

Temperature Sensor Distribution, Measurement Uncertainty, and Data Interpretation for Microprocessor Hotspots

Josef Miler¹, Keivan Eteessam-Yazdani², Mehdi Asheghi¹, Maxat Touzelbaev³, and Kenneth E.
Goodson¹

¹Dept. of Mechanical Eng., Stanford Univ., Stanford, California 94305

²Broadcom Corporation., Santa Clara, California 95054

³Glew Engineering Consulting Inc., Mountain View, California 94040

9 March 2012

Abstract:

Microprocessor hotspots are a major reliability concern with heat fluxes as much as 20 times greater than those found elsewhere on the chip. Chip hotspots also augment thermo-mechanical stress at chip-package interfaces which can lead to failure during cycling. Because highly localized, transient chip cooling is both technically challenging and costly, chip manufacturers are using dynamic thermal management (DTM) techniques that reduce hotspots by throttling chip power. While much attention has focused on methods for throttling power, relatively little research has considered the uncertainty inherent in measuring hotspots. The current work introduces a method to determine the accuracy and resolution at which the hotspot heat flux profile can be measured using distributed temperature sensors. The model is based on a novel, computationally-efficient, inverse heat transfer solution. The uncertainties in the hotspot location and intensity are computed for randomized chip heat flux profiles for varying sensor spacing, sensor vertical proximity, sensor error, and chip thermal properties. For certain cases the inverse solution method decreases mean absolute error in the heat flux profile by more than 30%. These results and simulation methods can be used to determine the optimal spacing of distributed temperature sensor arrays for hotspot management in chips.

Nomenclature

Variables

a	Width of chip, m
b	Length of chip, m
f_x	Sensor spatial frequency in the x-direction, m^{-1}
f_y	Sensor spatial frequency in the y-direction, m^{-1}
h	Convective heat transfer coefficient, $W/m^2 \cdot K$
k	Thermal conductivity, $W/m \cdot K$
N	Number of random heat flux profiles tested
Q''	Heat flux, W/m^2
R''_{th}	Chip vertical thermal resistance per unit area, $W/m^2 \cdot K$
t_0	Thickness of chip, m
T	Temperature, K

Subscripts

c	Circuit level
s	Sensor level
l	Low resolution
f	Full resolution
e	Includes sensor error
i	Index in x-direction spatial domain
j	Index in y-direction spatial domain

I. Introduction

As microprocessor manufacturers have adopted multi-core circuit architectures, the detection and management of temporal hotspots have become increasingly important for chip reliability and performance. While much attention has been given to increases in the overall chip power, hotspot heat fluxes are increasing even more rapidly for many applications [1]. Active portions of a microprocessor can produce as much as 20 times as much heat as inactive regions [2]. These high heat fluxes can cause elevated junction temperatures leading to electromigration and subsequent circuit failure. Furthermore, temperature non-uniformities in the chip can cause severe thermo-mechanical stress on the package leading to system failure. These challenges will be exacerbated in future processors that are expected to include many more processor cores integrated in three-dimensional geometries.

To date, chip cooling alone does not seem capable of addressing these challenges. Most cooling solutions are best suited to address relatively slow thermal phenomena occurring over large regions of the chip. It is especially difficult to directly address highly localized, dynamic hotspots with cooling solutions implemented in chip packaging. Thermal engineers are forced to overdesign the cooling solution to satisfy worse-case scenario conditions for a hotspot region. This can be both difficult and expensive, particularly because the cost of cooling solutions increases rapidly as a function of maximum local heat flux [1]. Various methods of dynamic, localized cooling (e.g. use of Peltier devices) are being investigated to address these difficult thermal requirements, but none have been adopted to date

An alternative overall approach to managing chip hotspots is to regulate the chip power output to maintain device temperature within specified limits. Such techniques are referred to as dynamic

thermal management (DTM) and have been a subject of intense investigation since being introduced by Brooks and Martonosi [3].

All DTM techniques fundamentally involve two steps: (1) interpreting temperature data from the chip and (2) responding to that data by reducing power. The majority of research has focused on the latter problem for DTM, specifically on finding innovative ways to locally regulate chip power. Proposed DTM techniques involve clock gating [4], Dynamic Frequency Control [5], DVFS [6], SMT thread reduction [7], and activity migration [2]. Much less attention has been given to designing temperature sensor arrays and interpreting the resulting thermal signals. Two important sources of uncertainty need to be considered for DTM applications. First, the thermal sensors used for DTM feedback are subject to error. Most DTM studies do not consider the effect of this error and thus provide overly optimistic results. Skadron et al [8] demonstrated that sensor error can cause significant performance reductions due to incorrect DTM triggering and reduced DTM threshold levels.

Discussions of uncertainty in DTM studies are typically limited to sensor error, but additional attention should be paid to the uncertainty caused by sensor placement. Because thermal sensors are not necessarily located at the chip hotspot, a DTM scheme must account for the temperature difference between the sensor location and the actual hotspot. Skadron et al [9] used an estimated spreading factor within a core to try to account for this discrepancy as an additional source of error. In their study, the spreading factor contributed an additional 2°C error in the temperature signal. To attempt to account for uncertainty in hotspot location and intensity, DTM methods are currently designed to be conservative, which causes reduced system performance.

To reduce the uncertainties associated with thermal sensing for DTM, a challenging optimization problem must be considered. Circuit designs with high circuit density but low sensor density suffer from increased uncertainty in the thermal profile. Increased uncertainty about hotspot location and magnitude requires more cautious DTM control algorithms, which diminishes performance metrics. Increasing sensor resolution improves DTM control algorithms but also reduces circuit density, ultimately reducing computational power. An optimization approach is required to find a design that maximizes computational power while maintaining the chip in reliable operating conditions.

This study endeavors to help address this challenging optimization problem by quantifying the uncertainty that should be accounted for in a DTM scheme given a particular thermal sensor array. We consider the generalized case of a grid-array of thermal sensors located some distance above an arbitrary heat flux profile. In order to better represent real applications, the thermal sensors are not necessarily located directly above known heat flux peaks. The chip heat flux profile is considered unknown, and the purpose of the thermal sensor array is to detect the regions of the chip that require dynamic power control.

We introduce a novel, computationally efficient, inverse heat transfer solution method and determine the accuracy to which it resolves the underlying heat flux profile. We consider cases with varying numbers of thermal sensors located with varying proximity to the circuitry level of the chip. Sensor error is also introduced to determine its effect on the estimated heat flux profile. For certain cases, the inverse solution method is shown to be susceptible to temperature sensor error. The results of these tests are compared to the uncertainty that results from treating the unprocessed thermal signal as a representation of the heat flux profile.

The approach taken here also has implications for the use of discrete thermal data in resolving the source of a hotspot. DTM schemes need not consider this uncertainty because the standard response to a hotspot is to throttle all activity in the vicinity. In chip development and production, however, thermal measurements are used to characterize the power distribution of the circuit design. For these tests, high resolution thermometry can be used (e.g. infrared microscopy [10]). The maximum spatial resolution at which these techniques can resolve neighboring hotspots is dictated by the resolution of the applied thermometry technique, the extent of thermal spreading in the chip, and the measurement error. The present study simulates the case of distinguishing two similar hotspot sources using discrete temperature measurements. For a given measurement error and chip configuration, there is a minimum spatial sampling frequency required to correctly resolve the source of a hotspot.

Section II of this paper presents the overall methodology used to simulate chip heat flux profiles and determine the uncertainty associated with a particular thermal sensor array. Section III presents the inverse heat transfer solution method derived for this study. The uncertainties in the heat flux profile associated with direct temperature interpretation and inverse solution method are presented in Section IV. Section V provides concluding remarks.

II. Methodology

A. Overall Simulation Methodology

The present study is based on a simplified conduction model for the chip. Figure 1 shows the model geometry. The chip is modeled as an isotropic, single-layer structure. The isotropic condition can be relaxed by transformation of the thermal conductivity and chip thickness [11]. The boundary condition on the top surface is convective heat transfer with a uniform heat

transfer coefficient. The boundary conditions on the four sidewalls are adiabatic. On the bottom surface, an arbitrary heat flux profile boundary condition is applied. The chip is 1 cm by 1cm and its thickness and thermal conductivity is varied in the simulations. The system operates in steady-state. This simplified model of the chip facilitates the generalized simulation methodology taken in this study which would be impractical with a highly discretized chip model.

Figure 2 shows the four main steps involved in the overall simulation methodology. The simulation begins by defining the geometry and system parameters and generating a randomized heat flux profile. The forward solution method is used to resolve the sensor-level, full-resolution temperature profile, $T_{s,f}$ (Figure 2b), based on the circuit-level, full resolution heat flux profile, $Q''_{c,f}$ (Figure 2a). A set of low-resolution temperature profiles, $T_{s,l}$ (Figure 2c), is created by interpolating the full-resolution temperature profile, $T_{s,f}$, at various spatial frequencies. Each low-resolution temperature profile represents the temperature profile that would be measured by a temperature sensor array of a particular spatial frequency. For example, for a temperature sensor spatial sampling frequency of 1000 m^{-1} (equivalent to nominal sensor spacing of 1mm), the low-resolution temperature profile, $T_{s,l}$, is a 10x10 grid on a 1cm by 1cm chip.

Random error is added to the low-resolution temperature profile, $T_{s,f}$, to simulate the measurement error introduced by real temperature sensors. The sensor error, T_{error} , is normally-distributed about the interpolated temperature value with a standard deviation that is specified relative to the maximum interpolated temperature. The sensor error at each index is calculated as:

$$T_{error}(i, j) = \sigma_{relative} T_{s,l,max} \Gamma \quad (1)$$

where $\sigma_{relative}$ is the standard deviation of the relative sensor error, $T_{s,l,max}$ is the maximum measured temperature, and Γ is a random number with a mean value of zero and a standard deviation of unity.

This study shows the results for standard deviations in the relative sensor error of 0, 0.5, and 1 percent. The case of 0 percent standard deviation in the relative sensor error is equivalent to no measurement error.

The sensor-level, low-resolution temperature profile with error, $T_{s,l,e} = T_{s,l} + T_{error}$, is used to calculate the circuit-level, heat flux profile, $Q''_{c,l,e}$ (Figure 2d), using a spatial sampling frequency domain, inverse heat transfer solution, described in detail in the next section. Because the inputted temperature profile is low resolution, the resulting heat flux profile, $Q''_{c,l,e}$, is also low resolution. To calculate the error resulting from the solution method, the low-resolution heat flux profile is interpolated to full resolution. The mean absolute error (MAE) is calculated by finding the difference between the correct profile and the calculated heat flux profile:

$$MAE = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |Q''_{c,f}(i, j) - Q''_{c,l,e}(i, j)| \quad (2)$$

where N_x and N_y are the total number of indices in the x and y directions, respectively. The mean absolute error is normalized by the average heat flux:

$$\text{Normalized MAE} = \frac{\frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |Q''_{c,f}(i,j) - Q''_{c,l,e}(i,j)|}{\frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} Q''_{c,f}(i,j)} \quad (3)$$

$$\text{Normalized MAE} = \frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |Q''_{c,f}(i,j) - Q''_{c,l,e}(i,j)|}{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} Q''_{c,f}(i,j)} \quad (4)$$

Figure 3 provides a block diagram of the simulation procedure. The procedure is repeated for numerous randomly-generated heat flux profiles and the results are averaged. The average MAE is plotted against the thermal sensor spatial sampling frequency.

In practice, an inverse heat transfer technique is not always used to interpret measured temperature profiles. Instead, the measured temperature profile is assumed to be representative of the chip heat flux profile. This technique is equivalent to treating the measured temperature profile as directly proportional to the heat flux profile:

$$Q''_c(i,j) = \frac{T_s(i,j)}{R''_{th}} \quad (5)$$

where R''_{th} is the chip vertical thermal resistance for unit area:

$$R''_{th} = \frac{t_0}{k} \quad (6)$$

This simplification results in additional uncertainty in the heat flux profile, the magnitude of which depends on the chip properties and boundary conditions. In this paper, this approach is

referred to as “direct interpretation” of the temperature profile and is compared to the inverse solution method in the results section.

B. Randomization of Heat Flux Profiles

The uncertainty in the calculated heat flux is dependent on the characteristics of the heat flux profile. Simple, well-spaced heat flux profiles are easier to resolve than overlapping, complicated heat flux profiles. To represent the most general case, the simulation is conducted over a set of heat flux profiles that contain varying degrees of complexities. The heat flux profiles are randomly generated to include between 1 and 15 hotspots which can vary in laterals dimension between 273 μm (equivalent to 7 grid cells) and 4.18 mm (equivalent to 107 grid cells). For reference, the chip is 1 by 1cm. The hotspots are created with soft edges; the edge of the hotspot spans 156 μm (equivalent to 4 grid cells) and has a linear slope from the value of the background heat flux to value of the hotspot heat flux. The background heat flux is 1 W/cm^2 and the maximum possible hotspot heat flux is 320 W/cm^2 . Hotspots are permitted to overlap with each other but not with the edge of the chip. For the first set of simulations, the hotspots have a random heat flux value between the background and the maximum heat flux. This is referred to as “variable heat flux”. For the second set of simulations, all hotspots have the maximum heat flux, referred to as “binary heat flux”. The case of binary heat flux represents a core that is either active or inactive. The case of variable heat flux represents a core for which the amount of activity is unknown. Since the variable heat flux case is most challenging from an uncertainty perspective, only select resulted are presented for binary heat flux cases.

The key result of each simulation is the mean absolute error (MAE) in the calculated heat flux profile. Because conduction through the chip is linear, the results are generalized by normalizing

the error in the heat flux profile by the input heat flux profile. Thus only the relative magnitude of the heat flux as compared to the background heat flux is relevant for consideration.

C. Resolution Study

A second study was conducted to quantify the ability of the inverse solution method to resolve a single hotspot from a group of neighboring hotspots. Two circuit-level heat flux profiles are created; the first heat flux profile consists of a single hotspot in the center while the second heat flux profile consists of 9 closely packed hotspots in the center. The average heat flux is the same in both cases. Figure 4 shows the two heat flux profiles. The circuit-level temperature profile resulting from the single-hotspot heat flux profile is calculated using the forward solution. The temperature profile is sampled at reduced spatial sampling frequency to simulate the signal from a thermal sensor array, as before.

Each solution method is used to deduce which of two possible heat flux profiles yielded the measured temperature profile. To do so, the inverse solution method is used to calculate the circuit-level heat flux profile. The results are compared to the two possible inputted heat flux profiles by calculating the mean absolute error. The profile resulting in the lower MAE represents the solution chosen by the inverse solution method. For example, if the MAE between the calculated heat flux profile and the single-hotspot heat flux profile is lower than the MAE between the calculated heat flux profile and the multi-hotspot heat flux profile, the inverse solution method chooses the single-hotspot heat flux profile. If the choice correctly corresponds to the actual inputted heat flux profile, the inverse solution method is correct. This procedure is conducted for all sensor spatial frequencies, and is also conducted for the direct interpretation method.

III. Introduction to Inverse Heat Transfer Solution in Spatial Frequency Domain

A. Inverse Heat Transfer Solution Method

To conduct the forward and inverse solutions needed for the overall simulation methodology, an analytical, spatial-frequency domain heat transfer analysis has been developed. This approach is more computationally efficient than finite-difference methods and thus facilitates rapid multi-parameter design optimization and possible integration into DTM schemes.

The thermal profile in the model geometry is defined by the heat diffusion equation. For each layer in the stack, the solution to the heat diffusion equation is given by:

$$\begin{aligned}
 T(x, y, z) = & A_0 + B_0 z \\
 & + \sum_{m=1}^{\infty} [A_m \cosh(\lambda_m z) + B_m \sinh(\lambda_m z)] \cos(\lambda_m x) \\
 & + \sum_{n=1}^{\infty} [A_n \cosh(\gamma_n z) + B_n \sinh(\gamma_n z)] \cos(\gamma_n y) \\
 & + \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} [A_{mn} \cosh(\beta_{mn} z) \\
 & + B_{mn} \sinh(\beta_{mn} z)] \cos(\lambda_m x) \cos(\gamma_n y)
 \end{aligned} \tag{7}$$

where:

$$\beta_{mn} = \sqrt{\lambda_m^2 + \gamma_n^2} \tag{8}$$

$$\lambda_m = \frac{m\pi}{a} \quad (9)$$

$$\gamma_n = \frac{n\pi}{b} \quad (10)$$

For the boundary conditions imposed in this model, Etessam-Yazdani [11] demonstrated a technique of representing this conduction problem as a two-port terminal network. The technique has been shown to be both accurate and fast for the forward heat transfer solution [12] and is adapted in this study for the inverse problem.

Figure 5 presents a schematic of the two-port terminal network for this system. The two-dimensional Fourier transforms of the heat flux profiles at the circuit and sensor levels are Q_c'' and Q_s'' , respectively. Similarly, T_c and T_s are the two-dimensional Fourier transforms of the temperature profiles at the circuit and sensor levels, respectively. The matrix A is a 2x2 matrix that relates T_c and Q_c'' to T_s and Q_s'' :

$$\begin{bmatrix} T_c(f_x, f_y) \\ Q_c''(f_x, f_y) \end{bmatrix} = A(f_x, f_y) \begin{bmatrix} T_s(f_x, f_y) \\ Q_s''(f_x, f_y) \end{bmatrix} \quad (11)$$

For radial spatial frequency $f_r > 0$:

$$A(f_x, f_y) = \begin{bmatrix} \cosh(2\pi f_r t_o) & \frac{\sinh(2\pi f_r t_o)}{2\pi f_r k} \\ 2\pi f_r k \sinh(2\pi f_r t_o) & \cosh(2\pi f_r t_o) \end{bmatrix} \quad (12)$$

And for $f_r = 0$:

$$A(f_x, f_y) = \begin{bmatrix} 1 & t_o/k \\ 0 & 1 \end{bmatrix} \quad (13)$$

where the radial spatial frequency f_r is defined as:

$$f_r = \sqrt{f_x^2 + f_y^2} \quad (14)$$

Further details on the derivation of the two-port terminal analysis are provided in [11].

Etessam-Yazdani et al [11] used the two-port terminal analysis to solve for the temperature as a function of the heat flux on the same level of the geometry, which represents the forward solution. In this study, the solution was modified to determine the heat flux profile on the circuit plane, Q_c'' , using the temperature profile on the sensor plane, T_s , which represents the inverse solution. From the two-port terminal analysis, the equation for Q_c'' is:

$$Q_c'' = A_{21}T_s + A_{22}Q_s'' \quad (15)$$

Applying the top boundary condition, $Q_s = hT_s$, and substituting the appropriate values of A_{ij} , the result for cases where $f_r > 0$ is:

$$Q_c = (2\pi f_r k \sinh(2\pi f_r t_o) + h * \cosh(2\pi f_r t_o))T_s \quad (16)$$

For which the inverse solution transfer function $G_{inv}(f_r)$ can be defined such that:

$$Q_c = G_{inv}(f_r) * T_s \quad (17)$$

and

$$G_{inv}(f_r) = 2\pi f_r k \sinh(2\pi f_r t_o) + h * \cosh(2\pi f_r t_o) \quad (18)$$

For cases where $f_r = 0$, the transfer function reduces to equal the heat transfer coefficient, h , and the equation is given as $Q_c = hT_s$.

B. High-Frequency Filtering

A filtering technique based on the forward solution transfer function is employed to reduce error in the inverse solution method. As shown by [13], the forward solution to the conduction problem yields a transfer function in the frequency domain that acts as a low pass filter. Physically this represents the attenuation of high spatial frequency components of the thermal signal via heat spreading in the chip.

The inverse transfer function has the form of a high-pass filter, as shown in Figure 6. The minimum of the transfer function occurs at $f_r = 0$ and increases rapidly as a function of f_r , thus amplifying the high frequency components of the temperature profile. The components of the temperature profile that are greater than the -3dB frequency of the forward solution transfer function, however, represent sensor noise. A filtering method has been developed to prevent this noise from propagating to the calculated heat flux profile. A low-pass filter is applied to the inverse transfer function with a filter cut-off frequency at the -3dB frequency of the forward solution transfer function. The filter has a soft roll-off. Figure 6c shows the filtered transfer function. This filtering technique dramatically improves the performance of the inverse solution method by decreasing sensitivity to high-frequency noise.

C. Solution Validation

The solution method was validated by comparison to COMSOL Multiphysics software using representative simulation parameters. The heat transfer coefficient was $10 \text{ W/m}^2\text{-K}$ and the thermal conductivity was 148 W/m-K . The simulated chip was 1cm by 1cm in lateral dimensions and 100 microns in thickness. A representative heat flux was applied in the COMSOL model and the temperature profile was resolved. The temperature profile was used as an input to the inverse solution method and the applied heat flux was calculated. The calculated heat flux matched the COMSOL heat flux at greater than 0.01% accuracy.

Additional testing was conducted to ensure the results for average heat flux error are independent of the number of random heat flux maps tested, N . Figure 7 shows the results for varying number of randomly generated heat flux maps for both varying heat flux and binary heat flux. The results are shown to be N -independent (i.e. independent of the number of random heat flux profiles) after 50 randomly generated heat maps. For all of the reported results, data was averaged for 50 heat maps ($N = 50$).

IV. Results

Figure 8 shows a representative distribution of mean absolute error for 50 randomized heat flux distributions. Results are reported for the case of variable heat flux and binary heat flux. These results provide a basis for understanding the effects of sensor spatial frequency on the calculated heat flux profile. Simulation parameters are typical of chip applications: the distance from the sensor array is $100 \text{ }\mu\text{m}$, the conductivity is 148 W/m-K and the heat transfer coefficient is $10,000 \text{ W/m}^2\text{-K}$. The sensor error is zero for this case. The mean value (shown in bold black) follows the expected trend of increased accuracy at higher spatial sampling frequency. A sampling frequency

of 2000 m^{-1} (approximately 500um sensor spacing) is required to achieve an average mean absolute error (MAE) below 25% for the variable heat flux case. At lower resolutions, the average MAE is dramatically higher. Significant deviations from the mean value are caused by variations between the randomized heat flux profiles.

The average MAE error is dependent on the heat transfer coefficient, the sensor error, and the proximity between thermal sensors and the circuit level. These effects are discussed in more details below. For clarity, only the average MAE is shown. The solid curves and dotted curves represent the average MAE for the inverse solution method and the direct temperature interpretation method, respectively.

The average MAE of the inverse solution method is dependent on whether the inputted heat flux profile is binary or variable. Figure 9 shows an approximately 65% drop in average MAE if the input heat flux is binary rather than variable. Since the heat flux cannot always be assumed to be binary, the remaining plots show results for variable heat flux.

Figure 10 shows the performance of the inverse solution method for varying heat transfer coefficients for variable heat flux with no sensor error. As expected, the average MAE for both methods is reduced by increasing heat transfer coefficient values. The direct interpretation method performs poorly at low heat transfer coefficients but makes significant improvements as the heat transfer coefficient is increased. The inverse method produces significantly lower average MAE and is less sensitive to changes in the heat transfer coefficient.

Figure 11 illustrates the difficulty of calculating the heat flux profile from temperature profiles containing sensor error. For the ideal case of zero sensor error, the inverse solution method outperforms the direct method by up to 50% MAE for variable heat flux. However, measurement

error causes the inverse method to diverge from the solution. For a case of 0.5% measurement error, the inverse solution is slightly better than the direct method for spatial frequencies up to about 3000 m^{-1} , at which point it diverges rapidly. For the case of 1% standard deviation in the sensor error, the direct interpretation method is superior for sensor spatial frequencies greater than 2000 m^{-1} . Similar trends are observed for the case of binary heat flux profiles as well.

Figure 12 presents the effect of vertical proximity between the sensor level and the circuit level. Average MAE results are shown for vertical distances between $1 \mu\text{m}$ and 1 mm for variable heat flux profiles. As the vertical proximity is reduced, modest improvements in MAE are observed for both the inverse and direct interpretation techniques with the exception of the $1 \mu\text{m}$ case where improvements in the direct interpretation method are approximately 0.8 normalized averaged MAE. For the extreme case of $1 \mu\text{m}$ of vertical proximity, the inverse and direct interpretation methods are comparable, but for all other cases the inverse solution significantly outperforms the direct interpretation method.

Figures 13 and 14 show the performance of the inverse solution method in resolving neighboring hotspots. The figures show the minimum sensor spatial frequency required to correctly differentiate between a single hotspot and a group of equivalent neighboring hotspots. The results are presented as a function of vertical proximity between the distributed thermal sensor array and the circuit plane, and a moving-average smoothing function is applied to remove discretization artifacts. The gray region of the plot shows the domain in which the inverse solution can correctly identify the underlying heat flux profile. A relatively low sensor spatial frequency is adequate when positioned in close proximity to the hotspot. Increasing the separation between the sensor array and the hotspot requires an increase in the sensor spatial frequency. Figures 13 and 14 show results for convective heat transfer coefficients of 10,000 and

50,000 W/m²-K, respectively. For a convective heat transfer coefficients of 10,000 W/m²-K at distances greater than approximately 240 μ m, the inverse solution method is unable to resolve the hotspot. Figure 14 shows that the limit of the inverse solution can be extended by increasing the heat transfer coefficient. For this case, the inverse solution method produces the correct results up to 300 μ m. For all cases shown, the direct interpretation method failed to correctly identify the single hotspot. The inverse technique is shown to be superior to the direct interpretation method for resolving neighboring hotspots. These results provide insight into the optimization of sensor vertical proximity and sensor spatial frequency for resolving neighboring hotspots.

IV. Summary and Concluding Remarks

This study investigates uncertainty and error propagation in distributed thermal sensor arrays in microprocessors. A novel, inverse heat transfer solution methodology is developed to provide a computationally efficient method for determining the heat flux profile at a remote level in a chip. The inverse solution method is used to determine the expected mean absolute error of the calculated heat flux profile in a chip. Several key conclusions are drawn.

- For systems with relatively low sensor spatial frequency such as typical microprocessors, large improvements in the accuracy of the calculated heat flux can be made by making relatively small improvements in the resolution of the sensor array. As the sensor array increases resolution, the uncertainty in the calculated heat flux is much reduced.
- For cases of very low sensor error, the proposed inverse solution technique more accurately calculates the heat flux profile than direct interpretation of the temperature profile.

- Depending on the system configuration and the magnitude of the sensor error, the inverse solution method can be inaccurate. This inaccuracy is mitigated by the proposed filtering method, but nonetheless represents a fundamental limitation of this technique.
- Direct interpretation of the temperature signal is shown to result in significant error in the calculated heat flux profile. Accounting for these errors in DTM techniques causes decreased computational performance and should therefore be considered during overall system design.

These conclusions regarding the nature of error propagation from distributed thermal sensor arrays can provide a basis for considering the difficult system-level optimization required for integrated circuit design. Sensor error, sensor spatial frequency, proximity between a sensor array and hotspots, and signal processing all affect hotspot uncertainty as well as circuit design. Each of these parameters can help improve DTM accuracy but can also pose costs for the performance of the circuit. Careful optimization of these parameters is necessary to maximize computational performance while ensuring reliable thermal conditions.

Acknowledgements

The authors gratefully acknowledge support from Advanced Micro Devices (AMD) Inc. as part of the Semiconductor Research Consortium (SRC) and further support from the Stanford Department of Mechanical Engineering Graduate Teaching and Research Fellowship.

Work Cited

- [1] S. H. Gunther, D. P. Group, I. Corp, and F. Binns, “Managing the Impact of Increasing Microprocessor Power Consumption,” *Intel Technology Journal*, vol. 1, pp. 1-9, 2001.
- [2] S. Heo, K. Barr, and K. Asanovic, “Reducing Power Density through Activity Migration,” in *Low Power Electronics and Design, 2003. ISLPED’03. Proceedings of the 2003 International Symposium on*, 2003, no. C, pp. 217–222.
- [3] D. Brooks and M. Martonosi, “Dynamic Thermal Management for High-Performance Microprocessors,” in *Proceedings of the 7th International Symposium on High-Performance Computer Architecture*, 2001, no. C.
- [4] E. Kursun, G. Reinman, S. Sair, A. Shayesteh, and T. Sherwood, “Low-overhead Core Swapping for Thermal Management,” *Power-Aware Computer Systems*, pp. 46–60, 2005.
- [5] A. Cohen, F. Finkelstein, A. Mendelson, R. Ronen, and D. Rudoy, “On Estimating Optimal Performance of CPU Dynamic Thermal Management,” *Computer architecture letters*, vol. 2, no. 1, pp. 6–6, 2003.
- [6] A. K. Coskun, R. Strong, D. M. Tullsen, and T. Simunic Rosing, “Evaluating the Impact of Job Scheduling and Power Management on Processor Lifetime for Chip Multiprocessors,” in *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*, 2009, pp. 169–180.
- [7] M. Gomaa, M. D. Powell, and T. Vijaykumar, “Heat-and-Run: Leveraging SMT and CMP to Manage Power Density through the Operating System,” in *ACM SIGARCH Computer Architecture News*, 2004, vol. 32, no. 5, pp. 260–270.
- [8] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan, “Temperature-aware Microarchitecture,” in *ACM SIGARCH Computer Architecture News*, 2003, vol. 31, no. 2, pp. 2–13.
- [9] K. Skadron, M. R. Stan, K. Sankaranarayanan, W. Huang, S. Velusamy, and D. Tarjan, “Temperature-aware Microarchitecture: Modeling and Implementation,” *ACM Transactions on Architecture and Code Optimization (TACO)*, vol. 1, no. 1, pp. 94–125, 2004.

- [10] L. Hom, A. Durieux, J. Miler, M. Asheghi, K. Ramani, and K. E. Goodson, "Calibration Methodology for Interposing Liquid Coolants Infrared Thermography of Microprocessors," in *ITHERM*, 2012.
- [11] K. Etessam-Yazdani, "Continuum and Subcontinuum Thermal Modeling of Electronic Devices and Systems," Carnegie Mellon University, 2006.
- [12] K. Etessam-Yazdani and H. Hamann, "Fast and Accurate Simulation of Heat Transfer in Microarchitectures Using Frequency Domain Techniques," *IPACK*, pp. 1-5, 2007.
- [13] K. Etessam-Yazdani, H. F. Hamann, and M. Asheghi, "Spatial Frequency Domain Analysis of Heat Transfer in Microelectronic Chips with Applications to Temperature Aware Computing," in *IPACK*, 2007.

Figures

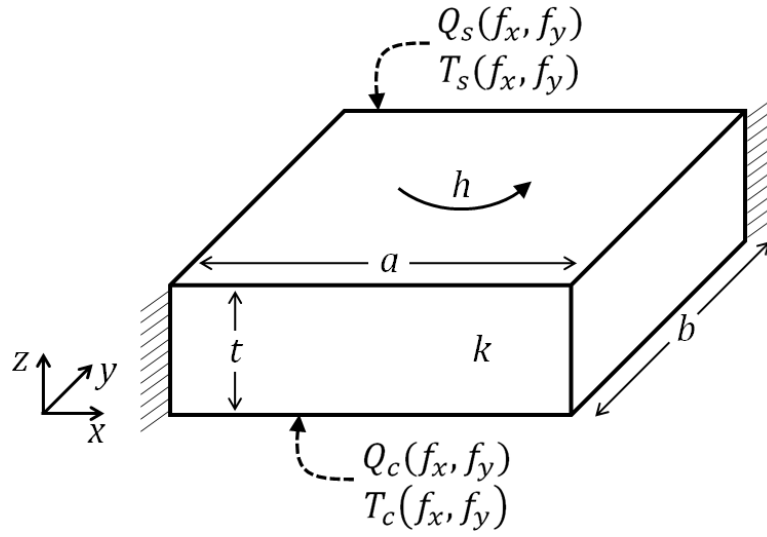


Figure 1: Schematic of model geometry. An arbitrary heat flux profile is applied on the bottom boundary. The boundary condition on all sidewalls is adiabatic; the boundary condition on the top surface is uniform heat transfer.

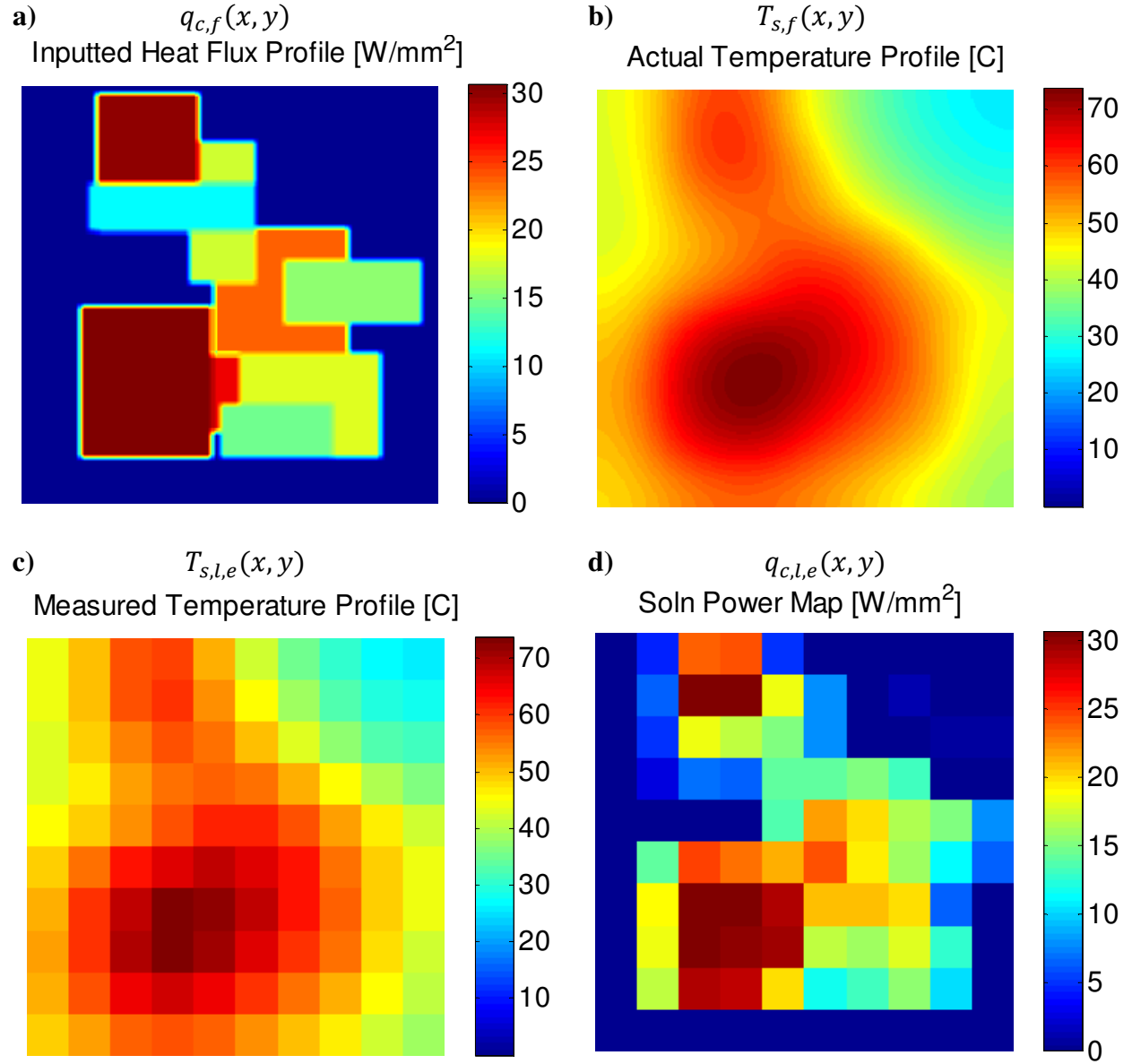


Figure 2: Representative images of each of the four main steps in the simulation methodology. The inputted heat flux profile (a) is used as a reference for determining the error in the calculated heat flux profile (d).

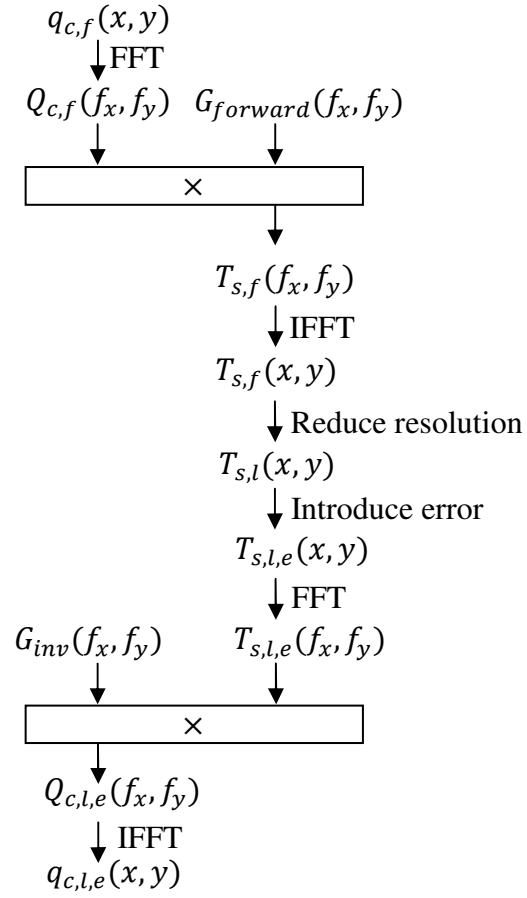


Figure 3: Block diagram of numerical approach used for determining hotspot detection accuracy. FFT and IFFT refer to the Fast Fourier Transform and the Inverse Fast Fourier Transform, respectively.

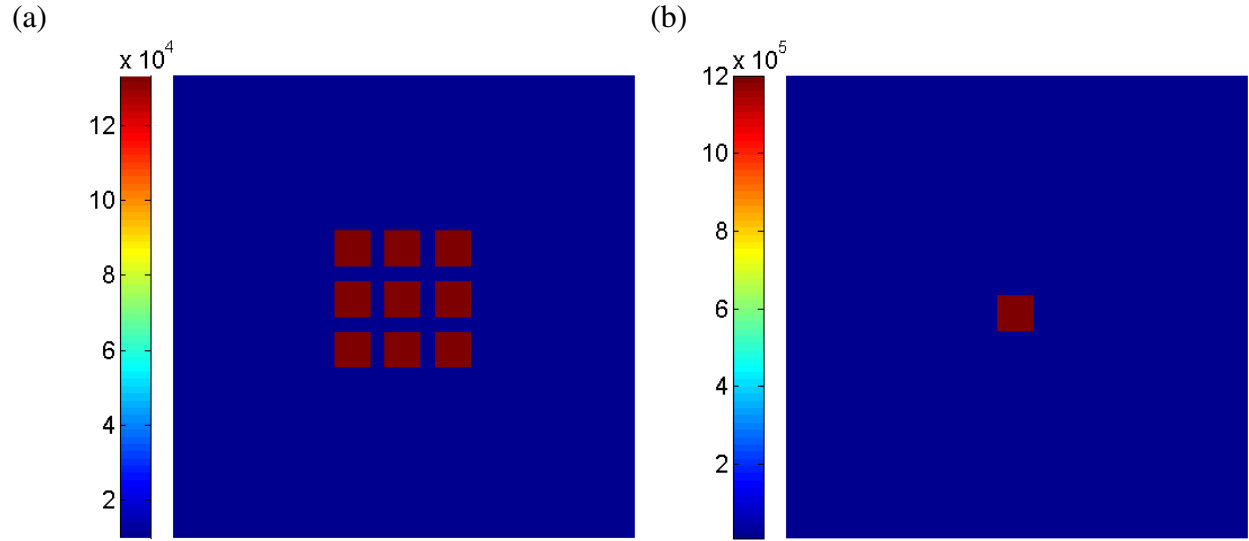


Figure 4: Heat flux profiles used for resolution study. Both heat flux profiles have equivalent average heat flux and produce similar temperature response profiles. The solution methods are tested for their ability to correctly resolve these heat flux profiles.

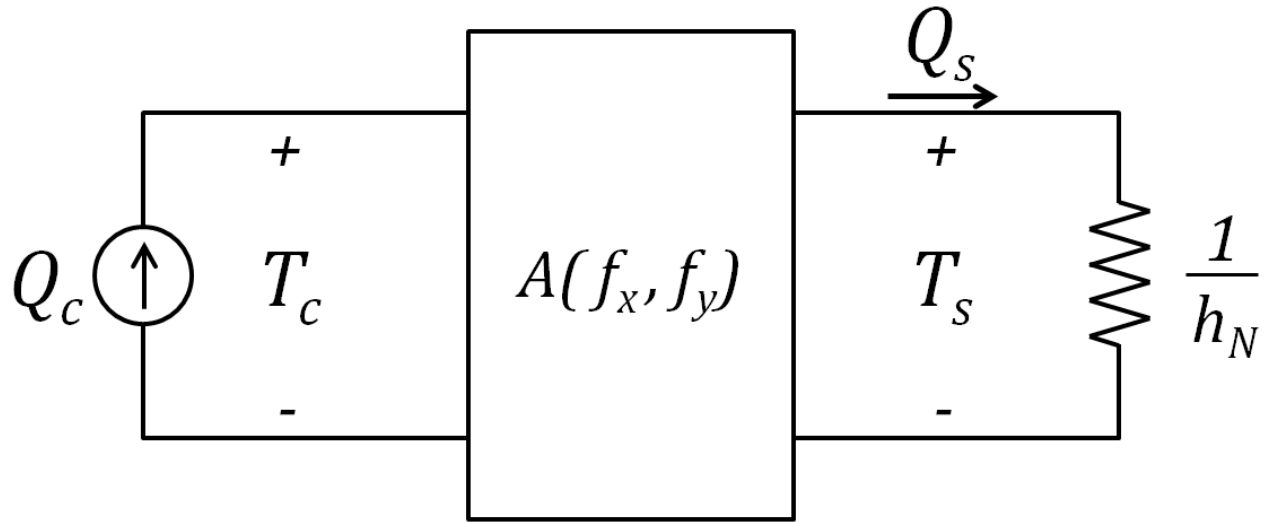
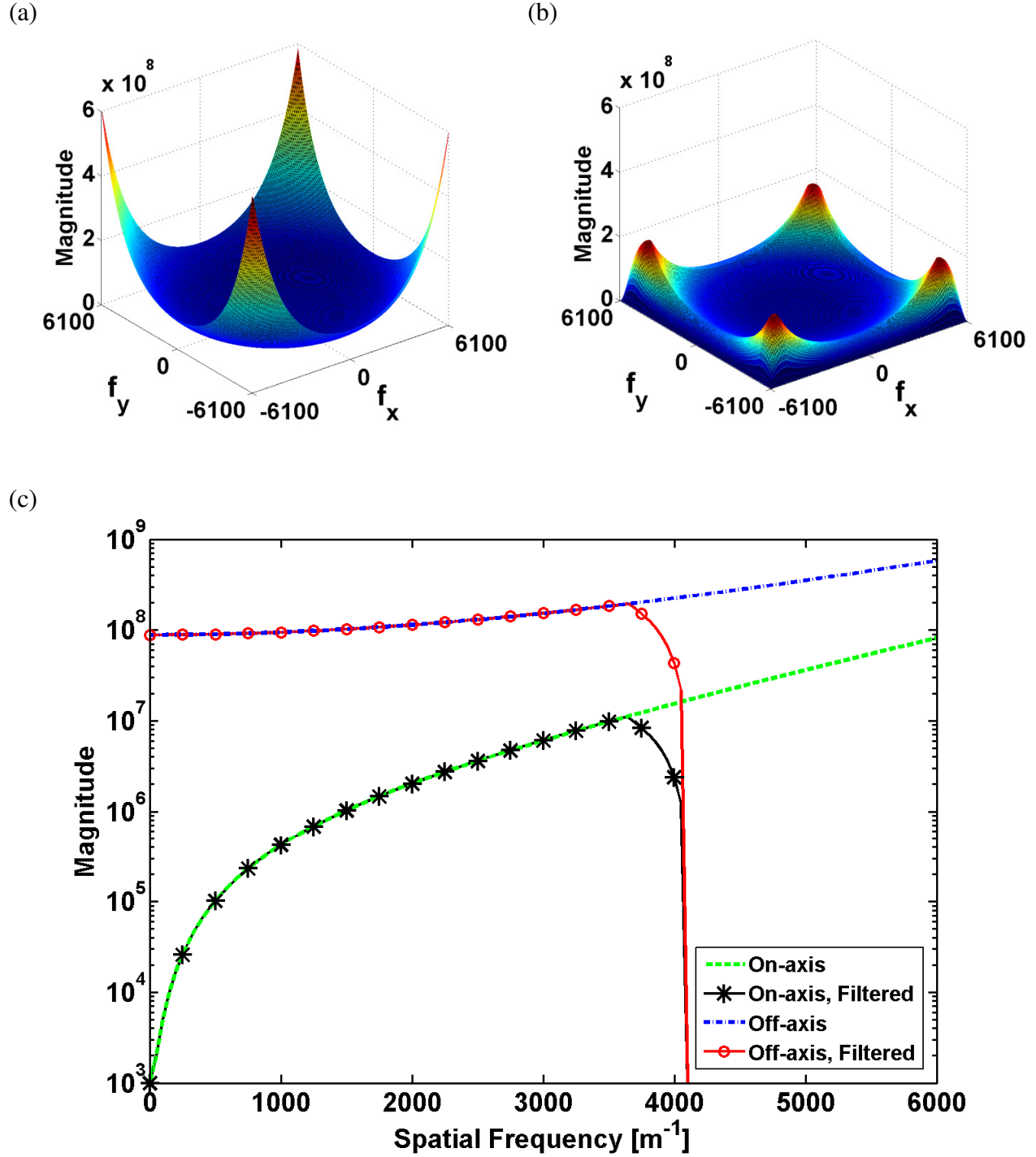


Figure 5: Schematic of two-port terminal network [11].



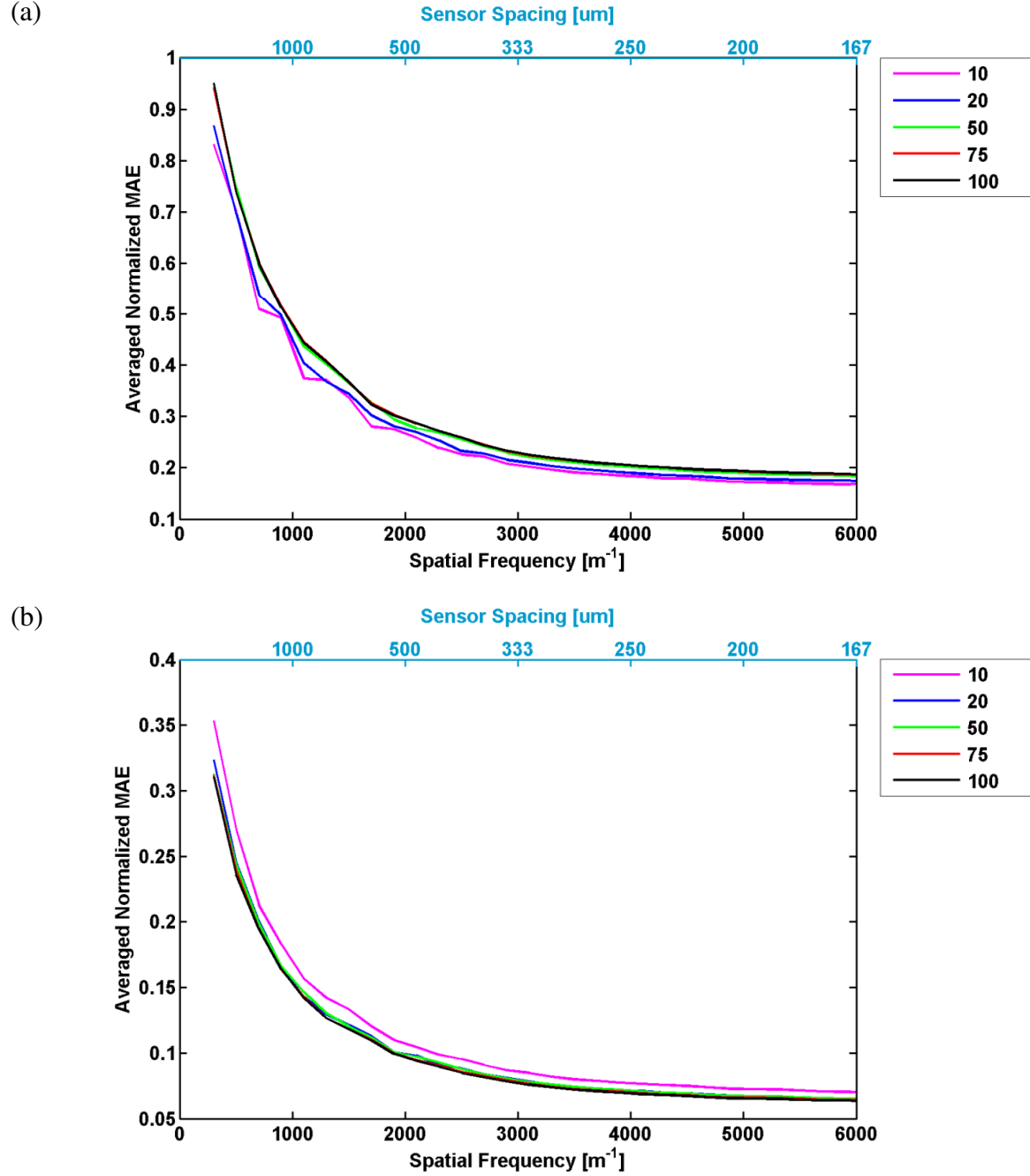


Figure 7: Average mean absolute error (MAE) for varying numbers of randomized heat flux profiles for (a) variable heat flux and (b) binary heat flux. Results for both cases are independent of the number of heat flux profiles for more than 50 heat flux profiles.

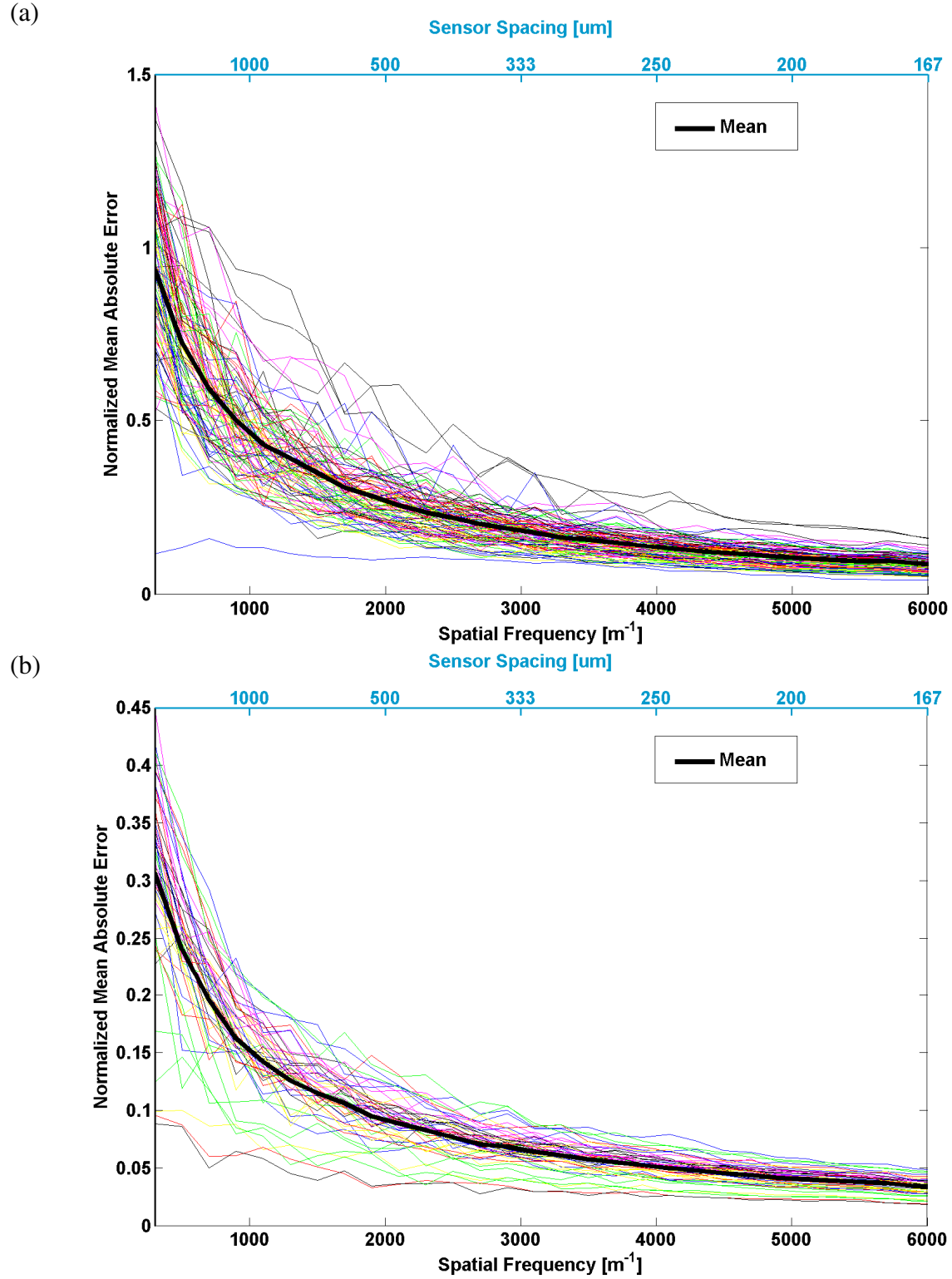


Figure 8: Demonstration of the averaging technique for (a) variable heat flux and (b) binary heat flux. Results for 50 heat flux profiles are shown. The bold black line indicates the average value.

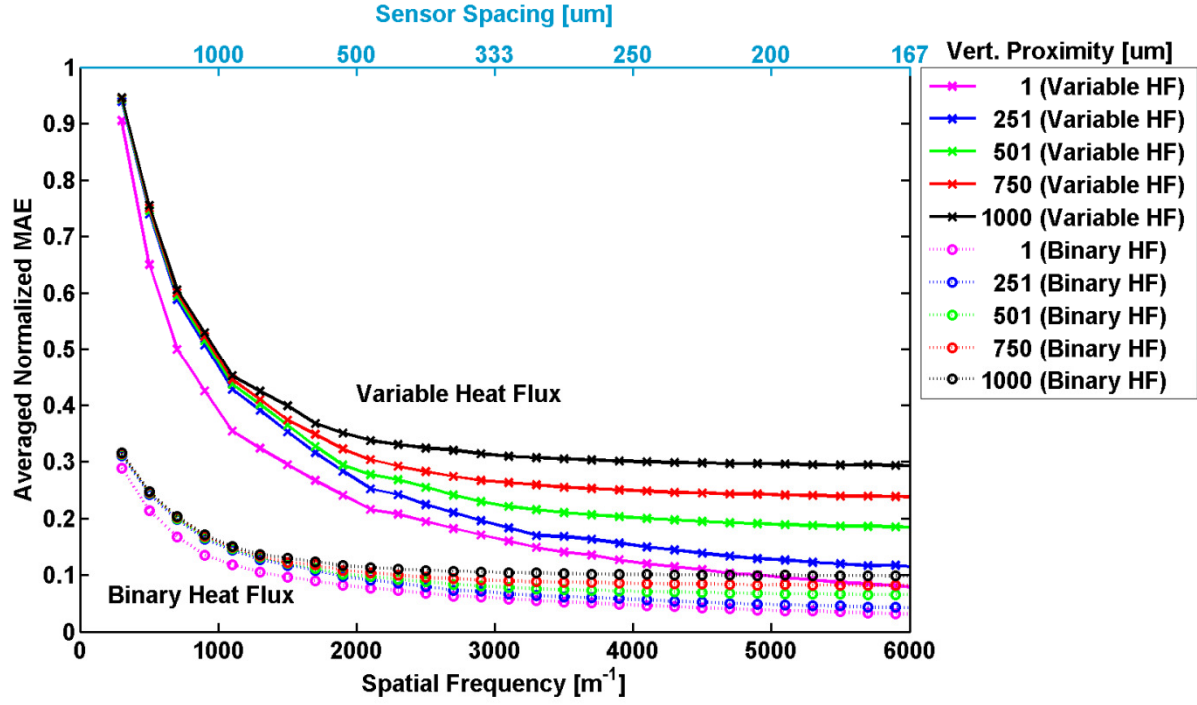


Figure 9: Effects on uncertainty of variable versus binary inputted heat flux profile for varying vertical proximity between sensor and circuit level. The binary heat flux profile results in substantially lower MAE.

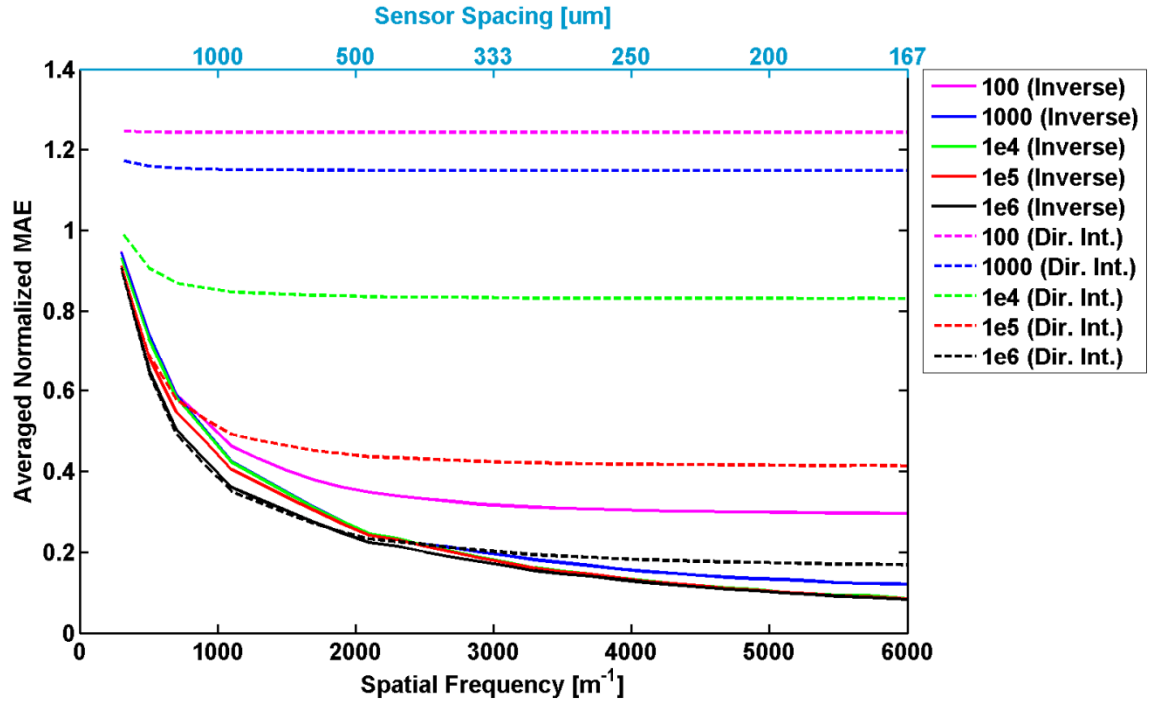
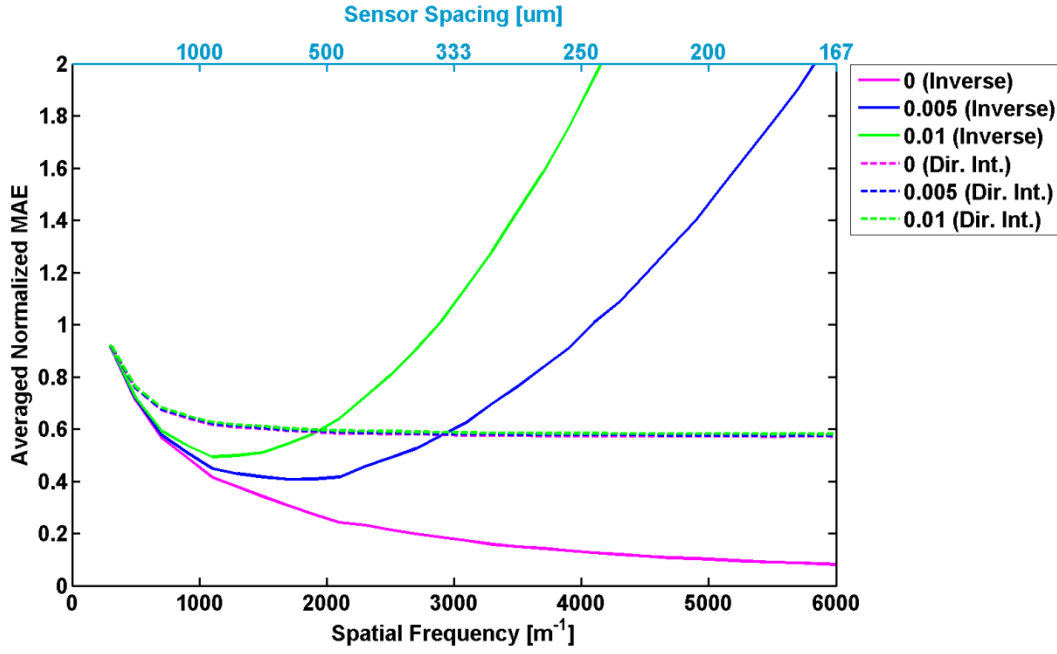


Figure 10: Uncertainty in calculated heat flux profile for varying convective heat transfer coefficient. The inverse solution method is much less sensitive to heat transfer coefficient than the direct interpretation method.

(a)



(b)

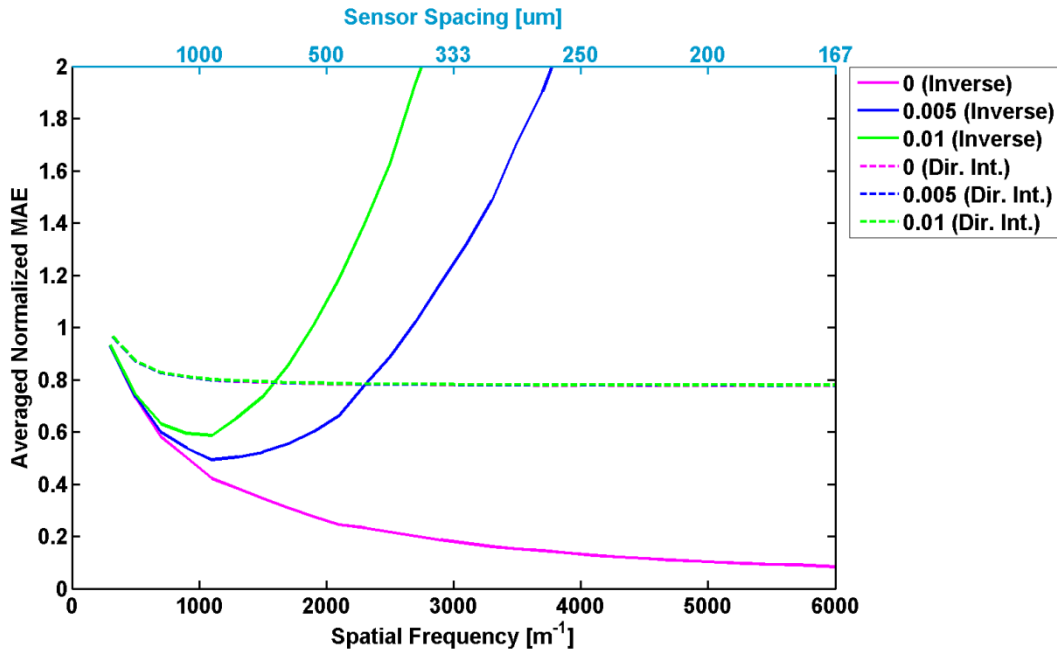


Figure 11: Uncertainty in calculated heat flux profile for varying sensor error at a vertical proximity of (a) 2.575 μm and (b) 7.53 μm . The inverse solution method is susceptible to sensor error at high spatial frequency. The MAE for the direct interpretation method is not affected by varying sensor error.

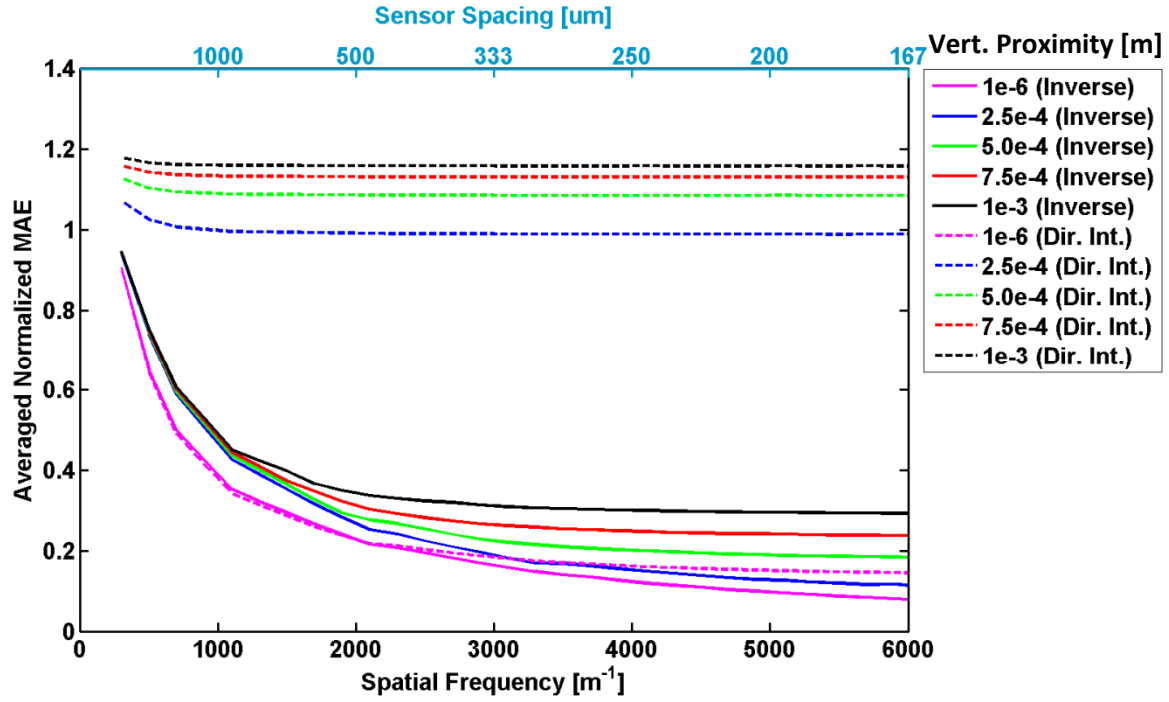


Figure 12: Uncertainty in calculated heat flux profile for varying vertical proximity between the sensor and circuit levels for zero sensor error. For most cases, large changes in vertical proximity yield modest improvements in heat flux uncertainty.

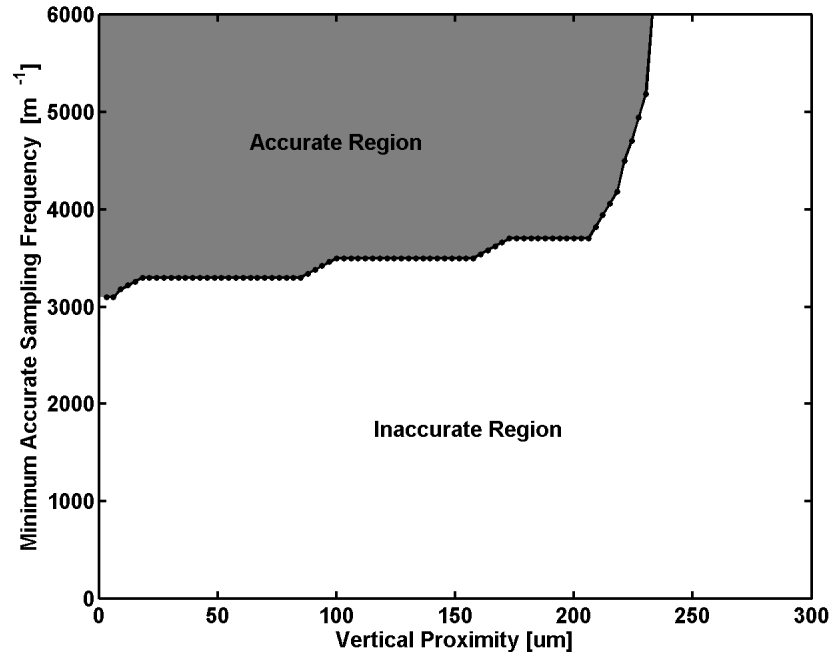


Figure 13: Plot of minimum accurate sampling frequency as a function of vertical proximity between chip and sensor level for heat transfer coefficient of 10^4 W/m²-K. The inverse solution method is accurate in the shaded region. The direct interpretation technique is inaccurate across the entire domain.

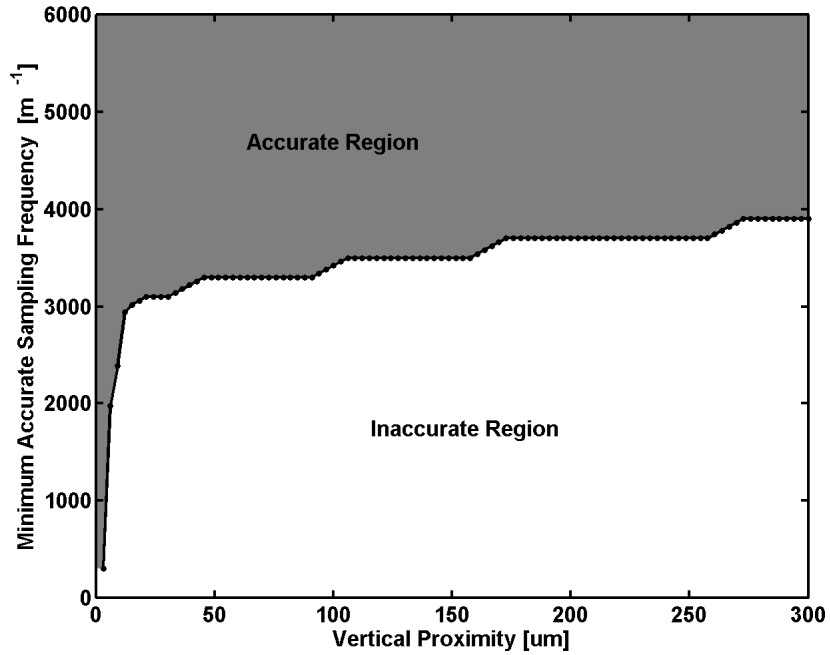


Figure 14: Plot of minimum accurate sampling frequency as a function of vertical proximity between chip and sensor level for heat transfer coefficient of 10^5 W/m²-K. The inverse solution method is accurate in the shaded region. The direct interpretation technique is inaccurate across the entire domain.