of \( f_c \cdot g_m \) and \( R_o \) can be shown to be proportional to \( (1/L_f^2) \cdot (NF_f \cdot W_f/L_f) \) and \( (W_f/N_f \cdot L_f) \), respectively, and by substituting them into (10) will result in \( F_{min} \) to be proportional to \( (W_f/L_f) \). From the above analysis, \( F_{min} \) is found to be independent of \( N_f \) and this explains why the measured NF_{min} results for the three total widths are so close. Furthermore, from the geometry dependence of \( F_{min} \), it is expected that transistors with the largest \( W_f \) will have the highest noise level, and this is clearly shown in measurement results when \( W_f \) of 24 μm exhibits the highest NF_{min} values.

From Fig. 6, it is also observed that for frequency range less than 5 GHz, all the extracted NF_{min} values are found to be overlapping and fluctuating without any clear trend with respect to \( W_f \) change. Hence, the extracted NF_{min} for less than 5 GHz cannot be used to study the \( W_f \) effect on the transistor noise. Furthermore, the NF_{min} for \( W_f \) of 4, 8, and 12 μm for the whole frequency range is also overlapping and, hence, the extracted NF_{min} results cannot be used to study the effect of \( W_f \). Due to the above observations, it is difficult to study the \( W_f \) effect of the transistor on its HF noise performance and, hence, there is a need to propose a new FOM to assist in this study for the \( W_f \) optimization.

C. Proposal of New FOM for HF Noise

In order to study the transistor’s \( W_f \) effect on its HF noise performance, a new FOM is proposed, as shown in (15), where \( F_{min} \) is in the minimum noise factor in (10) and \( R_n \) is the normalized noise resistance defined in a linear two-port noisy network [24]. In [25], the parameter \( R_n \) can be derived and simplified, as shown in (16), and the parameter \( \gamma \) is as discussed in (7). For long channel devices, the parameter \( \alpha \) is equal to unity and it will gradually decreases as the channel length reduces [25] as follows:

\[
FOM = F_{min} \cdot R_n
\]

\[
R_n = \frac{\gamma g_{d0}}{g_m} = \frac{\gamma}{\alpha} \cdot \frac{1}{g_m}
\]

\[
\alpha = \frac{g_m}{g_0}
\]

By multiplying the parameter \( F_{min} \) and \( R_n \) together, the resultant equation is as shown in (18) as follows:

\[
FOM = \frac{\gamma}{\alpha} \cdot \left( \frac{1}{g_m} \right) \cdot f \cdot K_1 \cdot \frac{\gamma}{\alpha} \cdot \sqrt{\frac{16\pi^2}{3g_m}} \cdot \left( R_g C_{gg} C_g + C_{gs}^2 (R_g + R_s + R_i) \right)
\]

\[
K_1 = \sqrt{P + R - 2C\sqrt{RP}}.
\]

By extracting the small-signal parameters for all the devices under comparison and substituting them back into (18), the calculated FOM can be obtained. Fig. 7 shows the measured and calculated FOM \( F_{\text{min}}' \cdot R_n \) versus frequency for different \( W_f \) transistors at a constant total width of 48, 120, and 240 μm. There are three important observations from this FOM versus \( W_f \) and total width change. Firstly, its value is found to be increasing with \( W_f \) and decreasing when the transistor’s total width increases. Note that for the noise figure in Fig. 6, only the NF_{min} trend with \( W_f \) is observed, there is no change in NF_{min} when the total width changes. The FOM behavior with respect to the total width and \( W_f \) can be explained by studying the geometry dependencies of term 1 and term 2 in (18). By substituting the geometry dependence of all the parameters into (18), it is clear that term 1 and term 2 are proportional to \( (L_s/N_f \cdot W_f) \) and \( (L_s^2/N_f) \), respectively. By adding the proportionality of these two terms, the resultant geometry dependence of the proposed FOM is found to be \( \left( \frac{L_s}{N_f \cdot W_f} \right) \cdot (1 + W_f \cdot L_g) \). Therefore, based on the above analysis for the proposed FOM, it is expected that at the same constant total width with \( W_f \) increases, the proposed FOM will increase and when the \( W_f \) is fixed with increasing total width, the proposed FOM will decrease. Hence, this analysis on the geometry dependence of the proposed FOM can explain the measurement results, as shown in Fig. 7. Secondy, it is also observed that the FOM has both an increasing and a decreasing trend with respect to frequency. The increasing and decreasing trend is mainly due to the contribution from term 2 and term 1 of (18), respectively. Note that in term 2, there exists a frequency variable so it will naturally contribute to the increasing trend of the FOM with frequency. As for term 1, it is, in fact, equivalent to \( R_n \), and in [19], it is found that \( R_n \) is inversely proportional to \( \gamma_{21} \) and as frequency increases, \( \gamma_{21} \) will increase and cause \( R_n \) to fall off with frequency. Finally, it is observed that the calculated FOM is closely matched to the measured FOM values and this confirms the derived FOM equation in (18) is correct and reliable.

In Fig. 7, the transistor with \( W_f \) of 4 μm for the three total widths shows the lowest FOM, and this implies that it has the lowest noise level when compared to the noise of the other \( W_f \) transistors. Note that such observation is not observable in Fig. 6, as the extracted NF_{min} for 4, 8, and 12 μm are all overlapping. Therefore, it is clear that the proposed FOM can help to differentiate the noise level behavior for a different \( W_f \) transistor, especially for a frequency range less than 5 GHz.

Hence, by only considering the HF noise optimization, the transistor’s \( W_f \) should be chosen as 4 μm, as they show the lowest noise. Note that \( f_{\text{cr}} \) of a transistor with \( W_f \) of 4 μm is the lowest while its \( f_{\text{MAX}} \) is at the highest level. Hence, from Fig. 4, if the optimum \( f_{\text{cr}} \) and \( f_{\text{MAX}} \) performance is required, \( W_f \) should be chosen at approximately 12 μm. From Fig. 7, the noise for \( W_f \) of 12 μm is only slightly higher and comparable to the 4-μm case. Hence, it can be concluded that for the optimization of \( f_{\text{cr}}, f_{\text{MAX}}, \) and HF noise, the optimum \( W_f \) to use for a transistor for this 0.18-μm CMOS technology is at approximately 12 μm.

IV. UNIT WIDTH OPTIMIZATION ON FLICKER NOISE

Flicker noise mainly affects the low-frequency performance of the transistor, but the impact of flicker noise cannot be neglected in certain RF circuits such as a mixer because it can up-convert the low-frequency noise to HF, which can affect the
Many theories have been presented to explain the flicker noise behavior. The two main theories are the random fluctuation of the carriers in the channel [26], [27], [28]–[30] and mobility fluctuation [31], [32]. Based on the carrier fluctuation [29], [30] and mobility fluctuation [31], [32] theory, the noise voltage power spectral density can be shown as in (20) and (21), respectively. Note that \( K_1 \) in (20) is bias independent, while in (21), \( KV_{GS} \) is a bias-dependent parameter. Their corresponding noise current power spectral density can be found using (22) as follows:

\[
S_{vg} = \frac{K_1}{C_{ox} W_{total} \cdot L_g} \frac{1}{f^c} \tag{20}
\]

\[
S_{vg} = \frac{K(V_{GS})}{C_{ox} W_{total} \cdot L_g} \frac{1}{f} \tag{21}
\]

\[
S_{id} = \frac{g_{m}^2}{2} \cdot S_{vg}. \tag{22}
\]

A. Experimental Results and Discussion

Fig. 8 shows the measured flicker noise versus frequency plot with varying \( W_f \) at a fixed total width of 48, 120, and 240 \( \mu \)m. The noise measurement is done at the maximum \( g_m \) biasing condition. It is observed that the \( W_f \) variation does not affect the flicker noise performance for the RF transistor at the same total width condition. Furthermore, the flicker noise current power spectral density \( S_{id} \) is also found to be increasing with increasing total width. In (20) or (21), it is clear that \( S_{vg} \) is inversely proportional to the total width of the transistor (\( W_{total} \)), while in (22), \( S_{id} \) is directly dependent on the transistor’s \( g_m \), which is also proportional to \( W_{total} \). Hence, the resultant geometry dependence of \( S_{id} \) can be derived and shown to be only the transistor’s total width (\( W_{total} \)), and this is clearly shown in the flicker noise measurement results. Since the measurement frequency for the flicker noise is low, the RF parasitic due to the \( W_f \) changes will not appear and affect the transistor’s flicker noise behavior. Therefore, it is expected that \( S_{id} \) shows no dependency with \( W_f \) at the same total width condition.

V. CIRCUIT APPLICATION DISCUSSION

From the analysis of \( f_{TR}, f_{MAX}, NF_{MIN} \), proposed FOM \( F_{MIN} \cdot R_{NB} \), and flicker noise spectral density, it shows that the unit width optimization technique can assist designers to select the optimized layout transistors for a specific circuit application such as an LNA, a VCO, or a mixer that required either low \( NF_{MIN} \), high \( f_{TR} \), or \( f_{MAX} \).

A. Transistor Selection for LNA Design

The LNA is usually the first stage of a receiver and it generally requires low noise figure and high-gain transistors in its circuit design. Based on the Friss equation [33], the first stage of the receiver will determine the whole receiver noise performance; hence, the noise in the LNA must be minimized. The transistor size used for this LNA design is usually large so as to provide enough gain to reduce the noise in the subsequent stages. The number of fingers for the chosen transistor must be large, hence, minimizes its gate resistance. This will cause the
amplified gate noise at the output of the LNA to be small. Therefore, the transistor’s $W_f$ to use for LNA design should be small so that maximum number of fingers can be achieved for a constant total width transistor. This conclusion coincides with the analysis that has been discussed in Section III.

B. Transistor Selection for VCO Design

For VCO design, the main considerations are low phase noise and low power. Since the transistor’s flicker noise will contribute to the output phase noise, the selected transistor in the VCO design should have low flicker noise spectral density. From Fig. 8, it can be observed that the flicker noise is directly dependent on the device size and drain current and it is independent of the transistor’s $W_f$. Therefore, small total width transistor size with small $W_f$ should be chosen for the VCO design so that it exhibits low flicker noise spectral density and high $f_{\text{MAX}}$, as presented in Figs. 8 and 4, respectively. Furthermore, using small total width for the transistor also ensure low power consumption for the VCO.

C. Transistor Selection for Mixer Design

The mixer is mainly used in the RF transceiver for up and down conversion of signals. In mixer design, the tradeoff for conversion gain, linearity, power consumption, and noise figure have to be studied so as to achieve the design specifications for the circuit. The noise figure for the mixer is generally large due to the switching term during the up and down conversion. By designing a mixer with low noise figure, the gain of the LNA can be low. Hence, from Fig. 7, a transistor with small $W_f$ can provide the lowest noise level and highest $f_{\text{MAX}}$, which allows higher operating frequency range for the mixer.

VI. Conclusion

In this paper, the trend of extracted $f_T$ and $f_{\text{MAX}}$ versus $W_f$ has been studied, and from the measurement results, it is clear that for the optimization of both FOMs, some tradeoff is required. As the measured $\text{NF}_{\text{min}}$ for $W_f$ of 4, 8, and 12 $\mu$m cannot reveal their noise behavior with respect to $W_f$, a new FOM $F_{\text{min}}$, has been proposed. From the calculated $F_{\text{min}}$, it has been clearly shown that it can be used to optimize the transistor noise when selecting the best $W_f$ value to be used in low-noise applications. The flicker noise spectral density with varying $W_f$ shows no optimization can be done for $W_f$, as at such low-frequency range, the RF parasitic will not appear and affect the transistor’s flicker noise. This experiment has shown that the optimization technique is feasible and can be used to help designers to select the optimized layout transistors that are optimized specifically for a certain application such as an LNA, a VCO, and a mixer circuit that have either low $\text{NF}_{\text{min}}$, high $f_T$, and $f_{\text{MAX}}$ requirements. Furthermore, by applying this optimization technique to existing and future technologies, the modeling engineers from foundry and IDM can understand the RF characteristics of the process and select to model only a small range of optimized $W_f$ transistors and shorten the model development time.

ACKNOWLEDGMENT

The authors would like to thank M. K. Raja, Institute of Microelectronics (IME), Singapore, and M. Nagarajan and T. C. Ng, both with Advanced RFIC (S) Pte. Ltd., Singapore, for their
helpful discussions. The authors are also grateful to Chartered Semiconductor Manufacturing Ltd. (Chartered), Singapore, for fabricating the test structures.

REFERENCES


Kiat Seng Yeo received the B.E. degree in electronics (Hons.) and Ph.D. degree in electrical engineering from Nanyang Technological University, Singapore in 1993 and 1996, respectively.

He is currently Sub-Dean and Associate Professor of electrical and electronic engineering with Nanyang Technological University (NTU), Singapore, where he is a recognized expert in CMOS technology and low-power CMOS integrated-circuit design. He is a consultant to multinational corporations. He has authored or coauthored over 200 papers. He has authored four books. His research interests include device modeling and RF IC design.

Prof. Yang serves on the Program Committee of several international conferences.

Zee Long Teng received the Diploma degree in electronics and computer engineering from Ngee Ann Polytechnic, Singapore, in 1998, and the Bachelor degree in computer science from the University of Adelaide, Adelaide, Australia, in 2006.

He was with the Technology Development Group and Photolithography Group, Chartered Semiconductor Manufacturing Ltd., as an Associate Engineer. He subsequently joined FTD Solutions as a Layout Engineer. He was then with the RF Device Modeling Group, Advanced RFIC (S) Pte. Ltd., Singapore, as a Device Modeling Engineer. He is currently a Computer-Aided Design (CAD) Engineer with the Institute of Microelectronics (IME), Singapore. He is involved in the process design kit (PDK) setting for integrated-circuit design, tool evaluation, and testing.

Pei Fern Ng received the Bachelor degree in electrical and electronics (with a major in electronics) from Nanyang Technology University (NTU), Singapore, in 2002.

In 2002, she joined the Electronic Design Automation (EDA) Group, Advanced RFIC (S) Pte. Ltd., Singapore, as an EDA Engineer, where she was involved in process design kit (PDK) development for various technology nodes and different processes. Her other involvements include test chip generation and customization of software for integrated-circuit designers. She is currently an Engineer with MediaTek Singapore Pte. Ltd., Singapore, where she supports both integrated-circuit designers and layout engineers in their problem analysis.