OUKA
The University of Osaka
Institutional Knowledge Archive

Title	The development of the fast optimization method in bulk crystal growth process using data science-integrated transport phenomena
Author(s)	竹原, 悠人
Citation	大阪大学, 2023, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/95946
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

The development of the fast optimization method in bulk crystal growth process using data science-integrated transport phenomena

Yuto Takehara

December 2023

The development of the fast optimization method in bulk crystal growth process using data science-integrated transport phenomena

A dissertation submitted to
THE GRADUATE SCHOOL OF ENGINEERING SCIENCE
OSAKA UNIVERSITY
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN ENGINEERING

BY

Yuto Takehara

December 2023

Abstract

In chemical engineering, the analysis and control of transport phenomena in the process are crucial as their effect on product quality is significant. In typical manufacturing processes, direct observation of the transfer phenomena is challenging. As a result, control and analysis have historically relied on experience-based black-box approaches and numerical simulations. However, there are challenges in this situation. The black-box approach leads to difficulties in providing clear reasons for control decisions and in scaling up due to the lack of understanding of transport phenomena in the process. Another problem is the long time for numerical simulation. Due to a long time, the optimization for a suggestion of control methods also takes a long time since numerical simulation is carried out many times. In addition to that, the identification of key phenomena is also difficult due to the nonlinear distributions of physical quantities, e.g. flow velocity, temperature, and concentration fields. Therefore, The objective of this study was to develop a rapid analysis method for the identification of crucial phenomena and a quick optimization approach for controlling these phenomena, even in real-time, by utilizing data science. In this study, the crystal growth process was selected as one of the processes involving the challenges mentioned above. The fast optimization for controlling transport phenomena, the identification of key phenomena, and fast analysis methods were achieved by Bayesian optimization, explainable AI, and physics informed neural networks, respectively. By combining these methods, the objective was achieved. Due to these methods, real-time analysis and suggestion of suitable control methods during experiments will be achieved. The proposed methods can be utilized in any other processes, which will raise the overall level of process control in chemical engineering.

Contents

1 General introduction **7**
1.1 Background and motivation . . . 7
1.2 Crystal growth process . . . 8
1.2.1 Czochralski (Cz) method for silicon (Si) crystals . . . 8
1.2.2 Top-Seeded Solution Growth (TSSG) method for silicon carbide (SiC) crystals . . . 9
1.3 Data science in crystal growth . . . 11
1.3.1 Previous applications: Neural networks . . . 11
1.3.2 Bayesian optimization . . . 12
1.3.3 Explainable AI . . . 13
1.3.4 Physics Informed Neural Networks . . . 14
1.4 Objective . . . 15
1.5 Thesis outline . . . 15

2 Bayesian optimization for the TSSG method of SiC crystals **16**
2.1 Numerical simulation . . . 16
2.1.1 Integrate Process Model (IPM) . . . 16
2.2 Bayesian optimization . . . 27
2.2.1 Gaussian process (GP) . . . 28
2.2.2 Kernel function and kernel trick . . . 30
2.2.3 Gaussian process regression (GPR) . . . 31
2.2.4 Acquisition functions . . . 32
2.3 Effectiveness of Bayesian optimization . . . 32
2.4 Gaussian process regression after the optimization . . . 34
2.5 The transport phenomena in the initial and the optimal state . . . 35
2.5.1 Suppression of the Marangoni convection . . . 35
2.5.2 Contributions of driving forces . . . 36
2.5.3 Transport phenomena near the seed crystal . . . 36
2.6 Current achievements, limitations, and future improvements . . . 36
2.6.1 Efficiency . . . 37

2.6.2 Application 37

3 Explainable machine learning for the analysis of the transport phenomena in the melt of TSSG method for SiC crystals 42

3.1 Concept of this study 42
3.2 Numerical method 43
3.2.1 Convolutional Neural Network (CNN) 43
3.2.2 Sensitivity tensor 45
3.2.3 SmoothGrad 45
3.3 The training and accuracy of CNN 46
3.4 Sensitivity analysis 46
3.4.1 Sensitivity below the seed center 46
3.4.2 Sensitivity below the seed edge 47
3.4.3 Sensitivity of the top 10% of the objective function 49
3.5 Current achievements, limitations, and future improvements 49

4 Physics Informed Neural Networks for Czochralski method of Si crystals 52

4.1 Numerical simulation 52
4.2 Physics Informed Neural Networks 54
4.2.1 Architecture and loss functions 54
4.2.2 Training process 58
4.3 Predictions by PINNs 59
4.4 Accuracy of PINNs 60
4.5 Current achievements, limitations, and future improvements 62

5 Summary and future works 66

5.1 Summary 66
5.2 Future works 67

Bibliography 68

List of Tables 78

List of Figures 79

Acknowledgements 82

List of Publications 84

List of International conference 85

List of domestic conference 86

Awards **88**

Chapter 1

General introduction

1.1 Background and motivation

In manufacturing processes, the analysis and control of transport phenomena are important since product quality depends on the transport phenomena, such as the distributions of fluid flow, temperature, and concentration. Generally, direct observations of the transport phenomena in the equipment are challenging, e.g. due to high-temperature conditions. Consequently, analysis and control methods have mainly relied on numerical simulation with transport phenomena and black-box control, (e.g. based on experiential knowledge and process control), respectively. Numerical simulation and black-box control enable a cost-effective visualization of transport phenomena that are challenging to observe directly and rapid control decisions without requiring a precise understanding of transport phenomena, respectively. In this study, we refer to the analysis of transport phenomena as a "white box" in contrast to a "black box."
However, several challenges delay the analysis and control of transport phenomena in manufacturing processes:

- Black-box approaches lead to difficulties in providing clear reasons for control decisions and in scaling up operations due to the lack of knowledge about the transport phenomena in the process.
- Numerical simulation for analysis of the transport phenomenon takes a long time.
- Suggesting control methods based on numerical simulations also needs a long time since the optimization of a number of parameters by numerical simulation is required.
- Identifying the key phenomena affecting product quality is challenging, as physical quantities show distributions and nonlinear behavior.

In order to overcome these challenges, fast and white-box suggestions of control methods combining the advantages of process control and numerical computation are required.

1.2 Crystal growth process

In this study, the crystal growth process is treated as an example process facing the above problems. While crystal growth can be broadly categorized into growth from melt, solution, and vapor phases, they all share a common characteristic: crystals grow from a fluid.
During the growth from the fluid phases, the crystal quality depends on the transport phenomena in the fluid, such as temperature, concentration, and flow velocity distributions. However, the direct in-situ observation of the transport phenomena is challenging, particularly when dealing with crystals with high melting points, such as semiconductors. Moreover, complex transport phenomena typically develop in growth systems, which means that numerical simulations take a long time. Therefore, the same difficulties mentioned in section 1.1 remain, requiring fast and white-box control methods.
In this study, we have focused on two specific crystal growth processes: the Czochralski method for silicon crystals and the Top-Seeded Solution Growth method for silicon carbide crystals.

1.2.1 Czochralski (Cz) method for silicon (Si) crystals

Si crystal is a widely used semiconductor and the foundation of today's digital technology. Si crystal is mainly produced by the Czochralski (Cz) method[1, 2]. In the Cz method, solid Si is heated by resistance heating until it melts in silica (SiO_2) crucible. Si crystal grows on the Si seed attached at the center of the free surface which is the cooler position (see Fig. 1.1). During the growth, the crystal is both rotated and pulled to control its radius, resulting in the production of Si ingots.
In this growth system, oxygen supplied from the crucible plays an important role. Interstitial oxygen in the Si crystal enhances its mechanical strength and influences the electrical properties of Si wafers[3, 4, 5]. Moreover, the oxygen precipitates in the Si crystals behave as sites for trapping metal impurities[6]. Thus, the control of oxygen concentration in the melt is important for stabilizing crystal quality.
In order to control the oxygen concentration in Si crystals, the understanding and control of the transport phenomena, e.g. fluid flow and temperature distributions, in the melt is important. On the other hand, the in-situ direct observation is difficult since the temperature in the Sim melt is higher than the melting point of Si (1685 K). Thus, a lot of numerical studies have been carried out for analysis of the transport phenomena in the Si melt. To have better control of the transport phenomena, various types of magnetic fields have been introduced ([7] for InSb crystals, [8, 9, 10]) and their effects on crystal quality and melt flow were studied[11, 12, 13, 14, 15, 16]. It was found that the application of magnetic fields, along with crucible and seed rotations, is beneficial in the control of transport structures in the melt.
At present, however, the challenges mentioned above remain in the Cz method. The numerical simulation takes a long time since the complex transport phenomena in the Si melt develop.

Moreover, another factor contributing to the long calculations is the necessity to perform computations for various geometries, as the liquid height decreases during crystal growth. Due to the long calculation time, the black-box control method without understanding the transport phenomena has been commonly employed during experiments. In addition, the optimization using numerical simulation also takes a long time. Thus, a fast prediction method of the transport phenomena in the melt is required.

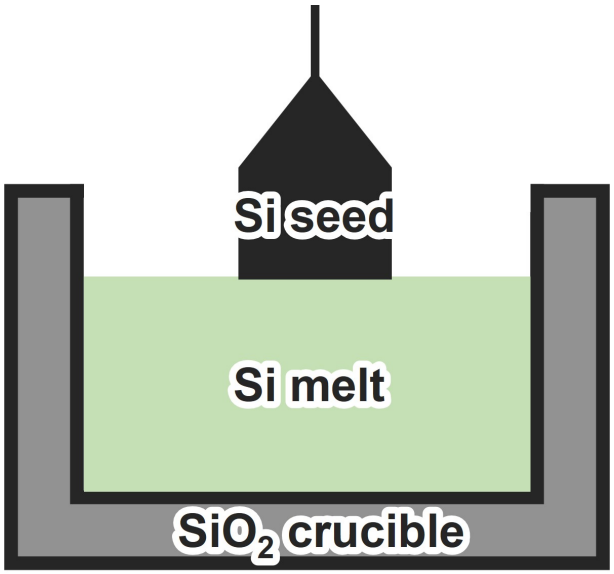

Figure 1.1. The schematic view of the Czochralski method for Si crystals.

1.2.2 Top-Seeded Solution Growth (TSSG) method for silicon carbide (SiC) crystals

SiC crystal is an attractive power semiconductor material due to its physical properties such as wide band gap and high thermal conductivity[17] (see table 1.1). The high operating temperature of SiC can be achieved by the wide band gap since the electron transition from the conduction to the valence bands is difficult. The cooling system can be also downsized because of the ease of exhausting heat by the high thermal conductivity of SiC. Moreover, low resistance is achieved due to the high breakdown field strength.
Currently, SiC crystals are mainly produced by the sublimation method (see [18]). In the sublimation method, solid SiC powder is sublimated at high-temperature conditions ($T > 2200$ K) and recrystallized on the SiC seed crystal attached to a cooler position (see Fig. 1.2). However, it is difficult to produce high-quality SiC crystals by using this sublimation method since the ingot purity depends on the purity of SiC powder and crystal defects develop due to high thermal stress. In addition to that, it is also difficult to maintain the thermal and chemical conditions

since slight modifications can change the growth behavior drastically[19].

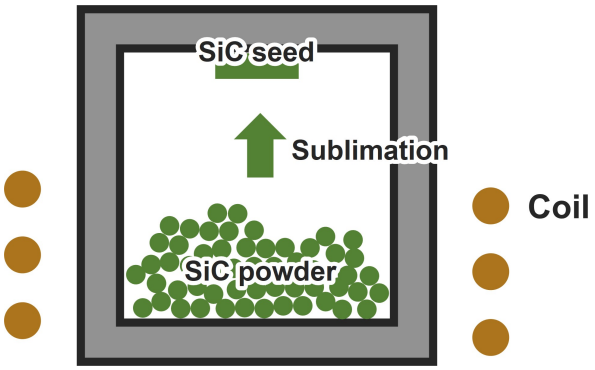

Figure 1.2. The schematic view of the sublimation method for SiC crystals

The TSSG method has been attractive as a promising technique for producing high-quality SiC crystals[18, 20, 21, 22]. In this method, SiC crystals are produced from Si-C solution (melt). Solid Si is heated by induction heating of the coil and melted in the carbon crucible. The carbon dissolved from the crucible is transported to the SiC seed crystal at the center of the free surface of the melt by convection. Then, the SiC crystal grows by supersaturation along the seed crystal (see Fig. 1.3). In the method, the crystal quality depends on carbon transport by the complex transport phenomena in the melt driven by the electromagnetic field from the coil, surface tension gradient, and Buoyancy force[23]. Thus, the understanding and control of the transport phenomena in the melt is essential for a stable supply of high-quality SiC crystals. However, direct in-situ observation of the transport phenomena is difficult since the temperature

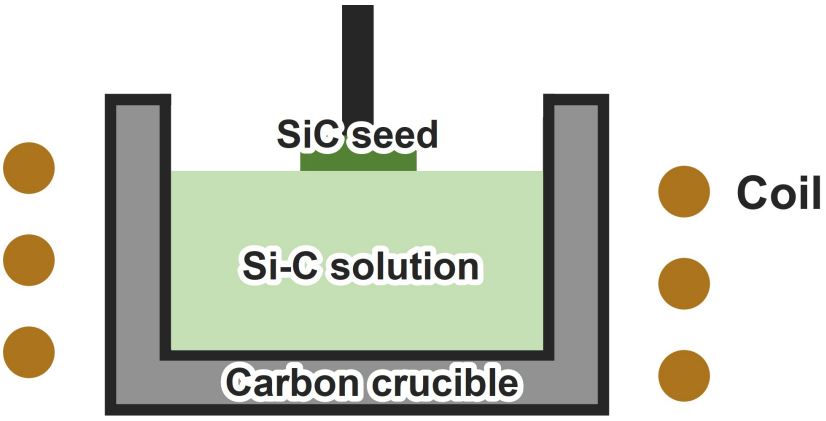

Figure 1.3. The schematic view of the TSSG method for SiC crystals

of the melt becomes about 2000 K during the growth of SiC crystals and liquid Si is opaque. Thus, numerical simulation and black-box control have been utilized.

Table 1.1 The physical properties of Si and SiC crystals[17]

Items	Si	4H-SiC
Band gap (eV)	1.12	3.3
maximum operating temperature (°C)	600	1580
Breakdown field strength (V/cm)	3×10^5	22×10^5
Thermal conductivity (W/cm· K)	1.5	4.9

Various research has been conducted for the TSSG system of SiC crystals by using numerical simulation[20, 24, 25, 26]. Yamamoto *et al.*[23] first conducted the global simulation of the TSSG method of SiC and found that surface tension gradients develop due to temperature difference along the free surface. The surface tension gradient causes the convection toward the seed crystal along the free surface called Marangoni convection which has an adverse effect on the crystal growth rates. Wang *et al.*[27] investigate the effect of the magnetic fields and found that the Marangoni convection can be suppressed by the magnetic fields and the seed rotation. Despite numerous studies, the problem mentioned above remains in the TSSG-SiC system. Due to the complex transport phenomena driven by Lorentz force by induction heating, Buoyancy, and surface tension coefficient, numerical simulation requires a long time. In addition to that, optimization of parameters is harder because the transport phenomena depend on many parameters. Thus, the optimal magnetic field for suppressing the Marangoni convection affecting crystal growth adversely had not been obtained due to the difficulty of the optimization. Moreover, the identification of key transport structures is also difficult since it shows distributions with non-linearity behavior.

1.3 Data science in crystal growth

1.3.1 Previous applications: Neural networks

In recent years, fast prediction and efficient optimization techniques in data science have attracted much attention. The application of data science to crystal growth processes has also recently become popular with the help of Python[28] and open source libraries[29, 30, 31, 32], and summarized in the literature[33]. In prior applications of data science to crystal growth, neural networks[34] were commonly utilized. Neural networks are mathematical models that mimic the behavior of neurons in the human brain. These networks enable fast predictions of results obtained from numerical simulations or experiments at unknown parameter values. Hence, in crystal growth, these prediction models work as substitutes for traditional numerical simulations and experiments. Tsunooka *et al.*[35] demonstrated a fast prediction model by neural network

for the TSSG-SiC system. In their research, the supersaturation and flow vectors in the melt were accurately predicted in a short time. Yu *et al.*[36] developed a prediction model with more input parameters by neural network and optimized the geometric design using the model. Isono *et al.*[37] optimized the flow distributions in the melt of the TSSG method by using a neural network.
However, a lot of training data is required for accurate models by neural networks, which does not lead to a decrease in the calculation time. Moreover, a neural network just learns the relationship between the input parameters and output parameters. Therefore, results predicted by neural networks do not always satisfy physical consistency and are called black-box prediction, which does not lead to white-box control. Thus, we have introduced the following methods for crystal growth in order to achieve fast and white-box suggestions of the control method.

1.3.2 Bayesian optimization

The fast and white-box suggestion control methods, the challenges mentioned above, is achieved by Bayesian optimization[38] with numerical simulation. The advantage of the optimization algorithm is the efficiency of searching the optimal parameter due to the balance of exploration and exploitation. The balance is achieved by a probabilistic prediction model called Gaussian process regression (GPR) (see section 2.2.3) and acquisition functions (see section 2.2.4).
In GPR, by assuming that the objective function follows Gaussian distribution, the mean and standard deviation values of the prediction are obtained as shown in Fig. 1.4. In the search space with few sampling points, the confidence bound becomes wide. On the other hand, in the search space with many sampling points, it becomes small. Thus, the uncertainty of the prediction can be expressed by the standard deviation from GPR. By selecting parameters based on GPR predictions, the balance of exploration and exploitation is achieved. When exploitation is preferred, the parameters showing a high value of the objective function in a region of low uncertainty are selected. On the contrary, when exploration is preferred, an objective function with the parameters showing a wide confidence bound is selected. The acquisition function plays this role automatically. Acquisition functions are often used to select parameters so that the probability of improving the objective function or its expected value is maximized.
The GPR is used only to determine the values of parameters and results at the parameter values are obtained by numerical simulation or experiment satisfying physical consistency. Thus, The rapid and white suggestion of control methods can be achieved. The optimization algorithm has been utilized in various areas due to its general utility[39], but the application to the crystal growth process had not been conducted.

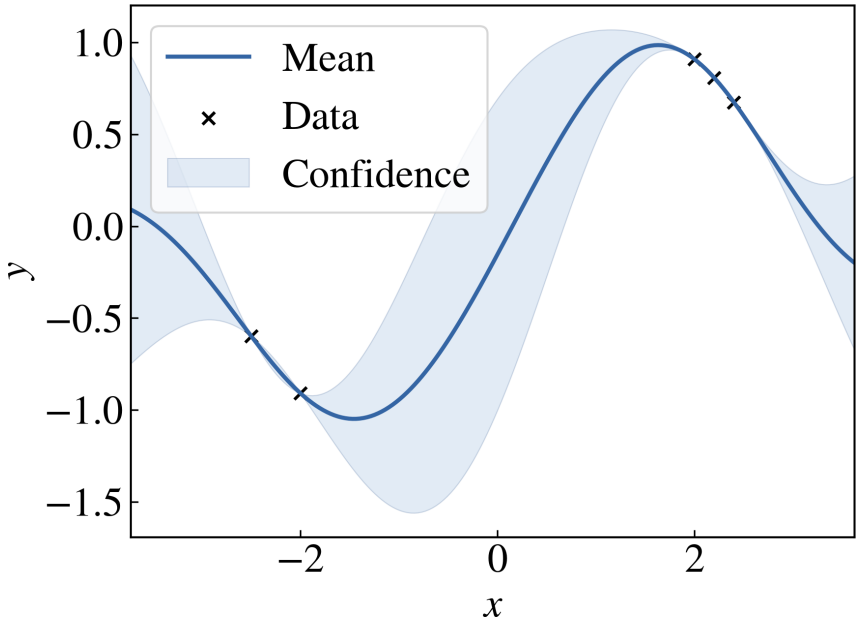

Figure 1.4. The example of Gaussian process regression. The true function is $y = \sin(x)$. The mean value of prediction is represented by the blue line. The confidence bound is obtained by $1.96\sigma_{\mathrm{GPR}}$, where σ_{GPR} is the standard deviation of the prediction.

1.3.3 Explainable AI

Explainable AI (XAI)[40, 41] makes it easy to analyze the transport phenomena showing nonlinearity distribution. The leading role of XAI is that it gives interpretations of predictions obtained from neural networks. XAI has been developed to make neural networks more transparent, especially in image recognition since their black-box nature poses anxiety when applied to practical systems. In the application for image recognition, the region affecting their recognition is usually highlighted on images by XAI, which leads to visualization of the basis of their recognition. As the importance of accountability has increased, various XAIs with different targets have been developed[42, 43, 44, 45, 46]. However, the application for transport phenomena in crystal growth had not been conducted.

Under this situation, in this study, a gradient-based explanation method called smoothGrad[47] was used as XAI. The choice of smoothGrad was motivated by its simplicity. In this method, the gradient of the model's output with respect to its input parameters called sensitivity was used for the explanation. High gradient values for specific input parameters indicate that these parameters have a significant influence on the model's output. As a result, the gradient works as a fundamental basis for understanding and interpreting the model's predictions.

We focused our attention on the concept of providing interpretations for neural networks. In order to detect migratory phenomena that affect crystal growth, we interpreted a neural network that learned the relationship between crystal quality and migratory phenomena. This approach

allows for the easy analysis of transport phenomena with non-linear distributions, which leads to the white-box control.

1.3.4 Physics Informed Neural Networks

Decreasing the calculation time by numerical simulation is achieved by fast and white-box predictions of Physics Informed Neural Networks (PINNs). The advantages of PINNs are summarized as follows:

- Prediction satisfies physical laws (white-box predictions)
- Prediction speed is as fast as that of a regular neural network.
- Applicable to shape changes
- After learning once, everyone can immediately make predictions simply by changing the input parameter values.

These advantages are achieved by the property of PINNs to learn the laws of physics.
PINNs have been attracted from the literature[48, 49] showing their effectiveness. Conventional neural networks learn the relationship between the input and output values in training data by minimizing differences between answer and predicted values called loss function. On the other hand, PINNs learn physical laws by minimizing loss functions indicating errors of physical laws. The loss function of PINNs, usually including partial differentiation, is calculated by auto-differentiation[50] of neural networks. After training, the predicted values by PINNs satisfy the physical laws.
Numerous studies have been conducted using PINNs, especially in fluid dynamics[51, 52, 53]. For instance, PINNs have been used to solve incompressible Navier-Stokes equations[54], Reynolds-averaged Navier-Stokes equations[55], buoyancy-induced flows[56]. Surrogate models have been also developed by PINNs[57, 58, 59, 60]. Moreover, various derivations of PINNs have been introduced[61, 62]. However, because researchers in data science and fluid dynamics, where PINNs have been studied actively, focus on investigating the potential of PINNs, most studies involve simple systems and have little application to complex systems in engineering.
Under this situation, we introduced PINNs to the crystal growth system in order to achieve fast predictions of the transport phenomena in the crystal growth system. Due to the advantages, PINNs allow fast and white-box prediction during experiments, which leads to a decrease in the calculation time. Thus, utilizing PINNs leads to fast and white-box suggestions of control methods.

1.4 Objective

The objective of this study is to develop fast and white-box control methods to overcome the problem mentioned in section 1.1 by data science (Bayesian optimization, XAI, and PINNs) and transport phenomena. The crystal growth system was selected as an example of a target system. Bayesian optimization allows the fast suggestion of the control method due to its efficiency. XAI makes it easy to analyze the transport phenomena by the automatic detection of key phenomena. The fast analysis can be achieved by PINNs because of its fast prediction satisfying physical laws. By combining these methods, the fast and white-box suggestions of control methods are achieved.

1.5 Thesis outline

Chapter 2 presents the fast and white-box suggestion of the control method by Bayesian optimization and numerical simulation in the TSSG method for SiC crystals. Numerical simulation for the TSSG method was first presented. Gaussian process regression and acquisition functions used in Bayesian optimization were then presented. The results of the optimization were shown, and the initial and optimal states obtained by the optimization with numerical simulation were analyzed. From the analysis, the melt flow affecting the crystal growth rates was discussed qualitatively.
In Chapter 3, XAI for analysis of the transport phenomena in the TSSG method for SiC crystal was introduced. Firstly, the concept of XAI was explained, and then the numerical method for Convolutional neural network (CNN), sensitivity tensor, and SmoothGrad utilized in the XAI were introduced. After that, the results of XAI were shown, and the efficiency of the proposed method was proved.
In Chapter 4, PINNs for the Czhocralski method of Si crystals were introduced. Numerical simulation for the Cz method and PINNs satisfying the physical laws of the Cz method were explained. The prediction examples of PINNs were compared with numerical simulation. The accuracy of PINNs was verified in the whole parameter region.
Chapter 5 concludes this thesis. Future works were also listed.

Chapter 2

Bayesian optimization for the TSSG method of SiC crystals

In this Chapter, the magnetic field and the rotation of the seed crystal were optimized for suppressing the Marangoni convection affecting crystal growth rates adversely by an effective optimization algorithm called Bayesian optimization and numerical simulation. As mentioned in section1.2.2, the Marangoni convection can be suppressed by the magnetic fields. However, optimizing the magnetic field has been challenging due to the need to perform numerous numerical simulations with varying parameter values and an effective optimization achieving optimization with few numerical simulations has been required. Thus, we employed Bayesian optimization for effective optimization with numerical simulation. As a result, the optimal magnetic field and the rotation speed of the seed crystal were obtained with a small number of numerical simulations by Bayesian optimization, and the analysis of the transport phenomena of the optimal state was carried out by numerical simulation, which leads to fast and white-box suggestions of control methods.

2.1 Numerical simulation

The numerical configuration of the TSSG method is shown in Figure 2.1. The configuration is the same as the previous studies[23, 63, 64, 65, 66]. The rod, heating wall, crucible, and bottom are made of graphite, while the shield is made of quartz. The gas region is filled with helium and the chamber is sealed. The depth of the SiC crystal , the meniscus and the shape change of the melt surface were not taken into account.

2.1.1 Integrate Process Model (IPM)

The global simulation of the TSSG method was separated into 3 steps by Integrate Process Model (IPM)[67, 68, 69];

1. Electromagnetic field by RF-coil
2. Temperature field in the whole furnace
3. Transport phenomena in the Si-C melt ,

since the length and time scales of the electromagnetic phenomena and thermal conduction in the whole domain and the transport phenomena in the Si-C melt. The time scales of the electromagnetic phenomena, conduction, and convection are expressed as

$$1/\omega \approx 10^{-5}\ \mathrm{s} \tag{2.1}$$

$$h_{\mathrm{chamber}}^2/\alpha_{\mathrm{He}} \approx 10^{2}\ \mathrm{s} \tag{2.2}$$

$$R_{\mathrm{c}}/U_{\mathrm{ref}} \approx 10^{-6}\ \mathrm{s}, \tag{2.3}$$

respectively, where ω is the frequency of the current in the coil, h_{chamber} the height of the chamber, α_{He} the thermal diffusivity of Helium, r_{c} the crucible radius, and the U_{ref} $(= \sigma_T \Delta T/\mu)$ the reference velocity of the Marangoni convection, σ_T the surface tension coefficient, ΔT the temperature difference in the Si-C melt, and μ_{Si} the viscosity of the liquid Si.
The outline of the IPM is shown in Fig. 2.2. In the first step, the Joule heat density and the Lorentz force are obtained. In the second step, the temperature field in the whole furnace is obtained by using the Joule heat density. Finally, the transport phenomena in the Si-C melt are obtained by using the Joule heat density, Lorentz force, and the temperature boundary conditions obtained from the previous steps. The numerical model and code were validated in Ph.D. thesis[70] and used in the literatures[23, 63, 64, 65, 66, 27]. The analysis method for each step is described in detail below.

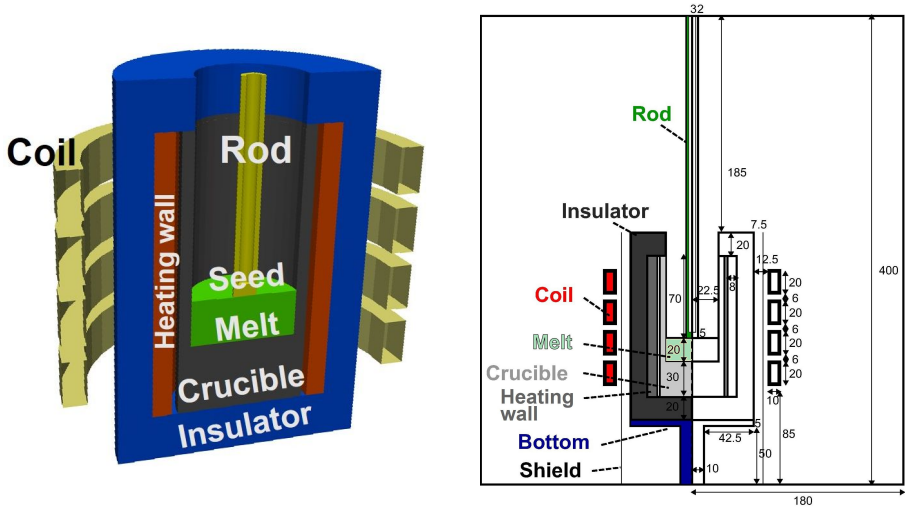

Figure 2.1. The 3-dimensional (left) and cross-sectional view of the numerical configuration (right) of the TSSG method utilized in this study. The length is written in the right figure, where the unit is mm.

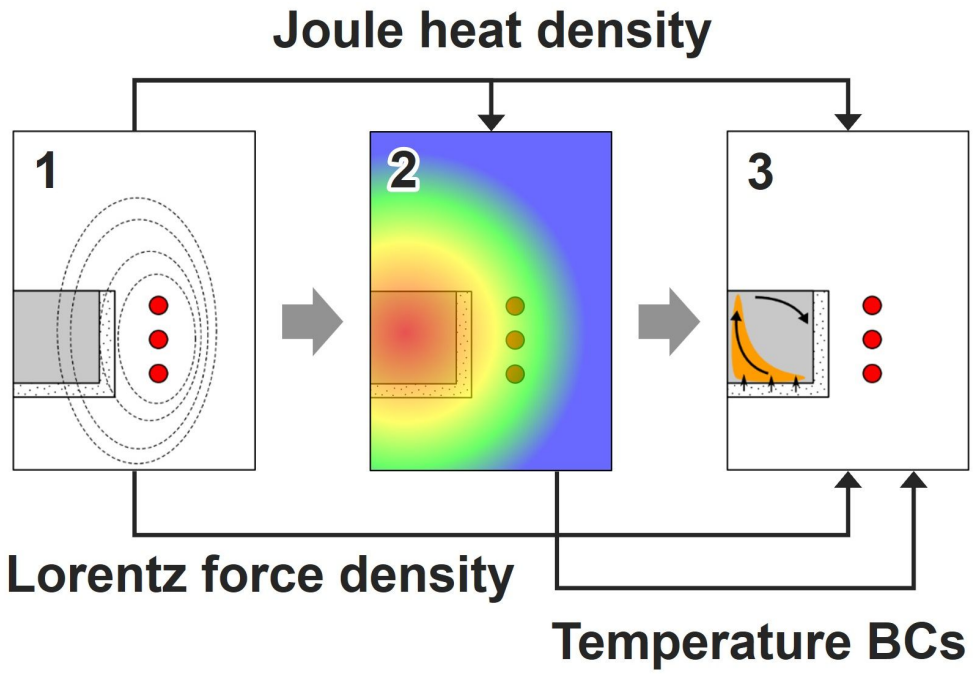

Figure 2.2. The outline of the Integrate Process Model (IPM). In step 1, Joule heat and Lorentz force densities are obtained and used in the following analysis. In step 2 temperature boundary conditions for the next step are obtained using the Joule heat density in step 1. In step 3, the detailed transport phenomena in the Si-C melt are calculated using the Joule heat and Lorentz force densities from step 1 and the temperature boundary conditions from step 2.

(i) Electromagnetic field by RF-coil

The governing equations for the electromagnetic field are derived from Maxwell's equations:

$$\nabla \cdot \boldsymbol{E}_{\mathrm{coil}} = \frac{\rho_{\mathrm{e}}}{\epsilon} = 0 \tag{2.4}$$

$$\nabla \cdot \boldsymbol{B}_{\mathrm{coil}} = 0 \tag{2.5}$$

$$\nabla \times \boldsymbol{E}_{\mathrm{coil}} = -\frac{\partial \boldsymbol{B}_{\mathrm{coil}}}{\partial t} \tag{2.6}$$

$$\nabla \times \boldsymbol{B}_{\mathrm{coil}} = \mu_{\mathrm{e}} \boldsymbol{J} \tag{2.7}$$

where $\boldsymbol{E}_{\mathrm{coil}}$ is the electric field, ρ_{e} the charge density, $\boldsymbol{B}_{\mathrm{coil}}$ the magnetic field, t the time, μ_{e} the magnetic permeability and the $\boldsymbol{J}$ the current density and the physical properites ρ_{e}, μ_{e}, and σ were assumed as constant. In the TSSG system, there is no charge ($\rho_{\mathrm{e}} = 0$), so RHS of Eq. (2.4) becomes 0. The source term of Eq. (2.5), $\epsilon \partial \boldsymbol{E}_{\mathrm{coil}} / \partial t$, is negligible, as it becomes significant only when the wavelength is nearly the same as or shorter than the length of the target system. Then, the vector and scalar potentials $\boldsymbol{A}_{\mathrm{coil}}$, ϕ_{coil} are introduced.

$$\boldsymbol{B}_{\mathrm{coil}} = \nabla \times \boldsymbol{A}_{\mathrm{coil}} \tag{2.8}$$

$$\boldsymbol{E}_{\mathrm{coil}} = -\nabla \phi_{\mathrm{coil}} - \frac{\partial \boldsymbol{A}_{\mathrm{coil}}}{\partial t} \tag{2.9}$$

Due to the definition of potentials, Eqs. (2.5), (2.6) are automatically satisfied.
On the other hand, Eq. (2.7) transforms as follows;

$$\begin{aligned} \nabla \times (\nabla \times \boldsymbol{A}_{\mathrm{coil}}) &= \mu_{\mathrm{e}} \boldsymbol{J} \\ \cancel{\nabla (\nabla \cdot \boldsymbol{A}_{\mathrm{coil}})} - \nabla^2 \boldsymbol{A}_{\mathrm{coil}} &= \mu_{\mathrm{e}} \boldsymbol{J} \\ \nabla^2 \boldsymbol{A}_{\mathrm{coil}} &= -\mu_{\mathrm{e}} \boldsymbol{J} \end{aligned} \tag{2.10}$$

The current density in the conductors is expressed as

$$\boldsymbol{J} = \sigma_{\mathrm{e}} \left(\boldsymbol{E}_{\mathrm{coil}} + \boldsymbol{u} \times \boldsymbol{B}_{\mathrm{coil}} \right) + \rho_{\mathrm{e}} \boldsymbol{u} \tag{2.11}$$

by Ohm's law, where σ_{e} the electrical conductivity, the $\boldsymbol{u}$ is the flow velocity. Since $(\boldsymbol{u} \times \boldsymbol{B}_{\mathrm{coil}})$ is 2 orders smaller than $\boldsymbol{E}_{\mathrm{coil}}$, $(\boldsymbol{u} \times \boldsymbol{B}_{\mathrm{coil}})$ is negligible[71]. Moreover, $(\rho_{\mathrm{e}} \boldsymbol{u})$ is also negligible due to the no charge ($\rho_{\mathrm{e}} = 0$). Hence, the current density in the conductor becomes

$$\boldsymbol{J} = \sigma_{\mathrm{e}} \boldsymbol{E}_{\mathrm{coil}} \tag{2.12}$$

With the assumption that the current density has only a circumferential component, $\boldsymbol{J}$ is expressed as

$$\boldsymbol{J} = \boldsymbol{e}_\theta J_\theta = \begin{cases} \boldsymbol{e}_\theta J_0 \cos \omega t & \text{in the coil} \\ \boldsymbol{e}_\theta \sigma_{\mathrm{e}} E_{\mathrm{coil},\theta} & \text{in the conductors (Ohm's law)} \\ 0 & \text{in the nonconductors} \end{cases}, \tag{2.13}$$

where J_0 is the peak current density and $\boldsymbol{e}_\theta$ is the unit vector toward to circumferential direction. This assumption leads to $E_{\mathrm{coil},r} = E_{\mathrm{coil},z} = 0$ from Ohm's law, $A_{\mathrm{coil},r} = A_{\mathrm{coil},z} = 0$ from Eq (2.10) and $B_{\mathrm{coil},\theta} = 0$ from the definition of the vector potential. Hence, Eq. (2.10) becomes

$$\begin{aligned} \nabla^2 A_{\mathrm{coil},\theta} &= -\mu_{\mathrm{e}} J_\theta \\ \frac{\partial}{\partial r} \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r A_{\mathrm{coil},\theta} \right) \right] + \frac{\partial^2 A_{\mathrm{coil},\theta}}{\partial z^2} &= -\mu_{\mathrm{e}} J_\theta, \end{aligned} \tag{2.14}$$

where r is the radial coordinate, the z the vertical coordinate, and $\partial^2 A_{\mathrm{coil},\theta} / \partial \theta^2 = 0$ with the axisymmetric assumption.
Here, the magnetic stream function ψ is introduced as follows;

$$\psi \equiv r A_{\mathrm{coil},\theta} \tag{2.15}$$

By introducing the magnetic stream function to Eq. (2.14),

$$\frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial \psi}{\partial r} \right) + \frac{\partial}{\partial z} \left(\frac{1}{r} \frac{\partial \psi}{\partial z} \right) = L(\psi) = -\mu_{\mathrm{e}} J_\theta \tag{2.16}$$

In order to obtain a time-harmonic solution expressed by sine waves, the magnetic stream function is expressed as;

$$\psi\left(r,z,t\right)=C\left(r,z\right)\cos\omega t+S\left(r,z\right)\sin\omega t, \tag{2.17}$$

where C and S is the in-phase and out-of-phase components of the solution, respectively. By introducing Eq. (2.17) to Eq. (2.14),

$$L\left(C\right)\cos\omega t+L\left(S\right)\sin\omega t=-\mu_{\mathrm{e}}J_{\theta} \tag{2.18}$$

is derived. By introducing Eq. (2.17) to Ohm's law with axisymmetric assumption, the current density in the conductor

$$\begin{aligned} J_{\theta} &= \sigma_{\mathrm{e}}E_{\theta} \\ &= -\sigma_{\mathrm{e}}\left(\cancel{-\nabla_{\theta}\phi_{\mathrm{coil}}}-\frac{\partial A_{\mathrm{coil},\theta}}{\partial t}\right) \\ &= -\sigma_{\mathrm{e}}\frac{1}{r}\frac{\partial\psi}{\partial t} \\ &= -\sigma_{\mathrm{e}}\left(C\sin\omega t-S\cos\omega t\right) \end{aligned} \tag{2.19}$$

is derived.
Thus, the governing equations of the electromagnetic field are derived as

$$\begin{aligned} \frac{\partial^2 C}{\partial r^2}+\frac{\partial^2 C}{\partial z^2}-\frac{1}{r}\frac{\partial C}{\partial r} &= \mu_e\sigma_e\omega S \quad \text{(in the condoctor)} \\ &= -\mu_e J_0 r \quad \text{(in the coil)} \\ &= 0 \qquad \text{(in the nonconductors)} \end{aligned} \tag{2.20}$$

$$\begin{aligned} \frac{\partial^2 S}{\partial r^2}+\frac{\partial^2 S}{\partial z^2}-\frac{1}{r}\frac{\partial S}{\partial r} &= -\mu_e\sigma_e\omega C \ \text{(in the condoctor)} \\ &= 0 \qquad \text{(in the coil)} \\ &= 0 \qquad \text{(in the nonconductors).} \end{aligned} \tag{2.21}$$

For the analysis, the weak coupling method was used for solving the governing equations in the multi-regions. The boundary conditions along the center axis and on the chamber are

$$C=S=0, \tag{2.22}$$

since the electromagnetic field goes to zero on the chamber and axisymmetry is assumed. On the interface of the i, j regions, the electromagnetic field is continuous. Thus the boundary

Table 2.1 The electromagnetic properties used in the TSSG system

property (symbol)	value	unit
magnetic permeability (μ_e)	1.2566×10^{-6}	Hm^{-2}
Peak current density (J_0)	6.285×10^{6}	Am^{-2}
Frequency of electric current (ω)	1.5708×10^{5} (25kHz)	rads^{-1}

conditions between the interface are

$$C_i = C_j \tag{2.23}$$

$$S_i = S_j \tag{2.24}$$

$$\frac{\partial C}{\partial \boldsymbol{n}_i} = -\frac{\partial C}{\partial \boldsymbol{n}_j} \tag{2.25}$$

$$\frac{\partial S}{\partial \boldsymbol{n}_i} = -\frac{\partial S}{\partial \boldsymbol{n}_j}, \tag{2.26}$$

where $\boldsymbol{n}_i,\ \boldsymbol{n}_j$ are the normal vector to the interface $i,\ j$ ($\boldsymbol{n}_i = -\boldsymbol{n}_j$).
The cycle-averaged Joule heat and Lorentz force densities are calculated using the following equations in IPM since it is difficult to resolve the time resolution due to the high frequency of the current in the coil.

$$Q = \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} \frac{J_\theta^2}{\sigma_e}\, dt = \frac{\sigma_e \omega^2}{2r^2}\left(C^2 + S^2\right) \tag{2.27}$$

$$\begin{aligned} \boldsymbol{F}_{\mathrm{E}} &= \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} \boldsymbol{J} \times \boldsymbol{B}_{\mathrm{coil}}\, dt \\ &= \left(\frac{\sigma_e \omega}{2r^2}\left(C\frac{\partial S}{\partial r} - S\frac{\partial C}{\partial r}\right), 0, \frac{\sigma_e \omega}{2r^2}\left(C\frac{\partial S}{\partial z} - S\frac{\partial C}{\partial z}\right)\right) \end{aligned} \tag{2.28}$$

Here, the peak current density J_0 is calculated from

$$\begin{aligned} J_0 = \frac{I_0}{S_{\mathrm{coil}}} &= \frac{\sqrt{2} I_{\mathrm{eff}}}{S_{\mathrm{coil}}} \\ &= \frac{\sqrt{2} \times 360\ \mathrm{A}}{81\ \mathrm{mm}^2} = 6.285 \times 10^6\ \mathrm{A/m}^2, \end{aligned} \tag{2.29}$$

where I_0 is the peak current, S_{coil} is the cross-sectional area where the current flows inside the coil, and I_{eff} is the effective current. The current only runs in the 1.5 mm area around the edge of the coil as shown in Fig. 2.3, so $S_{\mathrm{coil}} = \{10 \times 20 - (10 - 1.5 \times 2) \times (20 - 1.5 \times 2)\} = 81\ \mathrm{mm}^2$. The electromagnetic properties are listed in table 2.1. The numerical simulation of the electromagnetic field is conducted by 2-dimensional axisymmetry analysis.

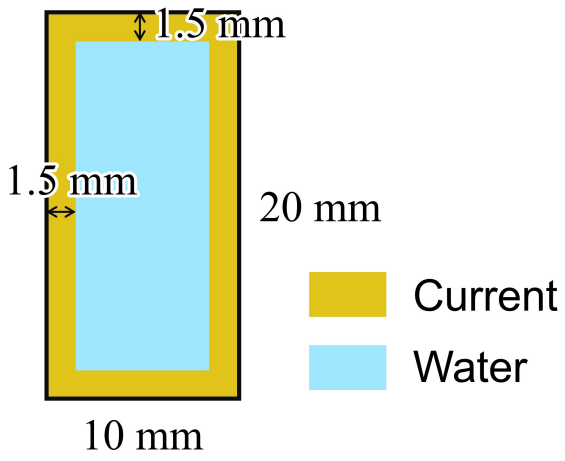

Figure 2.3. The region where the current (yellow) and water (light blue) run in a coil.

(ii) Temperature field in the whole furnace

The temperature field in the whole furnace is solved using the Joule heat density Q obtained from step (i). In the analysis, the weak coupling method was also used for considering multi-regions. Since radiative heat transfer is dominant due to the high temperature (about 2000 K) in the gas region and the melt region is small compared with the whole region, convective heat transfer is negligible. Thus, only radiation and thermal conduction were solved. The governing equation for thermal conduction is given by

$$\nabla \cdot (k \nabla T) + Q = 0, \tag{2.30}$$

where k is the thermal conductivity, T the temperature.

The governing equations for radiation are derived from these equations;

$$q_i = \frac{E_{\mathrm{b},i} - J_{\mathrm{rad},i}}{\rho_i / A_i \epsilon_i} = A_i \left(J_{\mathrm{rad},i} - G_i \right) \tag{2.31}$$

$$A_i G_i = A_i \sum_j J_{\mathrm{rad},i} F_{ij}, \tag{2.32}$$

where q_i is the net rate of heat transfer leaving surface i, $F_{i,j}$ the view factor from i to j surfaces and $E_{\mathrm{b},i}$, $J_{\mathrm{rad},i}$, ρ_i, A_i, ϵ_i, and G_i are the emissive power of black body, the radiosity, the reflectivity, the area, emissivity, and the irradiation of surface i. The derivation of the equations is mentioned in [72]. $E_{\mathrm{b},i}$ is calculated by Stefan-Boltzmann law ($E_{\mathrm{b},i} = \sigma_{\mathrm{s.b.}} T_i^4$, where T_i and $\sigma_{\mathrm{s.b.}}$ are the temperature of surface i and Stefan-Boltzmann coefficient).

Combining Eqs. (2.31) and (2.32),

$$q_i = A_i \left[J_{\mathrm{rad},i} - \sum_j F_{i,j} J_{\mathrm{rad},i} \right] \tag{2.33}$$

$$= \frac{A_i \epsilon_i}{\rho_i} \left(E_{\mathrm{b},i} - J_{\mathrm{rad},i} \right) \tag{2.34}$$

are obtained. Equation (2.33) is expressed as

$$J_{\mathrm{rad},i} = \sum_j F_{ij} J_{\mathrm{rad},i} = \frac{q_i}{A_i} \tag{2.35}$$

and by multiplying both sides by $(1 - \epsilon_i)$,

$$\begin{aligned} (1 - \epsilon_i) J_{\mathrm{rad},i} - (1 - \epsilon) \sum_j F_{ij} J_{\mathrm{rad},i} &= (1 - \epsilon_i) \frac{q_i}{A_i} \\ J_i - (1 - \epsilon_i) \sum_j F_{ij} J_{\mathrm{rad},i} &= \epsilon_i J_{\mathrm{rad},i} + \frac{\rho_i}{A_i} q_i \end{aligned} \tag{2.36}$$

is obtained, where $(1 - \epsilon_i) = (1 - \alpha_i) = \rho_i$ due to assuming gray and opaque surface.
Here, due to

$$\frac{\rho_i}{A_i} q_i = \epsilon_i \left(E_{\mathrm{b},i} - J_{\mathrm{rad},i} \right) \tag{2.37}$$

from Eq. (2.34), finally the governing equation for radiosity is given as

$$J_{\mathrm{rad},i} - (1 - \epsilon_i) \sum_j F_{ij} J_{\mathrm{rad},i} = \epsilon_i E_{\mathrm{b},i}. \tag{2.38}$$

The heat flux is derived from Eq. (2.33) with Eq. (2.38) as

$$\begin{aligned} \frac{q_i}{A_i} &= J_i - \sum_j F_{ij} J_{\mathrm{rad},i} \\ &= \epsilon_i \left(E_{\mathrm{b},i} - \sum_j F_{ij} J_{\mathrm{rad},i} \right). \end{aligned} \tag{2.39}$$

The temperature of surface T_i is obtained by Eq. (2.30) and then radiosity is calculated by (2.38) with T_i. Subsequently, the heat flux is calculated by Eq. (2.39).
In order to reduce the computational cost, the calculations of the radiation (Eqs. (2.38) and (2.39)) were performed by coarse-graining the neighboring grid as a single plane. The 3-dimensional calculation was conducted for analyzing the radiation.
The boundary conditions are given as follows;
On the chamber,

$$T = 313 \text{ K}, \tag{2.40}$$

On the solid interfaces,

$$T_i = T_j \tag{2.41}$$
$$k_i \nabla T \cdot \boldsymbol{n}_i = k_j \nabla T \cdot \boldsymbol{n}_j, \tag{2.42}$$

where i and j are the solid interfaces after coarse-graining, and on the gas-solid or liquid interfaces,

$$T_i = T_j \tag{2.43}$$
$$(-k\nabla T)_i = (-k\nabla T)_j + q_j, \tag{2.44}$$

where i and j are the gas and liquid or solid interfaces, respectively. The heat flux from the free surface of the Si-C melt is calculated as the area-averaged value by

$$q_{\mathrm{rad}} = -\frac{k}{S_{\mathrm{surface}}} \int_S \frac{\partial T}{\partial z} dS \tag{2.45}$$

where S_{surface} is the area of the free surface of the melt.

(iii) Transport phenomena in the Si-C melt

In this analysis, only the region of Si-C melt was considered. The melt was assumed to be an incompressible Newtonian fluid with no deformation of the free surface. All physical properties were assumed to be constant, and the Bousinesq approximation was used for buoyancy.
The cusp magnetic field $\boldsymbol{B}_{\mathrm{cusp}}$ same as Wang *et al.* [27] was applied to Si-C melt

$$\boldsymbol{B}_{\mathrm{cusp}} = \frac{B_0}{R_{\mathrm{c}}} \left(r, 1, 2(H_0 - z)\right), \tag{2.46}$$

where B_0 is the magnitude of the magnetic field, R_{c} the crucible radius, and H_0 the applied position of the magnetic field (see Fig. ?).
The Lorentz force by the cusp magnetic field is given by

$$\boldsymbol{F}_{\mathrm{cusp}} = \boldsymbol{J}_{\mathrm{cusp}} \times \boldsymbol{B}_{\mathrm{cusp}}, \tag{2.47}$$

where $\boldsymbol{J}_{\mathrm{cusp}}$ is the induced current density by the cusp magnetic field and is obtained by Ohm's law;

$$\begin{aligned} \boldsymbol{J}_{\mathrm{cusp}} &= \sigma_{\mathrm{e}} \left(\boldsymbol{E}_{\mathrm{cusp}} + \boldsymbol{u} \times \boldsymbol{B}_{\mathrm{cusp}}\right) \\ &= \sigma_{\mathrm{e}} \left(\boldsymbol{E}_{\mathrm{cusp}} + \boldsymbol{u} \times \boldsymbol{B}_{\mathrm{cusp}}\right) \\ &= \sigma_{\mathrm{e}} \left(-\nabla \phi_{\mathrm{cusp}} + \boldsymbol{u} \times \boldsymbol{B}_{\mathrm{cusp}}\right) \end{aligned} \tag{2.48}$$

where E_{cusp} is the electric field by the cusp magnetic field and ϕ_{cusp} is the electrostatic potential defined as $\boldsymbol{E}_{\mathrm{cusp}} = -\nabla \phi_{\mathrm{cusp}}$. The electrostatic potential is obtained by solving the following

equation derived from the continuity law of the current density;

$$\begin{aligned} \nabla \cdot \boldsymbol{J}_{\mathrm{cusp}} &= 0 \\ \nabla \cdot \left(-\nabla \phi_{\mathrm{cusp}} + \boldsymbol{u} \times \boldsymbol{B}_{\mathrm{cusp}}\right) &= 0 \\ \nabla^2 \phi_{\mathrm{cusp}} &= \nabla \cdot \left(\boldsymbol{u} \times \boldsymbol{B}_{\mathrm{cusp}}\right) \end{aligned} \tag{2.49}$$

Thus, the governing equations are given as

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\, \boldsymbol{u} = -\frac{1}{\rho_{\mathrm{Si}}} \nabla p + \nu \nabla^2 \boldsymbol{u} + \frac{\boldsymbol{F}_{\mathrm{E}}}{\rho_{\mathrm{Si}}} - \boldsymbol{g}\beta\,(T - T_0) + \frac{\boldsymbol{J}_{\mathrm{cusp}} \times \boldsymbol{B}_{\mathrm{cusp}}}{\rho_{\mathrm{Si}}} \tag{2.50}$$

$$\nabla \cdot \boldsymbol{u} = 0 \tag{2.51}$$

$$\frac{\partial T}{\partial t} + (\boldsymbol{u} \cdot \nabla)\, T = \alpha \nabla^2 T + \frac{Q}{\rho_{\mathrm{Si}} C_p} \tag{2.52}$$

$$\frac{\partial c}{\partial t} + (\boldsymbol{u} \cdot \nabla)\, c = D \nabla^2 c \tag{2.53}$$

$$\nabla^2 \phi_{\mathrm{cusp}} = \nabla \cdot \left(\boldsymbol{u} \times \boldsymbol{B}_{\mathrm{cusp}}\right), \tag{2.54}$$

where $\boldsymbol{u}$ is the melt velocity, t time, ρ_{Si} the density of Si melt, p pressure, ν kinetic viscosity of Si melt, $\boldsymbol{F}_{\mathrm{E}}$ the Lorentz force by current in the coil, $\boldsymbol{g}$ the gravitational acceleration, β the thermal expansion coefficient, T_0 the reference temperature, $\boldsymbol{J}_{\mathrm{cusp}}$ the current density by the cusp magnetic field, $\boldsymbol{B}_{\mathrm{cusp}}$ the cusp magnetic field, α the thermal diffusivity of the Si melt, Q the Joule heat density obtained from step (i), C_p the heat capacity, c the concentration of carbon in Si melt, D the diffusion coefficient of carbon in Si melt and ϕ the electrical potential.
The effect of the magnetic field generated by the current in the coil on the cusp magnetic field is negligible in the same discussion in Wang. *et al.*[27] as follows. Muiznieks *et al.*[73] introduced the estimation of the magnitude of the magnetic field by alternating current as

$$|\boldsymbol{B}_{\mathrm{E}}| = \sqrt{\frac{|2\boldsymbol{F}_{\mathrm{E}}|}{\sigma_{\mathrm{e}} \pi \omega R_{\mathrm{c}}}}, \tag{2.55}$$

where R_{c} is the crucible radius. $|\boldsymbol{B}_{\mathrm{E}}|_{\max} = 3$ mT is given under $|\boldsymbol{B}_{\mathrm{E}}|_{\max} = 19,700\ \mathrm{Nm}^{-3}$. The magnitude of the magnetic field by the current in the coil is much lower than the magnitude of the cusp magnetic field. Thus, the effect of the magnetic field from the current in the coil is negligible.
The boundary conditions for the velocity are given as follows;
On the crucible wall,

$$\boldsymbol{u} = \boldsymbol{0} \ \text{(no slip condition)} \tag{2.56}$$

and on the seed crystal,

$$\boldsymbol{u} = r\omega_{\mathrm{s}}, \tag{2.57}$$

where ω_s is the angular velocity of the seed crystal.
On the free surface,

$$\mu \frac{\partial u}{\partial z} = \sigma_T \frac{\partial T}{\partial r}, \tag{2.58}$$

where μ is the viscosity and σ_T is the thermal surface tension coefficient.
Along the center axis,

$$u_r = 0 \tag{2.59}$$

$$u_\theta = 0 \tag{2.60}$$

$$\frac{\partial u_z}{\partial r} = 0 \tag{2.61}$$

For the concentration of carbon, boundary conditions are expressed as

$$c = c_{\text{eq.}} \tag{2.62}$$

$$c_{\text{eq.}} = \frac{\rho_{\text{Si}}}{M_{\text{Si}}} \frac{x_{\text{e,eq.}}}{1 - x_{\text{e,eq.}}} \tag{2.63}$$

$$x_{\text{e,eq.}} = 10^{2.714 - \frac{10623}{T}} \tag{2.64}$$

on the crucible wall and seed surface, where $c_{\text{eq.}}$ is the equilibrium concentration, ρ_{Si} the density of liquid Si, M_{Si} the molar weight of Si, and $x_{\text{e,eq.}}$ the carbon molar fraction obtained by a linear approximation to the results of several experiments[74, 75].
On the free surface,

$$\frac{\partial c}{\partial z} = 0 \tag{2.65}$$

and along the center axis,

$$\frac{\partial c}{\partial r} = 0 \tag{2.66}$$

For the temperature, the Dirichlet boundary conditions obtained from step (ii) were used along the crucible wall and seed crystal surface. On the free surface, the Neumann boundary condition expressing the radiative heat transfer was used as follows;

$$-k \frac{\partial T}{\partial z} = q_{\text{rad}}. \tag{2.67}$$

As the initial condition, a fully developed flow (taken from Yamamoto *et al.*[23]) subjected to an applied cusp magnetic field was assumed. The crystal growth rate is defined as

$$v_{\text{g}} = -\frac{D M_{\text{SiC}}}{\rho_{\text{SiC}}} \boldsymbol{n} \cdot \nabla c|_{\text{seed}}, \tag{2.68}$$

where M_{SiC} is the molar weights of SiC, ρ_{SiC} is the density of SiC, and $\boldsymbol{n}$ is the normal vector directed to the crystal surface. The concentration gradient was calculated as the average value

during 100 seconds between 50 and 150 seconds from the initial state in order to consider concentration under the effect of the cusp magnetic field.
The definition (Eq. (2.68)) is derived from the assumption that mass transport is the rate-controlling step in three stages of crystal growth: 1. Mass transport, 2. Surface kinetics, 3. Elimination of latent heat, due to the low solubility of carbon to Si. The definition was validated by comparing the crystal growth rate obtained by experiment[76]. The physical properties are listed in Tabel 2.2. The numerical simulation was conducted by using OpenFOAM[77].

2.2 Bayesian optimization

Bayesian optimization is an optimization algorithm using Gaussian process regression and acquisition functions. The parameters to be optimized are magnetic and rotational parameters ($B_0,\ H_0,\ \omega_\mathrm{s}$) and objective function f was defined as

$$f = M^*_\mathrm{gr} - \sigma^*_\mathrm{gr}, \tag{2.69}$$

where M^*_gr and σ^*_gr are the normalized values of the mean and standard deviation of the crystal growth rate, respectively, defined as

$$M^* = M_\mathrm{gr}/M_0 \quad , \quad \sigma^* = \sigma_\mathrm{gr}/\sigma_0 \tag{2.70}$$

$$M_\mathrm{gr} = \frac{1}{L}\int_0^L v_\mathrm{g}\,(r)\ dr \quad , \quad \sigma_\mathrm{gr} = \sqrt{\frac{1}{L}\int_0^L \left\{v_\mathrm{g}\,(r) - M_\mathrm{gr}\right\}^2\ dr} \tag{2.71}$$

$$M_0 = M_\mathrm{gr}|_{\boldsymbol{B}=\boldsymbol{0},\omega_\mathrm{s}=0} \quad , \quad \sigma_0 = \sigma_\mathrm{gr}|_{\boldsymbol{B}=\boldsymbol{0},\omega_\mathrm{s}=0} \tag{2.72}$$

and L is the target length of the optimization. In this study, the condition $L = 4$ mm was used in order to eliminate the quite nonuniform region near the seed edge ($r = 4 \sim 5$ mm). The objective function is also expressed as

$$\begin{aligned} f &= M^* - \sigma^* \\ &= \frac{M_\mathrm{gr}}{M_0} - \frac{\sigma_\mathrm{gr}}{\sigma_0} \\ &= \frac{M_\mathrm{gr} - M_0}{M_0} + \frac{\sigma_0 - \sigma_\mathrm{gr}}{\sigma_0} \\ &= \frac{\Delta M}{M_0} + \frac{\Delta\sigma}{\sigma_0}. \end{aligned} \tag{2.73}$$

Therefore, the physical meaning of the objective function is the sum of the improvement ratios of the mean and standard deviation values. By maximizing the objective function, the parameters achieving high and uniform crystal growth rates are obtained since mean and standard deviation values indicate the magnitude and uniformity of the crystal growth rates, respectively.
The optimization algorithm is conducted by the following step.

1. The initial dataset $\mathcal{D}_{\mathrm{N}}=\left\{\left(\boldsymbol{x}^{(1)},\ y^{(1)}\right),\left(\boldsymbol{x}^{(2)},\ y^{(2)}\right),\ \ldots,\left(\boldsymbol{x}^{(N)},\ y^{(N)}\right)\right\}$ is made, where $\boldsymbol{x}$ is the parameters to be optimized, y is the result obtained from numerical simulation (or experiment) to be maximized (or minimized) and N is the number of the data of initial dataset.

2. The mean $\mu_{\mathrm{GPR}}\left(\boldsymbol{x}^{(N+1)}\right)$ and standard deviation $\sigma_{\mathrm{GPR}}\left(\boldsymbol{x}^{(N+1)}\right)$ values of the prediction are obtained by Gaussian process regression, where $\boldsymbol{x}^{(N+1)}$ is the parameters whose results are unknown (see section 2.2.3).

3. The parameter $\boldsymbol{x}_{\mathrm{next}}$ to be calculated (or experimented with) is selected by using the acquisition function (see section 2.2.4).

4. The values of the objective function $f(=y^{\mathrm{next}})$ with the parameter $\boldsymbol{x}_{\mathrm{next}}$ is obtained through numerical simulation (or experiment).

5. The data $\left(\boldsymbol{x}^{\mathrm{next}},\ y^{\mathrm{next}}\right)$ is added to the initial dataset.

6. Go to step 2.

The initial data set consisted of 8 ($= 2 \cdot 2 \cdot 2$) data combining the maximum and minimum values of the range of parameters shown in Table 2.3. Firstly, optimization for the discrete values of the parameters, as shown in Table 2.3, was performed. Secondly, optimization for the continuous values of the parameters was conducted in the same parameter ranges as the discrete parameters. Python library GpyOpt[80] was used for the Bayesian optimization.

2.2.1 Gaussian process (GP)

The input parameters and the output values are expressed by $\boldsymbol{x} = (x_1,\ x_2,\ \cdots, x_m)^\top$ and y, respectively, where m is the number of the parameters. In this study, $\boldsymbol{x} = (B_0,\ H_0,\ \omega_{\mathrm{s}})$, $m = 3$ and y is the objective function f. In the Gaussian process, y is expressed by the linear combination of weights w and the basis functions ϕ as

$$\begin{aligned} y &= w_0\phi_0(\boldsymbol{x}) + w_2\phi_2(\boldsymbol{x}) + w_3\phi_3(\boldsymbol{x}) + \cdots + w_H\phi_H(\boldsymbol{x}) \\ &= \sum_{i=0}^{H} w_i\phi_i(\boldsymbol{x}) \\ &= \boldsymbol{w}^\top\boldsymbol{\phi}(\boldsymbol{x}), \end{aligned} \tag{2.74}$$

where

$$\boldsymbol{w} = (w_0,\ w_1,\ \cdots,\ w_H)^\top \tag{2.75}$$

$$\boldsymbol{\phi}(\boldsymbol{x}) = (\phi_0(\boldsymbol{x}),\ \phi_1(\boldsymbol{x}),\ \cdots,\ \phi_H(\boldsymbol{x}))^\top, \tag{2.76}$$

and H is the number of the basis function.
For each data in the initial dataset, Eq. (2.74) is written as

$$\begin{pmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(N)} \end{pmatrix} = \begin{pmatrix} \phi_0(\boldsymbol{x}^{(1)}) & \phi_1(\boldsymbol{x}^{(1)}) & \cdots & \phi_H(\boldsymbol{x}^{(1)}) \\ \phi_0(\boldsymbol{x}^{(2)}) & \phi_1(\boldsymbol{x}^{(2)}) & \cdots & \phi_H(\boldsymbol{x}^{(2)}) \\ \vdots & \vdots & \cdots & \vdots \\ \phi_0(\boldsymbol{x}^{(N)}) & \phi_1(\boldsymbol{x}^{(N)}) & \cdots & \phi_H(\boldsymbol{x}^{(N)}) \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_H \end{pmatrix} \tag{2.77}$$

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{w}. \tag{2.78}$$

Here, assumption that $\boldsymbol{w}$ follows the Gaussian distribution with mean $\boldsymbol{0}$ and covariance matrix$\lambda^2 I$;

$$\boldsymbol{w} \sim \mathcal{N}\left(0,\ \lambda^2 I\right) \tag{2.79}$$

is introduced, where $\boldsymbol{I}$ is the identity matrix. The covariance matrix is $\lambda^2 I$, which means that the parameters are independent of each other.
Then Eq. (2.78) indicates the linear transformation of $\boldsymbol{w}$ following Gaussian distribution by the constant matrix $\boldsymbol{\Phi}$. Due to the property that linearly transformed values of a multivariate Gaussian distribution also follow a Gaussian distribution, $\boldsymbol{y}$ follows the Gaussian distribution. The expected value and the covariance matrix K of y are

$$\mathbb{E}\left[\boldsymbol{y}\right] = \mathbb{E}\left[\boldsymbol{\Phi}\boldsymbol{w}\right] = \boldsymbol{\Phi}\mathbb{E}\left[\boldsymbol{w}\right] = 0 \tag{2.80}$$

and

$$\begin{aligned} \boldsymbol{K} &= \mathbb{E}[\boldsymbol{y}\boldsymbol{y}^\top] - \cancel{\mathbb{E}[\boldsymbol{y}]\mathbb{E}[\boldsymbol{y}]^\top} \\ &= \mathbb{E}[(\boldsymbol{\Phi}\boldsymbol{w})\,(\boldsymbol{\Phi}\boldsymbol{w})^\top] \\ &= \mathbb{E}[\boldsymbol{\Phi}\boldsymbol{w}\boldsymbol{w}^\top\boldsymbol{\Phi}^\top] \\ &= \boldsymbol{\Phi}\mathbb{E}[\boldsymbol{w}\boldsymbol{w}^\top]\boldsymbol{\Phi}^\top \\ &= \lambda^2\boldsymbol{\Phi}\boldsymbol{\Phi}^\top, \end{aligned} \tag{2.81}$$

respectively. Here,

$$\begin{aligned} \lambda^2\boldsymbol{I} &= \Sigma\left(\boldsymbol{w}\right) = \mathbb{E}[\boldsymbol{w}\boldsymbol{w}^\top] - \cancel{\mathbb{E}[\boldsymbol{w}]\mathbb{E}[\boldsymbol{w}]^\top} \\ \mathbb{E}[\boldsymbol{w}\boldsymbol{w}^\top] &= \lambda^2\boldsymbol{I} \end{aligned} \tag{2.82}$$

is used. Thus, the Gaussian distribution that $\boldsymbol{y}$ follows is given as

$$\boldsymbol{y} \sim \mathcal{N}\left(0,\ \lambda^2\boldsymbol{\Phi}\boldsymbol{\Phi}^\top\right). \tag{2.83}$$

When Eq. (2.83) is satisfied with arbitrarily $\boldsymbol{x}$, the relationship between $\boldsymbol{x}$ and $\boldsymbol{y}$ follows Gaussian process. A more detailed explanation is written in the Japanese book[81].

2.2.2 Kernel function and kernel trick

Equation (2.83) indicates that the distribution of $\boldsymbol{y}$ is determined only by the covariance matrix ($\boldsymbol{K} = \lambda^2 \boldsymbol{\Phi}\boldsymbol{\Phi}^\top$). The i, j element of $\boldsymbol{K}$ is given as

$$\begin{aligned} K_{i,j} = \lambda^2 \left(\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{T}}\right)_{i,j} = \lambda^2 \sum_{k=0}^{H} \left\{\phi_k(\boldsymbol{x}^{(i)})\phi_k(\boldsymbol{x}^{(j)})\right\} &= \lambda^2 \boldsymbol{\phi}(\boldsymbol{x}^{(i)})^{\mathrm{T}} \boldsymbol{\phi}(\boldsymbol{x}^{(j)}) \\ &= k_{\mathrm{GP}}\left(\boldsymbol{x}^{(i)},\, \boldsymbol{x}^{(j)}\right), \end{aligned} \tag{2.84}$$

where k_{GP} is called the kernel function. This replacement of the covariance matrix by a kernel function is called the kernel trick and allows one to avoid representing the feature vector ϕ and the calculation of the covariance matrix. In this study, Matérn52 kernel was used.

- Matérn kernel

$$k_{\mathrm{GP}}\left(\boldsymbol{x},\, \boldsymbol{x}'\right) = \theta_0 \frac{2^{1-\nu_{\mathrm{M}}}}{\Gamma(\nu_{\mathrm{M}})} \left(\frac{l\sqrt{2\nu_{\mathrm{M}}}}{\theta_1}\right)^{\nu_{\mathrm{M}}} K_\nu \left(\frac{l\sqrt{2\nu_{\mathrm{M}}}}{\theta_1}\right) \tag{2.85}$$

$$l = |\boldsymbol{x} - \boldsymbol{x}'|, \tag{2.86}$$

  where θ is the hyper parameter, K_ν the Bessel function of the second kind and ν_{M} a parameter to determine the smoothness of the function. The Matérn kernel becomes the exponential kernel at $\nu_{\mathrm{M}} = 1/2$ and the RBF kernel at $\nu_{\mathrm{M}} = \infty$. In this study, $\nu_{\mathrm{M}} = 5/2$ was used based on a priori knowledge of the smoothness of the data.

- Matérn52 kernel ($\nu_{\mathrm{M}} = 5/2$)

$$k_{\mathrm{GP}}(\boldsymbol{x}, \boldsymbol{x}') = \theta_0 \left(1 + \frac{\sqrt{5}l}{\theta_1} + \frac{5l^2}{3\theta_1^2}\right) \exp\left(-\frac{\sqrt{5}l}{\theta_1}\right) \tag{2.87}$$

- exponential kernel

$$k_{\mathrm{GP}}\left(\boldsymbol{x},\, \boldsymbol{x}'\right) = \theta_0 \exp\left(-\frac{|\boldsymbol{x} - \boldsymbol{x}'|}{\theta_1}\right) \tag{2.88}$$

- Gaussian (RBF) kernel

$$k_{\mathrm{GP}}\left(\boldsymbol{x},\, \boldsymbol{x}'\right) = \theta_0 \exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{x}'\|^2}{2\theta_1^2}\right) \tag{2.89}$$

The hyperparameter $\boldsymbol{\theta} = (\theta_0,\ \theta_1)$ is obtained by maximum likelihood estimation (MLE). In MLE, $\boldsymbol{\theta}$ maximizing the probability of obtaining the dataset $\mathcal{D}_{\mathrm{N}} = (\boldsymbol{y}, \boldsymbol{X})$ is searched. The probability density function of obtaining $(\boldsymbol{y}, \boldsymbol{X})$ is given as

$$\begin{aligned} p\left(\boldsymbol{y}|\boldsymbol{X}\right) &= \mathcal{N}\left(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{K}\left(\boldsymbol{\theta}\right)\right) \\ &= \frac{1}{\left(\sqrt{2\pi}\right)^D \sqrt{|\boldsymbol{K}(\boldsymbol{\theta})|}} \exp\left(-\frac{1}{2}\left(\boldsymbol{y} - \boldsymbol{\mu}\right)^\top \boldsymbol{K}(\boldsymbol{\theta})^{-1}\left(\boldsymbol{y} - \boldsymbol{\mu}\right)\right). \end{aligned} \tag{2.90}$$

By taking the logarithm of both sides and $\boldsymbol{\mu} = 0$ from the preprocessing,

$$\begin{aligned} \log p\left(\boldsymbol{y}|\boldsymbol{X}\right) &= -\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{K}\left(\theta\right)^{-1}| - \frac{1}{2}\boldsymbol{y}^{\top}\boldsymbol{K}\left(\theta\right)^{-1}\boldsymbol{y} \\ &= -\frac{1}{2}\left(\log|\boldsymbol{K}\left(\theta\right)^{-1}| + \boldsymbol{y}^{\top}\boldsymbol{K}\left(\theta\right)^{-1}\boldsymbol{y}\right) + \text{const.} \end{aligned} \tag{2.91}$$

is given. Then, $\boldsymbol{\theta}^*$ is obtained by

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta}}{\text{argmax}}\left(\log p\left(\boldsymbol{y}|\boldsymbol{X}\right)\right). \tag{2.92}$$

The quasi-Newton method with BFGS[82] was used as the optimizer for the MLE in this study. The quasi-Newton method was conducted five times with random initial values due to the possibility of finding locally optimal solutions. Automatic Relevance Determination (ARD)[83] was not applied and the same length scales were used for different parameters.

2.2.3 Gaussian process regression (GPR)

GPR gives the prediction of $\left(\boldsymbol{x}^{(N+1)},\ \boldsymbol{y}^{(N+1)}\right)$. When the relationship between the $\boldsymbol{x}$ and $\boldsymbol{y}$ follows Gaussian process, $\boldsymbol{y}^* = [y^{(1)}, y^{(2)}, \cdots, y^{(N)}, y^{(N+1)}]^{\mathrm{T}}$ also follows the Gaussian distribution;

$$\boldsymbol{y}^* \sim \mathcal{N}\left(0, K^*\right), \tag{2.93}$$

where

$$\begin{aligned} K^* &= \begin{pmatrix} k_{\mathrm{gp}}\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(1)}\right) & \cdots & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(N)}\right) & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(N+1)}\right) \\ k_{\mathrm{gp}}\left(\boldsymbol{x}^{(2)}, \boldsymbol{x}^{(1)}\right) & \cdots & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(2)}, \boldsymbol{x}^{(N)}\right) & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(2)}, \boldsymbol{x}^{(N+1)}\right) \\ \vdots & \ddots & \vdots & \vdots \\ k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N)}, \boldsymbol{x}^{(1)}\right) & \cdots & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N)}, \boldsymbol{x}^{(N)}\right) & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N)}, \boldsymbol{x}^{(N+1)}\right) \\ k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N+1)}, \boldsymbol{x}^{(1)}\right) & \cdots & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N+1)}, \boldsymbol{x}^{(N)}\right) & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N+1)}, \boldsymbol{x}^{(N+1)}\right) \end{pmatrix} \\ &= \begin{pmatrix} K & k_{\mathrm{gp}}^* \\ k_{\mathrm{gp}}^{*\,\mathrm{T}} & k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N+1)}, \boldsymbol{x}^{(N+1)}\right) \end{pmatrix} \end{aligned} \tag{2.94}$$

$$k_{\mathrm{gp}}^* = [k_{\mathrm{gp}}\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(N+1)}\right), k_{\mathrm{gp}}\left(\boldsymbol{x}^{(2)}, \boldsymbol{x}^{(N+1)}\right), \cdots, k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N)}, \boldsymbol{x}^{(N+1)}\right)]^{\mathrm{T}} \tag{2.95}$$

and k_{gp} is the kernel function.
Generally, under the condition of

$$\begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}\right), \tag{2.96}$$

the conditional probability $p(\boldsymbol{y}_2|\boldsymbol{y}_1)$ is expressed as (see [84])

$$p\left(\boldsymbol{y}_2|\boldsymbol{y}_1\right) \sim \mathcal{N}\left(\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{y}_1 - \boldsymbol{\mu}_1), \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\right). \tag{2.97}$$

By comparing Eqs. (2.93, 2.97),

$$\begin{aligned} \boldsymbol{\mu}_1 &= 0 \\ \boldsymbol{\mu}_2 &= 0 \\ \boldsymbol{\Sigma}_{11} &= K \\ \boldsymbol{\Sigma}_{12} &= k_{\mathrm{gp}}^{*} \\ \boldsymbol{\Sigma}_{21} &= k_{\mathrm{gp}}^{\top} \\ \boldsymbol{\Sigma}_{22} &= k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N+1)}, \boldsymbol{x}^{(N+1)}\right) \end{aligned} \tag{2.98}$$

are obtained.
Thus, the conditional probability of obtaining y^* after obtaining $\boldsymbol{y}$ is

$$p\left(y^{(N+1)}|\boldsymbol{x}\right) \sim \mathcal{N}\left(k_{\mathrm{gp}}^{*\mathrm{T}} K^{-1}\boldsymbol{y}, k_{\mathrm{gp}}\left(\boldsymbol{x}^{(N+1)}, \boldsymbol{x}^{(N+1)}\right) - k_{\mathrm{gp}}^{*\mathrm{T}} K^{-1} k_{\mathrm{gp}}^{*}\right) \tag{2.99}$$

2.2.4 Acquisition functions

The parameter to be calculated ($\boldsymbol{x}_{\mathrm{next}}$) is determined by utilizing an acquisition function as

$$\boldsymbol{x}_{\mathrm{next}} = \underset{\boldsymbol{x}}{\mathrm{argmax}}\ A\left(\boldsymbol{x}\right). \tag{2.100}$$

Typical acquisition functions are upper confidence bound (UCB), Probability of Improvement (PI), and Expected Improvement (EI) defined as;

$$A_{\mathrm{UCB}}(\boldsymbol{x}) = \boldsymbol{\mu}_{\mathrm{GPR}}(\boldsymbol{x}) + \beta\boldsymbol{\sigma}_{\mathrm{GPR}}(\boldsymbol{x}) \tag{2.101}$$

$$A_{\mathrm{PI}}\left(\boldsymbol{x}\right) = p(y \geq y_{\max}) = \int_{y_{\max}}^{\infty} p\left(y^{(N+1)}|\boldsymbol{x}\right) dy \tag{2.102}$$

$$A_{\mathrm{EI}}(\boldsymbol{x}) = \mathbb{E}[\max\left(y - y_{\max},\ 0\right)] = \left(f^* - \mu_{\mathrm{GPR}}(\boldsymbol{x})\right)\Phi_{\mathrm{EI}}(\boldsymbol{Z}) + \sigma_{\mathrm{GPR}}\phi_{\mathrm{EI}}(\boldsymbol{Z}), \tag{2.103}$$

where f^* is the best value of the objective function, Φ_{EI} the cumulative density function, ϕ_{EI} the probability density function. and

$$\boldsymbol{Z} = \frac{f^* - u_{\mathrm{GPR}}(\boldsymbol{x})}{\sigma_{\mathrm{GPR}}(\boldsymbol{x})}. \tag{2.104}$$

In this study, Expected Improvement was used as the acquisition function.

2.3 Effectiveness of Bayesian optimization

Figure 2.4 shows the update history of M^*, σ^*, and the maximum value of the objective function $f_{\max}$. The values of M^* and σ^* were successfully improved as f increased. Figure 2.5 shows the

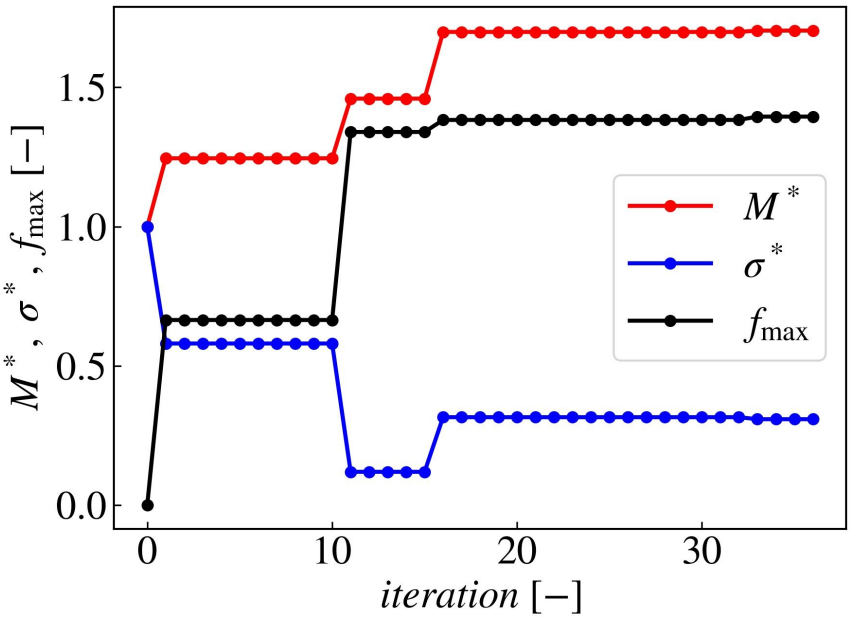

Figure 2.4. The update history of the maximum value of the mean M^*, minimum value of the standard deviation σ^* of the crystal growth rates, and maximum value of the objective function $f_{\max}$. The objective of the optimization is maximizing f $(= M^* - \sigma^*)$ which means maximizing M^*, minimizing S^*.

distribution of M^* and σ^*. The gray and colored dots represent the results of the whole parameter combination ($125 = 5 \times 5 \times 5$ cases) and 24 cases by the discrete Bayesian optimization. The star and cross symbols show the results of 5 cases by the continuous Bayesian optimization and the initial dataset, respectively. The upper left side indicates the high f due to the definition (Eq. (2.73)), which means high and uniform crystal growth rates. As shown in the figure, the upper left region was efficiently explored by Bayesian optimization. The optimal parameters of the discrete and continuous optimization were obtained with roughly 1/4 of the calculations required for a full parametric search (125 calculations). These optimal values are as follows:

$$\begin{pmatrix} B_0 \\ H_0 \\ \omega_s \end{pmatrix} = \begin{pmatrix} 0.05\ \mathrm{T} \\ 15\ \mathrm{mm} \\ 80\ \mathrm{rpm} \end{pmatrix} \tag{2.105}$$

and

$$\begin{pmatrix} B_0 \\ H_0 \\ \omega_s \end{pmatrix} = \begin{pmatrix} 0.046\ \mathrm{T} \\ 17\ \mathrm{mm} \\ 80\ \mathrm{rpm} \end{pmatrix}. \tag{2.106}$$

In the optimal states, the high and uniform crystal growth rates were achieved as shown in Fig. 2.6.

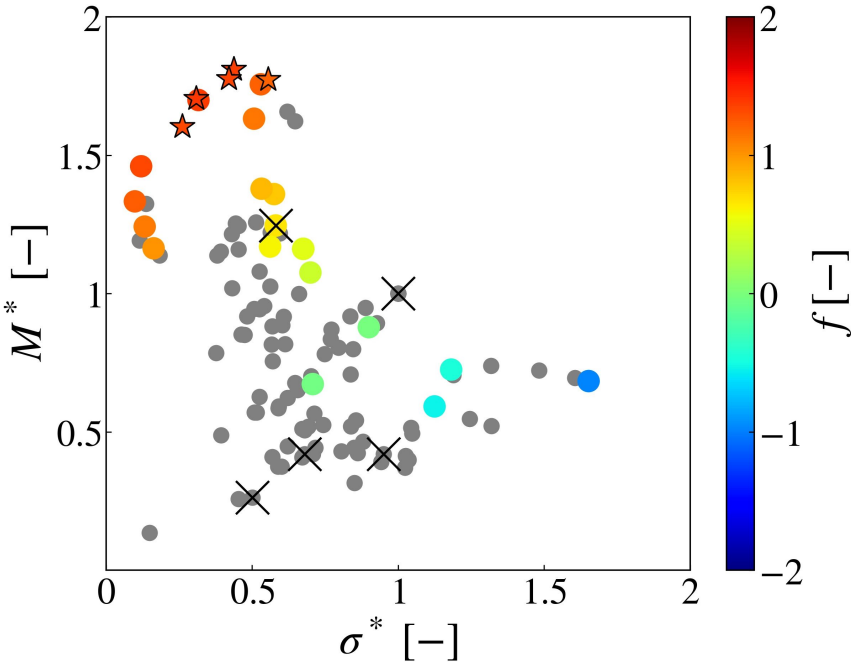

Figure 2.5. The distribution of M^* and σ^* (left). The gray dots represent the whole parameter combination (125 calculations) and the colored dots represent the 24 calculations by the discrete Bayesian optimization. The star symbols indicate five more iterations from the discrete-optimal parameters using the continuous Bayesian optimization algorithm. The circles with cross indicate the initial dataset for the Bayesian optimization.

2.4 Gaussian process regression after the optimization

The objective function f predicted through Gaussian process regression with the dataset $\mathcal{D}$ after continuous optimization is compared to the one obtained with the full parametric search, as shown in Fig. 2.7. The predicted f using the optimization results is roughly the same as those from all 125 results, especially around the optimal conditions. Figure 2.7 also shows the intensity B_0 and the position of the cusp center H_0 are sensitive factors in determining f (Fig. 2.7f). In contrast, f exhibits minimal sensitivity to changes in ω_s near its optimal states. Thus, the control of the magnetic field is more important than the rotation. Moreover, the robustness around the optimal parameters was observed in Fig. 2.7, which indicates that a sensitive control is not required.

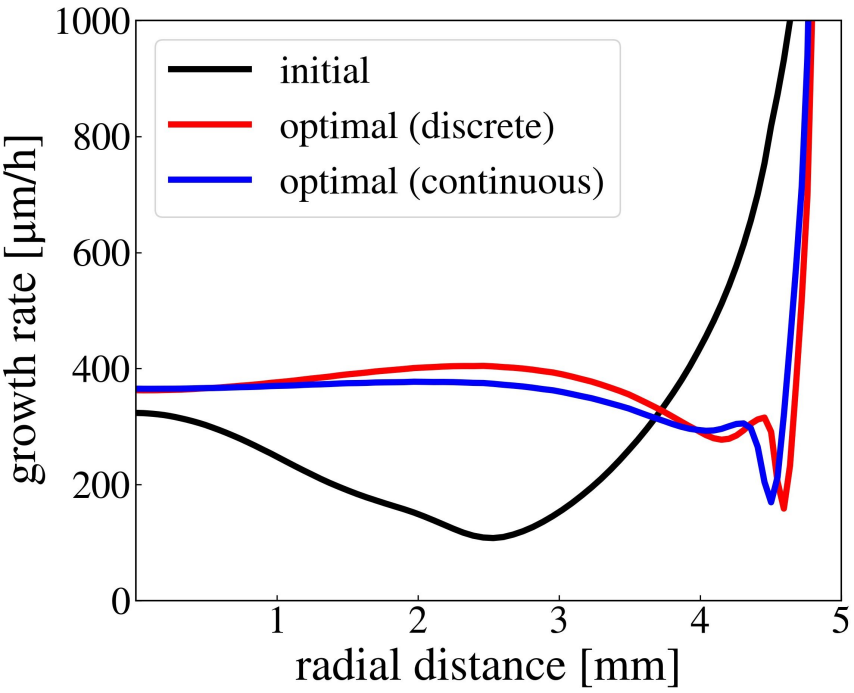

Figure 2.6. The crystal growth rates of the initial (black) and optimal states obtained from discrete (red) and continuous (blue) Bayesian optimizations.

2.5 The transport phenomena in the initial and the optimal state

2.5.1 Suppression of the Marangoni convection

The analysis of transport phenomena in the optimal state obtained from the continuous Bayesian optimization is presented in this section.

The time-averaged melt flow velocity, temperature, and concentration fields for 100 s after the initial 50 s in the initial and the optimal states are shown in Figs. 2.8 and 2.9. The downward flow derived from Marangoni convection was observed under the seed crystal in the initial state. On the other hand, in the optimal state, the downward flow transformed into the upward flow due to the suppression of Marangoni convection by the cusp magnetic field. The suppression of Marangoni convection is also confirmed by the radial velocity distribution along the free surface, as shown in Fig. 2.10. Negative radial velocity around the seed crystal ($r \sim 5$ mm) corresponds to the developed Marangoni convection. In the optimal state, the region with positive radial velocity is wider than in the initial state, indicating the effective suppression of Marangoni convection. Due to the upward flow under the seed crystal, the carbon transport to the seed crystal becomes more uniform in the optimal state than the initial state. Thus, the high and uniform crystal growth rate was achieved in the optimal state.

2.5.2 Contributions of driving forces

The contributions of each driving force term, i.e., the Lorentz forces by the induction heating, the gravity body force, and the surface tension gradient under the applied cusp magnetic field were analyzed. Figure 2.11 shows the flow distributions under the optimal condition, taking into account one of the Lorentz forces, surface tension, or buoyancy force. By comparing the magnitude of the velocities, it is found that the Buoyant convection is negligible, while the convection induced by the Lorentz force is dominant in the bulk region of the melt. Additionally, the figure indicates the upward flow under the crystal in the optimal state is mainly derived from the Lorentz force. In the melt flow considering only surface tension, Marangoni convection flows towards the outside of the melt after separating from the crystal surface. Consequently, the effect of Marangoni convection decreases under the crystal, enabling the upward flow generated by the Lorentz force in the optimal state to develop. The vortex observed near the seed edge in the optimal state is not generated in Fig. 2.11. Thus, the vortex developed by the combined effects of convection due to the Lorentz force and the surface tension.

2.5.3 Transport phenomena near the seed crystal

The oscillations of the crystal growth rate near the crystal edge under optimal conditions can be explained by the flow separation due to the vortex. Figure 2.12 shows the velocity gradient $\partial u_r/\partial z$, the concentration gradient $\partial C/\partial z$, and the concentration field near the seed crystal edge. As shown in the figure, the Marangoni convection separates from the crystal near the seed edge, where $\partial u_r/\partial z = 0$. Due to the separation, the minimum points of crystal growth rates in the optimal states were observed. Actually, the separation point corresponds to the minimum point of the crystal growth rate in the optimal state. Figure 2.12(left) also indicates that the radial velocity is increased near the separation point as shown by the red arrow. This increased velocity leads to a local maximum point in the crystal growth rates, adjacent to the minimum point.

2.6 Current achievements, limitations, and future improvements

As mentioned above, the number of trials required for optimization was significantly reduced by Bayesian optimization in the TSSG method of SiC crystals. This section clarifies what has been accomplished by this research and summarizes the current limitations of the proposed method and what is needed to make the method more applicable and efficient.

2.6.1 Efficiency

The efficiency of this method is discussed in this section. In the Bayesian optimization algorithm, prediction by Gaussian process regression and calculation of acquisition function are performed so fast that they are negligible compared to the time required for numerical simulations. Thus, the efficiency is evaluated by the needed time for the numerical simulation. The full parametric search takes 500 days (= 125 cases $\times$ 4 days/case), where 4 days correspond to the time required for a single calculation of the numerical simulation. On the other hand, Bayesian optimization takes only 96 days (= 24 cases $\times$ 4 days/case). Thus, the required time for the optimization can be reduced significantly. Moreover, unlike optimization methods based on gradients, such as the gradient descent method, the optimal state does not depend on initial values. The coding is easier since the gradients are not used.

For more efficient optimization, the number of iterations should be decreased. However, no further reduction in iterations can be expected since exploration should be performed on unobserved regions in parameter space. The only way to achieve faster optimization is to improve the speed of the numerical computation which is the rate-determining step in the optimization. Acceleration of numerical computations was performed by the Physics Informed Neural Network, as described in Chapter 4.

2.6.2 Application

As mentioned above, the number of trials required for optimization was significantly reduced by Bayesian optimization in the TSSG method of SiC crystals. Theoretically, this algorithm can be utilized in other systems. Indeed, our research group has also confirmed its effectiveness for different systems[85, 86, 87, 88].

The applicability depends on the number of parameters to be optimized. Generally, there are many parameters to be optimized in complex real systems. However, it is well known that optimizing a large number of parameters is difficult by Bayesian optimization. Hence, in order to apply the Bayesian optimization to a more complex system, it is necessary to reduce several parameters to a representative dimensionless number or to perform dimensional reductions before the optimization. Principal component analysis (PCA) and removing one of the strongly correlated parameters are the famous methods for dimensional reduction.

Table 2.2 The physical properties used in the TSSG system ([21, 78, 79]).

Materials	property (symbol)	value	unit
Si melt	Density (ρ)	2550	kgm^{-3}
	viscosity (μ)	9.0×10^{-4}	Pa s
	Electrical conductivity (σ_{e})	1.2×10^{6}	Sm^{-1}
	Thermal conductivity (k)	65	$\mathrm{Wm}^{-1}\mathrm{K}^{-1}$
	Specific heat (C_p)	1.0×10^{3}	$\mathrm{Jkg}^{-1}\mathrm{K}^{-1}$
	thermal surface tension coefficient (σ_T)	-2.5×10^{-4}	K^{-1}
	Thermal expansion coefficient (β)	1.4×10^{-4}	K^{-1}
	Emissivity (ϵ)	0.3	-
Graphite	Electrical conductivity (σ_{e})	7.54×10^{4}	Sm^{-1}
	Thermal conductivity (k)	$108 - 0.0864T$ $+2.304 \times 10^{-5}T^2$	$\mathrm{Wm}^{-1}\mathrm{K}^{-1}$
	Emissivity (ϵ)	0.66	-
Insulator	Electrical conductivity (σ_{e})	0	Sm^{-1}
	Thermal conductivity (k)	0.3	$\mathrm{Wm}^{-1}\mathrm{K}^{-1}$
	Emissivity (ϵ)	0.5	-
Helium	Electrical conductivity (σ_{e})	0	Sm^{-1}
	Thermal conductivity (k)	$6.35 \times 10^{-2} + 3.1 \times 10^{-4}T$ $-2.44 \times 10^{-8}T^2$	$\mathrm{Wm}^{-1}\mathrm{K}^{-1}$
Quartz	Thermal conductivity (k)	$1 + 1.35 \times 10^{-3}T$	$\mathrm{Wm}^{-1}\mathrm{K}^{-1}$
	Emissivity (ϵ)	0.4	-
Chamber	Emissivity (ϵ)	0.1	-
Coil	Emissivity (ϵ)	0.05	-

Table 2.3 The parameter ranges of the Bayesian optimization

B_0 [T]	H_0 [10^{-3} m]	ω_s [s^{-1}]
0	0	0
0.05	5	0.33
0.1	10	0.67
0.15	15	1.00
0.2	20	1.33

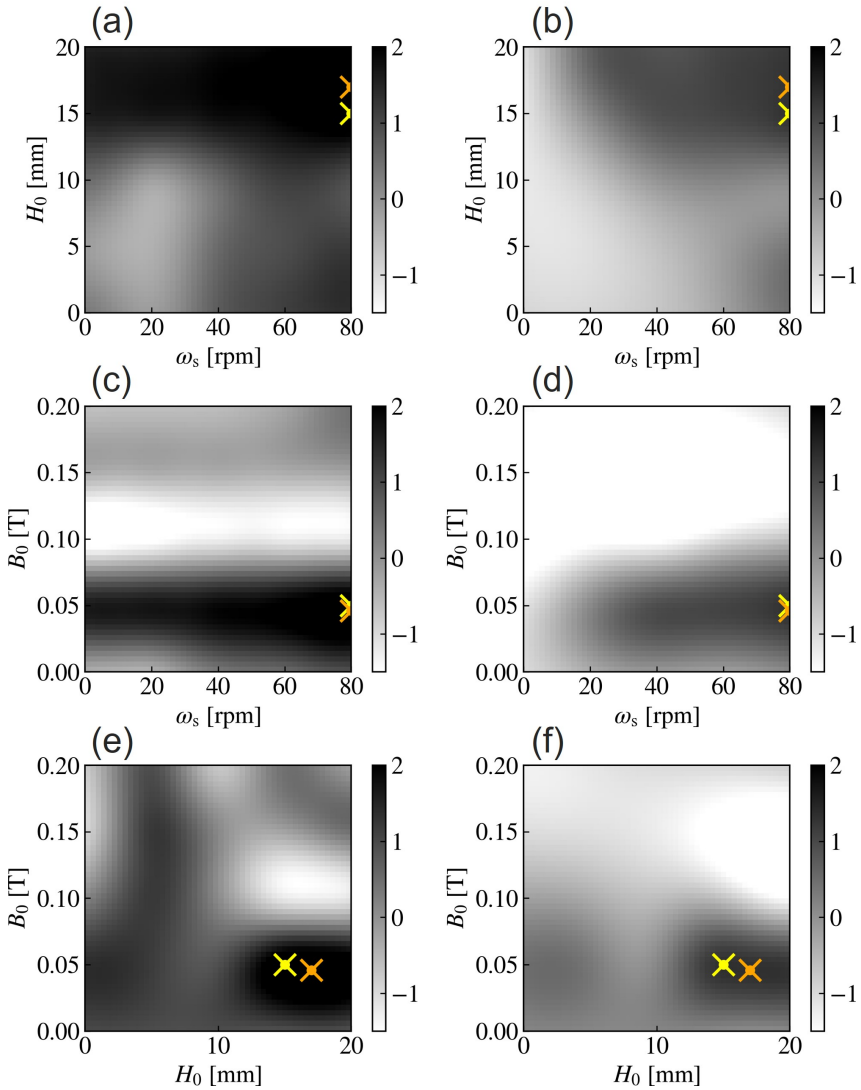

Figure 2.7. The objective function f predicted by Gaussian process regression using full parametric search (a, c, e) and one after optimizations (b, d, f). The yellow and orange cross represent the optimal parameters obtained by the discrete and continuous optimizations, respectively.

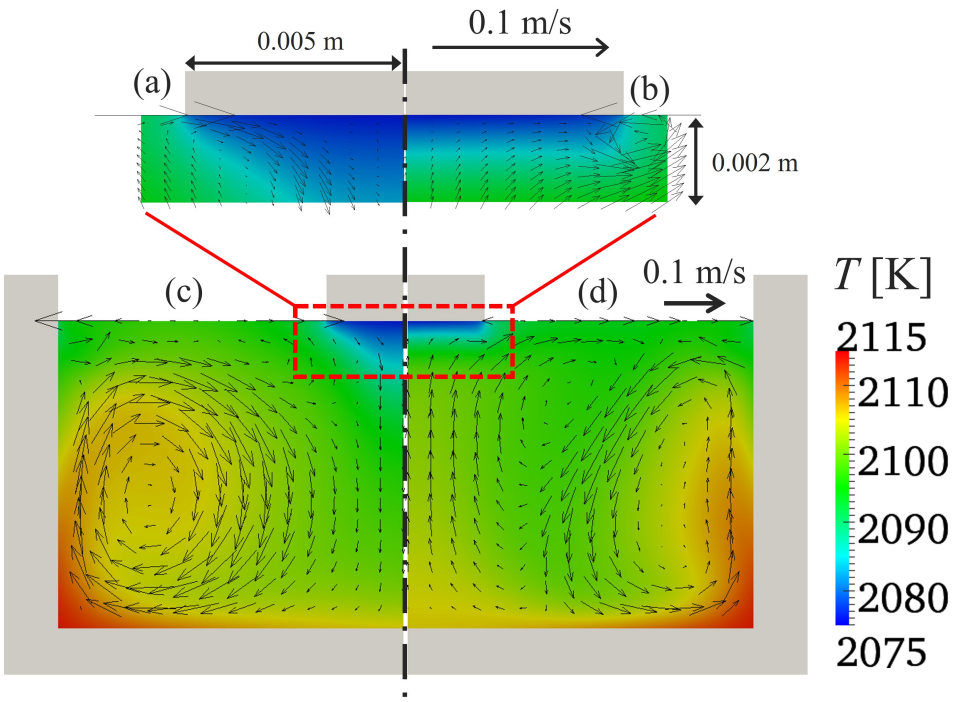

Figure 2.8. The temperature (color) and flow velocity (arrow) distributions in the initial (a, c) and the optimal states (b, d) and the zoomed region near the seed crystal (a, b).

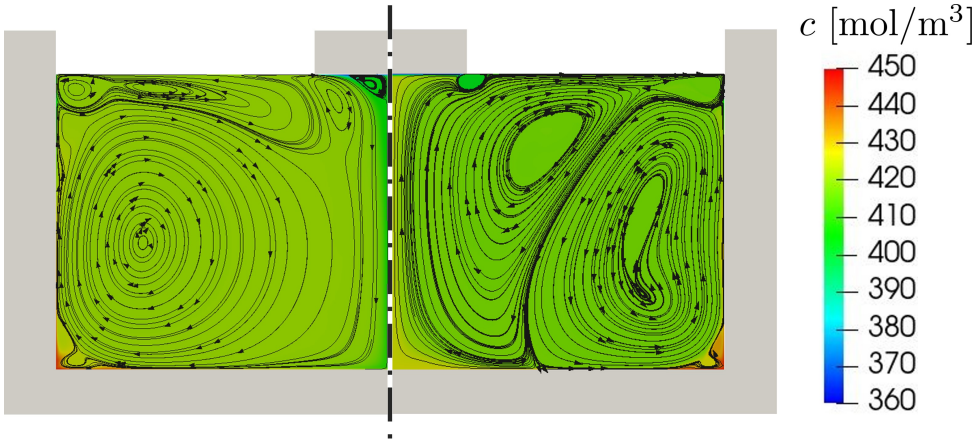

Figure 2.9. The concentration field (color) and streamline (lines with arrows) in the initial (left) and the optimal (right) states.

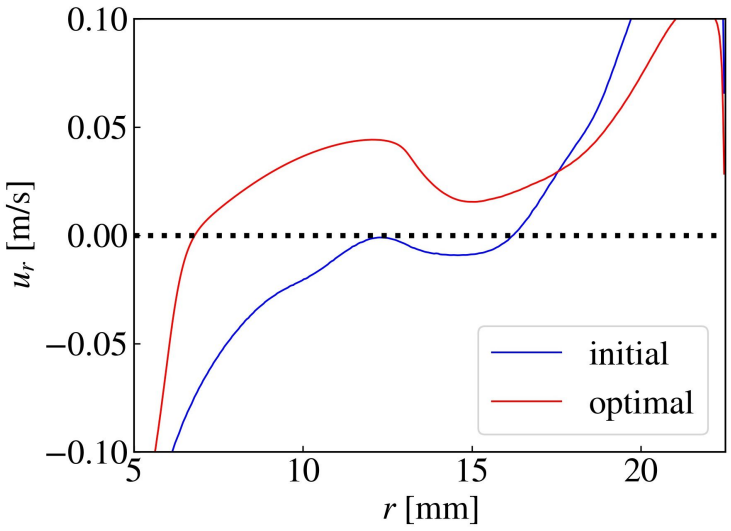

Figure 2.10. The radial velocity along the free surface in the initial (blue) and the optimal (red) states.

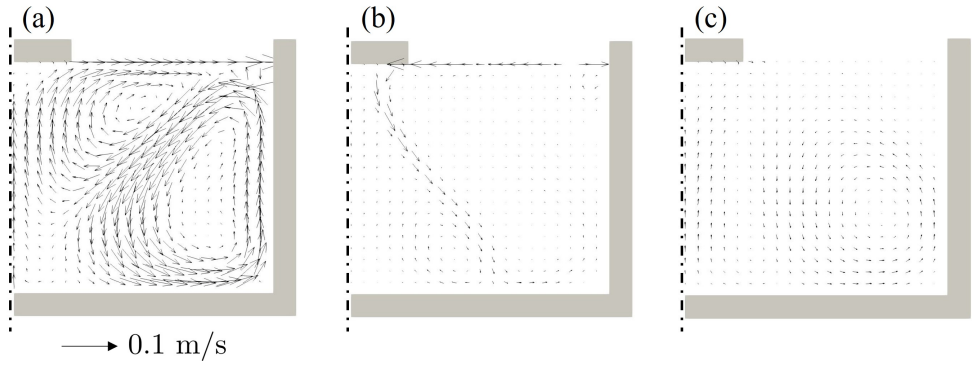

Figure 2.11. The velocity vectors representing the contributions of each driving forcing term under the optimal cusp magnetic field. (a) Lorentz force, (b) Surface tension, and (c) Buoyant force.

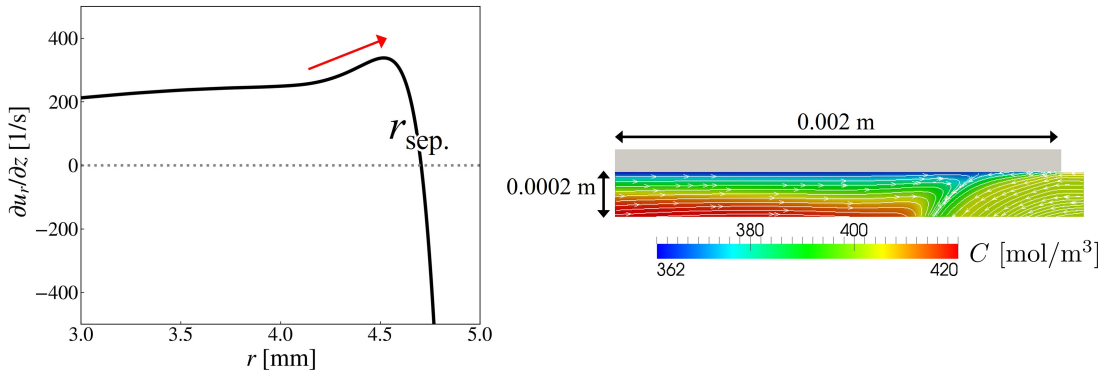

Figure 2.12. The velocity gradient along the seed surface near the seed edge (left) and the concentration field and the streamlines near the seed edge(right). $r_{\mathrm{sep.}}$ is the separation point of the melt flow ($\partial u_r/\partial z|_{r_{\mathrm{sep.}}} = 0$).

Chapter 3

Explainable machine learning for the analysis of the transport phenomena in the melt of TSSG method for SiC crystals

In this Chapter, the auto-detection technique of the transport phenomena affecting the crystal growth rates is introduced. The Bayesian optimization with numerical simulation introduced in the previous Chapter enables the fast and white-box suggestion of control methods with its explanation by numerical simulation. However, actually, there is still a challenge in the analysis of the transport phenomena. As shown in Fig. 2.8, the melt flow changes drastically by the optimized control. Therefore, it is difficult to identify the transport phenomena improving the crystal growth rates although the melt flow under the crystal was identified as the appropriate target based on previous experience and research. Due to the strong nonlinearities, melt flow in other regions may affect crystal growth rates. This type of analysis based on the observation and discovery of researchers leads to missing hidden phenomena. Thus, we introduced the auto-detection technique of the transport phenomena. This leads to whiter-box suggestions of control methods by easy analysis.

3.1 Concept of this study

The concept of this study is shown in Fig. 3.1. The relationship $f(\boldsymbol{U})$ between the melt flows $\boldsymbol{U}$ and the objective function (Eq. 2.73) is trained by a Convolutional neural network (CNN). The relationship trained by CNN is observed by the sensitivity analysis $\partial f(\boldsymbol{U})/\partial \boldsymbol{U}$. By the magnitude of the sensitivity, we find the important melt flow for the crystal growth rates.

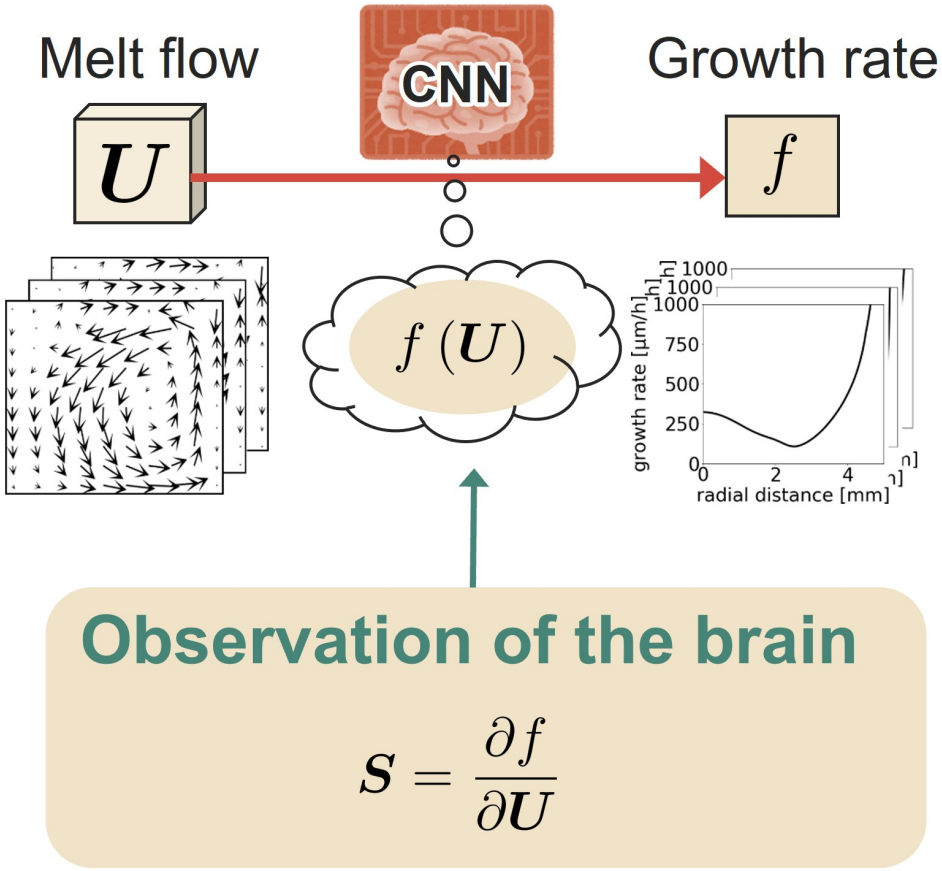

Figure 3.1. The concept of this study. Convolutional neural network learns the relationship $f(\boldsymbol{U})$ between the melt flow and the growth rate. Then, the relationship is observed by the sensitivity $\boldsymbol{S}$.

3.2 Numerical method

3.2.1 Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a type of neural network for handling two-dimensional and three-dimensional data, such as gray-scale images (height pixels × width pixels) or color images (height pixels × width pixels × RGB). Thus, CNNs are particularly utilized in image recognition[89]. In this study, the melt flow data $\boldsymbol{U}$ $[N_r \times N_z \times 3]$, where N_r and N_z are the numbers of the mesh of radial and axial directions, was treated as 3-dimensional color images. From the flow data $\boldsymbol{U}$ as input parameter, the objective function f (Eq. (2.73)) is predicted by CNN, where $\boldsymbol{U}$ is the time-averaged flow velocity for 100 s after the initial 50 s. The training was conducted by decreasing the following loss function:

$$L = \frac{1}{N} \sum_{i=1}^{N} \left(f^i_{\mathrm{CNN}} - f^i_{\mathrm{N.S.}} \right)^2, \tag{3.1}$$

where N represents the number of training data, i the data number, and f_{CNN}, $f_{\mathrm{N.S.}}$ are the predicted objective functions by CNN and numerical simulation, respectively. After the training, the value of the objective function predicted by CNN agrees with it obtained from numerical simulation. This indicates that the CNN establishes a mathematical expression that indicates the relationship between the melt flow and the objective function.

Figure 3.2 and Table 3.1 show the architecture of CNN utilized in this study. The input parameter is the time-averaged melt flow velocity $\boldsymbol{U}$. The full-size data ($480 \times 450 \times 3$) of the flow velocity is compressed to the size ($300 \times 300 \times 3$) by PIL library[90] in order to decrease the training cost. In the hidden layers, convolution, max pooling, and fully connected layers were used. In the convolution 1 and 2 layers, zero padding was applied to the output. ReLU function[91] was used as the activation function in each layer.

The training dataset consisted of 170 data points, with 125 data points covering all combinations of parameters listed in Table 2.3, and an additional 45 data points calculated randomly. The value of the learning rate was fixed at 0.001. The momentum optimizer[92] was used. The prediction accuracy was verified by Root Mean Squared Error (RMSE), Mean Squared Error (MAE), and coefficient of determination R^2 defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(f^i_{\mathrm{N.S.}} - f^i_{\mathrm{CFD}}\right)^2} \tag{3.2}$$

$$\mathrm{MAE} = \frac{1}{M}\sum_{i=1}^{M}|f^i_{\mathrm{CNN}} - f^i_{\mathrm{N.S.}}| \tag{3.3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{M}\left(f^i_{\mathrm{N.S.}} - f^i_{\mathrm{CNN}}\right)^2}{\sum_{i=1}^{M}\left(f^i_{\mathrm{N.S.}} - \overline{f^i_{\mathrm{N.S.}}}\right)^2}, \tag{3.4}$$

$$\tag{3.5}$$

respectively, where M is the number of data and $\overline{(\cdot)}$ indicates the averaged value.

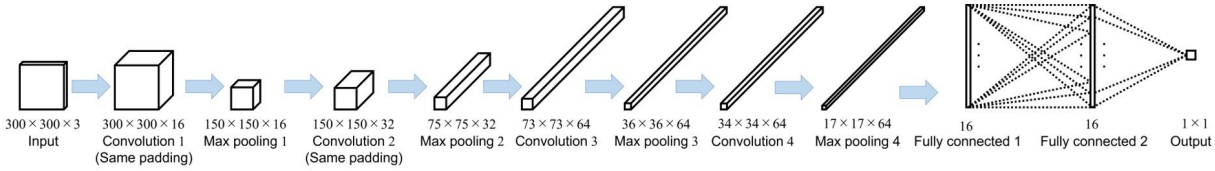

Figure 3.2. The architecture of CNN utilized in this study.

Table 3.1 The filter size and stride of CNN

Layer	(Filter size) × the number of filters	Stride	Output size
input			(300, 300, 3)
Convolution 1	$(3, 3, 3) \times 16$	(1, 1)	(300, 300, 16)
Max pooling 1	$(2, 2, 16) \times 16$	(2, 2)	(150, 150, 16)
Convolution 2	$(3, 3, 16) \times 32$	(1, 1)	(150, 150, 32)
Max pooling 2	$(2, 2, 32) \times 32$	(2, 2)	(75, 75, 32)
Convoltuion 3	$(3, 3, 64) \times 64$	(1, 1)	(73, 73, 64)
Max pooling 3	$(2, 2, 64) \times 64$	(2, 2)	(36, 36, 64)
Convolution 4	$(3, 3, 64) \times 64$	(1, 1)	(34, 34, 64)
Max polling 4	$(2, 2, 64) \times 64$	(2, 2)	(17, 17, 64)
Fully connected 1			16
Fully connected 2			1

3.2.2 Sensitivity tensor

After developing the CNN, the relationship between the objective function and flow velocity is formulated in CNN. The formulation is analyzed by calculating sensitivity $\boldsymbol{S}$ defined as

$$\boldsymbol{S}\left(\boldsymbol{U}\right) \quad = \quad \frac{\partial f_{\mathrm{CNN}}}{\partial \boldsymbol{U}}. \tag{3.6}$$

The sensitivity $\boldsymbol{S}$ is the 3rd order tensor, the same as the velocity data $\boldsymbol{U}$ and the i, j, k element of $\boldsymbol{S}$ is given as

$$S^{i,j,k} \quad = \quad \left(\frac{\partial f_{\mathrm{CNN}}}{\partial u_k^{i,j}}\right), \tag{3.7}$$

where $k = r, \theta, z$. For example, The high sensitivity $S_{i,j,k}$ shows that the effects of the melt flow $u_{i,j,k}$ at $(x, y) = (i, j)$ directed k on the objective function is high. Thus, the formulation $f(\boldsymbol{U})$ in CNN can be observed by the sensitivity.
The gradient was calculated by automatic differentiation of the neural network.

3.2.3 SmoothGrad

The raw sensitivity expressed as (Eq. 3.6)is usually noisy as mentioned in Smilkov *et al.*[47]. The sensitivity tensor $\boldsymbol{S}$ is smoothed[47] as

$$\widehat{\boldsymbol{S}} = \frac{1}{n}\sum_{1}^{n} \boldsymbol{S}(\boldsymbol{U} + \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}\right)), \tag{3.8}$$

where n is the number of the Gaussian noise and Σ is the covariance matrix. In this study, $n = 100$ and $\Sigma = \lambda \cdot \mathbf{I}$ were used, where I is the identity matrix and the variance.

3.3 The training and accuracy of CNN

Figure 3.3 shows the learning curves of the CNN when all available data is used for training. As shown in the figures, the training loss decreased and converged successfully.
Learning accuracy was calculated 10 times using 10 randomly selected cases as test data from the 45 cases obtained by randomly changing the parameters as shown in Fig. 3.4. The relationship between the objective functions obtained from the numerical simulation and the CNN is shown in Fig. 3.5. The gray dots and red crosses indicate the objective function f of the training and test data, respectively. When the accurate prediction is achieved, the test data is on the dotted line in the figure. For each randomly selected test data, the test data were on the dotted line, indicating that the accurate prediction model was developed. Moreover, RMSE, MAE, and R^2 of each test dataset were shown in Fig. 3.6. The average errors were also calculated incrementally while conducting 10 tests as shown in the solid lines in Fig. 3.6. The averaged errors against each test data converged, and the RMSE, MAE, and R^2 were 0.116, 0.0849, and 0.963, respectively. These values show that the accurate model was developed.

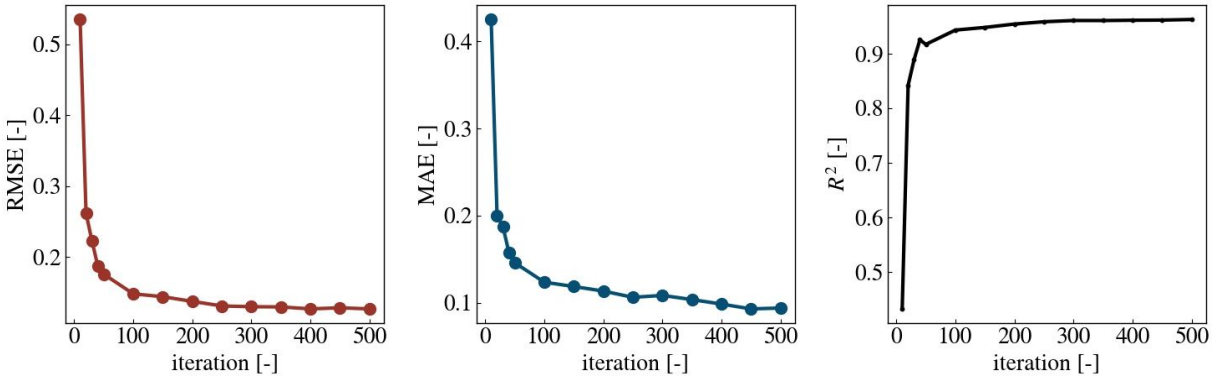

Figure 3.3. The learning curve of CNN.

3.4 Sensitivity analysis

3.4.1 Sensitivity below the seed center

The magnitudes and the components of the sensitivity of the melt flow in the initial and the optimal states are shown in Figs. 3.7 and 3.8, respectively. As shown in the figures, the regions below the seed crystal (A) and near the seed edge (C) have a high sensitivity value, indicating

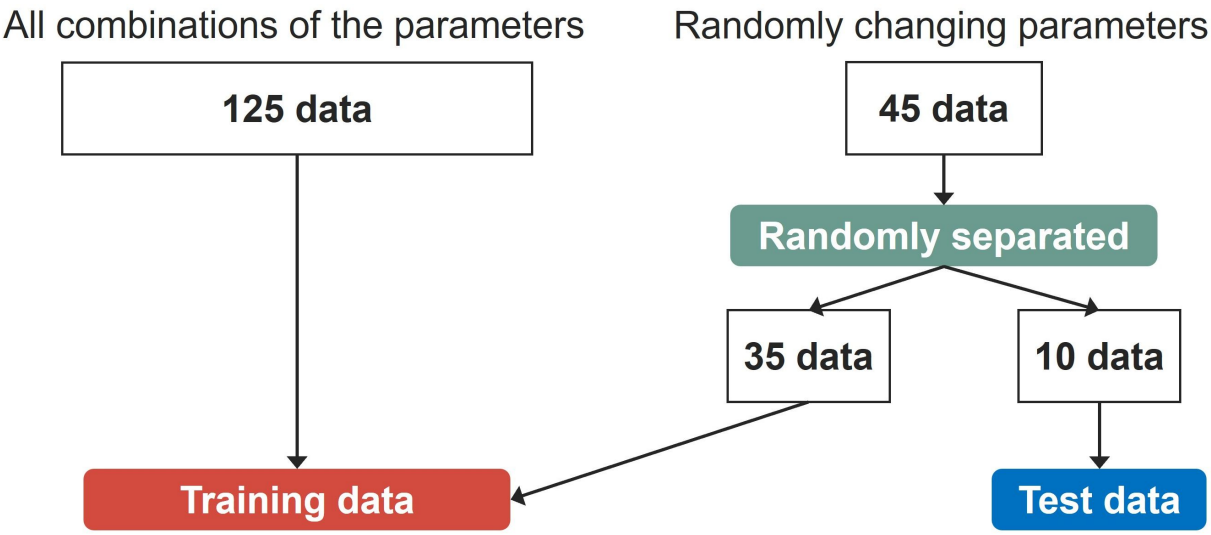

Figure 3.4. The structures of the training and test data. Random selections were repeated 10 times, and prediction accuracy was calculated for each training and test dataset.

that the melt flow in these regions has a significant effect on the objective function. In region A, S_z is positive as shown in Fig. 3.8. The positive $\hat{S}_z$ $(= \partial\hat{S}/\partial z)$ means that f increases as u_z increases. Thus, increasing u_z below the seed crystal is desired for improving the objective function in the initial state. On the other hand, in the optimal state, the sensitivity in region A does not show a high value. The decreasing sensitivity, as f is improved, suggests that the flow in that region is improved enough to have no effect on the f. Since the improved flow in region A of the optimal state was the upward flow, it is proved that the upward flow improved f. Previous studies have often mentioned that the upward flow under the seed crystal improves the crystal growth rate due to the uniform carbon transport [23, 66], and our proposed method successfully validates this observation.

3.4.2 Sensitivity below the seed edge

The sensitivity below the seed edge (region A) shows high negative values in the initial and optimal states as shown in Figs. 3.7 and 3.8. The negative S_z indicates that the downward flow below the seed edge improved f. The downward flow in the region is derived from the flow separation of the Marangoni convection. Thus, it is found that the flow separation of the Marangoni convection improved f. The flow separation was also observed in the optimal state obtained by optimizing the temperature boundary condition along the crucible by adjoint method[65]. This is because the flow separation prevents Marangoni convection from reaching the center of the crystal.

$\hat{S}_\theta$ also shows positive sensitivity at the seed crystal edge in the optimal state. The positive sensitivity $\hat{S}_\theta$ means that an increase in u_θ enhances f. Since the u_θ is driven by the seed

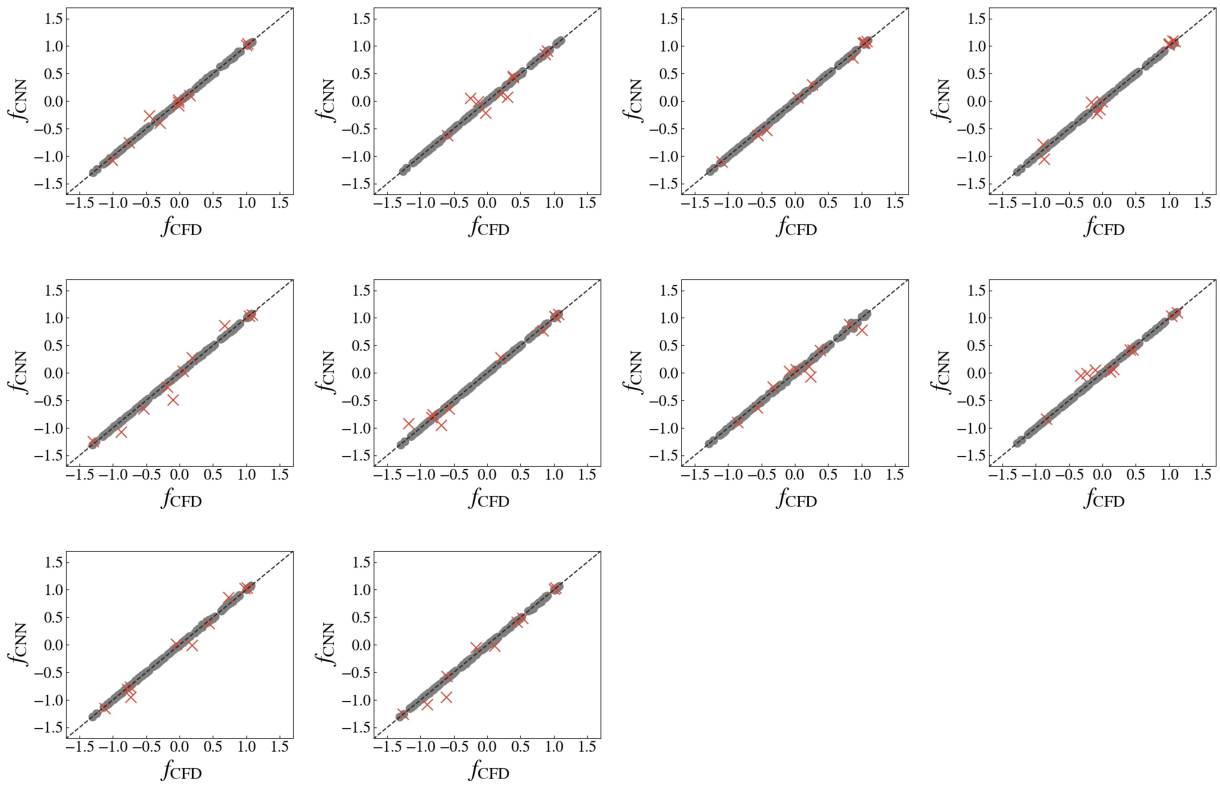

Figure 3.5. The relationship between the objective functions obtained from the numerical simulation and the CNN. The gray dots and red crosses indicate the training and test data, respectively.

rotation, faster seed rotation will improve the objective function.

The sensitivity in the optimal state suggests the presence of more favorable states than the optimal state, which cannot be achieved by the control method employed in this study.

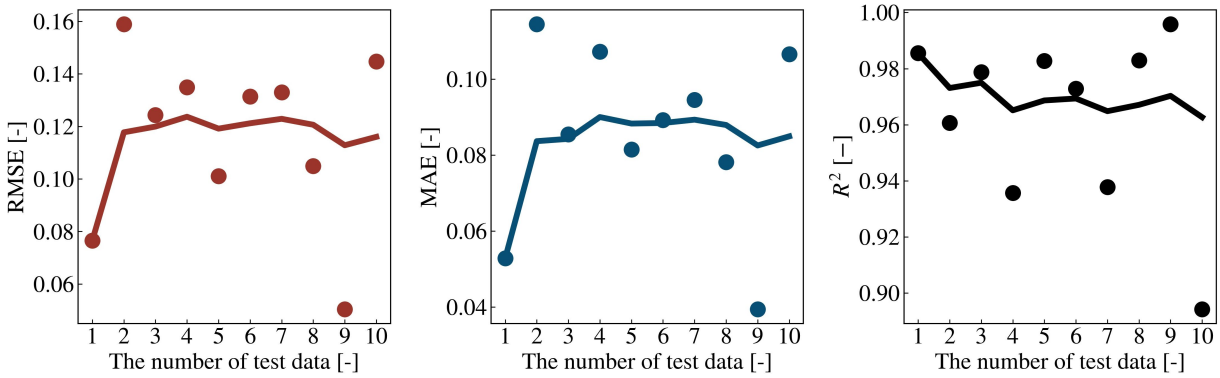

Figure 3.6. The RMSE, MAE, and R^2 values against each test data (points) and the mean values of the errors (line).

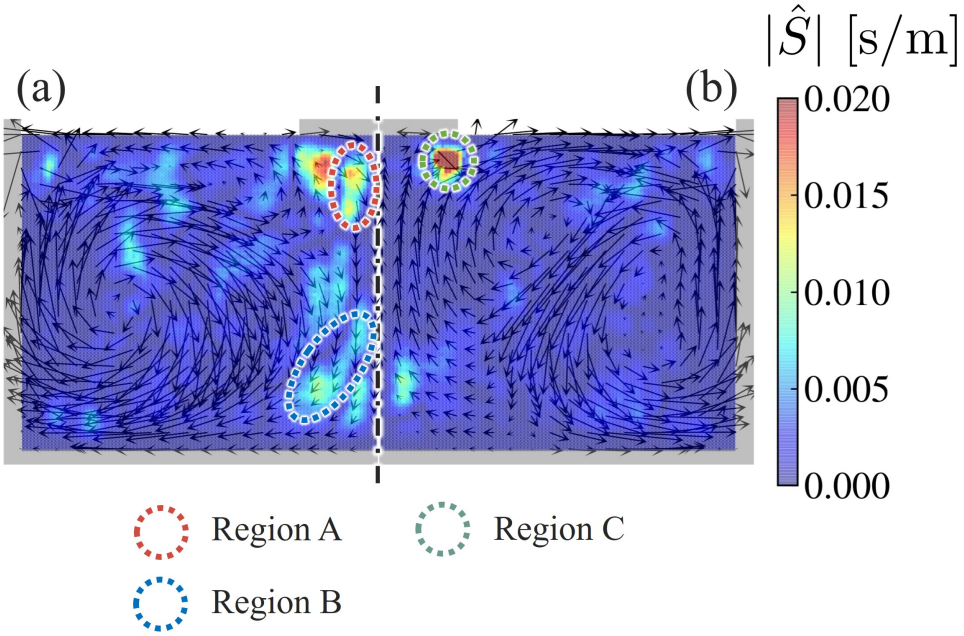

Figure 3.7. The magnitudes of the sensitivity of the melt flow in the initial (a) and the optimal (b) states.

3.4.3 Sensitivity of the top 10% of the objective function

Figures 3.9 and 3.10 show the distributions of the magnitude and the components of sensitivity, respectively, averaged within the top 10 % of the objective function. The high sensitivity value was observed at the seed edge similar to the optimal state. Moreover, the high sensitivity was also observed near the bottom center of the melt. The positive sensitivity $\hat{S}_z$ near the bottom indicates that an increase in u_z improves f. The high positive sensitivity can be understood by the streamlines through the region below the seed crystal, as shown in Fig. 3.11. All the carbon transported to the crystal by the melt flow follows the same trajectory. From the observation of the concentration field in the optimal state, carbon dissolved from the bottom center is transported to the crystal by the melt flow. Therefore, the sensitivity of the region is high.

3.5 Current achievements, limitations, and future improvements

The important phenomena in the melt of the TSSG-SiC system were detected automatically, which led to the assistance of the analysis for the white-box control. The proposed method is not an established and well-known one, but our original method. Although Parameter analyses using other XAI, such as SHAP, and LIME, methods have been reported[93], there are few studies of their application to physical fields as ours. This method can be applicable to other systems since a specific technique only for TSSG-SiC was not used. Indeed, XAI for fluid analysis has been

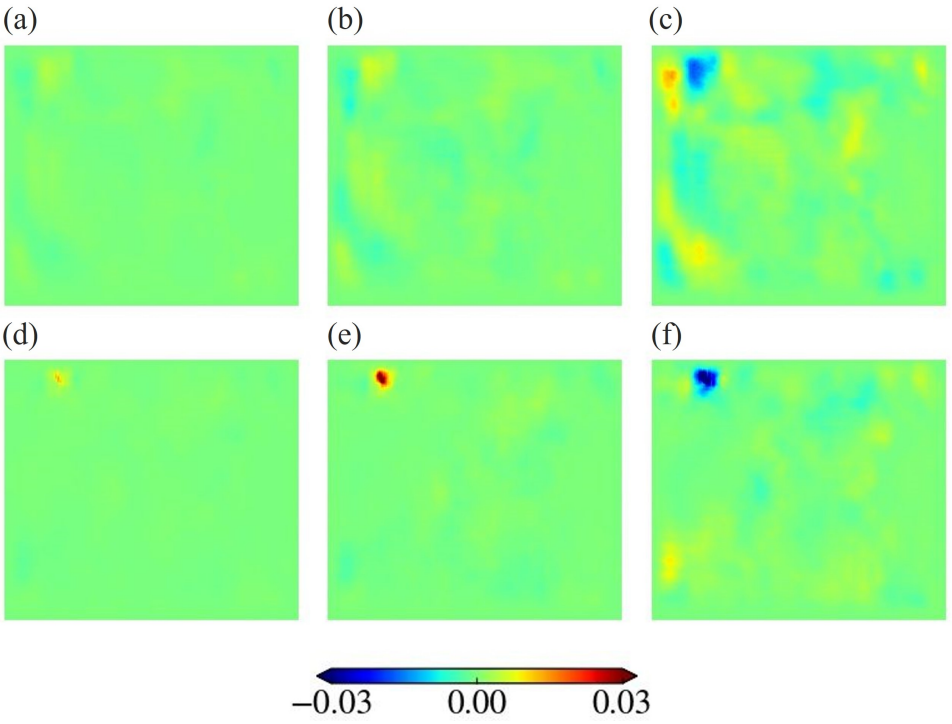

Figure 3.8. The components of the sensitivity, $\hat{S}_r$ (a, d), $\hat{S}_\theta$ (b, e), $\hat{S}_z$ (c, f) in the initial (a, b, c) and optimal (d, e, f) states.

reported, although it is an unreviewed manuscript[94].

The limitation of the proposed method is derived from the difficulty of generating a lot of training data and uncertainty in numerical simulation utilized for the development of the training data. In order to obtain the correct sensitivity, the training of CNN should be conducted accurately. For accurate training, a lot of accurate training data is required. On the other hand, it is difficult to develop such a training dataset of practical processes since numerical simulation takes a long time and involves uncertainties, e.g. physical properties. Thus, the difficulty of generation of certain datasets leads to the limitation of the application.

The problem about the long time of the numerical simulation was addressed in Chapter 4. However, the uncertainty of numerical simulation which has been discussed for a long time will remain and continue to be discussed in the field of computational science.

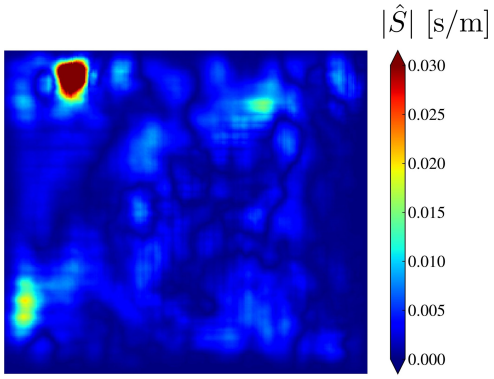

Figure 3.9. The averaged magnitudes of the sensitivity of the melt flow of the top 10 % of the objective function.

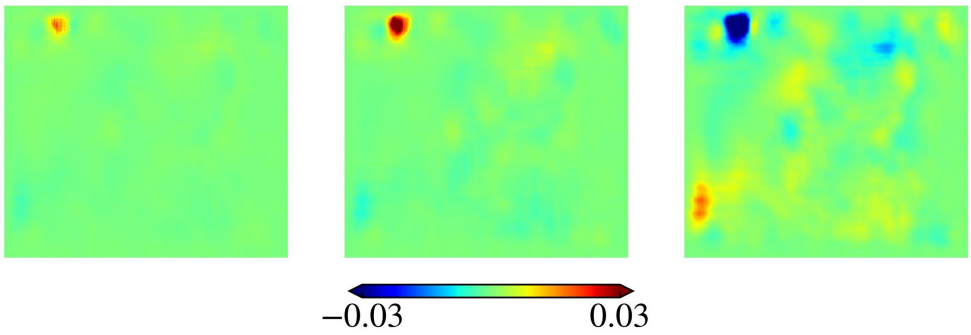

Figure 3.10. The components of the averaged sensitivity of the melt flow of the top 10 % of the objective function. $\hat{S}_r$ (left), $\hat{S}_\theta$ (center), $\hat{S}_z$ (right).

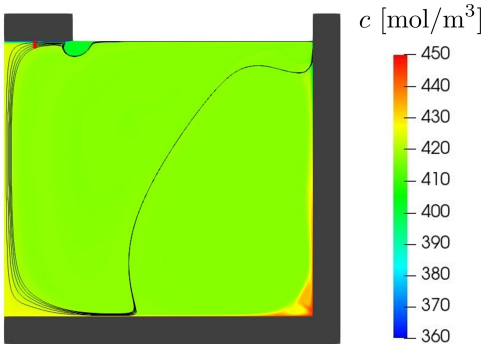

Figure 3.11. The streamlines and the concentration distribution of the optimal state. The streamlines through the red line below the seed crystal are shown. The streamlines and the concentration field in this figure were obtained by using the time-averaged melt flow and the concentration fields.

Chapter 4

Physics Informed Neural Networks for Czochralski method of Si crystals

In the previous studies, Bayesian optimization and XAI were introduced. Due to these methods, the latter two could be addressed. However, numerical simulation for the analysis of the transport phenomena still takes a long time, which leads to black-box control during experiments. Thus, in this Chapter, the fast prediction method is introduced for the fast analysis during experiments.

4.1 Numerical simulation

In this research, the Si melt region of the Cz method is considered. Figure 4.1 shows the schematic view of Si melt of the Cz method and the numerical domain. The 2-dimensional steady axisymmetric system was assumed since the flow velocity is dumped by the static vertical magnetic field. The bottom shape of the melt is expressed by a simple quadratic function and the configuration ranges are as follows;

$$0 \leq \quad r\ [\mathrm{m}] \quad \leq r_\mathrm{c} = 0.3, \tag{4.1}$$

$$\frac{h - h_\mathrm{s}}{r_\mathrm{c}^2} r^2 + (0.3 - h) = z_\mathrm{b} \leq \quad z\ [\mathrm{m}] \quad \leq 0.3. \tag{4.2}$$

The governing equations of the steady-state analysis of the Cz-Si melt are written as follows;

$$\begin{aligned} \boldsymbol{f}_U &= (f_u,\ f_w) \\ &= (\boldsymbol{u} \cdot \nabla)\, \boldsymbol{u} + \frac{1}{\rho} \nabla p - \nu \nabla^2 \boldsymbol{u} \\ &\quad + \boldsymbol{g} \beta\, (T - T_\mathrm{m}) - \frac{\boldsymbol{J} \times \boldsymbol{B}}{\rho} = 0 \end{aligned} \tag{4.3}$$

$$f_p = \nabla \cdot \boldsymbol{u} = 0 \tag{4.4}$$

$$f_T = (\boldsymbol{u} \cdot \nabla)\, T - \alpha \nabla^2 T = 0 \tag{4.5}$$

$$f_\phi = \nabla^2 \phi - \nabla \cdot (\boldsymbol{u} \times \boldsymbol{B}) = 0, \tag{4.6}$$

where $\boldsymbol{u}$ is the melt flow velocity, ρ the melt density, p the pressure, ν the kinetic viscosity, $\boldsymbol{g}$ the gravitational acceleration, β the thermal expansion coefficient, T the temperature, T_m the melting point of the liquid Si, $\boldsymbol{J}$ the current density, $\boldsymbol{B}$ (= (0, 0.3) T) the vertical magnetic field, α the thermal diffusivity, and ϕ the scalar potential of the electric field $\boldsymbol{E}_\mathrm{m}$ defined as $\boldsymbol{E}_\mathrm{m} = -\nabla\phi - \cancel{\partial \boldsymbol{A}_\mathrm{m}/\partial t} = -\nabla\phi$. The Boussinesq approximation was used for Buoyancy. The Poisson equation, Eq. (4.6), is derived from the continuity of the current density.

$$\nabla \cdot \boldsymbol{J} = 0 \tag{4.7}$$

$$\nabla \cdot (\boldsymbol{E}_\mathrm{m} + \boldsymbol{u} \times \boldsymbol{B} + \cancel{\rho_\mathrm{e}\boldsymbol{u}}) = 0 \qquad (\because \text{ Ohm's law}) \tag{4.8}$$

$$\nabla \cdot (-\nabla\phi + \boldsymbol{u} \times \boldsymbol{B}) = 0 \tag{4.9}$$

$$f_\phi = \nabla^2\phi - \nabla \cdot (\boldsymbol{u} \times \boldsymbol{B}) = 0 \tag{4.10}$$

Here $\rho_\mathrm{e}\boldsymbol{u}$ in Ohm's law is negligible assuming the electrical neutrality of Si melt.
The source term of Eq. (4.6) becomes

$$\nabla \cdot (\boldsymbol{u} \times \boldsymbol{B}) = 0, \tag{4.11}$$

in 2-dimensional axisymmetric analysis.
The boundary conditions for each velocity component are given as follows;

$$0 = g_u = \begin{cases} u & \text{along } S_\mathrm{seed},\ S_\mathrm{side},\ S_\mathrm{bottom},\ l_\mathrm{axis} \\ \frac{\partial u}{\partial z} & \text{along } S_\mathrm{surface} \end{cases}, \tag{4.12}$$

$$0 = g_w = \begin{cases} w & \text{along } S_\mathrm{seed},\ S_\mathrm{surface},\ S_\mathrm{side} \\ (w - U_z) & \text{along } S_\mathrm{bottom} \\ \frac{\partial w}{\partial r} & \text{along } l_\mathrm{axis} \end{cases}, \tag{4.13}$$

where S_seed, S_side, S_bottom, S_surface and l_axis are the boundaries along the seed crystal, side wall, bottom wall, free surface, and the center axis, respectively. U_z is the parameterized axial velocity by p_{u_z} and defined as

$$U_z = p_{u_z} \{\sin(2\pi r / r_\mathrm{c})\} \times 10^{-3}. \tag{4.14}$$

and utilized in the training of PINNs as described in section 4.2, even though it is zero in normal numerical simulations and the prediction. The sine function was chosen in order to keep the liquid volume balance in the melt.
The temperature boundary conditions were defined as follows;

$$0 = g_T = \begin{cases} T - T_\mathrm{m} & \text{along } S_\mathrm{seed} \\ \frac{\partial T}{\partial z} + \frac{\epsilon\sigma F}{k_T}\left(T^4 - T_\mathrm{amb}^4\right) & \text{along } S_\mathrm{surface} \\ T - \frac{T_\mathrm{r} - T_\mathrm{h}}{h_\mathrm{s}}(z - 0.3) + T_\mathrm{r} & \text{along } S_\mathrm{side} \\ T - \frac{T_\mathrm{h} - T_\mathrm{l}}{r_\mathrm{c}^2} r^2 + T_\mathrm{l} & \text{along } S_\mathrm{bottom} \\ \frac{\partial T}{\partial r} & \text{along } l_\mathrm{axis} \end{cases}, \tag{4.15}$$

where T_{m} is the melting point of the liquid Si, ϵ the emissivity, σ the Stefan-Boltzmann constant, F the view factor, k_T the thermal conductivity, T_{amb} the ambient temperature, h_{s} the height of the side wall of the crucible, r_{c} the crucible radius, and T_{l}, T_{r}, T_{h} the temperature at the bottom center, surface edge, and the bottom edge of the Si melt, respectively. The values of the view factor and ambient temperature are 0.3 and 1600 K, respectively.
The boundary condition for ϕ is derived from a non-penetration condition for current density as follows;

$$\boldsymbol{J} \cdot \boldsymbol{n} = 0 \tag{4.16}$$

$$(-\nabla\phi + \boldsymbol{u} \times \boldsymbol{B}) \cdot \boldsymbol{n} = 0 \tag{4.17}$$

$$g_\phi = \nabla\phi \cdot \boldsymbol{n} - (\boldsymbol{u} \times \boldsymbol{B}) \cdot \boldsymbol{n} = 0, \tag{4.18}$$

where $\boldsymbol{n} = (n_r,\ n_z)$ is the unit vector normal to the boundary surface. The second term of Eq. (4.18) becomes

$$(\boldsymbol{u} \times \boldsymbol{B}) \cdot \boldsymbol{n} = 0 \tag{4.19}$$

in 2-dimensional axisymmetric analysis. Therefore, due to Eqs. (4.11), (4.19), ϕ = const. in this study. The numerical simulation for the calculation of the accuracy of PINNs was conducted by OpenFOAM [77]. The values of the physical properties are shown in Table 4.1.

Table 4.1 Physical properties utilized in this study. ([16] and [95])

variable	parameter	value
g	gravitational acceleration vector [ms^{-2}]	-9.81
k	thermal conductivity [$\mathrm{Wm}^{-1}\mathrm{K}^{1}$]	67
T_{m}	melting point [K]	1685
α	thermal diffusivity [$\mathrm{m}^2\mathrm{s}^{-1}$]	2.564×10^{-5}
β	thermal expansion coefficient [T^{-1}]	$1.4{\times}10^{-4}$
ϵ	emissivity [−]	0.3
ν	kinetic viscosity [$\mathrm{m}^2\mathrm{s}^{-1}$]	$2.778{\times}10^{-7}$
ρ	density [kgm^{-3}]	2520
σ_{e}	electrical conductivity [Sm^{-1}]	1.2×10^6

4.2 Physics Informed Neural Networks

4.2.1 Architecture and loss functions

Physics Informed Neural Networks (PINNs) are neural networks learning physical laws, e. g. partial differential equations. The architecture of developed PINNs is shown in Fig. 4.2. The

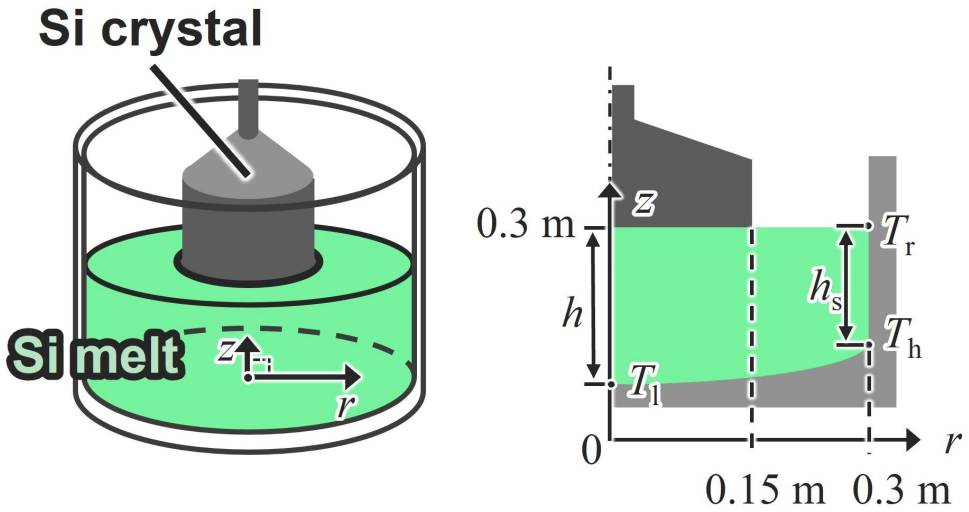

Figure 4.1. The whole domain (left) of Cz-Si and the numerical domain (right), showing the half cross-sectional view of the whole domain. The input parameters of PINNs are shown in the right figure. T_l, T_h, and T_r are the temperatures at the black dots. h and h_s represent the melt heights at the center and the edge of the melt, respectively.

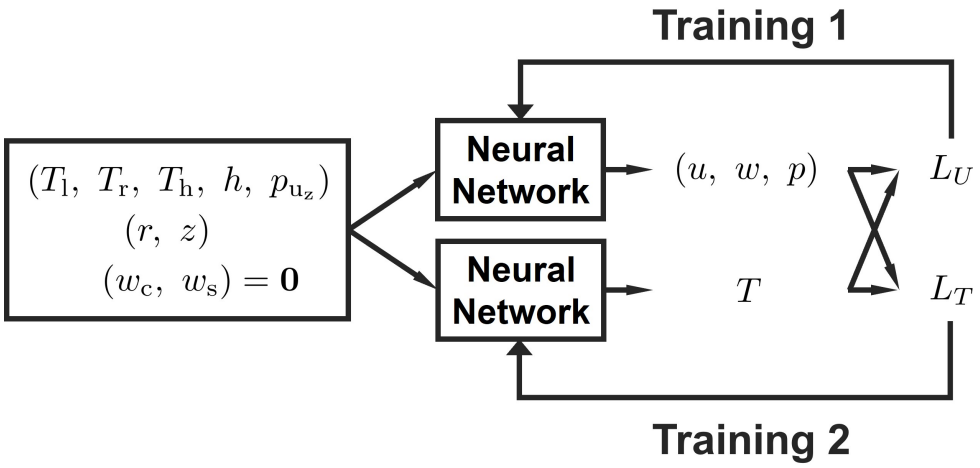

Figure 4.2. The architecture of PINNs developed in this study. The neural networks are trained separately. The input parameters $(w_c,\ w_s)$ represent the rotational speed of the crucible and the seed, respectively, and were included for future development considering the effect of seed and crucible rotation. However, in this study, they were kept constant at 0 rpm and did not contribute to the learning and prediction process.

neural networks for the prediction of the velocity (and pressure) and temperature were developed separately. Due to the separate development, the temperature is not learned to satisfy the physical laws of the velocity. These neural networks are called U- and T-nets, respectively in this study. On the other hand, the neural network for ϕ was not developed since ϕ = const. is always satisfied in this study. U-net in this study is not related to U-net[96] which is well known as the

architecture of convolutional network for image segmentation.

The input vector $\boldsymbol{x}$ is defined as

$$\boldsymbol{x} = (r,\ z,\ T_{\mathrm{l}},\ T_{\mathrm{r}},\ T_{\mathrm{h}},\ h,\ p_{u_z},\ \omega_{\mathrm{c}},\ \omega_{\mathrm{s}}), \tag{4.20}$$

where h is the melt heights (see Fig. 4.1), ω_{c} the rotational speed of the crucible, and ω_{s} the rotational speed of the seed crystal, and its ranges were defined as

$$\begin{aligned}
0 \leq \quad & r\ [\mathrm{m}] && \leq r_{\mathrm{c}} && (4.21)\\
z_{\mathrm{b}}(h) \leq \quad & z\ [\mathrm{m}] && \leq 0.3 && (4.22)\\
1685 \leq \quad & T_{\mathrm{l}}\ [\mathrm{K}] && \leq 1720 && (4.23)\\
T_{\mathrm{l}} \leq \quad & T_{\mathrm{r}}\ [\mathrm{K}] && \leq 1720 && (4.24)\\
T_{\mathrm{l}} \leq \quad & T_{\mathrm{h}}\ [\mathrm{K}] && \leq 1720 && (4.25)\\
0.108 \leq \quad & h\ [\mathrm{m}] && \leq 0.3 && (4.26)\\
-1 \leq \quad & p_{u_z}\ [\mathrm{m/s}] && \leq 1 && (4.27)\\
& \omega_{\mathrm{c}}\ [\mathrm{rpm}] && = 0 && (4.28)\\
& \omega_{\mathrm{s}}\ [\mathrm{rpm}] && = 0. && (4.29)
\end{aligned}$$

The rotational parameters $\omega_{\mathrm{c}},\ \omega_{\mathrm{s}}$ were always set to 0 rpm since rotation was not considered in this study. However, the rotational parameters were added in order to consider the rotation in the future. In the input layer, the values of the input vector are normalized from -1 to 1.

The fully connected neural networks with 8 hidden layers and 50 neurons in the layers were used for U- and T-nets. Tanh and identity functions were used as activation functions in all hidden and output layers, respectively. The weights of each network were initialized by glorot normal[97] defined as

$$w^i \sim \mathcal{N}\left(0,\ \frac{2}{n^i + n^{i+1}}\right), \tag{4.30}$$

where n^i is the number of columns of the weights of i-th layer.

The outputs of the U- and T-nets are velocity and pressure $(u,\ w,\ p)$ and temperature T, respectively. From the predicted outputs, the loss functions L_U and L_T are calculated. L_U and L_T are defined as

$$\begin{aligned}
L_U &= L_{f_U} + L_{g_U} + L_{f_p} && (4.31)\\
L_T &= L_{f_T} + L_{g_T}, && (4.32)
\end{aligned}$$

where

$$
\begin{aligned}
L_{f_U} &= \frac{w_{f_u}}{N_{f_U}} \sum_{j=1}^{N_{f_U}} \left(w_{\mathrm{SDF}}^{j} | f_u \left(\boldsymbol{X}_{f_u}^{j} \right) |^2 \right) \\
&+ \frac{w_{f_w}}{N_{f_U}} \sum_{j=1}^{N_{f_U}} \left(w_{\mathrm{SDF}}^{j} | f_w \left(\boldsymbol{X}_{f_w}^{j} \right) |^2 \right) \qquad (4.33) \\
L_{g_U} &= \frac{w_{g_u}}{N_{g_U}} \sum_{j=1}^{N_{g_U}} | g_u \left(\boldsymbol{X}_{g_u}^{j} \right) |^2 + \frac{w_{g_w}}{N_{g_U}} \sum_{j=1}^{N_{g_U}} | g_w \left(\boldsymbol{X}_{g_w}^{j} \right) |^2 \qquad (4.34) \\
L_{f_p} &= \frac{w_{f_p}}{N_{f_p}} \sum_{j=1}^{N_{f_p}} \left(w_{\mathrm{SDF}}^{j} | f_p \left(\boldsymbol{X}_{f_p}^{j} \right) |^2 \right) \qquad (4.35) \\
L_{f_T} &= \frac{w_{f_T}}{N_{f_T}} \sum_{j=1}^{N_{f_T}} \left(w_{\mathrm{SDF}}^{j} | f_T \left(\boldsymbol{X}_{f_T}^{j} \right) |^2 \right) \qquad (4.36) \\
L_{g_T} &= \frac{1}{N_{g_T}} \sum_{j=1}^{N_{g_T}} \left(w_{g_T}^{j} | g_T \left(\boldsymbol{X}_{g_T}^{j} \right) |^2 \right), \qquad (4.37)
\end{aligned}
$$

and w_i ($i = f_u,\ f_w,\ f_p,\ f_T,\ g_u,\ g_w,\ g_T$), N_i and $\boldsymbol{X}_i$ are the weight, the number of the training points, and the training points matrix for function i, respectively and $\boldsymbol{X}_{f_u} = \boldsymbol{X}_{f_w} = \boldsymbol{X}_{f_p}$. The training points for the loss functions of the governing equations were randomly sampled in the range of Eqs. (4.21~4.29) and for the boundary conditions in the range of Eqs. (4.23~4.29) on each boundary surface. The values of the weights were tuned during the training process (see section 4.2.2) manually.

L_i ($i = f_u,\ f_w,\ f_p,\ f_T,\ g_u,\ g_w,\ g_T$) (Eqs. (4.33)~(4.37)) represents the mean squared error (MSE) of function i. Since ($f = 0$) and ($g = 0$) mean that the governing equation and boundary condition, respectively, the mean squared errors of f and g indicate the magnitude of dissatisfaction with the governing equation and boundary condition, respectively. L_U and L_T indicate the simple weighted sum of the L_i.

w_{SDF} and w_{g_T} changing weights by the training point, in contrast to the other weights which have constant values.
w_{SDF} is the signed distance function defined as

$$
w_{\mathrm{SDF}}^{j} = \min \left\{ \min \left(r^j,\ r_{\mathrm{c}} - r^j \right),\ \min \left(z^j - z_{\mathrm{b}},\ 0.3 - z^j \right) \right\} / 0.15 \qquad (4.38)
$$

w_{SDF} goes to zero near the boundary surfaces as shown in Fig. 4.3. Thus the governing equations are learned only in the numerical domain not on the boundary surfaces by w_{SDF}.

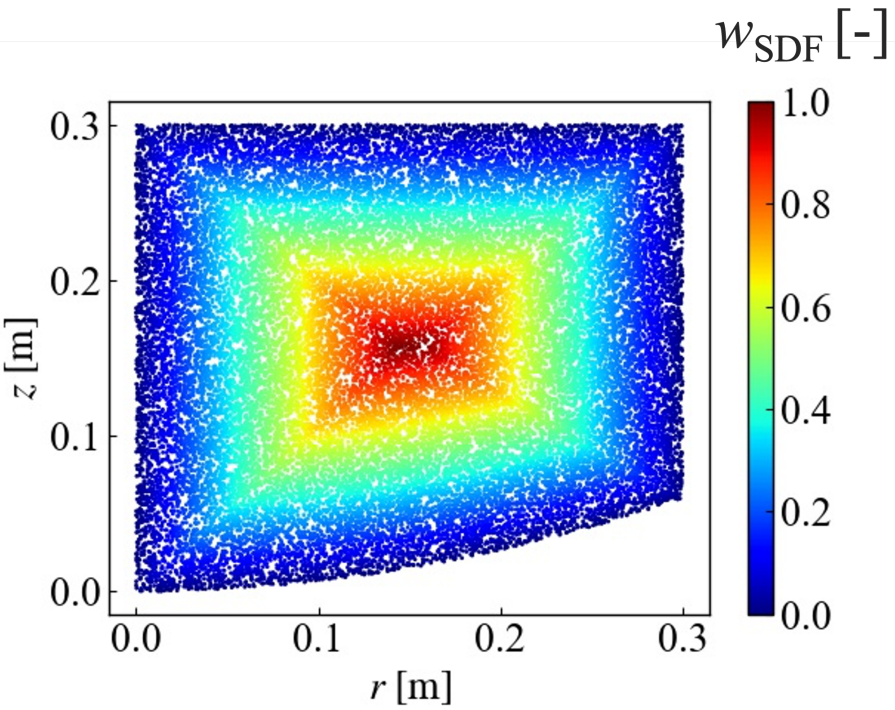

Figure 4.3. The distribution of singed distance function w_{SDF} at $h = 0$.

w_{gT} was defined on boundary S_{surface} as

$$w_{gT}^{j} = \begin{cases} 300 \times \frac{2\left(r^{j}-r_{\mathrm{s}}\right)}{r_{\mathrm{c}}-r_{\mathrm{s}}} & r_{\mathrm{s}} \leq r^{j} < r_{\mathrm{s}} + \frac{(r_{\mathrm{c}}-r_{\mathrm{s}})}{2} \text{ on } S_{\mathrm{surface}} \\ -300 \times \frac{2\left(r^{j}-r_{\mathrm{c}}\right)}{r_{\mathrm{c}}-r_{\mathrm{s}}} & r_{\mathrm{s}} + \frac{(r_{\mathrm{c}}-r_{\mathrm{s}})}{2} \leq r^{j} \leq r_{\mathrm{c}} \text{ on } S_{\mathrm{surface}} \end{cases} . \tag{4.39}$$

w_{gT} was used to prevent the model from learning two types of boundary conditions at both edges of the free surface.

4.2.2 Training process

The training was conducted by minimizing the loss functions L_U and L_T. However, it was difficult to train U- and T- nets simultaneously in the whole parameter range. Thus, the following training process was used;

1. The T-net was trained with the constraint of ($\boldsymbol{u} = \mathbf{0}$) which represents thermal conduction.

2. The U-net was trained with the parameters h and p_{u_z} fixed at 0, while keeping the T-net fixed.
   At the same time, the weights for the loss function suitable for the training were searched.

3. The p_{u_z} was not fixed, and then the U-net was trained with the parameter h fixed at 0, while keeping the T-net fixed.

4. The h was not fixed, and then the U-net was trained with the T-net fixed.

5. The T-net was trained with U-net fixed.

6. The U-net was trained with the T-net fixed, and the weights of the loss function L_U were optimized.

7. Finally, T-net was trained again.

During the training step, the U- and T-nets were trained interactively, and the number of the changing-input parameters was also gradually increased through transfer learning.

The Adam optimizer[98] was used. The learning rates of U- and T-nets were

$$(lr)_U = 0.001 \times 0.5^{iter./30000} \tag{4.40}$$

$$(lr)_T = 0.001 \times 0.5^{iter./10000}, \tag{4.41}$$

respectively, where $iter.$ is the number of training iterations.
The hyper-parameters are listed in Table 4.2.

4.3 Predictions by PINNs

Figure 4.4 shows the distributions of the velocity components and the temperature in the melt obtained from the numerical simulation and the PINNs, at $(T_\mathrm{l},\ T_\mathrm{r},\ T_\mathrm{h},\ p_{u_z})$ = (1690 K, 1710 K, 1715 K, 0 m/s) and varying melt heights h (= 0.3, , 0.25, 0.2 and 0.15 m). Moreover, the velocity and temperature profile along the vertical and horizontal lines under the same temperature boundary conditions as Fig. 4.4 at h = 0.3, 0.15 m are shown in Fig. 4.5.
The relative L2 errors of u, w, and T are 0.02042, 0.04867, and 1.797×10^{-8} at h = 0.3 m, 0.0192, 0.0414, 1.162×10^{-8} at h = 0.25 m, 0.0224, 0.03551, and 7.208×10^{-9} at h = 0.2 m, and 0.02525, 0.04609, and 6.542×10^{-9} at h = 0.15 m, respectively. Thus, accurate predictions were achieved even at different melt heights and the conditions shown in Fig. 4.4.
The predictions with different melt heights were carried out only by changing the value of the input parameter h, while mesh re-productions for each melt height were required in the numerical simulation. The time required for the prediction by PINNs was less than 0.1 seconds, while the numerical simulation took about 30 minutes. Furthermore, in this study, PINNs were developed without any answer-labeled training data generated by the numerical simulation, which is typically required in traditional machine-learning approaches. Therefore, in addition

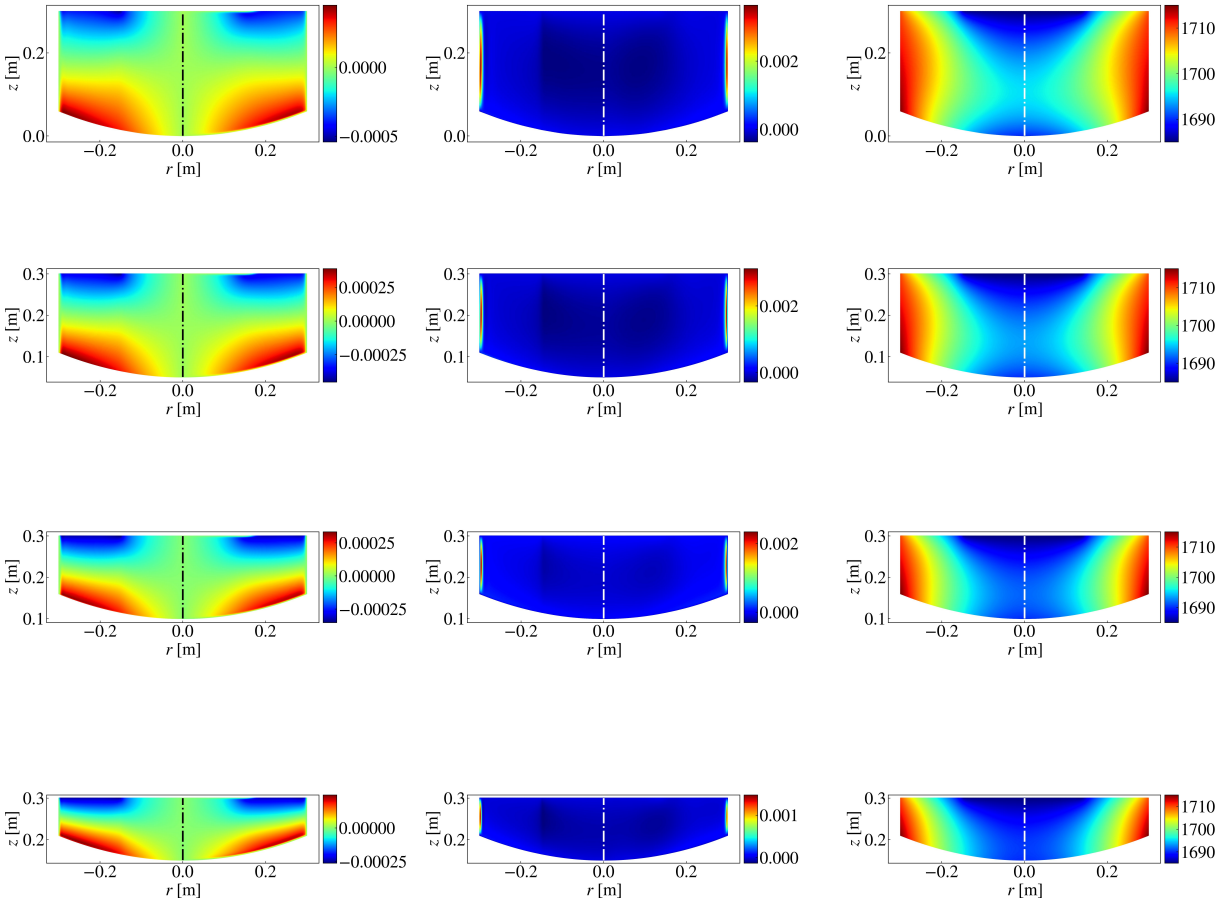

Figure 4.4. The distributions of the radial velocity u (a, d, g, j), the axial velocity w (b, e, h, k), and the temperature T (c, f, i, l) in the melt obtained from the numerical simulation (left) and PINNs (right) at the input parameter $(T_\mathrm{l},\ T_\mathrm{r},\ T_\mathrm{h},\ p_{u_z}) = (1690\ \mathrm{K},\ 1710\ \mathrm{K},\ 1715\ \mathrm{K},\ 0\ \mathrm{m/s})$, $h = 0.3$ (a, b, c), 0.25 (d, e, f), 0.2 (g, h, i) and 0.15 m (j, k, l).

to the significant reduction in time for prediction, the time and effort to be spent for the mesh reproduction, the evaluation of mesh dependency, and the development of the answer-labeled training dataset by a numerical simulation are eliminated. This is a significant gain.

4.4 Accuracy of PINNs

In the previous section, we discussed the efficiency of PINNs by verifying the prediction accuracy under only one temperature condition. However, it is necessary to validate the accuracy in the entire parameter space. The accuracy was evaluated by relative L2 error against the test data. The test data consists of equidistantly distributed T_l, T_r, and T_h values in the range of $(T_\mathrm{l} \leq T_\mathrm{r} \leq T_\mathrm{h})$ at the four distinct values of h $(= 0.3,\ 0.2,\ 0.25,\ \text{and } 0.15\ \mathrm{m})$.

The histograms shown in Fig. 4.6 indicate the frequencies of the relative L2 errors of the radial velocity u, the axial velocity w, and the temperature T at varying melt heights. Since the mode

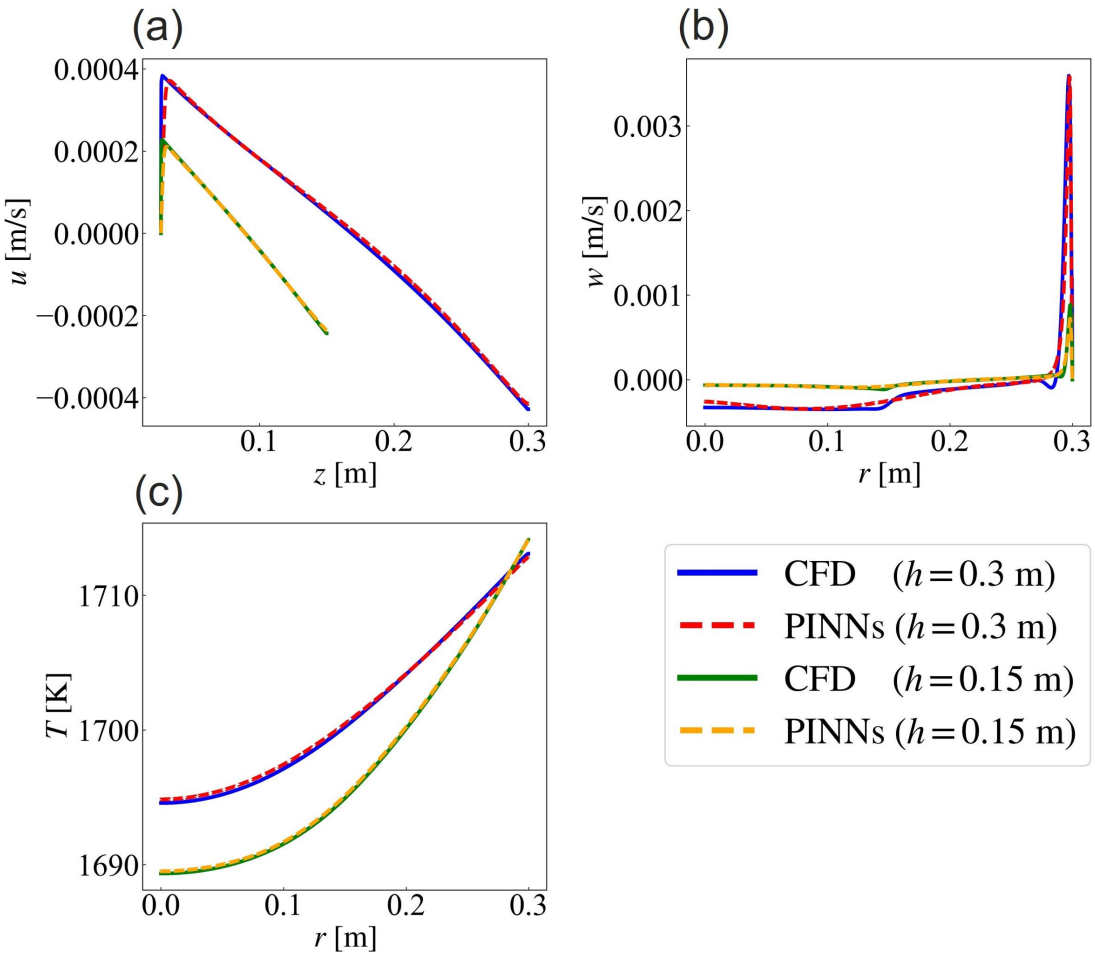

Figure 4.5. The radial velocity u (a) profile along the vertical line at $r = 0.2$ m, and the axial velocity w (b) and temperature T (c) profiles along the horizontal center line under the same input parameter as Fig. 4.4.

values of the errors are closely the same as those presented in Figs. 4.4 and 4.5, it is concluded that highly accurate PINNs were developed in the whole parameter space (even under different melt heights).

A small number of high error values were also observed. In order to investigate the parameters presenting high errors, the distributions of the relative L2 error in parameter spaces (T_l, T_r, T_h) at different melt heights are shown in Fig. 4.7. As shown in the figure, the relative L2 error tends to increase for cases closer to the edge of the parameter space. A few cases representing the high error values shown in Fig. 4.7 exist only around the edge of the parameter space. Therefore, accurate predictions with low errors were achieved over a wide range of the domain, except for around the edges of the parameter space.

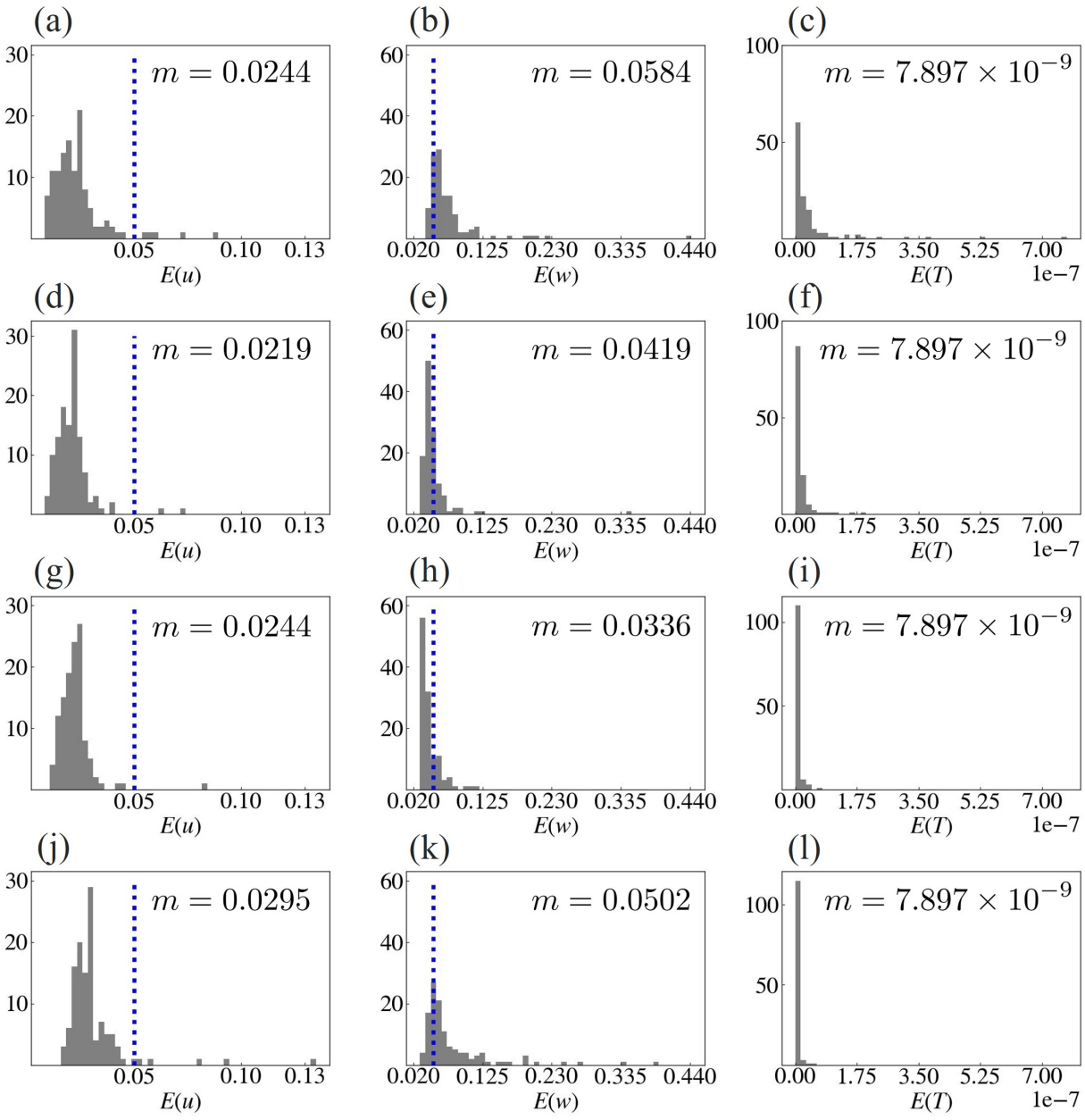

Figure 4.6. The histograms of the relative L2 error of u (a, d, g, j), w (b, e, h, k) and T (c, f, i, l) of different melt heights h = 0.3 m (a, b, c), 0.25 m (d, e, f), 0.2 m (g, h, i), 0.15 m (j, k, l). m indicates the mode value.

4.5 Current achievements, limitations, and future improvements

The fast prediction method of the steady-flow velocity and temperature distributions in the Si melt of Cz method with the assumption of 2-dimensional axisymmetry, no rotations, flat liquid and crystal surface, and the fixed vertical magnetic field. However, in the actual process,

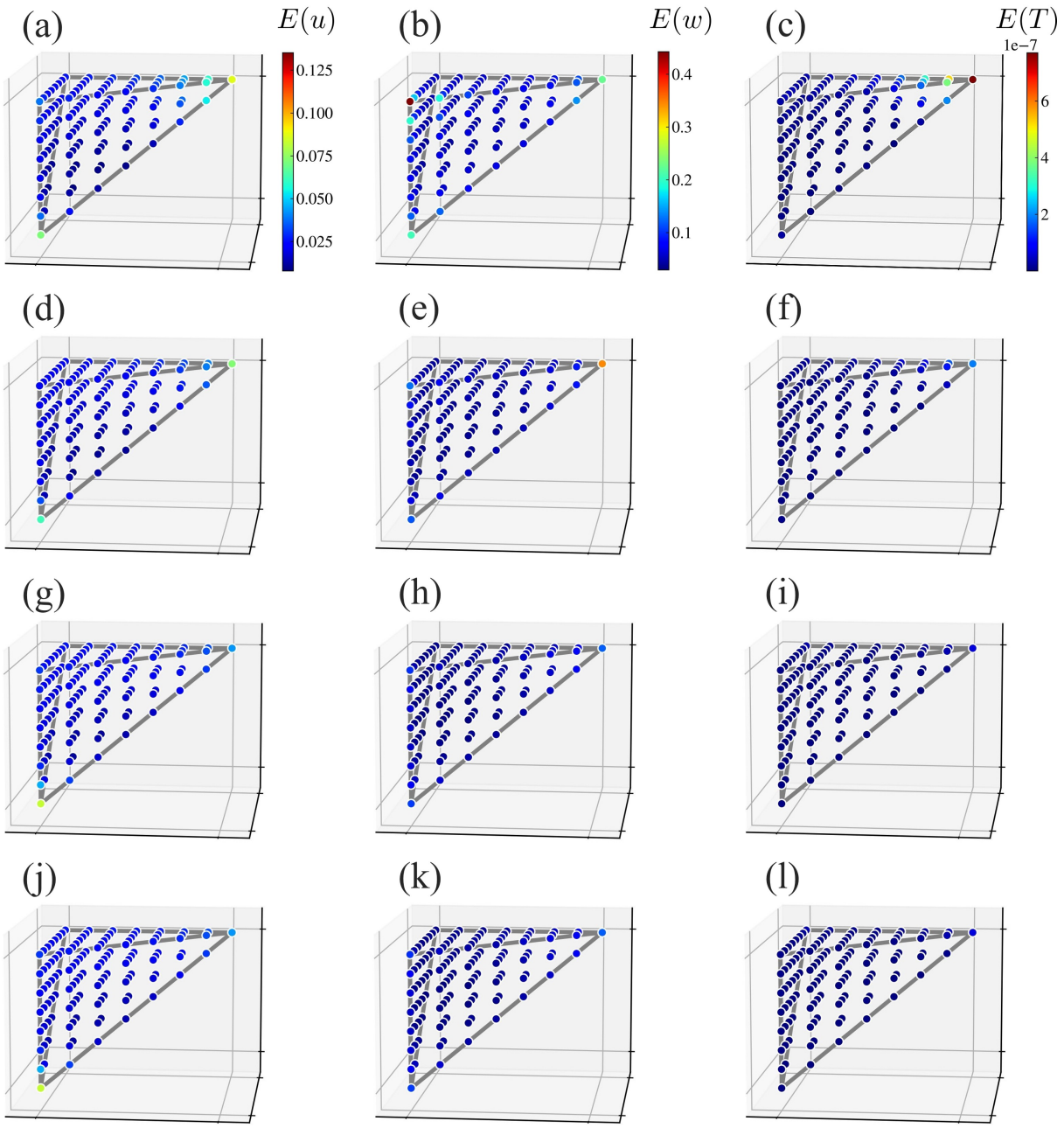

Figure 4.7. The histograms of the relative L2 error of u (a, d, g, j), w (b, e, h, k) and T (c, f, i, l) of different melt heights h = 0.3 m (a, b, c), 0.25 m (d, e, f), 0.2 m (g, h, i), 0.15 m (j, k, l). m indicates the mode value..

unsteady melt flow develops, the crucible and seed rotate, the horizontal or cusp magnetic field is utilized, and the liquid surface and crystal are not flat. Moreover, it may not be easy to measure the temperature utilized in the input parameters. For practical use of PINNs, efforts should be made to bring the targets closer to the actual system.

The most difficult thing is to introduce unsteady PINNs. For developing unsteady PINNs, the

parameter, time, is added to the input parameter. For parameters except time, different parameter values indicate different states. On the other hand, different times with the same values of the other parameters mean the same condition. Because of that, it is necessary to generate a large amount of training data at different times for the same condition, which exceeds the amount of memory on the GPU. Moreover, in order to predict time evolution, the state prior to the input time value must be accurately learned, which is not guaranteed in PINN training because random time are given at the same time. Because of the difficulties mentioned above, the application of PINNs to unsteady systems is difficult, and different strategies should be developed.

One possible approach is the combination of numerical simulation and PINNs. In solving unsteady problems by numerical simulation, it takes a long time to solve the Poisson equation of pressure. The source term of the Poisson equation changes at every time step. Thus, PINNs that predict the solution of the Poisson equation with different source terms will resolve the difficulty. PINNs that immediately predict the temperature field at different heating terms have been reported[99]. By utilizing the same method, PINNs for the Poisson equation will be developed, which enables to decrease in the time for unsteady numerical simulation.

Another difficulty is the training of PINNs itself. For application to real systems, many more parameters need to be added to the input parameters and multi-region calculations involving solids and fluids need to be considered. There are many know-hows for successful learning even in this study, and further know-how will be required for the application to real systems. For example, this includes guidelines for adjusting the weights of the loss function, methods for training multiple neural networks, and methods for generating training data. In addition to that, the know-how cannot be systematically summarized, which leads to the difficulty of the application to other systems. This problem will be solved when PINN know-how is developed as an academic field, as numerical computation has been developed.

Table 4.2 The hyper-parameters of PINNs. These values were optimized through a trial and error process to minimize relative L2 error against the validation dataset without clear criteria. w_i and N_i are the weights and the number of training points for loss function i, respectively.

network	parameter	value
T	layers	8
	neurons	50
	w_{f_T}	1×10^2
	w_{g_T} (on S_{seed}, S_{side}, S_{bottom})	1×10^4
	w_{g_T} (on l_{axis})	3×10^2
	w_{g_T} (on S_{surface})	see section 4.2.1
	N_{f_T}	3×10^4
	N_{g_T}	3×10^4
	optimizer	Adam
	activation function	tanh
	learning rate	$0.001 \times 0.5^{iter./30000}$
U	layers	8
	neurons	50
	w_{f_p}	1×10^6
	w_{f_u}	1×10^5
	w_{f_w}	1×10^7
	w_{g_u} (on S_{seed}, S_{side}, S_{bottom}, l_{axis})	1×10^2
	w_{g_u} (on S_{surface})	1
	w_{g_w} (on S_{seed}, S_{side}, S_{bottom})	1×10^2
	w_{g_w} (on S_{surface})	1×10^6
	w_{g_w} (on l_{axis})	1×10^1
	N_{f_U}	3×10^4
	N_{f_p}	3×10^4
	N_{g_U}	3×10^4
	optimizer	Adam
	activation function	tanh
	learning rate	$0.001 \times 0.5^{iter./10000}$

Chapter 5

Summary and future works

5.1 Summary

The fast and white-box suggestion of control methods was achieved by data science and transport phenomena, which enables

- White-box control based on the transport phenomena in the process
- Fast analysis of the transport phenomena even during experiments
- Fast suggestion of control methods
- Easy analysis assisted by AI for identifying key phenomena.

Thus, a revolutionary control system that immediately analyzes the transport phenomena during process operation and rapidly suggests control methods and key phenomena based on the results of the analysis may be accomplished.
The author first achieved the fast suggestion of the control method of the transport phenomena in the Cz-Si system by using Bayesian optimization. The optimal state achieving the high- and uniform-crystal growth rate was obtained with few trials. Moreover, the key melt flow to the improvement of the crystal growth rate was analyzed by numerical simulation.
The analysis of the transport phenomena can be assisted by AI by using the XAI proposed in this study. The XAI detected automatically the melt flow near the seed crystal and bottom center in the melt as the key flows to the improvements of the crystal growth rate. Due to the detection, in addition to the upward flow below the seed crystal which has been already detected by researchers, the flow separation of the Marangoni convection from the seed crystal and the upward flow near the bottom center were first found to be important for the improvements.
Finally, the author developed the fast and white-box prediction method of transport phenomena based on physical laws by using PINNs. This prediction method enables the fast analysis of

the transport phenomena. Thus, real-time analysis and suggestion of suitable control methods during experiments will be achieved by combining the proposed methods.
The developed analysis, prediction, and optimization methods can be utilized in any other process, although the target system is crystal growth in this study. Therefore, this study may lead to a new control methodology in chemical engineering.

5.2 Future works

The problems described in the final section of each chapter remain. Unsteady PINNs available for practical processes would solve the problem of application and efficiency of Bayesian optimization. Moreover, the fast prediction would enable the optimization even by full parametric search without optimization algorithms. Furthermore, the generation of training data for XAI can be conducted quickly by PINNs. In other words, improvements in PINNs solve the current difficulties. Because of that, it is important to achieve the same level of analytical capability in PINNs as in numerical calculations today.
In order to achieve them, the following works remain for that in the future:

- The fast prediction of the unsteady problem will be developed by the combination of numerical simulation and PINNs as mentioned in section 4.5.
- The methodology of learning PINNs will be systematized to have the same level of computational ability as numerical simulations.

These problems and tasks should be tackled in the future.

Bibliography

[1] J. Czochralski, Ein neues verfahren zur messung der kristallisationsgeschwindigkeit der metalle, Z. Physik. Chemic 92 (1918) 219–221. doi:https://www.degruyter.com/document/doi/10.1515/zpch-1918-9212/html.

[2] G. Mueller, The Czochralski Method-where we are 90 years after Jan Czochralski's invention, Cryst. Res. Technol. 42 (12) (2007) 1150–1161. doi:https://doi.org/10.1002/crat.200711001.

[3] S. Hu, W. Patrick, Effect of oxygen on dislocation movement in silicon, J. Appl. Phys. 46 (5) (1975) 1869–1874. doi:https://doi.org/10.1063/1.321883.

[4] K. Hoshikawa, H. Kohda, K. Ikuta, Improvement in CZ Silicon Wafer by Reducing Oxygen Impurity, Japanese J. Appl. Phys. 20 (S1) (1981) 241. doi:10.7567/JJAPS.20S1.241.

[5] K. Sumino, H. Harada, I. Yonenaga, The origin of the difference in the mechanical strengths of Czochralski-grown silicon and float-zone-grown silicon, Japanese J. Appl. Phys. 19 (1) (1980) L49. doi:10.1143/JJAP.19.L49.

[6] T. Tan, E. Gardner, W. Tice, Intrinsic gettering by oxide precipitate induced dislocations in Czochralski si, Appl. Phys. Lett. 30 (4) (1977) 175–176. doi:https://doi.org/10.1063/1.89340.

[7] A. Witt, C. Herman, H. Gatos, Czochralski-type crystal growth in transverse magnetic fields, J. Mater. Sci. 5 (1970) 822–824. doi:https://doi.org/10.1007/BF00562171.

[8] K. Hoshi, N. Isawa, T. Suzuki, Y. Ohkubo, Czochralski silicon crystals grown in a transverse magnetic field, J. Electrochem. Soc. 132 (3) (1985) 693. doi:10.1149/1.2113933.

[9] K. Hoshikawa, Czochralski silicon crystal growth in the vertical magnetic field, Japanese. J. Appl. Phys. 21 (9A) (1982) L545. doi:10.1143/JJAP.21.L545.

[10] H. Hirata, K. Hoshikawa, Silicon crystal growth in a cusp magnetic field, J. Cryst. Growth 96 (4) (1989) 747–755. doi:https://doi.org/10.1016/0022-0248(89)90631-3.

[11] V. Savolainen, J. Heikonen, J. Ruokolainen, O. Anttila, M. Laakso, J. Paloheimo, Simulation of large-scale silicon melt flow in magnetic Czochralski growth, J. Cryst. Growth 243 (2) (2002) 243–260. doi:https://doi.org/10.1016/S0022-0248(02)01441-0.

[12] V. Kalaev, Combined effect of DC magnetic fields and free surface stresses on the melt flow and crystallization front formation during 400 mm diameter Si Cz crystal growth, J. Cryst. Growth 303 (1) (2007) 203–210. doi:https://doi.org/10.1016/j.jcrysgro.2006.11.345.

[13] X. Liu, L. Liu, Z. Li, Y. Wang, Effects of static magnetic fields on thermal fluctuations in the melt of industrial CZ-Si crystal growth, J. Cryst. Growth 360 (2012) 38–42. doi:https://doi.org/10.1016/j.jcrysgro.2011.11.080.

[14] L. Liu, K. Kakimoto, 3D global analysis of Cz–Si growth in a transverse magnetic field with various crystal growth rates, J. Cryst. Growth 275 (1-2) (2005) e1521–e1526. doi:https://doi.org/10.1016/j.jcrysgro.2004.11.185.

[15] N. Ivanov, A. Korsakov, E. Smirnov, K. Khodosevitch, V. Kalaev, Y. N. Makarov, E. Dornberger, J. Virbulis, W. Von Ammon, Analysis of magnetic field effect on 3D melt flow in CZ Si growth, J. Cryst. Growth 250 (1-2) (2003) 183–188. doi:https://doi.org/10.1016/S0022-0248(02)02263-7.

[16] S. Jana, S. Dost, V. Kumar, F. Durst, A numerical simulation study for the Czochralski growth process of Si under magnetic field, Int. J. Eng. Sci. 44 (8-9) (2006) 554–573. doi:https://doi.org/10.1016/j.ijengsci.2006.02.005.

[17] Q. Xun, B. Xun, Z. Li, P. Wang, Z. Cai, Application of SiC power electronic devices in secondary power source for aircraft, Renewable and Sustainable Energy Reviews 70 (2017) 1336 – 1342. doi:https://doi.org/10.1016/j.rser.2016.12.035.

[18] D. Chaussende, P. Wellmann, M. Pons, Status of SiC bulk growth processes, J. Phys. D 40 (20) (2007) 6150. doi:10.1088/0022-3727/40/20/S02.

[19] C. Moulin, M. Pons, A. Pisch, P. Grosse, C. Faure, A. Basset, G. Basset, A. Passero, T. Billon, B. Pelissier, et al., SiC single crystal growth by sublimation: experimental and numerical results, in: Materials Science Forum, Vol. 353, Trans Tech Publications Ltd., Zurich-Uetikon, Switzerland, 2001, pp. 7–10. doi:https://doi.org/10.4028/www.scientific.net/MSF.353-356.7.

[20] K. Kusunoki, N. Okada, k. Kamei, K. Moriguchi, H. Daikoku, M. Kado, H. Sakamoto, T. Bessho, T. Ujihara, Top-seeded solution growth of three-inch-diameter 4H-SiC using convection control technique, J. Cryst. Growth 395 (2014) 68–73. doi:https://doi.org/10.1016/j.jcrysgro.2014.03.006.

[21] F. Mercier, J. M. Dedulle, D. Chaussende, M. Pons, Coupled heat transfer and fluid dynamics modeling of high-temperature SiC solution growth, J. Cryst. Growth 312 (2) (2010) 155–163. doi:10.1016/j.jcrysgro.2009.10.007.

[22] K. Murayama, T. Hori, S. Harada, S. Xiao, M. Tagawa, T. Ujihara, Two-step SiC solution growth for dislocation reduction, J. Cryst. Growth 468 (2017) 874–878. doi:https://doi.org/10.1016/j.jcrysgro.2016.11.100.

[23] T. Yamamoto, Y. Okano, T. Ujihara, S. Dost, Global simulation of the induction heating TSSG process of SiC for the effects of Marangoni convection, free surface deformation and seed rotation, J. Cryst. Growth 470 (470) (2017) 75–88. doi:10.1021/ja00040a007.

[24] M.-T. Ha, Y.-J. Shin, M.-H. Lee, C.-J. Kim, S.-M. Jeong, Effects of the Temperature Gradient Near the Crystal-Melt Interface in Top Seeded Solution Growth of SiC Crystal, Phys. Status Solidi(A) 215 (20) (2018) 1701017. doi:https://doi.org/10.1002/pssa.201701017.

[25] T. Umezaki, D. Koike, S. Harada, T. Ujihara, Analysis of the carbon transport near the growth interface with respect to the rotational speed of the seed crystal during top-seeded solution growth of SiC, Japanese. J. Appl. Phys. 55 (2016) 125601–1–125601–5. doi:10.7567/JJAP.55.125601.

[26] F. Mercier, S. Nishizawa, Comparative numerical study of the effects of rotating and traveling magnetic fields on the carbon transport in the solution growth of SiC crystals, J. Crst. Growth 362 (2013) 99–102. doi:10.1016/j.jcrysgro.2011.11.019.

[27] L. Wang, T. Horiuchi, A. Sekimoto, Y. Okano, T. Ujihara, S. Dost, Three-dimensional numerical analysis of Marangoni convection occurring during the growth process of SiC by the RF-TSSG method, J. Cryst. Growth 520 (2019) 72–81. doi:10.1016/j.jcrysgro.2019.05.017.

[28] T. E. Oliphant, Python for scientific computing, Computing in science & engineering 9 (3) (2007) 10–20. doi:10.1109/MCSE.2007.58.

[29] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, software available from tensorflow.org (2015).
URL https://www.tensorflow.org/

[30] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems 32, Curran Associates, Inc., 2019, pp. 8024–8035.

[31] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del R'ıo, M. Wiebe, P. Peterson, P. G'erard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T. E. Oliphant, Array programming with NumPy, Nature 585 (7825) (2020) 357–362. doi:10.1038/s41586-020-2649-2. URL https://doi.org/10.1038/s41586-020-2649-2

[32] J. D. Hunter, Matplotlib: A 2D graphics environment, Comput. Sci. Eng. 9 (3) (2007) 90–95. doi:10.1109/MCSE.2007.55.

[33] N. Dropka, M. Holena, Application of artificial neural networks in crystal growth of electronic and opto-electronic materials, Crystals 10 (8) (2020) 663. doi:https://doi.org/10.3390/cryst10080663.

[34] W. S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, The bulletin of mathematical biophysics 5 (4) (1943) 115–133. doi:https://doi.org/10.1007/BF02478259.

[35] Y. Tsunooka, N. Kokubo, G. Hatasa, S. Harada, M. Tagawa, T. Ujihara, High-speed prediction of computational fluid dynamics simulation in crystal growth, CrystEngComm 20 (41) (2018) 6546–6550. doi:10.1039/C8CE00977E.

[36] W. Yu, C. Zhu, Y. Tsunooka, W. Huang, Y. Dang, K. Kutsukake, S. Harada, M. Tagawa, T. Ujihara, Geometrical design of a crystal growth system guided by a machine learning algorithm, CrystEngComm 23 (14) (2021) 2695–2702. doi:https://doi.org/10.1039/D1CE00106J.

[37] M. Isono, S. Harada, K. Kutsukake, T. Yokoyama, M. Tagawa, T. Ujihara, Optimization of Flow Distribution by Topological Description and Machine Learning in Solution Growth of SiC, Adv. Theory Simul. 5 (9) (2022) 2200302. doi:https://doi.org/10.1002/adts.202200302.

[38] J. Snoek, H. Larochelle, R. P. Adams, Practical bayesian optimization of machine learning algorithms, in: F. Pereira, C. Burges, L. Bottou, K. Weinberger (Eds.), Advances in Neural Information Processing Systems, Vol. 25, Curran Associates, Inc., 2012.

[39] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. De Freitas, Taking the human out of the loop: A review of Bayesian optimization, Proceedings of the IEEE 104 (1) (2015) 148–175. doi:10.1109/JPROC.2015.2494218.

[40] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, G.-Z. Yang, XAI—Explainable artificial intelligence, Sci. Robot. 4 (37) (2019) eaay7120. doi:DOI:10.1126/scirobotics.aay7120.

[41] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, J. Zhu, Explainable ai: A brief survey on history, research areas, approaches and challenges, in: Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II 8, Springer, 2019, pp. 563–574. doi:https://doi.org/10.1007/978-3-030-32236-6_51.

[42] G. Vilone, L. Longo, Explainable artificial intelligence: a systematic review, arXiv preprint arXiv:2006.00093doi:https://doi.org/10.48550/arXiv.2006.00093.

[43] M. Sundararajan, A. Taly, Q. Yan, Axiomatic attribution for deep networks, in: International conference on machine learning, PMLR, 2017, pp. 3319–3328.
URL https://proceedings.mlr.press/v70/sundararajan17a.html

[44] M. T. Ribeiro, S. Singh, C. Guestrin, " why should I trust you?" Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1135–1144. doi:https://doi.org/10.1145/2939672.2939778.

[45] S. M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, Advances in neural information processing systems 30.
URL https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

[46] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: Visual explanations from deep networks via gradient-based localization, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 618–626.
URL https://openaccess.thecvf.com/content_iccv_2017/html/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.html

[47] D. Smilkov, N. Thorat, B. Kim, F. Viégas, M. Wattenberg, Smoothgrad: removing noise by adding noise, JarXiv preprint arXiv:1706.03825doi:https://doi.org/10.48550/arXiv.1706.03825.

[48] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707. doi:https://doi.org/10.1016/j.jcp.2018.10.045.

[49] M. Raissi, A. Yazdani, G. E. Karniadakis, Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations, Science 367 (6481) (2020) 1026–1030. doi:DOI:10.1126/science.aaw4741.

[50] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, J. M. Siskind, Automatic differentiation in machine learning: a survey, J. Mach. Learn. Res. 18 (2018) 1–43.
URL https://www.jmlr.org/papers/volume18/17-468/17-468.pdf

[51] S. Cai, Z. Mao, Z. Wang, M. Yin, G. E. Karniadakis, Physics-informed neural networks (pinns) for fluid mechanics: A review, Acta Mechanica Sinica 37 (12) (2021) 1727–1738.

[52] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, F. Piccialli, Scientific machine learning through physics–informed neural networks: Where we are and what's next, J. Sci. Comput. 92 (3) (2022) 88. doi:https://doi.org/10.1007/s10915-022-01939-z.

[53] S. Cai, Z. Wang, S. Wang, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks for heat transfer problems, J. Heat. Transfer 143 (6) (2021) 060801. doi:https://doi.org/10.1115/1.4050542.

[54] X. Jin, S. Cai, H. Li, G. E. Karniadakis, NSFnets (Navier-Stokes flow nets): Physics-informed neural networks for the incompressible navier-stokes equations, J. Comput. Phys. 426 (2021) 109951. doi:https://doi.org/10.1016/j.jcp.2020.109951.

[55] H. Eivazi, M. Tahani, P. Schlatter, R. Vinuesa, Physics-informed neural networks for solving Reynolds-averaged Navier–Stokes equations, Phys. Fluids 34 (7) (2022) 075117. doi:https://doi.org/10.1063/5.0095270.

[56] S. Cai, Z. Wang, F. Fuest, Y. J. Jeon, C. Gray, G. E. Karniadakis, Flow over an espresso cup: inferring 3-D velocity and pressure fields from tomographic background oriented schlieren via physics-informed neural networks, J. Fluid Mech. 915 (2021) A102. doi:https://doi.org/10.1017/jfm.2021.135.

[57] L. Sun, H. Gao, S. Pan, J.-X. Wang, Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data, Comput. Methods Appl. Mech. Eng. 361 (2020) 112732. doi:https://doi.org/10.1016/j.cma.2019.112732.

[58] S. Manavi, T. Becker, E. Fattahi, Enhanced surrogate modelling of heat conduction problems using physics-informed neural network framework, Int. Commun. Heat Mass Transf.

142 (2023) 106662. doi:https://doi.org/10.1016/j.icheatmasstransfer.2023.106662.

[59] J. C. Wong, C. Ooi, P.-H. Chiu, M. H. Dao, Improved surrogate modeling of fluid dynamics with physics-informed neural networks, arXiv preprint arXiv:2105.01838doi:https://doi.org/10.48550/arXiv.2105.01838.

[60] J. Oldenburg, F. Borowski, A. Öner, K.-P. Schmitz, M. Stiehm, Geometry aware physics informed neural network surrogate for solving Navier–Stokes equation (gapinn), Adv. Model. Simul. Eng. Sci. 9 (1) (2022) 8. doi:https://doi.org/10.1186/s40323-022-00221-z.

[61] H. Gao, L. Sun, J.-X. Wang, Phygeonet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain, J.Comput.Phys. 428 (2021) 110079. doi:https://doi.org/10.1016/j.jcp.2020.110079.

[62] L. Yang, X. Meng, G. E. Karniadakis, B-pinns: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data, J.Comput.Phys. 425 (2021) 109913. doi:https://doi.org/10.1016/j.jcp.2020.109913.

[63] L. Wang, T. Horiuchi, A. Sekimoto, Y. Okano, T. Ujihara, S. Dost, Numerical investigation of the effect of static magnetic field on the TSSG growth of SiC, J. Cryst. Growth 498 (2018) 140–147. doi:10.1021/ja00089a029.

[64] L. Wang, A. Sekimoto, Y. Takehara, Y. Okano, T. Ujihara, S. Dost, Optimal control of SiC crystal growth in the RF-TSSG system using Reinforcement Learning, Crystals 10 (9) (2020) 791. doi:https://doi.org/10.3390/cryst10090791.

[65] T. Horiuchi, . Wang, A. Sekimoto, Y. Okano, T. Ujihara, S. Dost, Adjoint-based sensitivity analysis for the optimal crucible temperature profile in the RF-Heating TSSG-SiC crystal growth process, J. Cryst. Growth 517 (2019) 59–63. doi:10.1016/j.jcrysgro.2019.04.001.

[66] T. Horiuchi, L. Wang, A. Sekimoto, Y. Okano, T. Yamamoto, T. Ujihara, S. Dost, The effect of crucible rotation and crucible size in top-seeded solution growth of single crystal silicon carbide, Cryst. Res. Tecnol. 54 (2019) 1900014(1)–1900014(5). doi:https://doi.org/10.1002/crat.201900014.

[67] P. Gresho, J. J. Derby, A finite element model for induction heating of a metal crucible, J. Cryst. Growth 85 (1-2) (1987) 40–48. doi:https://doi.org/10.1016/0022-0248(87)90202-8.

[68] J. J. Derby, L. J. Athertom, P. M. Gresho, An integrated process model for the growth of oxide crystals by the Czochralski method, J. Cryst. Growth 97 (3-4) (1989) 792–826. doi:https://doi.org/10.1016/0022-0248(89)90583-6.

[69] T. Munakata, I. Tanasawa, Study on silicon melt convection during the RF-FZ crystal growth process: II. numerical investigation, J. Cryst. Growth 206 (1-2) (1999) 27–36. doi:https://doi.org/10.1016/S0022-0248(99)00320-6.

[70] T. Yamamoto, Numerical investigation of surface tension gradient induced flow in various scales, PhD thesis.

[71] A. Mühlbauer, W. Erdmann, W. Keller, Electrodynamic convection in silicon floating zones, J. Cryst. Growth 64 (3) (1983) 529–545. doi:https://doi.org/10.1016/0022-0248(83)90338-X.

[72] J. Welty, G. L. Rorrer, D. G. Foster, Fundamentals of momentum, heat, and mass transfer, 6th edition, John Wiley & Sons, 2015.

[73] A. Muiznieks, A. Krauze, B. Nacke, Convective phenomena in large melts including magnetic field, J. Cryst. Growth 303 (2007) 211–220. doi:10.1016/j.jcrysgro.2006.12.051.

[74] J. Lefebure, J.-M. Dedulle, T. Ouisse, D. Chaussende, Modeling of the growth rate during top seeded solution growth of sic using pure silicon as a solvent, Cryst. Growth Des. 12 (2) (2012) 909–913. doi:https://doi.org/10.1021/cg201343w.

[75] F. Durand, J. Duby, Carbon solubility in solid and liquid silicon—a review with reference to eutectic equilibrium, Journal of Phase Equilibria 20 (1) (1999) 61–63. doi:https://doi.org/10.1361/105497199770335956.

[76] J. Lefebure, J. M. Deddule, T. Ouisse, D. Chaussende, Modeling of the growth rate during top seeded solution growth of sic using pure silicon as a solvent, Cryst. Growth. Des 12 (2011) 909–913. doi:10.1021/cg201343w.

[77] OpenFOAM, Openfoam, https://openfoam.org/ (Foundation).

[78] D.-X. Du, T. Munakata, Temperature distribution in an inductively heated CZ crucible, J. Cryst. Growth 283 (3-4) (2005) 563–575. doi:https://doi.org/10.1016/j.jcrysgro.2005.06.048.

[79] F. M. Faubert, G. S. Springer, Measurement of the thermal conductivity of helium up to 2100 k by the column method, J. Chem. Phys. 58 (10) (1973) 4080–4083. doi:https://doi.org/10.1063/1.1678965.

[80] Gpyopt, http://sheffieldml.github.io/GPyOpt/.

[81] D. Mochihashi, S. Oba, ガウス過程と機械学習, 講談社, 2019.

[82] J. D. Head, M. C. Zerner, A Broyden—Fletcher—Goldfarb—Shanno optimization procedure for molecular geometries, Chem. Phys. Lett. 122 (3) (1985) 264–270. doi:https://doi.org/10.1016/0009-2614(85)80574-1.

[83] D. J. MacKay, et al., Bayesian nonlinear modeling for the prediction competition, ASHRAE transactions 100 (2) (1994) 1053–1062. doi:https://doi.org/10.1007/978-94-015-8729-7_18.

[84] C. M. Bishop, N. M. Nasrabadi, Pattern recognition and machine learning, Vol. 4, Springer, 2006.

[85] R. Ghritli, Y. Okano, Y. Inatomi, A. Sekimoto, S. Dost, Estimation of the diffusion coefficient of gasb in insb melt using bayesian optimization and the iss experimental results, J. Cryst. Growth 573 (2021) 126280. doi:https://doi.org/10.1016/j.jcrysgro.2021.126280.

[86] R. Ghritli, Y. Okano, Y. Inatomi, S. Dost, A numerical study for the growth of ingasb crystals with a flatter interface by vertical gradient freezing under normal gravity and utilizing bayesian optimization, J. Chem. Eng. Japan 56 (1) (2023) 2222757. doi:https://doi.org/10.1080/00219592.2023.2222757.

[87] L. Wang, R. Isobe, Y. Okano, M. Kino-Oka, H. Matsuda, T. Tokura, Bayesian optimization for hydrodynamic characterization of a cylindrical orbitally shaken bioreactor with a bump at the bottom, J. Chem. Eng. Japan 54 (9) (2021) 493–499. doi:https://doi.org/10.1252/jcej.21we047.

[88] L. Wang, Y. Okano, I. Horiguchi, J. Zhang, Bayesian optimization for favourable suspension culture of orbitally shaken bioreactors with a hollow cylinder wall, J. Chem. Eng. Japan 56 (1) (2023) 2247046. doi:https://doi.org/10.1080/00219592.2023.2247046.

[89] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.
URL https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[90] A. Clark, et al., Pillow (pil fork) documentation, readthedocs.
URL https://buildmedia.readthedocs.org/media/pdf/pillow/latest/pillow.pdf

[91] A. F. Agarap, Deep learning using rectified linear units (relu), arXiv preprint arXiv:1803.08375doi:https://doi.org/10.48550/arXiv.1803.08375.

[92] I. Sutskever, J. Martens, G. Dahl, G. Hinton, On the importance of initialization and momentum in deep learning, in: S. Dasgupta, D. McAllester (Eds.), Proceedings of the 30th International Conference on Machine Learning, Vol. 28 of Proceedings of Machine Learning Research, PMLR, Atlanta, Georgia, USA, 2013, pp. 1139–1147.
URL https://proceedings.mlr.press/v28/sutskever13.html

[93] C. Zhang, Q. Cheng, Predicting melt-crystal interface position and shape during the manufacturing process of single crystal via explainable machine learning models, in: IOP Conference Series: Materials Science and Engineering, Vol. 1258, IOP Publishing, 2022, p. 012029. doi:10.1088/1757-899X/1258/1/012029.

[94] J. Liao, K. Misaki, M. Nakada, T. Uno, J. Sakamoto, Determination of significant 3d hemodynamic features for post-embolization recanalization in cerebral aneurysms through explainable artificial intelligencedoi:doi:10.20944/preprints202309.1632.v1.

[95] L. Liu, X. Liu, Y. Wang, Large-eddy simulation of melt turbulence in a 300-mm Cz–Si crystal growth, Int. J. Heat Mass Transf. 55 (1-3) (2012) 53–60. doi:https://doi.org/10.1016/j.ijheatmasstransfer.2011.08.038.

[96] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241. doi:https://doi.org/10.1007/978-3-319-24574-4_28.

[97] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: Proceedings of the thirteenth international conference on artificial intelligence and statistics, JMLR Workshop and Conference Proceedings, 2010, pp. 249–256.
URL https://proceedings.mlr.press/v9/glorot10a.html

[98] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980doi:https://doi.org/10.48550/arXiv.1412.6980.

[99] X. Zhao, Z. Gong, Y. Zhang, W. Yao, X. Chen, Physics-informed convolutional neural networks for temperature field prediction of heat source layout without labeled data, Engineering Applications of Artificial Intelligence 117 (2023) 105516. doi:https://doi.org/10.1016/j.engappai.2022.105516.

List of Tables

1.1 The physical properties of Si and SiC crystals[17] 11

2.1 The electromagnetic properties used in the TSSG system 21
2.2 The physical properties used in the TSSG system ([21, 78, 79]). 38
2.3 The parameter ranges of the Bayesian optimization 39

3.1 The filter size and stride of CNN 45

4.1 Physical properties utilized in this study. ([16] and [95]) 54
4.2 The hyper-parameters of PINNs. These values were optimized through a trial and error process to minimize relative L2 error against the validation dataset without clear criteria. w_i and N_i are the weights and the number of training points for loss function i, respectively. 65

List of Figures

1.1 The schematic view of the Czochralski method for Si crystals. 9
1.2 The schematic view of the sublimation method for SiC crystals 10
1.3 The schematic view of the TSSG method for SiC crystals 10
1.4 The example of Gaussian process regression. The true function is $y = \sin(x)$. The mean value of prediction is represented by the blue line. The confidence bound is obtained by $1.96\sigma_{\mathrm{GPR}}$, where σ_{GPR} is the standard deviation of the prediction. 13

2.1 The 3-dimensional (left) and cross-sectional view of the numerical configuration (right) of the TSSG method utilized in this study. The length is written in the right figure, where the unit is mm. 17
2.2 The outline of the Integrate Process Model (IPM). In step 1, Joule heat and Lorentz force densities are obtained and used in the following analysis. In step 2 temperature boundary conditions for the next step are obtained using the Joule heat density in step 1. In step 3, the detailed transport phenomena in the Si-C melt are calculated using the Joule heat and Lorentz force densities from step 1 and the temperature boundary conditions from step 2. 18
2.3 The region where the current (yellow) and water (light blue) run in a coil. . . . 22
2.4 The update history of the maximum value of the mean M^*, minimum value of the standard deviation σ^* of the crystal growth rates, and maximum value of the objective function $f_{\max}$. The objective of the optimization is maximizing f $(= M^* - \sigma^*)$ which means maximizing M^*, minimizing S^*. 33
2.5 The distribution of M^* and σ^* (left). The gray dots represent the whole parameter combination (125 calculations) and the colored dots represent the 24 calculations by the discrete Bayesian optimization. The star symbols indicate five more iterations from the discrete-optimal parameters using the continuous Bayesian optimization algorithm. The circles with cross indicate the initial dataset for the Bayesian optimization. 34
2.6 The crystal growth rates of the initial (black) and optimal states obtained from discrete (red) and continuous (blue) Bayesian optimizations. 35

2.7 The objective function f predicted by Gaussian process regression using full parametric search (a, c, e) and one after optimizations (b, d, f). The yellow and orange cross represent the optimal parameters obtained by the discrete and continuous optimizations, respectively. 39
2.8 The temperature (color) and flow velocity (arrow) distributions in the initial (a, c) and the optimal states (b, d) and the zoomed region near the seed crystal (a, b). 40
2.9 The concentration field (color) and streamline (lines with arrows) in the initial (left) and the optimal (right) states. 40
2.10 The radial velocity along the free surface in the initial (blue) and the optimal (red) states. 41
2.11 The velocity vectors representing the contributions of each driving forcing term under the optimal cusp magnetic field. (a) Lorentz force, (b) Surface tension, and (c) Buoyant force. 41
2.12 The velocity gradient along the seed surface near the seed edge (left) and the concentration field and the streamlines near the seed edge(right). $r_{\mathrm{sep.}}$ is the separation point of the melt flow ($\partial u_r/\partial z|_{r_{\mathrm{sep.}}} = 0$). 41

3.1 The concept of this study. Convolutional neural network learns the relationship $f(\boldsymbol{U})$ between the melt flow and the growth rate. Then, the relationship is observed by the sensitivity $\boldsymbol{S}$. 43
3.2 The architecture of CNN utilized in this study. 44
3.3 The learning curve of CNN. 46
3.4 The structures of the training and test data. Random selections were repeated 10 times, and prediction accuracy was calculated for each training and test dataset. 47
3.5 The relationship between the objective functions obtained from the numerical simulation and the CNN. The gray dots and red crosses indicate the training and test data, respectively. 48
3.6 The RMSE, MAE, and R^2 values against each test data (points) and the mean values of the errors (line). 48
3.7 The magnitudes of the sensitivity of the melt flow in the initial (a) and the optimal (b) states. 49
3.8 The components of the sensitivity, $\hat{S}_r$ (a, d), $\hat{S}_\theta$ (b, e), $\hat{S}_z$ (c, f) in the initial (a, b, c) and optimal (d, e, f) states. 50
3.9 The averaged magnitudes of the sensitivity of the melt flow of the top 10 % of the objective function. 51
3.10 The components of the averaged sensitivity of the melt flow of the top 10 % of the objective function. $\hat{S}_r$ (left), $\hat{S}_\theta$ (center), $\hat{S}_z$ (right). 51

3.11 The streamlines and the concentration distribution of the optimal state. The streamlines through the red line below the seed crystal are shown. The streamlines and the concentration field in this figure were obtained by using the time-averaged melt flow and the concentration fields. 51

4.1 The whole domain (left) of Cz-Si and the numerical domain (right), showing the half cross-sectional view of the whole domain. The input parameters of PINNs are shown in the right figure. T_l, T_h, and T_r are the temperatures at the black dots. h and h_s represent the melt heights at the center and the edge of the melt, respectively. 55

4.2 The architecture of PINNs developed in this study. The neural networks are trained separately. The input parameters (w_c, w_s) represent the rotational speed of the crucible and the seed, respectively, and were included for future development considering the effect of seed and crucible rotation. However, in this study, they were kept constant at 0 rpm and did not contribute to the learning and prediction process. 55

4.3 The distribution of singed distance function w_SDF at $h = 0$. 58

4.4 The distributions of the radial velocity u (a, d, g, j), the axial velocity w (b, e, h, k), and the temperature T (c, f, i, l) in the melt obtained from the numerical simulation (left) and PINNs (right) at the input parameter (T_l, T_r, T_h, p_{u_z}) = (1690 K, 1710 K, 1715 K, 0 m/s), h = 0.3 (a, b, c), 0.25 (d, e, f), 0.2 (g, h, i) and 0.15 m (j, k, l). 60

4.5 The radial velocity u (a) profile along the vertical line at r = 0.2 m, and the axial velocity w (b) and temperature T (c) profiles along the horizontal center line under the same input parameter as Fig. 4.4. 61

4.6 The histograms of the relative L2 error of u (a, d, g, j), w (b, e, h, k) and T (c, f, i, l) of different melt heights h = 0.3 m (a, b, c), 0.25 m (d, e, f), 0.2 m (g, h, i), 0.15 m (j, k, l). m indicates the mode value. 62

4.7 The histograms of the relative L2 error of u (a, d, g, j), w (b, e, h, k) and T (c, f, i, l) of different melt heights h = 0.3 m (a, b, c), 0.25 m (d, e, f), 0.2 m (g, h, i), 0.15 m (j, k, l). m indicates the mode value.. 63

Acknowledgements

Firstly, I would like to thank Professor Yasunori Okano. I appreciate his careful guidance, supervision, and encouragement. Prof. Okano always carefully revised my applications, abstracts for conferences, and papers, and gave me a lot of opportunities to attend international and domestic conferences. Moreover, he always gave proper advice when I struggled. Thanks to that, I could get research results and write this doctoral thesis. I can't write enough here to express my thanks for his help. I was very grateful and fortunate to have been provided with a platform and equipment to do research at my will.
I would also acknowledge the collaboration researcher, Prof. Sadik Dost at the University of Victoria. All of my papers were carefully modified by Prof. Dost. Thanks to his careful check, my paper could be accepted.
I am grateful to the researchers, Prof. Toru Ujihara and visiting assistant Prof. Kentaro Kutsukake at Nagoya University. They gave me a chance to conduct experiments with the TSSG method for SiC crystals in addition to advice for my research.
This research was financially supported by JSPS KAKENHI [Grant Numbers JP18H03839 and JP22KJ2057].
I appreciate Ms. Yuko Matsuzuki, secretary of Okano laboratory. She helped with many things about applications.
I appreciate Assistant Prof. Kento Kasahara at Osaka University. Dr. Kasahara kindly taught about network systems for utilizing computers from the outside. Thanks to him, we could use the computers from our home, which promotes our research.
I also appreciate Assistant Prof. Sho Yamaguchi for his kind advice for my career. He was kind enough to share my concerns about my career.
I appreciate Prof. Isao Murata of the Graduate School of Engineering, Osaka University. Even though he was my brother's supervisor, he was kind enough to take the time to advise me about the course, offering both practical and emotional support.
Mr. Nakai and Mr. Nakamura of Concurrent Systems helped maintenance of computers. They gave us a great deal of knowledge and assistance regarding the calculator beyond the range of their support. Thanks to that, I was able to use the computer without any worries.
I also appreciate Associate Professor Atsushi Sekimoto of Okayama University. Previously he was an Assistant professor in the Okano lab. at Osaka University. He also gave me sincere and

kind guidance about my research. In addition to that, I was invited to many dinners. Thanks to him, I could enjoy the research life of my master course.
I would also like to thank Nao Ishida and Rachid Mohamed Ghritli, graduates of Okano Lab. Their kindness encouraged me many times during my three years in the doctoral course. I would also like to thank Atsushi Onizuka and Junya Takayoshi, my "kouhai"s in the Okano lab. They were always my talkers.
My brother, Ryohei Takehara was always concerned about me and helped me with many things, including job hunting. My family also always cared about me.
I express my gratitude for the following: Drinks (Monster, Red Bull, Zone, Dr. Pepper), Keyboards (Keychron, Topre), Chair (okamura), Trackball mouses (logicool), "連続テレビ小説らんまん" (NHK) since they enhance the motivation for my research, and Ukulele (kiwaya) for relaxation. I also appreciate Assistant Professor Ban Takahiko.
Finally, I appreciate all the support I received from everyone.

List of Publications

1. Y. Takehara, A. Sekimoto, Y. Okano, T. Ujihara and S. Dost, "Bayesian optimization for a high- and uniform-crystal growth rate in the top-seeded solution growth process of silicon carbide under applied magnetic field and seed rotation", Journal of Crystal Growth, **532** (2020), 125437.

2. L. Wang, Y. Takehara, A. Sekimoto, Y. Okano, T. Ujihara and S. Dost, "Numerical Study of Three-Dimensional Melt Flows during the TSSG Process of SiC Crystal for the Influence of Input Parameters of RF-Coils and an External Rotating Magnetic Field", crystals, **10**(2) (2020), 111.

3. L. Wang, A. Sekimoto, Y. Takehara, Y. Okano, T. Ujihara and S. Dost, "Optimal Control of SiC Crystal Growth in the RF-TSSG System Using Reinforcement Learning", crystals, **10**(9) (2020), 791.

4. Y. Takehara, A. Sekimoto, Y. Okano, T. Ujihara and S. Dost, "Explainable machine learning for the analysis of transport phenomena in top-seeded solution growth of SiC single crystal", Journal of Thermal Science and Technology, **16** (2021), JTST0009.

5. Y. Takehara, Y. Okano and S. Dost, "Fast Prediction of Transport Structures in the Melt by Physics Informed Neural Networks during 'VMCz' Crystal Growth of Silicon", Journal of Chemical Engineering of Japan, **56** (2023), 2236656.

List of International conference

Oral presentation

1. Y. Takehara, A. Sekimoto, Y. Okano, T. Ujihara and S. Dost, "The Bayesian optimization for controlling transport phenomena in Top-Seeded Solution Growth of SiC crystal growth", The 2nd Pacific Rim Thermal Engineering Conference (PRTEC2019), Hyatt Regency Maui Resort & Spa, Maui, Hawaii, USA, December 13-17, 2019.

2. Y. Takehara, L.Wang, A. Sekimoto, Y. Okano, T. Ujihara and S. Dost, "Control of transport structures in TSSG of SiC by using a stack of insulator rings with different sizes and thermal conductivities", 14th World Congress in Computational Mechanics (WCCM) ECCOMAS Congress 2020, Paris, France, July 19-24, 2020 (online).

3. Y. Takehara, Y. Okano and S. Dost, "Control of melt temperature distribution in the Top-Seeded Solution Growth method of SiC crystals using insulators of different thermal conductivities", The 8th Asian Conference on Crystal Growth and Crystal Technology (CGCT-8), March 1-4, 2021 (online).

poster presentation

1. Y. Takehara and Y. Okano, "The Bayesian Optimization for a High- and Uniform-Crystal Growth Rate in the Top-Seeded Solution Growth Method Using Different Objective Functions", International Conference on Materials and Systems for Sustainability (ICMaSS 2021), Nagoya University, Nagoya, Japan, November 4-6, 2021 (online).

2. Y. Takehara and Y. Okano, "Development of the production process for high quality power semiconductor SiC crystals for an efficient energy system using data science", The Sixth International Symposium on Innovative Materials and Processes in Energy Systems (IMPRESS 2022), Universitat de Barcelona, Barcelona, Spain, October 25-28, 2022.

List of domestic conference

Oral presentation

1. 竹原悠人, 岡野泰則, "熱伝導率が異なる複数の断熱材を用いたTSSG法によるSiC結晶成長時の熱輸送の制御", 第68回応用物理学会春季学術講演会, 3月16-19日, 2021 (オンライン).

2. 竹原悠人, 岡野泰則, 異なる熱伝導率を有する断熱材を用いたTSSG法によるSiC結晶作製時の熱輸送の最適化", 第58回日本伝熱シンポジウム, 5月25-27日, 2021 (オンライン).

3. 竹原悠人, 岡野泰則, "Physics Informed Neural Networksを用いたSi結晶成長における熱流体の高速解析", 第35回計算力学講演会 (CMD2022), 11月16-18日, 2022 (オンライン).

4. 竹原悠人, 岡野泰則, "PINNs (Physics Informed Neural Networks) を用いた熱流体解析の高速化", 第83回応用物理学会秋季学術講演会, 東北大学川内北キャンパス+オンライン, 9月20-23日, 2022 (注目講演).

5. 竹原悠人, 岡野泰則, "PINNs (Physics Informed Neural Networks) を用いたシリコンバルク単結晶成長デジタルツインの構築", 化学工学会第53回秋季大会, 信州大学長野キャンパス+オンライン, 9月14-16日, 2022 (注目講演).

6. 竹原悠人, 岡野泰則, "TSSG法による炭化ケイ素結晶成長プロセスにおけるPhysics Informed Nerual Networkを用いた物性値の逆推定", 化学工学会第87年大会, 神戸大学鶴甲第1キャンパス+オンライン, 3月16-18日, 2022.

Poster presentation

1. 竹原悠人, 関本敦, 岡野泰則, 宇治原徹, Sadik Dost, "Top-Seeded Solution Growth 法によるSiC結晶成長のベイズ最適化を用いた対流制御", 第48回結晶成長国内会議 (JCCG-48), 大阪大学吹田キャンパス, 10月30日-11月-1日, 2019.

2. 竹原悠人, 岡野泰則, 宇治原徹, Sadik Dost, "説明可能機械学習を用いたTSSG-SiC結晶作製時の移動現象解析", 第49回結晶成長国内会議 (JCCG-49), 9月 9-11日, 2020 (オンライン).

Awards

1. 第21回化学工学会学生発表会優秀賞
2. 応用物理学会関西支部2021年度第1回講演会ポスター最優秀賞
3. 第10回（2021年度）応用物理学会関西支部関西奨励賞