

Investigations on the Optimal Support Vector Machine Classifiers for Predicting Design Feasibility in Analog Circuit Optimization

Jiho Lee and Jaeha Kim

Abstract—In simulation-based circuit optimization, many simulation runs may be wasted while evaluating infeasible designs, i.e. the designs that do not meet the constraints. To avoid such a waste, this paper investigates the use of support vector machine (SVM) classifiers in predicting the design's feasibility prior to simulation and the optimal selection of the SVM parameters, namely, the Gaussian kernel shape parameter γ and the misclassification penalty parameter C . These parameters affect the complexity as well as the accuracy of the model that SVM represents. For instance, the higher γ is good for detailed modeling and the higher C is good for rejecting noise in the training set. However, our empirical study shows that a low γ value is preferable due to the high spatial correlation among the circuit design candidates while C has negligible impacts due to the smooth and clean constraint boundaries of most circuit designs. The experimental results with an LC-tank oscillator example show that an optimal selection of these parameters can improve the prediction accuracy from 80 to 98% and model complexity by 10 \times .

Index Terms—Support vector machine classifier, analog design optimization, feasibility prediction, SVM parameter

I. INTRODUCTION

Simulation-based circuit optimizers find an optimal set of design parameter values (e.g. transistor sizes) for a given analog circuit by iteratively running circuit simulations with new design candidates [1-3]. One challenge that most simulation-based circuit optimizers face is the long execution time, due to a large number of simulation runs typically required to find the optimal design. Especially, when the design has many constraints to satisfy, the optimizer may waste simulation efforts merely in evaluating the infeasible designs. This is particularly the case when a designer prescribes the parameter ranges that span a much larger space than the actual feasible design space of the circuit, as illustrated in Fig. 1. For instance, when designing an LC oscillator, it is difficult to set the search ranges for the transistor sizes and inductance/capacitance values that can tightly encompass the designs that satisfy the start-up condition, frequency tuning range, etc. If the ranges are too wide, the optimizer may spend most of its time exploring the infeasible space, as it occupies a large portion of the design space.

This paper focuses on the problem of avoiding this waste of efforts in simulating the infeasible designs and investigates the use of support vector machine (SVM) classifiers to predict whether a chosen design candidate will satisfy the design constraints or not prior to running the simulation. SVM [4-6] is a popular classification algorithm that can identify via training the criteria of discriminating sample points distributed in a high-dimensional space. While the circuit optimizer selects

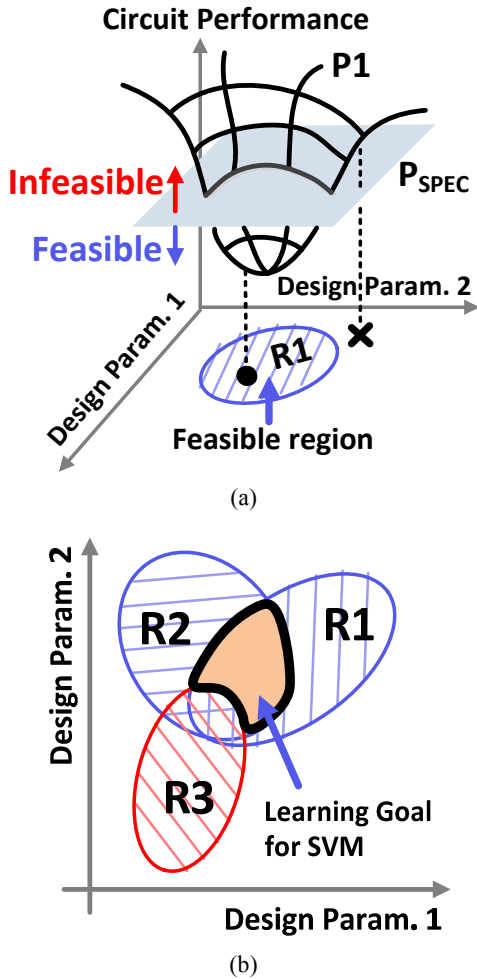


Fig. 1. (a) Illustration of a feasible region (R1) corresponding to a design constraint ($P_1 < P_{SPEC}$). performance constraints, (b) An SVM classifier aims to learn and predict the feasible space spanned by multiple design constraints (R1, R2, and R3)

design candidates and runs simulations for them, an SVM classifier can train itself regarding the feasible design space of the circuit and provide its prediction as to whether the design is feasible or not (Fig. 1). The SVM classifier can make an increasingly better prediction as the iteration proceeds and can prevent the optimizer from launching simulations for infeasible design candidates, saving the overall execution time.

However, the performance and accuracy of an SVM classifier can vary widely with its model parameters, such as kernel shape and penalty. These parameters are known to affect both the classification accuracy for the training data and prediction accuracy for the new data [7]. While some optimization approaches to determine the SVM parameters in image recognition applications have been previously proposed [8, 9], they are applicable only

when a large number of training samples are available and the training process may take a long time. Hence, these approaches are not suitable for predicting the feasibility of design candidates during circuit optimization where the training set accumulates incrementally and the SVM classifier must be trained repetitively as the optimization iteration proceeds. In prior works, the SVM parameters are determined by selecting the one which shows the best performance among trials [4-6].

This paper conducts an empirical study on the influences of the shape and penalty parameters on the SVM classifier's prediction accuracy and model complexity. Using a two-class SVM classifier with a soft margin as our baseline classifier, a series of experiments is conducted to investigate the influences of the classifier parameters on the performance of predicting the feasible design space of an LC oscillator with six design parameters. Contrary to the common belief that the larger shape parameter and larger penalty parameter lead to the more accurate and reliable classification, our results indicate that the SVM classifiers with small shape parameters delivers the better prediction performance despite their slower evaluation speed. It is primarily because most analog circuits exhibit strong spatial correlation across the design space [10]. On the other hand, the penalty parameter is found to have negligible influence because the training samples collected from circuit simulations have little noise and the feasible region boundaries formed by these samples are smooth and clean.

The remaining part of this paper is organized as follows. First, the background on the SVM classifier is briefly reviewed in Section 2, followed by a discussion of its key model parameters and their influences on the model complexity and prediction accuracy in Section 3. Then, Section 3 outlines the setup of our empirical study and Section 4 presents the results on the LC oscillator example. Finally, Section 5 concludes this paper.

II. SUPPORT VECTOR MACHINE (SVM) CLASSIFIERS

The classification problem is finding a function of the data point that the class or category of new data point can be computed from it. In machine learning approaches

with supervised-learning scheme, the function is usually found by determining its coefficient to classify a training data set, which includes the information of class for the data points.

The SVMs was first proposed with the idea of formulating classification problem as a maximization problem of classification margin, which is provided by a hyperplane that separates two group of data points.

Two-class soft margin SVM is used in this paper. It assumes the data points consists of two groups and allows some of the data points to be misclassified. The kernel trick for SVMs is proposed to handle the data which is not linearly separable. By using the kernel trick, the data points are projected into the space of higher dimension, where the projected points become linearly separable [7].

Eq. (1) shows the Lagrangian dual formulation for the problem that the SVM classifier solves to maximize classification margin. In (1), \mathbf{x}_i stands for the position vector of data point and y_i is the label of the data point whose value can be 1 or -1. The constraints in (1) requires the right classification of training data, while allowing some of the training data points to be misclassified for achieving larger classification margin and enhancing its prediction accuracy.

$$\begin{aligned} \text{Maximize } & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^N \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j) \\ \text{Subject to } & \sum_{i=1}^N \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C \\ & k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2\right\} \end{aligned} \quad (1)$$

$$(2)$$

The function (2) is called Gaussian kernel, which is exponentially decreasing function of the distance between two data points. By construction, it models the spatial correlation or resemblance between the data points, controlling the decreasing rate of correlation with the shape parameter γ . Gaussian kernel function is one of the most popular kernel functions, since it can provide projection to infinite dimensional space and can applied to problems with arbitrary nonlinear boundary.

The parameters γ and C are generally known to affect the performance in classifying the training data and predicting for new data. Since the kernel function can be

interpreted as the spatial correlation between data points, small γ is suitable for the data points with strong correlation and vice versa. The penalty parameter C controls the trade-off between maximizing classification margin and regularizing the noisy training data. It enables the SVM classifier to work for the data points with overlapping region between different classes. Usually, setting too small values for the γ and C causes underfitting problem, while overfitting problem arises when too large values are assigned for them.

III. OPTIMAL SVM CLASSIFIER FOR PREDICTING DESIGN FEASIBILITY OF ANALOG CIRCUITS

The smooth characteristic of analog performance surface makes the boundary of feasible design region for analog circuit to be clean and smooth. The feasible design region is formed from the performance functions and constraints which give upper bound or lower bound for the performance metric. Therefore, the smoothness or roughness of the boundary for feasible region is determined by that of the performance function. Also, the feasible design region and infeasible design region would be clearly distinguished. In other words, there are little design points that constructs overlapping region in which the feasible design points and infeasible design points are blended. In determining the parameters γ and C , these peculiarities need to be considered as well as the usual concerns for over-fitting and under-fitting problems.

The shape parameter γ of the Gaussian kernel first need to be small to reflect the smoothness. The continuous and smooth performance surface of analog circuit results in strong correlation between the feasibility of adjacent design points. The Gaussian kernel function models the strong correlation by making the decreasing rate of spatial correlation to be small. Certainly, too strong modeling for the correlation between design points may cause under-fitting problem, oversmoothing the boundary more than its natural smoothness. However, the main consideration in finding optimal values for γ is to find small value of it to best model the smoothness.

The penalty parameter C for misclassification is expected to have little influence because the boundary of feasible design regions will be clean. In other words, the smooth performance surfaces do not generate

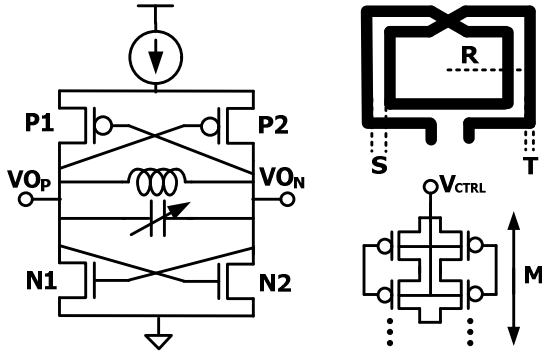


Fig. 2. An LC-tank voltage controlled oscillator. The LC tank consists of a spiral inductor and a MOS capacitor bank

overlapping regions for feasible design region and infeasible design. It implies there is little room for the penalty parameters to affect the performance of SVM classifier by rejecting noise in training data.

The small influence of the C will be exhibited near the pointed corners of feasible design region. When the design specification of analog circuit includes many performance metrics, the feasible design region can have sharp corners or edges which are formed during intersecting feasible design regions corresponding to individual performance metrics. In this locality, the penalty parameter can have large influence coinciding with too small shape parameter which over-smoothes the keen edges. In this case, some of the feasible region is trimmed out during training SVM classifier, being the misclassified data points. Nevertheless, the overall effect of penalty parameter is expected to be small since the great part of the boundary is smooth and clean.

IV. EXPERIMENT DESIGN

The overall experiment is performed by training SVM classifier with different parameters and comparing the corresponding prediction accuracy and model complexity. Two dimensional sweep analysis is performed for γ and C , where the sweep range for γ and C is 10^{-3} to 10^2 and 10^6 to 10^{10} , respectively. Two data sets are generated for training SVM and validating the prediction result. Each data set consists of design parameter values for the circuit and the label for feasibility, which is generated by measuring circuit performance with SPICE simulations.

The design parameters in data sets are generated by sampling their values uniformly in the range given in

Table 1. Range of design parameters for the LC-VCO and four sample design parameters

Parameters	Range	D1	D2	D3	D4
Wp [um]	0.480 – 480	148	52	54	140
Wn [um]	0.480 – 480	187	235	72	164
R [um]	15 – 50	16	25	21	15
T [um]	5.0 – 10	9.5	5.2	8.3	6.5
S [um]	2.0 – 3.0	2.2	2.4	2.1	2.7
M	50.0 – 450	390	280	341	275

Table 2. The performance and feasibility of design samples (D1-D4) in Table 1

Design	F_{MIN} [GHz]	F_{MAX} [GHz]	α	PN [dBc]	Feasibility
D1	5.7	7.5	3.5	-108	True
D2	5.5	7.2	3.3	-103	True
D3	5.6	7.5	1.9	-102	False
D4	6.5	8.5	3.0	-105	False
Constraint	< 6.0	> 7.0	> 2.0	< -100	

Table 1. The examples of sampled design points are listed in Table 1 (D1, .D2, D3 and D4). Total of 5000 design parameter values are drawn, which is large enough to reduce the influence of the size of training data and to focus on the effect of γ and C .

The model complexity and prediction accuracy is chosen to evaluate the prediction accuracy of SVM considering its potential application to analog design optimization. The prediction accuracy is defined by the ratio of the number of correctly predicted design point and total number of design points in validation data set. The model complexity is defined as the ratio of the number of support vectors and the number of total input data points.

The example circuit is an LC-tank voltage controlled oscillator (LC-VCO) [11]. Fig. 2 shows the schematic diagram of the circuit which consists of cross-coupled MOS pairs, LC tank and current source. The LC tank consists of spiral inductor and MOS varactor bank. There are six design parameters whose searching ranges are given in Table 1. It includes the widths for P1, P2 (Wp) and N1, N2 (Wn), the radius, spacing, track width of spiral inductor (R , S , T) and the multiplication factor of MOS varactor bank (M). The current source is realized by a PMOS transistor with fixed gate voltage

For the LC-VCO circuit, performance constraints are given on upper bound for minimum oscillation frequency (F_{MIN}), lower bound for maximum oscillation frequency (F_{MAX}), lower bound for startup criterion (α) and upper

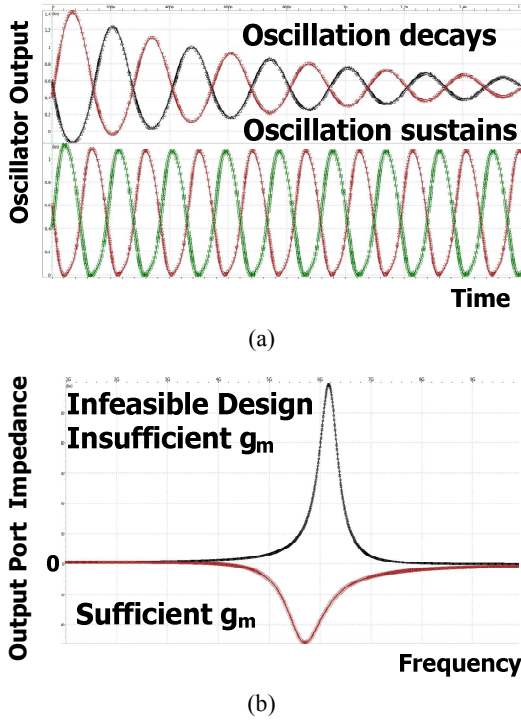


Fig. 3. (a) output voltage of oscillator, (b) output port impedance of a feasible design parameter and infeasible one

bound for phase noise (PN) at 10MHz offset from oscillation frequency. The oscillation frequency is controlled by changing the control voltage applied on the bulk node (V_{CTRL}) of the MOS varactor bank. The startup criterion is defined as (3), as the startup condition for oscillation is determined from the ratio of negative resistance of MOS pairs and the resistive loss of LC tank.

$$\alpha = g_m \cdot R_{TANK} \quad (3)$$

Although the power consumption of the circuit is also an important performance metric, it can be controlled easily by sizing PMOS current source and hence not considered by fixing the width of the PMOS transistor.

V. EXPERIMENT RESULT

Table 2 lists the circuit performances for example design parameters. The design points D1 and D2 are feasible while D3 and D4 are infeasible. The design D3 is infeasible since it violates the constraints on startup condition, α , while in D4, the minimum frequency constraint is not satisfied.

Fig. 3 shows the performance of circuit for a feasible

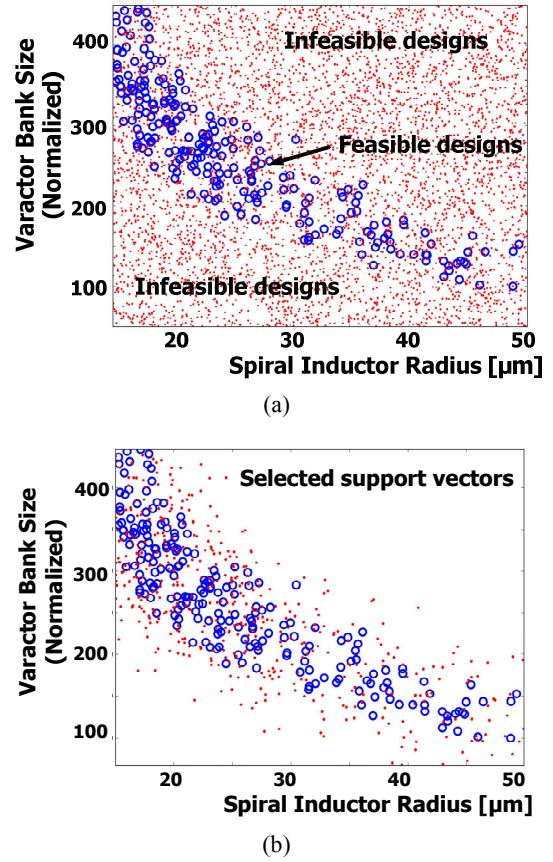


Fig. 4. Scatter plot of design parameters projected in 2D space. Circles represents feasible design parameters and dots represent infeasible ones (a) Training data set for SVM, (b) Support vectors determined by SVM

design and an infeasible one. The main difference between two designs is W_p and W_n , which determines the transconductance (g_m) provided by MOS pairs. Fig. 3(a) shows the output voltages on VO_p and VO_n over time. The oscillation sustains for the feasible design while the amplitude of oscillation decays for infeasible design. Fig. 3(b) shows the magnitude of output port impedance between VO_p and VO_n , which is also determined by the difference in g_m .

Fig. 4 shows the design parameters used for training data set and the support vectors selected after training SVM classifier. The circles and dots represent the feasible and the infeasible design points respectively. The support vectors are chosen from the boundary of the feasible design region as in Fig. 4(b). Among all of the input data points, only the support vectors are used in predicting feasibility of unknown design points and the number of support vectors determines the model complexity of the SVM classifier.

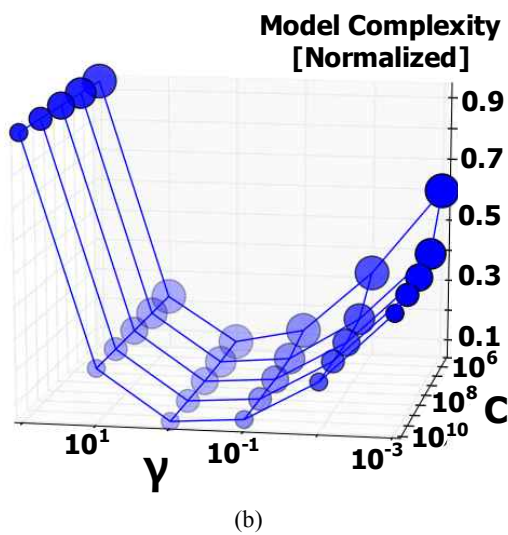
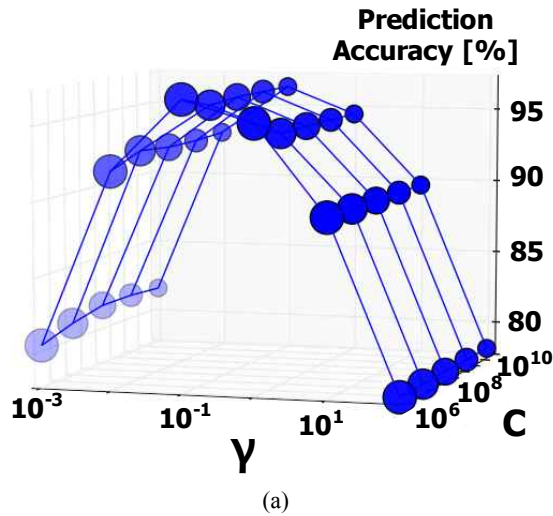


Fig. 5. The influence of SVM parameters on (a) prediction accuracy, (b) model complexity. The influence of shape parameter (γ) is much stronger than that of the penalty parameter (C)

Fig. 5(a) shows the influence of the parameters on the prediction accuracy. As expected, the shape parameter of kernel function mostly influences the prediction accuracy. The optimal value of shape parameter lies between 0.1 and 1.0, in which the smoothness in the boundary of the feasible region is modeled best.

The penalty parameter shows little influence on the prediction accuracy compared to that of the shape parameter. The underfitting problem, which is a usual consideration in machine learning algorithms, is expected to occur for the region where γ and C are smaller than 10^{-3} and 10^6 , respectively.

Fig. 5(b) shows the influence of the γ and C on the

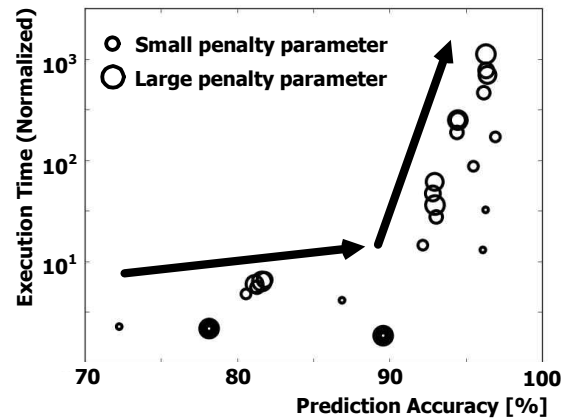


Fig. 6. The relationship between the prediction accuracy and execution time elapsed for training SVM and predicting feasibility based on it

resultant model complexity. Similar to prediction accuracy, the influence of the shape parameter is much stronger than that of the penalty parameter. The optimal value for shape parameter also lies between 0.1 and 1.0, where the prediction accuracy and the model complexity both show best performance.

Fig. 6 shows the relationship between prediction accuracy and execution time for SVM. The execution time includes the time elapsed for training SVM classifier and prediction. The overall trend is that the execution time increases as the prediction accuracy increase. For the high values of prediction accuracy about 97%, the execution time of SVM is rapidly increasing while the enhancement in prediction accuracy is little. It indicates that the misclassification penalty parameter trade-offs marginal enhancement in prediction accuracy and execution time of SVM classifier for feasibility prediction.

In using SVM, another point of practical interest is that how SVM's prediction accuracy affects the performance of circuit optimizer which uses it as a performance meta-model. Since this influence depends on the specific candidate selection algorithm used by a circuit optimizer, a notional representation of optimization process is useful to make a baseline analysis independent of the specific algorithm.

As the most basic approach for candidate selection is pure random sampling, a binomial distribution $B(n, p)$ can be used as an approximate of optimization process. That is, an optimization algorithm with SVM can be simplified as independent random trials of selecting n

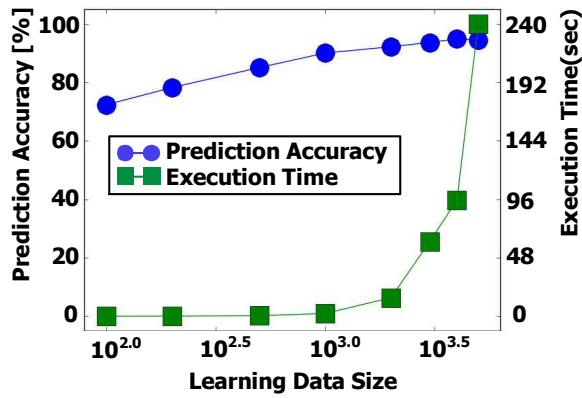


Fig. 7. The prediction accuracy and execution time of SVM. For learning data size larger than 10^{3.5}, the prediction accuracy increases marginally while the execution time shows rapid increase

candidate designs based on the SVM’s prediction accuracy p , where the trial success represents that a selected candidate design is proved to be feasible via simulation on it.

With this abstraction of an optimization process, the sensitivity of its execution number can be analyzed with respect to the prediction accuracy of SVM. For example, when the prediction accuracy of SVM in an optimizer changes from p to p' , its execution number would also change from n to n' . This change in optimization process can be represented by introducing another binomial distribution $B(n', p')$.

Since two optimization processes are assumed to be same in their candidate selection method and the only difference is the prediction accuracies of SVMs, they can give same quality of optimization when the average number of feasible designs among all sampled designs remains same. This can be expressed by Eq. (4), since the expectation of a binomial distribution is given by the product of the success probability (p) and the total number of trials (n).

$$np = n' p' \tag{4}$$

For example, when an optimization algorithm tries to evaluate 100 candidate designs with SVM accuracy of 60%, the trial number of optimizer can be reduce to 75 by enhancing the SVM’s prediction accuracy to 80% without degrading the average quality of optimized design chosen from the optimizer.

Fig. 7 shows the increase of prediction accuracy and execution time as the learning data set grows. For the size of learning data approximately larger than 10^{3.5}, the execution time rapidly increases while the prediction accuracy stops showing significant enhancement. Considering both execution time and prediction accuracy, it indicates that the learning data size should be set to 1000 - 3000 for training SVM most efficiently in optimizing the LC-VCO with the design requirements in Table 1, 2. The SVM parameters γ and C were set to 1.0 and 10⁶.

The circuit simulations and experiments in the paper are performed with TSMC 65nm technology process and Python package scikit-learn, which is implemented based on libsvm.

VI. CONCLUSION

This paper investigated the influences of SVM parameters for using it in predicting feasibility of analog design parameters with an LC-VCO circuit example. The effect of the penalty parameter and the shape parameter is analyzed by evaluating the prediction accuracy and the model complexity, which are important performances in using SVM classifier for simulation-based analog circuit optimization. The experimental result indicates that setting optimal value for the shape parameter for Gaussian kernel is the most important to achieve the high prediction accuracy and the low model complexity. Also, the trade-off between execution time and marginal enhancement in the prediction accuracy is exhibited. The result can be generalized for other types of analog circuits whose performance surface is smooth.

ACKNOWLEDGMENTS

This work was supported by the Technology Innovation Program (10049162, Battery-free general purpose RF remote controller using piezoelectric single crystal) funded By the Ministry of Trade, industry & Energy(MI, Korea). CAD tool licenses are supported by the IC Design Education Center (IDEC) in Korea.

REFERENCES

[1] G. Gielen and R. Rutenbar. “Computer-Aided

Design of Analog and Mixed-Signal Integrated Circuits," *Proceedings of IEEE*, pp. 1825-1852, Dec. 2000.

- [2] R. Phelps, et al. "ANACONDA: Simulation-based Synthesis of Analog Circuits via Stochastic Pattern Search," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, pp. 703-717, June, 2000.
- [3] S. Jung, J. Lee, and J. Kim. "Variability-aware, discrete optimization for analog circuits." *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, Vol. 33, No. 8, pp. 1117-1130, 2014.
- [4] De Bernardinis, et al. "Support vector machines for analog circuit performance representation." *Design Automation Conference DAC 2003. Proceedings. IEEE*, pp. 964-969, June, 2003.
- [5] T. Kieley and G. Gielen. "Performance modeling of analog integrated circuits using least-squares support vector machines", *Design, Automation and Test in Europe, DATE 2004. Proceedings, IEEE Computer Society*, pp.16 -20, Feb., 2004.
- [6] D. Boolchandani, et al. "Variability aware yield optimal sizing of analog circuits using SVM-genetic approach", *Symbolic and Numerical Methods, Modeling and Applications to Circuit Design (SM2ACD), 2010 XIth International Workshop on, IEEE*, pp.1-6, Oct., 2010.
- [7] C. M. Bishop. "Pattern recognition and machine learning" Vol. 4, No. 4, New York: springer. 2006.
- [8] P.B.C. de Miranda, et al. "Combining a multi-objective optimization approach with metalearning for SVM parameter selection," *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, pp. 2909- 2914, Oct. 2012.
- [9] O. Chapelle, et al. "Choosing multiple parameters for support vector machines." *Machine learning*, Vol. 46, pp. 131-159, 2002.
- [10] J. Kim, et al. "Discretization and discrimination methods for design, verification, and testing of analog/mixed-signal circuits," *Custom Integrated Circuits Conference (CICC), 2013 IEEE*, Vol. 1, No. 8, pp. 22-25, Sept. 2013.
- [11] D. Ham, et al. "Concepts and methods in optimization of integrated LC VCOs," *Solid-State Circuits, IEEE Journal of*, Vol. 36, No. 6, pp. 896-909, June, 2001.



Jiho Lee received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 2013. He is currently working toward the Ph.D. degree in Seoul National University. His research interests include design automation of analog/mixed-signal system, especially optimization, verification and testing of analog circuits.



Jaeha Kim is currently Assistant Professor at Seoul National University and his research interests include low-power mixed-signal systems and their design methodologies. He received the B.S. degree in electrical engineering from Seoul National University in 1997, and received the M.S. and Ph.D. degrees in electrical engineering from Stanford University in 1999 and 2003, respectively. Prior to joining Seoul National University in 2010, Dr. Kim was with Stanford University, CA as Acting Assistant Professor from 2009 to 2010, with Rambus, Inc., Los Altos, CA as Principal Engineer from 2006 to 2009, and with Inter-university Semiconductor Research Center (ISRC) in Seoul National University, Seoul, Korea as Post-doctoral Researcher from 2003 to 2006. From 2001 to 2003, he was with True Circuits, Inc., Los Altos, CA as Circuit Designer.