Yield-driven Iterative Robust Circuit Optimization Algorithm

Yan Li
Department of EECS,
Massachusetts Institute of Technology,
77 Mass Ave., Cambridge, MA 02139
liyan@mit.edu

Vladimir Stojanović
Department of EECS,
Massachusetts Institute of Technology,
77 Mass Ave., Cambridge, MA 02139
vlada@mit.edu

ABSTRACT

This paper proposes an equation-based multi-scenario iterative robust optimization methodology for analog/mixed-signal circuits. We show that due to local circuit performance monotonicity in random variations constraint maximization can be used to efficiently find critical constraints and worst-case scenarios of random process variations and populate them into a multi-scenario optimization. This algorithm scales gracefully with circuit size and is tested on both two-stage and fully differential folded-cascode operational amplifiers with a 90 nm predictive model. The improving yield-trends are confirmed across process and random variations with Hspice Monte-Carlo simulations.

Categories and Subject Descriptors
B.7.2 [Integrated Circuits]: Design Aids

General Terms
Algorithms

Keywords
Robust Circuit Optimization, Variability, Yield, Analog Circuits

1. INTRODUCTION

Traditional robust circuit design routine involves design iterations over process, voltage, and temperature (PVT) corners coupled with Monte-Carlo simulations for yield estimation. It is up to designers to interpret these estimation results to find design sensitivities, which is often complicated by the large dimensionality of design space, and results in long design time, over-design, and wasted power and area.

Equation-based circuit optimization enables easier global design-space exploration in a large-dimensional space and could potentially be used in design of robust circuits. Geometric programming (GP) has been used extensively in circuit design examples, from operational-amplifiers [1], [2] to phase-locked loops [3]. A straightforward approach to extend this methodology to robust design would be reformulating the optimization problem into a robust optimization problem. However, this robust optimization problem usually loses the tractability. To make the robust GP more tractable, [4] and [5] consider a special case of ellipsoidal uncertainty and use piecewise-linear convex approximation techniques. These techniques, however, do not scale well due to the number of piecewise-linear approximation terms, and are limited to circuits that are strictly GP compliant.

To circumvent the intractability in the robust optimization problem, and make use of the circuit optimization framework, in this paper, we propose an iterative yield-driven robust optimization algorithm. It inherits the iterative design fashion of a traditional robust design routine, but also takes advantage of the optimization engine’s global design space exploration power. The methodology gracefully extends the multi-scenario optimization used for design over PVT corners, by adding only the most critical random variations corners discovered through an efficient search step (worst-case analysis).

Other similar ‘worst-case analysis’ ideas have appeared in [6], [7] and [8]. To compute the ‘worst-case distance’, [6] uses local sensitivity analysis assuming performance linearization. In [7], the worst-case analysis is based on response surface modeling (RSM), which trades-off accuracy for size of design space and can become computationally expensive for large circuits. In [8], a robust taper is designed in a similar iterative fashion, but uses entirely different search technique due to a different problem structure. Without the dependence on RSM or local sensitivities, our algorithm is able to search globally for a robust design. Another advantage is that the size of the multi-scenario robustifying optimization problem scales linearly with the number of performance constraint, which in most cases, is much smaller than the number of variation variables. This makes our algorithm readily scalable to large circuits.

The remainder of the paper is organized as follows. Section 2 elaborates the details of the algorithm. Section 3 presents a two-stage and a folded-cascode operational amplifiers (op-amp) examples to show monotonic yield convergence and applicability of the algorithm for fast analog/mixed-signal robust circuit optimization.

2. ALGORITHM

2.1 Sources of Variability

In this paper, we focus on with-in die mismatch induced by random dopant fluctuations and line-edge roughness. Pelgrom’s model [9] and derivative work, to the first order show that the impact of process variations on deviations of threshold voltage $V_T$ and the relative current factor $\beta$ are inversely dependent on the area of the transistor, as

$$\sigma_{V_T} = \frac{AV_T}{\sqrt{W \cdot L}}, \quad \text{and} \quad \frac{\sigma_{\beta}}{\beta} = \frac{AV_T}{\sqrt{W \cdot L}}. \tag{1}$$
2.2 Proposed robust circuit optimization framework

The intuition behind this algorithm is illustrated in Fig. 1. In Fig. 1, (a)-(e) depict the impact of variations on feasible design space, the nominal design point, and how the algorithm pushes the nominal design point to be more robust in two iterations. Here, we assume a circuit design problem with underlying design variables as circuit sizing parameters (x-y plane, not shown here) and objective to minimize the power consumption, represented by different level contours. The feasible design space is constrained by a set of biasing and performance constraints simplified to a polygon in the figures. In (a), a nominal design point is marked with a dot and one of the constraints (e.g. gain) is active. Under variations, the gain constraint shifts with respect to the x-y plane (the dash line) and nominal design point now violates the design specification with some probability. Although the objective can also change under variations, it is not shown here for simplicity. In (c), the algorithm generates a variation-aware circuit scenario (adds constraints) to exclude the design space that would fail the gain specification under variations and push the design to a more robust design point, the triangle marker. In (d) and (e), the algorithm iterates once again, expanding the variation set.

While this type of iterative algorithm has been considered previously in other applications, e.g. [8], here we outline the necessary steps to map the algorithm to a circuit optimization scenario, by solving the inherent circuit biasing and large variation issues. The following subsections elaborate the algorithm flow as shown in Fig. 2 and map the algorithm steps on a two-stage op-amp circuit example in Fig. 3, to illustrate the specific circuit-related refinements.

2.2.1 Initial design

The initial design is a nominal circuit optimization problem without consideration of variations. It corresponds to the Robustified circuit optimization block without any added circuit scenarios in Fig. 2. The formulation is as the following:

\[
\begin{align*}
\text{minimize} & \quad f_0(x, \mu^*) \\
\text{subject to} & \quad f_i(x, \mu^*) \leq 1, i = 1, 2, \ldots, n, \\
& \quad g_j(x, \mu^*) = 0, j = 1, 2, \ldots, m, \\
& \quad h_k(x, \mu^*) \leq 1, k = 1, 2, \ldots, l, \\
\mu^* &= 0. 
\end{align*}
\]

where \( f_i(x, \mu^*) \) and \( g_j(x, \mu^*) \) are biasing constraints and \( h_k(x, \mu^*) \) are performance constraints. \( \mu^* = 0 \) means that no variation is considered here. In the two-stage op-amp example, we minimize a weighted sum of power and area, subject to the performance constraints illustrated in Table 1, and biasing constraints (KVL, KCL, transistor regions).

After the initial design, we make the optimization variation-aware in the following iterations by adding the critical variation direction information. To obtain the critical variation directions, the next subsection Constraint maximization performs a smart search and adds the information in the form of new circuit scenarios to the Robustified circuit optimization block.

2.2.2 Constraint maximization

Constraint maximization identifies the set of critical performance constraints that would violate the corresponding specifications under process variations. It involves solving the optimization problem defined in (3) for each performance constraint. This makes the al-
algorithm scale well, since the number of performance constraints
does not necessarily grow as the size of the circuit grows.

The problem can be formulated as

\[
\begin{align*}
\text{maximize} & \quad h_k(x^*, \mu), \quad k = 1, 2, \ldots, l, \\
\text{subject to} & \quad \mu \in U, \\
\text{Biasing constraints},
\end{align*}
\]

where \( h_k(x^*, \mu) \) is the performance constraint evaluated at
design point \( x^* \) from previous circuit optimization step, with
variation variable vector \( \mu \). The constraints are bounding and biasing
constraints. No probability distribution is assumed. The range for the
\( i \)-th variation variable can be written to be proportional to its
standard deviation as \( \mu_i = [-k \sigma_i, k \sigma_i] \), where \( \sigma_i \) is the standard
deviation of the \( i \)-th variation and \( k \) is a constant. In each iteration, the
variation set can be expanded by increasing \( k \). In the two-stage
op-amp example, for the bandwidth constraint maximization, the
objective in (3) becomes \( h_k(x^*, \mu) = \text{spec.} \omega_2/\omega_3(x^*, \mu) \).

The purpose of expanding the variation set in each iteration is to
take more variations into consideration and search for more critical
variation directions that would fail the design. Therefore, as iter-
ations go on, the robustified circuit optimization is aware of more
critical directions and pushes the design to become more and more
robust. By expanding the variation set iteratively, we achieve a de-
sign just meeting the yield target, avoiding over-design.

The key circuit-specific refinement in this step is to realize that
during the maximization, circuit biasing constraints have to be sat-
sified even in the presence of random variations for the circuit to
be realizible. Therefore, all circuit descriptions are defined with
the macro transistor model in Fig. 4. Fig. 4(a) reflects the effect of
threshold and current factor variations through the voltage source
\( \Delta V_T \) and the current source \( M_L \), defined in (1). We can further
simplify the model to Fig. 4(b), which has an aggregate standard
deviation accounting for both threshold and current factor vari-
ations, as shown in (4),

\[
\sigma_{V_T}^2 = \frac{A_B^2}{W \cdot L} \left( \frac{I_d}{g_m} \right)^2 + \frac{A_V \sigma_i}{W \cdot L}.
\]

Although (3) seems to be a general optimization problem, a close
examination on the dependence of circuit performances on process
variations reveals that circuit performances can be assumed to be
locally monotonic on process variations (but not circuit design pa-
rameters). This assumption holds reasonably well according to the
work in RSM, since linear regression model is a traditional way to
fit the circuit performance on variation variables [10] [11] and lin-
ear model is a monotonic function on fitting variables, the process
variations. Furthermore, the variation set is always set to be small
initially, i.e., \([-0.2 \sigma_{V_T}, 0.2 \sigma_{V_T}] \), and often does not need to be ex-
expanded to as large as \([-3 \sigma_{V_T}, 3 \sigma_{V_T}] \). In the two examples we show
in the next section, \( k \) only increases to around 1 with yield already
approaching 100\%. Under this small variation range, the accu-
tracy of the linear model is satisfactory, hence our assumption on
monotonicity is also reasonable.

To validate the assumption on this op-amp example, we let \( k =
[1, 2] \), with each transistor’s \( \Delta V_T \) ranging from \(-k \sigma_{V_T} \) to \( k \sigma_{V_T} \). For
each \( k \), we select random variation corners out of the total \( 2^k \)
corners and record the DC gain and unity-gain bandwidth of the design
under these variations. The performances show nicely monotonic
functions as we increase the \( k \), as shown in Fig. 5, where the x-
axes denotes equally spaced points along the line. A more general
test is to randomly select a line in the variation variable space, i.e.,
\( s, v \in [-20 \sigma_{V_T}, 20 \sigma_{V_T}] \) and let \( s + t \cdot v \) be the variation vector with
\( t \) as a varying scalar. Fig. 6 shows monotonic performance vari-
ations along different random lines. Because of the monotonicity
of performances on variation variable, the objective function in (3)
becomes monotonic together with bounding constraints, and the
optimal solution of the maximization problem can be found easily
with general gradient-based optimization solvers.

Although there is no theoretical proof showing the monotonicity
property on variations, it is confirmed in the two examples pre-
sented here. In the worst situation where the monotonicity does not
hold, the algorithm should still work, with the price of adding sub-
worst circuit scenarios during iterations. Those sub-worst circuit
scenarios can still help push the design to be more robust.

We also notice that since the optimization solver is inherently
a GP-based solver, it requires all design variables to be positive.
However, \( \Delta V_T \) range obviously includes negative values. A brute-
force way to overcome this problem is to prefix the sign of \( \Delta V_T \).
and maximize over all sign combinations, which makes the number of maximization grow exponentially with the number of variation variables. However, with the objective function being monotonic over variation range, it is easy to map the solution from the maximization done under the positive variations to a solution under the whole variation range. This is illustrated in Fig. 7, in which case two transistor variations are considered. Suppose we maximize \( \text{spec} \omega_o / \omega_p \) over the shadow area. Then, because of monotonicity, the optimal solution should be on one of the four vertices of the shadow area. If we obtain the optimal solution at \( O_1 \) and know that the objective function increases as \( \mu_{V_T} \) decreases, then if not restricted to positive range, the solution \( O_1 \) should shift to \( O_1' \). This simple mapping can help obtain the optimal solution to the maximization problem efficiently.

2.2.3 Robustified circuit optimization

With the critical variation directions obtained in constraint maximization, the next step is the Robustified circuit optimization block in Fig. 2, to achieve a new robust design. If the design has not been robust over the whole variation range, we should be able to identify a nonempty set of constraints whose right-hand sides are violated, i.e. \( h_k(x^*, \mu^*) > 1 \). Since each of the worst-case variation vectors \( \mu_i^* \) can cause a different biasing condition, a circuit scenario has to be instantiated for each of them. The new circuit scenario is setup such that it has the same topology and shares the optimization variables \( x \) with the nominal circuit, except that it uses the macro model in Fig. 4(b) for transistors, with variability vector having the sign of \( \mu_i^* \) and magnitudes parameterized by optimization variable \( x \).

The key point in the context of the circuit optimization, e.g. two-stage op-amp, is that the maximization solution \( \mu^* = \Delta V_T^* \) is used to determine only the polarity of \( \Delta V_T \), and the value of \( \Delta V_T \) in the macro model is a function of design variables as \( k \sigma_{V_T} \). This enables the optimizer to recognize that resizing the circuit will help decrease the variation and improve the worst performance. Here, \( k \) is the constant used to determine the range of the variation set in maximization, and \( \sigma_{V_T} \) is defined in (4). Therefore, the new scenario reflects the real situation when the nominal circuit is under variation with degraded performance. By doing multi-scenario optimization, we ensure that the degraded performance meets the specification, thus giving the nominal performance more margin to fail.

From the perspective of computation cost, adding clone circuits does create a larger optimization problem to solve. However, notice that the number of clone circuits is equal or less than the number of performance constraints, which in usual case is on the order of ten or less. Besides, the interior-point method used in GP-based solvers scales gracefully with increase in the number of constraints [12], since the number of optimization variables remains the same as in the nominal optimization problem.

2.2.4 Yield estimation

Yield estimation closes the loop in Fig 2. Although many fast yield estimation methods exist, for instance importance sampling [13], pseudo-noise analysis [14], here we use the direct Monte-Carlo sampling method performed in optimization domain. For each Monte-Carlo variability sampling point the optimizer solves a feasibility problem with all design variables fixed, checking the yield with proper biasing and performance constraints. In the result section we will see that the yield estimated in optimization domain follows the yield obtained from Hspice simulation, making these other simulation-based yield estimation methods possible to use in this context as well.

3. EXPERIMENTAL RESULTS

Two circuit examples are presented in this section: a two-stage and a folded-cascode op-amp. In the first example, we consider two different starting design points and show that in both cases, the algorithm is able to add more robustness to the circuit and increase the design yield. The second example has four times the number of variation variables as in the first example and the algorithm still behaves nicely. In both examples, the optimizer uses sigmoidal transistor models derived from a 90 nm predictive technology model.

3.1 A two-stage op-amp example

This is the example discussed in Section 2, with the schematic shown in Fig. 3. The specifications of the op-amp are listed in Table 1 with 2 pF load capacitance and design objective to minimize a weighted sum of gain and area. The variation sources under consideration are threshold voltage and current factor of total 8 transistors.

3.1.1 One-corner initial design

The initial nominal design is optimized under tt corner with 1 V supply voltage, 10 \( \mu \text{A} \) reference current and temperature at 298 K. The nominal design just meets the gain and bandwidth (\( \omega_o \)) constraints. In the first iteration, only these two constraints are found to violate the specifications after maximizations. Based on optimal solutions from maximization, we instantiate two circuit scenarios where the polarity of the \( \Delta V_T \) voltage sources are determined from the sign of maximization solutions. Then a 3-scenario circuit optimization is solved and we reach a robust design with gain and \( \omega_o \) improved to 311 and 163 MHz (measured without variability).

As iterations go on, the design margins increase at the same time, shown in the two columns under optimization in Table 2. Table 2 also shows the simulation results for verification. Because of the transistor model's multi-fold error, there is around 20% mismatch between Hspice simulations and optimization results. However, the simulation results do show the same trend of the improved performance and yields, with yields shown in the top two figures in Fig. 8 (the Hspice yields are calculated with respect to the simulated nominal design performance without variations). The cost of robustified design is increased power and area shown in the bottom two figures in Fig. 8.

Table 2: Robust two-stage op-amp designs in iterations from optimization and Hspice simulation

<table>
<thead>
<tr>
<th>k</th>
<th>DC gain</th>
<th>( \omega_o ) (MHz)</th>
<th>DC gain</th>
<th>( \omega_o ) (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>300</td>
<td>160</td>
<td>263</td>
<td>125</td>
</tr>
<tr>
<td>0.2</td>
<td>311</td>
<td>163</td>
<td>268</td>
<td>130</td>
</tr>
<tr>
<td>0.4</td>
<td>320</td>
<td>166</td>
<td>273</td>
<td>133</td>
</tr>
<tr>
<td>0.6</td>
<td>328</td>
<td>168</td>
<td>275</td>
<td>135</td>
</tr>
<tr>
<td>0.8</td>
<td>337</td>
<td>170</td>
<td>281</td>
<td>137</td>
</tr>
<tr>
<td>1</td>
<td>344</td>
<td>172</td>
<td>282</td>
<td>139</td>
</tr>
<tr>
<td>1.2</td>
<td>351</td>
<td>173</td>
<td>285</td>
<td>141</td>
</tr>
<tr>
<td>1.4</td>
<td>360</td>
<td>174</td>
<td>288</td>
<td>142</td>
</tr>
</tbody>
</table>
the second corner gradually increase to around 100%, shown in Fig. 9. The final robust design achieved in optimization is shown in Fig. 10, compared with the initial design. The cost to achieve the improvement is shown in the bottom two figures in Fig. 9. The simulations in Hspice of the second corner again show the same trend in yields in Fig. 9.

### 3.2 A fully differential folded-cascode op-amp with common-mode feedback

Fig. 11 shows the schematic of a fully differential folded-cascode op-amp with common-mode feedback (CMFB). The specifications considered in this example are listed in Table 4, with the objective to minimize a weighted sum of power and area. The variation sources under consideration are the threshold and current factor variations from total 32 transistors.

Table 5 shows the robust design evolution during iterations. The initial design has a transconductance value just on the lower bound of the specification and this critical constraint is found in the first maximization. To prevent this from happening, a circuit scenario is instantiated and along with the nominal circuit, it pushes the new robust design to a high transconductance value close to the upper bound. Therefore, in the second iteration, maximization finds another critical constraint, the transconductance upper bound constraint.

Table 4: The folded-cascode op-amp example: specifications for nominal design

<table>
<thead>
<tr>
<th>Specifications</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC gain</td>
<td>50 dB</td>
</tr>
<tr>
<td>Phase margin</td>
<td>60°</td>
</tr>
<tr>
<td>Gain</td>
<td>[0.5 mS, 0.6 mS]</td>
</tr>
<tr>
<td>CMFB DC gain</td>
<td>60 dB</td>
</tr>
<tr>
<td>CMFB unity-gain bandwidth (ω&lt;sub&gt;u&lt;/sub&gt;)</td>
<td>5.9 MHz</td>
</tr>
</tbody>
</table>

Table 3: Five-corner of the two-stage op-amp initial design.

<table>
<thead>
<tr>
<th>number</th>
<th>corner</th>
<th>temp (K)</th>
<th>vdd (V)</th>
<th>Iref (μA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>tt</td>
<td>298</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>ss</td>
<td>398</td>
<td>0.9</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>ff</td>
<td>233</td>
<td>1.1</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>fs</td>
<td>398</td>
<td>0.9</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>sf</td>
<td>233</td>
<td>1.1</td>
<td>8</td>
</tr>
</tbody>
</table>

Figure 11: The folded-cascode op-amp schematic.
Table 5: Folded-cascode op-amp: iterations of the robust designs from optimization: $g_m \in [0.5 \text{ mS}, 0.6 \text{ mS}]$

<table>
<thead>
<tr>
<th>$k$</th>
<th>violated constraints</th>
<th>$g_m$ (mS)</th>
<th>yield (%)</th>
<th>area ($\mu m^2$)</th>
<th>power (mW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$g_m \geq 0.5$</td>
<td>0.5</td>
<td>49</td>
<td>539</td>
<td>0.27</td>
</tr>
<tr>
<td>0.2</td>
<td>$g_m \leq 0.6 m$</td>
<td>0.59</td>
<td>69</td>
<td>544</td>
<td>0.30</td>
</tr>
<tr>
<td>0.4</td>
<td>None</td>
<td>0.58</td>
<td>84</td>
<td>627</td>
<td>0.31</td>
</tr>
<tr>
<td>0.6</td>
<td>None</td>
<td>0.57</td>
<td>92</td>
<td>639</td>
<td>0.32</td>
</tr>
<tr>
<td>0.8</td>
<td>None</td>
<td>0.55</td>
<td>94</td>
<td>668</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Table 6: The computational time breakdown per iteration step for the two-stage and folded-cascode robust design examples

<table>
<thead>
<tr>
<th></th>
<th>Two-stage op-amp</th>
<th>Folded-cascode op-amp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal design</td>
<td>10 s</td>
<td>17 s</td>
</tr>
<tr>
<td>Maximization</td>
<td>6 specs: 20 s</td>
<td>5 specs: 1 min</td>
</tr>
<tr>
<td>Redesign with 1-scenario</td>
<td>11 s</td>
<td>30 s</td>
</tr>
<tr>
<td>Redesign with 2-scenarios</td>
<td>20 s</td>
<td>40 s</td>
</tr>
</tbody>
</table>

straint, and instantiates another circuit scenario to guard the upper bound. The next robust design that comes out of a three-scenario circuit optimization gives a lower $G_m$ value and a higher yield. As iterations continue, no more constraint violations are found and the variation set keeps expanding, until the yield is satisfactory. As expected, the transconductance $G_m$ from the optimizer finally converges to the middle point of the specification range, i.e. 5.5 mS. Fig. 12 shows the Monte-Carlo performed in optimization domain of the initial and final robust design’s $G_m$ and the area and power consumption are shown in Table 5. Again, we can see the tradeoff between the circuit performance and area, power cost.

3.3 Computational efficiency

Optimizations ran on a server with 3.16 GHz Intel® Xeon® processor and 16 GB of memory running Linux. The time cost in one iteration consists of maximization time and multi-scenario redesign time. The approximate solving time for these two examples in one iteration is shown in Table 6. It takes 6 – 7 iterations (10 – 15 minutes) to achieve a robust design under random variations. The problem size of the two-stage op-amp redesign with 2-scenarios is around 200 variables and 600 constraints while for the folded-cascode op-amp redesign with 2-scenarios, it has around 800 variables and 2500 constraints.

4. CONCLUSION

In this paper we show that complicated and often intractable yield-driven robust circuit optimization can be performed in a scalable and tractable manner in a few iterative steps. The proposed algorithm is iterative and first decomposes the yield-driven robust optimization into minmax optimization followed by yield-estimation. We further show that the maximization step can be performed efficiently, due to the circuit-related monotonic properties on variations of the performance functions. We find that the key to success in this stage is to maintain the circuit biasing constraints during the maximization step. The multi-scenario optimization is used in the last step, where the key ingredient is to introduce only the relative values of the variability corners and let the solver figure out the updated absolute values based on the state of optimization variables. Lastly, the variability space bound is increased and algorithm iterates until the satisfactory yield has been achieved. Coupled with these key circuit optimization related insights, this iterative minmax optimization presents an efficient and scalable solution for yield-driven circuit optimization.

5. ACKNOWLEDGMENTS

Figure 12: Monte-Carlo check with the initial and final robust designs: $G_m$ of the folded-cascode op-amp.

The authors would like to thank Mar Hershenson, Sunderrajan Mohan, Dave Colleran, Almir Mutapović and Shyne Tseng at Magma Design Automation, Inc for all their technical discussions and support. This work is funded by the Center for Integrated Circuits and Systems at MIT.

6. REFERENCES