Optimal temperature profiles for post-exposure bake of photo-resist

Anders Hansson and Stephen Boyd

Information Systems Laboratory
Stanford University
Stanford, CA 94305–9510

ABSTRACT

In this paper it is shown how to compute optimal temperature profiles for post-exposure bake of photo-resist. The profiles are optimal in the sense that the worst case non-uniformity of the dissolution rate in the photo-resist is minimized. This yields uniform development profiles, which make over-development unnecessary. The optimal strategy turns out to be to heat and cool with maximum speed. This means that the only variable that remains to optimize in each specific case is the total time of baking. This is a significant reduction in optimization complexity, and it agrees with common industrial practice.

Keywords: Photo-Resist, Baking, Temperature-Profile, Linear Programming, Critical Dimensions

1. INTRODUCTION

In manufacturing of wafers it is of critical importance that the line-width of the features etched or deposited on the wafer is the desired one. The line-width depends on the temperature during the post-exposure bake. The objective in this work is to find optimal temperature profiles with respect to the critical dimension mentioned above.

Equations are known that relate the bake temperature to the dissolution rate of the resist profile during development. By considering uniformity specifications in this parameter it is shown that good performance can be achieved. The optimal temperature profile is obtained by iterating over linear programs.

The resist considered is a deep-ultraviolet acid hardening resist (Shipley XP-8843 or SNR 248). For this case the optimal strategy turns out to be to heat and cool with maximum speed. This means that the only variable that remains to optimize in each specific case is the total time of baking. This is a significant reduction in optimization complexity, and it agrees with common industrial practice.

The paper is organized as follows. In Section 2 the models used in the optimization are presented. In Section 3 the optimization problem is posed, and in Section 4 the solution is presented and discussed. Finally, in Section 5 some concluding remarks are given.

2. MODELS

In this section the models needed to simulate and optimize the temperature profiles for post-exposure bake are presented. The manufacturing steps considered are described in Figure 1. The first step is the exposure of the photo-resist in those locations where it is to be developed away. The exposure will result in a certain concentration of acid in the photo-resist, denoted by $C_a(x,z)$ where the coordinates x and z are related to the wafer and photo-resist as in Figure 2. The second step is the baking of the photo-resist. This is the step for which optimal temperature profiles will be computed. It results in a dissolution rate $R(x,z)$. The third step is the development of the photo-resist, in which the exposed and baked resist is dissolved. The result is a resist profile which can be described as a function $z(x)$.

Send correspondence to Anders Hansson, E-mail: andersh@isl.stanford.edu

Figure 1. The manufacturing steps relevant for design of optimal temperature profiles.

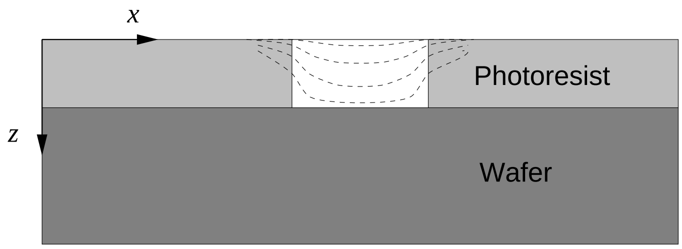

Figure 2. Coordinate-system orientation relative to the wafer and the photo-resist.

2.1. Exposure

To obtain the acid concentration $C_a(x,z)$ in the photo-resist from the exposure dose I_S it is necessary to model the illumination, the optics, the mask, and the optical absorption of photo-resist. It is assumed that the light source is circular and of uniform intensity and that the optics can be described with a coherent point spread function for clear aperture. Let

$$K(x) = \frac{J_1(|\kappa x|)}{|\kappa x|}$$

be the point spread function, and

$$J_0(x_1, x_2) = 2\frac{J_1(\kappa\sigma|x_1 - x_2|)}{\kappa\sigma|x_1 - x_2|}$$

the mutual intensity function, where $J_1(\cdot)$ is the Bessel function of first order. Then the surface light intensity is given by Hopkins equation:

$$I(x,0;t) = \int\int f(\eta_1)J_0(\eta_1,\eta_2)f^*(\eta_2)K(x-\eta_1)K^*(x-\eta_2)d\eta_1 d\eta_2$$

where

$$f(x) = \begin{cases} \sqrt{I_s}, & x \in X \\ 0, & x \notin X \end{cases}$$

is the object transmittance, i.e. the mask. The parameters are the source wave length $\lambda = 248 \times 10^{-9}$ m, the numerical aperture of the projection system $N_p = 0.42$, the coherence parameter $\sigma = 0.5$, the exposure dose $I_s = 30$ mJ/cm^2, and a normalization constant $\kappa = 2\pi N_p/\lambda$.

There are ways to compute the integral above efficiently. However, this is not the scope of this paper. Here it has been computed with brute force discretization. The resulting cross-section surface intensity for a mask with a square feature is seen in Figure 3.

The optical absorption of photo-resist can be described by the following partial differential equations[1,2]:

$$\begin{aligned} \frac{\partial I(x,z;t)}{\partial z} &= -I(x,z;t)\left(AC_a(x,z;t) + B\right) \\ \frac{\partial C_a(x,z;t)}{\partial t} &= -I(x,z;t)C_a(x,z;t)C \end{aligned}$$

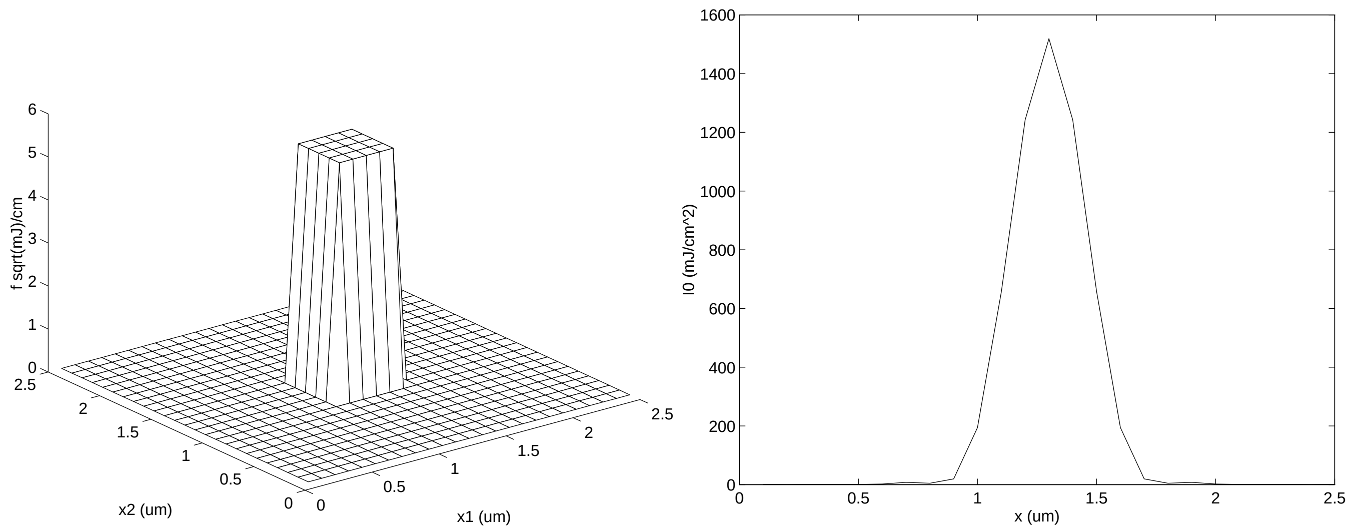

Figure 3. The mask and the resulting light surface intensity.

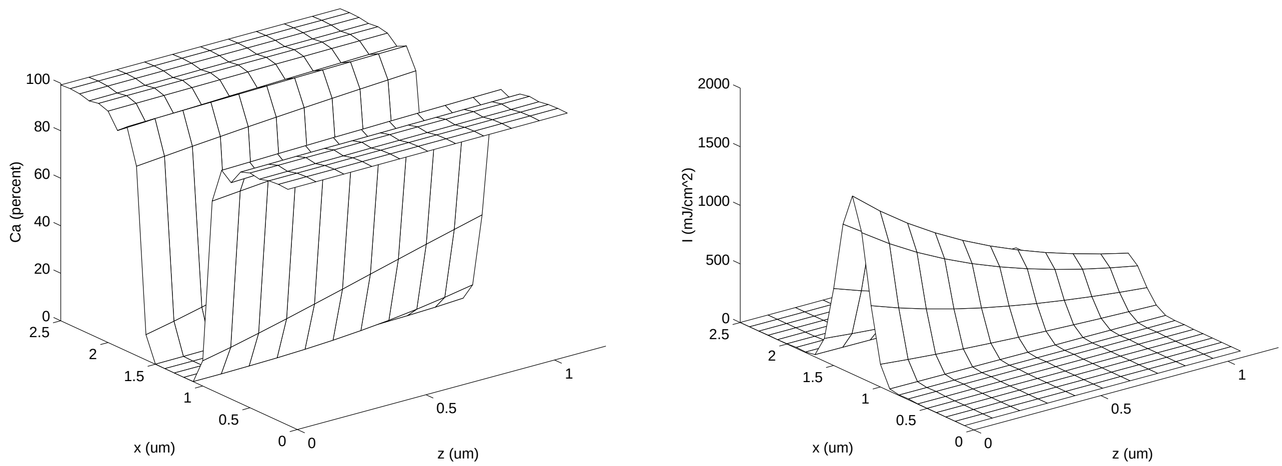

Figure 4. The resulting acid concentration and light intensity in the photo-resist after exposure.

where $I(x, z; t)$ is the light intensity in the photo-resist, and where $C_a(x, z; t)$ is the normalized acid concentration in the photo-resist. The boundary condition for the light intensity, $I(x, 0; t)$ is obtained from Hopkin's equation, and the boundary condition for the normalized acid concentration is given by $C_a(x, 0; t) = \exp(-I(x, 0; t)Ct)$. The initial conditions for the light intensity are given by $I(x, z; 0) = I(x, 0; t) \exp(-(A + B)z)$, and for the normalized acid concentration by $C_a(x, z; 0) = 1$. The parameters A, B, and C are the so-called Dill's parameters. In the computations made they were taken to be $A = -0.7\ \mu\text{m}^{-1}$, $B = 1.16\ \mu\text{m}^{-1}$, and $C = 0.0023\ \text{cm}^2/\text{mJ}$. For the same reason as mentioned before brute force discretization has been used to solve the equation. The results for an exposure time of 5 seconds is shown in Figure 4.

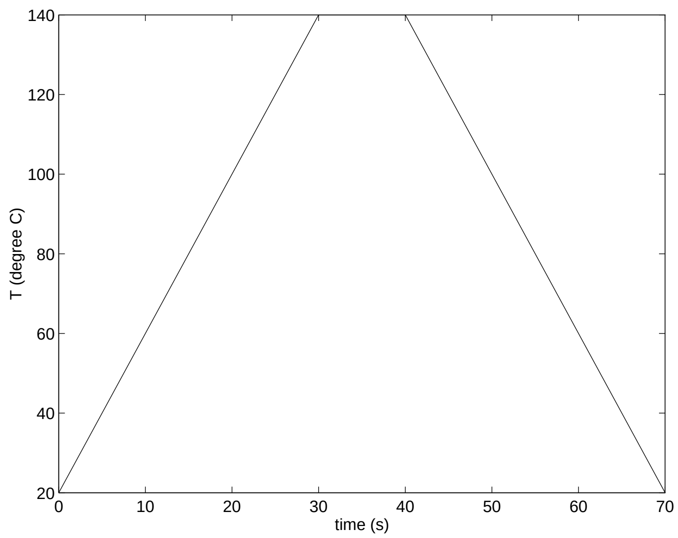

Figure 5. Temperature profile used for baking.

2.2. Bake

To model the bake of the photo-resist the local extent of cross-linking has to be related to the photo-generated acid concentration. The following equations describe the reaction[2]:

$$\begin{aligned}\frac{dC_{as}(x,z;t)}{dt} &= k_1(T)(1-C_{as}(x,z;t))C_a^m(x,z;t)\\ \frac{dC_a(x,z;t)}{dt} &= -k_2(T)C_a(x,z;t)\\ k_i(T(t)) &= k_{i0}\exp\left(-E_{ai}/kT(t)\right)\end{aligned}$$

where C_{as} is the normalized concentration of activated cross-linking sites, $T = T(t)$ is the bake temperature, k is Boltzmann's constant, $k_{10} = 6.56 \times 10^{11}$ s^{-1}, $k_{20} = 4600$ s^{-1}, $E_{a1} = 0.88$ eV, and $E_{a2} = 0.43$ eV. The initial conditions are $C_{as}(x,z;0) = 0$, and for C_a the value obtained from the exposure step. This equation can be solved with any standard ODE solver. In Figure 6 is shown the resulting acid concentration and concentration of activated cross-liking sites after baking with the temperature profile in Figure 5. A model for the dissolution rate is given by[2]:

$$R(x,z) = R_0\left(1 - p(C_{as}(x,z))/C_0\right)^{\alpha}$$

where C_{as} is the normalized concentration of activated cross-linking sites after bake, p is a 6:th order polynomial, $R_0 = 350$ Å/s, C_0 =6.3, and $\alpha = 6.5$. The dissolution rate is shown in Figure 7.

2.3. Development

A very simplified development model is given by

$$\frac{dz(x;t)}{dt} = R(x,z)$$

It can be solved with any standard ODE-solver. The resulting resist profile is shown in Figure 7.

2.4. Summary

The model described above has been validated assuming anti-reflective coating.[2] It is believed that the model of development could be more detailed. As will be seen later, this model is not of crucial importance in the optimization, since it is the dissolution rate that will be optimized.

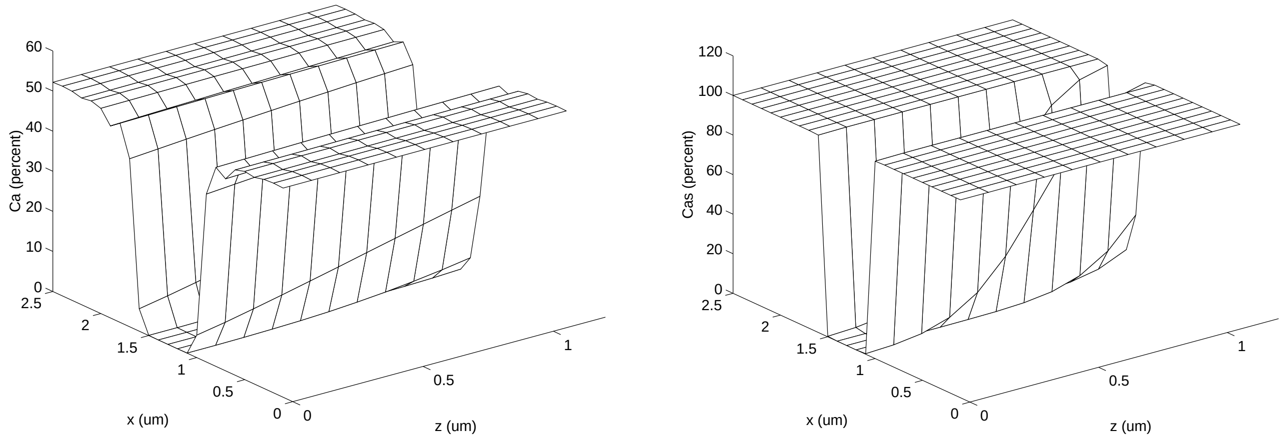

Figure 6. The resulting acid concentration and concentration of activated cross-liking sites after baking.

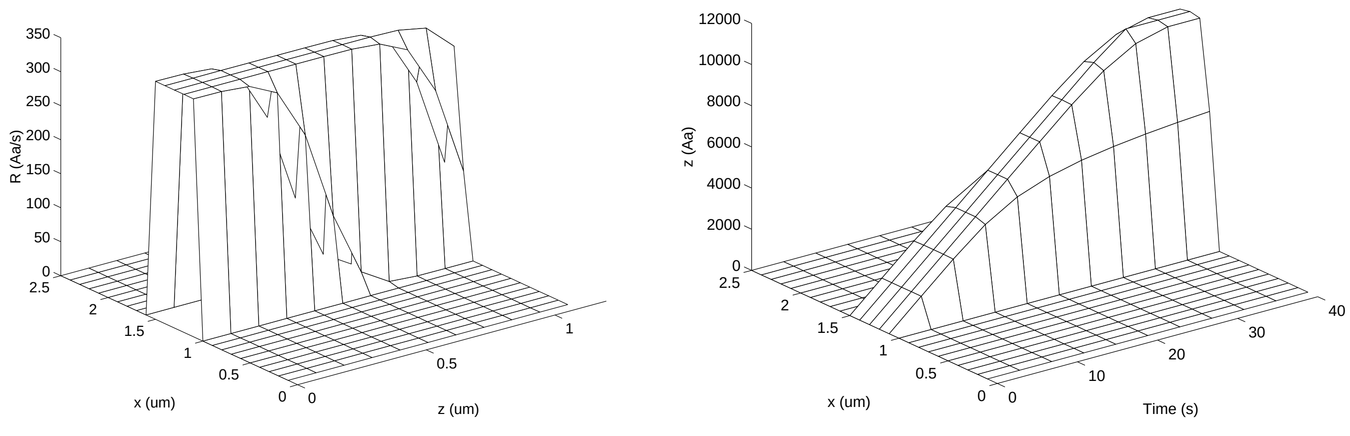

Figure 7. The dissolution rate obtained from baking and the resulting resist profile after development.

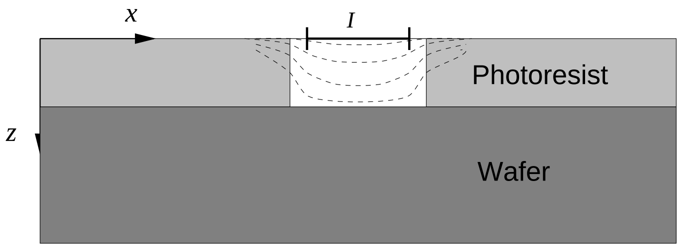

Figure 8. Definition of the interval I.

3. OPTIMIZATION PROBLEM

In this section the optimization problem will be formulated, and an algorithm for how to solve it will be proposed. The light intensity and exposure time are considered to be given. The objective is to find a temperature profile within a certain subset of admissible profiles that makes the resist profile as close as possible to a binary valued function. Unfortunately the model for development, simple as it may look, is non-linear, since the dissolution rate depends on z. However, a sufficient condition for the development profile to be binary valued is that the dissolution rate is uniform, and that is the criterion that will be considered. In case the differential equations for the bake had been linear, this problem could have been formulated as a Linear Program (LP) after discretization of the time-variable. Instead the resulting algorithm will linearize the bake differential equations around the temperature profile, discretize the resulting linear differential equation in time to get a linear difference equation. Then an LP is solved to get a modified temperature profile, and the whole procedure is repeated until it converges.

The optimization problem considered is to minimize the worst case non-uniformity of the dissolution rate:

$$\text{minimize}_{T\in\mathcal{T}} \max_{z} \left(\max_{x\in I} R(x,z) - \min_{x\in I} R(x,z) \right) \tag{1}$$

where I is defined in Figure 8 and where $\mathcal{T}$ is the set of T such that $T(t_0) = T_0$, $T(t_1) = T_1$, $T_{\min} \le T(t) \le T_{\max}$, and $S_{\min} \le dT(t)/dt \le S_{\max}$ for $t_0 \le t \le t_1$. The algorithm used can be stated as follows:

```
Compute surface intensity
Compute initial acid concentration
Initialize temperature trajectory
LOOP until converged{
   Solve bake ODE
   Linearize bake ODE
   Discretize linearized ODE
   Set up and solve LP
   Update temperature trajectory
}
```

The remaining part of this section will be devoted to the four latter steps in the loop above. The linearization of the bake ODE is straight forward, and results in a time-varying linear differential equation in new variables

$$\Delta x(t) = \begin{bmatrix} \Delta C_{as}(t) \\ \Delta C_a(t) \end{bmatrix} = \begin{bmatrix} C_{as}(t) - C^0_{as}(t) \\ C_a(t) - C^0_a(t) \end{bmatrix}; \quad \Delta R(t) = R(t) - R^0(t); \quad \Delta T(t) = T(t) - T^0(t)$$

where C^0_{as}, C^0_a, and R^0 is the solution to the nonlinear ODE for $T = T^0$. The linearized equation reads

$$\frac{d\Delta x(t)}{dt} = A(t)\Delta x(t) + B(t)\Delta T(t) \tag{2}$$

$$\Delta R(t) = C(t)\Delta x(t) \tag{3}$$

and it is an approximation to the original bake equations for small values of ΔT. The matrices $A(t)$, $B(t)$ and $C(t)$ are functions of C_{as}^0, C_a^0, and T^0 obtained by taking partial derivatives with respect to C_{as}, C_a, and T in the right hand sides of the bake-equation and the equation for the dissolution rate.[3]

The next step is to discretize this equation in time. Use the following approximations for $kh \leq t \leq (k+1)h$:

$$\begin{aligned} A(t) &= \frac{1}{2}(A(kh) + A((k+1)h)) =: \bar{A}(k) \\ B(t) &= \frac{1}{2}(B(kh) + B((k+1)h)) =: \bar{B}(k) \\ \Delta T(t) &= \Delta T(kh) + \frac{t-kh}{h}\left[\Delta T((k+1)h) - \Delta T(kh)\right] \end{aligned}$$

Consider the adjoint matrix differential equation

$$\frac{d\phi(t)}{dt} = -\phi(t)\bar{A}(k), \quad \phi(kh) = I$$

which has the solution $\phi(t) = \exp\{-\bar{A}(t-kh)\}$. Then the solution to the differential equation in (2) is given by

$$\Delta x(t) = \phi^{-1}(t)\left[\Delta x(kh) + \int_{kh}^{(k+1)h} \phi(s)B(s)\Delta T(s)ds\right], \quad kh \leq t \leq (k+1)h$$

Some computations show that the integral can be expressed as

$$\int_{kh}^{(k+1)h} \phi(s)B(s)\Delta T(s)ds = \int_0^h e^{-\bar{A}(k)s}\bar{B}(k)ds\Delta T((k+1)h) - \frac{1}{h}\int_0^h\int_0^s e^{-\bar{A}(k)\tau}\bar{B}(k)d\tau ds\Delta\tilde{T}$$

where $\Delta\tilde{T} = \Delta T((k+1)h) - \Delta T(kh)$. These integrals can be expressed in terms of a matrix exponential.[4] To this end introduce the matrix

$$D(k) = \begin{bmatrix} \bar{A}(k) & I & 0 \\ 0 & \bar{A}(k) & \bar{B}(k) \\ 0 & 0 & 0 \end{bmatrix}$$

and let

$$\begin{bmatrix} F_1(k) & G_1(k) & H_1(k) \\ 0 & F_2(k) & G_2(k) \\ 0 & 0 & F_3(k) \end{bmatrix} = e^{D(k)h}$$

Then it holds that

$$\Delta x((k+1)h) = F_1(k)\Delta x(kh) + G_2(k)\Delta T((k+1)h) - \frac{1}{h}H_1(k)\Delta\tilde{T}$$

Define

$$\begin{aligned} \Phi(k) &= \begin{bmatrix} F_1(k) & -\frac{1}{h}H_1(k) \\ 0 & 0 \end{bmatrix}; \quad \Gamma(k) = \begin{bmatrix} G_2(k) - \frac{1}{h}H_1(k) \\ 1 \end{bmatrix} \\ \Delta\xi(k) &= \begin{bmatrix} \Delta x(k) \\ \Delta T(kh) \end{bmatrix}; \quad \Delta u(k) = \Delta T((k+1)h) \\ C_\xi(k) &= \begin{bmatrix} C(kh) & 0 \end{bmatrix}; \quad \Delta y(k) = \Delta R(kh) \end{aligned}$$

Then

$$\begin{aligned} \Delta\xi(k+1) &= \Phi(k)\Delta\xi(k) + \Gamma(k)\Delta u(k) \\ \Delta y(k) &= C_\xi(k)\xi(k) \end{aligned}$$

The whole above procedure holds for any point (x, z); in particular it can be done for a grid of values. Let the values of $\Delta y(N)$ for these grid values be collected in the column matrix ΔY. It is now clear that the above difference equation can be iterated to construct a matrix Ψ such that with

$$\Delta U = \begin{bmatrix} \Delta u(0) \\ \vdots \\ \Delta u(N-1) \end{bmatrix}$$

it holds that

$$\Delta Y = \Psi \Delta U$$

which is a linear mapping from the change in the the temperature profile to the change in the dissolution rate. The true dissolution rate Y is obtained by adding the nominal Y^0 obtained by solving the bake equations for the nominal T^0, i.e. $Y = Y^0 + \Delta Y$. In the same way $U = U^0 + \Delta U$. Hence

$$Y - Y^0 = \Psi(U - U^0) \quad \text{or} \quad Y = \Psi U + Y^0 - \Psi U^0$$

So there is an affine mapping from U to Y, i.e. from the temperature profile to the dissolution rate. This is valid for small deviations from the nominal values (U^0, Y^0). This will now be used to set up the LP in standard form.

First notice that (1) is equivalent to

$$\begin{array}{ll} \text{minimize} & \max_z (\sigma(z) - \gamma(z)) \\ \text{subject to} & \gamma(z) \leq R(x,z) \leq \sigma(z), \quad x \in I \\ & T \in \mathcal{T} \end{array}$$

where $\sigma(z)$ and $\gamma(z)$ are functions of z. This can be further rewritten as

$$\begin{array}{ll} \text{minimize} & \tau \\ \text{subject to} & \sigma(z) - \gamma(z) \leq \tau, \forall z \\ & \gamma(z) \leq R(x,z) \leq \sigma(z), \quad x \in I \\ & T \in \mathcal{T} \end{array}$$

with τ as a new variable. If now only values of (x, z) on the grid are considered, γ is defined to be the column matrix of values of $\gamma(z)$ on this grid, σ is defined similarly, and if Ψ_I contains the rows in Ψ which give values of ΔY defined for $x \in I$, then the following LP is obtained:

$$\begin{array}{ll} \text{minimize} & \tau \\ \text{subject to} & \sigma - \gamma \leq \tau \mathbf{1} \\ & \gamma \leq \Psi_I U + Y_I^0 - \Psi_I U^0 \leq \sigma \\ & U \in \mathcal{U} \end{array}$$

where $\mathcal{U}$ defines the set of admissible U. In order to make sure that the linearization is a good approximation to the original problem the following constraint is a part of the definition of $\mathcal{U}$:

$$\Delta U_{\min} \leq U - U^0 \leq \Delta U_{\max}$$

Notice that all inequalities for column matrices defined above are to be interpreted as component-wise inequalities. Also notice that it might be better from a numerical point of view to replace the optimization variable U with ΔU.

4. EXAMPLES

In this section two examples will be presented. The only difference between them is the sample interval, which is 10 s in the first example and 5 s in the second example. In both examples the following values were used to constrain the temperature profiles: $T_0 = T_1 = T_{\min} = 20°\text{C}$, $T_{\max} = 200°\text{C}$, $S_{\min} = -5°\text{C/s}$, and $S_{\max} = 5°\text{C/s}$. The results are shown in Figures 9 and 10. It is seen that a uniform dissolution rate yields a binary valued development profile, which makes over-development unnecessary. Also the optimal strategy is to heat and cool with maximum speed. This means that the only variable that remains to optimize in each specific case is the total time of baking. This is a significant reduction in optimization complexity, and it agrees with common industrial practice. The optimal strategy reduced the development time by about 20% and the bake time by about 30% in these examples.

The most time consuming step in each iteration was to discretize the linearized differential equation in time. The reason for this is that a large matrix exponential has to be computed for each discrete time step. This is computationally very expensive. Most likely it should be possible to speed this up by designing a special purpose routine for the specific application of this paper.

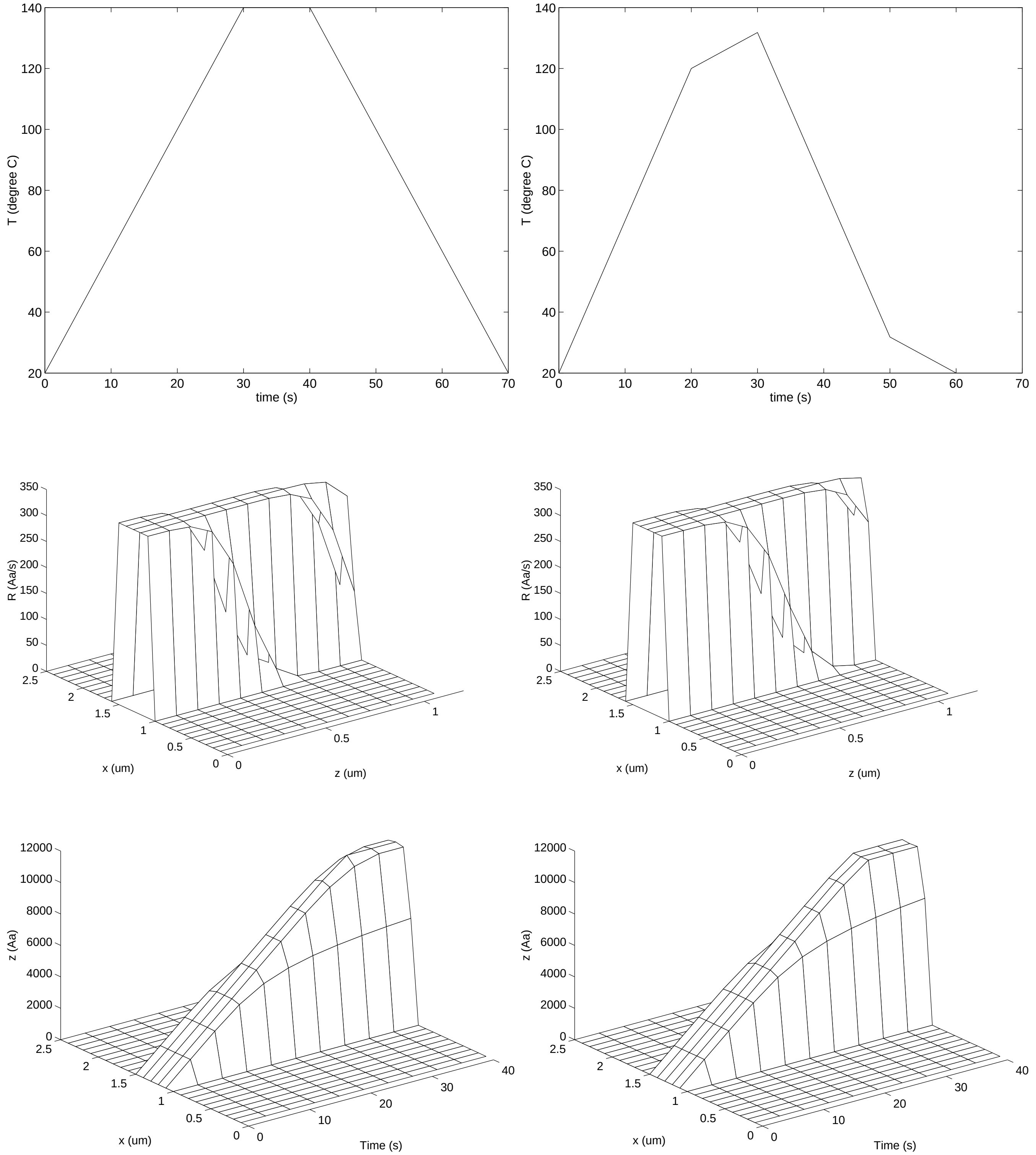

Figure 9. The initial and optimal temperature profiles, dissolution rates, and resist profiles for the case when the sample interval is 10 s.

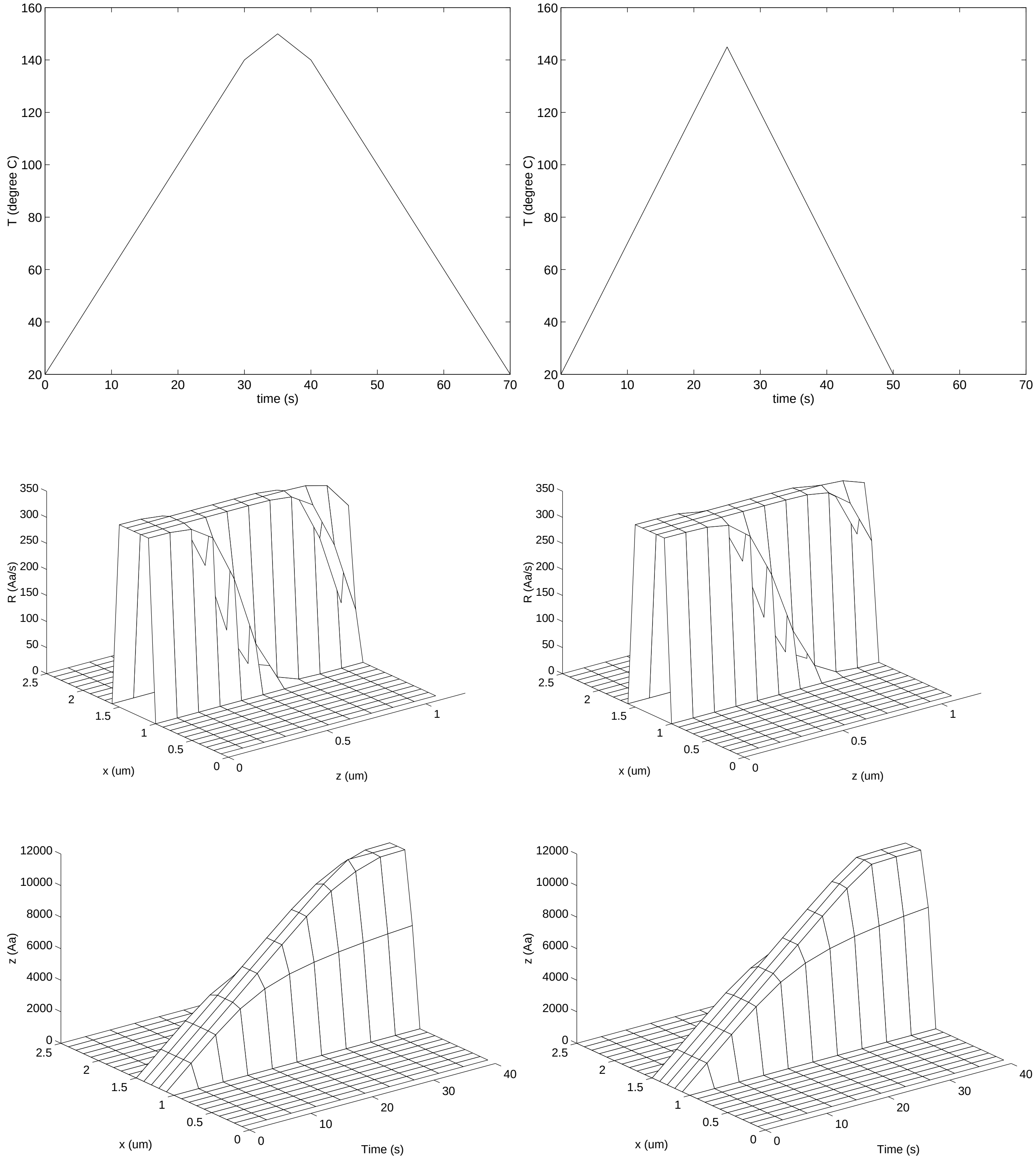

Figure 10. The initial and optimal temperature profiles, dissolution rates, and resist profiles for the case when the sample interval is 5 s.

5. CONCLUSIONS

In this paper it has been investigated how the resist development non-uniformity depends on the temperature profile during post-exposure bake. It has been seen that a uniform dissolution rate implies a binary-valued resist profile, and that a uniform dissolution rate can be obtained by solving a sequence of LPs. The optimal strategy for the examples considered was to heat and cool with maximum speed. Hence it is sufficient to optimize over the bake time to get the optimal solution. This is a one-parameter optimization, and hence this is a significant reduction in complexity. The optimal strategy reduced the bake time by about 30%. Shorter baking time, down to the optimal value, also results in shorter development time. Hence optimal temperature profiles for post-exposure bake increases throughput in the manufacturing of integrated circuits.

ACKNOWLEDGMENTS

Anders Hansson was supported by the Swedish Research Council for Engineering Sciences (under 95–838) and MURI (under F49620-95-1-0525) and Stephen Boyd was supported in part by AFOSR (under F49620-95-1-0318), NSF (under ECS-9222391 and EEC-9420565), and MURI (under F49620-95-1-0525).

REFERENCES

1. F. H. Dill, W. P. Hornberger, P. S. Hauge, and J. M. Shaw, "Characterization of positive photoresist," *IEEE Transactions on Electron Devices* **ED–22**(7), pp. 445–452, 1975.
2. R. A. Ferguson, J. M. Hutchinson, C. A. Spence, and A. R. Neureuther, "Modeling and simulation of a deep-ultraviolet acid hardening resist," *J. Vac. Sci Technol.* **B 8**(6), pp. 1423–1427, 1990.
3. R. W. Brockett, *Finite Dimensional Linear Systems*, Wiley, New Yourk, 1970.
4. C. F. van Loan, "Computing integrals involving the matrix exponential," *IEEE Transactions on Automatic Control* **AC-23**(3), pp. 395–404, 1978.