Comparison among Evolutionary Algorithms and Classical Optimization Methods for Circuit Design Problems

Angelo Marcello Anile
Department of Mathematics and Computer Science
University of Catania
V.le A. Doria 6, 95125 Catania, Italy
anile@dmi.unict.it

Vincenzo Cutello
Department of Mathematics and Computer Science
University of Catania
V.le A. Doria 6, 95125 Catania, Italy
cutello@dmi.unict.it

Giuseppe Nicosia
Department of Mathematics and Computer Science
University of Catania
V.le A. Doria 6, 95125 Catania, Italy
nicosia@dmi.unict.it

Rosario Rascunà
Department of Mathematics and Computer Science
University of Catania
V.le A. Doria 6, 95125 Catania, Italy
rascuna@dmi.unict.it

Salvatore Spinella
University of Calabria
Ponte P. Bucci 17/B
87036 Arcavacata di Rende, Italy
spins@unical.it

Abstract- This work concerns the comparison of evolutionary algorithms and standard optimization methods on two circuit design problems: the parameter extraction of device circuit model and the multi-objective optimization of an Operational Transconductance Amplifier. We compare standard optimization techniques and evolutionary algorithms in terms of quality of the solutions and computational effort, that is, objective function evaluations needed to compute them. The experimental results obtained show as standard techniques are robust with respect evolutionary algorithms, while the latter are more effective in terms of the standard metrics and function calls. In particular for the multiobjective problem, the observed Pareto front determined by evolutionary algorithms has a better spread of solutions with a larger number of nondominated solutions with respect to the standard multi-objective techniques.

1 Introduction

In this research work we compare experimentally the effectiveness of evolutionary algorithms and standard optimization techniques. As test bed we used two *Circuit Design Problems*: parameter extraction of device circuit model and optimization of a Operational Transconductance Amplifier. The first problem faced is a single-objective optimization problem, the latter is a multi-objective optimization problem. Both problems are real world applications which have been supplied by STMICROELECTRONICS.

Standard methods (or classical methods) are characterized by an analytical condition or gradient-based (e.g., *non linear least squared method*), nor approximation-based (e.g., *direct and statistical methods*). All classical methods find an individual solution in a single run.

Evolutionary algorithms are stochastic optimization methods population-based inspired by natural selection able to find several solutions in a single run, thus making them a good alternative to standard methods.

To assess the quality of experimental results has been used a class of metrics to characterize the various approaches and to make proper comparisons.

In section 2, we describe respectively the single-objective optimization problem concerning the *parameter extraction of a Inductor device circuit model*, the standard and evolutionary algorithms and metrics used, and the experimental results obtained.

In section 3, we present the design of an analog circuit, the *Operational Transconductance Amplifier*, modeled as multi-objective optimization problem. Analogously to the Inductor device circuit, in this section we describe the algorithms and metrics used, and report comparisons among the various algorithms.

Finally, section 4 concludes the article with some general remarks.

2 The inductor circuit

2.1 The problem of parameter extraction

Inductor devices hold a fundamental role in the radiofrequency field and it is important to develop models which represent the intrinsic characteristics correctly at every working frequencies.

The best values for the model parameters are found by fitting the measured data as closely as possible to the simulated data in the sense of a suitable weighted l^2 metric and this process is usually performed as a sequence of optimizations, usually based on the Levenberg-Marquard algorithm, which require a good initial guess and yield only local minima (corresponding to different set of parameters). In this context two problems arise. First, get robust a estimation for parameters when there are several measurement curves (e.g. different components of the Y matrix for small signal analysis).This seems to be the case when fitting compact models to MOS devices DC measurements [Keser00]. Second, how to choose the most convenient set of parameter values to obtain the best approximation for the circuit model. Microelectronics industries develop many circuit model to predict the behavior of this device. These models have as target a circuit simulator as SPICE and

their parameters are determined by a suitable optimization algorithm. The inductor has the following structure:

Shape	**Number of turns**	**Outer dim. (μm)**
Octag.	2.5	200.0
	Width (μm)	**Spacing (μm)**
	16.0	8.0
	SiO_2 Thick. (μm)	**Al Thick. (μm)**
	1.8	3.0

Although a circuit network may have any number of ports, network parameters can be explained most easily by considering a network with only two ports, an input port and an output port. To characterize the performance of such a network, any of several parameter sets can be used, each of which has certain advantages. Each parameter set is related to a set of four variables associated with the two-port model. Two of these variables represent the excitation of the network (independent variables), and the remaining two represent the response of the network to the excitation (dependent variables). If the network is excited by voltage sources V_1 and V_2, the network currents I_1 and I_2 will be related by the following equations (assuming the network behaves linearly):

$$\begin{aligned} I_1 &= y_{1,1}V_1 + y_{1,2}V_2 \\ I_2 &= y_{2,1}V_1 + y_{2,2}V_2 \end{aligned} \tag{1}$$

In this case, with port voltages selected as *independent variables* and port currents taken as *dependent variables*, the relating parameters are called short-circuit admittance parameters, or *y-parameters*. In the absence of additional information, four measurements are required to determine the four parameters $y_{1,1}$, $y_{1,2}$, $y_{2,1}$, $y_{2,2}$. Each measurement is made with one port of the network excited by a voltage source while the other port is short circuited. For example, $y_{2,1}$, the forward transadmittance, is the ratio of the current at port 2 to the voltage at port 1 with port 2 short circuited, as shown in equation 3.

$$y_{1,2} = \left.\frac{I_2}{V_1}\right|_{V_2=0 \ \text{(ouput short circuited)}}$$

Preliminary investigations have been carried out with commercial tools and in-house (STMICROELECTRONICS) optimization software and they have yielded different sets of parameters. These results are summarized in table 1. First column shows the initial estimation computed by the TMG fitting [Rinaudo98]. This fitting is based on a initial estimation of the parameter.

Comparison of previous data (table 1) shows large variance of parameter extraction. Causes of these behaviours could be non-homogeneous kind of variables; variables can converge with different speed rates; merit function of optimization on l^2 can find different balancing among errors; first and second derivative do not lead the optimization in useful regions. Previous remarks compel to consider the quality of results is sense of robustness.

Table 1: Preliminary results of parameter extraction.

	TMG init. estim.	ADS	Opsim
RL	0.1548	1.47	0.21
LL (10^{-9})	0.21	1.79	0.25
Rox1	20.41	10.27	4.60
Cox1 (10^{-12})	5.19	0.78	0.05
Rox2	17.67	8.87	0.0
Cox2 (10^{-12})	5.37	0.27	0.04
Rb1	16.49	0.56	6.57
Cb1 (10^{-12})	0.002	2.33	1.77
Rb2	0.6	0.58	0.03
Cb2 (10^{-12})	0.51	0.25	0.07
K1	0.6	0.3	0.3
K2	0.87	1.44	1.42

2.2 Uncertainty, Robustness, Confidence Limits

The concept of *uncertainty* summarizes various problem related to degree of model approximation, imprecisions on performing calculations, statistical representation of data.

A general practice is to include all sources of uncertainty into the statistical representation of data and evaluate the robustness of solution in terms of confident limits.

The term "robust" was coined in statistics by G.E.P. Box in 1953. General, referring to a parameter extraction for fitting a statistical model of data, it means "insensitive to small departures" from the idealized assumptions for which the data model is optimized. The word "small" can have two different interpretations, both important: either fractionally small departures for all data points, or else fractionally large departures for a small number of data points. It is the latter interpretation, leading to the notion of outlier points, that is generally the most stressful for statistical procedures.

In this work we used the M-estimate obtained by minimizing the mean absolute deviation, rather than the mean square deviation.

$$\min \sum_i |y_i - y(x_i)| \tag{2}$$

Here the tails of the distribution, although exponentially decreasing, are asymptotically much larger than any corresponding Gaussian.

$$Prob\{y_i - y(x_i)\} \sim e^{\frac{|y_i - y(x_i)|}{\sigma_i}} \tag{3}$$

This choice deals out the outlier points to get the requested robustness.

The comparisons in this work want to be more explicit regarding the precise meaning of these quantitative uncertainties, and to give further information about how quantitative confidence limits on fitted parameters can be estimated. Through the montecarlo simulation it is possible to repeat virtually an experiment and to get a quality measure of fitting robustness. The simulation starts with a initial fitting in order to identify a possible set of parameter $\widetilde{\mathbf{x}}$. This set of parameter is used to synthesize a new surrogated set of data $\mathcal{D}_{\widetilde{\mathbf{x}}}$ which are perturbed by a white noise. In this study the noise is a gaussian error with $\mu = 0$ and $\sigma = \frac{1}{10}$ of data

magnitude. This process mimes artificially the statistical properties of real data. Then the fitting is processed on this surrogated data to get a new set of parameters. This kind of artificial process is repeated many times to get a large class of parameter. Finally, classical statistics are performed on this class of parameter set and confidence limit on parameters are calculated from these simulations.

2.3 Metrics

The set of data is a sequence of complex matrix representing the y-parameter of the circuit network:

$$Y^f = \begin{pmatrix} y^f_{1,1} & y^f_{1,2} \\ y^f_{2,1} & y^f_{2,2} \end{pmatrix} \tag{4}$$

where the frequency f cover the range from 100 Mhz to 20 Ghz. The behavior of circuit is represented by a complex matrix function which has the frequency f as variable and $\mathbf{x}$ as parameters.

$$Y(f,\mathbf{x}) = \begin{pmatrix} y_{1,1}(f,\mathbf{x}) & y_{1,2}(f,\mathbf{x}) \\ y_{2,1}(f,\mathbf{x}) & y_{2,2}(f,\mathbf{x}) \end{pmatrix} \tag{5}$$

In our work we used two metrics. The first one is related to the robust estimation of parameter and compute *Absolute Error Deviation*

$$AED(\mathbf{x}) = \int \left\{ \sum_{i=1}^{2}\sum_{j=1}^{2} |real(y_{i,j}(f,\mathbf{x})) - real(y^f_{i,j})| \right\} df$$

$$+ \int \left\{ \sum_{i=1}^{2}\sum_{j=1}^{2} |imag(y_{i,j}(f,\mathbf{x})) - imag(y^f_{i,j})| \right\} df \tag{6}$$

A second class of comparison was performed minimizing the maximum absolute deviation

$$\min\{\max_{i,j} |real(y_{i,j}(f,\mathbf{x})) - real(y^f_{i,j})|, \\ \max_{i,j} |imag(y_{i,j}(f,\mathbf{x})) - imag(y^f_{i,j})|\} \tag{7}$$

This approach perform the best approximating parameter for the problem. The solution are different compared with the previous because the discrepancy of curve behaviors are considered rather then the deviation point by point.

2.4 Methodologies

2.4.1 The classical algorithms

In order to compare the robustness of the sets of parameters and also the computational effort the following four methods have been considered.

LSQ. The function LSQNONLIN of MATLAB (LSQ) with the default option of large scale optimization, which uses the subspace trust method based on the interior-reflective Newton method. The structure of the nonlinear least-squares problem $f(x) = \frac{1}{2}\|F(x)\|_2^2$ is exploited to enhance efficiency. In particular, an approximate Gauss-Newton direction, i.e., a solution s to $\min \|Js+F\|$ (where J is the Jacobian of F) is used to help define the subspace S. Second derivatives of the component function are not used. This is a sophisticated routine and similar algorithms are implemented in the commercial simulators used in the microelectronics industry.

DIRECT. The *DIRECT* method (DIR), which is a global search method described in [Jones] and applies to Lipschitz continuous functions and, after an initial implicit estimate of the Lipschitz constant chooses the potentially optimal rectangles and resamples them along the their axis. Afterward it divides these rectangles and proceed by sampling and dividing until a stop criterion is met. This method exploits the estimation of Lipschitz constant to balance global and local search and reaches quasi-global solution in large domain.

Snobfit. The *Snobfit* method (Snobfit) [Neumaier] produces a set of point for each step through successive divisions of the search region (branch) and builds local quadratic models. It combines local search using the best points with trust regions method and global search exploring regions not yet sampled.

Hybrid approach. The heuristic (HYB) combining few initial steps of DIRECT in order to obtain a reasonable initial guess and subsequently FGOALATTAIN. The first step detects a suitable region to start fgoalattain method, which can uses this initial information to set up good constraint system.

2.4.2 Evolutionary Algorithms

We used three evolutionary algorithms, SGA (Simple Genetic Algorithm), DE (Differential Evolution) and eGA (experimental Genetic Algorithm).

SGA. SGA use a real coded string to represents the variables. Simulated Binary Crossover (SBX), and polynomial mutation were used to create offsprings, a tournament selection with elitism replace the old population with the new one.

DE. Differential Evolution (DE) was introduced by Storn and Price [Price]. DE works as follow: after a random initialization the objective function is evaluated and the following steps are repeated until a termination condition is satisfied. Each individual is updated using a weighted difference of a number of selected parent solutions. If the offspring replaces the parent only if it improves the fitness value, otherwise the parent is copied in the new population. Using Storn and Price naming convention we used *DE/rand-to-best/1/bin*.

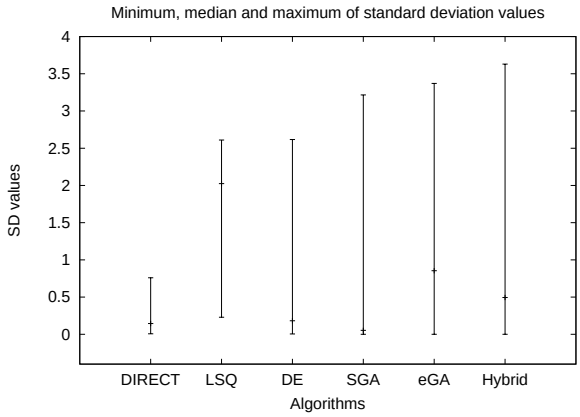

Figure 1: This figure shows the comparison of robustness among the methods on circuit parameters.

Table 3: Comparison on Absolute Error Deviation among Evolutionary Algorithms and Classical Methods.

Method	Best	mean(σ)	Avg. Pop.
eGA	**1.936982**	1.997899(0.063871)	**1.982449**
Snobfit	1.940544	n.a.	n.a.
DE	1.960735	**1.985943(0.030405)**	2.013328
Direct	2.006814	n.a.	n.a.
SGA	2.021452	2.049886(0.022182)	2.284572

eGA. eGA works the same way it does DE, but instead of using differences, it uses crossover. After performing crossover between two individual P_i and P_j, both offsprings O_1 and O_2 undergo crossover with the best individual found, but only one offsprings is kept per pair. The new generated offsprings are copied in the new population, then tournament selection with elitism is performed as usual. It is also present a polynomial mutation like SGA.

2.5 Results for the inductor

These results shows the best values obtained after performing 30 independent runs for each algorithms while the Montecarlo simulation used 1000 synthetic data set created adding a gaussian error with $\mu = 0$ and $\sigma = \frac{1}{10}$ of data magnitude in the initial data set. All methods for the two circuit design problems tackled use as termination condition the maximum number of objective function evaluations. In particular, for the inductor problem, the maximum number of function evaluations, T_{max}, has been fixed to 10^4.

The LSQ method has larger confidence limits in the Montecarlo simulation for the parameters estimated in Table 2 as we can see through the standard deviation of the parameters. The most robust parameter are found by DIRECT algorithm but it has an error higher than the other algorithms. Results for the Montecarlo simulation are summarized in Figure 1, showing minimum, median and maximum of parameters standard deviation.

In Table 3 the methods are compared on the minimization of the absolute error deviation in terms of best metric value reached, mean solution ($mean$) and standard deviation (σ). The best result and average fitness function value of population, averaged on 30 independent runs, are obtained by eGA; while the best mean solution is obtained by DE.

Table 4: Comparisons on minmax metric among Evolutionary Algorithms and Classical Methods.

Method	Best	mean(σ)	Avg. Pop.
DE	**7.723358e-03**	**7.726269e-03(2.73e-06)**	**7.752368e-03**
SGA	7.725642e-03	7.733349e-03(1.85e-06)	7.897869e-03
eGA	7.729536e-03	7.738457e-03(4.49e-05)	7.804943e-03
Snobfit	7.729214e-03	n.a.	n.a.
Direct	3.163156e-01	n.a.	n.a.

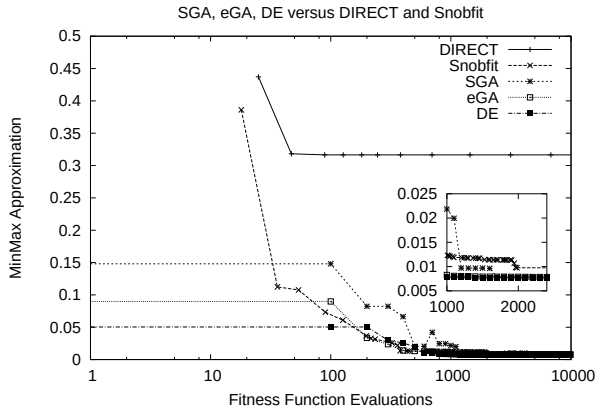

Figure 2: This figure shows the performance of the methods on minmax approximation.

The Table 4 shows a different task for the parameter extraction. In this case the maximum absolute error on the point of data set is minimized in order to determine the best approximation for the circuit model. The evolutionary algorithms and Snobfit have comparable results among them and are better than DIRECT.

In Figure 2 is shown the convergence for the best approximation and in Figure 3 the convergence on minimizing the absolute error deviation metric.

3 Operational Transconductance Amplifier

3.1 The circuit design problem

A considerably time is spent on device sizing of the analog circuit to satisfy the performance requirements. The main reason is the non-linear complex link between device size and performances [Hjam]. To improve the efficiency of the device design in the analog circuits, multiobjective approach has been proposed as alternative to the functional cost approach. The multiobjective formulation avoid the necessity of weighting different objectives in a single cost function which cannot phrase completely the analysis of the problem. In this study the optimization process is coupled to a circuit simulator (SPICE) in order to evaluate the circuit's performances [Phelps].

This case study proposes the MOS device sizing and the circuit net setting of an two-stage Operational Transconduc-

Table 2: Montecarlo simulation of the extraction parameter process. Synthetic sets has an additive Normal Error with mean 0 and standard deviation of $\frac{1}{10}$ of the range for each variables.

	Mean							STD						
	LSQ	DIR	HYB	SNO	eGA	SGA	DE	LSQ	DIR	HYB	SNO	eGA	SGA	DE
RL	4.88	1.51	1.03	1.03	1.03	1.03	1.02	2.50	0.15	0.16	0.018	0.030	0.028	0.030
LL (10^{-9})	2.41	1.53	1.72	1.25	1.22	1.22	1.22	2.04	0.09	0.73	0.005	0.012	0.011	0.011
Rox1	8.24	8.54	8.67	5.81	4.53	5.38	9.23	2.22	0.76	1.08	2.828	3.370	3.216	2.028
Cox1 (10^{-12})	0.38	0.25	0.96	0.09	0.10	0.10	0.10	0.23	0.05	0.15	0.051	0.001	0.001	0.008
Rox2	8.92	7.54	6.23	5.14	4.44	5.61	9.35	1.78	0.38	3.63	3.137	3.263	3.212	1.905
Cox2 (10^{-12})	0.42	0.06	0.003	0.08	0.09	0.09	0.09	0.25	0.004	0.02	0.025	0.002	0.001	0.005
Rb1	2	0.66	2.39	4.62	0.13	0.14	0.48	2.61	0.54	1.63	3.388	0.318	0.399	1.697
Cb1 (10^{-12})	0.76	0.92	0.26	0.48	0.01	0.01	0.18	0.24	0.04	0.26	0.315	0.081	0.033	0.310
Rb2	1.15	2.73	2.85	4.02	0.22	0.28	1.22	2.01	0.42	2.15	3.471	0.465	0.542	2.617
Cb2 (10^{-12})	5.3	0.93	0.68	0.47	0.02	0.03	0.22	0.27	0.42	0.23	0.336	0.107	0.087	0.363
K1	4.96	0.35	0.21	0.36	0.30	0.30	0.32	2.52	0.14	0.99	0.071	0.057	0.043	0.044
K2	4.93	1.44	1	1.25	1.31	1.31	1.27	2.47	0.13	0.0002	0.143	0.090	0.063	0.055

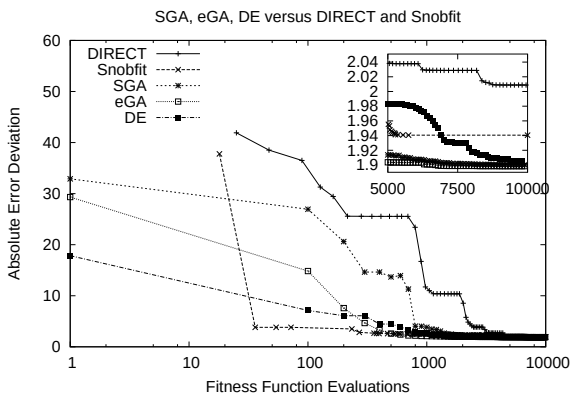

Figure 3: This figure shows the comparison of performance among the methods on minimizing the absolute error deviation.

tance Amplifier (OTA). The OTA is a useful device and it's used with few other devices to realize filters, comparators, wave generator, converters, etc.

The parameters and its ranges are showed in Table 5. The "W" parameters refers to the MOS channel width, L is referred to the MOS channel length, R (resistance) and C (capacity) are referred to the circuit net parameters. Minimum performance specifications are formulated with the constraints in Table 7. Finally, the conflicting objectives are showed in Table 6.

3.2 Objectives for the circuit performances

Many important performance metrics are to be considered in the OTA design. Those considered in this study are the following:

Low frequency gain: It is the gain at 100 Hz, that is the base of the amplification range

Unity Gain Frequency It's defined as the frequency range where the amplifier has at least the unity gain

Table 5: Parameters.

Parameters	Ranges	Unit
W1b = W1a	7 - 20	μm
W3	7 - 20	μm
W5	7 - 20	μm
L	0.525 - 0.875	μm
C	3 - 5	pF
R	20 - 40	KΩ
W4	7 - 20	μm
W2b = W2a	7 - 20	μm
I	1 - 15	μA

Table 6: Conflicting objectives for the problem.

Objectives	Specifications
Power Consumption	minimize
Total Width	minimize
Unity Gain Frequency	maximize
Gain at 100 Hz	maximize
Phase Margin	maximize

Phase Margin: The phase margin is a quality measure for the circuit because it is related to the parasitics effects, like cross coupling, which cause the failure in the attainment of the performance.

Area: The circuit's area is an important specification for the design because it is related to the yield of the manufacturing process. Generally, if it is possible increase the number of circuit per unit area than the yield of the manufacturing process increase and the ratio between number of functional failure and total number of circuit per unit area decrease. In our case we used an underestimation given by the sum of the MOS widths.

Power Consumption: the power consumption is today an important performance for all system in which the power is supplied by a battery.

Table 7: Constraints of the objective performances.

Objectives	Specifications	Unit
Unity Gain Frequency	> 31.221	Mhz
Gain at 100 Hz	> 64.118	dB
Phase Margin	> 60	Degree

3.3 MODirect

The MODirect method is an extension to the multiobjective case of the DIRECT algorithm [Jones]. The method is based on three operations: Lipschitz constant estimation, choice for potential optimality of domain subregions, domain subdivision. The choice for potential optimality is based on the estimation of Lipschitz constant for the objective function in a partition of the domain. This partition is build by hyperrectangles which are sampled in their centers in order to evaluate the value of the objective function. Therefore the estimation of Lipschitz constant leads to a possible choice of the hyperrectangles in the partition for a further sampling. In the main loop of the algorithm hyperrectangles are selected for sampling if they have a large area, an high Lipschitz constant estimation, and a good value of the function in their center. Formally it is possible to give the following definition for the single objective problem in one variable:

Definition 3.1 ***[Potential optimality relative to the objective i]*** *Let $\mathcal{S}$ be the set of hyperrectangles generated by the algorithm after k iterations, and let f^{min} and f^{max} be respectively the ideal and nadir points of the cone centered in $f(c_{\tilde{R}})$. An hyperrectangle $\tilde{R} \in \mathcal{S}$ with center $c_{\tilde{R}}$ and measure $\alpha(\tilde{R})$ is said potentially partial optimal relative to the i-th objective if there exists at least a Lipschitz constant $K_i^{lower} > 0$ such that*

$$f_i(\mathbf{c}_{\tilde{R}}) - K_i^{lower}\alpha(\tilde{R}) \leq f_i(\mathbf{c}_R) - K_i^{lower}\alpha(R) \quad (8)$$

$$f_i(\mathbf{c}_{\tilde{R}}) - K_i^{lower}\alpha(\tilde{R}) \leq f_i^{min} - \varepsilon|f_i^{min}|. \forall R \in \mathcal{S} \quad (9)$$

or a constant $K_i^{upper} > 0$ such that

$$f_i(\mathbf{c}_{\tilde{R}}) + K_i^{upper}\alpha(\tilde{R}) \leq f_i(\mathbf{c}_R) + K_i^{upper}\alpha(R) \quad (10)$$

$$f_i(\mathbf{c}_{\tilde{R}}) + K_i^{upper}\alpha(\tilde{R}) \leq f_i^{max} - \varepsilon|f_i^{max}|. \forall R \in \mathcal{S} \quad (11)$$

where $\varepsilon \sim 10^{-4}$ is a constant to control the clustering *during the search [Jones].*

This definition is easily extendible to the case of n variables.

In order to obtain the heuristic which extends the above definition to the multiobjective case, let us redefine the Pareto optimality in general terms of efficiency.

Definition 3.2 ***[Efficiency criterion]*** *A decision vector $x^* \in X$ is* efficient *with respect to the convex cone D if there does not exist another decision vector $x \in X$ such that $f(x^*) - f(x) \in D$*

The cone D is called *ordering cone* and if $D = R_+^n$ the efficiency criterion produces a partial ordering for the Pareto optimality criterion. This ordering is used by the algorithm as surrogate of linear ordering.

Remark 3.3 ***[Multiple estimation of the Lipschitz constants]*** *Starting from the conditions 8 and 10 in 3.1 it is possible to define the multiobjective optimality in terms of expected efficiency. For every objective i, from the above conditions we obtain estimates for K_i^{lower} in the form of an upper bound $\overline{K}_i^{lower} \geq 0$ and a lower bound $\underline{K}_i^{lower} \geq 0$ for K_i^{lower}. Analogously, for K_i^{upper} there will be an upper bound $\overline{K}_i^{upper} \geq 0$ and a lower bound $\underline{K}_i^{upper} \geq 0$.*

The heuristic criterion leading to the choice of the optimal hyperrectangles in the multiobjective is motivated by the potential increase of the expected efficiency.

Definition 3.4 ***[Multiobjective Potential optimality]*** *Given the estimations of the upper bounds and the lower bounds for the Lipschitz constant of every objective i in the cone centered in $f(c_{\tilde{R}})$, the hyperrectangle $\tilde{R}$ is said potentially optimal if*

$$\sqrt{\sum_{i=1}^{k}[\underline{K}_i^{lower}]^2} \leq \sqrt{\sum_{i=1}^{k}[\overline{K}_i^{lower}]^2} \quad (12)$$

or

$$\sqrt{\sum_{i=1}^{k}[\underline{K}_i^{upper}]^2} \leq \sqrt{\sum_{i=1}^{k}[\overline{K}_i^{upper}]^2} \quad (13)$$

Moreover, let f^{min} and f^{max} be respectively the ideal and nadir points of the cone centered in $f(c_{\tilde{R}})$. The choice of hyperrectangle $\tilde{R}$ leads to a non trivial improvement of objective functions

$$\sum_{i=1}^{k}[f_i(\mathbf{c}_{\tilde{R}}) - K_i^{lower}\alpha(\tilde{R})]^2 \leq \sum_{i=1}^{k}[f_i^{min} - \varepsilon|f_i^{min}|]^2 \quad (14)$$

or

$$\sum_{i=1}^{k}[f_i(\mathbf{c}_{\tilde{R}}) + K_i^{upper}\alpha(\tilde{R})]^2 \leq \sum_{i=1}^{k}[f_i^{max} - \varepsilon|f_i^{max}|]^2 \quad (15)$$

The above definition gives a heuristic rule to choose hyperrectangles which are potentially optimal in the sense of either increasing the efficiency of the objective vector or taking into account possible trade-off (the latter arises from considering both lower and upper bounds for the Lipschitz constant). Equations 14 and 15 can be interpreted as controlling the clustering nearby the optimal points.

It must be remarked that the criterion for multiobjective optimality uses the l_2 norm to synthetize the choice from several estimations (see Eqns. 12, 13, 14, 15). The l_α norm used in this synthesis characterizes the expectation of the Pareto front in terms of convexity. Possibly, the magnitude of α that is required can be related to a measure of the non-convexity of the Pareto front.

Afterwards the hyperrectangle will be subdivided in thirds along its widest sides based on a dominance sorting of function values $f(\mathbf{c} \pm \delta\mathbf{e}_i)$ with respect their efficiency. This strategy increases the attractiveness of searching near points with good function values into the large hyperrectangles.

3.4 Multiobjective Evolutionary Algorithms

Two well-known multiobjective evolutionary algorithms (MOEAs) [Deb01, CoelloEtAL] are compared, NSGA2

and SPEA2 with the multiobjective Direct algorithm (MO-Direct) in terms of hypervolume metric (see subsection 3.4.1). Both evolutionary algorithms use the constrained tournament selection to deal with unfeasible solution during the optimization process.

NSGA2 [Deb02] is a elitist evolutionary algorithm with a fast nondominated sorting procedure and a density estimation of the solutions provided by the crowding distance. SPEA2 has a fitness assignment scheme based on the Pareto dominance relation with a density estimation technique based on the "k-th nearest neighbour" [Zitzler].

3.4.1 Hypervolume

The *Hypervolume metric* [Zitzler98] compute the volume covered by a set of non dominated elements Q obtained by the algorithm at the end of the optimization process. It is computed as follow: for each solution $i \in Q$ an hypercube v_i is built between a reference point W (in this study W is the axis origin) and the solution i. The union of all the hypercubes is computed and its hypervolume (HV) is:

$$HV = volume(\cup_{i=1}^{|Q|} v_i) \tag{16}$$

A large value of HV is expected from a good algorithm, however, Veldhuizen [Veldhui], report that a such metric is not useful if the optimal Pareto front is not convex.

3.5 Results for the OTA

In this section, using the hypervolume metric, we provide a comparison of NSGA II, SPEA 2 and MOdirect to multipleobjective optimization of the Operational Transconductance Amplifier. Objective of our experimental protocol is to determine the algorithms that obtain large value of HV, that is good algorithms in term of a well-known metric.

The optimization process is done with a circuit simulator (SPICE). NSGA2 and SPEA2 parameters are: crossover probability $p_c = 0.9$, mutation probability $p_m = 1$. Both MOEAs use SBX crossover with index $\eta_c = 2$ and gaussian mutation with $\sigma \in \{0.01, 0.1, 0.4\}$. These parameter values have been obtained by a preliminary parameter tuning (not shown due to space limit), the gaussian mutation parameter, σ, seems to be the most relevant controlling value. Obviously, a more deep investigation is necessary to detect the "best" parameter values. Different performance due the value of σ are reported in terms of hypervolume in Table 8. Each algorithm is stopped after 600000 function evaluations. Table 8 clearly shows as the MOEAs obtain a larger value of HV with respect MOdirect. Bests results are obtained with $\sigma = 0.4$. Approximatively, the hypervolume value obtained by MOdirect is about an order less than NSGA II and SPEA 2.

The chart in Figure 3.5 shows the results which identify the Pareto front between unity gain frequency and gain at 100 Hz (two objective functions to maximize). The upper envelop of the point set approximates this tradeoff curve and points out that the 100 Hz gain decreases when the unity gain frequency increase. The chart shows also a sparse set

Table 8: Hypervolume metric.

σ	SPEA2	NSGA2	MO-Direct
0.01	1.032791e+08	1.038449e+08	
0.1	1.030949e+08	1.019888e+08	6.40256e+07
0.4	1.048202e+08	**1.051418e+08**	

of points which locate other tradeoff among these two objective and the other (Area, Power consumption, phase margin). It is easy to note as both MOEAs obtain wider pareto fronts with respet MOdirect. In particular, for the Frequency objective function, NSGA II and SPEA 2 reached non dominated solutions in the ranges $30 - -75$, where MOdirect obtain frequencies in the domain $30 - 55$. Analogously, for the Gain objective function, the multiobjective evolutionary algorithms outperform the deterministic optimization algorithm, MOdirect. Same conclusions can be reached ploting any other pair of objective functions.

It must be notice the different point distributions of the different algorithms. This phenomenon is related to the convexity or nonconvexity of the multiobjective problem in different region of the domain which affects the optimization strategie of the algorithms. It can be see a comparable result for NSGA2 and SPEA2 and a totally unlike result of MOdirect. In fact, the potentially optimal criterion of MOdirect (3.4) favours convex region of the domain and gives a low priority to the nonconvex regions. This property can explain away the reduce value of the hypervolume. We surmise that the results could be useful to characterize the particular solution sets of the Pareto optimal front which identifies a convex region for the multiobjective problem.

4 Conclusion

In this article we compared evolutionary algorithms and standard optimization methods facing single- and multi- objective optimization for circuit design problems. The single-objective optimization concerns parameter extraction of Inductor device circuit, while the multi-objective optimization refers to determination of the approximate Pareto Front of a Operational Transconductance Amplifier. Two real world applications supplied by STMICROELECTRONICS. Based only on these circuit design problems and on the above reported statistical analysis, we can make the following statements.

For the parameter extraction of device circuit model, using a montecarlo simulation, the standard method DIRECT is more robust than evolutionary algorithms. In terms of the standard metrics (Absolute Error Deviation and minimization of the maximum absolute deviation) and number of function calls evolutionary algorithms are more effective than standard optimization methods.

Moreover, for the multiobjective problem, the observed Pareto fronts determined by evolutionary algorithms have a better spread of solutions with a larger number of nondominated solutions with respect to the standard multi-objective techniques, Multi-Objective DIRECT (MO-DIRECT).

Finally, the experimental results obtained shows as evo-

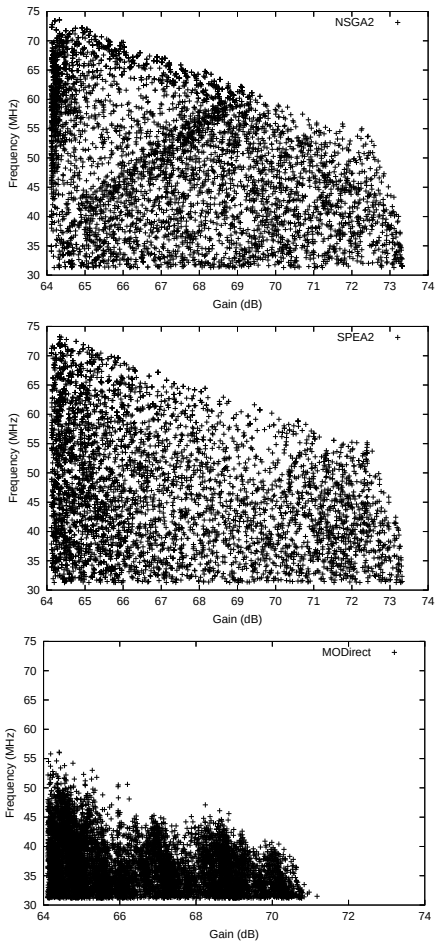

Figure 4: Frequency (MHz) vs. Gain (dB) for NSGA II, SPEA 2 and MOdirect algorithms.

lutionary algorithms in terms of quality solutions and computational effort are effective tools for circuit design problems.

Basically, multiobjective evolutionary algorithms have been designed on genetic algorithms or on evolutionary strategies framework; there exist few studies using genetic programming [Langton] (typically used for automated synthesis of analog eletrical circuits see for instance [KozaEtAl]). As future works, could be usefule to investigate the multiobjective genetic programming approach for the circuit design problems, and the synthesis of inductors and amplifiers optimizing the corresponding multiobjective problem and minimizing the size of the genetic programs.

Bibliography

[Keser00] Keser, M. and Joardar, K. (2000) “Genetic Algorithm Based MOSFET Model, Parameter Extraction,” Technical Proceedings of the International Conference on Modeling and Simulation of Microsystems, pp.341-344.

[Rinaudo98] Rinaudo, S. Privitera, G. Ferla, G. and Galluzzo, A. (1998) ”Small-Signal and Noise Modeling of Submicrometer Self aligned Bipolar Transistor,” Radio and Wireless Conference RAWCON 98, pp. 347-350.

[Jones] Jones, D. R. Perttunen, C. D. and Stuckman, B. E. (1993) “Lispchitz Optimization Without The Lipschitz Constant,” Journal of Optimization Theory and Application, vol. 79, pp.157-181.

[Neumaier] Huyer, W. and Neumaier A.(2005) “Snobit - Stable Optimization by Branch and Fit,” (submitted).

[Hjam] Hjamlmarson, E.(2003) “Studies on desing automation of analog circuits - the design flow,” Department of Electrical Engineering. Linlköpings universitet, Linlköpings, Sweden. PhD Thesis.

[Phelps] Phelps, R. *et al.* (2000) “Anaconda: Simulation based synthesis of analog circuits via stochastic pattern search,” IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, vol. 19, no. 6, pp. 703-717.

[Deb01] Deb, K. (2001) “Multi-Objective Optimization using Evolutionary Algorithms,” John Wiley & Sons.

[Deb02] Deb, K. Pratap, A. Agarwal, S. and Meyarivan T. (2002) “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA II,” IEEE Trans. on Evolutionary Computation, vol. 6, no. 2. pp.182-197.

[Zitzler] Zitzler, E. Laumannns, M. and Thiele, L. (2002) “Spea2: Improving the Strength Pareto Evolutionary Algorithm for MultiObjective Optimization,” Evolutionary Methods for Desing, Optimisation and Control. CIMNE, Barcelona, Spain, pp. 95-100.

[Veldhui] Veldhuizen, V. D. (1999) “Multiobjective Evolutionary Algorithms: Classification, Analyses, and New Innovations,” Ph.D. Thesis, Dayton, OH: Air Force Institute of Technology. Technical Report No. AFIT/DS/ENG/99-01.

[CoelloEtAL] Coello Coello, C. A. Van Veldhuizen, D. A. and Lamont, G. B. (2002) “Evolutionary Algorithms for Solving Multi-Objective Problems,” Kluwer Academic.

[Price] Price, K. Storn, R. M. and Lampinen, J. A. (2005) “Differential Evolution: A practical approach to global optimization,” Springer.

[Zitzler98] Zitzler, E. and Thiele, L. (1998) “Multiobjective optimization using evolutionary algorithms – A comparative case study,” in Parallel Problem Solving from Nature V (PPSN - V), pp. 292-301.

[Langton] Langton, W. B. (1995) “Pareto, Population partitioning, price, and genetic programming,” Technical Report Research Note RN/95/29, University College London, London, UK.

[KozaEtAl] Koza, J. R. Bennett, III F. H. Andre, D. and Keane, M. A. (1999) “Genetic Programming III,” Morgan Kaufmann Publishers.