Relation between Sample Area and Accuracy of STM: What Causes the Difficulty in Quantitative Elemental Analysis?

Ryohei Tanaka,† Tetsuya Kishimoto, Takayuki Yoshimura, Shu Kurokawa, Akira Sakai, Jun Kawai ‡

† Corresponding author: tanaka.ryohei.5r@kyoto-u.ac.jp
‡ Corresponding author: kawai.jun.3x@kyoto-u.ac.jp

A Monte Carlo numerical simulation is performed to clarify the reason why an atomic resolution microscope such as a scanning tunneling microscope (STM) cannot give accurately a ppm (parts per million) concentration of elements quantitatively. A relation between the sampling amount and the accuracy has been found. At least $10^9$ atoms are needed to be sampled in order to discriminate the 49 and 50 ppm elemental concentrations. This Monte Carlo result has also been applied to the elemental analysis methods such as ICP-AES (inductively-coupled plasma atomic emission spectrometry) and TXRF (total reflection X-ray fluorescence) to clarify the appropriate amount of sampling amount for required precision of quantitative elemental analysis.

Keywords: Atomic resolution microscopes; Quantitative elemental analysis; STM; AFM; TXRF

I. INTRODUCTION

Scanning probe microscopes (SPM), such as scanning tunneling microscope (STM), atomic force microscope (AFM), or other related methods, can observe a single atom, but it cannot analyze quantitatively even a ppm (parts per million) chemical element concentration. If we observe Si wafer surface where an impurity element was doped up to a concentration of a few hundreds of ppm, we cannot observe the dopant atoms in scanning probe microscope (SPM) images at all. One motivation of the present report is to clarify the reason of this. The total reflection X-ray fluorescence (TXRF) spectrometry is an elemental analysis method of achieving the detection limit down to a few pico-grams of transition metal elements for a square centimeter, which is equivalent to $10^9$ atoms in a square centimeter. We know empirically that it is not possible to find a dopant atom using the STM method when the concentration of which is $10^9$ atoms in a square centimeter. But why? The purpose of the present report is in order to theoretically clarify this reason of our empirical experience.

Another motivation of the present research is in order to know the relation between the accuracy and sampling amount when we use inductively-coupled plasma atomic emission spectrometry (ICP-AES). When we sample a powder (bulk) that is a mixture of various kinds of single elements, we want to know how many grams are appropriate for a required accuracy (mean value) and precision (standard deviation) of bulk analysis.

This kind of Monte Carlo numerical simulation is quite simple but requires several months of computer time for only one set of simulations. Therefore we cannot find similar research in the literature. However, we have found the following important publications related to the sampling amount or probing area and the elemental homogeneity.

Hoornaert et al. [1] assessed the homogeneity of powder samples of reference materials with the help of electron probe X-ray microanalysis (EPMA). They used Kolmogorov-Smirnov statistics coupled with the Akaike’s Information Criteria (AIC). Ro et al. [2] extended the results of Hoornaert...
et al. [1] to the representativeness of single particle EPMA analysis. Borkhodoev [3] studied mineral phase homogeneity using EPMA by applying the equation of Potts et al. [4]. Marques et al. [5] studied the grain size effect on the homogeneity (i.e., representativeness) for PIXE (particle-induced X-ray emission) as well as INAA (instrumental neutron activation analysis). Aerosol sampling is a similar situation, and Gysels et al. [6] who discussed sampling error of aerosols, in order to eliminate the observed concentration errors among different impactors, by comparing the size distribution of particles of each impactor stage through the non-parametric Kolmogorov-Smirnov statistics and EPMA diameter analysis of 2000 particles on every impactor stage, referring to air turbulence [7, 8].

The Monte Carlo simulation was developed during the Second World War in order to evaluate the critical mass of thermonuclear explosion, and thus was not open till 1949 by Metropolis and Ulam [9]. The title of this first paper was only “The Monte Carlo Method”, and thus attracted little attention at first. Then the interaction of individual molecules was published in 1953 [10], and the Monte Carlo method became used in chemical physics [11, 12]. It is true that we can reach the exact results by statistics, however the Monte Carlo method will provide the similar results without the functional analysis in mathematics. This is the reason we have used the Monte Carlo method in the present paper. A preliminary result of the present paper was reported in a different parameter [12].

II. NUMERICAL SIMULATION

In order to address the accuracy and precision of analysis of impurity surface atom, a Monte Carlo simulation was conducted. The algorithm used was a procedure made of steps (i) to (iv) as follows.

(i) A uniform random integer between 0 and, e.g., $10^6$ (when we want to simulate a ppm concentration) was generated by the Mersenne twister method in C++11 library.

(ii) The random integer generated in step (i) was counted when it was less than a particular value, e.g., 50 (when we simulated 50 ppm of concentration), in order to simulate the detection of an impurity atom.

(iii) The steps (i)–(ii) were repeated until, e.g., $10^{12}$ times.

(iv) The step (iii) was repeated, e.g., 10 times by changing the seed point of random integers.

At an early stage of our simulation, we generated random real numbers between 0.0 and 1.0, and counted the frequency of less than $50 \times 10^{-6}$ for 50 ppm simulation. However, small real numbers did not give an accurate probability. We found the integer random numbers gave satisfactory results.

The above steps (i)–(iv) are mathematically stated as follows. When simulating a sampling of $T$ atoms from the population whose concentration of impurity is $x$, the sampling concentration was determined by $T_x/T$, where $T$ denotes the frequency of integers between 0 and $x \times T - 1$. The population is a term in statistics against the sample. The number of sampling atoms, $T_x$ termed as the sample size in statistics, is set to $10^n$ ($0 < n < 13$). The calculation for $n = 13$ took a few weeks as a CPU time for a Windows computer, which meant that there existed the practical difficulty in the simulation for $n \geq 13$ due to a large computational time. Therefore, in the present study, $n$ was set to less than 13. After ten iterations in step (iv), the average and the standard deviation of $T_x/T$ were calculated for impurity concentrations of 50 and 500 ppb; 49, 50, and 500 ppm; and 50%. The average is a measure of accuracy; the standard deviation is a measure of precision.

III. RESULTS OF SIMULATION

Figure 1 is a sampling simulation for a highly-concentrated sample that has 50% impurity atoms. Results of ten trials are plotted. It was enough to sample $10^3$ atoms to obtain a satisfactory precision of surface impurity atoms. That is to say, the sample size of 1000 is enough to obtain the accurate concentration of 50%. On the other hand, for a 50-ppm dilute population as shown in Figure 2, $10^4$ atoms were needed to be sampled in order to obtain an impurity concentration satisfactorily, which made the relative standard deviation within 1%. Figure 3 shows a comparison of simulations for representative one

![Figure 1](image1.png)  
**Figure 1:** Sampling simulation of the analysis of impurity surface atom, whose concentration is 50%. Simulations were conducted 10 trials.

![Figure 2](image2.png)  
**Figure 2:** Sampling simulation of the analysis of impurity surface atom, whose concentration is 50 ppm. The simulation procedure was the same as Figure 1.
trials from 50% and 50 ppm. If the sample size was $10^7$ from the population concentration of 50 ppm, the accuracy became similar to the sample size of $10^4$ from the population concentration of 50% impurity.

Figure 4 shows the results of sampling simulations for 49- and 50-ppm impurity atoms. When the sample size was less than $10^8$ atoms, the standard deviation became larger than 1 ppm ($= 50$ ppm $- 49$ ppm). In order to discriminate the 49-ppm concentration from 50 ppm, at least $10^9$ atoms should be sampled from the population as shown in Figure 4.

Figure 5 is the dependence of the relative standard deviation as the change of the sample size. If a microscope that has an atomic resolution like STM is used for the quantitative analysis of chemical element concentration, a large number of atoms were needed to be sampled, whose amount depended on the sample concentration. It is practically difficult to scan such a large number of surface atoms using the STM.

IV. APPLICATION TO STM

Kurokawa et al. [13] reported an STM image of a hydrogen terminated Si(111) surface, where arsenic atoms were doped by $5 \times 10^{19}$ atoms cm$^{-3}$. Arsenic atoms were imaged as bright points with 0.1 nm height. The STM could detect arsenic atoms in the first and second surface layers. Numbers of arsenic atoms observed were 20 to 30 in 60 nm by 60 nm scan area. This scan area was equivalent to 28000 silicon atoms on the top surface. The silicon atom density was $5.0 \times 10^{22}$ atoms cm$^{-3}$, which was equivalent to 7.8 atoms nm$^{-2}$. Thus, 56 impurity arsenic atoms should be detected in their observations. However, the number of observed atoms was from 1/3 to 1/2 of the above estimated 56 atoms, and the error was by a factor of 2 to 3 times. This was equivalent to the sampling of $10^4$ to $10^5$ particles at 50 ppm in Figure 2. The observed large standard deviation was satisfactorily reproduced by the simulation of $10^4$ to $10^5$ sample sizes, which were not enough as a sample size.

V. APPLICATION TO ICP-AES AND TXRF

The above numerical results are easily applicable to the sampling amount of ICP-AES of a powder mixture. We will suppose that an analyte mixture consists of powders A and B, which are composed of chemical elements A and B, respectively. Here, we also suppose that these powders A and B have the same density, 2.0 g cm$^{-3}$. One spherical particle in both powders is assumed to be 0.1 μg, and thus the particle diameter becomes about 50 μm. Then, we can determine the minimum sample size of the powder particles in order to obtain a concentration with a required accuracy as follows.

According to the results of simulations, when the ratio of particle A in the population is 50%, sampling from 100 to 1000 particles, i.e., from 0.01 to 0.1 mg are enough for 1% standard deviation. In case that the concentration of particle A in the population is 50 ppm, we need to sample from 106 to 107 particles, i.e., from 0.1 to 1.0 g in order to achieve the same relative standard deviation. In many laboratories, an
SOP (standard operation procedure) has been used in such a way that the 250 mg should be sampled from the analyte and then the sample is analyzed by using ICP-AES. This sampling procedure is usually repeated five times. Then, from the results reported in the present work, we can analyze from ppm to % concentrations within a sufficient standard deviation.

The discussion in the present paper is also applicable to the TXRF analysis. Several hundred pico-grams of transition metals are put on the sample carrier within the area of 1 cm² and are analyzed by a TXRF spectrometer [14]. The X-ray counting duration is usually selected to be 10−30 minutes, and the peak intensity is several thousands of counts per the duration time. These counting statistics is enough. The detection limit is achieved down to the single pico-grams or several tens of pico-grams. Therefore, the measured concentration of analyte is only 10 times (several hundreds of pico-grams) larger than the detection limit (several tens of pico-grams). This is the reason of large standard deviation of TXRF results. The sample size should be at least two orders of magnitude larger than the detection limit in order to deliberately make lower the standard deviation, which is found from Figure 5. However, the 10⁸ atoms are large enough as the sample size to detect the ppm or ppb concentration of chemical elements.

VI. CONCLUSIONS

In order to rationalize the reason why an atomic resolution microscope cannot quantitatively analyze even a ppm concentration, Monte Carlo simulations were performed by changing the sample size. Based on the Monte Carlo simulation, accuracy (mean) and precision (standard deviation) have been estimated as a function of the sample size. When the impurity is 50 ppm, at least 10⁹ sample size is needed in order to discriminate from 49 ppm concentration. This kind of sample size can be read from the figures in the present report, and if other parameters are needed, the Monte Carlo simulation will give an appropriate sample size for any accuracy and precision before the experiment.

Acknowledgments

Thanks are due to Rene Van Grieken for informing us of the Ingham’s work, V. Ya. Borkhodoeev for fruitful discussion on EPMA analysis, and Anatoly G. Revenko for fruitful comments on the early stage of the present manuscript.

Note

This paper was presented at the 12th International Symposium on Atomic Level Characterizations for New Materials and Devices ’19 (ALC ’19), in conjunction with the 22nd International Conference on Secondary Ion Mass Spectrometry (SIMS-22), Miyako Messe, Kyoto, Japan, 20–25 October, 2019.

References


All articles published on e-J. Surf. Sci. Nanotechnol. are licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0). You are free to copy and redistribute articles in any medium or format and also free to remix, transform, and build upon articles for any purpose (including a commercial use) as long as you give appropriate credit to the original source and provide a link to the Creative Commons (CC) license. If you modify the material, you must indicate changes in a proper way.

Published by The Japan Society of Vacuum and Surface Science