Process Control in a High-Noise Environment Using a Limited Number of Measurements[1]

Leandro G. Barajas, Magnus B. Egerstedt, Edward W. Kamen, and Alex Goldstein
{LBarajas,Magnus,Ed.Kamen,Alex.Goldstein}@ece.gatech.edu
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332, U.S.A.

Abstract

In this paper, we develop a hybrid control algorithm that produces control values for processes where only a limited number of function evaluations are available for the control law generation. This situation arises, for example, in stencil printing processes in printed circuit board manufacturing, where the cost associated with multiple function evaluations is prohibitive. The proposed control algorithm is given by a modified version of a constrained conjugated-gradient method, transitioned into a windowed-smoothed block-form of the least-squares affine estimator.

1 Introduction

Process control has long been a highly successful branch of control theory, both in terms of industrial impact and theoretical development. (For example, see the classic reference [1]). However, as new technologies have emerged, from Printed Circuit Boards (PCB) to Micro-Electro-Mechanical Systems (MEMS), new challenges have presented themselves. In this paper we investigate one such challenge, concerning how to adjust the machine parameter settings in a stencil printing process (SPP) for PCB production. The proposed closed-loop control algorithm compensates for discrepancies associated with different print directions, recovers from faulty initial settings, and provides robustness by maintaining an acceptable performance in the presence of environmental variations or unscheduled process interrupts. To control the SPP in an closed-loop fashion has long been an evasive goal to the industry. In fact, even as some closed-loop control strategies have been proposed, their limited applicability and portability into production environments have prevented these methodologies from being widely adopted [2–4].

The outline of this paper is as follows: In Section 2 we present the mathematical process model under consideration and define the control problem at hand. This is followed by a description of the SPP, in Section 3. In Sections 4 and 5, the hybrid control strategy used for controlling the SPP is introduced, corresponding to one coarse tuning phase, and one local optimization phase. Finally, experimental results are given in Section 6.

2 The Control Problem

Consider a process governed by the control action $C \in \mathbb{R}$, with measurements $y \in \mathbb{R}$ of quality characteristics of the process given by

$$y = F(C) + v, \tag{2.1}$$

where F is an unknown function of the control variable C, and v is the process noise.[2] The control variable C can be thought of as a machine setting, and it is assumed that C can take on any value in the interval $[C_{min},\ C_{max}]$.

Now, given some desired value H_d for the quality characteristics, the objective is to determine a value C_d for the control variable such that $||F(C_d) - H_d||$ is suitably small. In many industrial processes, the value of the control variable C (the machine setting) is determined by the equipment operator using past experience obtained from running the process. However, in order to achieve a robust and effective performance in an on-line fashion, one obvious solution would be to generate feedback control laws of the form $C = G(y)$ for some function G, such that $||F(G(y)) - H_d||$ is made small during process operation. Unfortunately, in many industrial processes, the function F cannot be determined from a first-principles analysis.

Another complication is that any type of system identification approach relies on sufficient system excitation, i.e. either a large number of measurements must be made or large changes in the control variable are required. In the SPP application, each measurement corresponds to the printing of a

[1]This work was supported in part by the Georgia Institute of Technology Manufacturing Research Center Grants No. B01D01 and B01D07.

[2]This static model can for example be obtained by lifting the dynamic behavior of each run to a static map.

board, with a prohibitive cost and time investment associated with it. We are thus constrained by the process to assume that only a limited number of measurements are available for determining the control law.

If we formalize these observations, what we want to achieve is the following:

$$\min_{C_1,C_2,\ldots} \{M\}, \tag{2.2}$$

subject to

$$\begin{array}{l} y_k = F(C_k) + v_k \\ \|C_{k+1} - C_k\| \leq \delta \\ y_{M+p} \in \mathcal{Y}(H_d),\ p = 0,1,\ldots \\ \|C_{M+p} - C_M\| \leq \epsilon,\ p = 0,1,\ldots \end{array} \tag{2.3}$$

In other words, we want to reach a desirable process performance $y \in \mathcal{Y}(H_d)$ in a minimum amount of steps, M, thus achieving fast convergence, i.e. using few measurements. Secondly, we want to achieve this while bounding the variability of the control signals, i.e. $\|C_{k+1} - C_k\| \leq \delta$ for some given $\delta > 0$. For the SPP, no dynamical models are as of yet available due to the highly complex process dynamics [5, 6] and we need to bound the control variability in order to suppress the transient effects associated with changing the control value. The third point that is illustrated by the constraints is that we want the control laws to reach a desired operating point and then remain close to that point for all future control values, i.e. $\|C_{M+p} - C_M\| \leq \epsilon$, for some given $\epsilon > 0$. If F is known, this problem can be solved using dynamic programming, as shown in [7], but for the problem under consideration in this paper, no such assumptions can be made. Hence only suboptimal solutions can be obtained.

However, if the process measurements are considered noise-free and $F(C)$ is known and continuously differentiable, then the problem can be recursively addressed by Newton's approximation algorithm. If the starting point of the iteration, C_0, is close enough to the desired value of the control parameter,C_d, then the algorithm is guaranteed to converge [8, 9]. Another possibility could be to use linear multistep methods, predictor-corrector methods, or even Runge-Kutta methods. (A complete description of these methods and their convergence and stability properties can be found in [10].) However, these higher order methods will require several evaluations of $F(C)$, which is a luxury we cannot afford in this situation.

Now, if F is unknown, a more promising direction would be to use the classic Robbins-Monro algorithm for stochastic approximation [11], which can be used for recursive root finding of unknown functions, when only noisy observations are available. (Extensive descriptions of this and other related techniques can be found in [8, 12].) The Robbins-Monro algorithm can furthermore be augmented, as shown in [13–15], to produce a noise-constrained least mean squares algorithm, which is of particular relevance to the problem at hand.

Unfortunately, all of the previously mentioned algorithms rely on the assumption that (2.1) can be evaluated an arbitrary number of times. As we have pointed out, this is not possible in practice because each function evaluation requires the manufacturing of one product through the production line. In contrast to this, the algorithm proposed in this paper explicitly tries to address the following key points:

1. Minimize the number of function evaluations.
2. Produce adequate values for the control variable at all times.
3. Find and maintain an appropriate operational point for the process.
4. Avoid exciting higher order modes in the system by limiting the magnitude of the maximum step size.

The key idea behind the proposed algorithm is the observation that the problem is composed of two subproblems. The first involves driving the system to $y \in \mathcal{Y}(H_d)$ in as few steps as possible, and the second involves keeping the process in this region. It should thus be possible to decompose the control algorithm into two subparts as well, resulting in a switched, or hybrid control law. In particular, the following two control strategies will be used:

1. A constrained conjugated-gradient search for reaching the desired operational band is used in the first part of the control law.
2. A least-squares (LS) affine estimator for maintaining and fine-tuning the process once the process window has been encountered constitutes the second part.

3 Process Description

The goal of the SPP in Surface-Mount Technology (SMT) manufacturing of Printed Circuit Boards (PCB) is to apply an accurate and repeatable volume of solder paste deposits at precise locations [16–18]. Given that most of the defects in SMT manufacturing can be attributed to the SPP [2, 19, 20], then, this makes the SPP the most critical step in the process.

In order to solder components to a PCB, it is necessary to "print" solder bricks over the metallic contact pads on the PCB. Once this is successfully achieved, the components are placed on top of the solder bricks, pushing their leads into the paste. When the components have been attached, the solder paste is melted using either reflow soldering or vapor-phase soldering to create the electro-mechanical junctures. Finally, the manufactured PCBs are inspected and tested. Unfortunately, the SPP exhibits a number of features that make it extremely hard to control [2, 5, 6, 19]; some of the most relevant factors are the following:

1. Poorly understood process physics;
2. Difficulty in measuring key variables;

3. High-noise environment;
4. Limited number of measurements; and
5. Software/hardware implementation limitations.

The industry standard for measuring the quality characteristics of the process is solder-paste-volume deposition. However, as of yet, there are no machines that can directly measure the solder brick volume. Instead, algorithms are used to estimate the effective area and mean height of the solder paste deposit and therefore their product becomes the estimated volume. Given that under normal conditions, the area of the solder brick deposits do not change significantly with the modification of the control parameters, the height of the deposits become the measure of interest. Commonly, a direct sample mean of such values is used as quality characteristics. A more involved approach would be to assign different weights to each solder brick type so that problematic components can be given more importance in the quality characteristics generation process. Such a weighted scheme can be represented by

$$\overline{H}_W(n) = \sum_{i=1}^{Q} w_i h(n,i), \quad \sum_{i=1}^{Q} w_i = 1, \quad w_i \geq 0, \quad (3.1)$$

where Q is the number of solder bricks inspected in the n^{th} board, w_i is the weight assigned to the i^{th} solder brick, and $h(n,i)$ is the height of the i^{th} solder brick on the n^{th} board.

For steady-state performance evaluation, the weighted mean-squared error between the mean weighted height $\overline{H}_W(n)$ and the desired height H_d in board-by-board basis can be used. This representation, as shown in (3.2), is appropriate since it addresses two of the key factors in the process, i.e. the mean squared error and the variance of the process.

$$\begin{aligned} MSE_W(n) &= \mathrm{E}[(\overline{H}_W(n) - H_d)^2] \quad (3.2) \\ &= \underbrace{\mathrm{Var}(\overline{H}_W(n))}_{\text{AC ErrorTerm}} + \underbrace{(\mathrm{E}[\overline{H}_W(n)] - H_d)^2}_{\text{DC Error Term}}. \end{aligned}$$

4 Coarse Control Algorithm

In this section, we construct the first part of the switched control strategy proposed in Section 2. It is based on a quantization of the control values, i.e. C will only take on values in the set $\mathcal{S}(C_0, \Delta)$, where

$$\mathcal{S}(C_0, \Delta) = \{C \in [C_{min},\ C_{max}] \mid C = C_0 + k\Delta,\ k \in \mathbb{Z}\}. \quad (4.1)$$

What we want our iterative control algorithm to do is to generate a $C^\star \in \mathcal{S}(C_0, \Delta)$ such that

$$|F(C^\star) - H_d| \leq |F(C) - H_d|,\ \forall\, C \in \mathcal{S}(C_0, \Delta). \quad (4.2)$$

Note that in general $C^\star$ is not equal to C_d. [3]

[3] In fact, $C^\star$ belongs to a discrete set of equally spaced control values and it will minimize the steady-state error for that given set only, while C_d is a root of $F(C)$ that provides the desired average output value H_d for the zero-mean noise case.

We first assume that F in (2.1) is continuously differentiable and monotonically increasing or decreasing on $[C_{min},\ C_{max}]$. We furthermore assume that the noise term v is bounded over the interval $[-a,\ a]$, which implies that if F is monotone, then the only fundamental entity that our control algorithm needs to recover is the sign of the slope of F. However, this should be achieved in the presence of the noise term v, and for this, we let Δ denote the step-size in the algorithm, and we assume that Δ as well as the initial control value $C_0 \in [C_{min} + \Delta, C_{max} - \Delta]$ are chosen such that

$$\begin{aligned} F(C_0) \notin &\quad [F(C_0 + \Delta) - 2a, F(C_0 + \Delta) + 2a] \\ &\bigcup\ [F(C_0 - \Delta) - 2a, F(C_0 - \Delta) + 2a] \\ &\bigcup\ [H_d - 2a, H_d + 2a]. \quad (4.3) \end{aligned}$$

The interpretation here is that we have enough separation of our initial control values to be able to recover the slope of F at the same time as we start the iterative control process suitably far away from the desired operating point. We furthermore define δ_0 as

$$\delta_0 = H_d - y(C_0), \quad (4.4)$$

where we have used $y(C)$ as shorthand for $F(C) + v$ in (2.1).

In order to achieve the desired operational point, a weak-search algorithm based on a conjugated gradient scheme can be used:

Algorithm 4.1 *Let* C_0 *and* Δ *satisfy* (4.3), *and let* δ_0 *be given in* (4.4).
`Set` $C_1 = C_0 + \mathit{sign}(\delta_0)\Delta$
`Set` $m_1 = \mathit{sign}(y(C_1) - y(C_0))\mathit{sign}(C_1 - C_0)$
`Set` $i = 1$
`While` $|y(C_{i-1}) - H_d| > |y(C_i) - H_d|$
  `Set` $\delta_i = H_d - y(C_i)$
  `Set` $C_{i+1} = C_i + m_i \mathit{sign}(\delta_i)\Delta$
  `Set` $m_{i+1} = \mathit{sign}(y(C_{i+1}) - y(C_i))\mathit{sign}(C_{i+1} - C_i)$
  `Set` $i = i + 1$
`End`
`Set` $C^\star = C_{i-1}$.

We note that due to the monotonicity of F, Algorithm 4.1 will in fact terminate, and as long as we have enough separation around the final control value, i.e. that

$$\begin{aligned} F(C^\star) &\notin [F(C^\star + \Delta) - 2a,\ F(C^\star + \Delta) + 2a] \\ &\quad \bigcup\ [F(C^\star - \Delta) - 2a,\ F(C^\star - \Delta) + 2a], \\ C^\star &\in [C_{min} + \Delta,\ C_{max} - \Delta], \quad (4.5) \end{aligned}$$

then Algorithm 4.1 recovers the optimal control value in (4.2). Furthermore, since the first step of the algorithm consists of guessing the slope (setting $m_0 = 1$), maximally two steps (wrong initial guess and a recovery step) are wasted. Another step will be used for overshooting the desired value, which is needed for terminating the process. In other words, assume $k^\star \in \mathbb{Z}$ is such that $C_0 + k^\star\Delta \in \mathcal{S}(C_0, \Delta)$ solves (4.2). Then Algorithm 4.1 terminates in at most $k^\star$+3 steps with the terminating control value $C_0+k^\star\Delta$ as long as (4.5) is satisfied. [4]

As a consequence, we know that if there exists a $C \in \mathcal{S}(C_0,\Delta)$ such that $F(C) \in [H_d - a, H_d + a]$, then $C^\star$ satisfies

$$F(C^\star) \in [H_d - a, H_d + a].$$

The monotonicity assumption can roughly be thought of as a first order approximation of F, and Algorithm 4.1 is in fact general enough to capture second order approximations as well. To see this, assume that F is Lipschitz continuous and unimodal on $[C_{min},\ C_{max}]$, i.e. there exists a unique $C_m \in [C_{min},\ C_{max}]$ such that, $\forall\ C_1 \leq C_2 \leq C_m \leq C_3 \leq C_4$

$$\begin{aligned} m_1 &= \text{sign}\left(F(C_2) - F(C_1)\right) = \text{sign}\left(F(C_1) - F(C_m)\right), \\ m_2 &= \text{sign}\left(F(C_3) - F(C_m)\right) = \text{sign}\left(F(C_4) - F(C_3)\right), \\ m_1 &= -m_2. \end{aligned} \tag{4.6}$$

Here $C_1, C_2, C_3, C_4 \in [C_{min}, C_{max}]$ and C_m is the extremum value of F being either the maximum or minimum value depending on the concavity of the function. We furthermore assume that the initial control value C_0 and step-size Δ not only satisfy (4.3), but also that $F(C_0)$ is far enough from the extremum, i.e. that

$$\text{sign}\left(F(C_0+\Delta) - F(C_0)\right) = \text{sign}\left(F(C_0) - F(C_0 - \Delta)\right). \tag{4.7}$$

Under these assumptions we can directly apply Algorithm 4.1 to the unimodal case as well, since on each sides of the extremum (if it exists) F is monotone.

5 Steady State Optimization

Taking the final value of the Algorithm 4.1, $C^\star$, as the initial control value for the next stage, a local optimization procedure can be performed in order to fine-tune $C^\star$ such that the steady state error is minimized. For this, a Local Affine Least Squared (LALS) algorithm can be applied to the optimization of the process [21,22].

An interphasing step is necessary between these two stages in order to ensure that the switching between Algorithm 4.1 and the LALS algorithm is non-abrupt. For this, the values of the two final iterations of Algorithm 4.1 are kept in order to maintain the estimated slope direction in the presence of large disturbances. This is required because noisy measurements can otherwise invert the direction of the slope.

After this block-from LS-step, the algorithm changes to a windowed version that also preserves information about previous estimates. More specifically, if the length of the LS window is W, then the $(i-(W+1))^{th}$ estimated parameter vector θ is stored and used to maintain the slope of the estimator. The vector θ is the set of calculated coefficients for the affine or quadratic estimator depending on the process characteristics. If (4.5) holds, then an affine estimator can be used to determine the next control variable. However, if (4.5) does not hold, i.e. $F(C^\star)$ cannot generate the desired outputs, then a quadratic estimator can be used to minimize the steady state error even under such adverse conditions.

In other words, given the static input-output map $y = F(C) + v$, we want to find an affine estimate of F as $F(C) = \theta_0 + \theta_1 C$ using the last N output values. With a slight abuse of notation we denote these by $y_1,\ldots,y_N$ and let $C_1,\ldots,C_N$ denote the corresponding N last inputs. What this implies is that N defines the size of the sample window. Thence, in order to recover the affine estimator parameters $\theta = [\theta_0\ \theta_1]^{\mathrm{T}}$, the standard basis, given by the $N \times 2$-matrix M, can be formed from the unitary N-vector $b = [1\,1\,\ldots\,1]^{\mathrm{T}}$ and the control values vector $\mathbf{C} = [C_1\ C_2\ \ldots\ C_N]^{\mathrm{T}}$, arranged side by side such that $M = \left[\ b \ \vdots\ \mathbf{C}\ \right]$.

Now, it can be shown that the classic LS solution [23] to the over-determined problem $M\theta = \mathbf{y}$, $\theta = (M^{\mathrm{T}}M)^{-1}M^{\mathrm{T}}\mathbf{y}$, where $\mathbf{y} = [y_1\ y_2\ \ldots\ y_N]^{\mathrm{T}}$, satisfies the following equation

$$\theta = \frac{\mathbf{C}^{\mathrm{T}}\left[\left(\dfrac{\mathbf{C}b^{\mathrm{T}} - \mathbf{C}^{\mathrm{T}}b\mathbb{I}_N}{N\mathbb{I}_N - bb^{\mathrm{T}}}\right) \otimes \mathbf{y}\right]}{\mathbf{C}^{\mathrm{T}}(N\mathbb{I}_N - bb^{\mathrm{T}})\mathbf{C}}, \tag{5.1}$$

where $\mathbb{I}_N$ is the $N \times N$ identity matrix.[5]

Once θ has been estimated using the measurements $\mathbf{y}$, it is possible to calculate the next control value C_{N+1} needed to achieve the desired output H_d, i.e. $H_d = \theta_0 + \theta_1 C_{N+1}$. This also gives an estimation of the noise component of the measurements $\hat{v} = \mathbf{y} - \hat{\mathbf{y}}$. The expanded expression for the estimate of $\mathbf{y}$, $\hat{\mathbf{y}}$ in terms of the control values and the measurements is given by

$$\hat{\mathbf{y}} = \frac{\left[\mathbf{C}^{\mathrm{T}}\left(N\mathbb{I}_N - bb^{\mathrm{T}}\right)\mathbf{y}\right]\mathbf{C} + \left[\mathbf{C}^{\mathrm{T}}\left(\mathbf{C}b^{\mathrm{T}} - b^{\mathrm{T}}\mathbf{C}\mathbb{I}_N\right)\mathbf{y}\right]b}{\left[\mathbf{C}^{\mathrm{T}}(N\mathbb{I}_N - bb^{\mathrm{T}})\mathbf{C}\right]}. \tag{5.2}$$

Now, an estimate of the mean and variance can be obtained by the sample mean and sample variance of the noise vector $\hat{\mathrm{V}} = [\hat{v}_1\ \hat{v}_2 \cdots \hat{v}_N]^{\mathrm{T}}$ [23], such statistics are given by

$$\mu_{\hat{\mathrm{V}}} = \frac{1}{N} b^{\mathrm{T}}\hat{\mathrm{V}}, \tag{5.3}$$

$$\sigma^2_{\hat{\mathrm{V}}} = \frac{1}{N-1}(\hat{\mathrm{V}} - b\mu_{\hat{\mathrm{V}}})^{\mathrm{T}}(\hat{\mathrm{V}} - b\mu_{\hat{\mathrm{V}}}). \tag{5.4}$$

---

[4] It should be pointed out that the bound $k^\star + 3$ can be reduced to $k^\star + 2$ if we modify Algorithm 4.1 to let $|C_2 - C_1| = 2\Delta$ if the initial slope-guess was wrong.

[5] For the sake of clarity, it should be noticed that the partition matrix in 5.1 operates (using the Kronecker product ($\otimes$)) on the vector $\mathbf{y}$ of data samples.

A similar argument can be used in the quadratic regression case. In this case, the LS estimate becomes

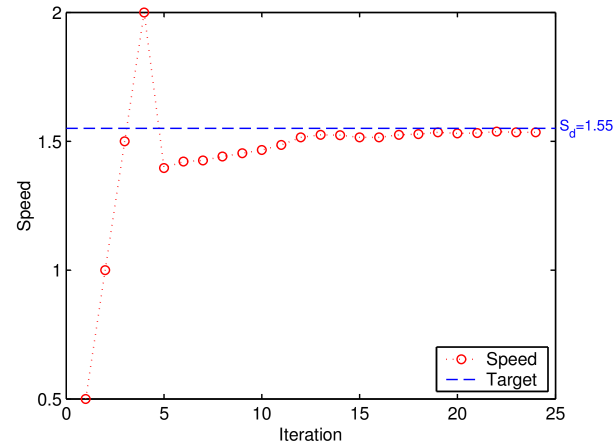
$$\hat{y} = \theta_0 + \theta_1 C_{N+1} + \theta_2 C_{N+1}^2, \quad (5.5)$$

which is obtained by including additional powers of $\mathbf{C}$ to the matrix M in a straight-forward manner.

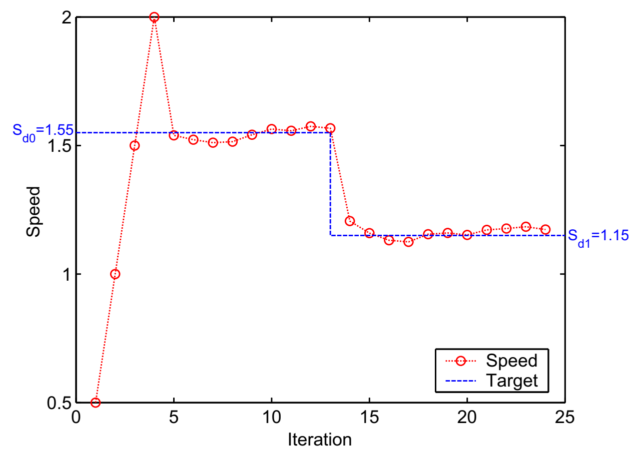
6 Experimental Results

The hybrid control algorithm proposed in this paper was applied to the stencil printing process in surface-mount technology manufacturing. The input control variable is taken to be the squeegee speed, and the output is the height of the deposited solder bricks. Figure 1(a) shows the complete set of control values generated by the hybrid algorithm going through all operational modes and its convergence to the optimum value. Additionally, the case when a large disturbance is introduced was considered. In this case, the output setting was changed by -25% after the 13^{th} iteration. Figure 1(b) shows how the algorithm was able to adapt to this sudden set-point change in only a few iterations.

Figure 2 shows the effect of the controller over the process in a real production run. This plot demonstrates how the controller discriminates between printing directions and by adjusting the weighted sample mean on a board-by-board basis independently in each direction it aligns the solder brick height distributions to the desired mean height. Depicted are the output histograms generated with and without the controller. It should be noted that the distribution in Fig.2(b) is located exactly on the target desired height H_d of 5.5mil (139.7μm). As marked for reference in Fig.2, the stencil thickness S_t in the experiment under consideration is 5mil (127μm).



(a) Constant desired height.



(b) Step applied to desired height.

Figure 1: Control variable response.

7 Conclusions

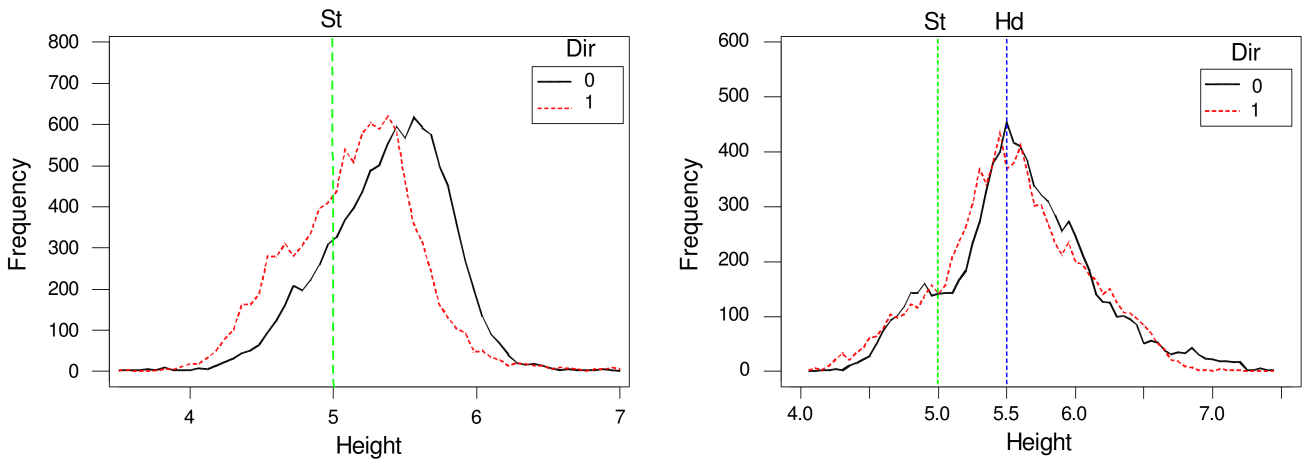
In this paper we present an algorithm for producing closed-loop control values in a printed circuit board application. The process is given by an unknown, and noisy, static input-output relation, and we have shown how to control such a system using only a small number of data samples. Since each measurement corresponds to the printing of a board it is vitally important (both from a cost and from a timing point of view) that the number of measurements are kept to a minimum.

Our solution to this problem consists of a hybrid control strategy that combines a constrained conjugate-gradient search with a least-squares solution. The soundness of the algorithm is stressed by the fact that it has been implemented and applied to numerically control a SPP in SMT, which is the first time that this has been done successfully without having to modify the process significantly. What our particular contribution has provided in this area is thus twofold. First of all, we have proved that it is possible to compensate for differences in solder brick heights in backward and forward printing directions, and to recover from abnormal equipment set points, which is potentially useful in a number of manufacturing applications. Secondly, we have solved a quite general control problem involving limited measurements, constrained control values, and significant amounts of process noise.

References

[1] H. Tore, *Process control in practice*. Lund: Studentlitteratur; [Bromley, Kent], Chartwell-Bratt, 1991.

[2] L. G. Barajas, E. W. Kamen, and A. Goldstein, "On-line enhancement of the stencil printing process," *Circuits Assembly*, no. March, pp. 32–36, March 2001.



(a) Without closed-loop control.

(b) With closed-loop control.

Figure 2: Process histograms.

[3] A. Lotfi and M. Howarth, *An intelligent closed-loop control of solder paste stencil printing stage of surface mount technology*. Nottingham: Nottingham Trent University. Department of Mechanical and Manufacturing Engineering, 1998.

[4] C. Chung and K. Srihari, "A "closed loop" system for the realtime control of solder paste stencil printing," *Journal Of Electronics Manufacturing*, vol. 5, no. 2, pp. 99–109, 1995.

[5] A. Johnson and A. Flori, "High density/fine feature solder paste printing," in *Electronics Assembly Process Exhibition and Conference, APEX*, San Diego, CA, 2001.

[6] R. Durairaj, T. A. Nguty, and N. N. Ekere, "Critical factors affecting paste flow during the stencil printing of solder paste," *Soldering & Surface Mount Technology*, vol. 13, no. 2, pp. 30 – 34, 2001.

[7] D. P. Bertsekas, *Dynamic programming and optimal control*, 2nd ed., ser. Athena Scientific optimization and computation series. Belmont, Mass.: Athena Scientific, 2000.

[8] H. J. Kushner and G. Yin, *Stochastic approximation algorithms and applications*. Springer, 1997.

[9] J. D. Lambert, *Numerical methods for ordinary differential systems: the initial value problem*. Wiley, 1991.

[10] K. E. Atkinson, *An introduction to numerical analysis*, 2nd ed. New York: Wiley, 1989.

[11] H. Robbins and S. Monro, "A stochastic approximation method," *Annals of Mathematical Statistics*, vol. 22, pp. 400–407, 1951.

[12] M. Csörgö and L. Horváth, *Weighted approximations in probability and statistics*, ser. Wiley series in probability and mathematical statistics. Probability and mathematical statistics. Chichester, England; New York: John Wiley & Sons, 1993.

[13] Y. Wei, S. Gelfand, and J. Krogmeier, "Noise-constrained least mean squares algorithm," *Signal Processing, IEEE Transactions on*, vol. 49, no. 9, pp. 1961–1970, 2001.

[14] Y. Maeda and Y. Kanata, "An algorithm for a least-square approximation problem of unknown systems," in *Industrial Electronics, Control and Instrumentation, 1991. Proceedings. IECON '91., 1991 International Conference on*, vol. 3, Dept. of Electr. Eng., Kansai Univ., Osaka, Japan, 1991, pp. 1881–1886.

[15] S. Kulkarni and C. Horn, "Convergence of the robbins-monro algorithm under arbitrary disturbances," in *Decision and Control, 1993., Proceedings of the 32nd IEEE Conference on*, vol. 1, Dept. of Electr. Eng., Princeton Univ., NJ, USA, 1993, pp. 537–538.

[16] S. Fujiuchi, "Fundamental study on solder paste for fine pitch soldering," in *Electronics Manufacturing Technology Symposium, 1991., Eleventh IEEE/CHMT International*, IBM Japan, Shiga, Japan, 1991, pp. 163–165.

[17] S. Venkateswaran, K. Srihari, J. Adriance, and G. Westby, "A realtime process control system for solder paste stencil printing," in *Electronics Manufacturing Technology Symposium, 1997., Twenty-First IEEE/CPMT International*, Dept. of Syst. Sci. & Ind. Eng., State Univ. of New York, Binghamton, NY, USA, 1997, pp. 62–67.

[18] L. Gopalakrishnan and K. Srihari, "Solder paste deposition through high speed stencil printing for a contract assembly environment," *Journal of Electronics Manufacturing*, vol. 8, no. 2, pp. 89–101, 1998.

[19] F. K. H. Lau and V. W. S. Yeung, "A hierarchical evaluation of the solder paste printing process," *Journal of Materials Processing Technology*, vol. 69, no. 1-3, pp. 79–89, 1997.

[20] J. Pan, G. Tonkay, R. Storer, R. Sallade, and D. Leandri, "Critical variables of solder paste stencil printing for micro-BGA and fine pitch QFP," in *Electronics Manufacturing Technology Symposium, 1999. Twenty-Fourth IEEE/CPMT*, Dept. of Ind. & Manuf. Syst. Eng., Lehigh Univ., Bethlehem, PA, USA, 1999, pp. 94–101.

[21] O. Stan and E. Kamen, "A local linearized least squares algorithm for training feedforward neural networks," *Neural Networks, IEEE Transactions on*, vol. 11, no. 2, pp. 487–495, 2000.

[22] M. Zohdy and B. Adamczyk, "Least squares approach to constrained global optimization," in *Decision and Control, 1991., Proceedings of the 30th IEEE Conference on*, Dept. of Electr. & Syst. Eng., Oakland Univ., Rochester, MI, USA, 1991, pp. 945–946.

[23] E. W. Kamen and J. Su, *Introduction to optimal estimation*, ser. Advanced textbooks in control and signal processing. London; New York: Springer, 1999.