

# On Combining Alternate Test with Spatial Correlation Modeling in Analog/RF ICs

Ke Huang\*, Nathan Kupp†, John M. Carulli Jr.‡, and Yiorgos Makris\*

\*Department of Electrical Engineering, The University of Texas at Dallas, Richardson, TX 75080

†Department of Electrical Engineering, Yale University, New Haven, CT 06511

‡Texas Instruments Inc., 12500 TI Boulevard, MS 8741, Dallas, TX 75243

**Abstract**—Statistical intra-die correlation has been extensively studied as a means for reducing test cost in analog/RF ICs. Generally known as alternate test, this approach seeks to predict the performances of an analog/RF chip based on low-cost measurements on the same chip and statistical models learned from a training set of chips. Recently, an orthogonal direction for leveraging statistical correlation towards reducing test cost of analog/RF ICs has also gained traction. Specifically, inter-die spatial correlation models learned from specification tests on a sparse subset of die on a wafer are used to predict performances on the unobserved die. In this work, we investigate the potential of combining these two statistical approaches, anticipating that the performance prediction accuracy of the joint correlation model will surpass the accuracy of its constituents. Experimental results on industrial semiconductor manufacturing data validate this conjecture and corroborate the utility of the combined performance prediction models.

## I. INTRODUCTION

Towards reducing the excessive cost of specification testing in analog/RF circuits, which requires complex test equipment and elaborate measurement procedures, various statistical approaches have been developed in the past decade. Generally known as “alternate test”, numerous methods have been proposed for accurately testing analog/RF devices without explicitly measuring costly performances. The underlying principle is to approximate these performances through correlation models based solely on low-cost measurements [1]–[3] or on a subset of performance measurements [4]<sup>1</sup>. Such alternate test methods, the basic principle of which is illustrated in Figure 1, leverage *die-level* correlations in order to reduce test cost. Specifically, if  $\mathbf{m}_{alt}$  denotes the low-cost measurement vector and  $m_{per}$  denotes the performances to be predicted, alternate test uses a set of training samples to learn the correlation function between  $\mathbf{m}_{alt}$  and  $m_{per}$ , denoted by  $f_1$ , and thereby to predict performances on new devices with  $\mathbf{m}'_{alt}$ , as  $\hat{m}_{per} = f_1(\mathbf{m}'_{alt})$ . While this approach achieves dramatic test cost reduction, it comes at the cost of increased test escapes and yield loss. Furthermore, the train-predict cycle makes an assumption of stationarity of process statistics, i.e., the trained regression model is assumed to be valid across all new devices, otherwise it needs to be periodically retrained.

Recently, an orthogonal direction for leveraging *wafer-level* statistical correlation towards reducing test cost has also

<sup>1</sup>Alternatively, as described in [5]–[7], a machine learning-based approach can be employed in order to learn classification boundaries which separate passing and failing populations of analog/RF devices in a multi-dimensional space of inexpensive measurements.

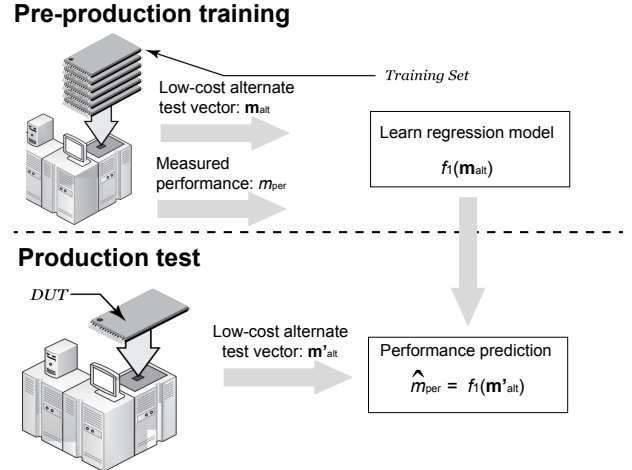


Fig. 1. Leveraging *die-level* measurement correlations for test cost reduction

attracted interest. An overview of this spatial correlation-based approach is shown in Figure 2. In this case, costly specification tests are not completely eliminated. Instead, they are only performed on a sparse subset of die on each wafer and, subsequently, used to build a spatial model  $f_2$ , which is then used to predict performances at unobserved die locations:  $\hat{m}_{per} = f_2(\mathbf{x})$ , where  $\mathbf{x}$  denotes the wafer's Cartesian coordinate  $\mathbf{x} = [x, y]$ . It should be noted that by building spatial correlation models on a per-wafer basis, this approach avoids making assumptions about process stationarity as in the case of alternate test. Along these lines, the expectation-maximization (EM) algorithm was used in [8] to estimate spatial wafer measurements, assuming that the data is generated by a multivariate normal distribution, or using the Box-Cox transformation in case it is not. The “Virtual Probe” (VP) approach [9], [10] modeled spatial variation via a Discrete Cosine Transform (DCT), which projects spatial statistics into the frequency domain. Similarly, the author of [11] laid the ground for applying Gaussian Process (GP) models to spatial interpolation of semiconductor data based on Generalized Least Square fitting and a structured correlation function. As recently shown in [12], using such GP models can dramatically improve both prediction accuracy and computational time, as compared to the VP model.

Each of these two statistical approaches has been shown to have high prediction accuracy and to effectively reduce the cost of specification testing. Herein, we investigate, for the first time, the potential of combining these two statistical ap-

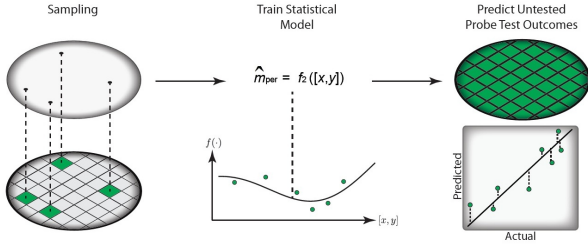


Fig. 2. Leveraging *wafer-level* spatial correlations for test cost reduction

proaches, expecting that the performance prediction accuracy of a joint correlation model will surpass the accuracy of its constituents. The proposed methodology, which is introduced in Section II, relies on a combined model for predicting performances, which incorporates the predictive power of both intra-die and inter-die correlations. In addition, we also introduce a screening step in order to verify that each of the constituent correlations truly exists. Indeed, not all performances are adequately predictable by both approaches. In fact, using a poor prediction model in the mix may not only fail to improve the accuracy of the combined model, but it might actually hurt it. Therefore, assessing effectiveness of the constituent models prior to combining them not only saves computation time but also safeguards the quality of the combined prediction model.

The proposed approach is experimentally assessed in Section III, using industrial semiconductor manufacturing data from an RF device. The reported results corroborate our expectation that when an RF measurement is deemed predictable by both models, a combined prediction model will significantly improve the prediction accuracy over the constituent models.

## II. PROPOSED APPROACH

### A. Overview

In this section, we describe in detail the proposed methodology for combining *die-level* (i.e. alternate test) and *wafer-level* (i.e. spatial correlation) models for analog/RF test cost reduction. Figure 3 illustrates an overview of the methodology, which consists of two stages, namely training and testing. Let  $\hat{m}_{per1}$  and  $\hat{m}_{per2}$  denote the estimated performance by *die-level* model  $f_1$  and *wafer-level* model  $f_2$ , respectively. During the training stage, we first learn  $f_1$  and  $f_2$  as explained in the previous section and as depicted in Figures 1 and 2, respectively. Then we assign a weight  $w_i$  to the  $i$ -th model,  $i \in \{1, 2\}$ , by solving the following optimization problem:

$$\underset{\mathbf{w}}{\text{minimize}} \quad \|\mathbf{m} - \mathbf{w} \cdot \mathbf{f}^T\| \quad (1)$$

where  $\|\cdot\|$  denotes the  $L2$  norm,  $\mathbf{w}$  denotes the weight vector of correlation functions,  $\mathbf{w} = [w_1, w_2]$ ,  $\mathbf{m}$  denotes the measurement vector of  $n$  samples used to assign the weights, and  $\mathbf{f}$  denotes the vector of considered correlation models,  $\mathbf{f} = [f_1, f_2]$ . In this work, we propose to solve the optimization problem in (1) using the ordinary least squares (OLS) method to learn the optimal weight vector  $\hat{\mathbf{w}}$ . Once  $\hat{\mathbf{w}}$  is learned, we can readily predict values of performances for untested die locations, as shown in the lower part of Figure 3:

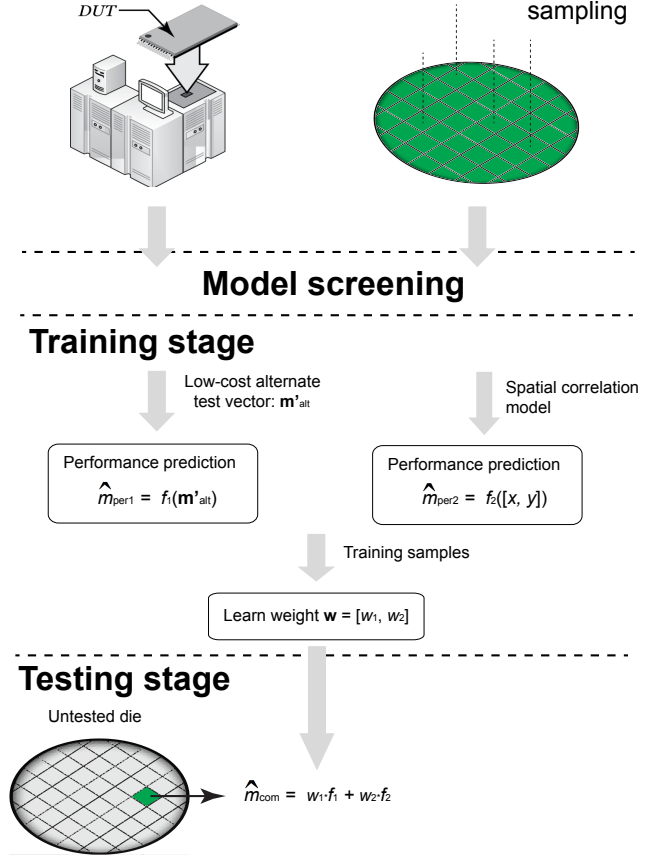


Fig. 3. Proposed approach: combined die- and wafer-level correlation models

$$\hat{m}_{com} = \hat{\mathbf{w}} \cdot \mathbf{f}^T \quad (2)$$

where  $\hat{m}_{com}$  denotes the predicted performances by the combined model. Thus, in order to obtain the combined estimation on a particular die using (2), we need to perform (a) low-cost measurements on the same die, and (b) specification tests on a sparse set of other die on the same wafer. The approach shown in equations (1) and (2) is, in essence, a regression combination approach, which assigns a weight to each considered model according to its performance, and predicts the final outcome using the estimated weights. Various methods to address this issue have been proposed, including Bayesian model averaging, which considers the weight as the posterior probability of the model [13], and weighting based on bootstrap or perturbation [14]. However, combining models may not always provide improvement in the prediction accuracy, especially when a subset of models have rather poor performances. To avoid this problem, we adopt the method proposed in [15], which introduces a model screening step. Such screening narrows down the list of candidate models, not only for saving computation time but also for removing poor models that would hurt the combined estimator.

### B. Model screening

As described previously, introducing models with poor performance in the combined estimator could degrade the

quality of the final prediction. Thus, it is important to evaluate the performance of each model before the combined estimator is constructed and applied. Authors in [15] proposed a model screening step, which narrows down the list of candidate models in the combined estimator. In this section, we present the details of model screening for each considered model.

1) *Screening of alternate test model*: The alternate test model consists of predicting a high-cost specification test  $m_{per}$  using low-cost alternate measurements  $\mathbf{m}_{alt}$ . To assess the effectiveness of the considered model, we propose to use a hold-out set  $S_1$  of  $n_1$  samples containing the specification test and the alternate measurement vector:  $S_1 = \{\mathbf{m}_{alt}^{(i)}, m_{per}^{(i)}\}$ ,  $i = 1, \dots, n_1$ . Then we split  $S_1$  into two equal sets  $S_{1t}$  and  $S_{1v}$  uniformly at random. The regression model for alternate test  $f_1(\mathbf{m}_{alt})$  is trained using  $S_{1t}$  and validated using  $S_{1v}$ . Finally, we compute the normalized root-mean-square (RMS) error in the validation set  $S_{1v}$ :

$$\epsilon_v = \sqrt{\frac{\sum_{i=1}^{n_1/2} (\hat{m}_{per}^{(i)} - m_{per}^{(i)})^2}{n_1/2}} / r_{per} \quad (3)$$

where  $r_{per}$  denotes the variation range of  $m_{per}$  in  $S_1$ , defined as  $r_{per} = \max(m_{per}) - \min(m_{per})$ , and  $\hat{m}_{per}^{(i)}$  is the estimated performance value for the  $i$ -th device in  $S_{1v}$ . We set a threshold value  $\theta_1$  for the considered performance  $m_{per}$ , such that the alternate test model is considered to be poor if  $\epsilon_v > \theta_1$ . Poor models are then screened out in the combined estimator.

2) *Screening of spatial model*: The spatial model predicts a performance on a die from measurements of the same performance on other die locations of the same wafer using a spatial correlation model  $f_2([x, y])$ . For each performance under consideration, the underlying question of assessing the spatial model is whether a systematic spatial correlation exists, or whether it is dominated by random noise. In the latter case, the performance can not be predicted by a spatial correlation model. In order to verify existence of systematic spatial patterns, we consider the measurement variation model as the sum of a systematic spatial component and a random component, as commonly modeled in the literature [16], [17]:

$$m(x, y) = g(x, y) + \epsilon \quad (4)$$

where  $m(x, y)$  is the considered measurement value, expressed as a function of a wafer's Cartesian coordinate  $(x, y)$ ,  $g(x, y)$  is the systematic spatial variation component, and  $\epsilon$  is the random component often modeled as  $\epsilon \sim \mathcal{N}(0, \sigma^2)$ . Notice that a constant term  $C$  can also be added to (4) to represent wafer-to-wafer and lot-to-lot shift. In this work, we assess the existence of systematic spatial variation component by computing the Pearson's correlation coefficient between two adjacent wafers  $i$  and  $j$ , defined by:

$$r_{i,j} = \frac{\sum_{k=1}^{n_m} (m_i(\mathbf{x}_k) - \bar{m}_i)(m_j(\mathbf{x}_k) - \bar{m}_j)}{\sqrt{\sum_{k=1}^{n_m} (m_i(\mathbf{x}_k) - \bar{m}_i)^2} \sqrt{\sum_{k=1}^{n_m} (m_j(\mathbf{x}_k) - \bar{m}_j)^2}} \quad (5)$$

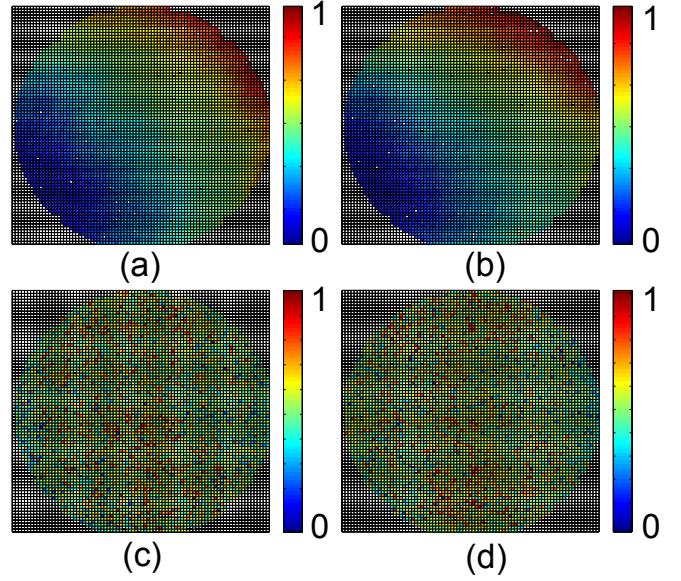


Fig. 4. Wafer map of two adjacent wafers for measurement with (a)(b) systematic spatial variation dominating random variation with  $r = 0.98$ , and (c)(d) systematic spatial variation dominated by random variation with  $r = 0.1$ .

where  $\mathbf{x}_k$  denotes the  $k$ -th Cartesian coordinate  $[x_k, y_k]$ ,  $m_i(\mathbf{x}_k)$  and  $m_j(\mathbf{x}_k)$  denote the performance value of the  $k$ -th Cartesian coordinate on the  $i$ -th and  $j$ -th wafers,  $\bar{m}_i$  and  $\bar{m}_j$  denote the mean value computed over all available die locations on the  $i$ -th and  $j$ -th wafers, and  $n_m$  is the number of available die on the wafer.

Our experience with production test data shows that the systematic spatial variation component  $g(x, y)$  remains very similar across adjacent wafers in manufacturing. Indeed, adjacent wafers in semiconductor manufacturing are fabricated in a very similar environment. As a consequence, the impact of wafer-to-wafer process variations and environmental variations, such as temperature, on these wafers is minimized. Based on this observation, it can be shown that in cases where  $g(x, y) \gg \epsilon$ , the correlation coefficient  $r$  between two adjacent wafers is large (we omit the derivation for brevity). Figure 4(a) and (b) show wafer maps of two adjacent wafers in manufacturing for a measurement with  $r = 0.98$ , and Figure 4(c) and (d) show the case of another measurement where  $r = 0.1$ . It can be observed by a simple visual inspection that  $r$  value is high for measurements having strong systematic spatial pattern and is low otherwise.

Based on the above observation, we set a threshold value  $\theta_2$  for spatial model screening, such that the model is considered to be poor if  $r_{i,j} < \theta_2$ , where  $r_{i,j}$  is the correlation coefficient between two adjacent wafers  $i$  and  $j$ .

We note that the choice of  $\theta_1$  and  $\theta_2$  is empirical, based on the acceptable levels of Test Escapes and Yield Loss, as we further discuss in Section III. We also note that the combined estimator, as well as the model screening approach, are independent of the underlying models; in other words, the proposed approach can be applied to combine any alternate test and spatial correlation models.

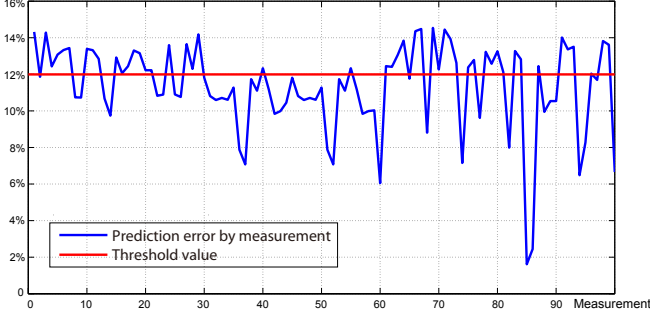


Fig. 5. Prediction error  $\epsilon_v$  in the hold-out set for alternate test model screening

### C. Prediction outcome evaluation

To evaluate effectiveness of the proposed method, we compute the mean RMS error across all predictions using (3):

$$\epsilon_{jh} = \sqrt{\frac{\sum_{i=1}^{n_{test}} (\hat{m}_{per}^{(i)} - m_{per}^{(i)})^2}{n_{test}}} / r_{per} \quad (6)$$

where  $\hat{m}_{per}^{(i)}$  denotes the predicted value on the  $i$ -th die location,  $\epsilon_{jh}$  represents the mean RMS error of predicting the  $h$ -th measurement for all unmeasured die locations on the  $j$ -th wafer,  $r_{per}$  denotes the range of variation for  $h$ -th measurement as defined in Section II-B1, and  $n_{test}$  denotes the number of predicted die locations on the  $j$ -th wafer. Then, we summarize the mean prediction error over all wafers as:

$$\epsilon_h = \frac{1}{N_{wafers}} \sum_{i=1}^{N_{wafers}} \epsilon_{ih} \quad (7)$$

where  $N_{wafers}$  denotes the number of considered wafers.

In order to gain insight about the prediction outcome, it is also worthwhile to compute the Test Escapes (TE) and Yield Loss (YL) incurred by applying the predictive model. For a particular measurement, let the indicator functions  $I_1^{(i)}/I_2^{(i)}$  be equal to '1' if the predicted value of the  $i$ -th die location passes/fails its specification, while the actual value fails/passes the specification, and let  $I_1^{(i)}/I_2^{(i)}$  be equal to '0' otherwise. Then the overall TE and YL are defined as:

$$\hat{TE} = \frac{1}{N} \sum_{i=1}^N I_1^{(i)} \quad (8)$$

$$\hat{YL} = \frac{1}{N} \sum_{i=1}^N I_2^{(i)} \quad (9)$$

where  $N$  denotes the total number of predicted die locations across all wafers.

## III. EXPERIMENTAL RESULTS

The proposed method is evaluated on high-volume manufacturing test data from an RF transceiver built in 65nm technology. Our dataset has a total of 291 wafers, each of which has approximately 2000 die characterized by 224 low-cost On-chip RF Built-in Tests (ORBiTs) obtained via on-chip sensors, and 101 high-cost RF specification tests. Both

low-cost and RF tests are performed at probe level. Thus, a statistically significant data-set of approximately 582,000 devices, each with 325 measurements, is used in our case study.

### A. Prediction with alternate test model

In this work, we use least-angle regression (LARS) [18] to construct alternate test models for each considered specification:  $\hat{m}_{per} = f_1(\mathbf{m}_{alt})$ . The LARS regression model automatically chooses a subset of variables in vector  $\mathbf{m}_{alt}$  which are most correlated with  $m_{per}$  in order to build the regression model  $f_1$ . Thus, this approach allows us to handle high-dimensional data in building the regression model without needing to perform a feature selection or dimensionality reduction analysis, which is very appropriate in our case since the dimension of  $\mathbf{m}_{alt}$  is 224. Details of the LARS regression model can be found in [18].

To perform screening of the alternate test models, as described in section II-B1, we generate the following data set: we choose the first wafer in our data set as the hold-out set  $S_1$  in order to assess the correlation between  $\mathbf{m}_{alt}$  and  $m_{per}$ :  $S_1 = \{\mathbf{m}_{alt}^{(i)}, m_{per}^{(i)}\}, i = 1, \dots, n_1$ , where  $n_1 \approx 2000$ . Then  $S_1$  is split into training set  $S_{1t}$  and validation set  $S_{1v}$  for the purpose of model screening. We empirically set the threshold value  $\theta_1 = 12\%$ , by observing that, for the majority of considered performances, this value results in a combined Test Escapes and Yield Loss of approximately 1000 ppm across our dataset. The prediction error  $\epsilon_v$  for each of the performances in  $S_1$  is computed using (3). Figure 5 shows  $\epsilon_v$  for all considered performances as well as the value of  $\theta_1$ . As may be observed, 52 RF specification tests have  $\epsilon_v$  smaller than  $\theta_1$  in our study. These measurements successfully pass model screening and are, therefore, forwarded to the combined model, as shown in Figure 3.

Once an alternate test model passes screening, we can readily use it to predict performances. As an example, Figure 6 and 7 show the prediction plot of devices from one wafer for RF measurements 85 and 86, respectively. Evidently, the predicted values correlate very well with the actual values of these performances. Nevertheless, further improvement can be achieved by combining these alternate test models with the spatial correlation models, as we demonstrate below.

### B. Prediction with spatial model

In this work, wafer-level spatial correlation models are built using the Gaussian Process-based approach described in [12] with a 10% die sample size. In screening out poor spatial models (i.e. performances that are not spatially correlated), the procedure described in section II-B2 is employed. Once again, we empirically set the threshold  $\theta_2 = 0.6$  by observing that, for the majority of considered performances, this value results in a combined Test Escapes and Yield Loss of approximately 1000 ppm across our dataset. The screening procedure for each measurement involves the following steps:

- *Step 1:* Randomly choose two adjacent wafers  $i$  and  $j$  produced in the same lot.

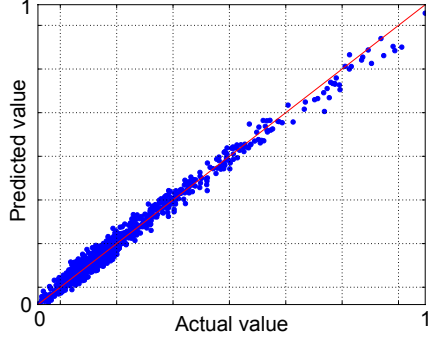


Fig. 6. Prediction plot for measurement 85 by alternate test model

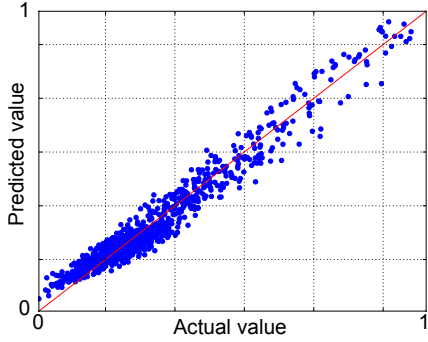


Fig. 7. Prediction plot for measurement 86 by alternate test model

- *Step 2:* Compute the correlation coefficient  $r_{i,j}$  as described in (5).
- *Step 3:* Repeat *Step 1* and *Step 2*  $n'$  times (where  $n'$  is a small integer greater than 1, in order to avoid making general decisions based on outlier wafers) and choose the maximum  $r_{i,j}$  value as the wafer correlation value.

Figure 8 shows the wafer correlation value  $r_{i,j}$  for all considered RF specification measurements as well as the value of  $\theta_2$ . As may be observed, a subset of 46 RF specification measurements have  $r_{i,j}$  value greater than  $\theta_2$ . These measurements successfully pass model screening and are, therefore, forwarded to the combined model, as shown in Figure 3.

### C. Prediction with combined model

The final combined model is used to predict performances for which both the alternate test and the spatial correlation models pass the respective screening step. In our experiment, the intersection of the two sets (of cardinality 52 and 46, respectively) contains 34 performances. For each of them, we use, again, the first wafer as the hold out set in order to assign the weight  $w_l$  for the  $l$ -th prediction model computed by (1), and equation (2) to predict the performances for the unobserved die locations on other wafers. For new wafers to be tested, low-cost alternate tests are taken on each die location, while specification tests are performed on 10% of the available die locations randomly sampled across the wafer, in order to train the spatial correlation model using GP as described in Section I. Figure 9 shows the mean RMS prediction error sorted in ascending order using all three methods, obtained by predicting on the remaining 90% die locations for all

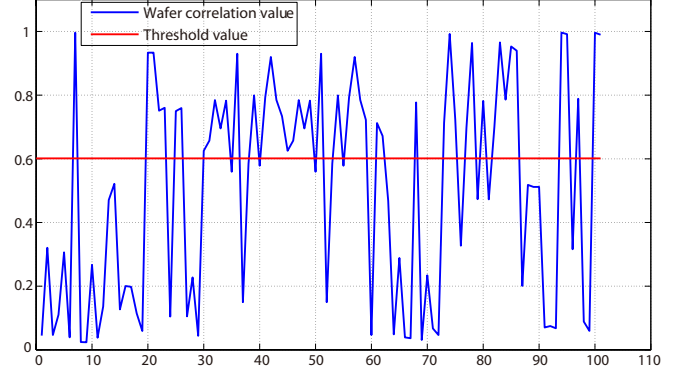


Fig. 8. Wafer correlation plot for spatial model screening

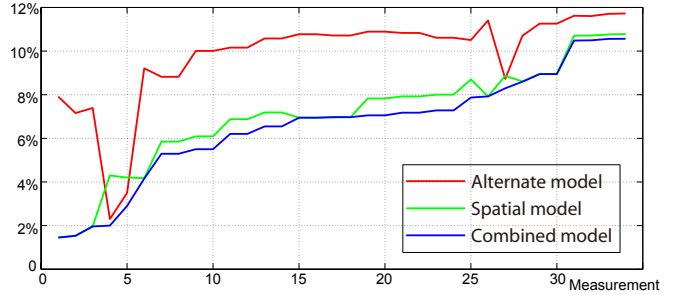


Fig. 9. Comparison of RMS prediction error for measurements passing screening of both models

wafers. As may be observed, the combined model consistently outperforms, or at least performs equally well as the best of the individual models. Figures 10 and 11 show the prediction plots of measurement 85 and 86 for the same devices as in Figure 6 and 7, using all three models. As may be observed by simple visual inspection, the combined model indeed provides the best prediction results.

Table I also shows this comparison in terms of RMS prediction error, as computed by (7). Once again, it can be observed that the combined model consistently provides the best prediction results, which justifies our choice to combine the two predictive models.

### D. Test escapes and yield loss improvement

To further elucidate the benefits of our approach, in the 2nd to 5th columns of Table II we compare the Test Escapes (TE) and Yield Loss (YL) of measurement 85 and 86, computed using (8) and (9) and projected in parts-per-million (*ppm*) for each individual model, as well as the combined model. Evidently, the proposed approach achieves a significant TE and YL improvement as compared to the individual models, thereby justifying the use of a combined model in predicting RF performances.

### E. Discussion

One might point out that the proposed approach incurs additional cost in predicting performances, as compared to traditional alternate test, since it requires that specification tests are performed on a subset of die on each wafer. In reality,

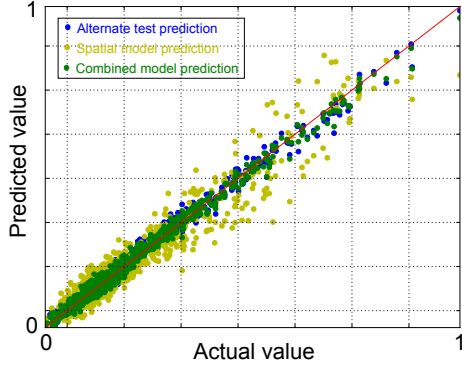


Fig. 10. Prediction plot for measurement 85 by different models

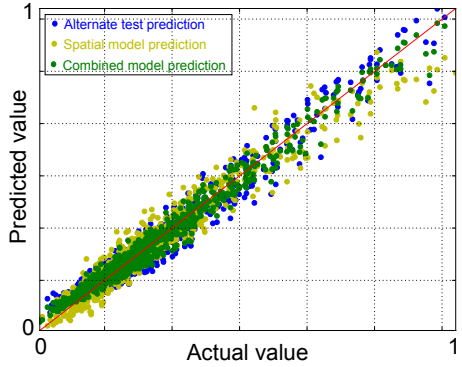


Fig. 11. Prediction plot for measurement 86 by different models

however, even in a traditional alternate test setting, a certain level of specification tests are also needed to ensure integrity of the models in the presence of process shifts, to deal with die for which performances cannot be predicted with high confidence (two-tier test [6]), and to monitor the process. With this in mind, and taking into account the achieved TE and YL reduction, which might give the extra nudge needed to meet test quality goals, we believe that the added cost is justifiable from a test economics point of view.

#### IV. CONCLUSION

Combining alternate test with spatial correlation modeling holds great promise in further reducing test cost in analog/RF integrated circuits without compromising test quality. Such merging, however, needs to be carefully orchestrated, taking into account that poorly performing constituents may jeopardize the effectiveness of the joint predictor and should, therefore, be screened out. Experimental results using test data from high-volume manufacturing assert that the joint prediction model proposed herein achieves lower RMS prediction error and, by extension, reduces Test Escapes and Yield Loss.

#### V. ACKNOWLEDGEMENT

This research has been carried out with the support of the National Science Foundation (NSF CCF-1149463) and the Semiconductor Research Corporation (SRC-1836.092).

TABLE I  
RMS PREDICTION ERROR COMPARISON

Prediction model	Measurement 85	Measurement 86
Alternate model	2.3%	3.5%
Spatial model	4.3%	4.2%
Combined model	<b>2%</b>	<b>2.9%</b>

TABLE II  
TEST ESCAPES (TE) AND YIELD LOSS (YL) (IN ppm) COMPARISON

	Measurement 85		Measurement 86	
	TE	YL	TE	YL
Alternate model	515	115	430	83
Spatial model	589	347	563	334
Combined model	<b>401</b>	<b>68</b>	<b>343</b>	<b>54</b>

#### REFERENCES

- [1] P. N. Variyam, S. Cherubal, and A. Chatterjee, "Prediction of analog performance parameters using fast transient testing," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 21, no. 3, pp. 349–361, 2002.
- [2] R. Voorakaranam, S. S. Akbay, S. Bhattacharya, S. Cherubal, and A. Chatterjee, "Signature testing of analog and RF circuits: Algorithms and methodology," *IEEE Trans. Circuits Syst. I*, vol. 54, no. 5, pp. 1018–1031, 2007.
- [3] H. Ayari, F. Azais, S. Bernard, M. Comte, V. Kerzerho, O. Potin, and M. Renovell, "Making predictive analog/RF alternate test strategy independent of training set size," in *Proc. IEEE International Test Conference*, 2012, Paper 10.1.
- [4] J. B. Brockman and S. W. Director, "Predictive subset testing: Optimizing IC parametric performance testing for quality, cost, and yield," *IEEE Trans. Semicond. Manuf.*, vol. 2, no. 3, pp. 104–113, 1989.
- [5] S. Biswas and R. D. Blanton, "Test compaction for mixed-signal circuits using pass-fail test data," in *IEEE VLSI Test Symposium*, 2008, pp. 299–308.
- [6] H.-G. Stratigopoulos and Y. Makris, "Error moderation in low-cost machine-learning-based Analog/RF testing," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 2, pp. 339–351, 2008.
- [7] H.-G. Stratigopoulos, P. Drineas, M. Slamani, and Y. Makris, "RF specification test compaction using learning machines," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 18, no. 6, pp. 998–1002, 2010.
- [8] S. Reda and S. R. Nassif, "Accurate spatial estimation and decomposition techniques for variability characterization," *IEEE Trans. Semicond. Manuf.*, vol. 23, no. 3, pp. 345–357, 2010.
- [9] H.-M. Chang, K.-T. Cheng, W. Zhang, X. Li, and K.M. Butler, "Test cost reduction through performance prediction using virtual probe," in *IEEE International Test Conference*, 2011, pp. 1–9.
- [10] W. Zhang, X. Li, F. Liu, E. Acar, R.A. Rutenbar, and R.D. Blanton, "Virtual probe: a statistical framework for low-cost silicon characterization of nanoscale integrated circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 30, no. 12, pp. 1814–1827, 2011.
- [11] F. Liu, "A general framework for spatial correlation modeling in VLSI design," in *Design Automation Conference*, 2007, pp. 817–822.
- [12] N. Kupp, K. Huang, J. Carulli, and Y. Makris, "Spatial correlation modeling for probe test cost reduction," in *IEEE/ACM International Conference on Computer-Aided Design*, 2012.
- [13] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, "Bayesian model averaging: A tutorial," *Statistical Science*, vol. 14, no. 4, pp. 382–417, 1999.
- [14] L. Breiman, "Bagging predictors," *Machine Learning*, no. 24, pp. 123–140, 1996.
- [15] Z. Yuan and Y. Yang, "Combining linear regression models: When and how?," *Journal of the American Statistical Association*, vol. 100, no. 472, pp. 1202–1214, 2005.
- [16] J.K. Kibarian and A. Strojwas, "Using spatial information to analyze correlations between test structure data," *IEEE Trans. Semicond. Manuf.*, vol. 4, no. 3, pp. 219–225, 1991.
- [17] W. Zhang, K. Balakrishnan, X. Li, D. Boning, and R. Rutenbar, "Toward efficient spatial variation decomposition via sparse regression," in *IEEE/ACM International Conference on Computer-Aided Design*, 2011, pp. 162–169.
- [18] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.