Grain Boundary Control in Colloidal Self-Assembly with Dynamic Programming

Xun Tang1, Yuguang Yang2, Michael A. Bevan2, and Martha A. Grover1,3

Abstract—We propose a Markov decision based dynamic programming method to manipulate the self-assembly of a quadrupole colloidal system for grain-boundary-free two-dimensional crystals. To construct the optimal control policy, we developed a Markov chain model, based on information extracted from a Langevin dynamics simulation model, which originated from a more complicated Brownian dynamics model. An infinite-horizon Markov decision process is defined, and the optimal control policy is solved with dynamic programming using policy iteration. Both the Markov chain Monte Carlo and the Langevin dynamics simulation results demonstrate that the control strategy is able to significantly accelerate the crystallization of a SiO2 colloidal self-assembly process for a grain-boundary-free, highly ordered crystal. Future work will focus on implementation of the control policy on the Brownian dynamics simulation and the experiments.

I. INTRODUCTION

Colloidal self-assembly refers to the process by which colloidal particles converge into an ordered structure with or without external interference. Periodically regular colloidal crystals, also called photonic crystals, have wide applications in various fields, including chemical and biological sensors [1], semiconductors [2], and photonic devices [3], due to their ability to manipulate the propagation of light. This unique feature of photonic crystals stimulated intense study on the techniques for its fabrication such as chemical stimuli-induced molecular interaction directed self-assembly, sedimentation, evaporation, adsorption, and external field directed self-assembly. Comprehensive literature reviews on the fabrication and applications of photonic crystals can be found in Ref. [4], [5], [6].

In spite of these many aforementioned techniques for photonic crystal fabrication, most self-assembling processes still suffer from defects, including vacancies and grain boundaries. Getting rid of these defects still remains one of the biggest challenges before the large-scale production of photonic crystals can be realized. In Ref. [7], the authors demonstrated that by periodically applying and removing the electric field during the self-assembly process, they were able to achieve a single-crystal monolayer in a microfluidic chamber. The authors’ findings give a good example of using a time-varying external electric field to achieve single domain colloidal crystals. However, the crystals are not achieved in any optimized way, and there is no feedback.

In addition, using feedback control policies has emerged as a viable approach [8], [9], [10]. In Ref. [8], model predictive control was applied to a stochastic simulation of colloidal self-assembly to rapidly achieve a highly ordered crystalline state. In Ref. [9], dynamic programming was applied to a similar stochastic colloidal self-assembly process as in Ref. [8] for a rapid formation of a highly ordered crystalline state. Besides, in Ref. [10], Juarez and Bevan designed a proportional controller to enable the connection between the real-time sensing of this dynamic assembly process (via a crystalline order parameter named C6) and the external tunable electric potential. They further demonstrated the development of the crystalline colloidal assembly from a fluid state under feedback control [10].

Recent study has shown the Markov chain model as a promising representative for micro scale processes [9], [11]. The robustness of a Markov decision process based control policy was demonstrated on a stochastic colloidal self-assembly process in Ref. [9] to produce a highly ordered colloidal crystal. In this paper, we introduce a Markov decision process based optimal control strategy to a SiO2 self-assembly system for grain-boundary-free, and highly ordered two-dimensional crystals. In our past work [8], [9], we studied optimal control on a system with 174 particles with a one-dimensional Langevin model. This system did not exhibit persistent grain boundaries due to the small number of particles; however, in this study, 210 particles are used, and this leads to grain boundary formation and persistence during the assembly process. Fig. 1 depicts results from a Brownian dynamics simulation [13]. In Fig. 1(a) the system is in a fluid-like state, while in Fig. 1(c) the system is in the desired crystalline state. The system in Fig. 1(b) has a grain boundary, a condensed yet defected state.

In this paper, we constructed Markov chain models from a two-dimensional Langevin dynamics model, which originated from a Brownian dynamics simulation, to simulate the dynamics of the colloidal assembly process. Based on the Markov chain model, dynamic programming was used to design an optimal control policy. The purpose of this paper is to introduce a control strategy to robustly control the formation of a single domain colloidal crystal in a system where grain-boundary formation is prevalent. Detailed descriptions of the experimental system, the simulation models and the analysis of the control results are given in the following sections.

II. PROBLEM STATEMENT

In order to study the reliability of a Markov decision process based optimal controller for the colloidal self-assembly
Fig. 1. (a) top view of dispersed colloidal particles in Brownian dynamics simulation; (b) colloidal ensemble with grain boundary; (c) colloidal ensemble without grain boundary.

In a colloidal assembly system, where the state variables are the coordinates of all the particles, the state space dimension can be prohibitively large. To reduce the state space dimension, we define an approximate system state using order parameters [14]. In the colloidal self-assembly system, a pair of order parameters is identified to describe the instantaneous state of the assembly process: $\psi_6$, which is a six-fold bond orientational order parameter, captures the occurrence of the grain boundary [12]; $R_g$, the radius of gyration, captures the condensation of the system [15]. A high $R_g$ value together with a low $\psi_6$ value indicates a fluid-like state of the system, while a low $R_g$ value together with a high $\psi_6$ value indicates a highly ordered, grain-boundary-free crystalline state. Therefore, the objective of the control is to achieve a state with the highest possible $\psi_6$ value and the lowest possible $R_g$ value at the same time. These order parameters can be both theoretically calculated with a dynamics model and experimentally measured with a laboratory microscope, and this enables optimal control over both a simulation model and a laboratory experiment.

With the help of a MATLAB program, $R_g$ and $\psi_6$ values can be accurately calculated online every 0.125 s from the particle coordinates captured by a microscope in the laboratory experiments. The definitions of these two order parameters are specified in the following section.

### III. Theory and Methods

In order to more effectively study the dynamics of the colloidal assembly system considered in our study, and to more conveniently construct a control policy, we developed mathematical models for the system simulation and prediction. Based on the physical analysis of the interactions inside the system, we constructed a Brownian dynamics (BD) model which is able to accurately describe the experiments. However, considering its high computational time, we further developed a Langevin dynamics (LD) model to simplify the simulation. To use the enable of dynamic programming for our control policy, we further developed Markov chain models based the Langevin dynamics model.

In a colloidal assembly system, the state variables are the coordinates of all the particles, the state space dimension can be prohibitively large. To reduce the state space dimension, we define an approximate system state using order parameters [14]. In the colloidal self-assembly system, a pair of order parameters is identified to describe the instantaneous state of the assembly process: $\psi_6$, which is a six-fold bond orientational order parameter, captures the occurrence of the grain boundary [12]; $R_g$, the radius of gyration, captures the condensation of the system [15]. A high $R_g$ value together with a low $\psi_6$ value indicates a fluid-like state of the system, while a low $R_g$ value together with a high $\psi_6$ value indicates a highly ordered, grain-boundary-free crystalline state. Therefore, the objective of the control is to achieve a state with the highest possible $\psi_6$ value and the lowest possible $R_g$ value at the same time. These order parameters can be both theoretically calculated with a dynamics model and experimentally measured with a laboratory microscope, and this enables optimal control over both a simulation model and a laboratory experiment.

With the help of a MATLAB program, $R_g$ and $\psi_6$ values can be accurately calculated online every 0.125 s from the particle coordinates captured by a microscope in the laboratory experiments. The definitions of these two order parameters are specified in the following section.

#### A. Brownian dynamics and Langevin dynamics model

In this paper, BD simulations in the canonical ensemble are performed for 210 colloidal particles for six different constant voltages: 0.3V, 0.4V, 0.45V, 0.5V, 0.7V, and 1V, where ‘V’ stands for the maximum possible voltage of 2 volts. The time step used to update the particle configuration is set as 0.1 ms. These simulations were run for $10^7$ time steps with an initial equilibration time of $2 \times 10^6$ time steps. Configurations were stored every $10^3$ time steps and used to calculate the order parameters $R_g$ and $\psi_6$. The parameters used in the BD simulation are summarized in Table 1, while the detailed construction of the BD simulation is elaborated in Ref. [16].

Although the BD model describes the dynamics of the system accurately, it takes hours to finish even a single realization of the simulation. Considering the simplicity of Langevin dynamics and its success at microscopic-level crystallization simulation [7], [8], [17], we further developed an LD model to shorten the computational time. The Langevin equation of motion specifying the evolution of the coordinate vector in order parameter space is given as:

$$\frac{d\mathbf{x}}{dt} = -\beta \mathbf{D}(x) \cdot \nabla \mathbf{W}(x) + \nabla \cdot \mathbf{D}(x) + \zeta(x,t)$$ (1)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2a$ (nm)$^2$</td>
<td>2920</td>
</tr>
<tr>
<td>$\kappa^{-1}$ (nm)$^2$</td>
<td>10</td>
</tr>
<tr>
<td>$\psi$ (mV)$^2$</td>
<td>-50.0</td>
</tr>
<tr>
<td>$\lambda^2$</td>
<td>1.31</td>
</tr>
<tr>
<td>$J_{CM}$</td>
<td>0.4667</td>
</tr>
<tr>
<td>$\epsilon_m / \epsilon_0$</td>
<td>78</td>
</tr>
<tr>
<td>$d_e$ (µm)$^2$</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 1. Parameters for Brownian Dynamics simulation: (a) colloidal particle size, (b) Debye screening length, (c) particle and wall Stern potential, (d) peak voltage applied to electrodes, (e) Clausius-Mosotti factor for an AC field frequency at 1 MHz, (f) medium dielectric permittivity, (g) electrode spacing.
where $x = (R_{g}, \psi_6)$ is the coordinate vector, tensor field $D(x)$ and scalar potential field $W(x)$ are diffusivity landscape (DL) and free energy landscape (FEL) which depend on the coordinates. $\xi$ is the random noise variable with zero mean and variance $\langle \xi_i(x,t)\xi_j(x,t') \rangle = 2D_{ij}(x)\delta(t - t')$, as required by the Einstein relation, where $i, j$ stand for $R_{g}$ and $\psi_6$ respectively. $\nabla$ is a gradient operator, and $\nabla \cdot$ is a divergence operator. The DL and FEL, as the input quantities of the LD equation, are derived from particle scale Brownian dynamics via the linear fitting method [18]. To evaluate the accuracy of the LD simulation model, we compared the mean first passage time of both order parameters, under several constant voltages, against the results from the BD simulation.

Fig. 2 shows the comparison of the time needed for the system to decrease by 1000 nm in the $R_{g}$ coordinate, from different initial points. In Fig. 2, the mean first passage time under 0.4V from Brownian dynamics is labeled as ‘BD04V’, and that from Langevin dynamics is labeled as ‘LD04V’, etc. The error bar in the plots shows the standard deviation at each point. The comparison demonstrates a good agreement between the LD and the BD model.

Fig. 3 shows the comparison of the distribution of the time needed for the system to increase by 0.2 in the $\psi_6$ coordinate, from different initial points. For example, the distribution labeled as ‘045V 0.15–0.35’ denotes the time distribution of a 1000-realization-simulation for the system to travel 0.2 in $\psi_6$ from $\psi_6 = 0.15$ to $\psi_6 = 0.35$ under a constant voltage of 0.45V. With these mean first passage time comparisons, we validate the accuracy of our Langevin dynamics.

B. Markov Chain Model

Considering the construction of the simulation model described in the above sections, together with the observation from the laboratory experiments, the state of the assembly process at each instant depends only on the previous system state and the voltage used at that state. Therefore, we can use a parameterized-time-independent Markov chain model to simulate the system, and at the same time to enable the development of an optimal control policy.

A first step in building a Markov chain model is to define the discrete state space $S$, transition matrix $P_a$, and the discrete actuation space $A$. In this SiO$_2$ colloidal self-assembly system, we discretize the order parameter $\psi_6$ into $N_1 = 40$ evenly spaced grid cells, with $\psi_6 \in [0, 1]$; and order parameter $R_{g}$ into $N_2 = 156$ evenly spaced grid cells, with $R_{g} \in [16200, 24000]$. Therefore, the final state space is discretized into $N = N_1 \times N_2$ grid cells. The choice of the bounds on these two order parameters are based on the effective space that can be visited by the system according to the free energy landscape analysis. The actuation space $A$ is defined as $A = \{0.3V, 0.4V, 0.45V, 0.5V, 0.7V, 1.0V\}$.

The transition matrix $P_a$ contains the probabilities $p_{a_{ij}}$ for the system to be in state $j$, given the current state of $i$ and action of $a$ after a time step of $\Delta t$. The $\Delta t$ value plays a significant role in determining the accuracy of the Markov chain model. By doing in-depth mean first passage time analysis for different $\Delta t$ values, $\Delta t = 10$ s is found to give an accurate Markov chain model to capture the dynamics of the system, and is used in all the following calculations. Given the $\Delta t$ value, the probability transition matrix $P_a$ is constructed by running 50 independent Langevin dynamics simulations at each of 2070 distinct initial points. These initial points are chosen to cover the effective space of the colloidal system. With all the LD simulations, the $N \times N$ transition matrix $P_a$ for each of the six actuators was achieved. A model characterized by the pair $\{S, P_a\}$ is referred to as a finite-state Markov chain model [19].

C. Markov Decision Process Based Dynamic Programming

A Markov decision process (MDP) is characterized by $\{T, S, A, P_a\}$ [20], where $S$, $A$, and $P_a$ are defined as in the previous section. $T$ is the collection of the discrete time epoch $k$. If $T$ is finite, the MDP is called finite-horizon MDP, and it is called infinite-horizon MDP if $T$ is infinite. Given the fact that an optimal control policy for a discounted infinite-horizon MDP is time-independent [18], we decided to adopt a discounted infinite-horizon MDP in our system for
the control policy. The control policy is solved with dynamic programming using policy iteration.

In infinite-horizon MDP, the optimization is achieved over an infinite number of time steps, and the objective function for an infinite-horizon MDP is defined as:

$$ J_a(x) = E \left\{ \sum_{k=0}^{\infty} \gamma^k R(x_k, a_k) \right\} $$

(2)

where $E$ is the expectation operator, $a \in A$ is the control action, $x \in S$ is the discrete system state, $k \in T$ is the discrete time instant, and $\gamma \in (0,1)$ is the discount factor introduced to ensure the convergence of the optimization [21]. $R(x, a)$ is the one-stage reward function obtained when the system is in state $x$ and a control action $a$ is taken according to the optimality criterion, and it follows that $R(x, a) : S \times A \rightarrow \mathbb{R}$.

The optimal value function and the optimal policy $a^*(x) \in A$ are defined as:

$$ J^*(x) = \sup_{a \in A} J_a(x) $$

$$ a^*(x) = \arg \{ \sup_{a \in A} J_a(x) \} = \arg J^*(x) $$

(3)

In our particular system, the one-stage-reward function is defined as:

$$ R(x_k, a_k) = \psi_6^2 $$

(4)

The optimal control policy is obtained by solving for the input $a$ trajectory that maximizes the sum of this one-stage-reward function over an infinite time horizon via policy iteration with discount factor $\gamma = 0.99$. The design of the reward function is aimed to achieve the highest possible $\psi_6$ value, which corresponds to a highly ordered, single domain crystalline state of the system. $R_g$ is not included in the reward function, because it is impossible to achieve a high value of $\psi_6$ without a low value of $R_g$.

IV. RESULT AND DISCUSSION

Based on the methods described in Section III, we constructed six Markov chain models for each of the six actuation voltages: 0.3V, 0.4V, 0.45V, 0.5V, 0.7V, and 1V. To evaluate the accuracy of these models, in Fig. 4, we compared the 300-second uncontrolled simulation results under each of these six voltages for both the Markov chain model and the LD model. The legend ‘MCMC’, and ‘LD’ stands for Markov chain Monte Carlo simulation and Langevin dynamics simulation respectively. ‘03V’ through ‘1V’ denotes the voltage under which the simulations are conducted.

All these trajectories are averages over 1000 independent simulations from the same initial point of $R_g = 20000$ nm, and $\psi_6 = 0.1$, which corresponds to a fluid-like state of the system. According to this comparison, although the Markov chain model tends to overpredict the dynamics in the aspect of $\psi_6$, the differences are slight. One noticeable phenomenon in these uncontrolled trajectories is, as the magnitude of the voltage increases, the system would reach a better, more crystalline state, which is defined by a high $\psi_6$ value and a low $R_g$ value. Intuitively, one would expect to use the highest possible voltage all the time to most efficiently form the crystals. However, experimental results show that under a high potential, like 0.7V and 1V, the particles are pushed together too quickly, which leaves less time for the system to self-correct for defects. Once the system gets stuck in a disordered state, it will take a long time for the system to relax through self-corrections, and this makes it impractical for a real world production. Therefore, using a high voltage all the time might not be the optimal policy.

With these six Markov chain models, a time-independent optimal control policy was developed using dynamic programming. Fig. 5 shows the control policy in a two-dimensional plot. The numbers in the color bar stand for the control actions in an increasing order, where 1 stands for voltage 0.3V, and 6 stands for voltage 1V. This policy is interpreted as a look-up table. To use it, for example, when the system is in state of $R_g = 1.8 \times 10^4$ nm and $\psi_6 = 0.4$, the voltage of 1V should be used, which is colored as red and labeled as 6 in Fig. 5. In the control policy, the highest voltage of 1V is optimal in most of the effective space; this observation meets the expectation in Fig. 4 that a high voltage would result in a higher crystalline state than a lower voltage. However, one important point in the control policy is the use of lower voltages in the region of $R_g \leq 1.7 \times 10^4$ nm and $\psi_6$ around 0.6. These lower voltages are used to give the system long enough relaxation time for self-corrections in order to avoid getting stuck in the disordered state.

The implementations of the control policy onto both the Markov chain model and the Langevin dynamics simulation demonstrate that the control policy shown in Fig. 5 is able to yield a better crystalline state than an uncontrolled process. Detailed comparison results are given in Fig. 6. In Fig. 6, we only showed the LD uncontrolled trajectories for comparison, considering the agreement between the uncontrolled LD and MCMC simulations illustrated in Fig. 4.

Fig. 6 shows the comparison of 1000-realization averaged controlled Langevin dynamics (LDcontrolled) and Markov chain Monte Carlo (MCcontrolled) simulations against the uncontrolled Langevin dynamics (LD03V through LD1V).
results over 300 seconds. All the trajectories are averages over 1000 independent realizations starting from the same initial point of $R_g = 20000$ nm and $\psi_6 = 0.1$.

From the comparison, we conclude the effectiveness of the control policy by noticing the system reached $\psi_6$ of around 0.95 and $R_g$ of around $1.65 \times 10^4$ nm after 300 seconds under the control. The highest possible $\psi_6$ value that can be achieved without control is around 0.8 under 1V. The comparison reveals an obvious divergence between the controlled trajectory and the 1V uncontrolled trajectory around the point of $\psi_6 = 0.5$ and $R_g$ around $1.7 \times 10^4$ nm due to the use of 0.3V. Because of this use of a low voltage, the system enjoys a boost in $\psi_6$ by correcting defects introduced by the high voltages. Therefore it leads to a higher final $\psi_6$ value after 300 seconds of simulation. All these observations meet the expectation that high voltages should be used most of the time according to Fig. 5, while low or intermediate voltages should also be used to provide corrections during the process. Besides, the controlled Markov chain Monte Carlo simulation and the controlled Langevin dynamics simulation demonstrate a high similarity with each other, which further confirms the accuracy of the Markov chain model in approximating the LD model.

In addition, the evolution of the order parameters over time for both the Markov chain Monte Carlo and the Langevin dynamics simulations are investigated for a better understanding of the performance with respect to time. Results are summarized in Fig. 7.

Fig. 7 shows the evolution of the order parameters over time, with the system started from a fluid-like state with a low $\psi_6$ value and a high $R_g$ value, and developed into a highly ordered crystalline state with a high $\psi_6$ value and a low $R_g$ value after 300 seconds of simulation. In Fig. 7, all these four order parameter trajectories are averaged over 1000 independent realizations. From this figure, we can see that both $R_g$ and $\psi_6$ reach a steady state after about 250 seconds of simulation in the LD model. While $\psi_6$ achieves its steady state after about 150 seconds, $R_g$ stabilizes after about 250 seconds in the MCMC simulation. These observations indicate the actual time needed for the system to reach a highly ordered, grain-boundary-free crystalline state is shorter than 300 seconds under the control policy.

Besides, the analysis of the averaged optimized control actions demonstrates that the use of high voltages for most of the time together with the use of low or intermediate voltages at times of around 70 seconds is able to give a better crystalline state. Both the optimized voltage trajectories from Markov chain Monte Carlo simulation (solid red line) and Langevin dynamics simulation (solid blue line) are summarized in Fig. 8.

Fig. 8 shows the 1000-realization optimized control action for both the Markov chain model and the LD model over 300 seconds of simulation. A similar trend can be observed in both of these two trajectories: at the beginning of the simulation, the highest voltage 1V was used to give the system a strong momentum to initialize the crystallization. However, when the system evolves to $R_g$ of $1.7 \times 10^4$ nm and $\psi_6$ around 0.4, a low voltage is sometimes used to
correct defects that are caused by the strong electric forces induced by applying a high voltage. This strategy is well demonstrated in the control policy given in Fig. 5. Once the defects get resolved, the system chooses to use higher and higher voltages towards the end of the simulation to further speed up the process. The use of high voltage in the end is also justified physically, that a strong electric force is better able to maintain the system in a highly ordered state than a lower voltage. However, the optimized control policy trajectories from the MCMC and the LD simulation shows a noticeable difference between each other. This is due to the fact that the number of the realizations that used a non-trivial control policy (a control that is not using a constant voltage V) differs in the two sets of 1000 realization simulations. Specifically, 361 out of 1000 realizations used a non-trivial control policy in the LD simulation, while 631 out 1000 realizations used a non-trivial control policy in the MCMC simulation. This observation further demonstrates that using a highest constant voltage throughout the simulation does not necessarily guarantee an optimal result, in both the MCMC and the LD simulation.

The ultimate goal of this work is to feasibly produce grain-boundary-free crystals in the lab with the control policy shown in this paper. Given the simulation results, our next step is to apply the control policy in the BD simulation and in the laboratory experiment for the production of single domain, defect-free SiO₂ colloidal crystals.

V. CONCLUSIONS

Using Markov decision process based dynamic programming for a SiO₂ colloidal self-assembly process, we are able to achieve a single domain, highly ordered colloidal crystal in simulations. From all the analysis of the controlled and uncontrolled MCMC and LD simulation results, we conclude that a Markov chain model is suitable for simulating the dynamics of the stochastic self-assembly system. We further conclude that the controlled system gives significantly better results than the uncontrolled system, and a grain-boundary-free crystalline state can be achieved within about 250 seconds of simulation time under the control. Given the required policy updating time, ∆t = 10 s, and the laboratory order parameter calculation time, 0.125 s, the in-lab control experiments are practical and are ongoing.

ACKNOWLEDGMENT

Acknowledgment is made to the National Science Foundation for support of this research, through collaborative grant Grant #1124678 (Grover) and #1124648 (Bevan) through the Cyber-Enabled Discovery and Innovation program.

REFERENCES