An SQP and Branch-and-Bound Based Approach for Discrete Sizing of Analog Circuits

Michael Pehl and Helmut Graeb
Technische Universitaet Muenchen
Germany

1. Introduction

Analog circuits form an important part in integrated circuits and in particular in ASICs (Application Specific Integrated Circuits). However, due to the high complexity, design of this part has become a bottle-neck in the design flow (Gielen, 2007; Rutenbar et al., 2007). To overcome this problem and to guaranty that the analog part can be designed in reasonable time even for future technologies, methods supporting automatic design of analog circuits must be advanced.

This chapter focuses on sizing of analog circuits. It starts from the point where a topology is given. The task now is to choose design parameters, e.g., lengths and widths of transistors, such that certain properties of the circuit are fulfilled.

Current tools to solve the sizing task mostly treat it as a continuous optimization problem and use, e.g., certain gradient-based approaches to solve the problem in the continuous domain (Graeb, 2007). However, many design parameters are discrete in reality, e.g., transistor multipliers (i.e., the number of transistors connected in parallel), or must be discretized for some practical purposes, e.g., transistor lengths and widths which should match to a manufacturing grid. Furthermore, for some future technologies as, e.g., FinFETs (Knoblinger et al., 2005), the transistor parameters must fulfill certain geometrical properties, and accordingly have to be discrete.

Considering discrete parameters, it is not sufficient to treat the sizing task as a continuous optimization problem and rounding the result. This can be followed from mathematical theory, where it is shown that continuous optimization with sub-sequent rounding might not solve the original discrete optimization task (i.e., a optimization task that considers discrete and continuous parameters) and leads to a suboptimal result (Li & Sun, 2006; Nemhauser & Wolsey, 1988). This can be confirmed by experiments.

To solve discrete optimization problems, statistical and evolutionary approaches have been proposed (Alpaydin et al., 2003; Cao et al., 2000; Gielen et al., 1990; Ochotta et al., 1996; Phelps et al., 2000; Somani et al., 2007). However, for practical approaches these tools are usually more slowly in comparison to deterministic gradient-based tools if a good initial solution can be given for the task (what is normally true for analog sizing). Even if statistical and evolutionary approaches might be the first choice if a global search is necessary, for many cases deterministic gradient-based approaches are more suitable. Deterministic approaches for discrete sizing of analog circuits have barely been published till today (Pehl & Graeb, 2009; Pehl et al., 2008). In this chapter a new deterministic gradient-based approach is presented. It
consists of Sequential Quadratic Programming and Branch-and-Bound.
For the approach in this chapter, the problem is sub-divided into a non-linear program (NLP) and a discrete program (DP). Afterward, a discrete program is modeled by a discrete quadratic program (DQP) to speed up the algorithm.
Before the algorithm is presented, it is shown in Section 2 how the task of analog sizing can be formulated as a discrete minimization program. The task is said to be solved if any parameter set is found, where sizing constraints as well as performance specifications are fulfilled.
Introducing a relaxation of the parameters (i.e., all parameters are considered to be continuously scalable), a non-linear, but continuous sub-problem can be defined, called the relaxed program. To solve this NLP, in Section 3.1 of the chapter a sequential quadratic programming (SQP) algorithm is introduced.
Obviously, the result of the relaxed program is a point in the relaxed - i.e., continuous - domain. So, in Section 3.2 of the chapter a Branch-and-Bound approach is introduced to find a discrete solution to the sizing task.
The algorithm based on SQP and Branch-and-Bound can be used to solve the discrete sizing problem. However, to improve the run time of the approach, in Section 3.3 of the chapter a modification to speed up the algorithm is described. In the modification, the quadratic model of the objective function - which is computed in the SQP algorithm - is used to get a discrete quadratic model of the original sizing task. By solving the discrete quadratic program a discrete point can be found which gives an approximation for the obtainable discrete solution. This approximation can be used to cut non-promising parts of the Branch-and-Bound tree and to speed up the algorithm.
Experimental results in Section 4 show that in contrast to continuous optimization with subsequent rounding the presented approach is able to find a discrete feasible solution in each test case. Furthermore, it can be seen that the modification described in Section 3.3 decreases the run time of the algorithm significantly without reducing the result quality.
Section 5 concludes and gives an outlook to future research.

2. Problem formulation

2.1 Sizing task

In the analog sizing step appropriate values for the design parameters $d$ of a given topology must be computed such that certain properties of the circuit are fulfilled. Typical design parameters are, e.g., lengths and widths of transistors, which were normally considered as continuous scalable in previous gradient-based approaches. However, in reality most parameters in the circuit sizing step are discrete, e.g., due to manufacturing grids, due to modern transistor types as FinFETs, or due to properties from the layout step.
For the approach presented in this chapter, the sizing task is formulated as a discrete optimization task, i.e., a sizing task considering scalable discrete and continuous parameters. For this purpose the vector of design parameters $d$ can be subdivided into three parts corresponding to different parameter classes:

1. Continuous parameters $d_c$ are used to model design parameters which do not require the consideration of any grid and which lie in an $N_c$-dimensional domain $D^{N_c}$ that is bounded by any upper bound $d_{c,U}$ and any lower bound $d_{c,L}$

$$d_c \in D^{N_c} = \{ d | d_{c,L} \leq d \leq d_{c,U} \}$$ (1)
2. Scalable discrete parameters $d_d$ are used to model design parameters which can only lie on a – not necessarily uniform – $N_d$-dimensional grid. These parameters $d_d$ are subset of a domain $\mathcal{D}^{N_d}$:

$$d_d \in \mathcal{D}^{N_d} = \bigtimes_{k=1}^{N_d} \mathcal{D}_i$$  \hspace{1cm} (2)

$\mathcal{D}_i$ is a set corresponding to the $i$-th discrete parameter $d_i$. Furthermore, $\mathcal{D}_i$ is ordered by a relation $<$ (Pehl & Graeb, 2009). Assuming $n_i$ discrete parameter values for parameter $d_i$, the ordered set can be formulated as:

$$d_i \in \mathcal{D}_i := (\mathcal{D},<) = \{d_{i,1}, ..., d_{i,k}, ..., d_{i,n_i}\}$$  \hspace{1cm} (3)

3. Non-scalable discrete parameters $d_x$ can be used to consider design options which can not be expressed by a scalable parameter and must be enumerated instead, e.g., the exchange of different technologies. This class of parameters is non-numerical in many cases. One way to consider this class of parameters, which fits to the approach presented in this chapter, is to define binary surrogate parameter for each design option. Assuming $n_i$ discrete design options $d_{i,1}, ..., d_{i,n_i}$ for a parameter $d_i$, i.e.,

$$d_i \in \mathcal{D}_i := \{d_{i,1}, ..., d_{i,k}, ..., d_{i,n_i}\}$$  \hspace{1cm} (4)

the values are collected in a vector $d_{x,i}$:

$$d_{x,i} = [d_{i,1}, ..., d_{i,k}, ..., d_{i,n_i}]^T$$  \hspace{1cm} (5)

Additionally the vector $d_{b,i}$ of surrogate design parameters is defined as:

$$d_{b,i} = [d_{b,1}, ..., d_{b,k}, ..., d_{b,n_i}]^T$$

$$d_{b,k} = \begin{cases} 1; \text{ when option } d_i \text{ should be chosen} \\ 0; \text{ otherwise} \end{cases}$$  \hspace{1cm} (6)

Thus, the vector of surrogate design parameters can be mapped to the value of the corresponding non-scalable discrete parameter $d_i$ by

$$d_i = d_{b,i}^T d_{x,i}$$  \hspace{1cm} (7)

To avoid that different options are chosen for the same parameter, an additional constraint must be added to the optimization task defined below for each non-scalable discrete parameter:

$$d_{b}^T d_{b} = 1$$  \hspace{1cm} (8)

The set of all binary variables corresponding to options for non-scalable parameters is assigned as $\mathcal{D}^{N_b}$.
In this chapter only continuous and scalable discrete parameters are used. However, using the binary surrogate parameters defined above, the approach in Section 3 can be applied accordingly. For continuous parameters $d_c$ and scalable discrete parameters $d_d$ the domain $D^N$ of the design parameters $d$ can be defined as:

$$d \in D^N = D^{N_c} \times D^{N_d}$$

Thus, vector $d$ is composed by a continuous part $d_c$ and a discrete part $d_d$:

$$d = \begin{bmatrix} d^T_c \ d^T_d \end{bmatrix}^T$$

The task of choosing a parameter point, such that certain circuit properties are fulfilled, now is formulated as:

$$\min_{d \in D^N} \varphi(d) \text{ s.t. } c(d) \geq 0$$

wherein $c(d)$ are sizing constraints, which ensure a reasonable sizing of the circuit (Graeb et al., 2001; Massier & Graeb, 2008). $\varphi(d)$ is the objective function, which maps a multi objective optimization task to a scalar minimization problem. The objective function for analog sizing should support improvement of any circuit property when the specification for a certain performance is fulfilled as well as when the specification is violated. To build up such a function, an error $\epsilon(d)$ for each performance $f_i(d)$ is defined, which is the normalized distance from the current performance value to the specification bound $f_{B,i}$ of the performance:

$$\epsilon(d) = \frac{f_i(d) - f_{B,i}}{f_{N,i}}$$

$f_{N,i}$ is a normalization factor which ensures that the values for all performances are comparable. Without loss of generality it can be assumed that $f_{B,i}$ is a lower bound for the performance such that

$$\epsilon(d) \geq 0 \text{ when specifications are fulfilled}$$
$$\epsilon(d) < 0 \text{ when specifications are not fulfilled}$$

This is illustrated in Figure 1. To support improvement of the performances when the specifications are violated as well as when the specifications are fulfilled, in this approach an exponential sum of the normalized errors for all $N_f$ performances is used:

$$\varphi(d) = \sum_{i=1}^{N_f} e^{-\epsilon_i(d)}$$

Although the given formulation leads to a Pareto-optimal point, i.e. a solution where one performance can not be further improved without deteriorating another performance, we choose to stop the optimization problem – and consider the sizing task solved – as soon as a point is found which fulfills all specifications. Thus, the minimization is stopped as soon as a point is found with:

$$\forall i = 1, ..., N_f \quad \epsilon(d) \geq 0$$
To set up the relaxation of a discrete optimization task, the domain for each discrete parameter is replaced by a continuous domain. As the domain for the discrete parameters in (3) can be an ordered, the lower bound \( d_{i,L} \) and upper bound \( d_{i,U} \) for a discrete parameter \( d_i \) can be defined as the first and the last element of the ordered set \( D_i \)

\[
d_{i,L} := d_{i,1} \quad \text{and} \quad d_{i,U} := d_{i,n_i}
\]

For all discrete parameters, the lower bounds can be collected in a vector \( d_{d,L} \)

\[
d_{d,L}^T = [...d_{i,1}...]
\]

and, respectively, the upper bounds can be collected in a vector \( d_{d,U} \)

\[
d_{d,U}^T = [...d_{i,n_i}...]
\]

Thus, a vector of lower and upper bounds for discrete and continuous parameters can be built up:

\[
d^T_L = \begin{bmatrix} d_{d,L}^T & d_{c,L}^T \end{bmatrix} ; \quad d^T_U = \begin{bmatrix} d_{d,U}^T & d_{c,U}^T \end{bmatrix}
\]

For the relaxed optimization task all parameter points must be in the domain

\[
D_{rel}^N = \{ d | d_L \leq d \leq d_U \}
\]

and the relaxed program can now be defined as:

\[
\min_{d \in D_{rel}^N} \varphi(d) \quad \text{s.t.} \quad c(d) \geq 0
\]

The relaxation of a problem is illustrated in Figure 2.
must be true.
The evaluation of the circuit performances in this approach is done by simulations. Thus, in the rest of this chapter it is assumed, that – even if parameters are discrete – simulation of the circuit is possible for each continuous point. In the future work the algorithm will be extended to use exclusively simulation results from discrete points.

3. Discrete sizing approach

3.1 Sequential Quadratic Programming

In the approach, presented in this paper, the relaxed optimization problem in (21) is solved by a Sequential Quadratic Programming (SQP) approach (e.g., (Nocedal & Wright, 1999)). SQP converts a constrained nonlinear optimization problem in the continuous domain into a sequence of unconstrained quadratic programming problems.

Using a vector of Lagrange multipliers \( \lambda \), the Lagrangian function of the problem can be given as

\[
\mathcal{L}(\mathbf{d}, \lambda) = \varphi(\mathbf{d}) - \lambda^T \mathbf{c}(\mathbf{d})
\]

and the optimization problem (21) can be reformulated as the unconstrained optimization task:

\[
\min_{\mathbf{d}, \lambda} \mathcal{L}(\mathbf{d}, \lambda)
\]  

(24)

From the first-order necessary optimality conditions for unconstrained optimization, it follows that \( \nabla_{\mathbf{d}, \lambda} \mathcal{L}(\mathbf{d}, \lambda) = 0 \) must hold true in the optimum point. Thus, to find the optimum of the Lagrangian function, Newton’s method can be applied to the gradient \( \nabla_{\mathbf{d}, \lambda} \mathcal{L}(\mathbf{d}, \lambda) \).

Using \( \Delta \mathbf{d} \) and \( \Delta \lambda \) for the change in parameters and \( \mu \) as a linearisation index, the equation system which must be solved can be given as:

\[
\begin{bmatrix}
\nabla_{\mathbf{d}}^2 \varphi(\mathbf{d}(\mu)) - \sum_i \lambda_{i,0} \mathbf{c}_i(\mathbf{d}(\mu)) - \nabla_{\mathbf{d}} \mathbf{c}^T \\
-\nabla_{\mathbf{d}} \mathbf{c}
\end{bmatrix}
\cdot
\begin{bmatrix}
\Delta \mathbf{d}^{(\mu+1)} \\
\Delta \lambda^{(\mu+1)}
\end{bmatrix}
= \begin{bmatrix}
-\nabla_{\mathbf{d}} \mathcal{L}(\mathbf{d}(\mu), \lambda(\mu) \\
\mathbf{c}(\mathbf{d}(\mu))
\end{bmatrix}
\]

(25)

or with \( \mathbf{H} = \nabla_{\mathbf{d}}^2 \varphi(\mathbf{d}(\mu)), \mathbf{J} = \nabla_{\mathbf{d}} \mathbf{c}(\mathbf{d}(\mu)), \mathbf{C}_i = \nabla_{\mathbf{d}}^2 \mathbf{c}_i(\mathbf{d}(\mu)), \) and \( \mathbf{g} = \nabla_{\mathbf{d}} \varphi(\mathbf{d}(\mu)) \).

\[
\begin{bmatrix}
\mathbf{H} - \sum_i \lambda_{i,0} \mathbf{C}_i & -\mathbf{J}^T \\
\mathbf{J} & 0
\end{bmatrix}
\cdot
\begin{bmatrix}
\Delta \mathbf{d} \\
\Delta \lambda
\end{bmatrix}
= \begin{bmatrix}
-\mathbf{g} \\
\mathbf{c}(\mathbf{d}_0)
\end{bmatrix}
\]

(26)
Algorithm 1: Branch and Bound($d_{inc}, D^N, \hat{D}^N_{rel}$)

**Input:** incumbent $d_{inc}$, parameter domain $D^N$, relaxed domain $\hat{D}^N_{rel}$

//compute minimum for given relaxed domain
$$d^* = \arg\min_{d \in \hat{D}^N_{rel}} \varphi(d) \text{ s.t.: } c(d) \geq 0$$

1

if No feasible solution in $\hat{D}^N_{rel}$ then
  // pruning by infeasibility
  return $d_{inc}$

2

else if $\varphi(d^*) \geq \varphi(d_{inc})$ then
  // pruning by value dominance
  return $d_{inc}$

3

else if $d^* \in D^N$ and $\varphi(d^*) < \varphi(d_{inc})$ then
  // pruning by optimality
  return $d^*$

4

else
  // branching
  Choose parameter $d_i$ for branching
  $\hat{D}^N_{down} = \{ d \in \hat{D}^N_{rel} | d_i \leq \lfloor d^*_i \rfloor \}$
  $\hat{D}^N_{up} = \{ d \in \hat{D}^N_{rel} | d_i \geq \lceil d^*_i \rceil \}$

5

$d_{inc} = \text{Branch and Bound}(d_{inc}, D^N, \hat{D}^N_{down})$

6

$d_{inc} = \text{Branch and Bound}(d_{inc}, D^N, \hat{D}^N_{up})$

7

return $d_{inc}$

8

9

10

11

12

13

14

15

If the second derivative of the Lagrangian function is convex, the optimization problem in (24) can be solved by iteratively solving the equation system in (26). The result describes the direction from the current point to the minimum of the quadratic model of the objective function subject to the constraints. In the SQP approach a model of the matrix is usually built up iteratively. This can be realized by different approaches, e.g., BFGS (Broyden Fletcher Goldfarb Shanno) update formula, which is used in this approach.

After computing the direction by solving (24), a step size is computed at the original relaxed program using line search. In this approach a Wolfe Powell step size algorithm is used. The solution which is computed by SQP on the relaxed program is obviously no discrete feasible point in general. In the next section of this chapter a Branch-and-Bound method is described, which can be used to find a discrete solution for the original sizing task based on the solution of the relaxed problem.

**3.2 Branch and Bound**

Branch and Bound (e.g., (Nemhauser & Wolsey, 1988)) is one of the most popular approaches in discrete optimization. In the form which is used in this work, it decomposes the discrete optimization task in a sequence of relaxed optimization tasks which are nonlinear but can be solved in the continuous domain. A description of the recursive method is given in Algorithm 1 and in the following.

The algorithm is primarily based on two principles:
1. As the domain of discrete points is a subset of its continuous relaxation (22), the optimum of the relaxed program is better than or equal to the continuous solution. For the minimization in (11) and with (21):

\[
\left( \min_{d \in \mathbb{D}^N_{rel}} \varphi(d) \text{ s.t. } c(d) \geq 0 \right) \leq \left( \min_{d \in \mathbb{D}^N} \varphi(d) \text{ s.t. } c(d) \geq 0 \right)
\] (27)

Consequently, the minimum of the relaxed program is a lower bound for the discrete minimum.

2. Each discrete point with objective function value better than the best discrete point so far, is an upper bound for the discrete solution of the optimization task.

Initially in each recursion the relaxed optimization task is solved in the current relaxed domain \(\hat{\mathbb{D}}^N_{rel}\) (Algorithm 1, line 1 and Figure 3(a)). In the approach presented in this chapter SQP is used at this point (Section 3.1). Following the first principle, the objective function value at the minimum of the relaxed task \(\varphi(d^*)\) is smaller than or equal to the value at the best discrete solution in the current sub-domain \(\varphi(d^*_{disc})\), i.e.,

\[
\left( \varphi(d^*) = \min_{d \in \hat{\mathbb{D}}^N_{rel}} \varphi(d) \text{ s.t. } c(d) \geq 0 \right) \leq \left( \varphi(d^*_{disc}) = \min_{d \in \mathbb{D}^N \cap \hat{\mathbb{D}}^N_{rel}} \varphi(d) \text{ s.t. } c(d) \geq 0 \right)
\] (28)

Thus, even if \(d^*_{disc}\) is not explicitly known at this point, the objective function value at the optimum of the relaxed sub-problem \(\varphi(d^*)\) is a lower bound for the sub-domain.

The minimum \(d^*\) which is computed for the relaxed optimization task is not necessarily discrete. Thus, one of the parameters \(d_i \in \mathbb{D}_i\) (3) which must be discretized is chosen, and a constraint is set on it to be greater than the next higher or smaller than the next lower discrete value (called branching). In Algorithm 1 (lines 10, 11) this is assigned by \(d_i \geq \lceil d_i^* \rceil\) and \(d_i \leq \lfloor d_i^* \rfloor\), with

\[
\lceil d_i^* \rceil = \max_{d \in \mathbb{D}_i} \text{ s.t. } d < d_i \quad \text{and} \quad \lfloor d_i^* \rfloor = \min_{d \in \mathbb{D}_i} \text{ s.t. } d > d_i
\] (29)

Figure 3(b) shows that adding one pair of such constraints can be considered as building up two new relaxed sub-problems with reduced parameter domain (\(\hat{\mathbb{D}}^N_{up}, \hat{\mathbb{D}}^N_{down}\) in Algorithm 1 and \(\hat{\mathbb{D}}^2_1, \hat{\mathbb{D}}^2_2\) in the example in Figure 3(b)). Typically, this is represented by a branching tree (Figure 4): If the parent node of this tree represents the current domain, branching is equivalent to adding two child nodes which correspond to the subsets \(\hat{\mathbb{D}}^N_{up}\) and \(\hat{\mathbb{D}}^N_{down}\). The edges of the branching tree correspond to the constraints, which are added to define the sub-problems.

For each sub-problem which is set up in the branching step, Algorithm 1 is executed recursively (Algorithm 1, lines 12, 13). Following the heuristic order in Algorithm 1, always \(\hat{\mathbb{D}}^N_{down}\) is explored before considering \(\hat{\mathbb{D}}^N_{up}\). Thus, in Figure 3 \(\hat{\mathbb{D}}^2_1\) is explored before considering \(\hat{\mathbb{D}}^2_2\). In the example, after computing the continuous solution of sub-domain \(\hat{\mathbb{D}}^2_1\) the sub-problem is further branched, as the continuous solution of the sub-problem is not element of the original discrete domain (Figure 3(c)).

If for each sub-problem branching constraints are added, until the solution of the relaxed sub-problem is a discrete point and thus a leaf of the search tree is reached, the discrete solution with the best objective function value is the optimum of the discrete optimization task. However, without further modification, the computational effort of the method is
The solution of the relaxed optimization task $d^*$ is a lower bound for the solution of the discrete minimization $d_{disc}^*$ which is not explicitly known at this point of time.

Branching leads to two sub-sets e.g., $\hat{D}_N^{down} = \hat{D}_1$ and $\hat{D}_N^{up} = \hat{D}_2$, and two corresponding sub-problems. The parameter which is chosen for branching, is computed by equation (30).

Following Algorithm 1, the lower sub-domain $\hat{D}_1$ is considered next. The relaxed solution of the optimization task in sub-domain $\hat{D}_1$ is non-discrete in parameter $d_2$. Thus branching constraints are added for this parameter and the sub-domains $\hat{D}_N^{down} = \hat{D}_3$, $\hat{D}_N^{up} = \hat{D}_4$ are generated.

Domain $\hat{D}_4$ is considered next. It does not include any feasible point and can be pruned by infeasibility.

Finally sub-domain $\hat{D}_2$ is explored. The continuous optimum in this region has a higher value than $d_{inc}$ from Figure 3(d). It can be pruned by value dominance.

Fig. 3. Illustration of Branch and Bound for a two-dimensional optimization task.
Fig. 4. Illustration of Branch and Bound from Figure 3 as a branching tree.

extremely high, as many non promising sub-problems must be solved. Thus pruning rules (e.g., (Nemhauser & Wolsey, 1988)) are introduced to reduce the run time of the algorithm. The pruning step can be understood as cutting non promising nodes from the search tree. Three pruning rules can be found:

1. If a relaxed sub-problem does not include any feasible solution, the corresponding node can be cut from the search tree. For this rule (known as pruning by infeasibility) it is considered that each discrete domain is subset of its relaxation (Algorithm 1, line 2, 3, and Figure 3(e)).

2. If a discrete solution $d_{inc}$ has been found which is smaller than the value of the current relaxed sub-problem, the node corresponding to the relaxed sub-problem can be cut off. This pruning by value dominance uses that – due to principle 1 from above – the relaxed program can not include a discrete point which is better than $d_{inc}$ (Algorithm 1, line 4, 5, and Figure 3(f)).

3. If the solution of the relaxed program is discrete and better than the best solution found so far, it is a new best solution ($d_{inc}$) for the discrete sizing task. However, at the same time it is a lower bound for the corresponding relaxed sub-problem which can not include any better point. Thus, no further branching is necessary in the sub-region. This is called pruning by optimality (Algorithm 1, line 6, 7, and Figure 3(d)).

The recursive approach described so far realizes a "Depth-First" search. For branching always the most fractional parameter is used (most fractional or most infeasible branching, e.g., Achtenberg et al. (2005)), i.e., assuming an index $i = 1, ..., N_d$ for the discrete parameters $d_i$ with value $d_i^*$, the branching index $i$ is chosen by

$$i = \arg \max_{i=1, ..., N_d} \left| 0.5 - \frac{d_i^* - \lfloor d_i^* \rfloor}{\lfloor d_i^* \rfloor - \lfloor d_i^* \rfloor} \right|$$

However, some problem specific properties can be used to speed up the algorithm in case of analog sizing. This is described in the following sub-section.
Algorithm 2: Modified Branch and Bound($d_{inc}, D^N, \hat{D}^N_{rel}$)

Input: incumbent $d_{inc}$, parameter domain $D^N$, relaxed domain $\hat{D}^N_{rel}$

1. if $d_{inc}$ solves the sizing task then
   // stop due to pruning rule 1'
   return $d_{inc}$
2. end
3. Run SQP (Section 3.1) on optimization problem
   \[
   \min_{d \in \hat{D}^N_{rel}} \varphi(d) \text{ s.t.: } c(d \geq 0)
   \]
   until the sizing task is solved or an optimum is found.
4. Get solution $d^*$ and quadratic model from SQP.
5. if $d^*$ does not solve the sizing task (i.e., no solution in $\hat{D}^N_{rel}$) then
   // pruning rule 1'
   return $d_{inc}$
6. else if $d^* \in D^N$ (i.e., $d^*$ solves the problem) then
   // pruning rule 2'
   return $d^*$
7. else
   // compute incumbent
   Get $d_{model}^*$ by solving the discrete quadratic optimization task (31) using Algorithm 1
8. if $d_{model}^*$ solves the sizing task then
   // stop
   return $d_{model}^*$
9. end
10. // branching
   Choose parameter $d_i$ for branching
   Compute $\lfloor d_i^* \rfloor$ and $\lceil d_i^* \rceil$ according to (34) and (35)
   Set $\hat{D}^N_{down} = \{ d \in \hat{D}^N_{rel} | d_i \leq \lfloor d_i^* \rfloor \}$
   Set $\hat{D}^N_{up} = \{ d \in \hat{D}^N_{rel} | d_i \geq \lceil d_i^* \rceil \}$
11. $d_{inc} = \text{Modified Branch and Bound}(d_{inc}, D^N, \hat{D}^N_{down})$
12. $d_{inc} = \text{Modified Branch and Bound}(d_{inc}, D^N, \hat{D}^N_{up})$
13. return $d_{inc}$
14. end

3.3 Modification of Branch and Bound

To reduce the computational effort of Branch and Bound, the most promising way is to improve the pruning and the branching heuristic. Certain properties of the underlying optimization problem can be used to speed up the process. The modified approach is shown in Algorithm 2 and explained in the following.

3.3.1 Consideration of the quadratic model

As the continuous solution of the relaxed program is computed by an SQP approach, beside the continuous solution a quadratic model of the objective function is computed during
Fig. 5. The quadratic model (right) and the continuous solution of the optimization task \( \mathbf{d}^* \) are computed for the original program (left) during SQP. Solving the quadratic model in the discrete domain gives a discrete solution \( \mathbf{d}^*_{\text{model}} \) which is an upper bound for the original program. In this example, \( \mathbf{d}^*_{\text{model}} \) is equal to the discrete optimum of the original task.

Solving the relaxed program. Furthermore, a linear model of the constraints is computed (Algorithm 2, lines 4, 5). As these models are good local approximations for the relaxed program, they are also a good local approximation for the discrete approach. Thus a quadratic optimization task with linear constraints can be formulated as a surrogate optimization task for (11):

\[
\min_{\mathbf{d} \in \hat{\mathcal{D}}} \frac{1}{2} \mathbf{d}^T \mathbf{H} \mathbf{d} + \mathbf{g}^T \mathbf{d} \quad \text{s.t.} \quad \mathbf{J} \mathbf{d} + c(\mathbf{d}_0) \geq 0 \tag{31}
\]

wherein \( \mathbf{H} \) and \( \mathbf{g} \) are the Hessian matrix and the gradient for the objective function at the solution point of the relaxed program, \( \mathbf{J} \) is the Jacobian matrix for the constraints and \( c_0 \) are the constraint values at this point. \( \hat{\mathcal{D}}^N \) is the set of discrete points \( \mathcal{D} \) in a relaxed sub-problem, i.e.,

\[
\hat{\mathcal{D}}^N = \mathcal{D}^N \cap \hat{\mathcal{D}}^N_{\text{up}} \quad \text{or} \quad \hat{\mathcal{D}}^N = \mathcal{D}^N \cap \hat{\mathcal{D}}^N_{\text{down}} \tag{32}
\]

By solving the program in (31) using the discrete domain of the original task, the discrete optimum \( \mathbf{d}^*_{\text{model}} \) for the approximation of the objective function can be found (Figure 5; Algorithm 2, line 11). Due to the second principle from Section 3.2, the value of \( \mathbf{d}^*_{\text{model}} \) in the original objective function – i.e., \( \varphi(\mathbf{d}^*_{\text{model}}) \) – is an upper bound for the discrete optimum if it is feasible for the original task. Consequently, sub-regions with a continuous solution worse than the discrete optimum of the model can be cut from the search tree. Thus, early pruning by pruning rule 2 is possible, as – in contrast to standard branch and bound – a discrete upper bound exists in the first branching node and not after discretizing all parameters, i.e., in the first leaf of the search tree. This fact is especially important if many discrete parameters exist. Additionally, solving the quadratic surrogate problem is computational much less expensive, as no circuit simulations are necessary, which cause the highest time consumption in solving the sizing task. Thus, the Branch and Bound algorithm from Section 3.2 can be used to solve the discrete quadratic program with linear constraints in (31).

3.3.2 Consideration of non-optimality

For analog sizing the SQP approach is stopped as soon as any point is found which fulfills specifications and constraints. Thus, the solution which is found is in general non-optimal in terms of the objective function. Taking into account that it is a binary decision if a certain point solves the sizing task or not, the branching rules can be reformulated.
Fig. 6. SQP can be stopped as soon as a continuous point $d^*$ is found which fulfills constraints and specifications (left). At this point the quadratic model (right) is set up. In the example, the objective function value at the discrete optimum of the quadratic model $d_{\text{model}}^*$ is better than the value at $d^*$ and specifications and constraints are fulfilled at $d_{\text{model}}^*$.

Assuming that there is at least one discrete solution for the sizing task, obviously only these sub-domains must be considered during Branch and Bound which include such a point. As – due to (22) – the discrete solutions must be also in the relaxed domain, all sub-domains can be cut which do not include a solution of the sizing task in their relaxation. This can be considered by reformulating pruning rule 1 as:

1'. If a relaxed sub-problem does not include any solution for the sizing task, the corresponding node can be cut from the search tree (Algorithm 2, lines 6, 7).

If pruning rule 1 is replaced by 1' the discrete point $d_{\text{inc}}$ – which represents a solution candidate – is only set up, if a discrete solution is found. The Branch and Bound algorithm can be stopped in this case. Thus, pruning rule 2 (pruning by value dominance) becomes redundant and can be left out. Pruning rule 1 is reformulated as a stop criterion, to set up the discrete solution correctly and to avoid insufficient computational effort when the discrete solution has been found:

3'. If any discrete solution for the sizing task has been found, no further branching is required (Algorithm 2, lines 1, 2 and 8, 9).

The modifications of the pruning rules have an even stronger influence if the quadratic model from Section 3.3.1 is considered. In this case, the quadratic model is set up once again in the point $d^*$ which is computed by SQP and solves the sizing task in the relaxed domain. The point $d^*$ can be non-optimal in terms of the objective function and thus in may cases the solution of the quadratic optimization problem in the relaxed domain is also a better solution for the underlying sizing task. The continuous solution of the quadratic model is of course not evaluated by simulation. However, as the quadratic model is set up once again at $d^*$, it is a locally better approximation of the objective function than the quadratic model used for the last SQP step. Thus, even the discrete optimum $d_{\text{model}}^*$ of the quadratic problem in (31) computed by use of the quadratic model at $d^*$ is often a better solution for the sizing task than $d^*$ itself (Figure 6).

Hence, in many cases the discrete solution of the model solves the sizing task in the initial node of Branch and Bound and Branch and Bound can be stopped after computing the discrete solution of the quadratic model (Algorithm 2, lines 12, 13).

If the initial solution of the quadratic model does not fulfill the specifications, the
non-optimality of the SQP solution \( d^* \) can also be used to improve the branching heuristic which has significant influence on the runtime of standard Branch and Bound. The gradient at a non-optimal point \( d^* \) is not equal to zero. Thus, it can be assumed that discrete solution candidates can be found in direction of degression of the objective function. The gradient \( g \) at the solution of the SQP algorithm has been already computed to improve the quadratic model and comes without additional cost. For the branching heuristic used in this approach, now the parameter which should be discretized and which corresponds to the gradient component \( g_i \) with the strongest influence to the objective function is discretized first (Algorithm 2, line 15).

In the "Depth First" search, then the sub-region is chosen which lies in direction of greatest improvement (Algorithm 2, line 16), i.e., assuming the next discrete values for the parameter \( d_i \) in domain \( D_i \) from (3) are \( d_a \) and \( d_b \), with

\[
d_a = \max_{d \in D_i} \text{s.t. } d < d_i \quad \text{and} \quad d_b = \min_{d \in D_i} \text{s.t. } d > d_i
\]

the rounding operator \( \lceil \cdot \rceil \) and \( \lfloor \cdot \rfloor \) in Algorithm 1 is modified such that

\[
[d_i] = \begin{cases} 
  d_a & \text{if } g_i > 0 \\
  d_b & \text{if } g_i \leq 0
\end{cases}
\]

and, respectively,

\[
[d_i] = \begin{cases} 
  d_a & \text{if } g_i < 0 \\
  d_b & \text{if } g_i \geq 0
\end{cases}
\]

Thus, discrete points in gradient direction are considered first during branch and bound.

### 4. Experimental results

To show the effectiveness and efficacy of the algorithm, the sizing process of three different circuits will be presented in this section. For each example, the results and the runtime of SQP with subsequent rounding, of SQP and modified Branch and Bound (BaB) without quadratic model, and of SQP and modified Branch and Bound considering the quadratic model (Section 3.3) is presented. The modified Branch and Bound algorithm considering the quadratic model is presented in Section 3.3 Algorithm 2. The modified Branch and Bound algorithm without the quadratic model is implemented identically, but the consideration of the quadratic model (lines 11 - 14 in Algorithm 2) is switched off. I.e., both Branch and Bound approaches stop as soon as a discrete solution for the sizing task is found. Branching in both Branch and Bound algorithms is realized according to (34) and (35).

The circuit in the first example is the Miller amplifier in Figure 7. For the sizing tasks the lengths, widths, and multipliers of the transistors are used as discrete parameters. The lengths of all transistors shall be equal. Furthermore some multipliers and transistor widths (e.g., multipliers and widths of the differential pair) are set equal to avoid mismatch effects. For transistor lengths and widths a 5\( \text{nm} \) manufacturing grid is assumed. The Miller capacitance is represented by a continuous parameter. A 0.5\( \text{pF} \) load capacitance and a 2\( \text{V} \) supply voltage are given for the circuit and the 45\( \text{nm} \) low power predictive technology (PTM; (Balijepalli et al., 2007; Cao et al., 2000; Zhao & Cao, 2006)) from (Nanoscale Integration and Modelling Group, Arizona State University, 2008) is used.

The simulated performance values of the amplifier before and after sizing are shown in Table 1. It can be seen from the results, that – as proposed in Section 1 – the continuous optimization
and subsequent rounding violates two specifications in this case. In contrast, the goal of the discrete sizing task was achieved if Branch and Bound with or without quadratic model has been used. The result quality of Branch and Bound with quadratic model is as good as the result quality achieved without the modification. However, the runtime comparison in Figure 8 clearly shows that the additional runtime for Branch and Bound considering the quadratic model presented in this paper, is significantly smaller, than without the modification and the additional cost compared to the optimization with subsequent rounding is neglectable in this case.

In the second example the sizing of the more complex amplifier in Figure 9, which is proposed in (Martins, 1998), is shown. For this example the 45nm high performance predictive technology model from (Nanoscale Integration and Modelling Group, Arizona State University, 2008) is used and again a 5nm manufacturing grid is assumed. The lengths of all transistors and the widths of transistors which are in the same current mirror or in the same differential pair are set equal. Additionally, some multipliers are set equal considering the symmetries of the circuit. Thus, 14 multipliers, 11 widths, and the length are considered as discrete parameters. Additionally, the compensation capacitance \(C_c\) and the bias voltages \(V_{bias,1}\) and \(V_{bias,2}\) are represented by continuous parameters. A 20\(pF\) load capacitance and a 2\(V\) supply voltage are given for the circuit. Again the sizing rules from (Massier & Graeb, 2008) are used which define 93 constraints in this case. Specifications and simulated performances
<table>
<thead>
<tr>
<th>Performance</th>
<th>Specification</th>
<th>Initial values</th>
<th>SQP + Rounding</th>
<th>SQP + Bab w/o quadratic model</th>
<th>SQP + Bab with quadratic model</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSRR [dB]</td>
<td>&gt; 135</td>
<td>134</td>
<td>138</td>
<td>139</td>
<td>137</td>
</tr>
<tr>
<td>Gain [dB]</td>
<td>&gt; 85</td>
<td>89</td>
<td>89</td>
<td>90</td>
<td>89</td>
</tr>
<tr>
<td>CMRR [dB]</td>
<td>&gt; 135</td>
<td>172</td>
<td>167</td>
<td>167</td>
<td>166</td>
</tr>
<tr>
<td>$f_{\text{transit}}$ [MHz]</td>
<td>&gt; 15</td>
<td>19</td>
<td>24</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>$\varphi$ [°]</td>
<td>&gt; 60</td>
<td>50</td>
<td>59 (violates spec)</td>
<td>61</td>
<td>60</td>
</tr>
<tr>
<td>SR (rising) [Volts]</td>
<td>&gt; 15</td>
<td>10</td>
<td>18</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>$</td>
<td>\text{SR (falling)}</td>
<td>$ [Volts]</td>
<td>&gt; 15</td>
<td>14</td>
<td>37</td>
</tr>
<tr>
<td>Area [(\mu m)^2]</td>
<td>&lt; 10</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Power [\mu W]</td>
<td>&lt; 50</td>
<td>56</td>
<td>52 (violates spec)</td>
<td>49</td>
<td>49</td>
</tr>
</tbody>
</table>

Table 1. Specification and performance values for Miller amplifier using 45nm PTM, 2V supply voltage, 1\mu A bias current

Fig. 9. Low-voltage low-power operational amplifier from (Martins, 1998)
Fig. 10. Runtime for up to 8 times parallelized algorithm on a 16 core 2.67GHz computer for sizing of amplifier in Figure 9 are presented in Table 2.

Also in this case specifications are violated when SQP and sub-sequent rounding is used and also here Branch and Bound with and without quadratic model solves the sizing task. The runtime comparison in this case shows that also here Branch and Bound using the quadratic model is much faster than without consideration of the quadratic model. As the number of discrete parameters is much higher in this case, also the runtime of Branch and Bound on the quadratic model is relatively large. Thus potential for further improvement of the algorithm can be seen: The runtime of the algorithm can be reduced if the Branch and Bound algorithm presented in (3.2) is advanced, which is used to find the discrete optimum on the quadratic model and needs approximately half of the computational time in this experiment.

The third example shows the sizing process for the sense amplifier from (Yeung & Mahmoodi, 2006) (see Figure 11). Considering the symmetry of the circuit, 5 multipliers, 5 transistor widths, and the transistor length are used as parameters. For the sizing process a 16nm low power PTM is used and a 2nm manufacturing grid is assumed. Specifications and results are listed in Table 3. For the simulation of the delay it is assumed that the inputs (bit line BL and negative bit line BLB) are preloaded to $V_{DD} = 1.5V$ and the input signal is a voltage reduction by 10mV at one of them. “Delay +” in Table 3 is defined as the time between the change of the input signal at the positive input BL and the point of time when the positive output reaches $0.95 \cdot V_{DD}$. Accordingly, the value of “Delay −” is defined as the time between the change of the signal at the negative input BLB and the point of time when the positive output reaches $0.05 \cdot V_{DD}$.

The results for this experiment show that in this case continuous optimization with subsequent rounding leads to a solution of the sizing task. This especially happens, if only a few or week constraints and specifications are defined and if only a small number of parameters is used. However, the additional runtime for the modified Branch and Bound approach is only a few seconds. Further analysis of the results shows, that the additional
Table 2. Specification and performance values for the amplifier in Figure 9 using 45nm PTM, 2V supply voltage, 20pF load capacity

<table>
<thead>
<tr>
<th>Performance</th>
<th>Specification</th>
<th>Initial values</th>
<th>SQP + Rounding</th>
<th>SQP + BaB w/o quadratic model</th>
<th>SQP + BaB with quadratic model</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSRR [dB]</td>
<td>&gt; 70</td>
<td>59</td>
<td>63 (violates spec)</td>
<td>77</td>
<td>79</td>
</tr>
<tr>
<td>Gain [dB]</td>
<td>&gt; 60</td>
<td>62</td>
<td>65</td>
<td>69</td>
<td>65</td>
</tr>
<tr>
<td>CMRR [dB]</td>
<td>&gt; 70</td>
<td>57</td>
<td>69.7 (violates spec)</td>
<td>81</td>
<td>83</td>
</tr>
<tr>
<td>$f_{\text{transit}}$ [MHz]</td>
<td>&gt; 150</td>
<td>92</td>
<td>274</td>
<td>256</td>
<td>221</td>
</tr>
<tr>
<td>$\varphi$ [°]</td>
<td>&gt; 60</td>
<td>85</td>
<td>65</td>
<td>62</td>
<td>60</td>
</tr>
<tr>
<td>SR (rising) $\frac{V}{V}$</td>
<td>&gt; 10</td>
<td>8</td>
<td>23</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>$</td>
<td>\text{SR (falling)}</td>
<td>$</td>
<td>&gt; 10</td>
<td>8</td>
<td>17</td>
</tr>
</tbody>
</table>

Fig. 11. Sense amplifier

runtime (approximately 30 seconds) is used for computing the gradient and setting up the quadratic model. In contrast to Branch and Bound with consideration of the quadratic model, Branch and Bound without the quadratic model has a significant higher runtime.

5. Conclusion

Sizing of analog circuits is one important task in the analog design flow. In this chapter a new deterministic and gradient-based method has been presented to solve this task. The method solves the relaxed, i.e., continuous sizing task using SQP. Discretization of the result is done by a subsequent Branch and Bound approach under consideration of the quadratic model which is computed during SQP. Additionally certain properties of the underlying sizing task are used to speed up the approach.

The experimental results show that SQP with subsequent rounding can not solve the sizing task in general. In contrast, SQP combined with Branch and Bound is a reasonable approach for sizing analog circuits with discrete parameters. Furthermore, the experimental results show, that the efficacy and efficiency of SQP and Branch and Bound can be increased...
Fig. 12. Runtime for up to 8 times parallelized algorithm on a 16 core 2.67GHz computer for sizing of the sense amplifier.

![Runtime chart]

Table 3. Specification and performance values for sense amplifier using 16nm PTM, 1.5V supply voltage, 1fF Load capacity.

<table>
<thead>
<tr>
<th>Performance</th>
<th>Specification</th>
<th>Initial values</th>
<th>SQP + Rounding</th>
<th>SQP + BaB w/o quadratic model</th>
<th>SQP + BaB with quadratic model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay $\frac{[\text{ps}]}{}$</td>
<td>&lt; 60</td>
<td>402</td>
<td>59</td>
<td>46</td>
<td>55</td>
</tr>
<tr>
<td>Delay $\frac{[\text{ps}]}{}$</td>
<td>&lt; 60</td>
<td>423</td>
<td>48</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td>Static power $\frac{[\mu\text{W}]}{}$</td>
<td>&lt; 5</td>
<td>27.5</td>
<td>1.4</td>
<td>1.4</td>
<td>2.0</td>
</tr>
<tr>
<td>Area $\frac{[\mu\text{m}^2]}{}$</td>
<td>&lt; 0.025</td>
<td>0.225</td>
<td>0.018</td>
<td>0.022</td>
<td>0.019</td>
</tr>
</tbody>
</table>

significantly, if the modifications in Section 3.3 are used.

The task presented so far is not able to solve the discrete sizing task, if the circuit performances can only be evaluated for discrete points. Thus, in the future work this problem will be tackled. Additionally, the experiments have shown, that the runtime of the algorithm can be reduced by accelerating the Branch and Bound approach which is used to solve the quadratic model. Also the consideration of non-scalable discrete parameters mentioned in Section 2 is an open task for the future work.

6. References


www.intechopen.com


This book highlights key design issues and challenges to guarantee the development of successful applications of analog circuits. Researchers around the world share acquired experience and insights to develop advances in analog circuit design, modeling and simulation. The key contributions of the sixteen chapters focus on recent advances in analog circuits to accomplish academic or industrial target specifications.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: