

Article

Optimal Control of SiC Crystal Growth in the RF-TSSG System Using Reinforcement Learning

Lei Wang ¹, Atsushi Sekimoto ^{1,*} , Yuto Takehara ¹, Yasunori Okano ¹, Toru Ujihara ² and Sadik Dost ³

¹ Department of Materials Engineering Science, Osaka University, Toyonaka 560-8531, Osaka, Japan; wanglei8657@cheng.es.osaka-u.ac.jp (L.W.); takehara.yuto@cheng.es.osaka-u.ac.jp (Y.T.); okano@cheng.es.osaka-u.ac.jp (Y.O.)

² Department of Materials Science, Nagoya University, Chikusa-ku 464-8603, Nagoya, Japan; ujihara@nagoya-u.jp

³ Crystal Growth Laboratory, University of Victoria, Victoria, BC V8W 3P6, Canada; sdost@uvic.ca

* Correspondence: asekimoto@cheng.es.osaka-u.ac.jp

Received: 7 August 2020; Accepted: 3 September 2020; Published: 7 September 2020



Abstract: We have developed a reinforcement learning (RL) model to control the melt flow in the radio frequency (RF) top-seeded solution growth (TSSG) process for growing more uniform SiC crystals with a higher growth rate. In the study, the electromagnetic field (EM) strength is controlled by the RL model to weaken the influence of Marangoni convection. The RL model is trained through a two-dimensional (2D) numerical simulation of the TSSG process. As a result, the growth rate under the control of the RL model is improved significantly. The optimized RF-coil parameters based on the control strategy for the 2D melt flow are used in a three-dimensional (3D) numerical simulation for model validation, which predicts a higher and more uniform growth rate. It is shown that the present RL model can significantly reduce the development cost and offers a useful means of finding the optimal RF-coil parameters.

Keywords: SiC crystal growth; TSSG method; flow control

1. Introduction

Silicon carbide (SiC) crystal is a promising semiconductor material of power devices and the radio-frequency (RF) top-seeded solution growth (TSSG) method that has been used to produce high-quality SiC crystals. However, the unstable growth and slow growth rates of SiC crystals prevent the utilization of the TSSG method to grow large single crystals in industrial setups. In the RF-TSSG process, maintaining the uniform growth rate along the seed within a certain range can stabilize crystal morphology, and the melt flow during the crystal growth plays an important role in the change of growth rate. To improve the quality of SiC crystals, we have conducted numerical simulations to shed light on the phenomena governing the SiC growth in this process [1–5]. Most studies have carried out simulations of the melt flow using a particular condition, that is, under certain control parameters of the TSSG system, such as a fixed input power for the RF-coil, prescribed boundary conditions, the magnetic field strength, and seed rotation rate.

It is standard for a design parameter in a crystal growth system to be optimized one by one through numerical simulations. However, the design parameters involved are many and may have a combined effect on the melt flow. Thus, it would be inefficient and expensive to optimize the whole system by means of numerical simulations alone. For instance, in the TSSG process of SiC crystals with the control of static magnetic fields, Takehara et al. [6] applied a Bayesian optimization to determine the optimal set-up of a cusp magnetic field and seed rotation for high- and uniform-crystal

growth rates. However, with the increase of control parameters to be taken into account, this approach becomes costly because the algorithm requires a large amount of data set. To this end, a fast and proper design of the TSSG system for an optimal control of the growth rate is required. To the best of our knowledge, the optimization and control models developed previously were only for the Czochralski growth process. The first-principle is used to establish a prediction model that relates input design parameters to output crystal parameters [7–11]. Some studies have focused on developing a model that represent the relationship between the changes of the crystal radius and the crystal slope angle at the meniscus section [12–14]. A proportional integral derivative (PID) design has been used to control the growth process based on the developed models [15,16]. Basically, all those models introduced earlier are based on a series of differential equations but do not fully involve the physics of fluid mechanics. In a more complicated crystal growth process, however, the improvement of the accuracy of the models remains a challenge, and PID control still has deficiencies in dealing with nonlinear and high-dimensional problems.

In recent years, the application of machine learning has received a notable attention in fluid control problems. This is because this approach allows to examine completely different cases and gets results faster. The combination of appropriate machine learning tools and fluid mechanics knowledge can be used to directly optimize the control strategy, to reduce or even eliminate the artificial control modeling and design, and to change the traditional approach. In the field of crystal growth, Dropka et al. [17] designed and trained artificial neural networks (ANNs) in directional solidification of silicon to identify the relation and the optimum combination of magnetic fields and growth parameters through the data of 2D CFD simulation. As expected, the accuracy of the model naturally depends on the amount and accuracy of the available data set in a given system. On the other hand, Reinforcement Learning (RL), which is one of the machine learning tools recently widely utilized in the field of optimal control of fluid flows [18–21], can automatically discover the optimal control strategies without any prior knowledge. This approach presents itself as a powerful tool in general in modelling, and would naturally be beneficial for modelling crystal growth techniques. To this end, in the present study, we introduce the RL technique to the TSSG system to control the melt flow during the SiC crystal growth process. In the RL model, an ‘agent’ tries to learn the policy to maximize a ‘reward’ function through an certain ‘environment’. The environment can be any stochastic process. For example, the numerical simulation of the SiC melt flow can be taken as the ‘environment’ in this study. The agent first obtains the state of the simulation (environment). Then the agent performs ‘actions’ to affect the time evolution of the simulating melt flow (environment). After receiving the reward from the state of the environment controlled by actions, the agent completes one control loop. In the TSSG process, maximizing both the growth rate and its uniformity simultaneously is essential for growing high quality crystals. This is the main objective of the present study.

2. Methodology

2.1. Computational Fluid Mechanics Model

Figure 1 shows the whole computational domain, Dimensions in the figure are in mm. The present simulation of SiC crystal growth in the RF-TSSG process is based on an Integrate Process Model (IPM) developed by Gresho and Derby et al. [22,23]. The IPM solves the process in three steps: (i) the coil-induced electromagnetic field; (ii) heat generation and heat transfer in the furnace; and (iii) the melt flow in the crucible. A 2D numerical simulation for the melt flow was taken as an interactive environment in the reinforcement learning process as explained later.

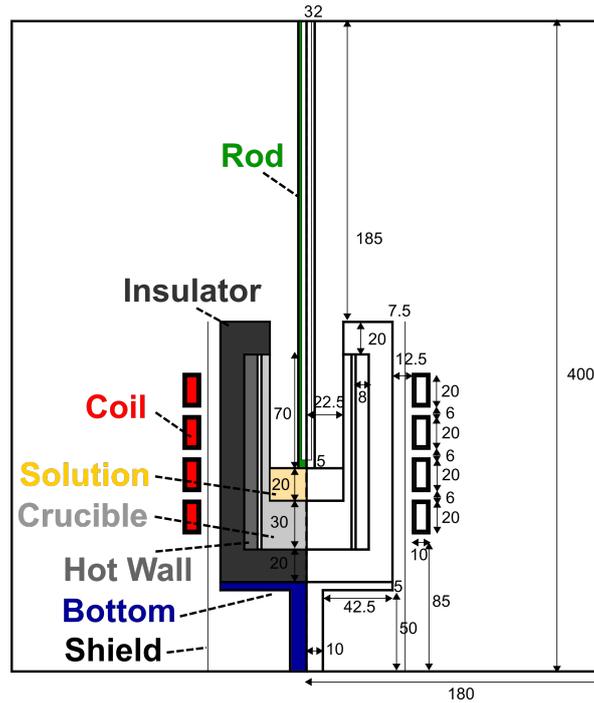


Figure 1. Schematic view of the computational domain used for the TSSG system. Dimensions in the figure are in mm.

2.1.1. Electromagnetic Field

In RF-heating, the frequency of the electric current in the coil is too high to resolve the time resolution by numerical simulations. Thus, IPM uses period-averages for calculating the densities of the Lorentz force and heat generation. The Lorentz force and heat generation are calculated by:

$$F_E = \frac{\omega}{2\pi} \int_0^{\frac{2\pi}{\omega}} J \times B dt$$

$$= \left(\frac{\sigma_e \omega}{2r^2} \left(C \frac{\partial S}{\partial r} - S \frac{\partial C}{\partial r} \right), 0, \frac{\sigma_e \omega}{2r^2} \left(C \frac{\partial S}{\partial z} - S \frac{\partial C}{\partial z} \right) \right). \quad (1)$$

$$Q = \frac{\omega}{2\pi} \int_0^{\frac{2\pi}{\omega}} \frac{J_0^2}{\sigma_e} dt = \frac{\sigma_e \omega^2}{2r^2} (C^2 + S^2), \quad (2)$$

where ω is the frequency of electric current in the coil, σ_e is the electric conductivity, and J_0 the peak current in the coil, C and S are time-independent in-phase and out-phase amplitudes of the magnetic stream function.

2.1.2. Heat Transfer in the Furnace

The steady-state conductive and radiative heat transfers are considered for computing the temperature field in the furnace. The associated governing equations are given by:

$$\nabla \cdot (k \nabla T) + Q = 0. \quad (3)$$

$$J_i - (1 - \epsilon_i) \sum_j F_{ij} J_j = \epsilon_i E_{b,i}$$

$$\frac{q_i}{A_i} = \epsilon_i (E_{b,i} - \sum_j F_{ij} J_j) \quad (4)$$

where k is the thermal conductivity, T temperature, Q the Joule heat generation density, J_i the radiosity, ϵ_i emissivity, F_{ij} the view factor, $E_{b,i}$ the emissive power of a black body and q_i / A_i the heat flux.

2.1.3. The Melt Flow

The governing equations of the melt flow are the well-known continuity, momentum balance, energy balance, and mass transport equations that take the following forms,

$$\nabla \cdot \mathbf{u} = 0, \quad (5)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{u} + a \frac{\mathbf{F}_E}{\rho} - \mathbf{g} \beta (T - T_{ref}), \quad (6)$$

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \alpha \nabla^2 T + \frac{Q}{\rho C_p}, \quad (7)$$

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \nabla c = D \nabla^2 c, \quad (8)$$

where \mathbf{u} is the flow velocity vector, ρ density, p pressure, μ the kinematic viscosity, a represents the control value that will be explained later, \mathbf{g} the gravitational acceleration, β the thermal expansion coefficient, T_{ref} reference temperature, α thermal diffusivity, C_p the specific heat, D the diffusion coefficient, and c the carbon concentration. The initial and boundary conditions of the melt flow are obtained from the results of the heat transfer simulation in the furnace, and the overall simulation is actually coupled with the computation of electromagnetic field, heat transfer in the furnace, which are described in Sections 2.1.1, 2.1.2 and 2.1.3. The physical properties and boundary conditions are referred to [1].

2.2. Reinforcement Learning

Reinforcement learning involves an agent built by ANNs interacting with an environment. In this study, numerical simulation is regarded as the interactive environment through three steps: the agent makes an observation of the state s_t (an array of fluid flow variable obtained from the simulation), imposes the action a_t on the simulation, and computes a reward r_t from the controlled simulation. Here, t is the discrete time step when the interaction takes place. The optimal control problem is aiming to learn an optimal policy that maximizes the expected cumulative reward.

$$R_{max} = \max \mathbb{E} \left[\sum_{t=0}^{\tau} \gamma^t r_t | a_t = \pi_{\Theta}(s_t) \right] \quad (9)$$

where γ is a discount factor, π_{Θ} is the policy function described by ANN (Θ is the weights).

The current RL model training is based on an episode, which means that the model will learn active control strategy in a limited time before analyzing the obtained results and resume learning with a new episode. The sketch in Figure 2 presents the one episode of the learning process, interacted with the simulation of fully developed melt flow (started from 700 s [1]) in the TSSG process. The RL agent interacts with the 2D melt flow simulation via a state inquiry, and an action decision is made at every $T = 0.25$ s during the simulation, and one episode training lasts 5.0 s. The states in the current simulation are the supersaturation near the seed with 50 sample points, which are shown in Figure 2. Wang et al. [5] reported that the melt flow can be controlled by the RF-coil induced electromagnetic field, and the contribution of electromagnetic field is described as a source term in Equation (6). The value of a in the initial case without control is equal to 1.0. To simplify the calculation process, a reference case (at the RF-coil frequency 25 kHz and current 360 A) that performed in [1] is directly used as an initial case in the present study. According to the computational results of the effect of electromagnetic field in [5], applying the Lorentz force twice in magnitude (compared to the initial case) is detrimental to crystal growth. Therefore, the output action range a is limited between 0 and 2 when the model training is carried out. It should be noted that the parameters of the RF-coil control the electromagnetic field, and change the heat generation and temperature boundary conditions. For simplicity, we set the control value a representing the parameters of the RF-coil that ideally control electromagnetic

field strength, and the heat generation and temperature boundary conditions are assumed not to be changed in the training process, the solution of the changing of temperature boundary condition under control during the training process is discussed in later Section 3.2.

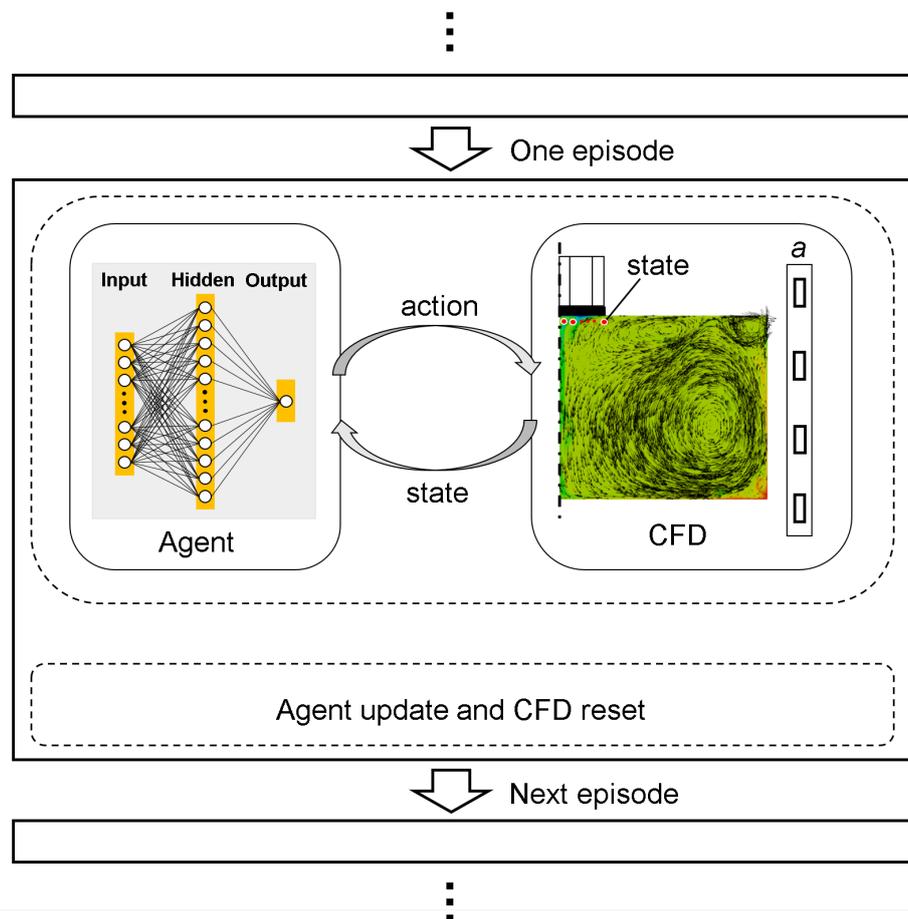


Figure 2. Sketch of the reinforcement learning process in simulation environments for one episode.

Due to the aim of improving the growth rate, the instantaneous reward function, r_t , consists of two contributions: the growth rate gradient (uniformity) and the growth rate,

$$r_t = -0.001 \langle |G_x| \rangle_T + \langle |v_g| \rangle_T. \quad (10)$$

$$v_g = -\frac{DM_{SiC}}{\rho_{SiC}} \mathbf{n} \cdot \nabla c \quad (11)$$

$$G_x = \frac{\partial v_g}{\partial x} \quad (12)$$

where T is the per action duration time, v_g is the growth rate, M_{SiC} is the molar weight of SiC, ρ_{SiC} is the crystal density, and G_x is the growth rate gradient along the seed interface. The growth rate of the reference case around the seed edge is extremely high compared with those on the other interface positions [1], and thus calculating G_x through the full seed radius makes the training process very difficult to converge. Thus, a partial growth interface (0–3 mm) is used in Equation (12). To balance the growth rate contribution, a factor of 0.001 is set in Equation (10) due to the large order difference between G_x and v_g . The agent consists of simple feed-forward ANNs with a hidden layer of

512 neurons. The discount factor γ is set as 0.95. Proximal Policy Optimization PPO algorithm [24] is used to update the agent, it belongs to the policy gradient class. The details of the policy gradient can be found in the article [25].

3. Results and Discussion

3.1. Growth Rate Improvement through Lorentz Force Control

By adapting the algorithm and hyperparameters mentioned in Section 2.2, we performed a robust RL training. The reward of every training episode is shown in Figure 3. As seen, the reward increases quickly after 40 episodes and it is converged after 100 episodes. Therefore, we consider that the policy is the close to the optimal one can be obtained after 100 episodes. The policy at the 120th episode was chosen to control the initial case over 50 s (which is 10 times longer than training time). Figure 4 shows the value of the action a in 50 s. The red point in the figure is the initial case ($a = 1.0$). For the case with control, it is clear that the Lorentz force was enhanced, the enhancement is about 1.5 times to 1.8 times the initial value ($a = 1.5$ to $a = 1.8$). Meanwhile, there is no obvious pattern of the action sequences, according to our previous study [3], the results may be due to the effect of the unstable flow generated by electromagnetic and interfacial forces.

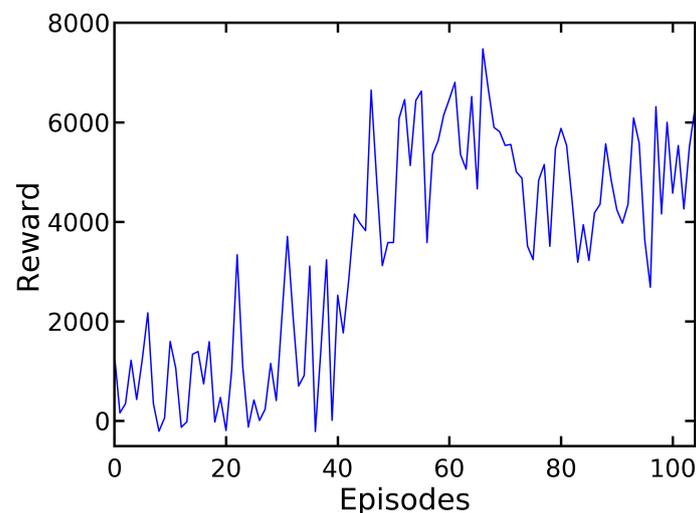


Figure 3. Illustration of the reward in the learning process.

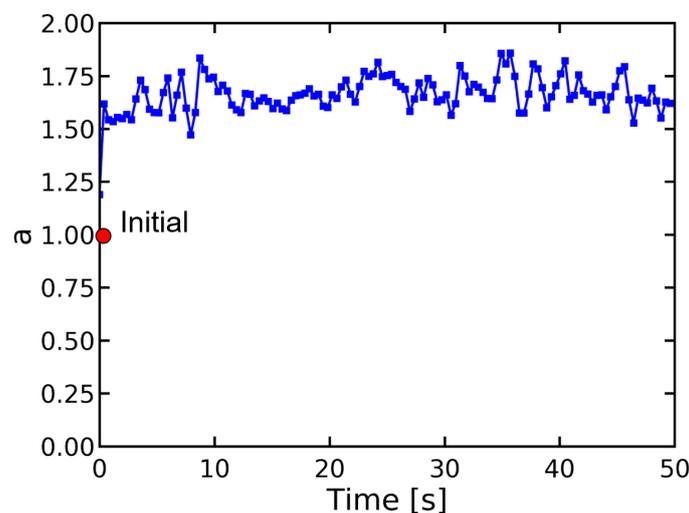


Figure 4. The control a of Lorentz force intensity for the 2D case (red point $a = 1$ is initial case).

The growth rate uniformity and its value for the initial and controlled cases are compared in Figure 5. Although the growth rate gradient in the case without and with control partially overlaps during the simulation time, the gradient in the case with control is generally smaller than that in the case without control. The smaller growth rate gradient indicates that the uniformity is improved. On the other hand, the value of growth rate in the case with control increased significantly (almost twice) compared with that in the case without control, as seen in Figure 5b. The results indicate that the present policy for the Lorentz force can increase the growth rate and improve its uniformity at the same time.

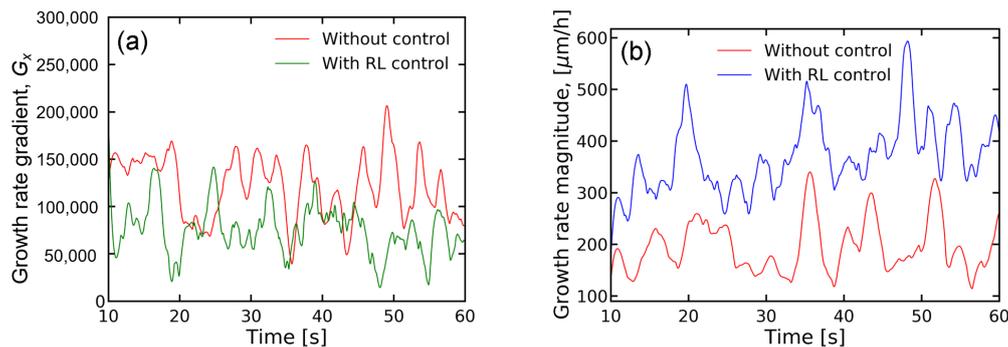


Figure 5. Time evolution of the growth rate gradient (a) and growth rate value (b).

Figure 6 shows the time-averaged temperature field in the melt without and with optimal control. The hottest part is located at the bottom corner of the crucible wall and the lowest temperature region is near the seed. In the case with control, the temperature field beneath the seed is flatter. The time-averaged flow velocity and supersaturation in the melt are presented in Figure 7, which more directly shows the comparison between the initial and controlled cases. In the computations, the supersaturation S is calculated by using $(c - c_{eq})/c_{eq}$, where c_{eq} is the equilibrium concentration [1]. The flow patterns of the two cases are very similar, which are characterized as the electromagnetic convection induced by Lorentz force in the main region of the melt and Marangoni convection along the free surface [4]. The directions of Marangoni convection are towards the crystal and crucible on the free surface due to the low temperature region in the vicinity of the crystal and upper corner of the crucible wall as seen in Figure 6. In the initial case, the Marangoni effect gives rise to a strong downward flow near the seed. This is the reason we predict non-uniform growth rate and a lower rate in this case. In the case with control, the downward flow is weakened significantly by the effect of the upward flow induced by the Lorentz force. Thus, as seen in Figure 7, we predict a more uniform supersaturation distribution below the seed.

The surface supersaturation along the crystal radius is quantitatively plotted in Figure 8. The predicted results of the cases without and with control are presented. We see that, after applying the by trained RL model, the supersaturation value and uniformity are apparently improved. It should be noticed that the growth rate non-uniformity in the 2D simulation is overestimated compared to that in the 3D simulation [1,3]. Thus, the result of optimal control in Figure 8 could be different in 3D simulations.

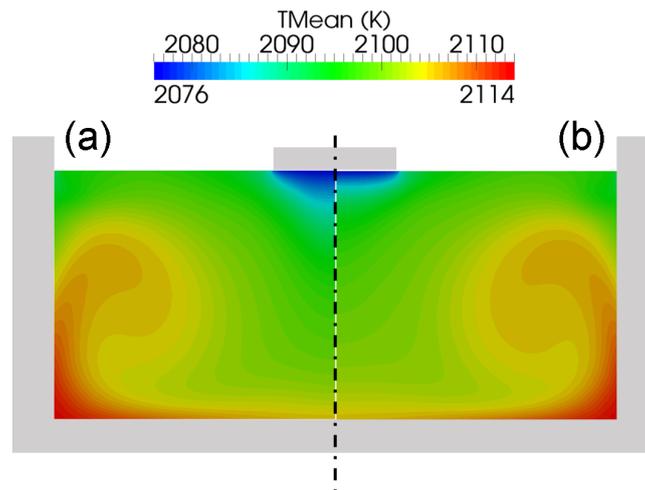


Figure 6. Time-averaged temperature field in the melt for the 2D case (a) without control and (b) with optimal control.

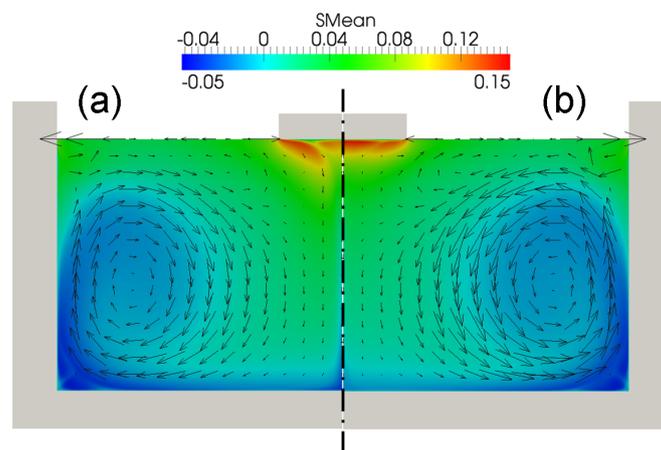


Figure 7. Time-averaged velocity vectors and supersaturation distribution in the melt for the 2D case (a) without control and (b) with optimal control.

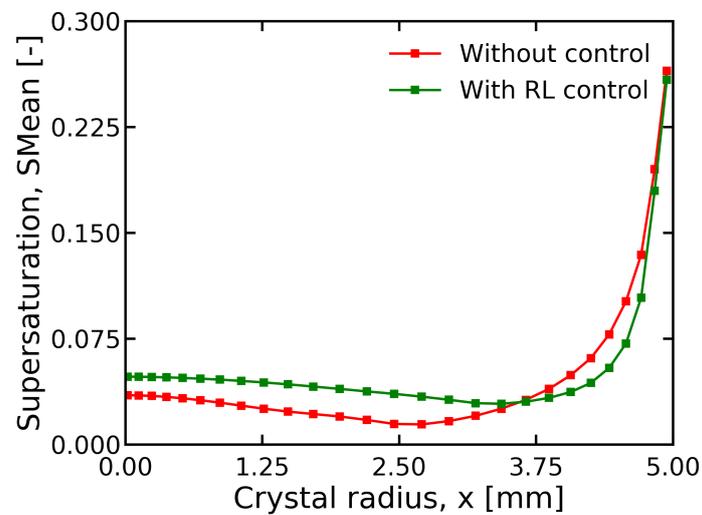


Figure 8. Time-averaged supersaturation profile along the crystal radius for the 2D case.

3.2. Discussion of the Optimal Control

In the previous section, RL is trained to optimally control the melt flow by directly adjusting the Lorentz force strength. If we consider a real case about changing the Lorentz force, the natural way is to change the frequency and current of the RF-coil. Figure 9 shows the relationship between the Lorentz force, heat generation density, and the frequency and current of the RF-coil. We can input different parameters with time according to the action of Figure 4. However, it arises some problems. For instance, adjusting the RF-coil not only changes the electromagnetic field but also changes the heat transfer in the furnace, which cannot ensure the accuracy of the control model. More importantly, the input parameters are usually fixed values before carrying out the crystal growth, which means the real-time control of the Lorentz force is quite difficult. Therefore, we implemented a compromised method to guarantee the same temperature boundary condition and a constantly optimized parameter. As in Figure 4, the optimized a fluctuates between 1.5 and 1.8, so that optimal electromagnetic field strength should in the range of 1.5 and 1.8 times the initial electromagnetic strength. Due to the reason that Marangoni induced downward flow is overestimated in 2D simulation, here, the minimum value of $a = 1.5$ is chosen as the optimized parameter in the 3D case. Figure 9 shows the frequency and current dependency of Lorentz force and heat generation. The value of heat generation should be close ($Q_{max} \approx 1.32 \times 10^7 \text{ W/m}^3$) for keeping the similar boundary condition between the initial and optimized case, and Lorentz force is 1.5 times the initial case ($Fe_{max} \approx 29,000 \text{ N/m}^3$). According to the calculation results in Figure 9, the optimized input parameters were then estimated:

- (1) The initial case (at 25 kHz the coil frequency, 360 A the coil current).
- (2) The optimized case (at 18 kHz the coil frequency, 390 A the coil current).

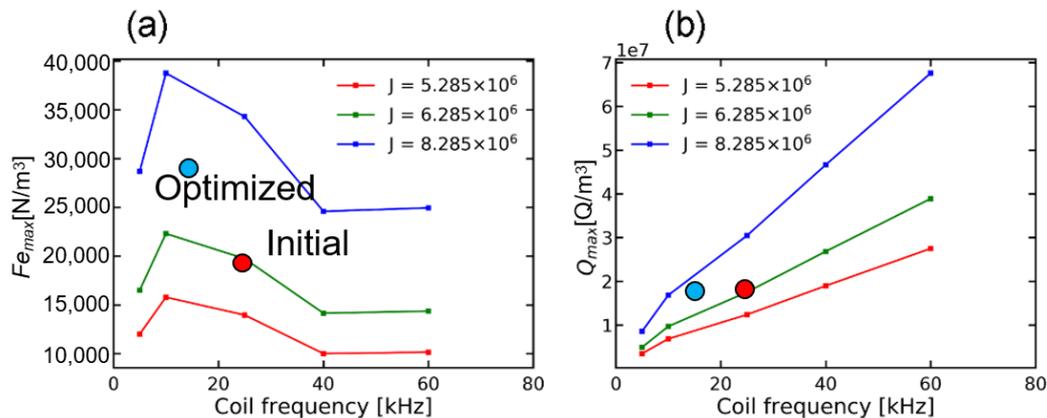


Figure 9. (a) The maximum value of Lorentz force density in the melt at various frequencies and current densities. (b) The maximum of the heat generation density in the whole calculation domain.

In order to validate the optimization method, the initial and optimized parameters are applied in the 3D system. The computed Lorentz force density and temperature along the crucible for the initial and optimized cases are shown in Figure 10. As seen from the comparison of those two cases, the maximum value of Lorentz force density in the optimized case is 28,200 N/m³, which is around 1.5 times that in the initial case 19,137 N/m³, this is in line with our proposed optimization plan ($a = 1.5$). The temperature distributions along the wall and seed in Figure 10b,d are similar, and the temperature difference between the seed and crucible walls is almost the same. Thus, the thermocapillary numbers of the initial and optimized cases are very close, $Re_{\sigma} \approx 4.8 \times 10^5$ [3], which means that at 18 kHz and 390 A, only the electromagnetic flow is enhanced without changing Marangoni convection significantly.

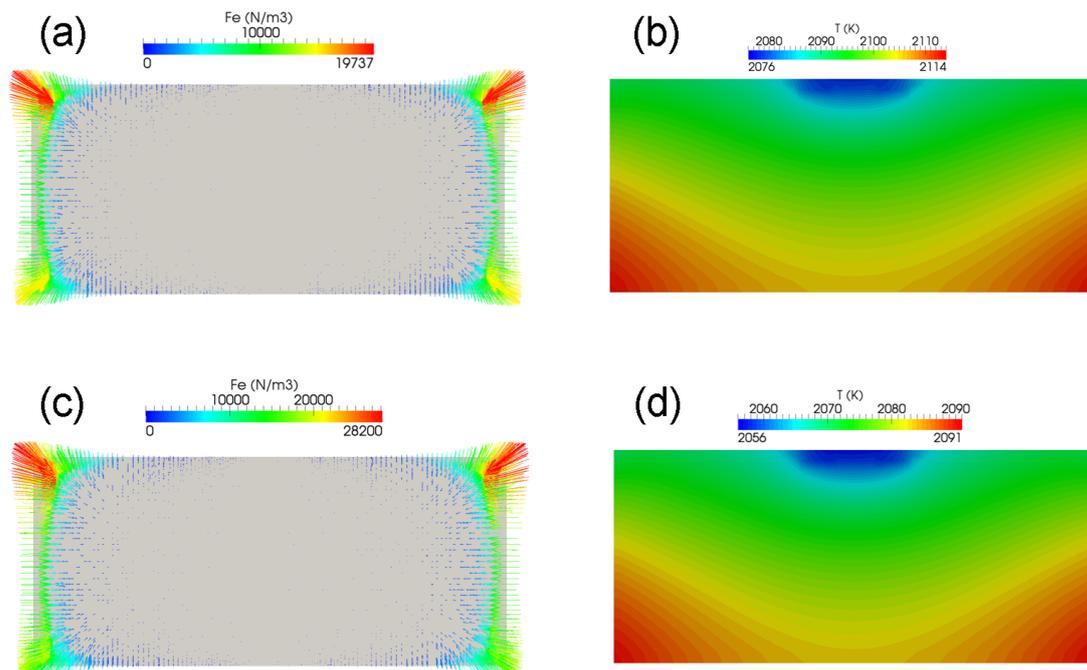


Figure 10. Lorentz force density (a,c) and temperature distribution (b,d) in the melt flow: (a,b) 25 kHz, 360 A; (c,d) 18 kHz, 390 A.

The time-averaged flow velocity and supersaturation in the 3D melt for the initial and optimized cases are presented in Figure 11. As seen, the Marangoni induced downward flow is significantly weakened in Figure 11b. In the enlarged view of the region below the seed, the flow velocity is apparently weaker in the optimized case. Usually, as the Lorentz force gets stronger the flow becomes more unstable. However, in this case, the flow is more stable for the stronger Lorentz force. This is due to the competition between the Marangoni downward flow and the electromagnetically induced upward flow. The results indicate that the effect of Marangoni flow is reduced, and the melt flow becomes more stable under the optimized RF-coil parameters.

The supersaturation along the seed diameter is plotted in Figure 12. It is clear that in the optimized case we predict more uniform supersaturation and higher supersaturation distribution on the crystal surface. These results show the validation of the present optimization model. It should be noticed that present optimized parameters are rough values that are close to the accurate optimal condition since they were chosen based on an estimation of the optimal control strategy discovered by the RL model in Section 3.1. It would be more accurate if we say that proposing EM/Re_σ^2 is in the range of 0.006 and 0.0072 is an optimization range for the selection of input variables. Here, EM/Re_σ^2 represents the ratio of electromagnetic and Marangoni forces [5].

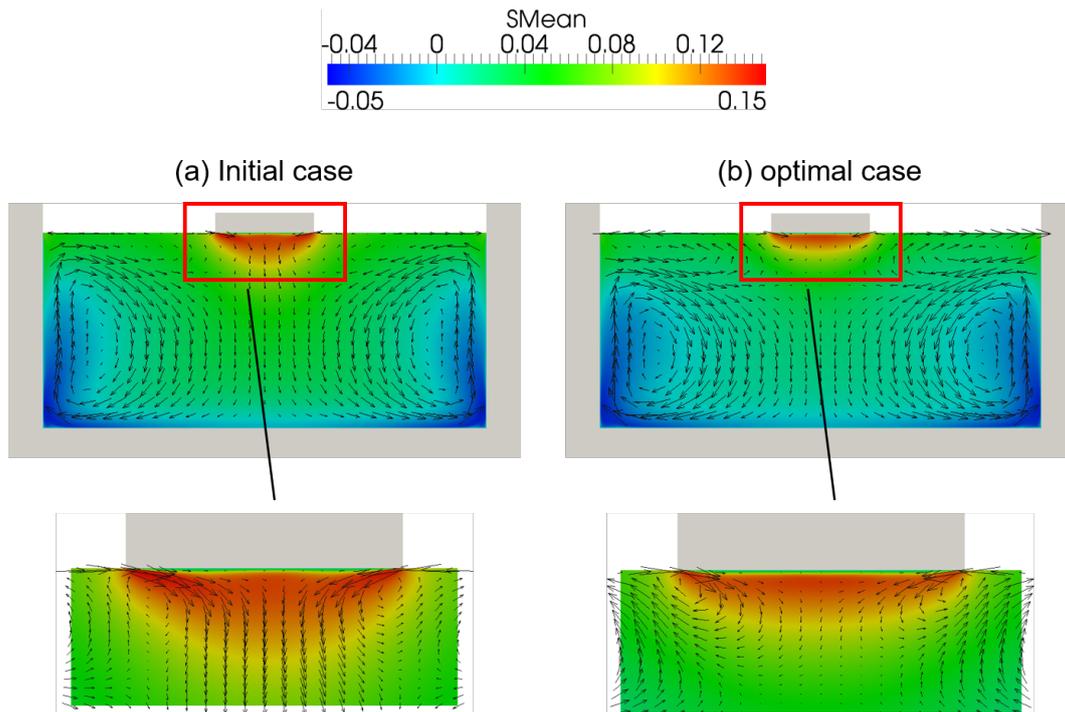


Figure 11. Time-averaged velocity vectors and supersaturation distribution in the melt for the 3D case (a) 25 kHz, 360 A and (b) 18 kHz, 390 A.

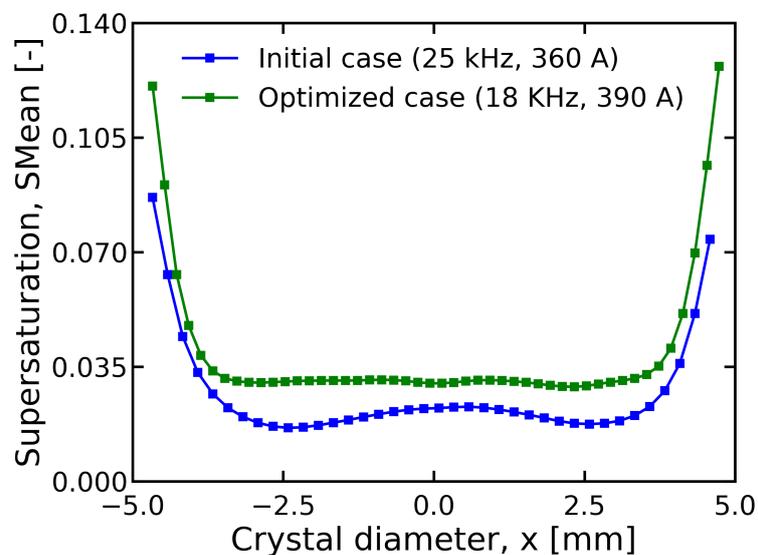


Figure 12. Time-averaged supersaturation profile along the crystal diameter for the 3D case.

4. Conclusions

A reinforcement learning model was developed to optimally control the top-seeded solution growth (TSSG) process. The model was trained by the 2D numerical simulation, and it improved the growth rate (supersaturation) of SiC crystal through the automatic control of electromagnetic field strength. The model accuracy was validated using a constant optimized parameter value (at 18 kHz and 390 A) to the 3D system according to the obtained optimal strategy range by the RL model in the 2D TSSG process. The selected optimized parameter enhanced the electromagnetic field without significantly changing the heat generation, and the supersaturation along the crystal diameter is also improved.

Based on the characteristics of the RL model that is capable of high-dimensional output, other parameters such as seed and crucible rotations, RF-coil position, and external magnetic fields, can be simultaneously optimized and the model needs to be validated by experiments in the future work, to test the potential of RL in the field of crystal growth.

Author Contributions: Conceptualization, L.W. and Y.O.; methodology, L.W.; investigation, L.W.; resources, A.S. and Y.O.; data curation, L.W.; writing—original draft preparation, L.W.; writing—review and editing, A.S. and Y.T. and Y.O. and T.U. and S.D.; supervision, A.S. and Y.O.; project administration, Y.O. and T.U.; funding acquisition, Y.O. and T.U. All authors have read and agreed to the published version of the manuscript.

Funding: The research work was financially supported by Grant-in-Aid for Scientific Research (A) (JSPS KAKENHI Grant Number JP18H03839 and JP20H00320) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Acknowledgments: The authors gratefully acknowledge the computational resources provided by the Research Institute for Information Technology at Kyushu University.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yamamoto, T.; Okano, Y.; Ujihara, T.; Dost, S. 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* **2017**, *470*, 75–88. [\[CrossRef\]](#)
2. Yamamoto, T.; Adkar, N.; Okano, Y.; Ujihara, T.; Dost, S. Numerical investigation of the transport phenomena occurring in the growth of SiC by the induction heating TSSG method. *J. Cryst. Growth* **2017**, *474*, 50–54. [\[CrossRef\]](#)
3. Wang, L.; Horiuchi, T.; Sekimoto, A.; Okano, Y.; Ujihara, T.; Dost, S. Three-dimensional numerical analysis of Marangoni convection occurring during the growth process of SiC by the RF-TSSG method. *J. Cryst. Growth* **2019**, *520*, 72–81. [\[CrossRef\]](#)
4. Wang, L.; Horiuchi, T.; Sekimoto, A.; Okano, Y.; Ujihara, T.; Dost, S. Numerical investigation of the effect of static magnetic field on the TSSG growth of SiC. *J. Cryst. Growth* **2018**, *498*, 140–147. [\[CrossRef\]](#)
5. Wang, L.; Takehara, Y.; Sekimoto, A.; Okano, Y.; Ujihara, T.; Dost, S. 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* **2020**, *10*, 111–123. [\[CrossRef\]](#)
6. Takehara, Y.; Sekimoto, A.; Okano, Y.; Ujihara, T.; Dost, S. 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. *J. Cryst. Growth* **2020**, *532*, 125437. [\[CrossRef\]](#)
7. Gevelber, M.; Stephanopoulos, G. Dynamics and control of the Czochralski process: I. Modelling and dynamic characterization. *J. Cryst. Growth* **1987**, *84*, 647–688. [\[CrossRef\]](#)
8. Ng, J.; Dubljevic, S. Optimal control of convection–diffusion process with time-varying spatial domain: Czochralski crystal growth. *J. Process Control* **2011**, *21*, 1361–1369. [\[CrossRef\]](#)
9. Ng, J.; Dubljevic, S. Optimal boundary control of a diffusion–convection–reaction PDE model with time-dependent spatial domain: Czochralski crystal growth process. *Chem. Eng. Sci.* **2012**, *67*, 111–119. [\[CrossRef\]](#)
10. Abdollahi, J.; Izadi, M.; Dubljevic, S. Model predictive temperature tracking in crystal growth processes. *Comput. Chem. Eng.* **2014**, *71*, 323–330. [\[CrossRef\]](#)
11. Zheng, Z.; Seto, T.; Kim, S.; Kano, M.; Fujiwara, T.; Mizuta, M.; Hasebe, S. A first-principle model of 300 mm Czochralski single-crystal Si production process for predicting crystal radius and crystal growth rate. *J. Cryst. Growth* **2018**, *492*, 105–113. [\[CrossRef\]](#)
12. Winkler, J.; Neubert, M.; Rudolph, J. Nonlinear model-based control of the Czochralski process I: Motivation, modeling and feedback controller design. *J. Cryst. Growth* **2010**, *312*, 1005–1018. [\[CrossRef\]](#)
13. Winkler, J.; Neubert, M.; Rudolph, J. Nonlinear model-based control of the Czochralski process II: Reconstruction of crystal radius and growth rate from the weighing signal. *J. Cryst. Growth* **2010**, *312*, 1019–1028. [\[CrossRef\]](#)
14. Neubert, M.; Winkler, J. Nonlinear model-based control of the Czochralski process III: Proper choice of manipulated variables and controller parameter scheduling. *J. Cryst. Growth* **2012**, *360*, 3–11. [\[CrossRef\]](#)

15. Gevelber, M. Dynamics and control of the Czochralski process III. Interface dynamics and control requirements. *J. Cryst. Growth* **1994**, *139*, 271–285. [[CrossRef](#)]
16. Gevelber, M. Dynamics and control of the Czochralski process IV. Control structure design for interface shape control and performance evaluation. *J. Cryst. Growth* **1994**, *138*, 286–301. [[CrossRef](#)]
17. Dropka, N.; Holena, M. Optimization of magnetically driven directional solidification of silicon using artificial neural networks and Gaussian process models. *J. Cryst. Growth* **2017**, *471*, 53–61. [[CrossRef](#)]
18. Novati, G.; Mahadevan, L.; Koumoutsakos, P. Controlled gliding and perching through deep-reinforcement-learning. *Phys. Rev. Fluids* **2019**, *4*, 093902. [[CrossRef](#)]
19. Rabault, J.; Kuchta, M.; Jensen, A.; Reglade, U.; Cerardi, N. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *J. Fluid Mech.* **2019**, *865*, 281–302. [[CrossRef](#)]
20. Viquerat, J.; Rabault, J.; Kuhnle, A.; Ghraieb, H.; Hachem, E. Direct shape optimization through deep reinforcement learning. *arXiv* **2019**, arXiv:1908.09885.
21. Fan, D.; Yang, L.; Triantafyllou, M.; Karniadakis, G. Reinforcement Learning for Active Flow Control in Experiments. *arXiv* **2020**, arXiv:2003.03419.
22. Gresho, P.; Derby, J. A finite element model for induction heating of a metal crucible. *J. Cryst. Growth* **1987**, *85*, 40–48. [[CrossRef](#)]
23. Derby, J.; Atherton, L.; Gresho, P. An integrated process model for the growth of oxide crystals by the Czochralski method. *J. Cryst. Growth* **1989**, *97*, 792–826. [[CrossRef](#)]
24. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal Policy Optimization Algorithms. *arXiv* **2017**, arXiv:1707.06347.
25. Duan, Y.; Chen, X.; Houthoofd, R.; Schulman, J.; Abbeel, P. Benchmarking deep reinforcement learning for continuous control. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 20–22 June 2016; pp. 1329–1338.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).