

## ABSTRACT

ANDREW, TRICITY MARIE. Lattice-Based Surrogate Models for Two-Dimensional Particle Systems with Internal Collisions. (Under the direction of Mansoor Haider.)

Direct numerical simulation is a useful tool in modeling systems of particles. Results from direct simulations over many realizations can be used to develop probabilistic surrogate models that are able to reproduce key features of the full direct simulations. Lattice-based models are commonly used in biological applications such as biotransport in tissues or cancer modeling. These models are useful in analyzing diffusion in crowded media. This dissertation investigates the connection between continuous model simulations and lattice-based modeling approaches. We developed two continuous model simulations of particle movement: Model 1, which simulated a two-dimensional interacting particle system, and Model 2, which simulated single particle random walks in a two-dimensional representative volume element (RVE) with obstacles. For both models we used continuous simulation results to develop probabilistic surrogate models.

In Model 1, we modeled twenty-seven particles undergoing internal collisions with other particles and with domain boundaries. We considered increasing values of particle radii and tracked state changes (number of particles per subdomain) for the different subdomains. These results were used to develop a surrogate Markov chain (MC) model. We observed linear relationships between subdomain states and the particle radius for each subdomain type, in addition to good agreement between the continuous and MC model results. This approach also allowed us to estimate uncertainty for quantities of interest (subdomain states).

In Model 2, we simulated single particle random walks in a 2D RVE with fixed obstacles. We considered three different obstacle configurations—equally spaced obstacles, four internal obstacles, and multisize obstacles. For all three models, we considered a fixed particle radius with varying obstacle radii. In the equally spaced obstacles case, we also considered a fixed obstacle radius with varying particle radius. Subdomain types were delineated based on obstacle locations, and their transition probabilities were determined and used to build the surrogate lattice model. For both the continuous and surrogate models and for all obstacle configurations, we observed a linear relationship between the mean squared displacement and time. As in previous studies, the mean squared displacement was used to estimate diffusivities for both the continuous and surrogate models. In the equally spaced obstacles model, we connected the simulated effective diffusivity to a theoretical estimate of diffusivity, which was found using a Taylor Series expansion for random walks on a 2D

lattice. Our continuous and surrogate models were in good agreement for all obstacle configurations, with the equally spaced obstacles model having the best agreement between continuous and surrogate results. This study demonstrated that lattice-based surrogate models can be used to represent direct simulation of diffusion in continuous media with a prescribed arrangement of obstacles.

The modeling approach developed in this dissertation can be applied to a variety of materials with different obstacle arrangements, shapes, and symmetries. Internal geometric properties can be directly incorporated into the modeling and simulation process in order to enable the calculation of diffusivity and develop a corresponding surrogate model.



© Copyright 2020 by Tricity Marie Andrew

All Rights Reserved

Lattice-Based Surrogate Models for Two-Dimensional Particle Systems with Internal  
Collisions

by  
Tricity Marie Andrew

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Applied Mathematics

Raleigh, North Carolina

2020

APPROVED BY:

---

Alen Alexanderian

---

Kevin Flores

---

Ralph Smith

---

Mansoor Haider  
Chair of Advisory Committee

## **DEDICATION**

To my dad, Rob Andrew, who never doubted me and always had the best dad-jokes.

## BIOGRAPHY

Tricity Andrew grew up in Oklahoma. Her father named her "Tricity" because if she ever ran for President, she could be "Elec-Tricity." She has no political aspirations but enjoys having a unique name. In high school she was over-involved, participating in cross-country, track, mock trial, constitution team, school plays, a drama troupe, choir, ballet classes, and jazz classes.

As an undergraduate she attended The University of Tulsa where she majored in Mathematics with minors in Biology and Business Administration. While in college she was on the cross country and track teams. She finished her B.S. in 2013, graduating *magna cum laude*. In 2015 she began her graduate studies in Applied Mathematics at North Carolina State University. Her Ph.D. was partially funded by the National Science Foundation Graduate Research Fellowship (NSF GRF).

In her free time, Tricity enjoys running, especially at all the local social runs. She ran 10 marathons while in graduate school. She also enjoys spending time with her husband, Daniel, and her dog, Lila.

## ACKNOWLEDGEMENTS

First I would like to thank my advisor Mansoor Haider for all the guidance, encouragement, and detailed feedback over the last 5 years. Thanks for both the research advice and life advice. Thank you for being so generous with your time and support.

I'm grateful to the entire Raleigh running community that I met through different social runs and races (especially Runologie, Trophy, and Sir Walter Running). You made my favorite hobby even better. Thanks especially for all the runs when you let me rant for miles (literally), and for giving me a fun thing to look forward to every day. Special thanks to Allaire Welk, Alan Freeman, Jason Honeycutt, Justin Kearney, Candida Leigh,

Thanks to my math friends (and Modelers' Anon), especially Mallory McMahon, Michael Lavigne, Jared Cook, and Katie Ahrens.

Thank you to family members who supported me even though they had no idea what I was doing. Thank you for being interested in my research even when it was confusing.

Thanks to my dog, Lila, who moved to NC with me. The unconditional love and endless games of fetch made every day a little bit better.

Finally, I wish to thank Daniel. I never knew I wouldn't be the nerdiest person in the room at my family gatherings, but thanks to you, I'm now maybe only second nerdiest. Thank you for all your support and encouragement.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship (NSF GRF) under Grant No. 555650. I also received support from the Doctoral Dissertation Completion Grant through the NCSU Graduate School.

# TABLE OF CONTENTS

<b>LIST OF TABLES</b> . . . . .	<b>vii</b>
<b>LIST OF FIGURES</b> . . . . .	<b>x</b>
<b>Chapter 1 Introduction and Background</b> . . . . .	<b>1</b>
1.1 Two-dimensional interacting particle system . . . . .	2
1.1.1 Model descriptions . . . . .	3
1.2 Single particle random walks in a periodic representative volume element . . . . .	5
1.2.1 Model descriptions . . . . .	8
<b>Chapter 2 Two-dimensional interacting particle system: development of a Markov chain model</b> . . . . .	<b>12</b>
2.1 Model descriptions . . . . .	12
2.1.1 Continuous Model of Particle Interactions . . . . .	13
2.1.2 Markov Chain Model . . . . .	14
2.2 Results . . . . .	17
2.2.1 Continuous Model Results . . . . .	18
2.2.2 Markov Chain Model Results . . . . .	22
2.3 Discussion and Conclusions . . . . .	26
<b>Chapter 3 Single particle random walks in a periodic representative volume element</b> . . . . .	<b>28</b>
3.1 Introduction and Background . . . . .	28
3.1.1 Example of random walks on a 1D lattice . . . . .	28
3.2 Continuous Model Description . . . . .	30
3.2.1 Model for equally spaced obstacles . . . . .	34
3.2.2 Model for four internal obstacles . . . . .	35
3.2.3 Model for multisize obstacles . . . . .	36
3.3 Surrogate lattice model description . . . . .	37
3.3.1 Lattice model for equally spaced obstacles . . . . .	37
3.3.2 Lattice model for four internal obstacles . . . . .	43
3.3.3 Lattice model for multisize obstacles . . . . .	46
<b>Chapter 4 Model calibration and diffusivity estimation</b> . . . . .	<b>51</b>
4.1 Commitment index for the continuous model . . . . .	52
4.2 Estimating diffusivity using mean squared displacement and Taylor Series expansions on a lattice . . . . .	53
4.2.1 Previous work on using mean squared displacement to estimate diffusivity . . . . .	54
4.2.2 Derivation of linear relation between mean squared displacement and time on a 2D lattice . . . . .	55
4.2.3 Deriving the diffusion equation for the equally spaced obstacles model . . . . .	66
4.2.4 Deriving the diffusion equation for the multisize obstacles model . . . . .	68

4.2.5	Relating estimated diffusivity and theoretical diffusivity . . . . .	70
4.2.6	Comparing the ratio of $D_{\text{surr}}$ to $D_{\text{TS}}$ to estimate $\Delta x$ . . . . .	73
4.2.7	A theoretical estimate for $\Delta x$ . . . . .	74
<b>Chapter 5</b>	<b>Results for single particle random walks in a periodic RVE . . . . .</b>	<b>80</b>
5.1	Continuous Model Results . . . . .	82
5.1.1	Equally spaced obstacles . . . . .	82
5.1.2	Model for four internal obstacles . . . . .	91
5.1.3	Multisize obstacles . . . . .	98
5.2	Lattice Model Results . . . . .	102
5.2.1	Model for equally spaced obstacles . . . . .	102
5.2.2	Four internal obstacles . . . . .	111
5.2.3	Multisize obstacles . . . . .	117
5.3	Model comparison . . . . .	125
5.4	Conclusions . . . . .	129
<b>BIBLIOGRAPHY</b>	<b>. . . . .</b>	<b>133</b>

## LIST OF TABLES

Table 2.1	Values of the parameters $\mu$ and $\sigma$ when the truncated normal distribution (2.13) subject to the constraint (2.15) was fit to the direct simulation (continuous) model and to the Markov chain (surrogate) model. . . . .	24
Table 2.2	Slope, intercept and coefficient of determination ( $R^2$ ) for a linear regression analysis of the relationship between the average number of particles per subdomain $\bar{\eta}_\alpha$ , $\alpha = L, I, C$ and the particle radius $R$ for the continuous and the Markov chain models. . . . .	25
Table 4.1	Notation for different variables and parameters used in calculating the diffusivity. . . . .	71
Table 4.2	Resulting average distance $d_{\text{approx}}$ to a neighboring subdomain based on different values of the obstacle radius $R_{\text{obs}}$ . The value for $R_{\text{obs}}$ refers to the no obstacle case. . . . .	79
Table 5.1	Equally spaced obstacles model. Transition probabilities by subdomain for particle radius $R_{\text{part}} = 1.0$ and obstacle radius $R_{\text{obs}} = 1.0$ for one realization ( $M_c = 0$ ). . . . .	85
Table 5.2	Equally spaced obstacles model. Slope of mean squared displacement versus time for increasing number of realizations when particle radius $R_{\text{part}} = 1.0$ and for varied obstacle radius $R_{\text{obs}}$ . . . . .	88
Table 5.3	Equally spaced obstacles model. Transition probabilities by type for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ for after 2000 realizations (for $M_c = 0$ ). . . . .	89
Table 5.4	Equally spaced obstacles model. Slope of mean squared displacement versus time for increasing number of realizations when obstacle radius $R_{\text{obs}} = 2.0$ and for varied particle $R_{\text{part}}$ . . . . .	90
Table 5.5	Four internal obstacles model. Slope of mean squared displacement versus time for increasing number of realizations when particle radius $R_{\text{part}} = 1.0$ and for varied obstacle radius $R_{\text{obs}}$ . . . . .	95
Table 5.6	Four internal obstacles model. Transition probabilities for center subdomain (C) by transition type for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ for after 2000 realizations (for $M_c = 0$ ). . . . .	97
Table 5.7	Four internal obstacles model. Transition probabilities for onewall subdomain (I) by transition type for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ for after 2000 realizations (for $M_c = 0$ ). . . . .	97
Table 5.8	Four internal obstacles model. Transition probabilities for corner subdomain (L) by transition type for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ for after 2000 realizations (for $M_c = 0$ ). . . . .	97
Table 5.9	Multisize obstacles model. Slope of mean squared displacement versus time for increasing number of realizations when particle radius $R_{\text{part}} = 1.0$ , $R_{\text{obs}_1} = 1.0$ , and for varied obstacle radius $R_{\text{obs}_2}$ . . . . .	100



Table 5.10	Multisize obstacles model. Transition probabilities for center subdomain by transition type for particle radius $R_{\text{part}} = 1.0$ , $R_{\text{obs}_1} = 1.0$ and varying obstacle radius $R_{\text{obs}_2}$ after 2000 realizations (for $M_c = 0$ ). . . .	102
Table 5.11	Equally spaced obstacles. Slope values for $\langle r^2 \rangle$ versus time for continuous model and surrogate model with the best $M_c$ values. . . . .	104
Table 5.12	Equally spaced obstacles model. Transition probabilities by type for optimal $M_c$ for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ after 2000 realizations. . . . .	107
Table 5.13	Equally spaced obstacles model. Table showing commitment index $M_c$ , probability of staying in current subdomain $q$ , and diffusivities obtained via simulation for the continuous and surrogate models ( $D_{\text{cont}}$ and $D_{\text{surr}}$ , respectively) for each obstacle radius value. . . . .	109
Table 5.14	Equally spaced obstacles model. Table showing diffusivity and effective $\Delta x$ . . . . .	110
Table 5.15	Equally spaced obstacles model. Resulting average distance $d_{\text{approx}}$ to a neighboring subdomain based on different values of the obstacle radius $R_{\text{obs}}$ . The value for $R_{\text{obs}}$ refers to the no obstacle case. . . . .	111
Table 5.16	Four internal obstacles model. Slope values for $\langle r^2 \rangle$ versus time for continuous model and surrogate model with the best $M_c$ values. . .	112
Table 5.17	Four internal obstacles model. Transition probabilities for center subdomain (C) by transition type for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ for after 2000 realizations (for $M_c = 4$ ). .	114
Table 5.18	Four internal obstacles model. Transition probabilities for onewall subdomain (I) by transition type for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ for after 2000 realizations (for $M_c = 4$ ). .	114
Table 5.19	Four internal obstacles model. Transition probabilities for corner subdomain (L) by transition type for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ for after 2000 realizations (for $M_c = 4$ ). .	114
Table 5.20	Four internal obstacles model. Table showing the diffusivities obtained via simulation for the continuous and surrogate models ( $D_{\text{cont}}$ and $D_{\text{surr}}$ , respectively) for each obstacle radius value (commitment index $M_c = 4$ for all cases). . . . .	117
Table 5.21	Multisize obstacles model. Slope values for $\langle r^2 \rangle$ versus time for continuous model and surrogate model with the best $M_c$ values. $R_{\text{part}} = 1.0$ , $R_{\text{obs}_1} = 1.0$ . . . . .	119
Table 5.22	Multisize obstacles model. Transition probabilities for center subdomain by transition type for particle radius $R_{\text{part}} = 1.0$ , $R_{\text{obs}_1} = 1.0$ and varying obstacle radius $R_{\text{obs}_2}$ for after 2000 realizations (for $M_c = 0$ ). .	120
Table 5.23	Multisize obstacles model. Table showing commitment index $M_c$ , probability of staying in current subdomain $q$ , and diffusivities obtained via simulation for the continuous and surrogate models ( $D_{\text{cont}}$ and $D_{\text{surr}}$ , respectively) for each obstacle radius value. . . . .	120
Table 5.24	Diffusivity across obstacle radii for all models considered. . . . .	126
Table 5.25	Diffusivity across obstacle radii and corresponding solid volume fraction for equally spaced obstacles model and Mackie-Mears relation.	129

Table 5.26	Diffusivity across obstacle radii and corresponding solid volume fraction for four internal obstacles model and Mackie-Mears relation.	129
Table 5.27	Diffusivity across obstacle radii and corresponding solid volume fraction for multisized obstacles model and Mackie-Mears relation.	129

## LIST OF FIGURES

Figure 1.1	Model 1. (a) Continuous model configuration at $t = 0$ . (b) Illustration of surrogate lattice model. (c) State transition probabilities for building a lattice-based surrogate model. . . . .	4
Figure 1.2	Single particle random walks on a 2D lattice with varying obstacle concentrations. Mean squared displacement, $\langle r^2 \rangle$ is linearly with time. Reprinted from Biophysical Journal, Vol 66, M.J. Saxton, "Anomalous diffusion due to obstacles: a Monte Carlo Study", Page No. 396, Copyright 1994, with permission from Elsevier [32]. . . . .	6
Figure 1.3	Single particle random walks on a 2D lattice with varying obstacle concentrations. The estimate for normalized diffusivity, $D^*$ , depends upon the number of realizations included in the average of mean squared displacement. Reprinted from Biophysical Journal, Vol 72, M.J. Saxton, "Single-particle tracking: the distribution of diffusion coefficients", Page No. 1745, Copyright 1997, with permission from Elsevier [31]. . . . .	7
Figure 1.4	Plot of $\langle r^2 \rangle$ versus time in 3D systems studied in <i>Vilaseca</i> , showing that, for large timesteps, the relationship between $\langle r^2 \rangle$ versus time is linear. Reproduced from Ref. [35] with permission from the PCCP Owner Societies. . . . .	8
Figure 1.5	Model 2. Domains for deterministic model with numbered subdomains for particle radius $R_{\text{part}} = 1.0$ and obstacle radius $R_{\text{obs}} = 2.0$ . (a) Equally spaced obstacles model.(b) Four internal obstacles model.	10
Figure 1.6	Model 2. Domain for deterministic model of multisize obstacles with numbered subdomains for particle radius $R_{\text{part}} = 1.0$ , the first set of obstacle radius $R_{\text{obs}_1} = 1.0$ and the second set of obstacle radius $R_{\text{obs}_2} = 2.0$ . . . . .	10
Figure 2.1	Two dimensional dynamic model for perfectly elastic collisions between 27 particles and four rigid walls in a square domain: (left) the initial state ( $t = 0$ ) for a sample realization, (right) an intermediate state ( $t > 0$ ). . . . .	13
Figure 2.2	Partitioning of the direct simulation domain into 9 square subdomains of equal area: (left) subdomains are delineated into three types based on the number of subdomain boundaries, (right) illustration of adjacency relationships between the 9 subdomains. . . . .	14
Figure 2.3	Illustration of the Markov chain model for tracking transition probabilities in the number of particles per subdomain at a snapshot in time when $N_s = 10$ . Note that transition probabilities are calculated separately for each of the three subdomain types $\alpha = L, I, C$ using equation (2.10). . . . .	17

Figure 2.4	Illustration of the stationary response for the average number of particles per subdomain in the case of 27 particles over the 9 subdomains shown in Fig. 2.2. Results are shown for a single realization in the case $R = 0.5$ and $\Delta t = 0.0125$ ( $N = 2 \times 10^5$ ). . . . .	19
Figure 2.5	Illustration of the effect on histograms for the number of particles per subdomain $\bar{\eta}_C$ as the number of realizations is increased. Results are shown for the center subdomain in the case $R = 0.5$ . . . . .	19
Figure 2.6	Histograms and fitted normal distributions for mean values of the number of particles per subdomain $\bar{\eta}_\alpha$ with varying subdomain type ( $\alpha = L, I, C$ ) and particle radius $R$ obtained by direct simulation using the continuous model. . . . .	20
Figure 2.7	Effects of varying the maximum number of states per subdomain $N_s$ in the Markov chain model. Stationary distributions obtained from (2.12) were fit with the probability density function (2.13) subject to the constraint (2.15) in the case $R = 0.5$ : (a) Center ( $\alpha = C$ ), (b) Corner ( $\alpha = L$ ), (c) One-wall ( $\alpha = I$ ). . . . .	21
Figure 2.8	Distributions for the number of particles in each subdomain type ( $\alpha = L, I, C$ ) with varying particle radius $R$ . Stationary distributions $\bar{\pi}_\alpha$ from the Markov Chain (surrogate) model are compared to data obtained using the (continuous) direct simulation model. The surrogate model was fit using the truncated normal distribution (2.13) subject to the constraint (2.15). . . . .	21
Figure 2.9	Distributions for the number of particles in each subdomain grouped by subdomain type ( $\alpha = L, I, C$ ) to illustrate effects of varying particle radius $R$ . Stationary distributions $\bar{\pi}_\alpha$ from the Markov Chain (surrogate) model were fit using the truncated normal distribution (2.13) subject to the constraint (2.15). . . . .	23
Figure 2.10	Linear regression fits for the relationship between the average number of particles per subdomain $\bar{\eta}_\alpha$ , $\alpha = L, I, C$ and the particle radius $R$ , delineated by subomain type. . . . .	24
Figure 3.1	Illustrations of a particle moving in a one-dimensional random walk, having equal probability of moving one unit to the right or left. (a) The particle's first step. (b) Molecule locations possible after $N = 4$ steps. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer. Diffusion. In: <i>Foundations of Applied Mathematics</i> by M.H. Holmes, Springer Science + Business Media, LLC (2009) [15]. . . . .	29
Figure 3.2	Illustrations of 5000 realizations for a particle's random walk with $N = 100$ time steps. Top: The different paths taken during the 5000 realizations. Bottom: Spatial distribution of the particle's location after the last timestep. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer. Diffusion. In: <i>Foundations of Applied Mathematics</i> by M.H. Holmes, Springer Science + Business Media, LLC (2009) [15]. . . . .	30

Figure 3.3	Domains for deterministic model with numbered subdomains for particle radius $R_{\text{part}} = 1.0$ and obstacle radius $R_{\text{obs}} = 2.0$ . (a) Equally spaced obstacles model.(b) Four internal obstacles model. . . . .	31
Figure 3.4	Domain for deterministic model of multisize obstacles with numbered subdomains for particle radius $R_{\text{part}} = 1.0$ , the first set of obstacle radius $R_{\text{obs}_1} = 1.0$ and the second set of obstacle radius $R_{\text{obs}_2} = 2.0$ . . . . .	31
Figure 3.5	Domain diagram for four internal obstacles case with subdomain types labeled. Solid lines represent the original domain, or the RVE, and dashed lines represent a portion of the extended domain. . . .	32
Figure 3.6	Illustration of a particle's random walk path in the four internal obstacle case. The larger black circles are obstacles, the red circle indicates the particle's initial location, and the light orange circles depict the random walk path that the particle takes. . . . .	33
Figure 3.7	Illustrations of periodic boundary conditions, with exaggerated step-sizes for clarity. (a) The particle moves from location (1) to location (2). Location (2) is on the boundary, but the center of the particle is still in the RVE, so the particle is not considered to have crossed the boundary yet. (b) The particle moves from location (2) to location (3). Location (3) is over the boundary, so the particle reenters the RVE from the opposite side of the domain. (c) Particle moves from location (3) to location (3*). . . . .	33
Figure 3.8	Domain diagram for multisize obstacles to illustrate which obstacle radii are varied. In simulations $R_{\text{part}}$ and $R_{\text{obs}_1}$ are fixed at 1.0 while $R_{\text{obs}_2}$ is varied. . . . .	35
Figure 3.9	Equally spaced obstacles model. (a) Continuous domain with solid red lines indicating RVE subdomains and dashed gray lines indicating subdomains in the extended domain. Solid blue circles are obstacles in the original domain, and solid gray circles are obstacles in the extended domain. The center subdomain (subdomain 5) is highlighted in yellow, with transition types indicated with a green arrow. Only cardinal direction transitions are possible (in addition to staying in the current subdomain), so only one arrow is shown. (b) Lattice representation with red rings representing possible subdomains in the original domain and dashed gray rings indicating possible subdomains in the extended domain. Distinct transition types from subdomain 5 indicated with green arrows and numbers. . . . .	36

- Figure 3.10 Continuous domain for four internal obstacle case. RVE is shown with red boundary lines and extended domain is shown with gray, dashed boundary lines. Obstacles are indicated by solid blue circles in RVE and solid gray circles in extended domain. Distinct transition types indicated with green arrow. (a) Illustration of how subdomains are numbered for tracking transition probabilities. (b) Illustration of transitions from the center subdomain (subdomain 5), with the center subdomain highlighted in yellow. Note that transitions in any cardinal direction should occur with the same probability, this is illustrated with a single green arrow indicating there is only one type of transition possible. From subdomain 5 there are two independent transition probabilities. (c) Illustration of transitions from a subdomain 2, a onewall subdomain, with subdomain 2 highlighted in yellow. From subdomain 2 there are five independent transition probabilities. (d) Illustrations of transitions from subdomain 1, a corner subdomain, with subdomain 1 highlighted in yellow. From subdomain 1 there are five independent transition probabilities. . . 41
- Figure 3.11 Lattice representation for four internal obstacle case with red rings indicating possible subdomains in the original domain and dashed gray rings indicating possible subdomains in the extended domain. Possible transitions are indicated with solid blue lines. Distinct transition types indicated with green arrows and numbers. (a) Lattice representation. (b) Illustration of transitions from subdomain 5, the center subdomain (2 distinct transition types). (c) Illustration of transitions from a subdomain 2, a onewall subdomain (5 distinct transition types). (d) Illustrations of transitions from subdomain 1, a corner subdomain (5 distinct transition types). . . . . . 42
- Figure 3.12 Multisize obstacles model. (a) Continuous grid with solid red lines indicating subdomains in the original domain and dashed gray lines indicating subdomains in the extended domain. Solid blue and green circles indicate obstacles in the original domain, with the smaller green circles indicating obstacles with obstacle radius fixed at  $R_{obs_1} = 1.0$ , and the larger blue circles indicating obstacles with varying obstacle radius  $R_{obs_2}$ . Solid gray circles are obstacles in the extended domain, with the circle size smaller and larger indicating radius  $R_{obs_1}$  and  $R_{obs_2}$ , respectively. Subdomain 1 is highlighted in yellow, with distinct transition types indicated with a green arrow. (b) Lattice representation with solid red rings indicating possible subdomains in the original domain and dashed gray rings indicating possible subdomains in the extended domain. Distinct transition types from subdomain 1 indicated with green arrows and numbers. 47

Figure 4.1	Commitment index, $M_C$ , illustration. (a) The particle (solid black circle) has just crossed the boundary between subdomains (dashed gray line). We call this the 1st-step location. Possible particle 2nd-step locations are indicated by the solid black ring. (b) Three possible 2nd-step locations are indicated with the solid red circles. We illustrate some possible next steps (3rd-step locations) with red rings. (c) For a specific 2nd-step location (solid red circle), the 1st-step location and ring of possible 2nd-step locations are colored gray (as these steps have passed). A red ring illustrates the particles' remaining possible 3rd-step locations. (d) Possible 4th-step locations are indicated with blue rings. . . . .	52
Figure 4.2	Illustration of probabilities for the three cases for independent events $h_i(N)$ and $h_i(M)$ , where $p_x$ is the probability of moving horizontally. Cases (ii) and (iii) are associated with the same probability, $(p_x/2)^2$ . . . . .	59
Figure 4.3	Illustration of possible linear distances for a particle to leave the subdomain at when the particle is at a particular location. . . . .	71
Figure 4.4	Subdomain $\Omega$ with a particle located at an arbitrary position $(\xi, \eta)$ . Boundaries are denoted as $\Gamma_i$ for $i = 1, 2, 3, 4$ . . . . .	74
Figure 4.5	Subdomain $\Omega$ with a particle located at an arbitrary position $(\xi, \eta)$ . . . . .	75
Figure 4.6	Close up of subdomain $\Omega$ to show distances $R_{\text{part}}$ , $R_{\text{obs}}$ , and $r$ . Distance $r$ must be included in calculation of average distance a particle is from a neighboring subdomain. . . . .	76
Figure 4.7	Subdomain regions used in calculating the average distance to a neighboring subdomain. (a) General dimensions and location of excluded regions (gray). (b) Region 1 is highlighted in yellow. . . . .	77
Figure 4.8	Subdomain regions used in calculating the average distance to a neighboring subdomain, where gray regions are the exclusion zone. (a) Region 2 is highlighted in yellow. (b) Region 3 is highlighted in yellow. . . . .	77
Figure 5.1	Domains for continuous model with numbered subdomains for particle radius $R_{\text{part}} = 1.0$ and obstacle radius $R_{\text{obs}} = 2.0$ . (a) Equally spaced obstacles model. (b) Four internal obstacles model. . . . .	81
Figure 5.2	Domain for continuous model of multisize obstacles with numbered subdomains for particle radius $R_{\text{part}} = 1.0$ , and obstacle radii $R_{\text{obs}_1} = 1.0$ and $R_{\text{obs}_2} = 2.0$ . . . . .	81

Figure 5.3	Equally spaced obstacles model. Transitions from subdomain 6 are shown to illustrate convergence of transition probabilities. (a) Simulation domain with subdomain 6 highlighted in yellow. Red boundary lines indicate subdomains in the original domain, and grey boundary lines indicated subdomains in the extended domain. Blue circles represent obstacles. (b) Plot of all transition probabilities. At this scale, only the probability of staying in the current subdomain is visible. (c) Plot of transition probability for staying in subdomain 6. (d) Plot of transition probabilities for transitions from subdomain 6 to subdomains in cardinal directions (i.e. transitions to subdomains 3, 4, 5, 9). . . . .	83
Figure 5.4	Equally spaced obstacles model. Effect of increasing number of realizations on mean squared displacement for $R_{\text{part}} = 1.0$ and $R_{\text{obs}} = 1.0$ . (a) Mean squared displacement averages for increasing number of realizations, (b) Linear fit of mean squared displacement averages for increasing number of realizations. . . . .	86
Figure 5.5	Equally spaced obstacles model. Slope of $\langle r^2 \rangle$ vs. time for increasing values of obstacle radius $R_{\text{obs}}$ , where $R_{\text{part}} = 1.0$ . . . . .	87
Figure 5.6	Equally spaced obstacles model. Diagram showing transition types based on subdomain type. Distinct transitions are indicated with green arrows, and transition types are numbered in the surrogate model representation. In the continuous model, red lines indicate boundaries between subdomains in the original RVE, dashed lines indicate subdomain boundaries in the extended domain, solid blue circles indicate obstacles in the original RVE, and solid gray circles indicate obstacles in the extended domain. In the surrogate model, red circles correspond to subdomains in the original RVE, gray dashed circles correspond to subdomains in the extended domain, and blue lines indicate possible transitions. . . . .	88
Figure 5.7	Equally spaced obstacles model. Slope of $\langle r^2 \rangle$ vs. time for increasing values of particle radius $R_{\text{part}}$ , where $R_{\text{obs}} = 2.0$ . . . . .	90
Figure 5.8	Equally spaced obstacles model. Comparing slope of $\langle r^2 \rangle$ vs. time for varied particle radius value and varied obstacle radius values. . . . .	91
Figure 5.9	Four internal obstacles model. Transitions from subdomain 6 are shown to illustrate convergence of transition probabilities. (a) Simulation domain with subdomain 6 highlighted in yellow. Red boundary lines indicate subdomains in the original domain, and grey boundary lines indicated subdomains in the extended domain. Blue circles represent obstacles. (b) Plot of probability for staying in subdomain 6. (c) Plot of transition probabilities for cardinal direction transitions (i.e. transitions to subdomains 3, 4, 5, and 9). (d) Plot of transition probabilities for transitions from subdomain 6 to subdomains in diagonal directions (i.e. transitions to subdomains 1, 2, 7, and 8, where 2 and 8 are blocked by obstacles and remain at zero). . . . .	92



Figure 5.10	Four obstacle model for $R_{\text{part}} = 1.0$ and $R_{\text{obs}} = 1.0$ . Effect of increasing number of realizations on mean squared displacement. (a) Mean squared displacement averages for increasing number of realizations, (b) Linear fit of mean squared displacement averages for increasing number of realizations. . . . .	94
Figure 5.11	Four internal obstacles model. Slope of $\langle r^2 \rangle$ vs. time for increasing values of obstacle radius $R_{\text{obs}}$ , where $R_{\text{part}} = 1.0$ . . . . .	95
Figure 5.12	Four internal obstacles model. Diagram showing transition types based on subdomain type. Distinct transitions are indicated with green arrows, and transition types are numbered in the surrogate model representation. In the continuous model, red lines indicate boundaries between subdomains in the original RVE, dashed lines indicate subdomain boundaries in the extended domain, solid blue circles indicate obstacles in the original RVE, and solid gray circles indicate obstacles in the extended domain. In the surrogate model, red circles correspond to subdomains in the original RVE, gray dashed circles correspond to subdomains in the extended domain, and blue lines indicate possible transitions. . . . .	96
Figure 5.13	Multisize obstacles model. Diagram showing the continuous model RVE. . . . .	98
Figure 5.14	Multisize obstacle model with $R_{\text{part}} = 1.0$ , $R_{\text{obs}_1} = 1.0$ , and $R_{\text{obs}_2} = 1.0$ . Effect of increasing number of realizations on mean squared displacement. (a) Mean squared displacement averages for increasing number of realizations, (b) Linear fit of mean squared displacement averages for increasing number of realizations. . . . .	99
Figure 5.15	Multisize obstacles model. Slope of $\langle r^2 \rangle$ versus time for increasing values of obstacle radius $R_{\text{obs}_2}$ , where $R_{\text{part}} = 1.0$ , $R_{\text{obs}_1} = 1.0$ . . . . .	100
Figure 5.16	Multisize obstacles model. Diagram showing transition types based on subdomain type. Distinct transitions are indicated with green arrows, and transition types are numbered in the surrogate model representation. For continuous models, red lines indicate boundaries between subdomains in the original RVE, dashed lines indicate subdomain boundaries in the extended domain, solid blue circles indicate obstacles in the original RVE, solid gray circles indicate obstacles in the extended domain. For surrogate models, red circles correspond to subdomains in the original RVE, gray dashed circles correspond to subdomains in the extended domain, and blue lines indicate possible transitions. . . . .	101
Figure 5.17	Equally spaced obstacles. $\langle r^2 \rangle$ versus time for continuous model and resulting surrogate model when commitment index $M_c$ is varied. Results shown for particle radius $R_{\text{part}} = 1.0$ and obstacle radius $R_{\text{obs}} = 1.0$ . . . . .	103

Figure 5.18	Equally spaced obstacles. Slope of $\langle r^2 \rangle$ versus time for an increasing number of realizations, along with a cumulative average of the slope value. Shown for particle radius 1.0 and obstacle radius 2.0. (a) $M_c = 3$ , (b) $M_c = 4$ , (c) $M_c = 5$ , (d) $M_c = 6$ . . . . .	104
Figure 5.19	Equally spaced obstacles. Slope of $\langle r^2 \rangle$ vs. time for increasing number of realizations, along with a cumulative average of the slope value. For $R_{\text{part}} = 1.0$ . (a) $R_{\text{obs}} = 0.5, M_c = 7$ , (b) $R_{\text{obs}} = 1.0, M_c = 6$ , (c) $R_{\text{obs}} = 1.5, M_c = 5$ , (d) $R_{\text{obs}} = 2.0, M_c = 5$ , (e) $R_{\text{obs}} = 2.5, M_c = 4$ , and (f) $R_{\text{obs}} = 3.0, M_c = 4$ . . . . .	105
Figure 5.20	Equally spaced obstacles. Slope of $\langle r^2 \rangle$ versus time for continuous model and resulting surrogate model with best $M_c$ values. Results for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ . . . . .	106
Figure 5.21	Equally spaced obstacles model. Estimated diffusivity for continuous and surrogate models for $R_{\text{part}} = 1.0$ . (a) $R_{\text{obs}} = 0.5, M_c = 7$ , (b) $R_{\text{obs}} = 1.0, M_c = 6$ , (c) $R_{\text{obs}} = 1.5, M_c = 5$ , (d) $R_{\text{obs}} = 2.0, M_c = 5$ , (e) $R_{\text{obs}} = 2.5, M_c = 4$ , (f) $R_{\text{obs}} = 3.0, M_c = 4$ . . . . .	108
Figure 5.22	Four internal obstacles model. Slope of $\langle r^2 \rangle$ vs. time over differing number of realizations, along with a cumulative average of the slope value. For $R_{\text{part}} = 1.0$ . (a) $R_{\text{obs}} = 0.5, M_c = 4$ , (b) $R_{\text{obs}} = 1.0, M_c = 4$ , (c) $R_{\text{obs}} = 1.5, M_c = 4$ , (d) $R_{\text{obs}} = 2.0, M_c = 4$ , (e) $R_{\text{obs}} = 2.5, M_c = 4$ , and (f) $R_{\text{obs}} = 3.0, M_c = 4$ . . . . .	113
Figure 5.23	Four internal obstacles model. Slope of $\langle r^2 \rangle$ versus time for continuous model and resulting surrogate model with best $M_c$ values ( $M_c = 4$ for all cases). Results for particle radius $R_{\text{part}} = 1.0$ and varying obstacle radius $R_{\text{obs}}$ . . . . .	115
Figure 5.24	Four internal obstacles model. Estimated diffusivity for continuous and surrogate models for $R_{\text{part}} = 1.0$ when $M_c = 4$ . (a) $R_{\text{obs}} = 0.5$ , (b) $R_{\text{obs}} = 1.0$ , (c) $R_{\text{obs}} = 1.5$ , (d) $R_{\text{obs}} = 2.0$ , (e) $R_{\text{obs}} = 2.5$ , (f) $R_{\text{obs}} = 3.0$ . . .	116
Figure 5.25	Multisize obstacles model. Slope of $\langle r^2 \rangle$ vs. time over differing number of realizations, along with a cumulative average of the slope value. For $R_{\text{part}} = 1.0$ . (a) $R_{\text{obs}} = 0.5, M_c = 7$ , (b) $R_{\text{obs}} = 1.0, M_c = 6$ , (c) $R_{\text{obs}} = 1.5, M_c = 5$ , (d) $R_{\text{obs}} = 2.0, M_c = 5$ , (e) $R_{\text{obs}} = 2.5, M_c = 4$ , and (f) $R_{\text{obs}} = 3.0, M_c = 4$ . . . . .	118
Figure 5.26	Multisize obstacles model. Slope of $\langle r^2 \rangle$ versus time for continuous model and resulting surrogate model with best $M_c$ values. Results for $R_{\text{part}} = 1.0, R_{\text{obs}_1} = 1.0$ , and varying $R_{\text{obs}_2}$ . . . . .	119
Figure 5.27	Multisize obstacles model. Estimated effective diffusivity for continuous and surrogate models for $R_{\text{part}} = 1.0, R_{\text{obs}_1} = 1.0$ (based on simulation results). (a) $R_{\text{obs}_2} = 0.5, M_c = 6$ . (b) $R_{\text{obs}_2} = 1.0, M_c = 6$ . (c) $R_{\text{obs}_2} = 1.5, M_c = 6$ . (d) $R_{\text{obs}_2} = 2.0, M_c = 6$ . (e) $R_{\text{obs}_2} = 2.5, M_c = 6$ . (f) $R_{\text{obs}_2} = 3.0, M_c = 6$ . . . . .	121
Figure 5.28	Multisize obstacles model. Horizontal and vertical mean squared displacements for the surrogate model. $R_{\text{part}} = 1.0, R_{\text{obs}_1} = 1.0$ . (a) $R_{\text{obs}_2} = 0.5, M_c = 6$ . (b) $R_{\text{obs}_2} = 1.0, M_c = 6$ . (c) $R_{\text{obs}_2} = 1.5, M_c = 6$ . (d) $R_{\text{obs}_2} = 2.0, M_c = 6$ . (e) $R_{\text{obs}_2} = 2.5, M_c = 6$ . (f) $R_{\text{obs}_2} = 3.0, M_c = 6$ . . . .	123

Figure 5.29	Multisize obstacles model. Ratio of horizontal diffusivity to vertical diffusivity $R_{\text{part}} = 1.0$ , $R_{\text{obs}_1} = 1.0$ . (a) $R_{\text{obs}_2} = 0.5$ , $M_c = 6$ . (b) $R_{\text{obs}_2} = 1.0$ , $M_c = 6$ . (c) $R_{\text{obs}_2} = 1.5$ , $M_c = 6$ . (d) $R_{\text{obs}_2} = 2.0$ , $M_c = 6$ . (e) $R_{\text{obs}_2} = 2.5$ , $M_c = 6$ . (f) $R_{\text{obs}_2} = 3.0$ , $M_c = 6$ . . . . .	124
Figure 5.30	Multisize obstacles model. RVE illustrating case where $R_{\text{obs}_1} = 1.0$ and $R_{\text{obs}_2} = 3.0$ (to scale). . . . .	125
Figure 5.31	Comparison of diffusivity versus obstacle radius for all models considered. (a) Continuous models only. (b) Continuous and surrogate models. . . . .	126
Figure 5.32	Comparison of effective diffusivity for continuous simulations and corresponding Mackie-Mears relation. (a) Equally spaced obstacles model. (b) Four internal obstacles model. (c) Multisize obstacles. (d) All models considered and Mackie-Mears relation. (a)-(c) plotted at a different scale. . . . .	128

## CHAPTER

# 1

## INTRODUCTION AND BACKGROUND

In the modeling of multiscale or heterogeneous systems, direct numerical simulation is one tool that can be used to analyze and quantify stationary properties of the system. Such an approach may be considered when the full coupling of governing equations at different scales is difficult due to computational complexity or the lack of appropriate theoretical models describing interactions among system variables at the macro- or meso-scale [10, 1]. Examples of direct simulation approaches include the use of representative domains (volume elements) [13, 14, 19] or agent-based models [2, 16, 25]. Both approaches have been applied to simulate and identify complex or heterogeneous interactions on microscopic scales, and have seen wide use in a variety of applications. Direct simulation results can be used to develop probabilistic or lattice-based surrogate models. In direct simulation models, cells can be represented as discrete entities and used to model applications such as cancer progression, cartilage development, wound healing, and other tissue models [5, 28, 3, 6, 24]. Using direct simulation (off-lattice models) we can model particles (or cells) as discrete objects and incorporate rules for how they move in space or interact with other particles. While direct simulation requires significant computational power, these methods give insight into biological processes and can inform more computationally efficient lattice models.[3]. Lattice-based models are commonly used to study cancer and other tissue models, such as the Cellular Potts Model (CPM) in the biomedical field [3, 17]. In the CPM,

a cell is defined for a region of multiple lattice sites [12, 30, 34]. This lattice model allows for study of biological cell dynamics and has been applied to various biomedical problems, such as epidermal biology, cancer, angiogenesis, and vasculogenesis [3, 29]. The connection between off-lattice and lattice-based models is ripe for further investigation.

In this dissertation, we explore the relationship between continuous and lattice-based surrogate models, investigated in the context of two models: (1) a two-dimensional interacting particle system, and (2) single particle random walks in a two-dimensional periodic representative volume element (RVE) with circular obstacles. In both models we use results obtained from direct simulations to systematically develop corresponding probabilistic surrogate models. Model 1 can be thought of as an initial motivating problem, and is presented in Chapter 2. In Model 1 we identify state-change probabilities in the continuous model in order to build a surrogate Markov chain (MC) model. Model 2 is presented in Chapters 3 - 5. In Model 2, we track different types of transitions in the continuous model in order to build a surrogate lattice model that preserves particle diffusivity, tailoring the approach to the different obstacle configurations. Throughout this thesis the term surrogate model will be used in the context described above, and we note that the same term is widely used to characterize response surface models which are more statistical in their nature and different from the models developed herein. Since Model 1 is delineated by distinct Markovian states (number of particles in a subdomain), it is possible to build a corresponding Markov chain model. However, in Model 2 we are most interested in the mean squared displacement and its relation to diffusion, and, precluding the possibility of distinct states, we cannot construct a corresponding Markov chain model.

## **1.1 Two-dimensional interacting particle system**

Many interacting particle systems exhibit structure in their stationary response due to properties of the particles, a finite set of rules governing their interactions, or features or constraints within the simulation domain. Direct simulation of particle interactions over multiple realizations of the initial system configuration can potentially provide detailed sets of information for quantities of interest. In particular, such data sets obtained from direct simulation can enable quantification of both the mean properties of stationary system variables as well as uncertainty due to inherent stochasticity. These statistical properties of the system can be used to develop deterministic models at coarser scales by studying the mean or expected values of key system variables. Alternatively, these data sets can be used to develop probabilistic surrogate models [4] that preserve the statistical features of these system variables using a simpler model representation. Model 1 in this study explores the

latter idea in the context of a two-dimensional model of colliding circular particles within a square domain.

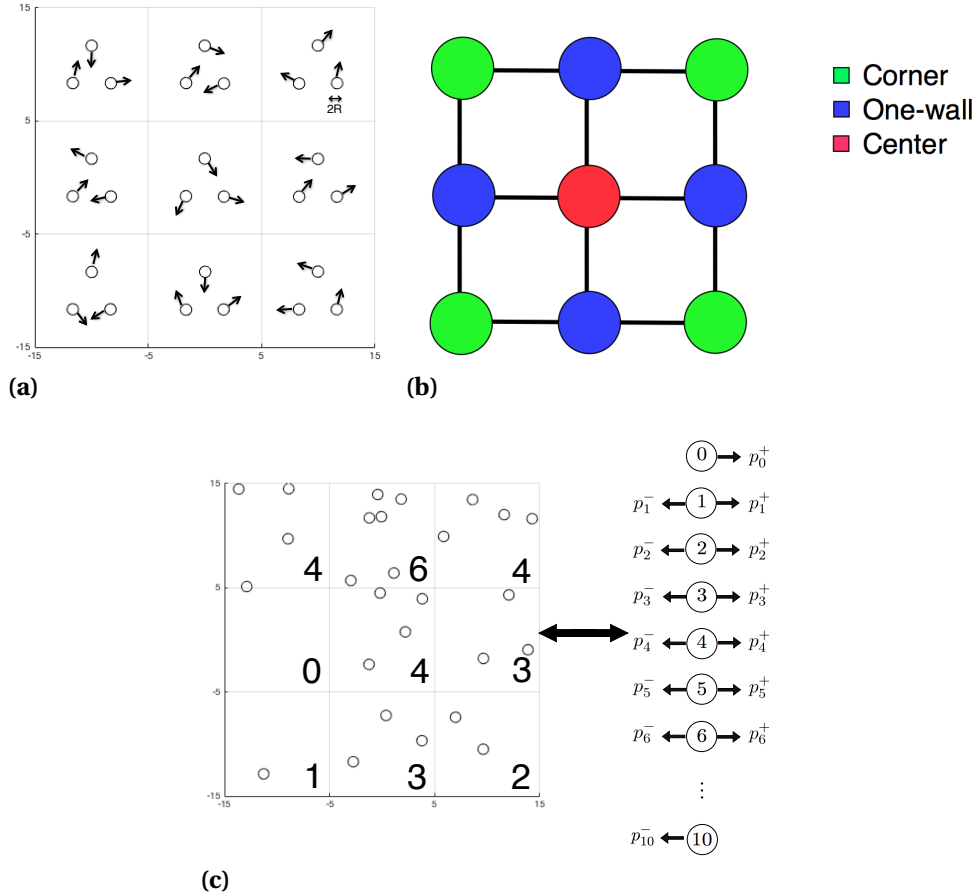
When quantities of interest can be tracked by classification into states at each point in time, and transitions between these states at successive times, Markov chains [26, 33] provide a framework for developing probabilistic surrogate models. In many systems, statistical properties of the particle interactions become stationary, i.e., independent of time, when dynamics are simulated over long time periods and over many realizations. In this regime, distributions for key quantities of interest can be identified and represented via a small number of parameters, provided that appropriate probability density functions can be identified. For example, a stationary variable exhibiting Gaussian-like features can be represented via the mean and the standard deviation in a normal distribution. Overall, the success of this approach will depend on several factors. First, the direct simulation model must be robust and efficient enough to exhibit stationary statistical properties over appropriate time scales and over many realizations. Second, the number and nature of the states identified for use in the Markov chain (MC) model must be such that the surrogate model efficiently reproduces key features of the full continuous (directly simulated) model. And, lastly, the data (histograms) for stationary quantities of interest obtained using the surrogate MC model should be self-consistent in that a family of distributions can be identified to accurately reproduce them via curve-fitting.

Model 1 in this thesis investigates the aforementioned approach in the context of a two-dimensional model of interacting circular particles within a square domain. The particles are assumed to all have the same radius and to interact with each other and with the rigid walls of the square domain via perfectly elastic collisions. System dynamics are captured using a computational model that enforces conservation of momentum for all collisions. The stationary quantity of interest is chosen as the number of particles within each of nine coarser subdomains, classified based on three types of geometric features. Direct simulation of the continuous model is used to develop a surrogate Markov Chain model and evaluate its accuracy. The resulting models are then used to quantify mean properties and uncertainty of the stationary variable as the particle radius is varied.

### 1.1.1 Model descriptions

In the continuous model, we directly simulate particle interactions within a square, 2D domain with fixed boundary walls. We consider twenty-seven interacting particles with equal mass and radius ( $R$ ) (Fig. 1.1: Left). We consider the square domain as a union of nine non-overlapping square subdomains of equal area and with three types (Fig. 1.1). We

have four corner ( $L$ ) subdomains ( $i = 1, 3, 7, 9$ ), four one-wall ( $I$ ) subdomains ( $i = 2, 4, 6, 8$ ), and one center ( $C$ ) subdomain ( $i = 5$ ). The state of each subdomain at any given time is the number of particles with centers in a particular subdomain. Initially, the particles are evenly distributed among the nine subdomains (Fig. 1.1: Left). The particles have a constant initial speed  $|\nu| = 8$ , but the initial directions are chosen randomly based on a uniform distribution. Conservation of linear momentum is used to determine particle locations when a collision occurs with another particle or with the boundary walls.



**Figure 1.1** Model 1. (a) Continuous model configuration at  $t = 0$ . (b) Illustration of surrogate lattice model. (c) State transition probabilities for building a lattice-based surrogate model.

Within the continuous model, we track the state of each subdomain and its state changes in order to investigate the time evolution of the number of particles per subdomain and identify a stationary response. We create a time series recording the number of particles

(state) for all subdomains. Using this time series, we can identify all state changes within the system. We then determine the probabilities for transitions between the number of particles per subdomain at any given time. These probabilities are specialized for each subdomain type, so we pool the probabilities by subdomain type. We consider  $M$  ( $\leq 27$ ) the maximum number of particles present in a subdomain at any particular time,  $\alpha = L, I, C$  the subdomain type, and  $j = 0, \dots, M$  the state of a particular subdomain. Then  $p_\alpha^{j+}$  is the probability that subdomain  $\alpha$  gains a particle when it is in state  $j$ ;  $p_\alpha^{j-}$  is the probability that subdomain  $\alpha$  loses a particle when it is in state  $j$ . The probability that a subdomain of type  $\alpha$  remains in state  $j$  is:

$$p_\alpha^{j*} = 1 - p_\alpha^{j+} - p_\alpha^{j-}, \quad \alpha = L, I, C, \quad j = 0, \dots, N_s, \quad (1.1)$$

where  $N_s$  ( $\leq M$ ) is the maximum number of particles per subdomain to be accounted for via the states in the Markov chain model (Fig. 1.1 (c)). From these probabilities, we can assemble tridiagonal transition matrices for the Markov chain model:  $P_\alpha$  where  $\alpha = L, I, C$ . The stationary distributions of the three types of transition probability vectors ( $\vec{\pi}_L, \vec{\pi}_I, \vec{\pi}_C$ ) are computed as the solutions to the linear algebraic equations:

$$\vec{\pi}_C = \vec{\pi}_C P_C, \quad \vec{\pi}_I = \vec{\pi}_I P_I, \quad \vec{\pi}_L = \vec{\pi}_L P_L. \quad (1.2)$$

The solutions of (2.12) constitute a surrogate Markov chain model for the directly simulated continuous model.

## 1.2 Single particle random walks in a periodic representative volume element

Diffusion is a key transport mechanism in biological systems. Biopolymer structure affects the diffusivity of solutes within biological tissues; however, the relationships between biopolymer structure of the tissue and diffusivity are not well understood. For example, in cartilage tissue engineering applications, native cells (or stem cells) are seeded into biomaterial scaffolds [11, 20]. The cells need nutrients and growth factors to both sustain themselves and further construct the extracellular matrix via cellular uptake of these solutes [24]. With the prevalence of osteoarthritis, this method of cell therapy needs further development [11]. Ideally, replacement scaffolds should have similar properties to native cartilage tissue [27]. A wide variety of scaffold materials are used (e.g., gels, native biopolymers, woven fibers), and thus the scaffold can have many different microscopic features, properties, shapes,



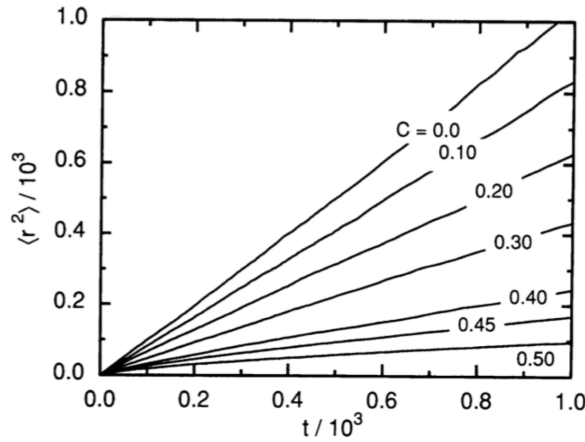
and sizes that all affect the diffusivity of the system [9, 21, 36]. Understanding the scaffold micro-architecture is important in developing successful treatments to repair articular cartilage [23]. Diffusive transport strongly affects design outcomes for the engineered tissue and is also relevant for understanding biotransport in native tissues [7].

Previous studies have investigated diffusion via direct simulation by considering single particle movements on a lattice [32, 35]. In this thesis, we consider random walks on a continuous domain, a more general case enabling incorporation of a variety of local geometric features of a material. A surrogate lattice-based model will be built by using transition probabilities observed in the continuous model simulations, based on identifying subdomains tailored to the geometry of the simulation domain.

Normal diffusion is the term used to describe diffusion that occurs on a homogenous domain. The mean-squared displacement is calculated by averaging displacement trajectories over a large number of realizations (e.g., typically 2000 realizations for this study). In normal diffusion, the relationship between mean-squared displacement,  $\langle r^2 \rangle$ , and the diffusion constant,  $D$ , is

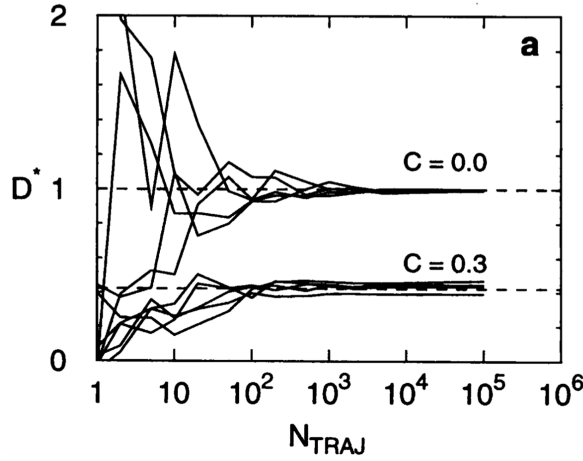
$$\langle r^2 \rangle = 4D t, \quad (1.3)$$

where  $t$  is the time. [32, 31, 35]



**Figure 1.2** Single particle random walks on a 2D lattice with varying obstacle concentrations. Mean squared displacement,  $\langle r^2 \rangle$  is linearly with time. Reprinted from Biophysical Journal, Vol 66, M.J. Saxton, "Anomalous diffusion due to obstacles: a Monte Carlo Study", Page No. 396, Copyright 1994, with permission from Elsevier [32].

In prior lattice studies of single particle random walks, it was shown that the associated mean-squared displacement grows linearly with time, for a large number of time steps. [32].



**Figure 1.3** Single particle random walks on a 2D lattice with varying obstacle concentrations. The estimate for normalized diffusivity,  $D^*$ , depends upon the number of realizations included in the average of mean squared displacement. Reprinted from Biophysical Journal, Vol 72, M.J. Saxton, "Single-particle tracking: the distribution of diffusion coefficients", Page No. 1745, Copyright 1997, with permission from Elsevier [31].

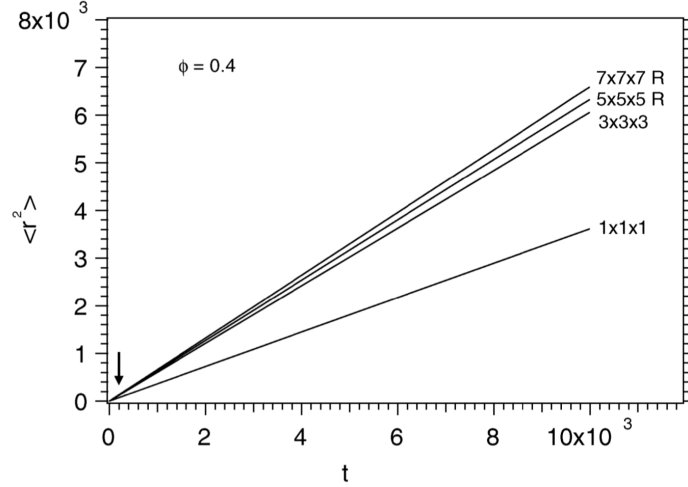
*Saxton* considered simulations performed on a 2D lattice, with varying obstacle concentrations,  $C$  [32]. Typically, *Saxton* averaged between  $10^4$  and  $10^5$  displacement trajectories in order to calculate the mean-squared displacement [31]. Figure 1.2 shows the mean squared displacement as a function of time. In this figure, the obstacle concentration  $C$  varies from 0.0 to 5.0. For the no obstacle case,  $C = 0.0$ , we observe that the curve for mean squared displacement versus time has the greatest slope and that the particle traveled the furthest relative to its initial location. In this work, simulations were also performed on a 2D lattice, and diffusivity was estimated according to equation 1.3. Figure 1.3 shows the normalized diffusivity,  $D^*$ , which is scaled such that  $D^* = 1$  in the no obstacle case ( $C = 0.0$ ). For an increasing number of realizations (or trajectories),  $N_{\text{TRAJ}}$ , the values of  $D^*$  are observed to reach stationary values (Fig. 1.3)

*Vilaseca* considered obstructed diffusion on a three-dimensional lattice in order to study movement in crowded media, showing that obstacles lead to anomolous diffusion for early time steps [35]. Anomolous diffusion is described by the equation

$$\langle r^2(t) \rangle = (2d)\Gamma t^\alpha, \quad (1.4)$$

where  $d$  is the dimension,  $\Gamma$  is a generalized transport coefficient (anomolous diffusion coefficient), and  $\alpha$  is the anomolous diffusion exponent. With sufficient timesteps,  $\alpha = 1$  and normal diffusion is also observed in this case. Figure 1.4 illustrates the relationship

between  $\langle r^2 \rangle$  versus time for a large number of timesteps, such that the relationship is linear and the system exhibits normal diffusion. When  $\alpha = 1$ , we can use equation 1.3 to study the diffusivity of the system.



**Figure 1.4** Plot of  $\langle r^2 \rangle$  versus time in 3D systems studied in *Vilaseca*, showing that, for large timesteps, the relationship between  $\langle r^2 \rangle$  versus time is linear. Reproduced from Ref. [35] with permission from the PCCP Owner Societies.

In developing Model 2 of this thesis, continuous simulations are analyzed for a sufficient number of timesteps such that the system exhibits normal diffusion. We expand upon on previous studies by tracking different types of transition probabilities, depending on the obstacle arrangement and size. We are then able to build corresponding surrogate lattice models. In the continuous model, a particle can move in any direction  $\theta \in [0, 2\pi]$  with a spatial stepsize  $\Delta x$  at each timestep  $\Delta t$ . We consider the continuous domain as being composed of subdomains, and we track transition probabilities between these subdomains. The resulting lattice models more directly relate to the previous studies of diffusivity, as discussed above. Depending on the model studied, we track between two and twelve different types of transitions.

### 1.2.1 Model descriptions

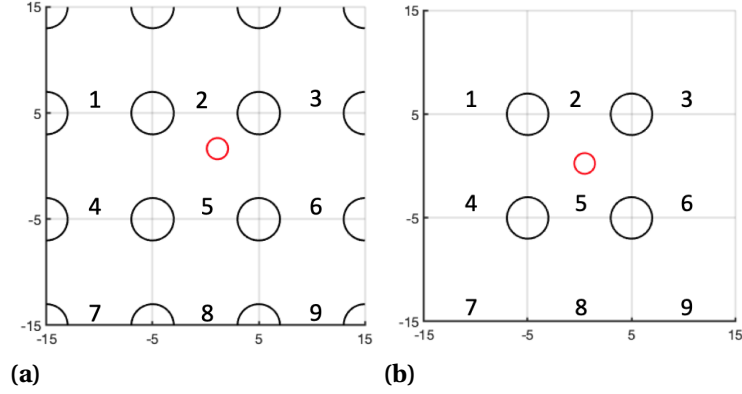
In the continuous model, we simulate a single circular particle moving via random walks on a 2D domain with obstacles. Simulations are performed on a representative volume element (RVE) with periodic boundary conditions (BCs). Fixed obstacles are added to the system, and all collisions between the particle and obstacles are assumed to be perfectly

elastic. In this model, our states are the subdomain in which the particle is located at a given point in time. We model a random walk by having the particle choose a random new direction at each new timestep with a prescribed fixed spatial step size. Over long time periods and many realizations, continuous model results are used to calculate stationary values for the probability of state transitions as a particle moves to an adjacent subdomain.

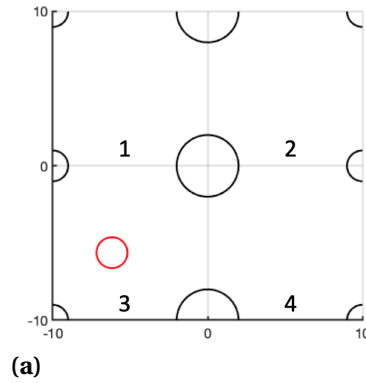
We aim to consider three different obstacle arrangements: (i) four circular obstacles located completely inside (internal to) the RVE, (ii) an arrangement of full and partial obstacles on the RVE that generates a regular double-periodic medium, and (iii) equally spaced obstacles such that the middle column of obstacles is varied in size relative to the other obstacles. Note that model (ii) represents a more isotropic material as compared to models (i) and (iii). These obstacle arrangements are visualized in Figs. 1.5-1.6. In developing this model we aim to: (a) analyze the effects of particle size, obstacle size, and obstacle arrangement on the diffusivity in 2D RVEs, (b) perform direct simulations using continuous off-lattice models to obtain baseline probabilities for surrogate lattice-based models, and (c) use baseline probabilities to develop and analyze a surrogate 2D lattice model on a 3x3 or a 2x2 periodic lattice. Previous studies have used random walks on a lattice model to investigate properties of diffusivity, but, to our knowledge, none have made connections to simulations on a continuous domain. These studies have identified a linear relationship between the mean squared displacement  $\langle r^2 \rangle$  and time [32], [31], where the constant of proportionality depends on the underlying structure of the domain. These studies have primarily been on two or three dimensional lattices (i.e., discrete domains), while the models proposed here are developed on a continuous domain with the lattice model serving as a surrogate model. We use continuous model simulations and surrogate lattice model results to study diffusive transport in varying domain geometries.

In this dissertation we use continuous model simulation results to develop probabilistic surrogate models. In Chapter 2 we present Model 1, which models a two-dimensional interacting particle system. The particles have perfectly elastic collisions with other particles and with domain boundaries. We use results from the continuous model to develop a surrogate Markov chain model using state change probabilities. We analyze and compare continuous and surrogate model results for varying particle radii. Model 1 can be thought of as an initial motivating problem for Model 2.

In Chapter 3 we introduce Model 2, which models single particle random walks in a two-dimensional RVE with circular obstacles. We have three different obstacle configurations: equally spaced obstacles, four internal obstacles, and multisize obstacles. For each obstacle configuration we compare results for varying obstacle radii. We first describe the continuous models, and then describe how we use continuous results to identify transition probabilities



**Figure 1.5** Model 2. Domains for deterministic model with numbered subdomains for particle radius  $R_{\text{part}} = 1.0$  and obstacle radius  $R_{\text{obs}} = 2.0$ . (a) Equally spaced obstacles model. (b) Four internal obstacles model.



**Figure 1.6** Model 2. Domain for deterministic model of multisize obstacles with numbered subdomains for particle radius  $R_{\text{part}} = 1.0$ , the first set of obstacle radius  $R_{\text{obs}_1} = 1.0$  and the second set of obstacle radius  $R_{\text{obs}_2} = 2.0$ .

for the surrogate lattice-based models. In Chapter 4 we discuss model calibration and diffusivity estimation. We develop a parameter referred to as the *commitment index*,  $M_c$ , which is used to determine when a particle commits to a new subdomain. We also discuss computation of effective diffusivities from simulation results, as well as determining and comparing this to a theoretical estimate of the diffusivity. In Chapter 5 we discuss results from Model 2, beginning with continuous model simulations and then discussing surrogate model results for all obstacle configurations and obstacle radii values considered. We then compare the results for the three different obstacle configurations and their diffusivities. We end Chapter 5 with overall conclusions of this dissertation.

## CHAPTER

## 2

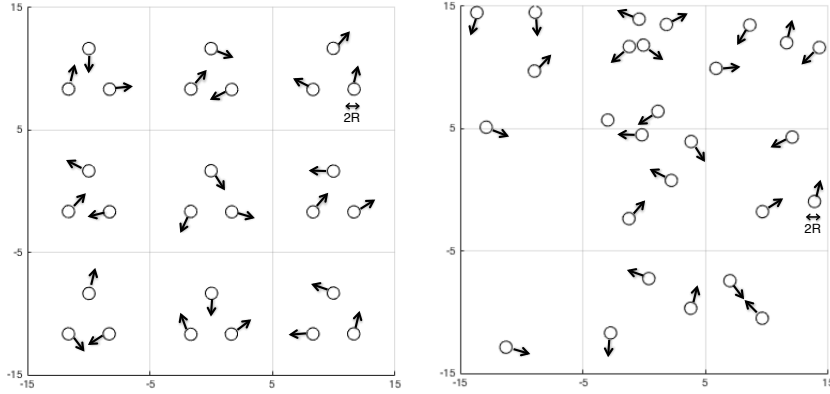
# TWO-DIMENSIONAL INTERACTING PARTICLE SYSTEM: DEVELOPMENT OF A MARKOV CHAIN MODEL

### **2.1 Model descriptions**

This chapter describes the development of Model 1, as initially outlined in Chapter 1. Model 1 directly simulates a system of interacting circular particles within a square domain. These particles undergo perfectly elastic collisions with boundary walls and with other particles, and the continuous simulation results are used to develop a surrogate Markov chain model. We first describe the continuous model used to directly simulate particle interactions (Sec. 2.1.1) and then outline how the results can be used to develop a surrogate Markov chains model, to accurately capture the system dynamics in an alternate manner (Sec. 2.1.2).

### 2.1.1 Continuous Model of Particle Interactions

We consider a square domain in two spatial dimensions with four rigid walls. Within this domain, 27 circular particles having identical mass and radius ( $R$ ) are undergoing motion involving dynamic collisions with the other particles and with the four walls of the domain (Fig. 2.1-right). To facilitate development of the surrogate model (Sec. 2.1.2), the square domain is viewed as the union of nine, non-overlapping square subdomains of equal area. At  $t = 0$ , the 27 particles are partitioned into nine groups of three particles each and initially placed, as illustrated, into each of the nine subdomains (Fig. 2.1-left). The initial direction of motion for each particle is chosen randomly based on a uniform distribution and prescribed a constant initial speed  $|\mathbf{v}|$ . Once the system is set into motion ( $t > 0$ ), the particles are assumed to exhibit perfectly elastic collisions with each other and with the four rigid walls of the square domain, i.e. all collisions conserve linear momentum.

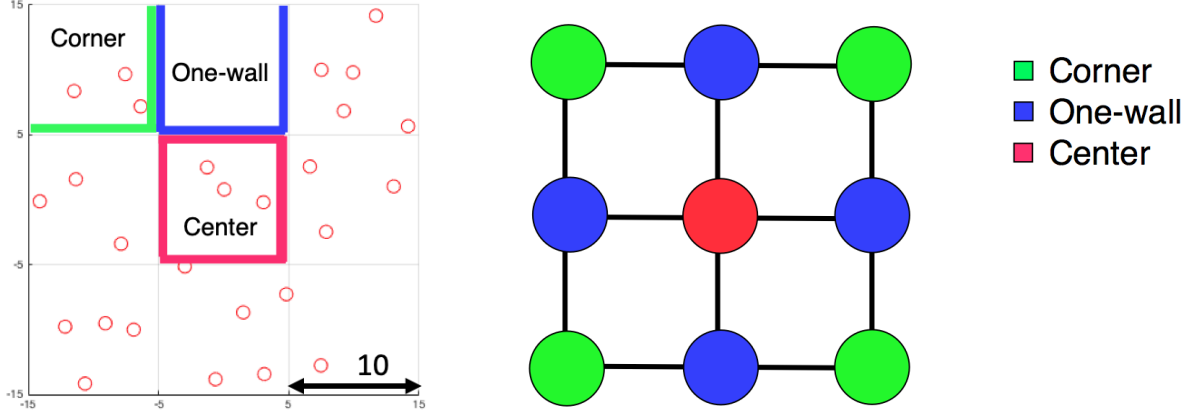


**Figure 2.1** Two dimensional dynamic model for perfectly elastic collisions between 27 particles and four rigid walls in a square domain: (left) the initial state ( $t = 0$ ) for a sample realization, (right) an intermediate state ( $t > 0$ ).

Our goal is to investigate time evolution of the number of particles per subdomain and delineate the stationary response of this quantity of interest based on the subdomain type and particle radius  $R$ . The three subdomain types are referred to as *Corner* (two exterior walls), *One-wall* (one exterior wall) and *Center* (no exterior walls) (Fig. 2.2-left). The entire simulation domain thus consists of four Corner subdomains, four One-wall subdomains, and one Center subdomain, along with adjacency relationships (Fig. 2.2-right).

Based on the system characteristics described above, we can make the following observations:





**Figure 2.2** Partitioning of the direct simulation domain into 9 square subdomains of equal area: (left) subdomains are delineated into three types based on the number of subdomain boundaries, (right) illustration of adjacency relationships between the 9 subdomains.

1. As the particle radius  $R \rightarrow 0$ , the expected value of the number of particles per subdomain in the stationary response should tend to 3.
2. As the particle radius  $R$  increases above zero, peripheral exclusion zones along the four domain walls limit the area that is available to be occupied by each particle in the Corner and One-wall subdomains.
3. Observation#2 implies that the expected value for the number of particles per subdomain should be greatest in the Center subdomain, smaller in the One-wall subdomains, and smallest in the Corner subdomains.

While these three observations can be made about this interacting particle system, the precise manner in which expected values of the number of particles per subdomain in the stationary response vary with both  $R$  and the subdomain type is not readily ascertained, a priori.

### 2.1.2 Markov Chain Model

We aim to identify and quantify system properties of the directly simulated continuous model via time series  $\eta_i(t_k)$  ( $i = 1, \dots, 9$ ,  $t_k = k\Delta t$ ,  $k = 0, \dots, N$ ), that track the number of particles per subdomain, where  $N$  is the total number of time steps. A particle is counted as being located in the  $i$ th subdomain at time  $t_k$  if its center is located in that subdomain at time  $t_k$ . Introduce a second time series  $c_i(t_k) = \eta_i(t_k) - \eta_i(t_{k-1})$  ( $i = 1, \dots, 9$ ,  $k = 1, \dots, N$ ). Assuming that  $\Delta t$  is sufficiently small, this second time series will take one of three values,

i.e.,

$$c_i(t_k) = \begin{cases} 0 & \text{if } \eta_i(t_k) = \eta_i(t_{k-1}) \\ -1 & \text{if } \eta_i(t_k) = \eta_i(t_{k-1}) - 1 \\ +1 & \text{if } \eta_i(t_k) = \eta_i(t_{k-1}) + 1 \end{cases}, \quad i = 1, \dots, 9, k = 1, \dots, N. \quad (2.1)$$

When  $\Delta t$  is small, a state change in which the number of particles in a subdomain changes by more than one particle is a rare event. Such instances are tracked in the implementation and adverse affects can be remedied by decreasing  $\Delta t$  until the resulting state changes universally obey (2.1).

The following indicator function is also introduced to determine when the  $i$ th subdomain contains  $j$  particles at time  $t_k$ :

$$\gamma_i^j(t_k) = \begin{cases} 1 & \text{if } \eta_i(t_k) = j \\ 0 & \text{if } \eta_i(t_k) \neq j \end{cases}, \quad i = 1, \dots, 9, j = 0, \dots, M, k = 1, \dots, N, \quad (2.2)$$

where  $M (\leq 27)$  is the maximum number of particles present in a subdomain at any particular time. Two additional indicator functions are then used to determine when the  $i$ th subdomain (containing  $j$  particles at time  $t_k$ ) gained or lost (respectively) one particle from a neighboring subdomain as time advanced from  $t = t_{k-1}$  to  $t = t_k$ :

$$\beta_i^{j+}(t_k) = \begin{cases} \gamma_i^j(t_k) & \text{if } c_i(t_k) = +1 \\ 0 & \text{if } c_i(t_k) \neq +1 \end{cases}, \quad i = 1, \dots, 9, j = 0, \dots, M, k = 1, \dots, N, \quad (2.3)$$

$$\beta_i^{j-}(t_k) = \begin{cases} \gamma_i^j(t_k) & \text{if } c_i(t_k) = -1 \\ 0 & \text{if } c_i(t_k) \neq -1 \end{cases}, \quad i = 1, \dots, 9, j = 0, \dots, M, k = 1, \dots, N. \quad (2.4)$$

The quantities in (2.2)-(2.4) are then used to determine probabilities for transitions between the number of particles per subdomain (states) as now outlined. Let  $G_i^j$  track the number of time steps for which the  $i$ th subdomain contains  $j$  particles during the time interval  $[t_1, t_N]$ :

$$G_i^j = \sum_{k=1}^N \gamma_i^j(t_k), \quad i = 1, \dots, 9, j = 0, \dots, M. \quad (2.5)$$

Then, let  $\delta_i^{j+}$  and  $\delta_i^{j-}$  count the number of time steps where the  $i$ th subdomain is in state  $j$  and gains or loses (respectively) one particle:

$$\delta_i^{j+} = \sum_{k=1}^N \beta_i^{j+}(t_k), \quad \delta_i^{j-} = \sum_{k=1}^N \beta_i^{j-}(t_k), \quad i = 1, \dots, 9, j = 0, \dots, M. \quad (2.6)$$

Using (2.5)-(2.6), it then follows that the probabilities of the  $i$ th subdomain gaining or

losing a particle when in a particular state  $j$  are, respectively:

$$\gamma_i^{j+} = \frac{\delta_i^{j+}}{G_i^j}, \quad \gamma_i^{j-} = \frac{\delta_i^{j-}}{G_i^j}, \quad i = 1, \dots, 9, \quad j = 0, \dots, M. \quad (2.7)$$

These probabilities are now specialized to the subdomains depicted in Fig. 2.2, i.e. four corner ( $L$ ) domains ( $i = 1, 3, 7, 9$ ), four one-wall ( $I$ ) domains ( $i = 2, 4, 6, 8$ ) and one center ( $C$ ) domain ( $i = 5$ ). Pooled probabilities for each of the three subdomain types are calculated as:

$$p_L^{j+} = \frac{\sum_{i=1,3,7,9} \delta_i^{j+}}{\sum_{i=1,3,7,9} G_i^j}, \quad p_I^{j+} = \frac{\sum_{i=2,4,6,8} \delta_i^{j+}}{\sum_{i=2,4,6,8} G_i^j}, \quad p_C^{j+} = \frac{\delta_5^{j+}}{G_5^j}, \quad j = 0, \dots, M, \quad (2.8)$$

for gaining a particle and as:

$$p_L^{j-} = \frac{\sum_{i=1,3,7,9} \delta_i^{j-}}{\sum_{i=1,3,7,9} G_i^j}, \quad p_I^{j-} = \frac{\sum_{i=2,4,6,8} \delta_i^{j-}}{\sum_{i=2,4,6,8} G_i^j}, \quad p_C^{j-} = \frac{\delta_5^{j-}}{G_5^j}, \quad j = 0, \dots, M, \quad (2.9)$$

for losing a particle.

By pooling data from (2.8)-(2.9) for each of the three subdomain types, the probability that a subdomain of type  $\alpha$  remains in state  $j$  in the Markov chain model is:

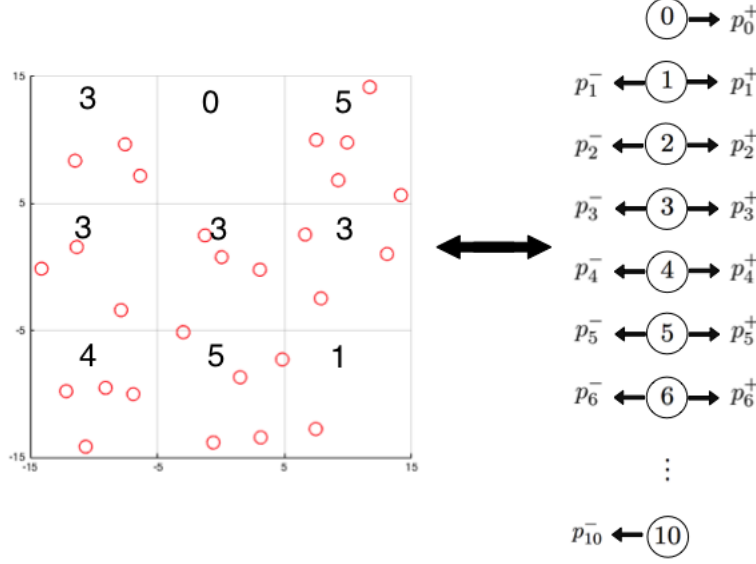
$$p_\alpha^{j*} = 1 - p_\alpha^{j+} - p_\alpha^{j-}, \quad \alpha = L, I, C, \quad j = 0, \dots, N_s, \quad (2.10)$$

where  $N_s (\leq M)$  is the maximum number of particles per subdomain to be accounted for via the states in the Markov chain model. As  $N_s$  increases past 8-9 particles per subdomain, the occurrence of such states in a subdomain becomes quite rare. Hence, the model is calibrated to choose a value for  $N_s$  beyond which the associated results exhibit little sensitivity to further increasing  $N_s$ . For illustration, the process of tracking state transitions in the case  $N_s = 10$  is shown in Fig. 2.3.

The probabilities in equations (2.8)-(2.10) are used to assemble three tridiagonal transition matrices for the Markov Chain model:

$$P_\alpha = \begin{bmatrix} p_\alpha^{0*} & p_\alpha^{0+} & 0 & \dots & 0 \\ p_\alpha^{1-} & p_\alpha^{1*} & p_\alpha^{1+} & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & p_\alpha^{(N_s-1)-} & p_\alpha^{(N_s-1)*} & p_\alpha^{(N_s-1)+} \\ 0 & \dots & 0 & p_\alpha^{N_s-} & p_\alpha^{N_s*} \end{bmatrix}, \quad \alpha = L, I, C. \quad (2.11)$$

Lastly, stationary distributions of the three types of transition probability vectors ( $\vec{\pi}_C, \vec{\pi}_I,$



**Figure 2.3** Illustration of the Markov chain model for tracking transition probabilities in the number of particles per subdomain at a snapshot in time when  $N_s = 10$ . Note that transition probabilities are calculated separately for each of the three subdomain types  $\alpha = L, I, C$  using equation (2.10).

$\vec{\pi}_L$ ) are computed as the solutions to the linear algebraic equations:

$$\vec{\pi}_C = \vec{\pi}_C P_C, \quad \vec{\pi}_I = \vec{\pi}_I P_I, \quad \vec{\pi}_L = \vec{\pi}_L P_L. \quad (2.12)$$

The solutions of (2.12) then, effectively, constitute a surrogate Markov chain model for the directly simulated continuous model. These solutions also facilitate a detailed investigation of both the expected values and uncertainty for the number of particles per subdomain in the stationary regime as both the subdomain type and the particle radius are varied.

## 2.2 Results

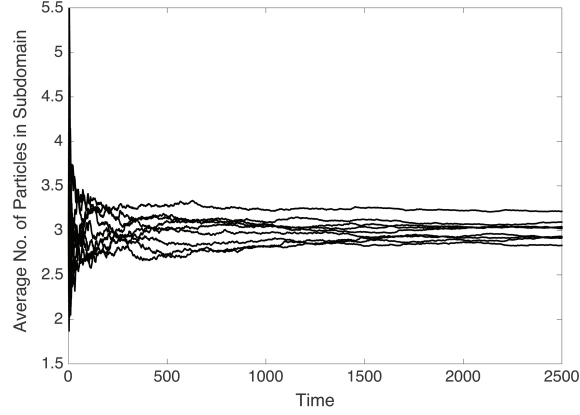
We first illustrate some key properties of the continuous model used to directly simulate particle interactions and discuss its calibration (Sec. 2.2.1). Results are then compared to those obtained using the surrogate Markov chain model, and also used to study the stationary response for the number of particles per subdomain, quantifying both expected values and uncertainty with increasing particle radius (Sec. 2.2.2).

### 2.2.1 Continuous Model Results

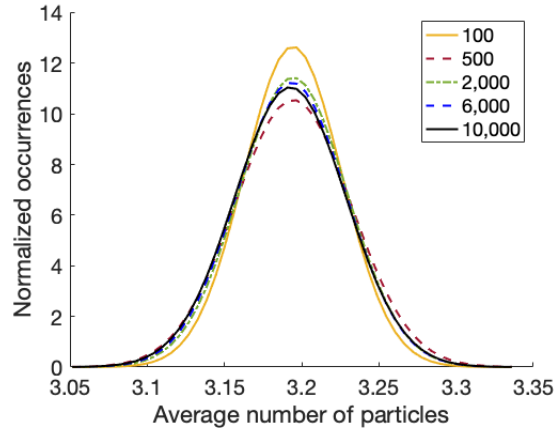
To quantify uncertainty in the continuous model, simulations were run over a large number of realizations of the initial particle configuration on a square of side length 30. This length scale was normalized relative to a particle radius  $R$ , i.e. with  $0 < R \leq 1$ . A large number of time steps was also required to be in a regime where the average number of particles per subdomain exhibited a stationary response. This process was then repeated as the particle radius  $R$  was varied in the range  $[0.1, 0.9]$  to determine effects of particle size on the quantity of interest. For each realization, each of the nine subdomains initially contained three particles. These initial particle positions were not varied across realizations. Particles were prescribed random initial velocities that were fixed in the sense that they were all prescribed the same initial speed  $|\mathbf{v}| = 8$ , but their initial directions were drawn from a uniform distribution. At each time step, the positions of all 27 particles were updated and the center of each particle was used to determine if the subdomain in which the particle resided at the prior time step had changed. Conservation of linear momentum was used to determine particle locations when a collision with another particle or with one of the four rigid walls occurred over the duration of one time step. The resulting set of particle locations was then used to compute the time series  $\eta_i(t_k)$  in Sec. 2.2 that track the number of particles per subdomain as time advances.

It was determined that a value of  $N = 2 \times 10^5$  (time steps), corresponding to  $\Delta t = 0.0125$ , was sufficient for yielding data exhibiting the stationary properties needed to build the surrogate Markov chain model. This value of  $N$  was used in all subsequent simulations. The stationary nature of the average number of particles in each of the nine subdomains is illustrated for a single realization in Fig. 2.4 in the case  $R = 0.5$ . Mean values of  $\bar{\eta}_\alpha$  ( $\alpha = L, I, C$ ) of  $\eta_i(t_k)$  ( $t_k = k\Delta t, k = 1, \dots, N$ ) were obtained by averaging over all times and then pooling data for each of the three subdomain types, i.e. corner ( $\alpha = L \leftrightarrow i = 1, 3, 7, 9$ ), one-wall ( $\alpha = I \leftrightarrow i = 2, 4, 6, 8$ ) and center ( $\alpha = C \leftrightarrow i = 5$ ) across all realizations. As illustrated in Fig. 2.5, 6,000 realizations were sufficient to stabilize variation in this statistic for the quantity of interest. Specifically, in this regime the mean and standard deviation in the normal distribution stabilized their values to 3.19 and 0.036, respectively, for the case shown in Fig. 2.4. Consequently, 6,000 realizations were used in all subsequent simulations.

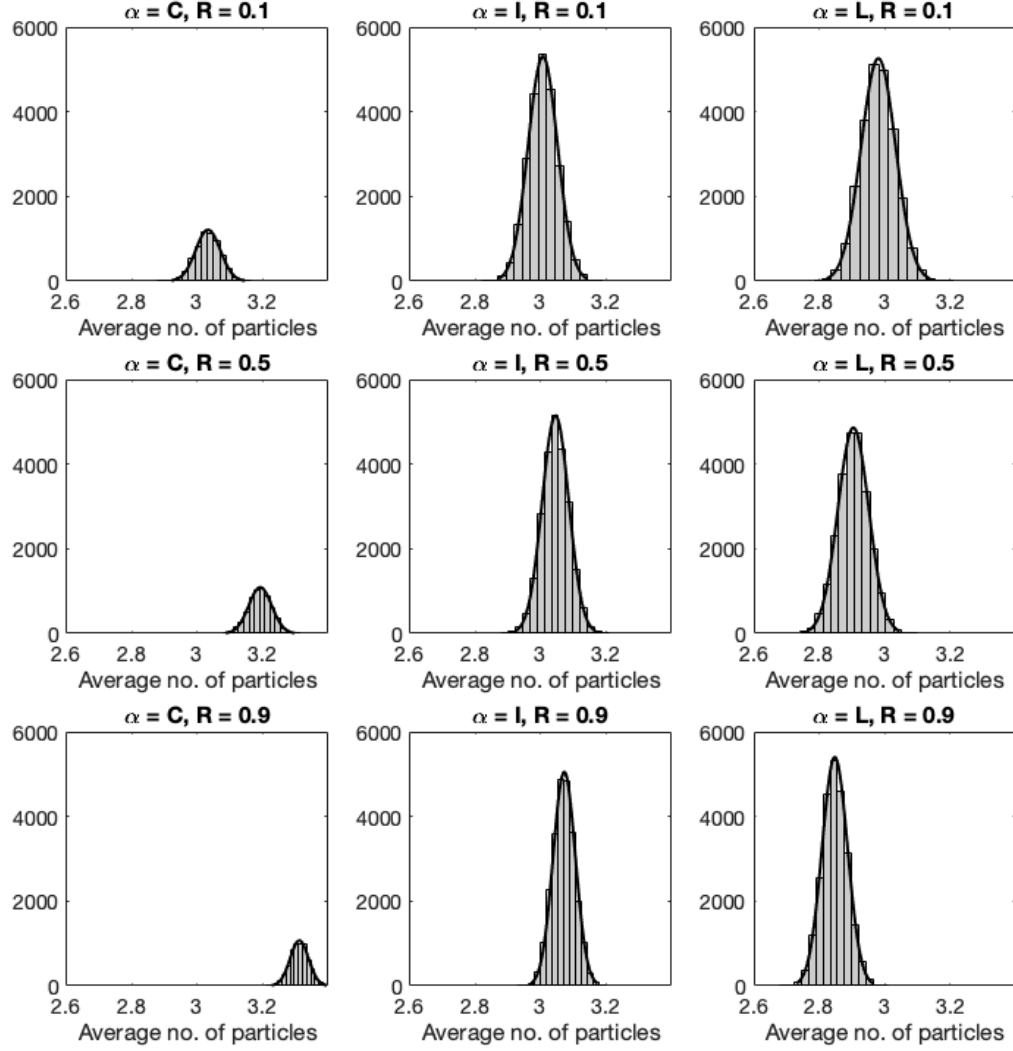
Via direct simulation using the continuous model, histograms for the average number of particles per subdomain  $\bar{\eta}_\alpha$  were determined as the particle radius  $R$  was varied between 0.1 and 0.9. Results pooled for each of the three subdomain types indicated that this statistic appeared to follow a normal distribution (Fig. 2.6). These results also illustrate observation 3 from Sec. 2.1, i.e. that the Center ( $\alpha = C$ ) subdomain (left column) has, on average, a greater



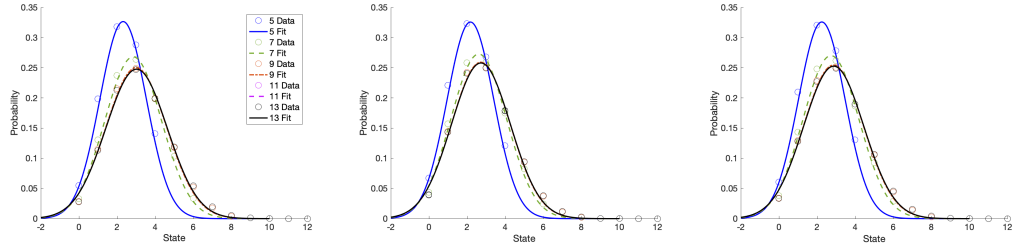
**Figure 2.4** Illustration of the stationary response for the average number of particles per subdomain in the case of 27 particles over the 9 subdomains shown in Fig. 2.2. Results are shown for a single realization in the case  $R = 0.5$  and  $\Delta t = 0.0125$  ( $N = 2 \times 10^5$ ).



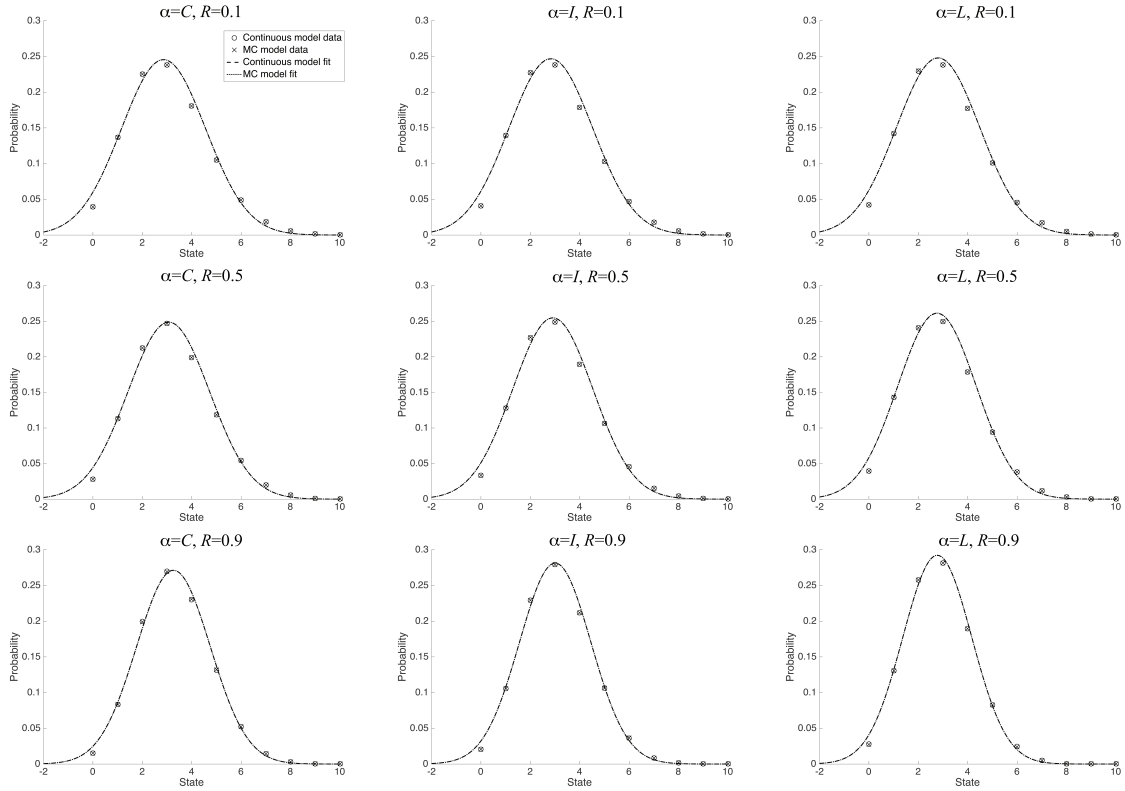
**Figure 2.5** Illustration of the effect on histograms for the number of particles per subdomain  $\bar{\eta}_C$  as the number of realizations is increased. Results are shown for the center subdomain in the case  $R = 0.5$ .



**Figure 2.6** Histograms and fitted normal distributions for mean values of the number of particles per subdomain  $\bar{n}_\alpha$  with varying subdomain type ( $\alpha = L, I, C$ ) and particle radius  $R$  obtained by direct simulation using the continuous model.



**Figure 2.7** Effects of varying the maximum number of states per subdomain  $N_s$  in the Markov chain model. Stationary distributions obtained from (2.12) were fit with the probability density function (2.13) subject to the constraint (2.15) in the case  $R = 0.5$ : (a) Center ( $\alpha = C$ ), (b) Corner ( $\alpha = L$ ), (c) One-wall ( $\alpha = I$ ).



**Figure 2.8** Distributions for the number of particles in each subdomain type ( $\alpha = L, I, C$ ) with varying particle radius  $R$ . Stationary distributions  $\tilde{\pi}_\alpha$  from the Markov Chain (surrogate) model are compared to data obtained using the (continuous) direct simulation model. The surrogate model was fit using the truncated normal distribution (2.13) subject to the constraint (2.15).



number of particles than the One-wall ( $\alpha = I$ ) subdomains (middle column) which, in turn, have a greater number of particles than the Corner ( $\alpha = L$ ) subdomains (right column). It is also observed (Fig. 2.6) that these differences become more pronounced as the particle radius  $R$  was increased from a value of 0.1 to a value of 0.9.

Overall, these results for the continuous model provide data, obtained via direct simulation into the stationary regime over many realizations, for formulation of the Markov chain (surrogate) model described in Sec. 2.2. They also illustrate fundamental statistical properties for the quantity of interest with respect to particle radius  $R$  and subdomain type  $\alpha = L, I, C$  that will also be used to evaluate accuracy of the surrogate model.

## 2.2.2 Markov Chain Model Results

Based on direct numerical simulations performed using the continuous model, the Markov Chain model was formulated by calculating the transition probabilities in (2.8)-(2.11). The matrices in (2.11) were then used to determine the three stationary distributions  $\vec{\pi}_\alpha, \alpha = L, I, C$  in (2.12). A qualitative analysis of the results indicated that a *truncated normal distribution* [18] was well-suited for fitting a probability density function to all three stationary distributions. This distribution is represented as:

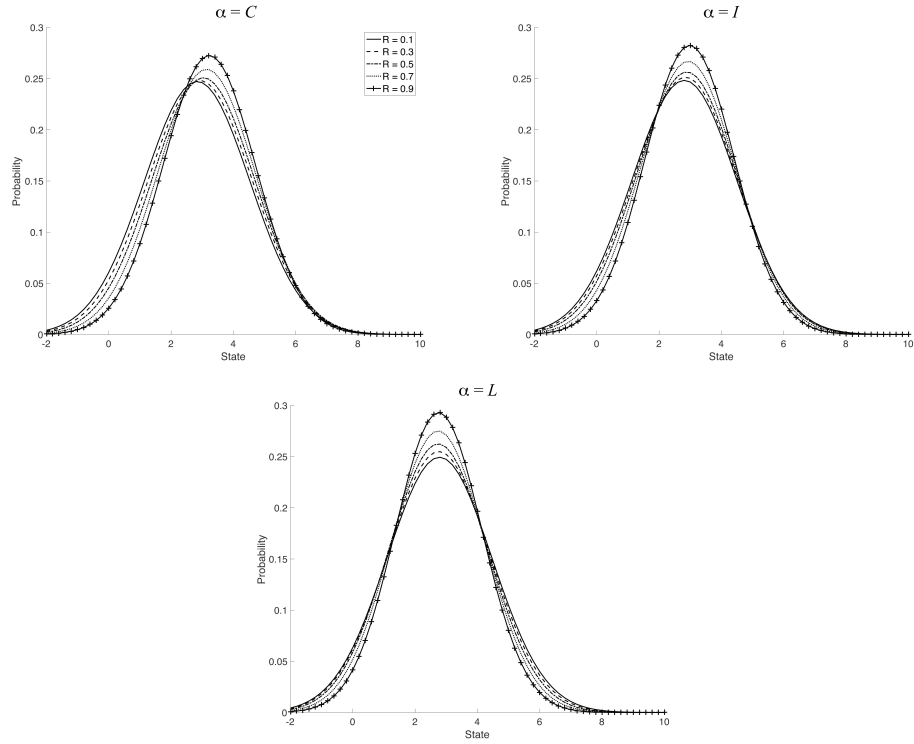
$$f(x; \mu, \sigma, a, b) = \frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{\sigma\left(\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right)}, \quad a \leq x \leq b \leq \infty, \quad (2.13)$$

where:

$$\phi(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right) \text{ and } \Phi(x) = \frac{1}{2} \left(1 + \exp\left(\frac{x}{\sqrt{2}}\right)\right). \quad (2.14)$$

In (2.14),  $\phi(\xi)$  and  $\Phi(x)$  are the probability density function and the cumulative distribution function of a standard normal distribution, respectively and  $[a, b]$  is the range of the random variable being considered.

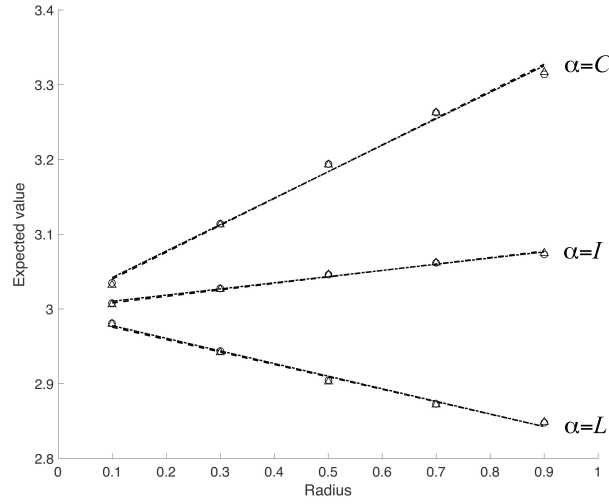
The bounds for the range of our random variable, which is the number of particles per subdomain (quantity of interest), were prescribed as  $[a, b] = [0, N_s]$ , recalling that  $N_s$  is the maximum number of states per subdomain in the Markov chain model. An appropriate choice for  $N_s$  was determined by increasing its value until curve fits of the truncated normal distribution to the stationary distributions in the Markov chain model  $\vec{\pi}_\alpha$  stabilized their shape. It should be noted that the occurrence of states becomes rarer as the number of particles per subdomain increases. The small probability values in  $\vec{\pi}_\alpha$  at these larger values make the process of fitting the truncated normal distribution quite sensitive to the method of curve-fitting. To address this challenge, the distributions were fit to the data using constrained optimization via the “fmincon” function in MATLAB (using the SQP



**Figure 2.9** Distributions for the number of particles in each subdomain grouped by subdomain type ( $\alpha = L, I, C$ ) to illustrate effects of varying particle radius  $R$ . Stationary distributions  $\vec{\pi}_\alpha$  from the Markov Chain (surrogate) model were fit using the truncated normal distribution (2.13) subject to the constraint (2.15).

**Table 2.1** Values of the parameters  $\mu$  and  $\sigma$  when the truncated normal distribution (2.13) subject to the constraint (2.15) was fit to the direct simulation (continuous) model and to the Markov chain (surrogate) model.

Continuous Model						
Radius	Center		One-wall		Corner	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
0.1	2.8578	1.7037	2.8287	1.6967	2.8002	1.6904
0.3	2.9632	1.6869	2.8681	1.6670	2.7754	1.6473
0.5	3.0708	1.6559	2.9104	1.6247	2.7545	1.5924
0.7	3.1758	1.5869	2.9603	1.5463	2.7554	1.5057
0.9	3.2613	1.4926	3.0074	1.4453	2.7682	1.3981
Markov Chain model						
Radius	Center		One-wall		Corner	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
0.1	2.8593	1.7050	2.8301	1.6989	2.8007	1.6927
0.3	2.9634	1.6881	2.8690	1.6686	2.7766	1.6496
0.5	3.0714	1.6573	2.9110	1.6260	2.7558	1.5944
0.7	3.1754	1.5873	2.9597	1.5469	2.7557	1.5071
0.9	3.2582	1.4922	3.0057	1.4457	2.7672	1.3991



**Figure 2.10** Linear regression fits for the relationship between the average number of particles per subdomain  $\bar{\eta}_\alpha$ ,  $\alpha = L, I, C$  and the particle radius  $R$ , delineated by subdomain type.

option) with incorporation of the following constraint:

$$\mu + \frac{\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} - \sum_{j=0}^{N_s} j \pi_\alpha^j = 0, \quad \alpha = L, I, C. \quad (2.15)$$

**Table 2.2** Slope, intercept and coefficient of determination ( $R^2$ ) for a linear regression analysis of the relationship between the average number of particles per subdomain  $\bar{\eta}_\alpha$ ,  $\alpha = L, I, C$  and the particle radius  $R$  for the continuous and the Markov chain models.

	Continuous			Markov Chain		
	Slope	Intercept	$R^2$	Slope	Intercept	$R^2$
Center	0.3539	3.0067	0.9929	0.3595	3.0039	0.9939
Onewall	0.0826	3.0021	0.9895	0.0868	2.9995	0.9921
Corner	-0.1688	2.9946	0.9928	-0.1663	2.9923	0.9903

Equation (2.15), with  $a = 0$  and  $b = N_s$ , expresses the condition that the mean of the truncated normal distribution (2.14) should equal the expected value for the Markov chain model when  $N_s$  states are included in the surrogate model.

An illustration of the process of refining  $N_s$ , including fits to the stationary distributions based on (2.13)-(2.15), is shown in Fig. 2.7. Based on this process, all subsequent results presented herein employed a value of  $N_s = 13$ . The resulting Markov chain stationary distributions ( $\vec{\pi}_\alpha$  from (2.12)), corresponding continuous model data, and fits to both using the truncated normal distribution (2.13) were determined for each subdomain type ( $\alpha = L, I, C$ ) and with varying particle radius  $R$  in the range 0.1-0.9 (Fig. 2.8). Excellent agreement is observed between stationary distributions calculated using the continuous model and the (surrogate) Markov chain model. The corresponding values of the estimated parameters  $\mu$  and  $\sigma$  for the fits shown in Fig. 2.8 (see Table 2.1) demonstrate excellent agreement between the continuous and surrogate (MC) model for all cases shown in Fig. 2.8. Effects of increasing the particle radius  $R$  for each of the three subdomain types resulted in significant shifts to the right in the distributions for the Center ( $\alpha = C$ ) and One-wall ( $\alpha = I$ ) subdomains, and much less pronounced variation for the Corner ( $\alpha = L$ ) subdomains (Fig. 2.9).

Lastly, a regression analysis was performed to determine empirical relations between the mean value of the number of particles per subdomain ( $\bar{\eta}_\alpha$ ) and the particle radius  $R$  for each of the three subdomain types ( $\alpha = L, I, C$ ) (Fig. 2.10, Table 2.2). Results demonstrated that linear regression provided excellent fits ( $R^2 > 0.98$ ) for all three subdomain types. In the Center ( $\alpha = C$ ) and Corner ( $\alpha = L$ ) cases, the slope and intercept in the regression fits for the continuous model and the (surrogate) Markov chain model agreed to within 1.6% relative error. In the case of the One-wall subdomains ( $\alpha = I$ ), the relative error was larger at a value of 5.1%, but it is noted that the slopes for this case are closer to zero as compared to the other two subdomain cases. The loss of available area as the particle radius  $R$  is increased is reflected in the negative slope for the Corner subdomains where the area available to

be occupied by particles is significantly reduced as  $R$  increases. This is countermanded by the relatively large positive slope for the Center subdomain for which no available area is lost with increasing particle radius. Conversely, as  $R \rightarrow 0$ , all three regression lines appear to converge to a value of 3.00 particles per subdomain, as evidenced by the intercept values in Table 2.2.

## 2.3 Discussion and Conclusions

The study considered in this chapter investigated a two-dimensional system of circular particles interacting, via perfectly elastic collisions, with each other and with the four walls of a square domain. By partitioning the simulation domain into 9 equal square subdomains, statistical properties of the system were delineated based on three subdomain types with differing geometric features. By taking the number of particles per subdomain as the quantity of interest, a surrogate model was formulated based on Markov chains. The states in the Markov chain model were the number of particles per subdomain, in each of the three subdomain categories, with transitions occurring between adjacent subdomains.

Excellent agreement between the directly simulated continuous model and the surrogate (Markov) model was achieved by tracking 14 states, i.e. significantly less than the total of 27 particles. Statistics for the quantity of interest indicated that a truncated normal distribution was well-suited to capturing statistical properties in estimates of the number of particles per subdomain type. Expected values were found to vary linearly with increasing particle radius  $R$ , and increased with  $R$  for subdomains with zero walls or one wall, but decreased for subdomains with two walls. Results of this type can be used to estimate the aggregate density of particles, via direct simulations in a representative domain, since it is unclear that such measures can be predicted a priori, except in the limit  $R \rightarrow 0$ . Indeed, the uniform density in this limiting case is consistent with Brownian motion (diffusion) on a square with insulated boundaries, yet the density profile as  $R$  increases is difficult to predict analytically. Beyond expected values, the modeling approach presented herein has the advantage that it includes an estimate of uncertainty via the use of truncated normal distributions that were demonstrated to be well-suited for analyzing the quantity of interest.

The approach employed in this study may have potential application in other systems where aggregate properties are unknown but can be investigated via direct simulations of discrete entities on a representative domain. Success will depend on the ability to accurately and efficiently compute solutions using the direct model on time scales and for enough realizations to exhibit stationary statistical properties for the quantities of interest. More specifically, development of a surrogate model using Markov chains will require that the

quantities of interest have a readily identified set of states and that the notion of transitions between these states can be easily defined. When states are continuous rather than discrete quantities, it is possible that a set of discrete states can still be defined by binning the random variable into adjoining sub-ranges of the continuous variable. The viability of this approach for such systems needs further investigation. When accurate Markov chain surrogate models can be developed, they also have the potential to serve as useful tools for accelerating portions of simulations for complex systems while also providing a quantitative framework for uncertainty quantification.

Overall, the methods and approach developed in this study for a simpler two-dimensional system may have potential utility in multiscale modeling of more complex systems with discrete entities (e.g. particles, polymers or biological cells) exhibiting non-trivial dynamical interactions that can be directly and efficiently simulated on a representative domain.

## CHAPTER

### 3

# SINGLE PARTICLE RANDOM WALKS IN A PERIODIC REPRESENTATIVE VOLUME ELEMENT

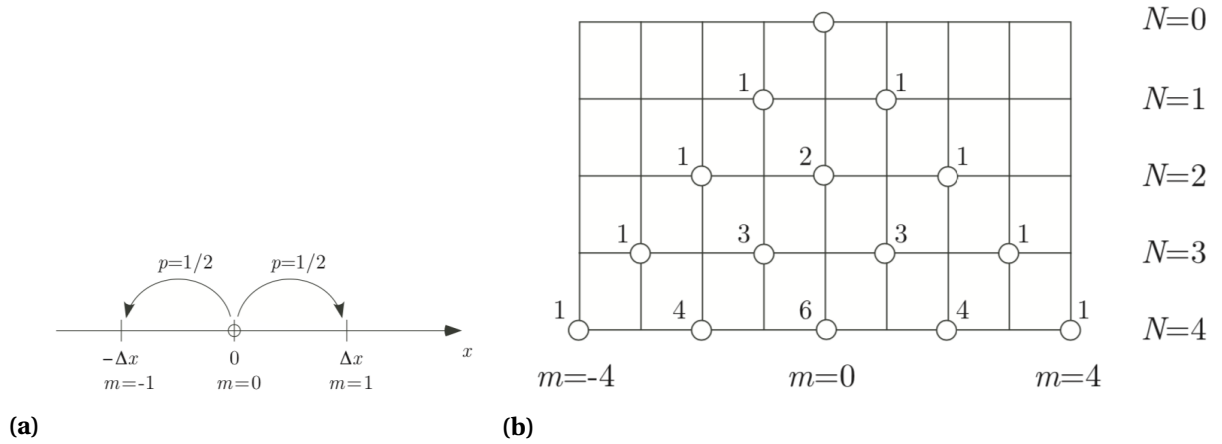
## 3.1 Introduction and Background

Moving on from the initial problem in Chapter 2 (Model 1), we now consider single particle random walks in a two-dimensional periodic representative volume element with internal obstacles (Model 2). We begin with an initial example of a particle with a random walk on a 1D lattice, which we use to motivate our 2D random walk models. In our continuous simulations, the particle can move in any direction in a 2D representative volume element (RVE). We use continuous model simulations to build transition probabilities for the surrogate model of random walks on a 2D lattice.

### 3.1.1 Example of random walks on a 1D lattice

We start by considering the simpler case of a single particle random walk on a 1D lattice. In this instance, the particle begins at location  $m = 0$  and takes a step either left or right

with equal probabilities (e.g. it may not stay in the same location), changing its location to  $m = -1$  or  $m = 1$ , respectively, after one time step. [15]. Figure 3.1 illustrates this scenario. When this experiment is repeated many times over a fixed number of time steps, the resulting histogram of spatial distances traveled for each random walk can be fit by a normal distribution (Fig. 3.2).

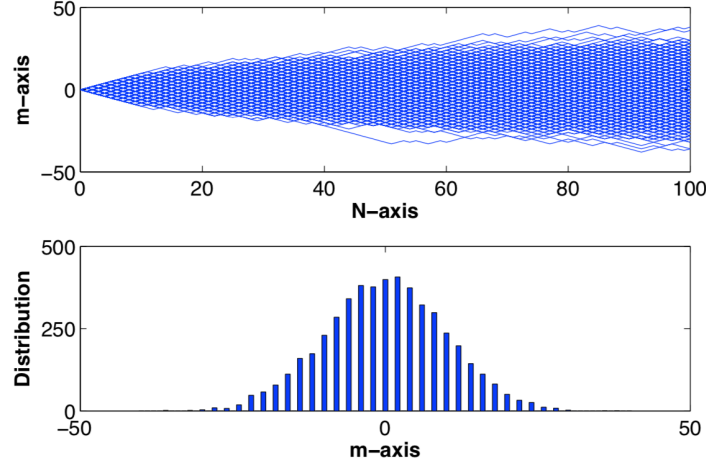


**Figure 3.1** Illustrations of a particle moving in a one-dimensional random walk, having equal probability of moving one unit to the right or left. (a) The particle's first step. (b) Molecule locations possible after  $N = 4$  steps. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer. Diffusion. In: *Foundations of Applied Mathematics* by M.H. Holmes, Springer Science + Business Media, LLC (2009) [15]

In contrast with the example given above, our simulations are in two dimensions and on a continuous domain. We first perform simulations on a continuous (off-lattice) domain from which we obtain baseline probabilities to build a surrogate 2D lattice model on a periodic lattice.

Compared to the continuous model, the periodic lattice model is more analogous to the example in Figure 3.1; however, in contrast to this example, the periodic lattice model allows for movement in two dimensions. We use the mean squared displacement to estimate the diffusivity for both the deterministic and surrogate models. We aim to analyze the effects of particle size, obstacle size, and obstacle arrangement on the diffusivity in a 2D representative volume element (RVE).



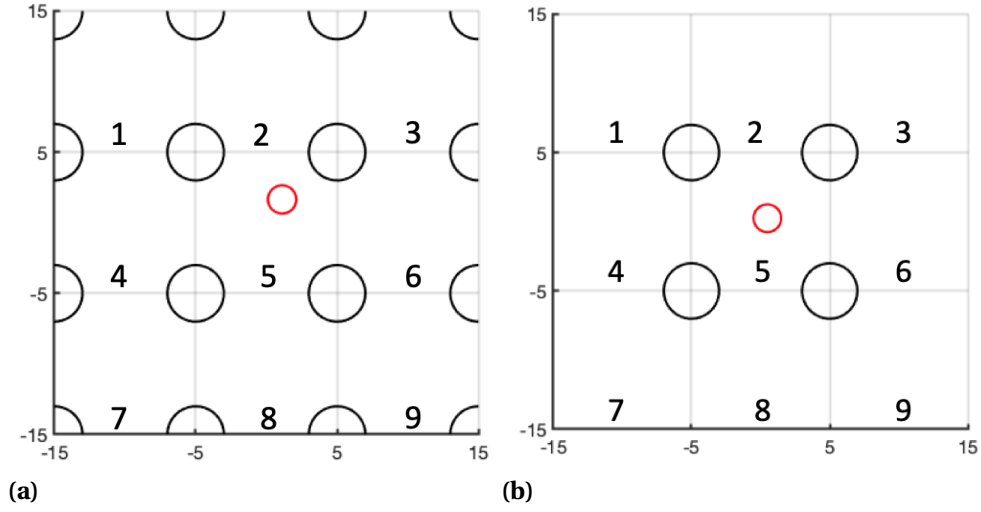


**Figure 3.2** Illustrations of 5000 realizations for a particle's random walk with  $N = 100$  time steps. Top: The different paths taken during the 5000 realizations. Bottom: Spatial distribution of the particle's location after the last timestep. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer. Diffusion. In: *Foundations of Applied Mathematics* by M.H. Holmes, Springer Science + Business Media, LLC (2009) [15].

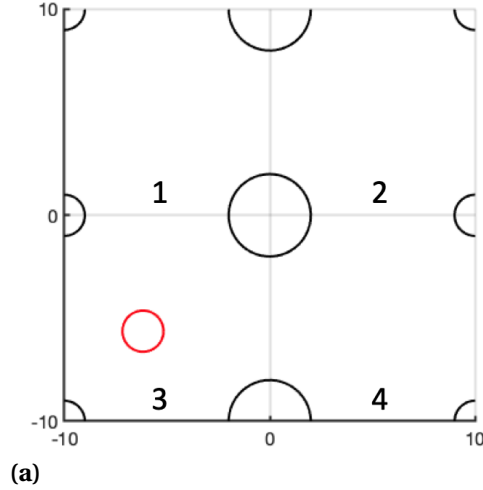
## 3.2 Continuous Model Description

We consider a single diffusing particle which moves via random walks in two spatial dimensions. In the continuous model, we directly simulate the particle's movement in a 2D periodic representative volume element (RVE). The domain has periodic boundary conditions (PBCs), and stationary obstacles with which the particle collides. The RVE has several different obstacle arrangements that serve as models for varying degrees of solid phase density (or volume fraction) and material anisotropy. Domains with greater symmetry will be useful in comparing to previous studies of diffusion on lattice-based models as well as theoretical relations for diffusion in porous media [32], [31], [35]. Figures 3.3 and 3.4 show the three different obstacle arrangements considered: (i) *equally spaced obstacles* (Fig. 3.3(a)), (ii) *four internal obstacles* (Fig. 3.3(b)), and (iii) *multisize obstacles* (Fig. 3.4). With these different arrangements, case (i) models the most isotropic of the materials, case (ii) models a material that is more transversely isotropic, and case (iii) introduces properties of a composite material.

We can divide the RVE into nine subdomains of equal size for cases (i) and (ii) and into four subdomains of equal size in case (iii). Due to differences in obstacle arrangements, subdomain types are most distinct (from other subdomain types) in the four internal obstacle case (case (ii)). For the four internal obstacles model, the domain can be divided into three distinct subdomain types – termed *Center* (C), *Onewall* (I), and *Corner* (L) – as



**Figure 3.3** Domains for deterministic model with numbered subdomains for particle radius  $R_{\text{part}} = 1.0$  and obstacle radius  $R_{\text{obs}} = 2.0$ . (a) Equally spaced obstacles model. (b) Four internal obstacles model.

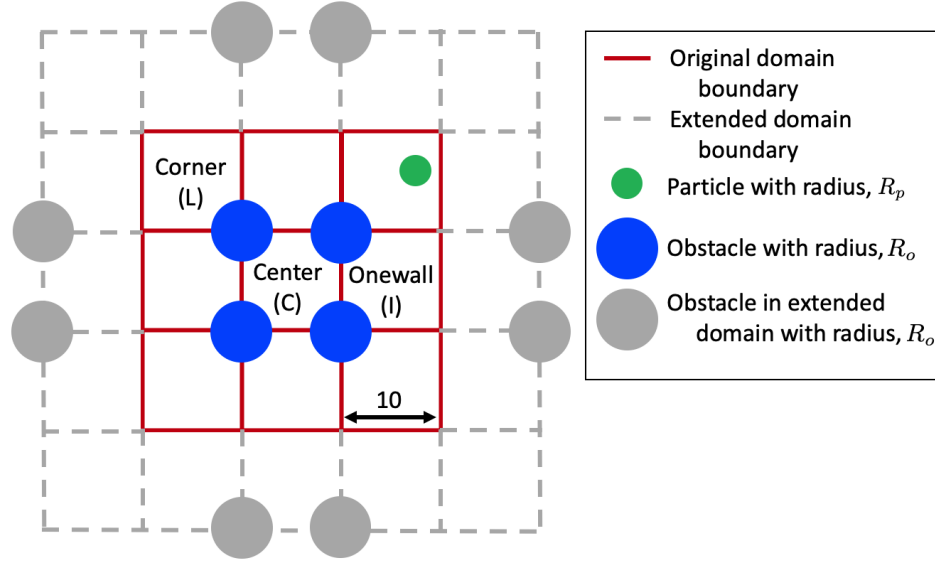


**Figure 3.4** Domain for deterministic model of multisize obstacles with numbered subdomains for particle radius  $R_{\text{part}} = 1.0$ , the first set of obstacle radius  $R_{\text{obs}_1} = 1.0$  and the second set of obstacle radius  $R_{\text{obs}_2} = 2.0$ .

determined by subdomain locations in the original domain. Subdomain types are illustrated in Figures 3.3-3.4.

We run simulations on a 2D RVE with periodic boundary conditions. Each subdomain is of length  $l = 10$ , making the total RVE of size  $30 \times 30$  for cases (i) and (ii) and size  $20 \times 20$  for case (iii). We will use the cardinal directions to describe the different boundary walls of the domain and particle movements to neighboring subdomains (i.e. north, east, south,

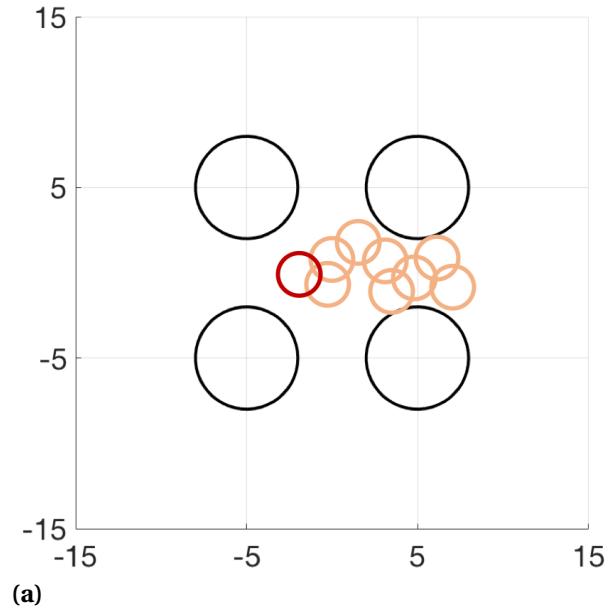
west).



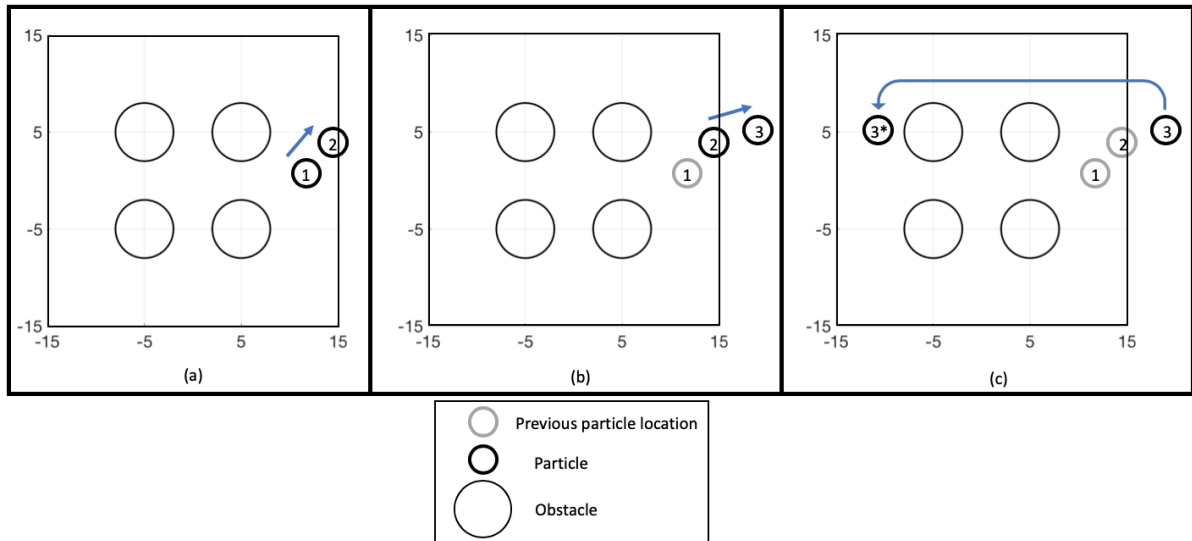
**Figure 3.5** Domain diagram for four internal obstacles case with subdomain types labeled. Solid lines represent the original domain, or the RVE, and dashed lines represent a portion of the extended domain.

Figure 3.6 shows an illustration of a possible random walk in the four internal obstacle case. At  $t = 0$ , the particle is located in the center subdomain (subdomain 5). At each timestep of size  $\Delta t = 0.05$ , a new direction is chosen from  $\theta \in [0, 2\pi]$  via a uniform sampling. The particle now moves a distance of  $\Delta x = v \Delta t$  in the direction of  $\theta$ . If the particle collides with an obstacle, we simulate perfectly elastic collisions, with the obstacles remaining stationary but the particle bouncing off of the obstacle in the appropriate direction. Effects of the periodic boundary conditions (PBC) on particle motion are shown in Figure 3.7. If a particle reaches the edge of a boundary wall in the initial RVE, it reenters the simulated domain on the opposite side but at the same position along the boundary. For example, if the particle exits from the eastern side of the domain, it will reenter on the western side of the domain at the same vertical location where the particle exited from the opposite side.

The state of the system at time  $t$  is considered to be the subdomain in which the particle is located. Subdomains are numbered in order to track particle location and state changes, according to the number of subdomains present (No. 1-9 for the equally spaced obstacles model (Fig. 3.3(a)) and four internal obstacles model (Fig. 3.3(b)); No. 1-4 for the multisize obstacles model (Fig. 3.4)). Tracking state changes is important for building the transition probabilities used in the surrogate model, discussed in greater detail later in this chapter



**Figure 3.6** Illustration of a particle's random walk path in the four internal obstacle case. The larger black circles are obstacles, the red circle indicates the particle's initial location, and the light orange circles depict the random walk path that the particle takes.



(a)

**Figure 3.7** Illustrations of periodic boundary conditions, with exaggerated stepsizes for clarity. (a) The particle moves from location (1) to location (2). Location (2) is on the boundary, but the center of the particle is still in the RVE, so the particle is not considered to have crossed the boundary yet. (b) The particle moves from location (2) to location (3). Location (3) is over the boundary, so the particle reenters the RVE from the opposite side of the domain. (c) Particle moves from location (3) to location (3\*).

(see sections 3.3.1.1, 3.3.2.1, and 3.3.3.1).

We investigate how the slope of mean squared displacement  $\langle r^2 \rangle$  versus time changes with different obstacle configurations and with increasing obstacle radii. We compare these results for the continuous and surrogate models. After all simulations are completed, the resulting mean-squared displacement is used to estimate the diffusivity of the system. We compare these results to the Mackie-Mears relation for diffusivity in porous media  $D^N[\phi^S]$ ,

$$D^N[\phi^S] = D_0^N \left( \frac{1 - \phi^S}{1 + \phi^S} \right)^2, \quad (3.1)$$

where  $D_0^N$  is the diffusivity in a free solution (one without obstacles), and  $\phi^S$  is the solid volume fraction of a mixture [22]. The solid volume fraction can be written as,

$$\phi^S = \frac{A_{\text{obs}}}{A_{\text{RVE}}}. \quad (3.2)$$

where  $A_{\text{obs}}$  is the two-dimensional area of obstacles, and  $A_{\text{RVE}}$  is the two-dimensional area of the total domain. In calculating  $A_{\text{obs}}$  and  $A_{\text{RVE}}$ , we assume the RVE has a fixed width,  $w$  (that cancels out when we calculate the solid volume fraction  $\phi^S$ ).

### 3.2.1 Model for equally spaced obstacles

To build the equally spaced obstacles model (Fig. 3.3(a)), obstacles are first placed at the corners of the center subdomain—these obstacles are located entirely within the RVE. Partial obstacles are then added along the boundary so that there is equal spacing between obstacles. The addition of partial obstacles makes our domain more isotropic than the other domains considered in this study (though not truly isotropic). Since this domain is the most isotropic of our domains, it is the most comparable to previous studies of diffusivity based on lattice models. This will be discussed more in Chapter 4. All obstacles are chosen to have the same size with radius  $R_{\text{obs}}$ .

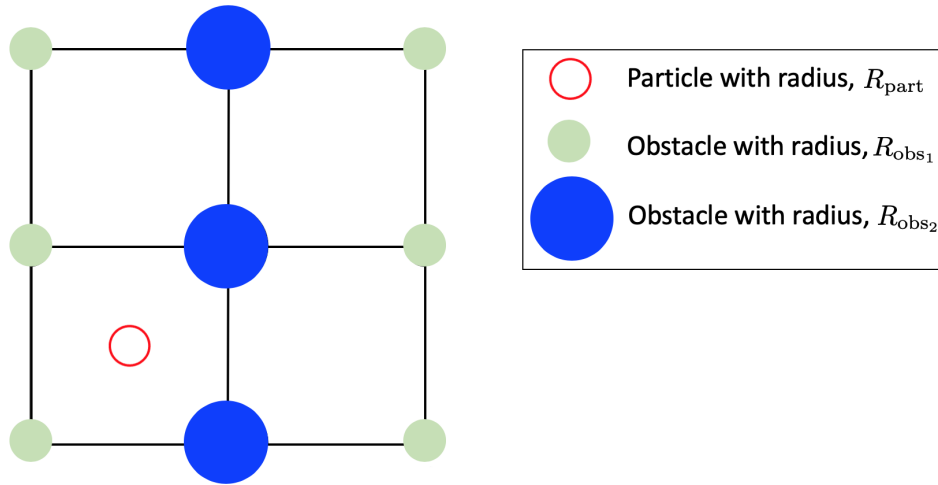
We consider a single particle with radius  $R_{\text{part}}$  which moves in a random walk with a speed of  $v = 8$ . For the deterministic model, we simulate many realizations over  $N_t = 1 \times 10^7$  timesteps and consider the plot of the mean-squared displacement  $\langle r^2 \rangle$  versus time. The slope of mean-squared displacement versus time can first be used as a proxy for diffusivity. In practice, 1500 to 2000 realizations on a High Performance Cluster are sufficient for the slope to reach stationary values. We first fix the particle radius as  $R_{\text{part}} = 1.0$  and vary the obstacle radius  $R_{\text{obs}}$  from 0.5 to 3.0 in increments of 0.5. This range is chosen to allow particle movement between obstacles—an obstacle radius greater than 3.0 greatly inhibits

transitions between subdomains.

After these simulations are completed, we also consider the case in which we fix the obstacle radius at  $R_{\text{obs}} = 2.0$  and vary the particle radius  $R_{\text{part}}$  from 0.5 to 2.5 in increments of 0.5. For the equally spaced obstacles model, we do not vary the particle radius because early results using the four internal obstacles case indicated that the system was more sensitive to varying the particle radius.

### 3.2.2 Model for four internal obstacles

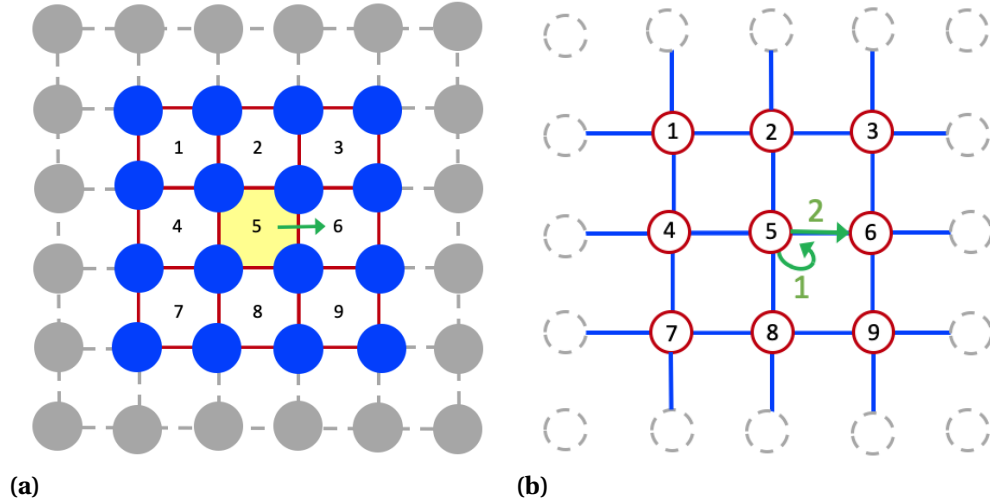
In the four internal obstacles model (Fig. 3.3(b)), four obstacles are placed at the corners of the center subdomain—these obstacles are located entirely within the RVE. Obstacle locations can also be thought of as being located at the intersection of subdomain boundary lines. Again, all obstacles are chosen to have the same size with radius  $R_{\text{obs}}$ . As in the equally spaced obstacles case, we first set the particle radius to  $R_{\text{part}} = 1.0$  and vary the obstacle radius,  $R_{\text{obs}}$  from 0.5 to 3.0. After these simulations are carried out, we also consider the case in which we fix the obstacle radius at  $R_{\text{obs}} = 2.0$  and vary the particle radius  $R_{\text{part}}$  from 0.5 to 2.5. We use the mean-squared displacement to estimate the diffusivity of the system under both scenarios (i.e. varying only the particle radius or varying only the obstacle radius).



**Figure 3.8** Domain diagram for multisize obstacles to illustrate which obstacle radii are varied. In simulations  $R_{\text{part}}$  and  $R_{\text{obs}_1}$  are fixed at 1.0 while  $R_{\text{obs}_2}$  is varied.

### 3.2.3 Model for multisize obstacles

For the multisize obstacle model (Fig. 3.4), we want to structure the RVE such that we can have one set of obstacles with fixed obstacle radius  $R_{\text{obs}_1}$  and another set of obstacles with an obstacle radius  $R_{\text{obs}_2}$  that we vary. The simplest way to structure this model and maintain symmetry to allow for simpler bookkeeping in the RVE is to reduce the domain from a  $3 \times 3$  to a  $2 \times 2$  domain. Figure 3.8 illustrates this modified domain. We consider the domain as having "columns" of obstacles, where all obstacles in the same column have the same radius. The first set of obstacles we denote as having radius  $R_{\text{obs}_1}$  and the second set of columns as having radius  $R_{\text{obs}_2}$ . We fix the particle radius at  $R_{\text{part}} = 1.0$ , the first set obstacle radius values at  $R_{\text{obs}_1} = 1.0$ , and vary the other obstacle radius values  $R_{\text{obs}_2}$  from 0.5 to 3.0. We, again, use the mean-squared displacement to estimate the diffusivity of the system. For this model we do not consider fixing the obstacle radii and varying the particle radius, due to the added complexity of two different sizes of obstacles and because our simulations on the equally spaced obstacles model and the four internal obstacles model both indicated the system was more sensitive to varying the obstacle radius.



**Figure 3.9** Equally spaced obstacles model. (a) Continuous domain with solid red lines indicating RVE subdomains and dashed gray lines indicating subdomains in the extended domain. Solid blue circles are obstacles in the original domain, and solid gray circles are obstacles in the extended domain. The center subdomain (subdomain 5) is highlighted in yellow, with transition types indicated with a green arrow. Only cardinal direction transitions are possible (in addition to staying in the current subdomain), so only one arrow is shown. (b) Lattice representation with red rings representing possible subdomains in the original domain and dashed gray rings indicating possible subdomains in the extended domain. Distinct transition types from subdomain 5 indicated with green arrows and numbers.

### 3.3 Surrogate lattice model description

We construct the surrogate lattice models by using the transition probabilities obtained from the continuous model simulations. The process by which we create transition probabilities is explained in sections 3.3.1.1, 3.3.2.1, and 3.3.3.1 for the equally spaced obstacles model, four internal obstacles model, and multisize obstacles model, respectively.

After collecting transition probabilities from the continuous model, we build the lattice model to keep track of the following:

- (i) The initial particle location (subdomain number and corresponding  $(x, y)$  position).
- (ii) The current particle location at each time step (subdomain number and corresponding  $(x, y)$  position).
- (iii) The previous subdomains the particle was in.
- (iv) Transitions to the extended domain, which is used to keep track of net particle displacement.

The lattice model relies on transition probabilities to determine if a particle changes subdomains. At each timestep we draw a random number uniformly from 0 to 1 to determine the particle's action at that timestep. We run the lattice model for  $N = 1 \times 10^7$  timesteps and compare lattice model results for mean squared displacement and diffusivity with those of the continuous model.

We mimic the domain of the continuous model, with subdomains in the continuous model corresponding to possible subdomains in the surrogate lattice model. Hence, the number of subdomains in the lattice model is the same as the number of subdomains in the corresponding continuous model. This means that we have a  $3 \times 3$  lattice for the equally spaced obstacles model (case (i)) and the four internal obstacles model (case (ii)), and a  $2 \times 2$  lattice for the multisize obstacles model (case (iii)). Although case (i) could be modeled with a  $1 \times 1$  lattice, for flexibility in comparing with our other models and versatility in using the RVE to capture varied obstacle arrangements, we use a  $3 \times 3$  lattice.

#### 3.3.1 Lattice model for equally spaced obstacles

We first construct the lattice model for the equally spaced obstacles case (Fig. 3.3(a)). Figure 3.9(a) shows the continuous domain and the extended domain with subdomains numbered 1-9, with distinct transitions from the center subdomain indicated with green arrows. Due to the extensive symmetries in this system, all subdomains are the same in terms of boundary



edges they share with neighboring subdomains and locations of obstacles. Further, based on symmetries within each subdomain, we can aggregate transition probabilities by transition type. Transition types are affected by the length of the boundary edge one subdomain shares with a neighboring subdomain. Figure 3.12(b) shows the resulting lattice model for the equally spaced obstacle case, with distinct transitions from subdomain 5 indicated with green arrows. As particles in this system are most likely to stay in their current subdomain, we expect transition probability No. 1 in Figure 3.9(b) to be close to 1.

Consider transitions from the center subdomain (subdomain 5), which is highlighted in Figure 3.9(b). When the particle is in the center subdomain, we make the following observations of possible transitions:

- (i) The particle can stay in its current subdomain (in this case, subdomain 5).
- (ii) The particle cannot transition diagonally from this subdomain since subdomain 5 has obstacles at each of its corners. These obstacles block any possible diagonal transitions.
- (iii) If the particle is leaving its current subdomain, the particle can only move in a cardinal direction (north, east, south, west). For each possible cardinal direction, the unobstructed boundary edge between subdomain 5 and the neighboring subdomain (subdomains 2, 4, 6, or 8) is the same regardless of which cardinal direction the particle moves. For example, the probability of transitioning from subdomain 5 to subdomain 2 (a transition to the north) is the same as the probability of transitioning from subdomain 5 to subdomain 4 (a transition to the west).

Based on the symmetries in our continuous model, for each obstacle radius and particle radius configuration we have two resulting probabilities –(1) the probability that a particle stays in its current subdomain, and (2) the probability of the particle moving to a neighboring subdomain in a cardinal direction. The probability of diagonal transitions should be zero, as obstacles are always at the corners blocking those transitions.

### 3.3.1.1 Tracking transition probabilities for the equally spaced obstacles model

We consider the state of the system to be the subdomain in which the particle is located. To visualize subdomains and particle locations, recall Figure 3.9. We aim to identify and quantify system properties of the directly simulated continuous model via the *timecourse*  $\eta(t_k)$  ( $t_k = k\Delta t$ ,  $k = 0, \dots, N$ ). This tracks the subdomain the particle is in (its state) over the total timecourse, where  $N$  is the total number of time steps. A particle is counted as being located in the  $i$ th subdomain ( $i = 1, \dots, 9$ ) at time  $t_k$  if its center is located in that

subdomain at time  $t_k$ . We create a  $9 \times 9$  matrix  $S^{(k)}$  to count all transitions that have occurred up to the current timestep  $k = 1, \dots, N$ , (we initialize  $S^{(0)}$  is a  $9 \times 9$  zero matrix). The rows of  $S^{(k)}$  correspond with the subdomain in which the particle was located, and the columns of  $S^{(k)}$  correspond with the subdomain to where the particle moves.  $S^{(k)}$  is updated at each timestep according to the formula,

$$S^{(k)}(\eta(t_{k-1}), \eta(t_k)) = S^{(k-1)}(\eta(t_{k-1}), \eta(t_k)) + 1, \quad k = 1, \dots, N \quad (3.3)$$

For example, if the particle is in subdomain 1 (a corner subdomain) (Fig. 3.9(b)) for timesteps  $k = 1, \dots, 10$ , and moves to subdomain 2 (the onewall subdomain to the east) in the next timestep ( $k = 11$ ), then we have the following recorded:

$$\begin{aligned} \eta(1) &= \eta(2) = \dots = \eta(10) = 1 \\ \eta(11) &= 2 \\ S^{(11)}(1, 1) &= 10 \\ S^{(11)}(1, 2) &= 1 \\ S^{(11)} &= \begin{bmatrix} 10 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 \end{bmatrix} \end{aligned} \quad (3.4)$$

We want to build a vector  $\vec{p}^E$  of transition probabilities, with superscript  $E$  indicating probabilities are for the equally spaced obstacle model. In order to begin creating the vector  $\vec{p}^E$ , we first create the vector  $\vec{T}_i^E$ ,  $i = 1, \dots, 9$  to count transitions that occurred for each subdomain  $i$ . After timestep  $k$  we use the rows of  $S^{(k)}$  to create vector  $\vec{T}_i^E$  ( $i = 1, \dots, 9$ ) by grouping transition sums by transition types. For the equally spaced obstacles model all subdomains have two possible transitions types—staying in its current subdomain or leaving in a cardinal direction (as discussed above). Hence  $\vec{T}_i^E$  has dimensions of  $1 \times 2$  for  $i = 1, \dots, 9$ , which corresponds to the number of independent transition probabilities (note that the number of transition probabilities will be different for the other two models).

Specifically, at the current time  $t_k$  we construct  $\vec{T}_i^E$  using matrix  $S^{(k)}$ , as follows:

$$\vec{T}_i^E = \left[ S^{(k)}(i, i), \sum_{j \neq i} S^{(k)}(i, j) \right], \quad i = 1, \dots, 9. \quad (3.5)$$

The index  $j$  indicates other subdomains besides subdomain  $i$  that the particle could move to. These possible subdomains  $j$  are dependent on which subdomain  $i$  the particle is

currently in.

For example, for subdomain 5, the particle can stay in its current subdomain ( $i = 5$ ) or it can move in a cardinal direction to subdomains  $i = 2, 4, 6$ , or  $8$ . Vector  $\vec{T}_5^E$  counts when the particle stayed in subdomain 5 or moved from subdomain 5 in a cardinal direction, and is created from  $S$  by:

$$\vec{T}_5^E = [S^{(k)}(5, 5), S^{(k)}(5, 2) + S^{(k)}(5, 4) + S^{(k)}(5, 6) + S^{(k)}(5, 8)].$$

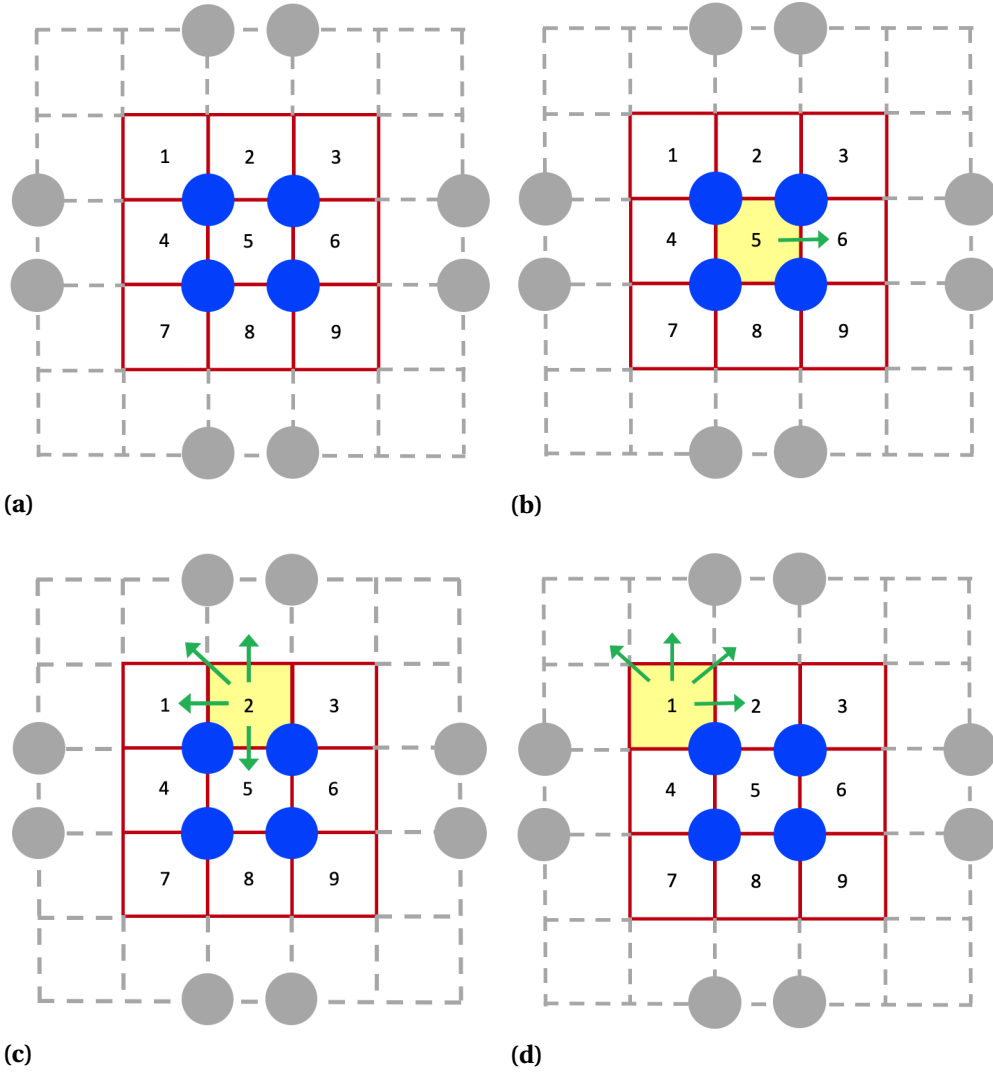
The same process is followed for all  $\vec{T}_i^E$  for  $i = 1, \dots, 9$ . For example, for each subdomain,  $\vec{T}_i^E(1, 1)$  counts the number of times a particle was in subdomain  $i$  and stayed in subdomain  $i$ , and  $\vec{T}_i^E(1, 2)$  counts the number of times a particle was in subdomain  $i$  and moved in a cardinal direction to one of the four possible subdomains for  $i = 1, \dots, 9$ .

Due to the symmetries in the equally spaced obstacles case, we create the transition probability vector  $\vec{p}^E$  via:

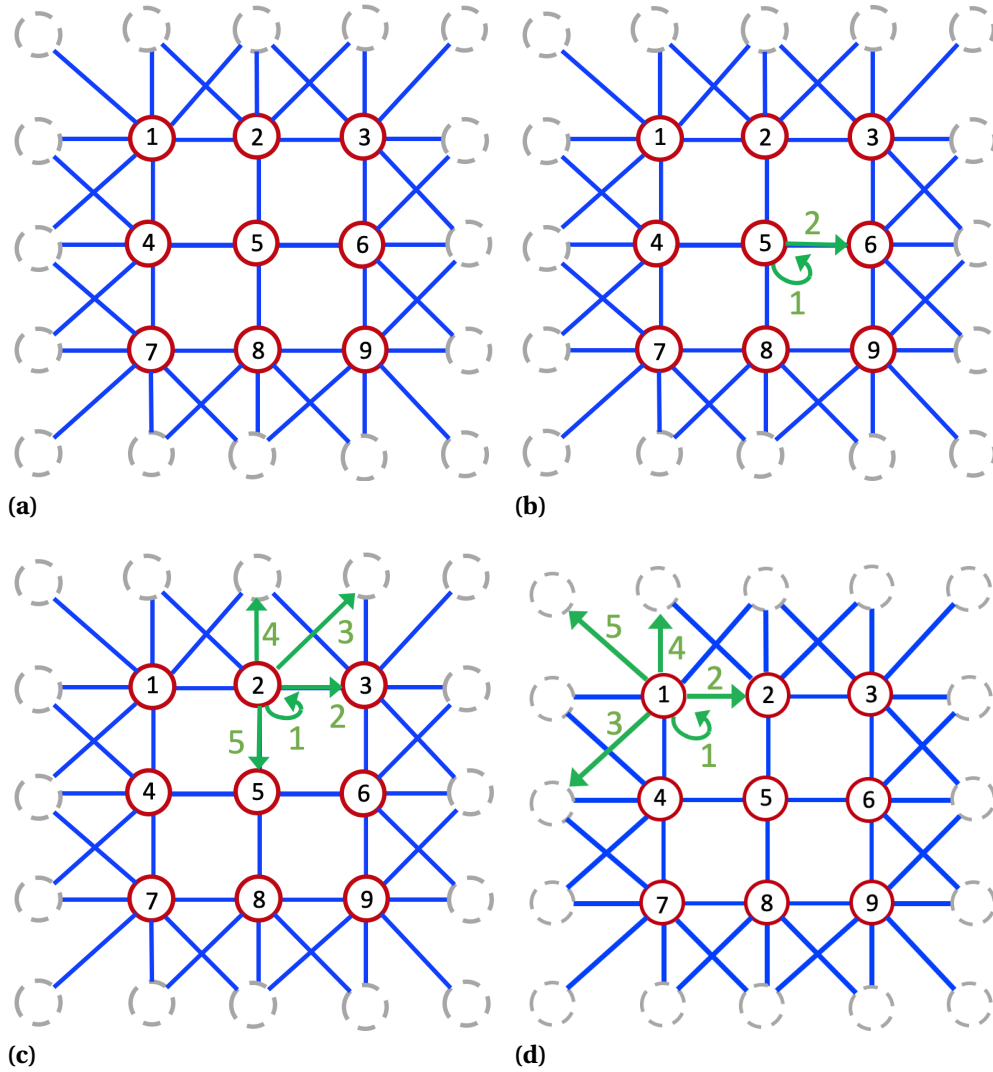
$$\vec{p}^E = \frac{1}{N} \sum_{i=1}^N \vec{T}_i^E. \quad (3.6)$$

For the equally spaced model,  $\vec{p}^E$  is a  $1 \times 2$  matrix created by averaging the transition sums for each subdomain.

For example, in our resulting transition probability vector,  $\vec{p}^E$  the value at  $\vec{p}^E(1, 1)$  indicates the probability that a particle will stay in its current subdomain, and the value at  $\vec{p}^E(1, 2)$  indicates the probability that a particle will move in a cardinal direction to one of the four possible subdomains. There are more types of transitions for the four internal obstacle model and the multisize obstacles model, and these differences will be discussed further in sections 3.3.2.1 and 3.3.3.1. In our simulations, we check the value of the transition probabilities along the timecourse to ensure that the transition probabilities are reaching stationary values.



**Figure 3.10** Continuous domain for four internal obstacle case. RVE is shown with red boundary lines and extended domain is shown with gray, dashed boundary lines. Obstacles are indicated by solid blue circles in RVE and solid gray circles in extended domain. Distinct transition types indicated with green arrow. (a) Illustration of how subdomains are numbered for tracking transition probabilities. (b) Illustration of transitions from the center subdomain (subdomain 5), with the center subdomain highlighted in yellow. Note that transitions in any cardinal direction should occur with the same probability, this is illustrated with a single green arrow indicating there is only one type of transition possible. From subdomain 5 there are two independent transition probabilities. (c) Illustration of transitions from a subdomain 2, a onewall subdomain, with subdomain 2 highlighted in yellow. From subdomain 2 there are five independent transition probabilities. (d) Illustrations of transitions from subdomain 1, a corner subdomain, with subdomain 1 highlighted in yellow. From subdomain 1 there are five independent transition probabilities.



**Figure 3.11** Lattice representation for four internal obstacle case with red rings indicating possible subdomains in the original domain and dashed gray rings indicating possible subdomains in the extended domain. Possible transitions are indicated with solid blue lines. Distinct transition types indicated with green arrows and numbers. (a) Lattice representation. (b) Illustration of transitions from subdomain 5, the center subdomain (2 distinct transition types). (c) Illustration of transitions from a subdomain 2, a onewall subdomain (5 distinct transition types). (d) Illustrations of transitions from subdomain 1, a corner subdomain (5 distinct transition types).

### 3.3.2 Lattice model for four internal obstacles

We now construct the lattice model for the four internal obstacle case (Fig. 3.3(b)). Figure 3.10(a) shows the continuous domain and the extended domain with subdomains numbered 1-9. Utilizing the symmetries present in the RVE, we aggregate the transition probabilities by transition type. Transition types are described in Figure 3.10 based off of transitions from a particular subdomain to another neighboring subdomain.

First consider transitions from the center subdomain, which is highlighted in Figure 3.10(b). When the particle is in the center subdomain (or subdomain 5), we make the following observations of possible transitions:

- (i) The particle can stay in its current subdomain (in this case, subdomain 5).
- (ii) The particle cannot transition diagonally from this subdomain since subdomain 5 has obstacles at each of its corners.
- (iii) If the particle is leaving its current subdomain, the particle can only move in a cardinal direction (north, east, south, west). For each possible cardinal direction, the unobstructed boundary edge between subdomain 5 and the neighboring subdomain (subdomains 2, 4, 6, or 8) is the same regardless of which cardinal direction the particle moves. For example, the probability of transitioning from subdomain 5 to subdomain 2 (a transition to the north) is the same as the probability of transitioning from subdomain 5 to subdomain 4 (a transition to the west).

In summary, there are only two possible types of transitions when a particle is in the center subdomain—either (1) the particle stays in the current subdomain, or (2) the particle transitions to another subdomain in a cardinal direction.

Next we consider transitions from a onewall subdomain. Figure 3.10(c) illustrates transitions from subdomain 2, a onewall subdomain. For transitions from subdomain 2, we can identify five distinct transition types:

- (i) The particle can stay in its current subdomain (in this case, subdomain 2).
- (ii) The particle can transition west to subdomain 1 or east to subdomain 3 with equal probabilities, because the unobstructed subdomain edges between subdomain 2 and subdomain 1 or subdomain 3 are of equal length.
- (iii) Similarly, the particle can move northwest to subdomain 7 or northeast to subdomain 9 with equal probabilities.

- (iv) Transitions west from subdomain 2 to subdomain 8 have a distinct transition probability.
- (v) Transitions east from subdomain 2 to subdomain 5 have a distinct transition probability.

Finally, we consider transitions from a corner subdomain. Figure 3.10(d) illustrates transitions from subdomain 1, a corner subdomain. For transitions from subdomain 1, we can identify five distinct transition types:

- (i) The particle can stay in its current subdomain (in this case, subdomain 1).
- (ii) The particle can transition east to subdomain 2 or south to subdomain 4 with equal probability, because the unobstructed subdomain edge between subdomain 1 and subdomain 2 or subdomain 4 is of equal length.
- (iii) Similarly, the particle can move north to subdomain 7 or west to subdomain 3 (in the extended domains) with equal probabilities.
- (iv) Similarly, the particle can move southwest to subdomain 6 or northeast to subdomain 8 with equal probability.
- (v) Lastly, the particle can move northwest to subdomain 9 with a distinct probability.

Now that we have identified transition types in the continuous domain, we can construct a lattice representation, as illustrated in Figure 3.11. In the lattice model, rings represent the particle's location (e.g. ring 5 in the lattice model corresponds with a particle being in subdomain 5 (the center subdomain) in the continuous model). Figure 3.11(b) illustrates transitions in the lattice model from the center subdomain (subdomain 5). A particle can either stay in its current subdomain (transition type 1) or move to another subdomain in a cardinal direction (transition type 2).

Figure 3.11(c) shows the lattice representations of transitions from onewall subdomains and Figure 3.11(d) shows lattice representations of transitions from the corner subdomains. Again, the particle will be most likely to stay in its current subdomain. The onewall subdomain has 5 independent transition types, and the corner subdomain has five independent transition types.

### 3.3.2.1 Tracking transition probabilities for the four internal obstacles model

Tracking transition probabilities for the four internal obstacle model begins similarly to the equally spaced model, as discussed in section 3.3.1.1. We want to build vectors  $\vec{p}_l^F$  of

transition probabilities, with superscript  $F$  indicating probabilities are for the four internal obstacles model and subscript  $l = L, I, C$  to indicate transitions occurring from the corner, onewall, and center subdomains, respectively. In order to begin creating the vectors  $\vec{p}_l^F$  for  $l = L, I, C$ , we first create the vectors  $\vec{T}_i^F$ ,  $i = 1, \dots, 9$  to count transitions that occurred for each subdomain  $i$ . The  $9 \times 9$  matrix  $S^{(k)}$  is created in the same manner as in the equally spaced obstacles case (Sec. 3.3.1.1) for timesteps  $k = 1, \dots, N$ . We use  $S^{(k)}$  to create vectors  $\vec{p}_i^F$  to count transitions that occurred over the total timecourse when the particle was in subdomain  $i$  for  $i = 1, \dots, 9$ . Since the subdomain types for this model are different, the dimensions of  $\vec{T}_i^F$  vary based on the subdomain type.

Recall Figure 3.11 (b) - (d) for transition types grouped by subdomain type. Subdomain 5 is the only center subdomain ( $l = C$ ), and transitions for this subdomain are similar to those in the equally spaced obstacles model. The vector  $\vec{T}_5^F$  has dimensions  $1 \times 2$  and is calculated by

$$\vec{T}_5^F = \left[ S^{(k)}(5, 5), \sum_{j=2,4,6,8} S^{(k)}(5, j) \right]. \quad (3.7)$$

The index  $j$  indicates other subdomains besides subdomain 5 that the particle could move to. These possible subdomains are  $j = 2, 4, 6$ , and  $8$ .

To consider transitions occurring from a onewall subdomain ( $l = I$ ), it is simpler to consider transitions possible when in subdomain 2. As illustrated in Figure 3.11(c), there are five possible transition types when the particle is in subdomain 2, with the most likely being for the particle to remain in subdomain 2. We then have four other transition types. The vector  $\vec{T}_2^F$  has dimensions  $1 \times 5$  and is calculated by

$$\vec{T}_2^F = \left[ S^{(k)}(2, 2), \sum_{j=1,3} S^{(k)}(2, j), \sum_{j=7,9} S^{(k)}(2, j), S^{(k)}(2, 8), S^{(k)}(2, 5) \right]. \quad (3.8)$$

For the remaining onewall subdomains ( $j = 4, 6, 8$ ), these transitions are counted in the same manner, following the transition types as numbered in Figure 3.11(c).

To consider transitions occurring from a corner subdomain ( $l = L$ ), it is simpler to consider transitions possible when in subdomain 1 (Fig. 3.11(d)). When in subdomain 1, there are 5 possible transition types, with the most likely being that the particle remain in subdomain 1. The vector  $\vec{T}_1^F$  has dimension  $1 \times 5$  and is calculated by

$$\vec{p}_1^F = \left[ S^{(k)}(1, 1), \sum_{j=2,4} S^{(k)}(1, j), \sum_{j=6,8} S^{(k)}(1, j), \sum_{j=3,7} S^{(k)}(1, j), S^{(k)}(1, 9) \right]. \quad (3.9)$$



For the remaining corner subdomains ( $j = 3, 7, 8$ ), these transitions are counted in the same manner, following the transition types as numbered in Figure 3.11(d).

Now that we have vectors  $\vec{T}_i^F$  ( $i = 1, \dots, 9$ ) tracking transitions for each subdomain, we want to create vectors  $\vec{p}_l^F$  for the transition probabilities by subdomain type ( $l = L, I, C$ ). For the center subdomain, this is the vector  $\vec{T}_5^F$  divided by the sum of components of the vector  $\vec{T}_5^F$  (or the total number of times the particle was in subdomain 5). We denote the  $i$ th component of the vector  $\vec{T}_5^F$  as  $T_5^F(i)$ . Hence, transition probability vector  $\vec{p}_C^F$  has dimension  $1 \times 2$  and is written as

$$\vec{p}_C^F = \frac{\vec{T}_5^F}{\sum_{i=1}^9 T_5^F(i)}. \quad (3.10)$$

For the onewall subdomains, we create the probability vector  $\vec{p}_I^F$  by summing up the different onewall transition vectors,  $\vec{T}_i^F$  for  $i = 2, 4, 6, 8$  and dividing by the sum of these vectors (the total number of times the particle was in a onewall subdomain). Hence, the transition probability vector  $\vec{p}_I^F$  has dimension  $1 \times 5$  and is written as

$$\vec{p}_I^F = \frac{\sum_{j=2,4,6,8} \vec{T}_j^F}{\sum_{i=1}^9 \sum_{j=2,4,6,8} T_j^F(i)}. \quad (3.11)$$

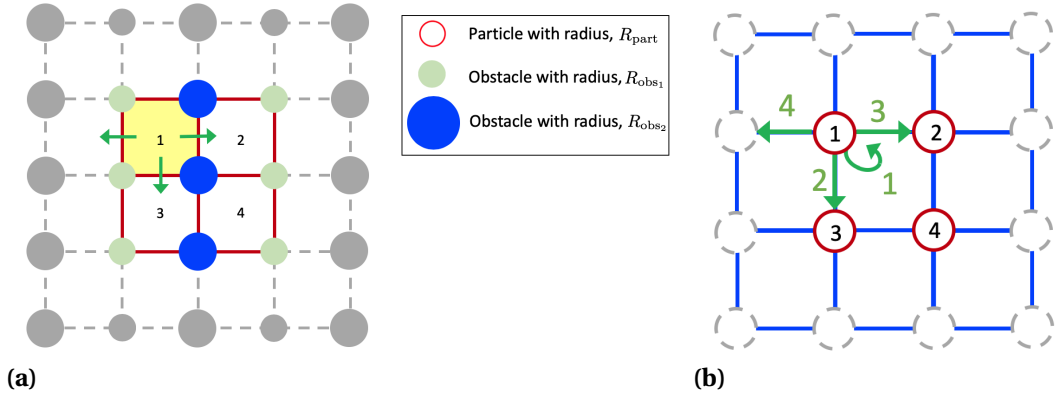
For the corner subdomains, we create the probability vector  $\vec{p}_L^F$  by summing up the different corner transition vectors,  $\vec{T}_i^F$  for  $i = 1, 3, 7, 9$  and dividing by the sum of these vectors (the total number of times the particle was in a corner subdomain). Hence, the transition probability vector  $\vec{p}_L^F$  has dimension  $1 \times 5$  and is written as

$$\vec{p}_L^F = \frac{\sum_{j=1,3,7,9} \vec{T}_j^F}{\sum_{i=1}^9 \sum_{j=1,3,7,9} T_j^F(i)}. \quad (3.12)$$

In our simulations, we check the value of the transition probabilities along the timecourse to ensure that the transition probabilities are reaching stationary values.

### 3.3.3 Lattice model for multisize obstacles

We similarly construct the lattice model for the multisize obstacles case (Fig. 3.4). Recall for this obstacle configuration, due to the inherent symmetries of the system, we reduce the domain to a  $2 \times 2$  grid. Figure 3.12(a) shows the continuous domain and the extended domain with subdomains numbered 1-4, with distinct transitions from subdomain 1 indicated with green arrows.



**Figure 3.12** Multisize obstacles model. (a) Continuous grid with solid red lines indicating subdomains in the original domain and dashed gray lines indicating subdomains in the extended domain. Solid blue and green circles indicate obstacles in the original domain, with the smaller green circles indicating obstacles with obstacle radius fixed at  $R_{\text{obs}_1} = 1.0$ , and the larger blue circles indicating obstacles with varying obstacle radius  $R_{\text{obs}_2}$ . Solid gray circles are obstacles in the extended domain, with the circle size smaller and larger indicating radius  $R_{\text{obs}_1}$  and  $R_{\text{obs}_2}$ , respectively. Subdomain 1 is highlighted in yellow, with distinct transition types indicated with a green arrow. (b) Lattice representation with solid red rings indicating possible subdomains in the original domain and dashed gray rings indicating possible subdomains in the extended domain. Distinct transition types from subdomain 1 indicated with green arrows and numbers.

Due to the inherent symmetry in the system, all subdomains have the same transition types, although these transition types occur in different directions. For example, subdomain 1 has the same transition probabilities as subdomain 2 but these transition probabilities occur in opposite directions with respect to the horizontal axis. Based on the symmetries of this RVE, we can aggregate transition probabilities by transition type. Transition types are affected by the length of the boundary edge one subdomain shares with a neighboring subdomain. Figure 3.12(b) shows the resulting lattice model for the multisize obstacle case, again with distinct transitions from subdomain 1 indicated with green arrows. As particles in this system are most likely to stay in their current subdomain, we expect transition probability 1 in Figure 3.12(b) to be close to 1.

Consider transitions from subdomain 1, which is highlighted in Figure 3.9(b). When the particle is in subdomain 1, we make the following observations of possible transitions:

- (i) The particle can stay in its current subdomain (in this case, subdomain 1).
- (ii) The particle cannot transition diagonally from this subdomain since subdomain 1 has obstacles at each of its corners. These obstacles block any possible diagonal transitions.

- (iii) The particle can move east to subdomain 2 in the original domain with a distinct transition probability, since this boundary has two obstacles with radius  $R_{\text{obs}_2}$ .
- (iv) The particle can move west to subdomain 2 in the extended domain with a distinct transition type, since this boundary has two obstacles with radius  $R_{\text{obs}_1}$ .
- (v) the particle can move south to subdomain 3 in the original domain or north to subdomain 3 in the extended domain with the same probability due to the boundary edge having one obstacle with radius  $R_{\text{obs}_1}$  and one obstacle with radius  $R_{\text{obs}_2}$ .

Based on the symmetries in our continuous model, for each obstacle radius configuration we have four resulting probabilities. The probability of diagonal transitions should be zero, as obstacles are always at the corners blocking those transitions.

After collecting transition probabilities from the continuous model, we build the lattice model to keep track of the following:

- (i) The initial particle location (subdomain number and corresponding  $(x, y)$  position).
- (ii) The current particle location at each time step (subdomain number and corresponding  $(x, y)$  position).
- (iii) The previous subdomains the particle was in.
- (iv) Transitions to the extended domain, used to keep track of net particle displacement.
- (v) A counter to distinguish between horizontal transitions within the original domain versus horizontal transitions to the extended domain (further explained below).

Since this RVE is composed of  $2 \times 2$  subdomains, and, in the horizontal direction, there is a differing obstacle arrangement to the east versus to the west, we need a counter to distinguish between horizontal transitions within the original domain and horizontal transitions to the extended domain. For example, if a particle is located in subdomain 1 and moves east to subdomain 2 (in the original domain), then that boundary edge has two obstacles with radius  $R_{\text{obs}_2}$ . However, if a particle is located in subdomain 1 and moves west to subdomain 2 (in the extended domain), then that boundary edge has two obstacles with radius  $R_{\text{obs}_1}$ . We do not need such a counter for vertical transitions, since the boundary in the vertical direction between any two subdomains has one obstacle with radius  $R_{\text{obs}_1}$  and one obstacle with radius  $R_{\text{obs}_2}$ .

We run the surrogate lattice model for  $N = 1 \times 10^7$  timesteps and compare lattice model results for mean squared displacement with those of the continuous model.

### 3.3.3.1 Tracking transition probabilities for the multisize obstacles model

Tracking transition probabilities for the multisize obstacles model begins similarly to that of the equally spaced model, as discussed in section 3.3.1.1. We want to build vectors  $\vec{p}^M$  of transition probabilities, with superscript  $M$  indicating probabilities are for the multisize obstacles model. In order to begin creating the vectors  $\vec{p}^M$ , we first create the vectors  $\vec{T}_i^M$ ,  $i = 1, \dots, 4$  to count transitions that occurred for each subdomain  $i$ .

As we can see from Figure 3.12, since this domain is  $2 \times 2$ , when either (1) the particle moves in a cardinal direction (e.g. north) to a subdomain in the original domain, or (2) when the particle moves in the opposite direction (e.g. south) to a subdomain in the extended domain, it moves to the same numerically numbered subdomain. For example, when a particle is in subdomain 1, it can move west into the extended domain and into subdomain 2, or it can move east, staying in the original domain and moving into subdomain 2. With north-south transitions, the obstacle configuration between subdomains is the same (one obstacle of radius  $R_{\text{obs}_1}$  and one obstacle of  $R_{\text{obs}_2}$ ), but with east-west transitions the obstacle configuration between subdomains is different (either two obstacles with radius  $R_{\text{obs}_1}$  or two obstacles with radius  $R_{\text{obs}_2}$ ). Due to this difference, we must keep track of when east-west transitions occur to subdomains outside the original domain. To track this difference, we make the matrix  $S^{(k)}$  a  $4 \times 5$  matrix, with the rows indicating the subdomain the particle was in, and the columns indicating the subdomain the particle went to. The 5th column in  $S^{(k)}$  indicates an east-west transition occurred in which the particle moved outside the original domain.

Other than the column differentiating between in-domain east-west and out-of-domain east-west transitions,  $S^{(k)}$  is then created in the same manner as in the equally spaced obstacles case (Sec. 3.3.1.1) for timesteps  $k = 1, \dots, N$ . We use  $S^{(k)}$  to create vectors  $\vec{p}_i^M$  to count transitions that occurred over the total timecourse when the particle was in subdomain  $i$  for  $i = 1, \dots, 4$ .

Recall Figure 3.12 (a) - (b) for transition types by subdomain type. We only have one type of subdomain, but with the obstacles located at different corners. Consider transitions from subdomain 1. The vector  $\vec{T}_1^M$  is calculated by

$$\vec{T}_1^M = [S^{(k)}(1, 1), S^{(k)}(1, 2), S^{(k)}(1, 3), S^{(k)}(1, 4) + S^{(k)}(1, 5)]. \quad (3.13)$$

A similar ordering is followed for the remaining three subdomains. For vector  $\vec{T}_i^M$ ,  $i = 1, 2, 3, 4$ , we can describe the elements in the following manner:

- (i) the element in the (1, 1) position indicates the number of times a particle stayed in its current subdomain.

- (ii) the element in the (1,2) position indicates the number of times a particle moved either north or south (note that we do not need to track if it moves within the original domain since the obstacle configuration is the same for both boundary edges – the boundary edge has one obstacle of radius  $R_{\text{obs}_1}$  and one obstacle of radius  $R_{\text{obs}_2}$ ).
- (iii) the element in the (1,3) position indicates the number of times a particle moved either east or west but remained in the original domain (the boundary edge has two obstacles of radius  $R_{\text{obs}_2}$ ).
- (iv) the element in the (1,4) position indicates the number of times a particle moved either east or west but moved to the extended domain (the boundary edge has two obstacles of radius  $R_{\text{obs}_1}$ ).

Now that we have vectors  $\vec{T}_i^M$  ( $i = 1, \dots, 4$ ) tracking transitions for each subdomain, we want to create vectors  $\vec{p}^M$  for the transition probabilities. Due to the symmetries in the multisize obstacles model, we create the transition probability vector  $\vec{p}^M$  by

$$\vec{p}^M = \frac{\sum_{j=1}^4 \vec{T}_j^M}{\sum_{i=1}^4 \sum_{j=1}^4 T_j^M(i)}. \quad (3.14)$$

In our resulting transition probability vector,  $\vec{p}^M$  the value at  $\vec{p}^M(1,1)$  indicates the probability that a particle will stay in its current subdomain, and the values at  $\vec{p}^M(1,i)$  for  $i = 2, 3, 4$  indicate the probabilities that a particle will move according to that transition type. For the transition probability at  $\vec{p}^M(1,2)$ , this indicates the probability of a particle moving either north or south, and these two transitions occur with equal probability. In our simulations, we check the value of the transition probabilities along the timecourse to ensure that the transition probabilities are reaching stationary values.

## CHAPTER

# 4

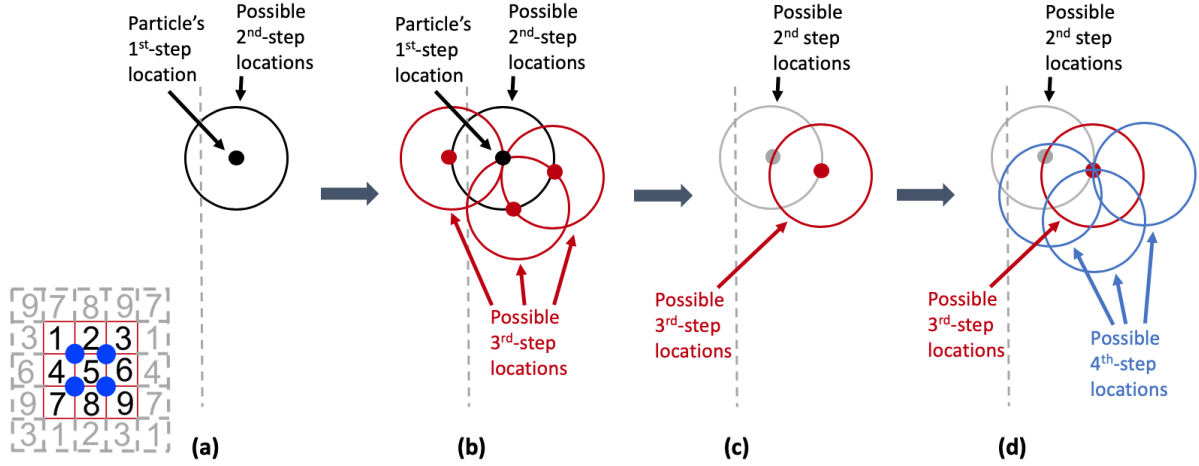
## MODEL CALIBRATION AND DIFFUSIVITY ESTIMATION

In this chapter, we introduce some additional concepts used to build and assess the continuous and surrogate models for 2D random walks: a commitment index  $M_c$ , and several topics relating to model diffusivity. Since the particle exhibits a random walk in our continuous model,  $M_c$  allows us to determine when a particle commits to a new subdomain. Without including  $M_c$ , the probability of a particle leaving a particular subdomain is, effectively, too high due to frequent crossing and recrossing over subdomain boundaries. Incorporating  $M_c$  into our models allows us to identify transition probabilities for the surrogate model that yield results in better agreement with the continuous model.

For diffusivity, we discuss estimating the effective diffusivity from the mean squared displacement,  $\langle r^2 \rangle$ , simulation results, as well as estimates for  $\Delta x$  used in comparing our simulated effective diffusivities with theoretical diffusivity estimates for diffusion models on a 2D lattice. We use equations discussed in literature which relate  $\langle r^2 \rangle$  to the diffusion constant  $D$ , which we will refer to as simulated diffusivity or effective diffusivity ( $D_{\text{cont}}$  and  $D_{\text{surr}}$  for the continuous and surrogate models, respectively). We use a 2D Taylor Series expansion to find the diffusion equation and theoretical diffusion constant,  $D_{\text{TS}}$ . We then use limits involving particle movement on a lattice in two dimensions to relate our simulated

diffusivity  $D_{\text{surr}}$  to the theoretical diffusivity  $D_{\text{TS}}$ .

## 4.1 Commitment index for the continuous model



**Figure 4.1** Commitment index,  $M_C$ , illustration. (a) The particle (solid black circle) has just crossed the boundary between subdomains (dashed gray line). We call this the 1st-step location. Possible particle 2nd-step locations are indicated by the solid black ring. (b) Three possible 2nd-step locations are indicated with the solid red circles. We illustrate some possible next steps (3rd-step locations) with red rings. (c) For a specific 2nd-step location (solid red circle), the 1st-step location and ring of possible 2nd-step locations are colored gray (as these steps have passed). A red ring illustrates the particles' remaining possible 3rd-step locations. (d) Possible 4th-step locations are indicated with blue rings.

In comparing preliminary results from the continuous and surrogate models, we observe that we need to include a *commitment index*,  $M_C$  to indicate when a particle has "committed" to its new subdomain. Due to the random direction at each time step, a particle frequently crosses and returns across the boundary line between subdomains over a few time steps. We can think of this boundary recrossing as added noise in our calculation of transition probabilities. Introducing  $M_C$  allows us to mitigate some of that noise. For an illustration of  $M_C$  and why it is needed, see Figure 4.1, as explained below:

- (i) In Figure 4.1(a), we consider a particle that has just crossed the boundary between subdomains. The particle is just inside the new boundary, and a ring of possible particle locations after the next step includes some locations that result from recrossing the boundary between subdomains. We will call the particle's location at this point the 1st-step location, as this is its first step into the new subdomain.

- (ii) In Figure 4.1 (b), we further consider the possible steps relative to the particle's 1st-step location, as illustrated with three solid circles indicating possible 2nd-step locations. A ring of possible next steps (3rd-step locations) can be drawn around each 2nd-step location.
- (iii) In Figure 4.1 (c) we choose one of the possible 2nd-step locations (shown in gray). In this case, we can reduce the possible 3rd-step locations to the remaining solid red ring.
- (iv) In Figure 4.1 (d) we illustrate the remaining possible 3rd-step locations, and after an additional step, several 4th-step locations are illustrated with blue rings.

With each successive step in which the particle stays in the new subdomain, the set of possible next-steps resulting in a return back into the previous subdomain becomes smaller and smaller. For each case of our model (i.e. particle and radius configurations), we need to calibrate the commitment index by identifying its best value. We do this by simulating the continuous model and considering varying values of the commitment index. For example, for  $M_C = 5$ , the particle must remain in its new subdomain for 5 consecutive steps to be recorded as being in the new subdomain.

When evaluating a particular value for the commitment index,  $M_C$ , we construct transition probabilities using that value of  $M_C$  and run the lattice model for  $1 \times 10^7$  timesteps. We plot the mean-squared displacement for the lattice model and compare the resulting slope to that of the mean squared displacement of the continuous model. In practice, by comparing commitment index values ranging from three to seven, we are able to calibrate the commitment index for all cases considered, i.e. obstacle radius and particle radius configuration.

## 4.2 Estimating diffusivity using mean squared displacement and Taylor Series expansions on a lattice

In general, our method of comparison for the continuous model and surrogate model is via slopes of the mean squared displacement versus time. However, we would also like to use these results to estimate the diffusivity of the system for both the continuous and surrogate models. Furthermore, we can compare these estimated diffusivities with the theoretical diffusivity obtained using a Taylor Series expansion for random walks on the lattice. In cases with a high degree of symmetry, we can relate the diffusivity obtained via mean squared



displacement with that obtained via a Taylor Series expansion—for our models, this includes the equally spaced obstacles model (case (i)) and the multisize obstacles model (case (iii)).

The equally spaced obstacles model (case(i)) is the only model with equal probabilities of moving in the vertical and horizontal directions. We will also consider the theoretical and estimated diffusivities for the multisize obstacles model by including different probabilities of moving in the vertical direction or in the horizontal direction. The four internal obstacles case is too anisotropic to compare with a theoretical diffusivity obtained via a Taylor Series expansion.

#### 4.2.1 Previous work on using mean squared displacement to estimate diffusivity

First, we estimate diffusivity using the mean squared displacement  $\langle r^2 \rangle$  for both the continuous and surrogate models. In normal diffusion in two dimensions, the mean squared displacement can be related to the diffusivity by the equation

$$\langle r^2 \rangle = 4D t, \quad (4.1)$$

where  $t$  is time and  $D$  is the diffusion constant [32], [31]. Other studies have considered anomalous diffusion (diffusion in which movement is hindered or obstructed), as described by

$$\begin{aligned} \langle r^2 \rangle &= \Gamma t^\alpha \\ &= 4D(t)t, \end{aligned} \quad (4.2)$$

where  $\Gamma$  is a diffusivity constant,  $\alpha$  is a constant ( $\alpha > 0$ ), and  $D(t)$  is the diffusivity as a function of  $t$  [8]. Note that when  $D(t)$  is a constant, this reduces to equation (4.1). Generally, equation (4.2) holds when there are only a small number of timesteps considered, and, over longer times, this becomes a linear relationship between  $\langle r^2 \rangle$  and  $t$ . [8]. Since our goal is building a corresponding surrogate model, we will focus on the long term scenario in which we have a linear relationship between  $\langle r^2 \rangle$  and time. To ensure that the continuous and surrogate models are in the regime where this relationship is linear, we fit the resulting mean squared displacement to a power law. We need the constant  $\alpha \approx 1$  in order to consider our model as exhibiting normal diffusion (we find that  $\alpha \approx 1$  for all cases when the number of timesteps is  $N = 1 \times 10^7$ ).

Once we verify the model exhibits normal diffusion, we can estimate the diffusivity by

solving equation (4.1) for  $D$ :

$$D = \frac{\langle r^2 \rangle}{4t}. \quad (4.3)$$

Note for our case,  $N$  is the number of time steps of size  $\Delta t = 0.05$ , so  $t = N\Delta t$ . We can then estimate the diffusion constant  $D$ , by the following equation:

$$D = \frac{\langle r^2 \rangle}{4N\Delta t}. \quad (4.4)$$

## 4.2.2 Derivation of linear relation between mean squared displacement and time on a 2D lattice

In order to relate the estimated diffusivity obtained by mean squared displacement with the theoretical diffusivity obtained by a Taylor Series expansion (described in the next section), we need a relation between the mean squared displacement,  $\langle r^2 \rangle$  and the number of timesteps  $N$ . In this section, we will denote the mean squared displacement as  $\langle r_K^2(N) \rangle$ , with  $K$  referring to the number of single particle realizations of a random walk on the lattice, and  $N$  being a particular timestep. We will return to the notation  $\langle r^2 \rangle$  for mean squared displacement, starting in section 4.2.5.1.

### 4.2.2.1 Properties and notation for random walks on a 2D lattice

Consider  $K$  realizations of a single particle undergoing a random walk on a 2D lattice. In the surrogate model, the transition possibilities represent the following moves: (1) stay in current subdomain, (2) move to an adjacent subdomain in the horizontal direction, and (3) move to an adjacent subdomain in the vertical direction. If the particle moves, it can only move in a cardinal direction, thus we consider the next location as being due to steps in either the  $x$  or  $y$  directions. Let  $x_i(N)$  and  $y_i(N)$  denote the horizontal and vertical positions of the  $i$ th particle at timestep  $N$ , respectively. We can write them as

$$\begin{aligned} x_i(N) &= x_i(N-1) + h_{x_i}(N), \text{ for } i = 1, \dots, K, \\ y_i(N) &= y_i(N-1) + h_{y_i}(N), \text{ for } i = 1, \dots, K, \end{aligned} \quad (4.5)$$

where  $h_{xi}$  and  $h_{yi}$  are the possible displacements with associated stepsizes of  $\Delta x$  and  $\Delta y$ , respectively. In any one timestep, we have three possibilities for displacement,

$$\begin{aligned} \text{(i)} \quad & h_{xi}(N) = \pm\Delta x, \quad h_{yi}(N) = 0, \\ \text{(ii)} \quad & h_{xi}(N) = 0, \quad h_{yi}(N) = \pm\Delta y, \\ \text{(iii)} \quad & h_{xi}(N) = h_{yi}(N) = 0, \end{aligned} \tag{4.6}$$

where cases (i) (horizontal transition) and (ii) (vertical transition) occur with equal probability, and case (iii) (no transition) occurs with the highest probability.

#### 4.2.2.2 Some useful limits in evaluating mean squared displacement

In order to evaluate the mean squared displacement,

$$\langle r_K^2(N) \rangle = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N) + y_i^2(N), \tag{4.7}$$

we first identify a few limits that will arise in the derivation to follow:

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N), \tag{4.8}$$

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N)h_{xi}(M) \text{ when } N \neq M, \tag{4.9}$$

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N)h_{xi}(M) \text{ when } N = M. \tag{4.10}$$

We will evaluate the above limits for the horizontal transition case ( $h_{xi}$ ), but similar results also hold for the vertical transition case ( $h_{yi}$ ). We now evaluate each of the limits in equations (4.8)-(4.10).

##### 4.2.2.2.1 Evaluating equation (4.8)

We first consider the average displacement in the horizontal direction, over all realizations, as  $K \rightarrow \infty$ ,

$$\lim_{K \rightarrow \infty} \sum_{i=1}^K h_{xi}(N). \tag{4.11}$$

Let  $k_{x1}$ ,  $k_{x2}$ , and  $k_{x3}$  be the following:

- (i)  $k_{x1}$  : the number of times that  $h_{x1}(N) = 0$ ,

- (ii)  $k_{x2}$  : the number of times that  $h_{x1}(N) = \Delta x$ ,
- (iii)  $k_{x3}$  : the number of times that  $h_{x1}(N) = -\Delta x$ .

We can write

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} [k_{x1}(0) + k_{x2}(\Delta x) + k_{x3}(-\Delta x)] \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} [k_{x2}(\Delta x) + k_{x3}(-\Delta x)] \\
&= \lim_{K \rightarrow \infty} (\Delta x) \left( \frac{k_{x2}}{K} - \frac{k_{x3}}{K} \right).
\end{aligned} \tag{4.12}$$

For the equally spaced obstacles problem, moving in either the  $(\Delta x)$  and  $-(\Delta x)$  directions is equally likely. Let  $p_x$  be the probability that the particle moves in the horizontal direction. Then for very large values of  $K$ , the ratio

$$\frac{k_{x2}}{K} = \frac{p_x}{2},$$

and the ratio

$$\frac{k_{x3}}{K} = \frac{p_x}{2}.$$

We then write

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) \\
&= \lim_{K \rightarrow \infty} (\Delta x) \left( \frac{k_{x2}}{K} - \frac{k_{x3}}{K} \right) \\
&= \lim_{K \rightarrow \infty} (\Delta x) \left( \frac{p_x}{2} - \frac{p_x}{2} \right) \\
&= 0.
\end{aligned} \tag{4.13}$$

Following the same process for the vertical direction, we let  $k_{y1}$ ,  $k_{y2}$ , and  $k_{y3}$  be the following:

- (i)  $k_{y1}$  : the number of times that  $h_{y1}(N) = 0$ ,
- (ii)  $k_{y2}$  : the number of times that  $h_{y1}(N) = \Delta y$ ,
- (iii)  $k_{y3}$  : the number of times that  $h_{y1}(N) = -\Delta y$ .

Similarly, we find that

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{yi}(N) = 0. \quad (4.14)$$

#### 4.2.2.2.2 Evaluating equation (4.9)

We next want to show that, if  $N \neq M$ , then

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) h_{xi}(M) = 0. \quad (4.15)$$

We have three types of cases for  $h_{xi}(N)$  and  $h_{xi}(M)$ :

- (i)  $h_{xi}(N) = \Delta x$ , and  $h_{xi}(M) = -\Delta x$  (or vice versa),
- (ii)  $h_{xi}(N) = h_{xi}(M) = \Delta x$  or  $h_{xi}(N) = h_{xi}(M) = -\Delta x$ ,
- (iii)  $h_{xi}(N) = 0$  and/ or  $h_{xi}(M) = 0$ .

To describe the frequency of these cases, we define  $k_{x4}$ ,  $k_{x5}$ , and  $k_{x6}$  to be the following, respectively:

- (i)  $k_{x4}$ : the number of times that  $h_{xi}(N) = \Delta x$ , and  $h_{xi}(M) = -\Delta x$  (or vice versa),
- (ii)  $k_{x5}$ : the number of times that  $h_{xi}(N) = h_{xi}(M) = \Delta x$  or  $h_{xi}(N) = h_{xi}(M) = -\Delta x$ ,
- (iii)  $k_{x6}$ : the number of times that  $h_{xi}(N) = 0$  and/ or  $h_{xi}(M) = 0$ .

We illustrate the combinations of probabilities for the product in equation (4.15) in Figure 4.2. We now consider the limit:

$$\begin{aligned} \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) h_{xi}(M) \\ = \lim_{K \rightarrow \infty} \frac{1}{K} [k_{x4}(-(\Delta x)^2) + k_{x5}(\Delta x)^2 + k_{x6}(0)] \\ = \lim_{K \rightarrow \infty} (\Delta x)^2 \left[ \frac{k_{x5}}{K} - \frac{k_{x4}}{K} \right]. \end{aligned} \quad (4.16)$$

Reviewing Figure 4.2, we observe equal probabilities for events corresponding with  $k_{x4}$  and  $k_{x5}$ , and we expect

$$\lim_{K \rightarrow \infty} \frac{k_{x4}}{K} = \lim_{K \rightarrow \infty} \frac{k_{x5}}{K}. \quad (4.17)$$

		$h_{xi}(N)$		
		0	$\Delta x$	$-\Delta x$
$h_{xi}(M)$	0	$(1 - p_x)^2$ case (iii)	$\left(\frac{p_x}{2}\right)(1 - p_x)$ case (iii)	$\left(\frac{p_x}{2}\right)(1 - p_x)$ case (iii)
	$\Delta x$	$\left(\frac{p_x}{2}\right)(1 - p_x)$ case (iii)	$\left(\frac{p_x}{2}\right)^2$ case (ii)	$\left(\frac{p_x}{2}\right)^2$ case (i)
	$-\Delta x$	$\left(\frac{p_x}{2}\right)(1 - p_x)$ case (iii)	$\left(\frac{p_x}{2}\right)^2$ case (i)	$\left(\frac{p_x}{2}\right)^2$ case (ii)

**Figure 4.2** Illustration of probabilities for the three cases for independent events  $h_i(N)$  and  $h_i(M)$ , where  $p_x$  is the probability of moving horizontally. Cases (ii) and (iii) are associated with the same probability,  $(p_x/2)^2$ .

Hence,

$$\begin{aligned}
\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) h_{xi}(M) \\
= \lim_{K \rightarrow \infty} (\Delta x)^2 \left[ \frac{k_{x5}}{K} - \frac{k_{x4}}{K} \right] \\
= 0.
\end{aligned} \tag{4.18}$$

For the vertical direction, we define  $k_{y4}$ ,  $k_{y5}$ , and  $k_{y6}$  be the following:

- i)  $k_{y4}$  : the number of times that  $h_{yi}(N) = \Delta y$ , and  $h_{yi} = -\Delta y$  (or vice versa),
- ii)  $k_{y5}$  : the number of times that  $h_{yi}(N) = h_{yi}(M) = \Delta y$  or  $h_{yi}(N) = h_{yi}(M) = -\Delta y$ ,
- iii)  $k_{y6}$  : the number of times that  $h_{yi}(N) = 0$  and/ or  $h_{yi}(M) = 0$ .

We can then follow a similar line of reasoning as in the horizontal case to determine that

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{yi}(N) h_{yi}(M) = 0, \tag{4.19}$$

when  $N \neq M$ .

#### 4.2.2.2.3 Evaluating equation (4.10)

In the case  $N = M$ , and with  $k_{x1}$ ,  $k_{x2}$ , and  $k_{x3}$  defined as above (in the evaluation of equation (4.8)),

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) h_{xi}(M) \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}^2(N) \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} [k_{x1}(0) + k_{x2}(\Delta x)^2 + k_{x3}(-\Delta x)^2] \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} [k_{x1}(0) + (k_{x2} + k_{x3})(\Delta x)^2] \\
&= \lim_{K \rightarrow \infty} \left( \frac{k_{x2} + k_{x3}}{K} \right) (\Delta x)^2.
\end{aligned} \tag{4.20}$$

Recalling that  $p_x$  was defined above to be the probability of moving either horizontal direction, then

$$\lim_{K \rightarrow \infty} \left( \frac{k_{x2} + k_{x3}}{K} \right) = p_x. \tag{4.21}$$

Hence, for  $N = M$ ,

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) h_{xi}(M) \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}^2(N) \\
&= \lim_{K \rightarrow \infty} \left( \frac{k_{x2} + k_{x3}}{K} \right) (\Delta x)^2 \\
&= p_x (\Delta x)^2.
\end{aligned} \tag{4.22}$$

We can again follow a similar line of reasoning to obtain a similar result in the vertical direction.

#### 4.2.2.2.4 Summary of limits in equations (4.8) - (4.10)

In summary, we have derived the following values for the limits introduced in equations (4.8), (4.9), and (4.10), respectively:

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) = 0, \quad (4.23)$$

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) h_{xi}(M) = 0 \text{ when } N \neq M, \quad (4.24)$$

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_{xi}(N) h_{xi}(M) = p_x(\Delta x)^2 \text{ when } N = M. \quad (4.25)$$

#### 4.2.2.3 Limits using induction to evaluate mean squared displacement

In considering the mean squared displacement on a lattice, there are two additional limits involving movements of particles that can be simplified:

$$\lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N) h_{xi}(M) = 0 \text{ when } M > N, \quad (4.26)$$

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N) = N p_x(\Delta x)^2. \quad (4.27)$$

We will evaluate the above limits for the horizontal transition case ( $x_i$ ), but similar results also hold for the vertical transition case ( $y_i$ ). To derive these limits we will use equations (4.23) - (4.25), as well as equation (4.5) for a particle's possible horizontal locations:

$$x_i(N) = x_i(N-1) + h_{xi}(N), \quad (4.28)$$

##### 4.2.2.3.1 Evaluating equation (4.26)

We use induction to show that

$$\lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N) h_{xi}(M) = 0 \quad (4.29)$$

when  $M > N$ .

When  $N = 0$ , the left hand side of (4.29) becomes

$$\lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(0) h_{xi}(M), \quad M > 0. \quad (4.30)$$



The value of  $x_i(N) = 0$  for all particles  $i = 1, \dots, K$ , since this indicates the initial starting location on the lattice. Hence,

$$\lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N) h_{xi}(M) = 0, \text{ when } N \neq M. \quad (4.31)$$

When  $N = 1$ , equation (4.29) becomes

$$\begin{aligned} & \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(1) h_{xi}(M), \quad M > 1 \\ &= \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K (x_i(0) + h_{xi}(1)) h_{xi}(M) \\ &= \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K (x_i(0) h_{xi}(M) + h_{xi}(1) h_{xi}(M)) \\ &= \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(0) h_{xi}(M) + \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K h_{xi}(1) h_{xi}(M) \\ &= \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K 0 h_{xi}(M) + \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K h_{xi}(1) h_{xi}(M) \\ &= \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K h_{xi}(1) h_{xi}(M). \end{aligned} \quad (4.32)$$

Applying our result in equation (4.24) to the last line above, it then follows that:

$$\lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(1) h_{xi}(M) = 0. \quad (4.33)$$

For  $N > 1$ , we employ mathematical induction, first assuming that

$$\lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N) h_{xi}(M) = 0, \text{ when } M > N, \quad (4.34)$$

holds for  $N > 1$ . We now show that it also holds for  $N + 1 > 2$ :

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N+1)h_{xi}(M) \\
&= \frac{2}{K} \sum_{i=1}^K (x_i(N) + h_{xi}(N+1))h_{xi}(M) \\
&= \frac{2}{K} \sum_{i=1}^K (x_i(N)h_{xi}(M) + h_{xi}(N+1)h_{xi}(M)) \\
&= \frac{2}{K} \sum_{i=1}^K x_i(N)h_{xi}(M) + \frac{2}{K} \sum_{i=1}^K h_{xi}(N+1)h_{xi}(M).
\end{aligned} \tag{4.35}$$

Using the inductive hypothesis in equation (4.34) and equation (4.24) (i.e. both terms are zero), we then obtain the result:

$$\lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N+1)h_{xi}(M) = 0 \quad \text{for } M > N. \tag{4.36}$$

#### 4.2.2.3.2 Evaluating equation (4.27)

We now use mathematical induction to show that

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N) = N p_x(\Delta x)^2. \tag{4.37}$$

When  $N = 1$ , equation (4.37) becomes

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(1) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K (x_i(0) + h_i(1))^2. \tag{4.38}$$

We note that  $x_i(0)$  is simply the starting location, which is recorded as  $x_i(0) = 0$ . Hence,

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(1) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_i^2(1) = p_x(\Delta x)^2, \tag{4.39}$$

by (4.25).

When  $N = 2$ , equation (4.37) becomes

$$\begin{aligned}
\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(2) &= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K (x_i(1) + h_i(2))^2 \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K (x_i^2(1) + 2x_i(1)h_i(2) + h_i^2(2)) \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(1) + \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(1)h_i(2) + \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_i^2(2).
\end{aligned} \tag{4.40}$$

Using (4.25), we can rewrite this as

$$\begin{aligned}
\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(2) &= 2p_x(\Delta x)^2 + \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(1)h_i(2) \\
&= 2p_x(\Delta x)^2 + 0,
\end{aligned} \tag{4.41}$$

by equation (4.26).

We now consider the inductive step, assuming that, for  $N > 2$ ,

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N) = Np_x(\Delta x)^2, \tag{4.42}$$

holds for  $N$ . We now show that it also holds for  $N + 1$ :

$$\begin{aligned}
\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N+1) &= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K (x_i(N) + h_i(N+1))^2 \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K (x_i^2(N) + 2x_i(N)h_i(N+1) + h_i^2(N+1)) \\
&= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N) + \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N)h_i(N+1) \\
&\quad + \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K h_i^2(N+1).
\end{aligned} \tag{4.43}$$

Using the inductive hypothesis in equation (4.42) and equation (4.25), we obtain:

$$\begin{aligned}
\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N+1) &= Np_x(\Delta x)^2 + \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N)h_i(N+1) + p_x(\Delta x)^2 \\
&= (N+1)p_x(\Delta x)^2 + \lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N)h_i(N+1).
\end{aligned} \tag{4.44}$$

Using (4.26), the last term above is zero, so that:

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N+1) = (N+1)p_x(\Delta x)^2. \quad (4.45)$$

Thus

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N+1) = Np_x(\Delta x)^2, \quad \text{for all } N \geq 1. \quad (4.46)$$

#### 4.2.2.3.3 Summary of limits in equations (4.26) and (4.27)

In summary,

$$\lim_{K \rightarrow \infty} \frac{2}{K} \sum_{i=1}^K x_i(N)h_{xi}(M) = 0 \quad \text{for } M > N, \quad (4.47)$$

$$\lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N+1) = Np_x(\Delta x)^2. \quad (4.48)$$

Both (4.47) and (4.48) also hold in the  $y$  direction due to similar reasoning.

#### 4.2.2.4 Evaluating the mean squared displacement

We now want to use the properties from Sections 4.2.2.2 and 4.2.2.3 to evaluate the mean squared displacement of  $K$  random walkers,  $\langle r_K^2(N) \rangle$ , in order to show that the mean squared displacement is linear in time. Recall equation (4.7):

$$\langle r_K^2(N) \rangle = \frac{1}{K} \sum_{i=1}^K x_i^2(N) + y_i^2(N), \quad (4.49)$$

Substituting our results in (4.48) into (4.49), the mean squared displacement becomes

$$\begin{aligned} \langle r_K^2(N) \rangle &= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N) + y_i^2(N) \\ &= Np_x(\Delta x)^2 + Np_y(\Delta y)^2 \\ &= N[p_x(\Delta x)^2 + p_y(\Delta y)^2]. \end{aligned} \quad (4.50)$$

Since in the equally spaced lattice model,  $\Delta x = \Delta y$ , we can simplify (4.50) as follows:

$$\begin{aligned}
\langle r_K^2(N) \rangle &= \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K x_i^2(N) + y_i^2(N) \\
&= N [p_x(\Delta x)^2 + p_y(\Delta y)^2] \\
&= N [p_x(\Delta x)^2 + p_y(\Delta x)^2] \\
&= N(p_x + p_y)(\Delta x)^2.
\end{aligned} \tag{4.51}$$

In (4.51),  $p_x$  and  $p_y$  are the probabilities of making a transition in the horizontal and vertical directions, respectively. Hence, the mean squared displacement,  $\langle r_K^2(N) \rangle$ , grows linearly in time; i.e. it increases in linear proportion to the current number of timesteps  $N$ . In the surrogate lattice model, we denote  $q$  as the probability of staying in the current subdomain, and thus  $1 = q + p_x + p_y$ . We can rewrite equation (4.51) as

$$\langle r_K^2(N) \rangle = N(1 - q)(\Delta x)^2 \tag{4.52}$$

In this section, we've shown that mean squared displacement grows linearly in time. We will come back to these results after deriving the diffusion equation using a Taylor Series expansion in the next section.

### 4.2.3 Deriving the diffusion equation for the equally spaced obstacles model

In addition to approximating the diffusivity using the mean squared displacement, we also want to derive an equation for the theoretical diffusivity in the lattice model. We aim to compare the theoretical diffusivity (discussed in this section) to the estimated diffusivities, obtained using the mean squared displacement from our simulations of the surrogate model. Among the models considered, the equally spaced obstacles model is the most reasonable model to compare diffusivities since the particle has equal probabilities of moving to subdomains in the horizontal or vertical direction.

We begin by using a Taylor Series expansion to derive the diffusion equation for a particle which can move in cardinal directions in two dimensions. Let  $u(x, y, t)$  be the density of particles (assuming a fixed volume and uniform particle mass), and  $q$  be the probability that a particle stays in its current subdomain. We assume all four directions (north, east, south, west) occur with equal probability, and the probability of a transition to a neighboring lattice point is  $(1 - q)$ . The resulting particle density at location  $(x, y)$  after a

single timestep  $\Delta t$  is given by

$$u(x, y, t + \Delta t) = qu(x, y, t) + \frac{1}{4}(1 - q)[u(x + h, y, t) + u(x - h, y, t) + u(x, y + k, t) + u(x, y - k, t)], \quad (4.53)$$

where  $h$  horizontal step size and  $k$  is the vertical step size. Multiplying both sides of equation (4.53) by 4, we have,

$$4u(x, y, t + \Delta t) = 4qu(x, y, t) + (1 - q)[u(x - h, y, t) + u(x + h, y, t) + u(x, y + k, t) + u(x, y - k, t)]. \quad (4.54)$$

Consider the following Taylor Series expansions:

$$u(x, y, t + \Delta t) = u(x, y, t) + u_t(x, y, t)(\Delta t) + \frac{1}{2}u_{tt}(x, y, t)(\Delta t)^2 + \frac{1}{6}u_{ttt}(x, y, t)(\Delta t)^3 + \quad (4.55)$$

$$u(x + h, y, t) = u(x, y, t) + u_x(x, y, t)(\Delta x) + \frac{1}{2}u_{xx}(x, y, t)(\Delta x)^2 + \frac{1}{6}u_{xxx}(x, y, t)(\Delta x)^3 + \quad (4.56)$$

$$u(x - h, y, t) = u(x, y, t) - u_x(x, y, t)(\Delta x) + \frac{1}{2}u_{xx}(x, y, t)(\Delta x)^2 - \frac{1}{6}u_{xxx}(x, y, t)(\Delta x)^3 + \quad (4.57)$$

$$u(x, y + k, t) = u(x, y, t) + u_y(x, y, t)(\Delta y) + \frac{1}{2}u_{yy}(x, y, t)(\Delta y)^2 + \frac{1}{6}u_{yyy}(x, y, t)(\Delta y)^3 + \quad (4.58)$$

$$u(x, y - k, t) = u(x, y, t) - u_y(x, y, t)(\Delta y) + \frac{1}{2}u_{yy}(x, y, t)(\Delta y)^2 - \frac{1}{6}u_{yyy}(x, y, t)(\Delta y)^3 + \quad (4.59)$$

Using these Taylor Series expansions in equation (4.54), we have

$$\begin{aligned} & 4[ u + u_t(\Delta t) + \frac{1}{2}u_{tt}(\Delta t)^2 + \frac{1}{6}u_{ttt}(\Delta t)^3 + \mathcal{O}((\Delta t)^4) ] \\ &= 4qu + (1 - q)[ u + u_x(\Delta x) + \frac{1}{2}u_{xx}(\Delta x)^2 + \frac{1}{6}u_{xxx}(\Delta x)^3 + \mathcal{O}((\Delta x)^4) \\ &+ u - u_x(\Delta x) + \frac{1}{2}u_{xx}(\Delta x)^2 - \frac{1}{6}u_{xxx}(\Delta x)^3 + \mathcal{O}((\Delta x)^4) \\ &+ u + u_y(\Delta y) + \frac{1}{2}u_{yy}(\Delta y)^2 + \frac{1}{6}u_{yyy}(\Delta y)^3 + \mathcal{O}((\Delta y)^4) \\ &+ u - u_y(\Delta y) + \frac{1}{2}u_{yy}(\Delta y)^2 - \frac{1}{6}u_{yyy}(\Delta y)^3 + \mathcal{O}((\Delta y)^4) ]. \end{aligned} \quad (4.60)$$

Simplifying the above equation, we obtain

$$\begin{aligned} & 4[ u_t(\Delta t) + \frac{1}{2}u_{tt}(\Delta t)^2 + \frac{1}{6}u_{ttt}(\Delta t)^3 + \mathcal{O}((\Delta t)^4) ] \\ &= (1 - q)[ u_{xx}(\Delta x)^2 + \mathcal{O}((\Delta x)^4) + u_{yy}(\Delta y)^2 + \mathcal{O}((\Delta y)^4) ], \end{aligned} \quad (4.61)$$

which can be reduced to

$$u_t(\Delta t) + \mathcal{O}((\Delta t)^2) = \frac{(1-q)}{4} [u_{xx}(\Delta x)^2 + u_{yy}(\Delta y)^2 + \mathcal{O}((\Delta y)^4) + \mathcal{O}(\Delta x)^4]. \quad (4.62)$$

For our model we assume a uniform 2D square lattice, so that  $\Delta x = \Delta y$ . Using this fact and dividing by  $\Delta t$ , we can rewrite equation (4.62) as

$$u_t + \mathcal{O}(\Delta t) = \frac{(1-q)}{4} \left( u_{xx} \frac{(\Delta x)^2}{(\Delta t)} + u_{yy} \frac{(\Delta x)^2}{(\Delta t)} + \mathcal{O}\left(\frac{(\Delta x)^4}{(\Delta t)}\right) \right). \quad (4.63)$$

Finally, we can write the diffusion equation,

$$u_t \approx \frac{(1-q)(\Delta x)^2}{4(\Delta t)} (u_{xx} + u_{yy}), \quad (4.64)$$

where the diffusion constant,  $D$ , is given by

$$D = \frac{(1-q)(\Delta x)^2}{4(\Delta t)}. \quad (4.65)$$

#### 4.2.4 Deriving the diffusion equation for the multisize obstacles model

Due to obstacle arrangements in the multisize obstacles model, the particle has a different diffusivity in the vertical direction than in the horizontal direction. We follow the same procedure for deriving the diffusion equation as in section 4.2.3, but this time we denote

- (i)  $q_1$  is the probability that the particle stays in its current subdomain;
- (ii)  $q_2$  is the probability that the particle moves to a subdomain in the horizontal direction;
- (iii)  $q_3$  is the probability that the particle moves to a subdomain in the vertical direction.

Note that  $q_1 + q_2 + q_3 = 1$ . We can then write the relation,

$$\begin{aligned} u(x, y, t + \Delta t) = & q_1 u(x, y, t) + \frac{q_2}{2} u(x + h, y, t) + \frac{q_2}{2} u(x - h, y, t) + \frac{q_3}{2} u(x, y + k, t) \\ & + \frac{q_3}{2} u(x, y - k, t). \end{aligned} \quad (4.66)$$

Using the Taylor Series expansions in equations (4.55) - (4.59), we can rewrite the above equation as

$$\begin{aligned}
& u + u_t(\Delta t) + \frac{1}{2}u_{tt}(\Delta t)^2 + \frac{1}{6}u_{ttt}(\Delta t)^3 + \mathcal{O}((\Delta t)^4) \\
&= q_1 u + \frac{q_2}{2} \left( u + u_x(\Delta x) + \frac{1}{2}u_{xx}(\Delta x)^2 + \frac{1}{6}u_{xxx}(\Delta x)^3 + \mathcal{O}((\Delta x)^4) \right) \\
&+ \frac{q_2}{2} \left( u - u_x(\Delta x) + \frac{1}{2}u_{xx}(\Delta x)^2 - \frac{1}{6}u_{xxx}(\Delta x)^3 + \mathcal{O}((\Delta x)^4) \right) \\
&+ \frac{q_3}{2} \left( u + u_y(\Delta y) + \frac{1}{2}u_{yy}(\Delta y)^2 + \frac{1}{6}u_{yyy}(\Delta y)^3 + \mathcal{O}((\Delta y)^4) \right) \\
&+ \frac{q_3}{2} \left( u - u_y(\Delta y) + \frac{1}{2}u_{yy}(\Delta y)^2 - \frac{1}{6}u_{yyy}(\Delta y)^3 + \mathcal{O}((\Delta y)^4) \right).
\end{aligned} \tag{4.67}$$

Simplifying, we have

$$\begin{aligned}
& u + u_t(\Delta t) + \mathcal{O}((\Delta t)^2) \\
&= q_1 u + q_2 u + q_3 u + \frac{q_2}{2} (u_{xx}(\Delta x)^2 + \mathcal{O}((\Delta x)^4)) + \frac{q_3}{2} (u_{yy}(\Delta y)^2 + \mathcal{O}((\Delta y)^4)).
\end{aligned} \tag{4.68}$$

Using the fact that  $q_1 + q_2 + q_3 = 1$ , and dividing by  $\Delta t$ , we can rewrite equation (4.68) as,

$$u_t + \mathcal{O}((\Delta t)) = \frac{q_2}{2} u_{xx} \frac{(\Delta x)^2}{(\Delta t)} + \frac{q_3}{2} u_{yy} \frac{(\Delta y)^2}{(\Delta t)} + \frac{\mathcal{O}((\Delta x)^4 + (\Delta y)^4)}{(\Delta t)}. \tag{4.69}$$

For our model, we assume a uniform 2D square lattice so that  $\Delta x = \Delta y$ . We simplify to find the diffusion equation for the multisize obstacles,

$$u_t = \frac{(\Delta x)^2}{2(\Delta t)} (q_2 u_{xx} + q_3 u_{yy}). \tag{4.70}$$

We can then consider the horizontal diffusivity,  $D_H$ , to be

$$D_H = \frac{q_2(\Delta x)^2}{2(\Delta t)}, \tag{4.71}$$

and the vertical diffusivity,  $D_V$ , to be

$$D_V = \frac{q_3(\Delta x)^2}{2(\Delta t)}. \tag{4.72}$$



#### 4.2.4.1 Consistency between the multisize and equally spaced obstacles models

We should be able to obtain the diffusion equation for the equally spaced obstacles model (eq. (4.64)) using the resulting diffusion equation for the multisize obstacles model (eq. (5.15)) by setting  $q_2 = q_3$ , since, in the equally spaced obstacles model, there is an equal probability of moving vertically or horizontally. If we let  $q_2 = q_3$ , we can rewrite equation (5.15) as

$$u_t = \frac{(\Delta x)^2}{2(\Delta t)} (q_2 u_{xx} + q_2 u_{yy}). \quad (4.73)$$

We can simplify this as

$$u_t = \frac{q_2(\Delta x)^2}{2(\Delta t)} (u_{xx} + u_{yy}). \quad (4.74)$$

The relation

$$q_1 + q_2 + q_3 = 1 \Rightarrow q_2 + q_3 = 1 - q_1. \quad (4.75)$$

When  $q_2 = q_3$ , we can write

$$1 - q_1 = 2q_2 \Rightarrow q_2 = \frac{(1 - q_1)}{2}. \quad (4.76)$$

Using this result in equation (4.74), we have

$$u_t = \frac{(1 - q_1)(\Delta x)^2}{4(\Delta t)} (u_{xx} + u_{yy}), \quad (4.77)$$

which is the diffusion equation for the equally spaced obstacles model (eq. (4.64)), as obtained in section 4.2.3.

## 4.2.5 Relating estimated diffusivity and theoretical diffusivity

We will now use the notation  $D_{\text{surr}}$  and  $D_{\text{TS}}$  as the diffusion constants for the surrogate model and theoretical calculation, respectively. Table 4.1 contains a summary of the notation used in this chapter. Recall from Chapter 3, that the estimate for the diffusion constant from the surrogate mean squared displacement simulations is

$$D_{\text{surr}} = \frac{\langle r_{\text{surr}}^2 \rangle}{4N\Delta t}, \quad (4.78)$$

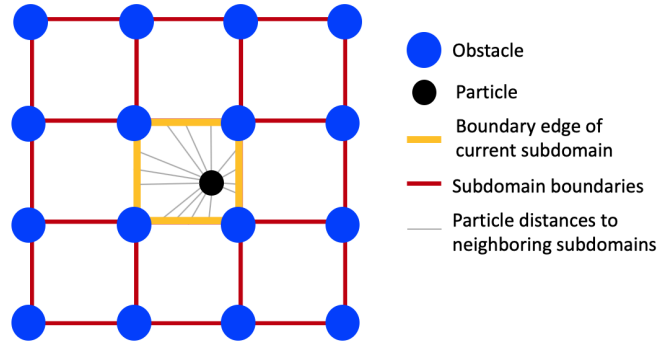
and that the estimate for the diffusion constant using a Taylor Series expansion (eq. 4.65) is

$$D_{\text{TS}} = \frac{(1 - q)(\Delta x)^2}{4\Delta t}. \quad (4.79)$$

**Table 4.1** Notation for different variables and parameters used in calculating the diffusivity.

Notation	Description
$D_{\text{surr}}$	Diffusivity for surrogate model obtained via simulation results
$D_{\text{TS}}$	Diffusivity for surrogate model obtained via Taylor Series expansion
$\langle r_{\text{surr}}^2 \rangle$	Mean squared displacement of surrogate model, obtained via simulation results
$N$	Number of time steps for surrogate model
$q$	Probability that the particle stays in the current subdomain
$p_x$	Probability that a particle moves to a subdomain to the east or to the west (i.e. the particle makes a horizontal transition)
$p_y$	Probability that a particle moves to a subdomain to the north or to the south (i.e. the particle makes a vertical transition)

In (4.79)  $\Delta x$  is the appropriate step-size for the lattice model. We want to relate these two estimates for diffusivity. Our first step is making an appropriate choice for  $\Delta x$  in equation (4.79), which is a nontrivial task.



**Figure 4.3** Illustration of possible linear distances for a particle to leave the subdomain at when the particle is at a particular location.

#### 4.2.5.1 Selecting the appropriate $\Delta x$ for theoretical diffusivity

In order to continue relating  $D_{\text{surr}}$  and  $D_{\text{TS}}$ , we need to choose an appropriate  $\Delta x$ . In the continuous model, a particle at any arbitrary location in a subdomain has many different net distances it could travel to reach the next subdomain. Figure 4.3 illustrates possible linear distances to neighboring subdomains given a particular particle location. From this figure, we can see that the particle's distance to a neighboring subdomain is dependent

upon where it is located in the current subdomain and over which boundary it will cross. Our aim in this section will be to determine an approximate average value for  $\Delta x$ .

In the continuous model each subdomain has side length  $L = 10$ , and our stepsize in time is  $\Delta t = 0.05$ . When the particle is at a node representing a particular subdomain in the continuous model, we can record the particle's location (for the purposes of the surrogate model) as the  $(x, y)$  coordinates at the center of that subdomain in the continuous model. Thus, a natural inclination for the surrogate model stepsize in space is  $\Delta x = \Delta y = L = 10$ . However, based on Figure 4.3, it is unclear if  $\Delta x = L = 10$  is the best choice, and we need to determine the appropriately sized  $\Delta x$  for estimating the diffusivity  $D_{\text{TS}}$ .

We use results from earlier in this chapter (sec. 4.2.2.4) to relate the two equations for the diffusion constant (i.e. eq. (4.78) and (4.79)). Our diffusivity  $D_{\text{surr}}$  comes from the mean squared displacement results based on the surrogate model, which we will refer to as  $\langle r_{\text{surr}}^2 \rangle$ . Recall, from equation (4.7) that :

$$\langle r_{\text{surr}}^2 \rangle = \frac{1}{K} \sum_{i=1}^K (x_i^2 + y_i^2), \quad (4.80)$$

where  $x_i$  and  $y_i$  are horizontal and vertical locations, respectively, and  $K$  is the number of realizations. As  $K \rightarrow \infty$ , we recall from equation (4.52) that:

$$\begin{aligned} \langle r_{\text{surr}}^2 \rangle &= \frac{1}{K} \sum_{i=1}^K (x_i^2 + y_i^2) \\ &= N(1-q)(\Delta x)^2, \end{aligned} \quad (4.81)$$

where  $N$  is the number of timesteps. Solving for  $N$ ,

$$N = \frac{\langle r_{\text{surr}}^2 \rangle}{(1-q)(\Delta x)^2}. \quad (4.82)$$

We can substitute  $N$  from equation (4.82) into equation (4.78), yielding

$$\begin{aligned} D_{\text{surr}} &= \frac{\langle r_{\text{surr}}^2 \rangle}{4 \left( \frac{\langle r_{\text{surr}}^2 \rangle}{(1-q)(\Delta x)^2} \right) \Delta t} \\ &= \frac{\langle r_{\text{surr}}^2 \rangle}{\langle r_{\text{surr}}^2 \rangle} \cdot \frac{(1-q)(\Delta x)^2}{4\Delta t} \\ &= \frac{\langle r_{\text{surr}}^2 \rangle}{\langle r_{\text{surr}}^2 \rangle} \cdot D_{\text{TS}}. \end{aligned} \quad (4.83)$$

This result shows us that  $D_{\text{surr}} = D_{\text{TS}}$ , but we must still choose an appropriate  $\Delta x$ .

In summary, after choosing an appropriate value of  $\Delta x$ , we expect

$$\begin{aligned} D_{\text{surr}} &\approx D_{\text{TS}} \\ &\approx \frac{(1-q)(\Delta x)^2}{4\Delta t} \end{aligned} \quad (4.84)$$

We will estimate  $\Delta x$  in two ways:

- (i) Comparing the ratio of  $D_{\text{surr}}$  to  $D_{\text{TS}}$ . Since we determine that  $D_{\text{surr}} \approx D_{\text{TS}}$ , if we determine the ratio by which these two values differ using an initial  $\Delta x = 10$ , we can use that ratio to approximate the effective  $\Delta x$ .
- (ii) Using line integrals to find the average distance for an arbitrarily located particle to a neighboring subdomain.

#### 4.2.6 Comparing the ratio of $D_{\text{surr}}$ to $D_{\text{TS}}$ to estimate $\Delta x$

We relate  $\Delta x$  to  $L$  via a constant  $\alpha$  where

$$\Delta x = \frac{L}{\alpha} . \quad (4.85)$$

We can then rewrite equation (4.84) as

$$\begin{aligned} D_{\text{surr}} &\approx D_{\text{TS}} \\ &\approx \frac{(1-q)(\Delta x)^2}{4\Delta t} \\ &\approx \frac{(1-q)(\frac{L}{\alpha})^2}{4\Delta t} \\ &\approx \frac{(1-q)L^2}{4\alpha^2\Delta t} \\ &\approx \frac{1}{\alpha^2} \frac{(1-q)L^2}{4\Delta t} \\ &\approx \frac{1}{\alpha^2} D_{\text{TS}}|_{\Delta x=10} , \end{aligned} \quad (4.86)$$

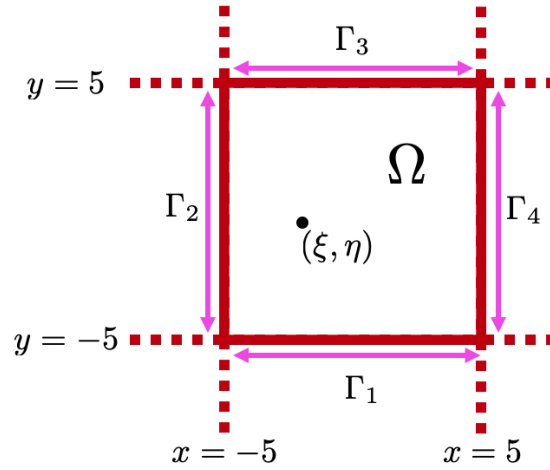
where  $D_{\text{TS}}|_{\Delta x=10}$  indicates  $D_{\text{TS}}$  when  $\Delta x = L = 10$ . Hence, once we determine the ratio,

$$\frac{D_{\text{TS}}|_{\Delta x=10}}{D_{\text{surr}}} \approx \alpha^2, \quad (4.87)$$

we can use the estimated  $\alpha$  to find the appropriate  $\Delta x$  for the Taylor Series approximation by equation (4.85).

### 4.2.7 A theoretical estimate for $\Delta x$

We use a set of boundary integrals to evaluate a particle's average distance from the next subdomain in order to help determine the appropriate  $\Delta x$  for the lattice model. We describe the particle as being located at an arbitrary position  $(\xi, \eta)$  in a subdomain,  $\Omega$ . We first evaluate the average distance when there are no obstacles and then consider the case where there are obstacles. For the no obstacles case, the particle radius does not need to be considered because (1) the center of the particle determines its subdomain location, and (2) when there are no obstacles, there is nothing for the particle to collide with. However, in the case with obstacles, we need to consider the area that will be excluded due to particle-obstacle collisions.



**Figure 4.4** Subdomain  $\Omega$  with a particle located at an arbitrary position  $(\xi, \eta)$ . Boundaries are denoted as  $\Gamma_i$  for  $i = 1, 2, 3, 4$ .

#### 4.2.7.1 Average distance between subdomains (no obstacles)

For the no obstacle case, Figure 4.4 shows a subdomain with its boundary edges and particle location labeled. We introduce the following line integrals, written in terms of the

particle's distance to all points along each boundary:

$$\begin{aligned}
\text{(i) For } \Gamma_1 : I_1 &= \int_{-5}^5 \sqrt{(s-\xi)^2 + (-5-\eta)^2} ds, \\
\text{(ii) For } \Gamma_2 : I_2 &= \int_{-5}^5 \sqrt{(-5-\xi)^2 + (s-\eta)^2} ds, \\
\text{(ii) For } \Gamma_3 : I_3 &= \int_{-5}^5 \sqrt{(s-\xi)^2 + (5-\eta)^2} ds, \\
\text{(ii) For } \Gamma_4 : I_4 &= \int_{-5}^5 \sqrt{(5-\xi)^2 + (s-\eta)^2} ds.
\end{aligned} \tag{4.88}$$

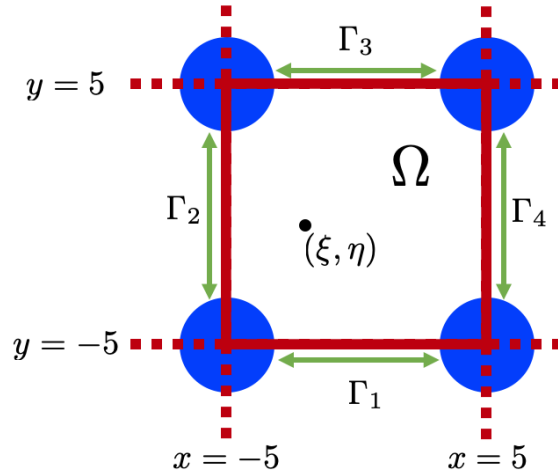
Since the length of each boundary edge is  $L = 10$ , the total length of the subdomain's boundary is  $4L = 40$ . Hence, we can represent the average distance that the particle travels from  $(\xi, \eta)$  to any neighboring subdomain as

$$\frac{1}{40} (I_1 + I_2 + I_3 + I_4). \tag{4.89}$$

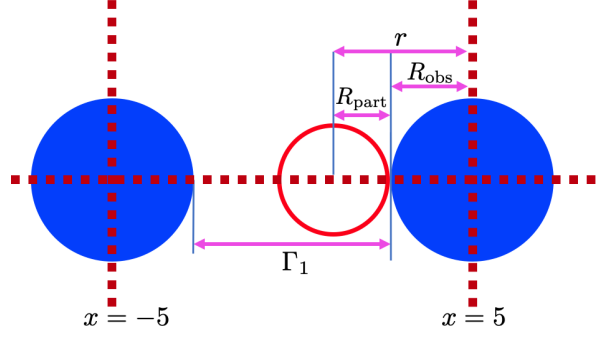
Over the entire subdomain, the average distance from a particle to any neighboring subdomain (according to equation (4.89)) is

$$d_{\text{avg}} = \frac{1}{100} \int_{-5}^5 \int_{-5}^5 \left( \frac{I_1 + I_2 + I_3 + I_4}{40} \right) d\xi d\eta. \tag{4.90}$$

Using Maple, we find an approximate value of  $d_{\text{avg}} \approx 6.5176$ .



**Figure 4.5** Subdomain  $\Omega$  with a particle located at an arbitrary position  $(\xi, \eta)$ .



**Figure 4.6** Close up of subdomain  $\Omega$  to show distances  $R_{\text{part}}$ ,  $R_{\text{obs}}$ , and  $r$ . Distance  $r$  must be included in calculation of average distance a particle is from a neighboring subdomain.

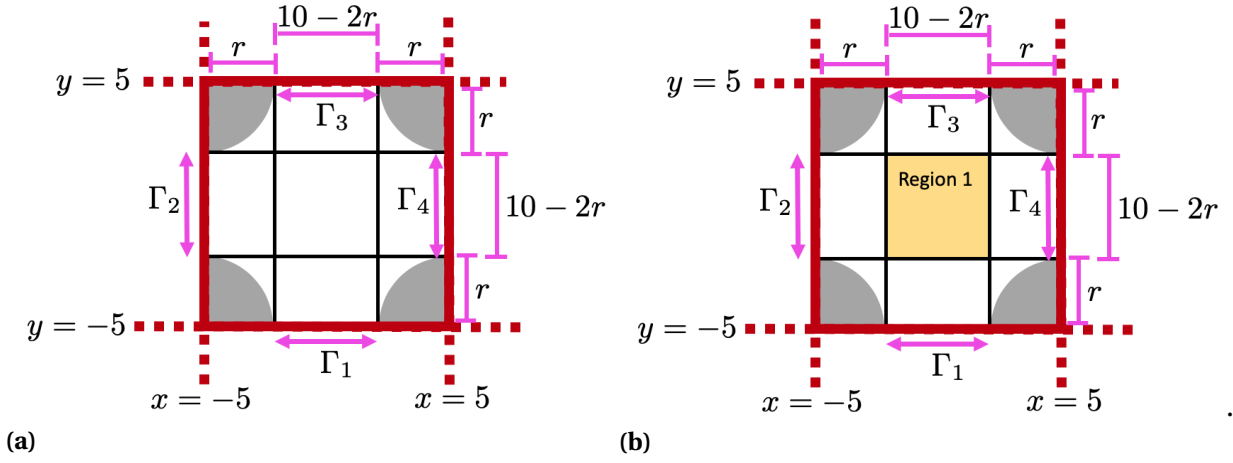
#### 4.2.7.2 Average distance between subdomains with obstacles

For the case with equally spaced obstacles, Figure 4.5 shows the domain with boundary edges and particle location labeled. In our calculations, we need to consider both the obstacle radius  $R_{\text{obs}}$  and the particle radius  $R_{\text{part}}$ . This will lead to an excluded distance  $r$  where  $r = R_{\text{obs}} + R_{\text{part}}$  (see Fig. 4.6). We make the following observations regarding the different boundaries:

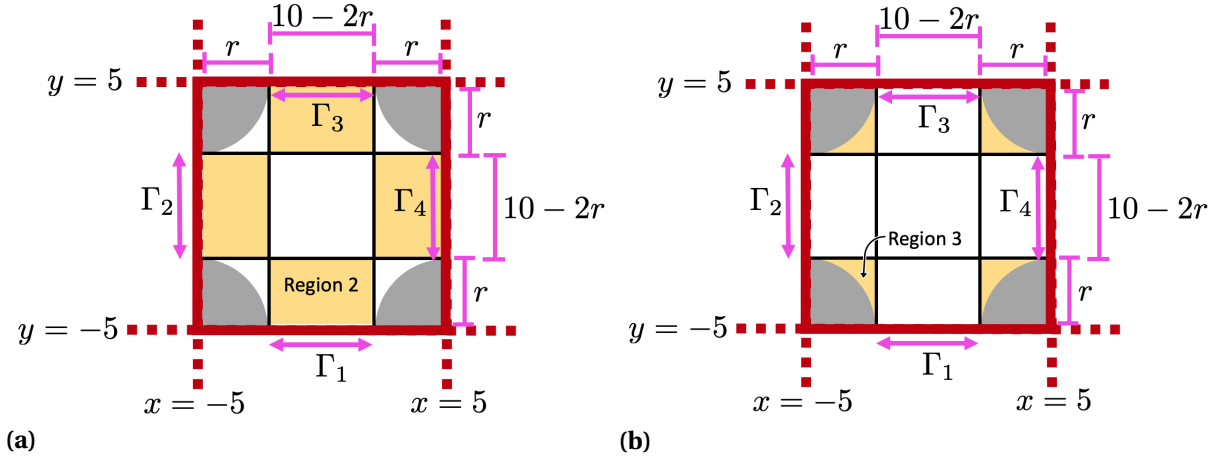
- (i)  $\Gamma_1$  has a  $y$ -axis boundary at  $y = -5$ , and varies along the  $x$ -axis for  $x \in [-5 + r, 5 - r]$ .
- (ii)  $\Gamma_2$  has a  $x$ -axis boundary at  $x = -5$ , and varies along the  $y$ -axis for  $y \in [-5 + r, 5 - r]$ .
- (iii)  $\Gamma_3$  has a  $y$ -axis boundary at  $y = 5$ , and varies along the  $x$ -axis for  $x \in [-5 + r, 5 - r]$ .
- (iv)  $\Gamma_4$  has a  $x$ -axis boundary at  $x = 5$ , and varies along the  $y$ -axis for  $y \in [-5 + r, 5 - r]$ .

We introduce the following line integrals, written in terms of the particle's distance to all points along each boundary:

$$\begin{aligned}
 \text{(i) For } \Gamma_1 : I_1 &= \int_{-5+r}^{5-r} \sqrt{(s-\xi)^2 + (-5-\eta)^2} ds \\
 \text{(ii) For } \Gamma_2 : I_2 &= \int_{-5+r}^{5-r} \sqrt{(-5-\xi)^2 + (s-\eta)^2} ds \\
 \text{(ii) For } \Gamma_3 : I_3 &= \int_{-5+r}^{5-r} \sqrt{(s-\xi)^2 + (5-\eta)^2} ds \\
 \text{(ii) For } \Gamma_4 : I_4 &= \int_{-5+r}^{5-r} \sqrt{(5-\xi)^2 + (s-\eta)^2} ds
 \end{aligned} \tag{4.91}$$



**Figure 4.7** Subdomain regions used in calculating the average distance to a neighboring subdomain. (a) General dimensions and location of excluded regions (gray). (b) Region 1 is highlighted in yellow.



**Figure 4.8** Subdomain regions used in calculating the average distance to a neighboring subdomain, where gray regions are the exclusion zone. (a) Region 2 is highlighted in yellow. (b) Region 3 is highlighted in yellow.

For simplicity we did not consider Region 3 (Fig. 4.8) since evaluation of the distance integrals is more complicated and the particle spends little time in this region. Since the boundary edge on each side is of length  $l = 10 - 2r$ , the total length of the subdomain's boundary is  $4(10 - 2r)$ . Hence, we can represent the average distance that the particle travels from  $(\xi, \eta)$  to any neighboring subdomain as

$$I^*(\xi, \eta) = \frac{(I_1 + I_2 + I_3 + I_4)}{4(10 - 2r)}. \quad (4.92)$$



To compute an average distance across the entire subdomain, we need to decompose the subdomain into three different regions, as illustrated in Figures 4.7 and 4.8. Regions are drawn based on the excluded area, as the bounds on  $\xi$  and  $\eta$  depend on where obstacles are located. We will use line integrals to determine arbitrary distances to the boundary edge given a particle's location in both Region 1 and Region 2.

For Region 1, the bounds of both  $\xi$  and  $\eta$  are from  $-5 + r$  to  $5 - r$ , and the total area of that region is  $(10 - 2r)^2$ . We can then describe the average distance from an arbitrary point in Region 1 as

$$J_1 = \frac{1}{(10 - 2r)^2} \int_{-5+r}^{5-r} \int_{-5+r}^{5-r} I^*(\xi, \eta) d\eta d\xi. \quad (4.93)$$

For Region 2, we pick one of the 4 subregions to evaluate the distance, as the distance to boundaries will be the same for all regions due to symmetry of the subdomain. We choose the subregion located at the bottom of the subdomain, with bounds  $\eta \in [-5 + r, 5 - r]$  and  $\xi \in [-5, -5 + r]$ . We can describe the average distance from an arbitrary point in Region 2 as:

$$J_2 = \frac{1}{r(10 - 2r)} \int_{-5+r}^{5-r} \int_{-5}^{-5+r} I^*(\xi, \eta) d\eta d\xi. \quad (4.94)$$

We then estimate the average distance from a particle to any neighboring subdomain, over the whole subdomain, using the weighted average:

$$d_{\text{approx}} = m_1 J_1 + m_2 J_2, \quad (4.95)$$

where  $m_1$  and  $m_2$  are weights determined by the ratios of the respective areas in Regions 1 and 2. The total area we are estimating over is  $a = (10 - 2r)^2 + 4(10 - 2r)r$ , so

$$m_1 = \frac{(10 - 2r)^2}{(10 - 2r)^2 + 4(10 - 2r)r}, \text{ and } m_2 = \frac{4r(10 - 2r)}{(10 - 2r)^2 + 4(10 - 2r)r}. \quad (4.96)$$

Using Maple, we find values for  $d_{\text{approx}}$  based on the different obstacle radius  $R_{\text{obs}}$  values. Table 4.2 shows the values of the average distance a particle would have to travel to reach another subdomain. We note that the range of values for  $d_{\text{approx}}$  in Table 4.2 is between 5.58 and 6.51. Based off the subdomains in the continuous model, we would expect the minimum for distances between subdomains to be  $\frac{L}{2} = 5$ , since that is the distance from the center of one subdomain to the boundary of a neighboring subdomain, and we would expect the maximum distance between subdomains to be  $\frac{L}{2}\sqrt{2} \approx 7.07$ . We note that the computed values for  $d_{\text{approx}}$  in Table 4.2 are all between 5 and 7.07.

We will come back to these calculations in Chapter 5 when evaluating results based on

**Table 4.2** Resulting average distance  $d_{\text{approx}}$  to a neighboring subdomain based on different values of the obstacle radius  $R_{\text{obs}}$ . The value for  $R_{\text{obs}}$  refers to the no obstacle case.

Obstacle radius $R_{\text{obs}}$	0	0.5	1.0	1.5	2.0	2.5	3.0
$r = R_{\text{obs}} + R_{\text{part}}$	0	1.5	2.0	2.5	3.0	3.5	4.0
$d_{\text{approx}}$	6.5176	6.1108	5.9658	5.8257	5.6952	5.5792	5.4822

simulations using both the continuous model and the surrogate (2D) lattice model.

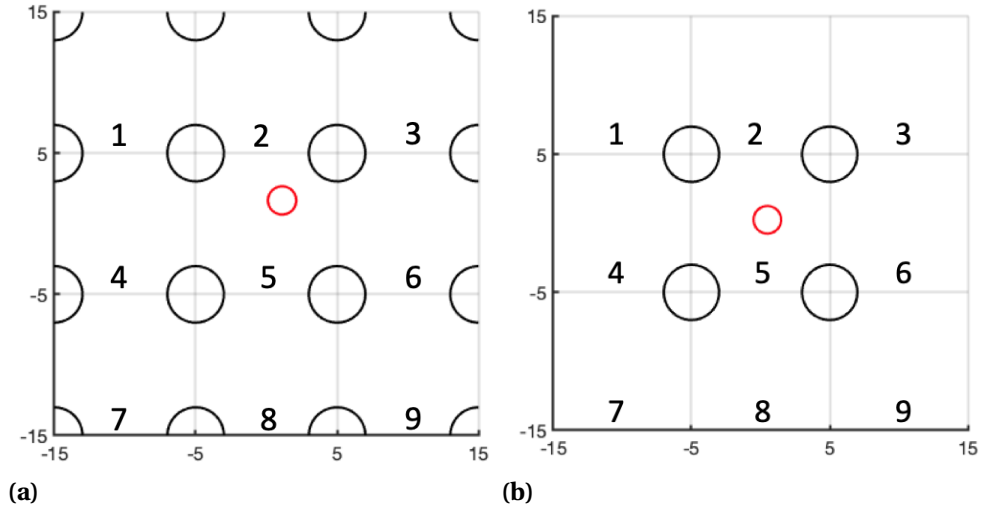
## CHAPTER

# 5

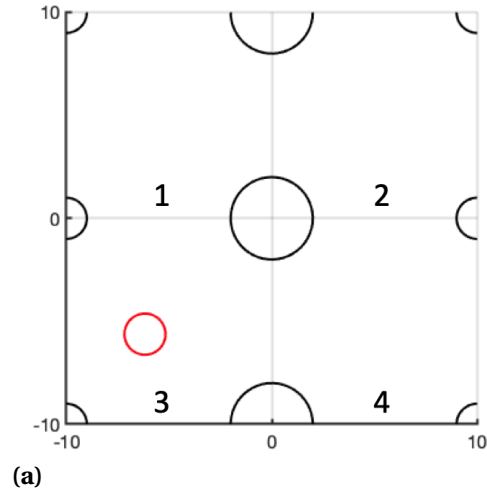
## RESULTS FOR SINGLE PARTICLE RANDOM WALKS IN A PERIODIC RVE

We now consider simulations based on the continuous and surrogate lattice models, as described in chapters 3 and 4. Continuous model domains are illustrated again in Figures 5.1 and 5.2. We compare the results for the continuous model simulations for each obstacle arrangement and then compare with simulation results from their resulting surrogate lattice model simulations. Throughout our simulations, a sufficient number of timesteps and realizations were used to reach stationary values ( $10^7$  timesteps and 2000 realizations were typical). At first we do not consider the commitment index  $M_c$  (section 4.1), since identifying optimal values of  $M_c$  is inherently tied to surrogate model simulations. Thus, in the initial discussion of the continuous model and its transition probabilities,  $M_c$  is effectively zero (i.e. we count a transition as having occurred at any time that the particle crosses the boundary between subdomains). Thus, our initial transition probabilities for a particle staying in its current subdomain are *lower* than they will be in the best set of lattice models, due to the particle crossing the boundary between subdomains repeatedly before more fully committing to a new subdomain. In a later section, the optimal values for  $M_c$  are determined through evaluations of the lattice model. Preliminary transition probabilities are created from the continuous simulation results for varied values of  $M_c$ .

For each obstacle configuration and radius value, optimal values of  $M_c$  is determined according to which value gives  $\langle r^2 \rangle$  versus time results that are in closest agreement with the continuous model. We then compare the resulting surrogate model with its corresponding continuous model.



**Figure 5.1** Domains for continuous model with numbered subdomains for particle radius  $R_{\text{part}} = 1.0$  and obstacle radius  $R_{\text{obs}} = 2.0$ . (a) Equally spaced obstacles model. (b) Four internal obstacles model.



**Figure 5.2** Domain for continuous model of multisize obstacles with numbered subdomains for particle radius  $R_{\text{part}} = 1.0$ , and obstacle radii  $R_{\text{obs}_1} = 1.0$  and  $R_{\text{obs}_2} = 2.0$ .

## 5.1 Continuous Model Results

We now consider results for the three obstacle configurations: equally spaced obstacles (sec. 5.1.1), four internal obstacles (sec.5.1.2), and multisize obstacles (sec. 5.1.3), as depicted in Figure 5.1 and 5.2.

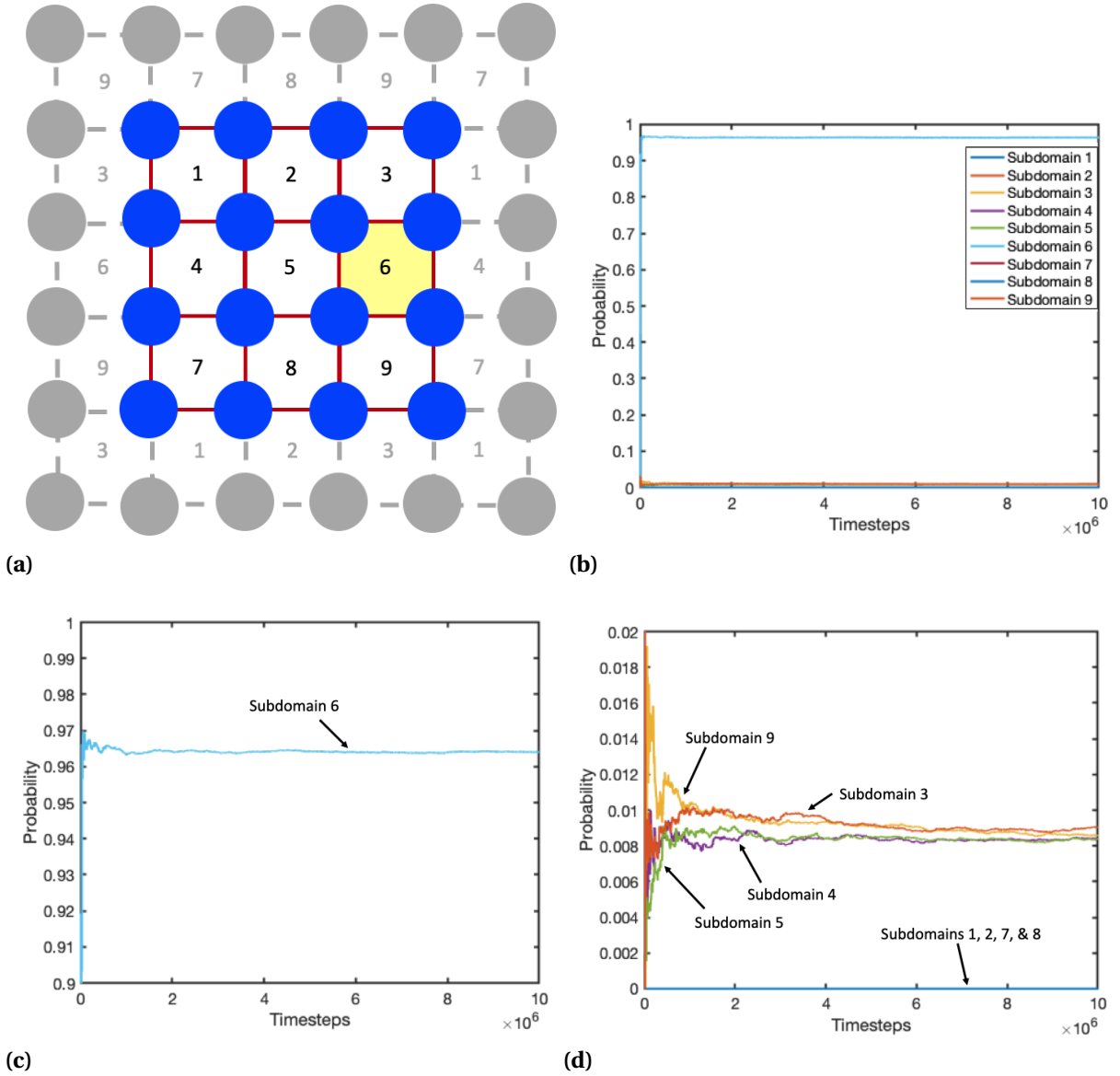
### 5.1.1 Equally spaced obstacles

For the continuous model simulations, we first consider the model with equally spaced obstacles (Fig. 5.1). The following two cases were considered: (1) fixing the particle radius to  $R_{\text{part}} = 1.0$  and varying the obstacle radius  $R_{\text{obs}}$  from 1.0 to 3.0, and (2) fixing the obstacle radius to  $R_{\text{obs}} = 2.0$  and varying the particle radius  $R_{\text{part}}$  from 1.0 to 2.5.

#### 5.1.1.1 Fixed particle radius and varying obstacle radius

Before considering the mean squared displacement, we first verified that there were a sufficient number of time steps to ensure that the transition probabilities in the continuous model reached stationary values. These transition probabilities were later averaged over many realizations to build the surrogate model discussed later on in this chapter.

Figure 5.3 illustrates transition probabilities for transitions from subdomain 6 over the course of one realization. Transition probabilities were calculated after pooling all transitions across realizations, but results in this figure are useful for illustrating that  $10^7$  timesteps is sufficient for probabilities to approach stationary values. Figure 5.3(a) shows the domain for the equally spaced obstacles model, with subdomain 6 highlighted in yellow. Figure 5.3(b) is scaled such that all transition probabilities are plotted, however, since the probability of staying in the current subdomain (subdomain 6) is much greater than for other types of transitions, only the probability of staying the current subdomain is visible. Figure 5.3(c) is scaled to better illustrate the probability that the particle stays in subdomain 6. From this plot we see that the transition probability varied a lot over early timesteps but that it approaches a stationary value at later timesteps (the actual recorded value at the end was 0.96398). Figure 5.3(d) is scaled to illustrate cardinal direction transitions, over a finer and smaller range (e.g. small fluctuations in the probability appear amplified). From this figure we observe that the four cardinal direction transitions also vary a lot initially but eventually approach stationary values. These values are all between 0.008 and 0.009, and they were averaged together across all realizations in creating the actual transition probability for the surrogate model. We also observe that cardinal direction transitions are much less likely than staying in the current subdomain, so these transition



**Figure 5.3** Equally spaced obstacles model. Transitions from subdomain 6 are shown to illustrate convergence of transition probabilities. (a) Simulation domain with subdomain 6 highlighted in yellow. Red boundary lines indicate subdomains in the original domain, and grey boundary lines indicated subdomains in the extended domain. Blue circles represent obstacles. (b) Plot of all transition probabilities. At this scale, only the probability of staying in the current subdomain is visible. (c) Plot of transition probability for staying in subdomain 6. (d) Plot of transition probabilities for transitions from subdomain 6 to subdomains in cardinal directions (i.e. transitions to subdomains 3, 4, 5, 9).

probabilities take a greater number of timesteps to reach stationary values. The remaining transition probabilities remain at zero throughout the timecourse (appearing as a solid light blue line in the graph). These results also verify the consistency of our continuous model, i.e. obstacles are inhibiting particle movement and the particle does not make diagonal transitions.

An example of transition probabilities after one realization is shown in Table 5.1. In this example, the particle radius  $R_{\text{part}} = 1.0$ , and the obstacle radius  $R_{\text{obs}} = 1.0$ . The rows indicate the subdomain a particle came from, and the columns indicate the subdomain a particle went to during the transition to a new subdomain. Consider three distinct groupings of probabilities by type:

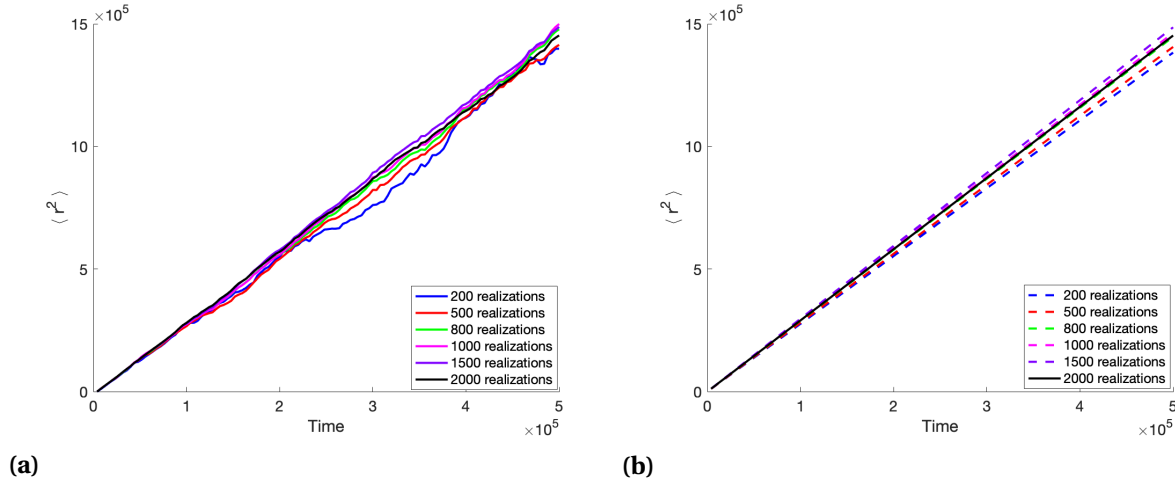
- (i) *Staying in the current subdomain.* Along the main diagonal are the transition probabilities for a particle staying in its current subdomain. These probabilities of staying are approximately 0.96 or 0.97 (close to one), indicating that a particle is much more likely to stay in a given subdomain than to travel to another subdomain, within a given time step.
- (ii) *Moving to a neighboring subdomain in a cardinal direction.* Transition probabilities for cardinal direction transitions are the next most likely at approximately 0.008 or 0.009. For example, considering transition from subdomain 6, the corresponding cardinal-direction subdomains are subdomains 3, 4, 5, and 9.
- (iii) Transitions that should not occur (*diagonal transitions*). Diagonal transitions should not occur due to obstacles blocking such transitions. Therefore, these corresponding transition probabilities should be zero. For example, considering transitions from subdomain 6, the corresponding diagonal transitions are subdomains 1, 2, 7, and 8, and these values were all recorded as zero.

**Table 5.1** Equally spaced obstacles model. Transition probabilities by subdomain for particle radius  $R_{\text{part}} = 1.0$  and obstacle radius  $R_{\text{obs}} = 1.0$  for one realization ( $M_c = 0$ ).

		Subdomain after transition								
		1	2	3	4	5	6	7	8	9
Subdomain before transition	1	<b>0.96544</b>	0.00884	0.00863	0.00870	0.00000	0.00000	0.00874	0.00000	0.00000
	2	0.00841	<b>0.96490</b>	0.00804	0.00000	0.00872	0.00000	0.00000	0.00852	0.00000
	3	0.00869	0.00855	<b>0.96554</b>	0.00000	0.00000	0.00902	0.00000	0.00000	0.00944
	4	0.00912	0.00000	0.00000	<b>0.96560</b>	0.00956	0.00916	0.00918	0.00000	0.00000
	5	0.00000	0.00899	0.00000	0.00891	<b>0.96453</b>	0.00855	0.00000	0.00884	0.00000
	6	0.00000	0.00000	0.00858	0.00842	0.00838	<b>0.96398</b>	0.00000	0.00000	0.00905
	7	0.00834	0.00000	0.00000	0.00837	0.00000	0.00000	<b>0.96425</b>	0.00872	0.00862
	8	0.00000	0.00872	0.00000	0.00000	0.00881	0.00000	0.00894	<b>0.96474</b>	0.00914
	9	0.00000	0.00000	0.00920	0.00000	0.00000	0.00929	0.00889	0.00918	<b>0.96375</b>



After establishing stationary values for the transition probabilities when using  $10^7$  timesteps, we performed many realizations of the continuous simulation. We determined a sufficient number of realizations for the slope of  $\langle r^2 \rangle$  versus time to reach stationary values. Due to MATLAB memory constraints, the continuous model was run on a high performance cluster (HPC) in batches of 50 realizations. We ran enough realizations to yield stationary values in quantities of interest, such as the slope of the linear relationship between mean squared displacement,  $\langle r^2 \rangle$  and time (Our simulations used a timestep  $\Delta t = 0.05$ , so time = timestep  $\times 0.05$ ). These results were then pooled and averaged across all subdomains by transition type. Figure 5.4 shows the effects of increasing the number of realizations on the plot of mean squared displacement versus the time (with  $R_{\text{part}} = 1.0$  and  $R_{\text{obs}} = 1.0$ ). Based on these results, we determined that 2000 realizations was sufficient to: (1) consider the relationship between mean squared displacement and time to be linear, and (2) for the slope to have reached a stationary value.

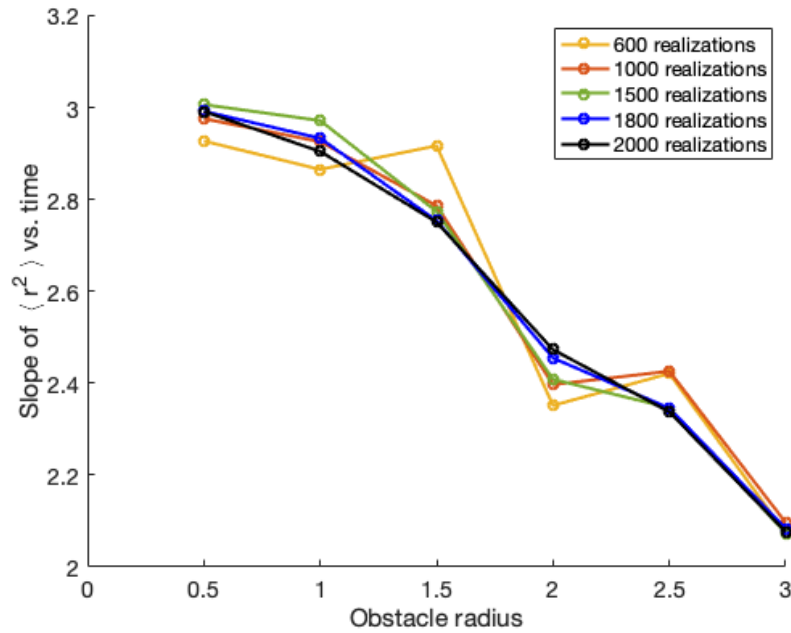


**Figure 5.4** Equally spaced obstacles model. Effect of increasing number of realizations on mean squared displacement for  $R_{\text{part}} = 1.0$  and  $R_{\text{obs}} = 1.0$ . (a) Mean squared displacement averages for increasing number of realizations, (b) Linear fit of mean squared displacement averages for increasing number of realizations.

In the same manner, we determined that 2000 realizations was sufficient to reach stationary values for all obstacle radius values. We analyze the slope of  $\langle r^2 \rangle$  versus time for (1) an increasing number of realizations and (2) an increasing obstacle radius size. Figure 5.5 and Table 5.2 show the resulting slope values. We observe that, as the number of realizations approaches 2000, the slope value approaches a stationary value. We also observe that the

slope strictly decreases as the obstacle radius increases, which verifies that larger obstacles can inhibit outward drift of the particle.

Figure 5.5 illustrates how the slope of mean squared displacement changes as the obstacle radius is varied. The slope appears to reach a stationary value for each obstacle radius by 2000 realizations, at which point the slope is monotonically decreasing as the obstacle radius increases. This decrease is expected, as more of the domain is obstructed by obstacles, hindering the particle movement, and the particle is more likely to be occasionally “trapped” between a few obstacles.



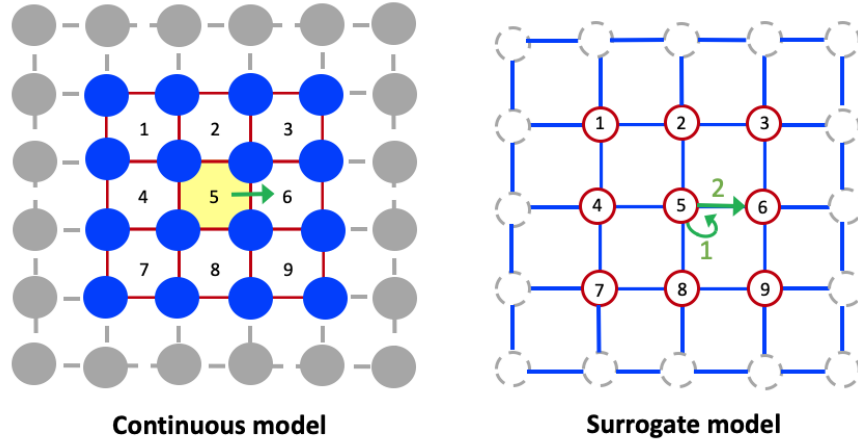
**Figure 5.5** Equally spaced obstacles model. Slope of  $\langle r^2 \rangle$  vs. time for increasing values of obstacle radius  $R_{\text{obs}}$ , where  $R_{\text{part}} = 1.0$ .

After determining a sufficient number of realizations, we returned to using our results to determine the transition probabilities needed to build the lattice model. For the equally spaced obstacles model, we have two distinct probabilities: (1) the probability that a particle remains in its current subdomain, and (2) the probability that a particle moves to another subdomain in a cardinal direction. Aggregating by transition type was done after all realizations were completed. Figure 5.6 illustrates the different transition types. Our resulting probabilities are summarized in Table 5.3. As the obstacle radius increases, the probability of staying in the current subdomain increases, due to a greater obstacle size

**Table 5.2** Equally spaced obstacles model. Slope of mean squared displacement versus time for increasing number of realizations when particle radius  $R_{\text{part}} = 1.0$  and for varied obstacle radius  $R_{\text{obs}}$ .

		Number of realizations				
		600	1000	1500	1800	2000
Obstacle radius	0.5	2.9272	2.9758	3.0065	2.9920	2.9915
	1.0	2.8653	2.9256	2.9717	2.9331	2.9044
	1.5	2.9169	2.7864	2.7737	2.7537	2.7507
	2.0	2.3515	2.3970	2.4085	2.4549	2.4731
	2.5	2.4210	2.4261	2.3465	2.3455	2.3376
	3.0	2.0733	2.0964	2.0735	2.0836	2.0765

being more likely to inhibit a particle's movement. In section 5.2.1 we will revisit transition probabilities with the appropriate commitment index  $M_c$  when discussing the surrogate model results.



**Figure 5.6** Equally spaced obstacles model. Diagram showing transition types based on subdomain type. Distinct transitions are indicated with green arrows, and transition types are numbered in the surrogate model representation. In the continuous model, red lines indicate boundaries between subdomains in the original RVE, dashed lines indicate subdomain boundaries in the extended domain, solid blue circles indicate obstacles in the original RVE, and solid gray circles indicate obstacles in the extended domain. In the surrogate model, red circles correspond to subdomains in the original RVE, gray dashed circles correspond to subdomains in the extended domain, and blue lines indicate possible transitions.

The rows in Table 5.3 correspond with the transition probability vector  $\vec{p}^E$  (section

**Table 5.3** Equally spaced obstacles model. Transition probabilities by type for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  for after 2000 realizations (for  $M_c = 0$ ).

Obstacle radius	Probability by transition type	
	1	2
0.5	0.96149	$9.6233 \times 10^{-3}$
1.0	0.96503	$8.7426 \times 10^{-3}$
1.5	0.96834	$7.9142 \times 10^{-3}$
2.0	0.97166	$7.0840 \times 10^{-3}$
2.5	0.97527	$6.1814 \times 10^{-3}$
3.0	0.97971	$5.0726 \times 10^{-3}$

3.3.1.1) . For example, for obstacle radius  $R_{\text{obs}} = 0.5$ ,

$$\vec{p}^E = [0.96149, 0.0096233]. \quad (5.1)$$

Note that for the equally spaced obstacles case, there are four possible subdomains corresponding with a cardinal direction transition, and each has the transition probability  $\vec{p}^E(2)$  (calculated to be  $\vec{p}^E(2) \approx 0.0096$  when  $R_{\text{obs}} = 1.0$ ), so, when confirming that all possible transition probabilities add to one, we verify

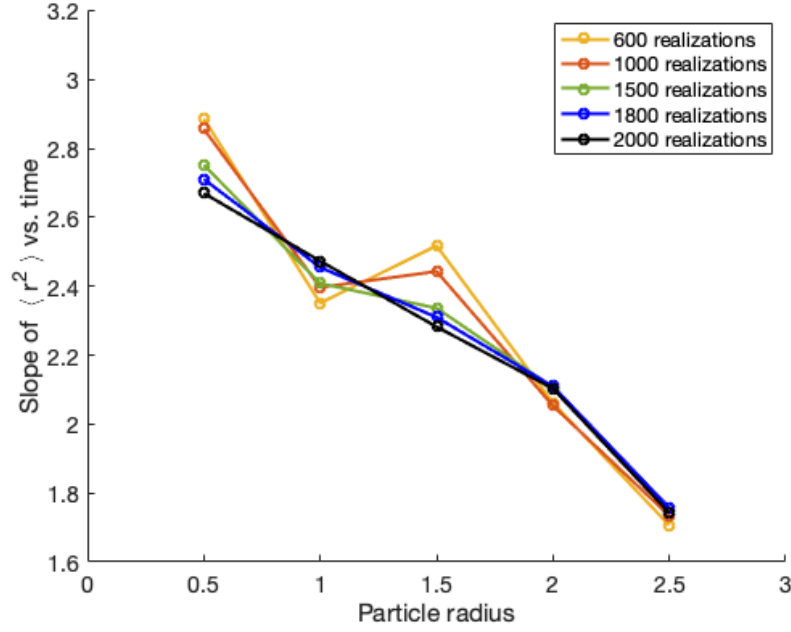
$$p^E(1) + 4p^E(2) = 1. \quad (5.2)$$

Continuing with the  $R_{\text{obs}} = 1.0$  case, we can account for the slight difference from 1 due to rounding error, and these sums are within 0.02% of 1.00 for all cases. Note that in the actual surrogate model, resulting transition probabilities are taken directly from the continuous model results in MATLAB, so less rounding error is involved.

### 5.1.1.2 Fixed obstacle radius and varying particle radius

After determining the relationship between obstacle radius and slope of the resulting mean squared displacement, we investigated how changing the particle radius affects this result. We fixed the obstacle radius at  $R_{\text{obs}} = 2.0$  and varied the particle radius  $R_{\text{part}}$  from 0.5 to 2.5. Increasing the particle radius beyond  $R_{\text{part}} = 2.5$  caused the gap between obstacles to be too small for the particle to travel in between the obstacles, and the particle is trapped in its initial subdomain. Figure 5.7 shows the relationship between increasing particle radius and the slope of the mean squared displacement versus time. These slope values are

summarized in Table 5.4



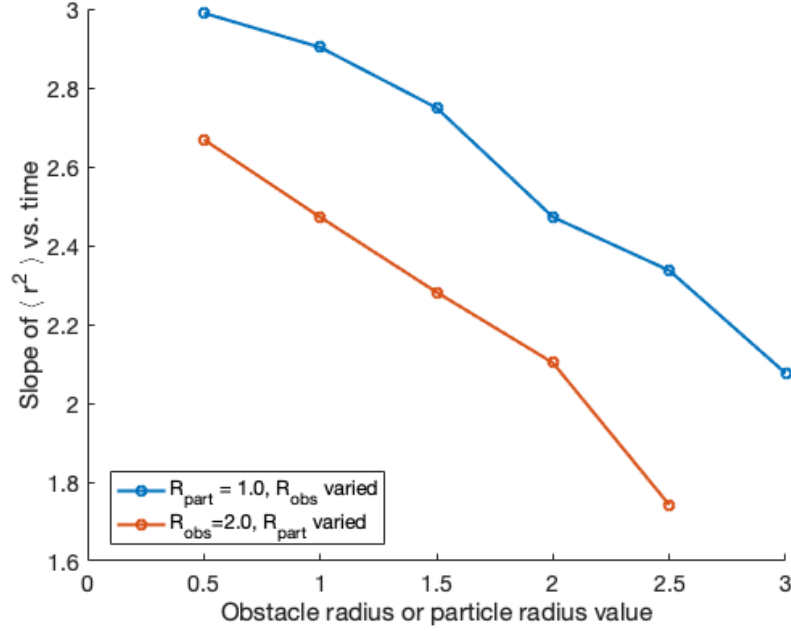
**Figure 5.7** Equally spaced obstacles model. Slope of  $\langle r^2 \rangle$  vs. time for increasing values of particle radius  $R_{\text{part}}$ , where  $R_{\text{obs}} = 2.0$ .

**Table 5.4** Equally spaced obstacles model. Slope of mean squared displacement versus time for increasing number of realizations when obstacle radius  $R_{\text{obs}} = 2.0$  and for varied particle  $R_{\text{part}}$ .

		Number of realizations				
		600	1000	1500	1800	2000
Particle radius	0.5	2.8865	2.8568	2.7523	2.7111	2.6703
	1.0	2.3515	2.3970	2.4085	2.4549	2.4731
	1.5	2.5177	2.4432	2.3366	2.3101	2.2825
	2.0	2.0615	2.0533	2.1070	2.1095	2.1031
	2.5	1.7061	1.7337	1.7544	1.7569	1.7420

In Figure 5.8 we compare the results from varying the obstacle radius (with  $R_{\text{part}} = 1.0$ ) and varying the obstacle radius (with  $R_{\text{obs}} = 2.0$ ). We observe that the model has similar sensitivity to varied particle radius as to varied obstacle radius. For the sake of brevity, in

the remaining continuous models (four internal obstacles, multisize obstacles), we will focus on varying the obstacle radius size for a fixed particle radius. We will also focus on varying obstacle radius value for the surrogate models.



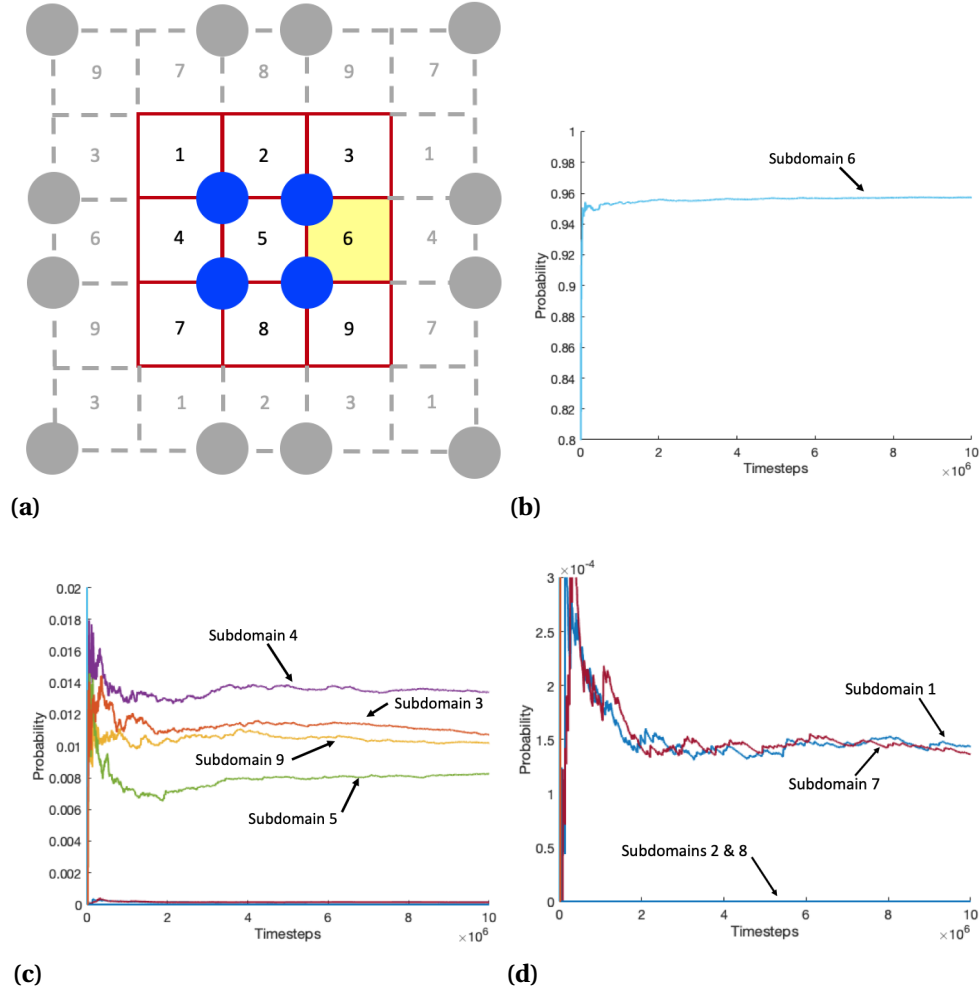
**Figure 5.8** Equally spaced obstacles model. Comparing slope of  $\langle r^2 \rangle$  vs. time for varied particle radius value and varied obstacle radius values.

## 5.1.2 Model for four internal obstacles

### 5.1.2.1 Fixed particle radius and varying obstacle radius

We next considered the model with four internal obstacles (Fig. 5.9(a)). We only considered the case of fixed particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  between 1.0 to 3.0. We again first verified that a sufficient number of time steps were used so that the transition probabilities in the continuous model reached stationary values. These transition probabilities were later averaged over all realizations to build the surrogate model discussed later in on this chapter.

To explain the convergence of transition probabilities, consider transitions from subdomain 6 (a one-wall subdomain), as illustrated in Figures 5.9(a) - (d). Figure 5.9(a) illustrates the continuous domain, with subdomain 6 highlighted. Figure 5.9(b) is best for visualizing



**Figure 5.9** Four internal obstacles model. Transitions from subdomain 6 are shown to illustrate convergence of transition probabilities. (a) Simulation domain with subdomain 6 highlighted in yellow. Red boundary lines indicate subdomains in the original domain, and grey boundary lines indicated subdomains in the extended domain. Blue circles represent obstacles. (b) Plot of probability for staying in subdomain 6. (c) Plot of transition probabilities for cardinal direction transitions (i.e. transitions to subdomains 3, 4, 5, and 9). (d) Plot of transition probabilities for transitions from subdomain 6 to subdomains in diagonal directions (i.e. transitions to subdomains 1, 2, 7, and 8, where 2 and 8 are blocked by obstacles and remain at zero).

the probability of staying in the current subdomain. This figure is summarized below:

- (i) The probability of staying in subdomain 6 appears to reach a stationary value by  $1 \times 10^7$  timesteps.
- (ii) At this scale on the  $y$ -axis, the eight remaining transition probabilities appear near zero. Hence, we need to consider a finer scale.

The remaining eight transition probabilities are much smaller in magnitude, and are visualized in Figure 5.9(c), which is summarized below:

- (i) The transition from subdomain 6 to subdomain 4 corresponds with a transition in the horizontal direction, and is not hindered by any obstacles; thus, this is the next greatest magnitude of transition probability.
- (ii) The transitions from subdomain 6 to subdomains 3 and 9 correspond to vertical transitions, with only one obstacle along the boundary between that subdomain and subdomain 6. Thus, these are the next 2 greatest probabilities. As mentioned in Chapter 3, transition probabilities of the same type are averaged; e.g. for transitions from subdomain 6 to a vertical-direction subdomain (e.g. subdomains 3 and 9).
- (iii) Finally, the transition from subdomain 6 to subdomain 5 is a transition in the horizontal direction that is hindered by two obstacles, making this transition have the next greatest probability.

In Figure 5.9(c) the magnitude on the  $y$ -axis is small, and these transition probabilities are at or near stationary values.

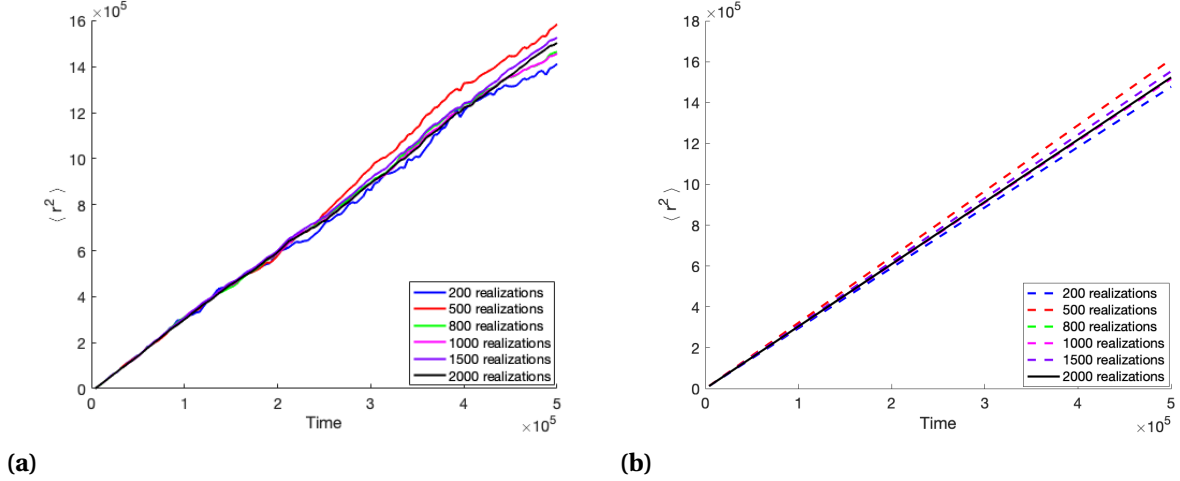
Finally, Figure 5.9(d) illustrates the remaining transition probabilities and is summarized below:

- (i) In this instance, transitions to subdomains 1 and 7 correspond with diagonal transitions. These transition probabilities are very small due to a particle having a much higher likelihood of a transition across the wider subdomain boundary for cardinal directions than at the corner.
- (ii) The transition probabilities from subdomain 6 to subdomains 8 and 2 are both zero, which is expected as these transitions correspond to transitions in the diagonal direction that are blocked by obstacles. This provides a good consistency check that the obstacle collision simulations are correctly implemented.



In Figure 5.9(d), the magnitude on the y-axis is now very small, and these transition probabilities are also at or near stationary values.

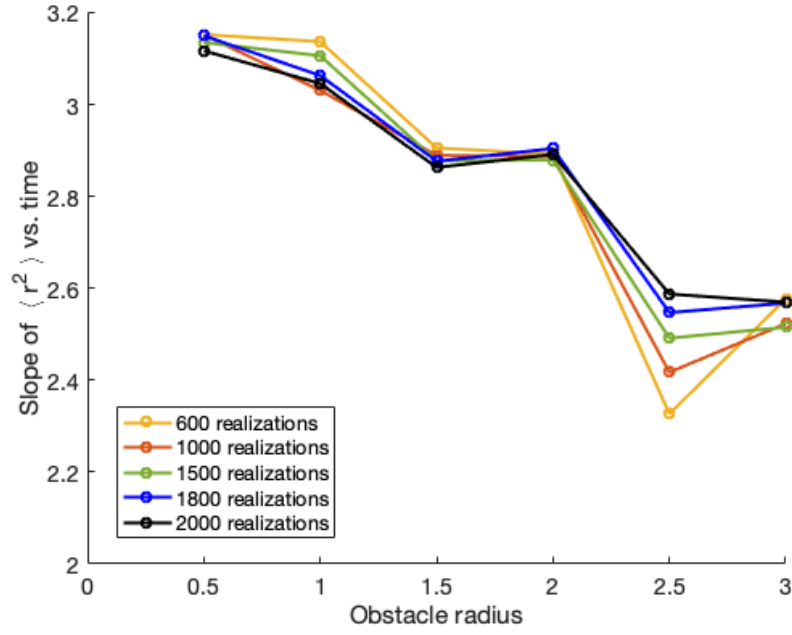
Based on these results, we determined that  $10^7$  timesteps was sufficient for transition probabilities to reach stationary values. Similar results regarding magnitude and convergence of transition probabilities by transition type occurred for the other subdomains.



**Figure 5.10** Four obstacle model for  $R_{\text{part}} = 1.0$  and  $R_{\text{obs}} = 1.0$ . Effect of increasing number of realizations on mean squared displacement. (a) Mean squared displacement averages for increasing number of realizations, (b) Linear fit of mean squared displacement averages for increasing number of realizations.

As in the equally spaced obstacles case, we ran the continuous model on a high performance cluster (HPC) in batches of 50 realizations. We ran enough realizations to yield stationary values in quantities of interest. These results were then pooled and averaged by subdomain type. Figure 5.10 shows how increasing the number of realizations affects the plot of mean squared displacement versus time, for  $R_{\text{part}} = 1.0$  and  $R_{\text{obs}} = 1.0$ . Based on these results, we determine that 2000 realizations was sufficient: (1) to consider the relationship between mean squared displacement and time to be linear and (2) for the slope to have reached a stationary value.

In the same manner, we determined that 2000 realizations was sufficient for stationarity for all obstacle radius values. We analyzed the slope of  $\langle r^2 \rangle$  versus time for (1) an increasing number of realizations and (2) an increasing obstacle radius size. Figure 5.11 and Table 5.5 show the resulting slope values. We observe that, as the number of realizations approaches 2000, the slope value approaches a stationary value. We also observe that the slope mostly

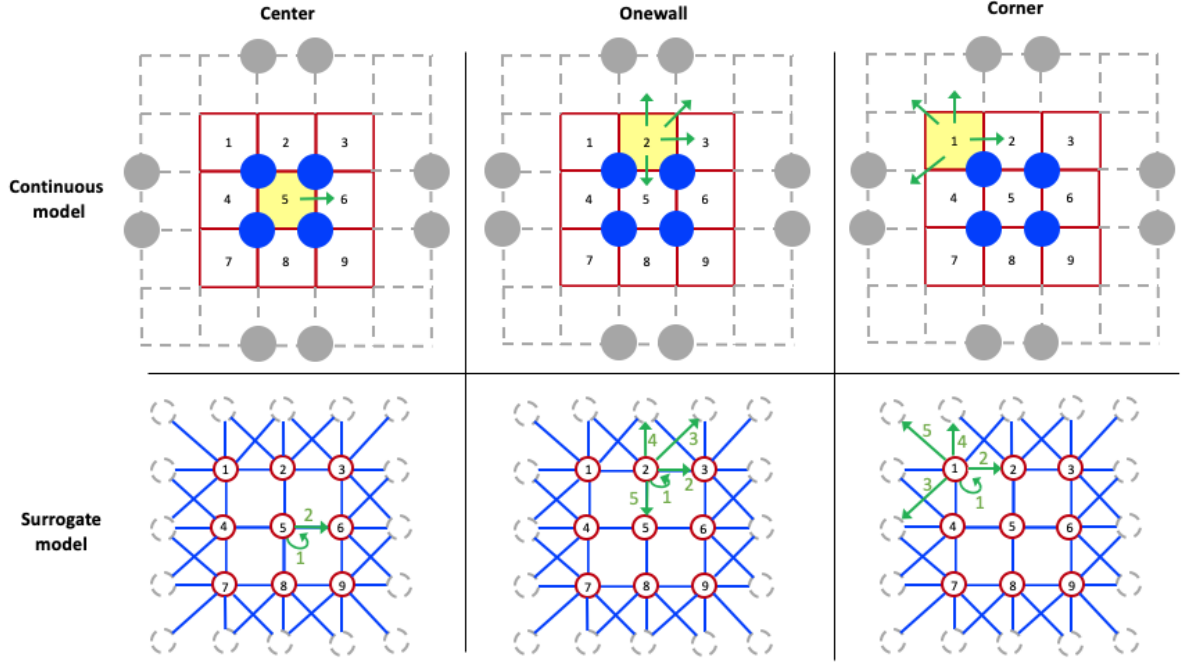


**Figure 5.11** Four internal obstacles model. Slope of  $\langle r^2 \rangle$  vs. time for increasing values of obstacle radius  $R_{\text{obs}}$ , where  $R_{\text{part}} = 1.0$ .

**Table 5.5** Four internal obstacles model. Slope of mean squared displacement versus time for increasing number of realizations when particle radius  $R_{\text{part}} = 1.0$  and for varied obstacle radius  $R_{\text{obs}}$ .

		Number of realizations				
		600	1000	1500	1800	2000
Obstacle radius	0.5	3.1524	3.1519	3.1350	3.1498	3.1170
	1.0	3.1371	3.0308	3.1063	3.0633	3.0459
	1.5	2.9055	2.8895	2.8775	2.8767	2.8631
	2.0	2.8922	2.8870	2.8793	2.9043	2.8914
	2.5	2.3267	2.4173	2.4915	2.5472	2.5876
	3.0	2.5768	2.5233	2.5154	2.5678	2.5700

decreases as the obstacle radius increases, meaning the particle does not travel as far when larger obstacles are involved. The value of the slope does increase between  $R_{\text{obs}} = 1.5$  and  $R_{\text{obs}} = 2.0$  from 0.1432 to 0.1446.



**Figure 5.12** Four internal obstacles model. Diagram showing transition types based on subdomain type. Distinct transitions are indicated with green arrows, and transition types are numbered in the surrogate model representation. In the continuous model, red lines indicate boundaries between subdomains in the original RVE, dashed lines indicate subdomain boundaries in the extended domain, solid blue circles indicate obstacles in the original RVE, and solid gray circles indicate obstacles in the extended domain. In the surrogate model, red circles correspond to subdomains in the original RVE, gray dashed circles correspond to subdomains in the extended domain, and blue lines indicate possible transitions.

After determining a sufficient number of realizations, we returned to using our results to build the transition probabilities to be used in the surrogate lattice model. For the four internal obstacles model, we have twelve distinct probabilities (as detailed in section 3.3.2.1), with two probabilities for transitions from center the subdomain, five probabilities for transitions from onewall subdomains, and five probabilities for transitions from corner subdomains. Aggregating by transition type was carried out after all realizations were completed. Transition types are labeled by subdomain type in Figure 5.12. Our resulting probabilities (when  $M_c = 0$ ) are summarized in Tables 5.6 - 5.8. The rows in these tables

**Table 5.6** Four internal obstacles model. Transition probabilities for center subdomain (C) by transition type for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  for after 2000 realizations (for  $M_c = 0$ ).

Obstacle radius	Probability by transition type	
	1	2
0.5	0.96160	$9.5959 \times 10^{-3}$
1.0	0.96492	$8.7682 \times 10^{-3}$
1.5	0.96826	$7.9345 \times 10^{-3}$
2.0	0.97167	$7.0811 \times 10^{-3}$
2.5	0.97520	$6.1998 \times 10^{-3}$
3.0	0.97971	$5.0736 \times 10^{-3}$

**Table 5.7** Four internal obstacles model. Transition probabilities for onewall subdomain (I) by transition type for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  for after 2000 realizations (for  $M_c = 0$ ).

Obstacle radius	Probability by transition type				
	1	2	3	4	5
0.5	0.95538	$1.1107 \times 10^{-2}$	$1.3266 \times 10^{-4}$	$1.2929 \times 10^{-2}$	$9.2035 \times 10^{-3}$
1.0	0.95671	$1.0741 \times 10^{-2}$	$1.3599 \times 10^{-4}$	$1.3365 \times 10^{-2}$	$8.1691 \times 10^{-3}$
1.5	0.95790	$1.0446 \times 10^{-2}$	$1.4010 \times 10^{-4}$	$1.3862 \times 10^{-2}$	$7.0601 \times 10^{-3}$
2.0	0.95877	$1.0249 \times 10^{-2}$	$1.4911 \times 10^{-4}$	$1.4554 \times 10^{-2}$	$5.8756 \times 10^{-3}$
2.5	0.95930	$1.0093 \times 10^{-2}$	$1.5779 \times 10^{-4}$	$1.5536 \times 10^{-2}$	$4.6658 \times 10^{-3}$
3.0	0.95941	$1.0060 \times 10^{-2}$	$1.7029 \times 10^{-4}$	$1.6809 \times 10^{-2}$	$3.3223 \times 10^{-3}$

**Table 5.8** Four internal obstacles model. Transition probabilities for corner subdomain (L) by transition type for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  for after 2000 realizations (for  $M_c = 0$ ).

Obstacle radius	Probability by transition type				
	1	2	3	4	5
0.5	0.95236	$1.0912 \times 10^{-2}$	$1.3042 \times 10^{-4}$	$1.2712 \times 10^{-2}$	$1.2923 \times 10^{-4}$
1.0	0.95314	$1.0390 \times 10^{-2}$	$1.3129 \times 10^{-4}$	$1.2844 \times 10^{-2}$	$1.3243 \times 10^{-4}$
1.5	0.95352	$9.8880 \times 10^{-3}$	$1.3399 \times 10^{-4}$	$1.3149 \times 10^{-2}$	$1.3379 \times 10^{-4}$
2.0	0.95382	$9.4388 \times 10^{-3}$	$1.3689 \times 10^{-4}$	$1.3448 \times 10^{-2}$	$1.3630 \times 10^{-4}$
2.5	0.95400	$8.9760 \times 10^{-3}$	$1.4082 \times 10^{-4}$	$1.3811 \times 10^{-2}$	$1.4138 \times 10^{-4}$
3.0	0.95379	$8.5664 \times 10^{-3}$	$1.4594 \times 10^{-4}$	$1.4322 \times 10^{-2}$	$1.4506 \times 10^{-4}$

correspond to values of the transition probability vector  $\vec{p}_i^F$  (section 3.3.2.1 ), for  $i = C, I, L$  (center, onewall, corner), with Table 5.6 corresponding to  $\vec{p}_C^F$  , Table 5.7 corresponding to  $\vec{p}_I^F$  , Table 5.8 corresponding to  $\vec{p}_L^F$  . For example, the transition probability vector for transitions from the center subdomain, when  $R_{\text{obs}} = 1.0$  is

$$\vec{p}_C^F = [0.9659, 0.0088]. \quad (5.3)$$

Note that for the four internal obstacles case, the number of possible transitions depends on the subdomain type, with some transition types corresponding to movements to multiple other subdomains. This means that when confirming that all possible transition probabilities add to one, we verify

$$\vec{p}_C^F(1) + 4\vec{p}_C^F(2) = 1, \quad (5.4)$$

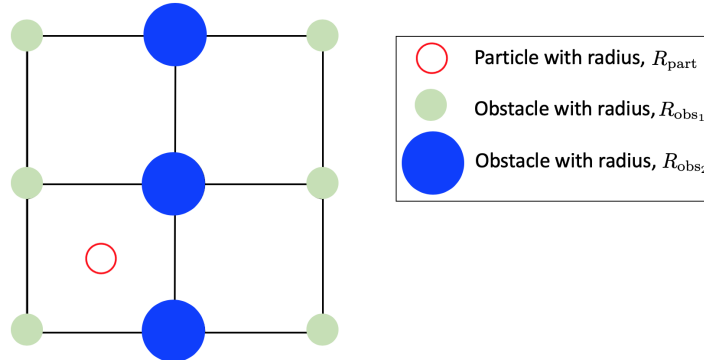
$$\vec{p}_I^F(1) + 2\vec{p}_I^F(2) + 2\vec{p}_I^F(3) + \vec{p}_I^F(4) + \vec{p}_I^F(5) = 1, \quad (5.5)$$

$$\vec{p}_L^F(1) + 2\vec{p}_L^F(2) + 2\vec{p}_L^F(3) + 2\vec{p}_L^F(4) + \vec{p}_L^F(5) = 1, \quad (5.6)$$

for the center ( $C$ ), onewall ( $I$ ), and corner ( $L$ ) subdomains, respectively.

From these tables, we observe that the particle is most likely to remain in its current subdomain for all subdomain types and obstacle radius values, with probabilities ranging from 0.95 to 0.98. Our next most likely type of transition involves transitions in a cardinal direction, with transition probabilities near 0.01. Finally, the rarest type of transition are diagonal transitions, which have transition probabilities near 0.0001 or 0.0002.

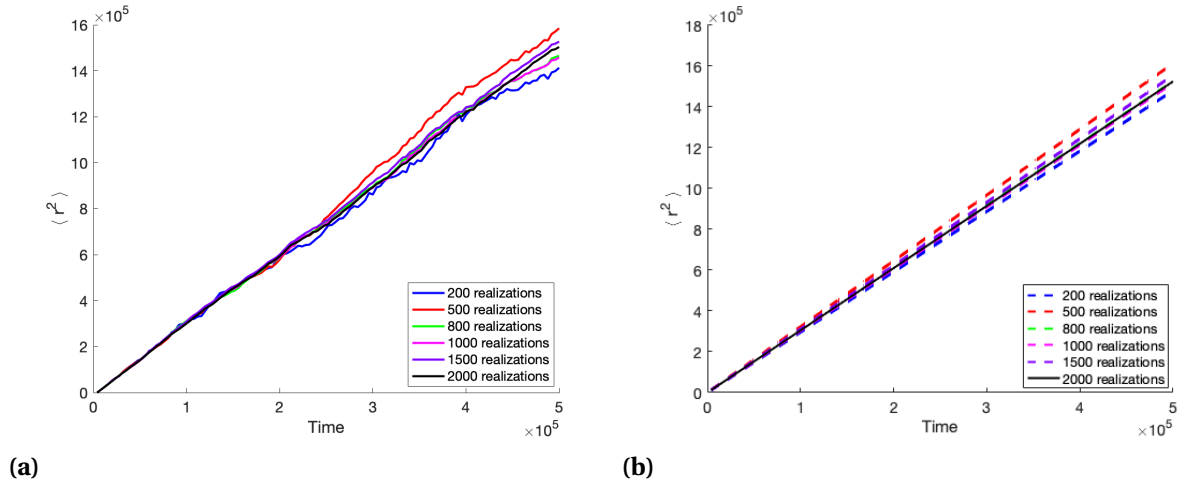
### 5.1.3 Multisize obstacles



**Figure 5.13** Multisize obstacles model. Diagram showing the continuous model RVE.

As described in Section 3.2.3, for the multisize obstacle model, we use a  $2 \times 2$  RVE instead of the  $3 \times 3$  RVE used in the four internal obstacles and equally spaced obstacles models. Figure 5.13 illustrates the domain for the multisize obstacles model. We fixed the particle radius  $R_{\text{part}} = 1.0$  and fixed the first "column" of obstacles radii to  $R_{\text{obs}_1} = 1.0$ . For the remaining column of obstacles, the obstacle radius  $R_{\text{obs}_2}$  was varied from 0.5 to 3.0. Since the results were more sensitive to varying the obstacle radius in the four internal obstacles and the equally spaced obstacles models, we fixed the particle radius  $R_{\text{part}} = 1.0$  and only varied the obstacle radius  $R_{\text{obs}_2}$  for the multisize obstacles case.

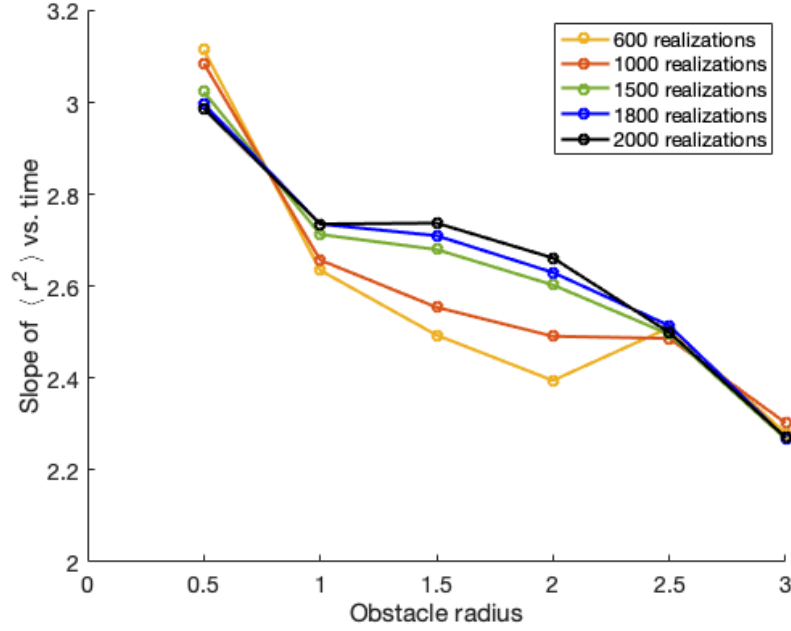
We verified the number of timesteps needed to reach stationarity for the multisize obstacles model, and we concluded that  $10^7$  timesteps was again sufficient. The adequate number of realizations for the multisize case was also determined in the same manner as in the other cases, and found to be 2,000. For brevity, we do not show an example of transition probabilities reaching stationary values over one realization. We again observed that the probability of staying was significantly larger than the other transition probabilities. Figure 5.14 shows the effect of increasing number of realizations on the plot of mean squared displacement versus time.



**Figure 5.14** Multisize obstacle model with  $R_{\text{part}} = 1.0$ ,  $R_{\text{obs}_1} = 1.0$ , and  $R_{\text{obs}_2} = 1.0$ . Effect of increasing number of realizations on mean squared displacement. (a) Mean squared displacement averages for increasing number of realizations, (b) Linear fit of mean squared displacement averages for increasing number of realizations.

Figure 5.15 and Table 5.9 compare how the slope of mean squared displacement versus time changes as the second column of obstacle radii ( $R_{\text{obs}_2}$ ) increases, and as the number

of realizations increases. The slope appears to reach a stationary value for each obstacle radius by 2000 realizations. Considering just the 2000 realizations case, the slope is mostly decreasing, except for at  $R_{\text{obs}_2} = 1.5$ , where there is a very slight increase.

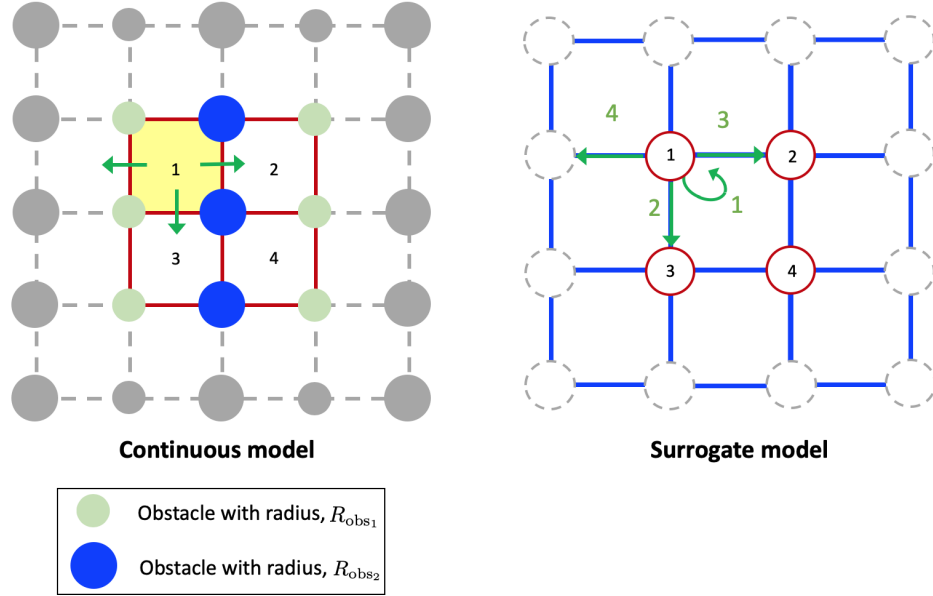


**Figure 5.15** Multisize obstacles model. Slope of  $\langle r^2 \rangle$  versus time for increasing values of obstacle radius  $R_{\text{obs}_2}$ , where  $R_{\text{part}} = 1.0$ ,  $R_{\text{obs}_1} = 1.0$ .

**Table 5.9** Multisize obstacles model. Slope of mean squared displacement versus time for increasing number of realizations when particle radius  $R_{\text{part}} = 1.0$ ,  $R_{\text{obs}_1} = 1.0$ , and for varied obstacle radius  $R_{\text{obs}_2}$ .

		Number of realizations				
		600	1000	1500	1800	2000
Obstacle radius	0.5	0.1557	0.1543	0.1512	0.1498	0.1494
	1.0	0.1317	0.1329	0.1356	0.1368	0.1368
	1.5	0.1247	0.1277	0.1340	0.1355	0.1369
	2.0	0.1198	0.1246	0.1301	0.1315	0.1331
	2.5	0.1255	0.1243	0.1248	0.1257	0.1250
	3.0	0.1140	0.1151	0.1134	0.1135	0.1136

After determining a sufficient number of realizations, we returned to using our results to determine the transition probabilities needed to build the surrogate lattice model. For the multisize obstacles model, we have four distinct probabilities (as detailed in 3.3.3.1). Aggregating by transition type was done after all realizations were completed. Transition types are shown in Figure 5.16. Our resulting probabilities (when the commitment index  $M_c = 0$ ) are summarized in Table 5.10.



**Figure 5.16** Multisize obstacles model. Diagram showing transition types based on subdomain type. Distinct transitions are indicated with green arrows, and transition types are numbered in the surrogate model representation. For continuous models, red lines indicate boundaries between subdomains in the original RVE, dashed lines indicate subdomain boundaries in the extended domain, solid blue circles indicate obstacles in the original RVE, solid gray circles indicate obstacles in the extended domain. For surrogate models, red circles correspond to subdomains in the original RVE, gray dashed circles correspond to subdomains in the extended domain, and blue lines indicate possible transitions.

For the multisize obstacles case, transition type 2 has two possible subdomains to which the particle could transition. In this case, we need to verify that

$$\vec{p}^M(1) + 2\vec{p}^M(2) + \vec{p}^M(3) + \vec{p}^M(4) = 1. \quad (5.7)$$

We observe that the particle is most likely to remain in its current subdomain for all subdomain types and obstacle radius values, with probabilities ranging from 0.74 to 0.92. We start seeing a difference in transition probabilities by transition type once  $R_{\text{obs}_2} > 1.0$ ,



**Table 5.10** Multisize obstacles model. Transition probabilities for center subdomain by transition type for particle radius  $R_{\text{part}} = 1.0$ ,  $R_{\text{obs}_1} = 1.0$  and varying obstacle radius  $R_{\text{obs}_2}$  after 2000 realizations (for  $M_c = 0$ ).

Obstacle radius $R_{\text{obs}_2}$	Probability by transition type			
	1	2	3	4
0.5	0.96324	$9.1974 \times 10^{-3}$	$9.9067 \times 10^{-3}$	$8.4528 \times 10^{-3}$
1.0	0.96503	$8.7331 \times 10^{-3}$	$8.7534 \times 10^{-3}$	$8.7496 \times 10^{-3}$
1.5	0.96661	$8.3340 \times 10^{-3}$	$7.5820 \times 10^{-3}$	$9.1406 \times 10^{-3}$
2.0	0.96803	$7.9905 \times 10^{-3}$	$6.3746 \times 10^{-3}$	$9.6181 \times 10^{-3}$
2.5	0.96926	$7.6896 \times 10^{-3}$	$5.0753 \times 10^{-3}$	$1.0288 \times 10^{-3}$
3.0	0.97029	$7.4213 \times 10^{-3}$	$3.6445 \times 10^{-3}$	$1.1226 \times 10^{-3}$

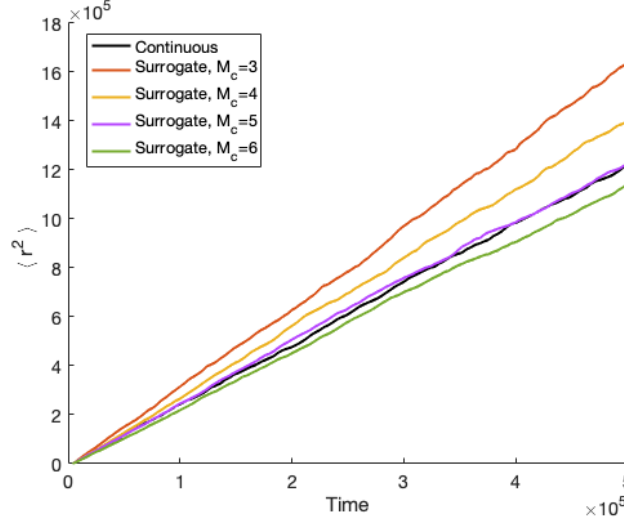
and then transitions of type 3 have a greater magnitude.

## 5.2 Lattice Model Results

The surrogate lattice model is built using transition probabilities determined based on simulations using the continuous model. We initially ran the surrogate model using transition probabilities obtained when the commitment index  $M_c = 0$  (introduced in sec. 4.1), but this lead to disagreement between the continuous and surrogate models. Consequently, we then tested increasing values for  $M_c$  until the surrogate model was in better agreement with the continuous models. Once the appropriate value for  $M_c$  was selected, we ran 2000 realizations for the surrogate models and then compared our results to those obtained using the continuous models. Since our continuous results showed that our models were more sensitive to varying obstacle radius than to varying particle radius, we only considered the surrogate models for the case of varying obstacle radius.

### 5.2.1 Model for equally spaced obstacles

Surrogate model simulations began by identifying the optimal values of  $M_c$  for varying obstacle radius, with the particle radius fixed at  $R_{\text{part}} = 1.0$ . For each obstacle radius  $R_{\text{obs}}$  we ran the surrogate model multiple times using probabilities obtained using the continuous model as the commitment index  $M_c$  was varied from 3 to 8. We determined the value for  $M_c$  by comparing the continuous and surrogate plots of mean squared displacement  $\langle r^2 \rangle$  versus time. Figure 5.17 shows several such results for the case where  $R_{\text{part}} = 1.0$  and  $R_{\text{obs}} = 2.0$ . In this instance, we observe that  $M_c = 5$  yields the best results for the surrogate



**Figure 5.17** Equally spaced obstacles.  $\langle r^2 \rangle$  versus time for continuous model and resulting surrogate model when commitment index  $M_c$  is varied. Results shown for particle radius  $R_{\text{part}} = 1.0$  and obstacle radius  $R_{\text{obs}} = 1.0$ .

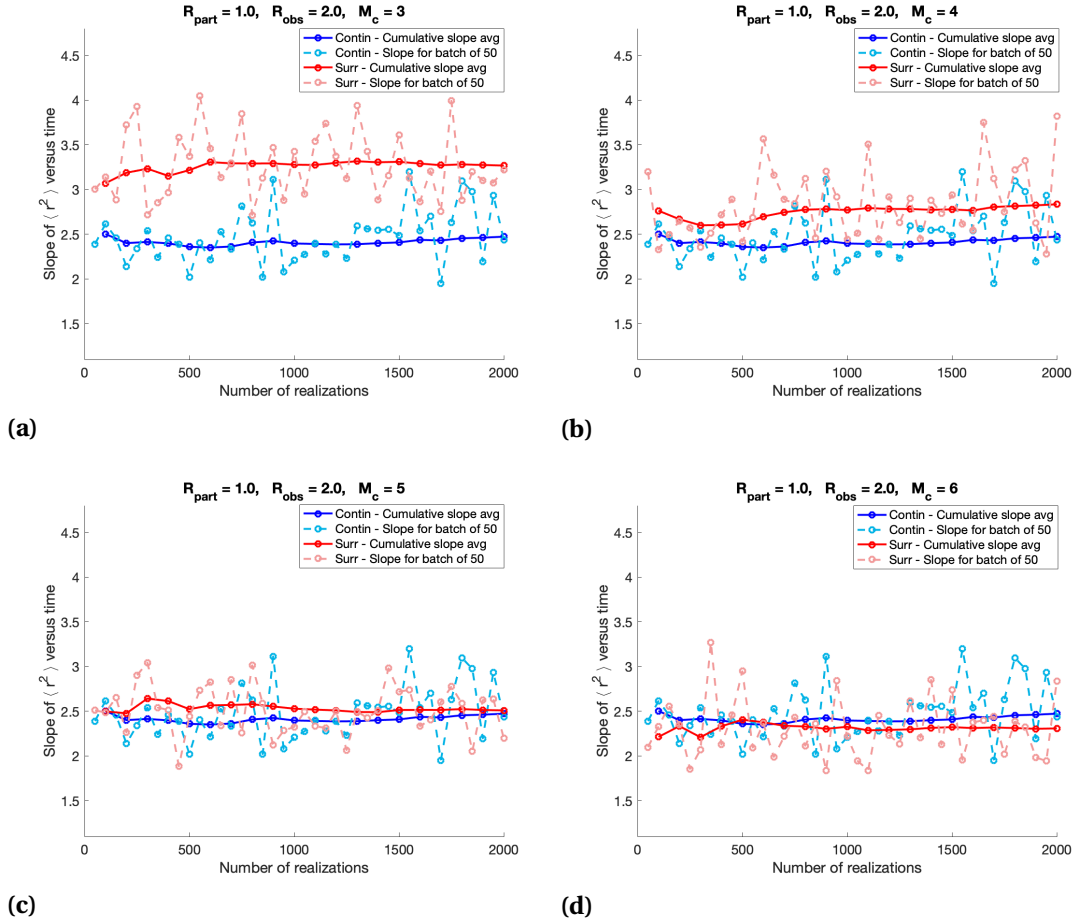
model.

We consider the resulting slope of  $\langle r^2 \rangle$  versus time to be a quantity of interest. Figure 5.18 shows the resulting average slope over different batches of 50 realizations, and a cumulative average slope for up to 2000 realizations. These figures demonstrate that these slopes reach stationary values with an increasing number of realizations, and that  $M_c = 5$  is the best choice for the commitment index in the case of  $R_{\text{part}} = 1.0$  and  $R_{\text{obs}} = 2.0$ .

Figure 5.19 visualizes these results for the best values of  $M_c$  for all cases in which the obstacle radius was varied. Note that the best value for  $M_c$  ranges from  $M_c = 4$  ( $R_{\text{obs}} = 3$ ) up to  $M_c = 7$  ( $R_{\text{obs}} = 0.5$ ). Based on these results, it appears that our slope values for both the continuous and surrogate cases have reached stationary values. We also observe very good agreement between the continuous and surrogate models.

In Figure 5.20 we show the resulting slope of  $\langle r^2 \rangle$  versus time for increasing obstacle radius values, showing results for the continuous model and the surrogate model with the best  $M_c$  value. We summarize these slope values and their percent difference in Table 5.11. The slope values for the continuous and surrogate model have a correlation coefficient of  $R = 0.9863$ .

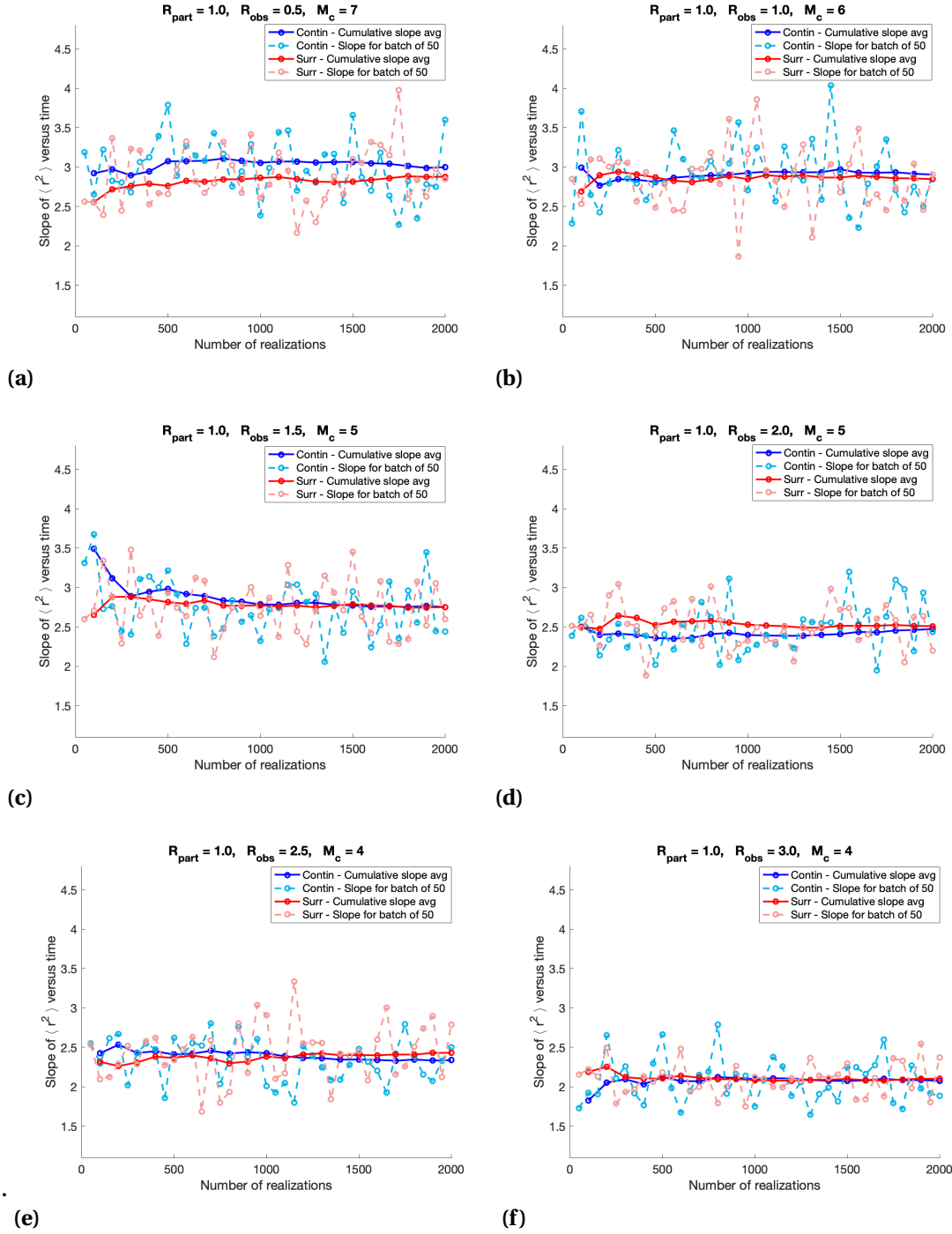
After finding the appropriate  $M_c$  values, we are able to update the transition probabilities accordingly. Our resulting  $M_c$  values and probabilities are summarized in Table 5.12. As the obstacle radius increases, the probability of staying in the current subdomain increases, due to a greater obstacle size being more likely to inhibit a particle's outward movement.



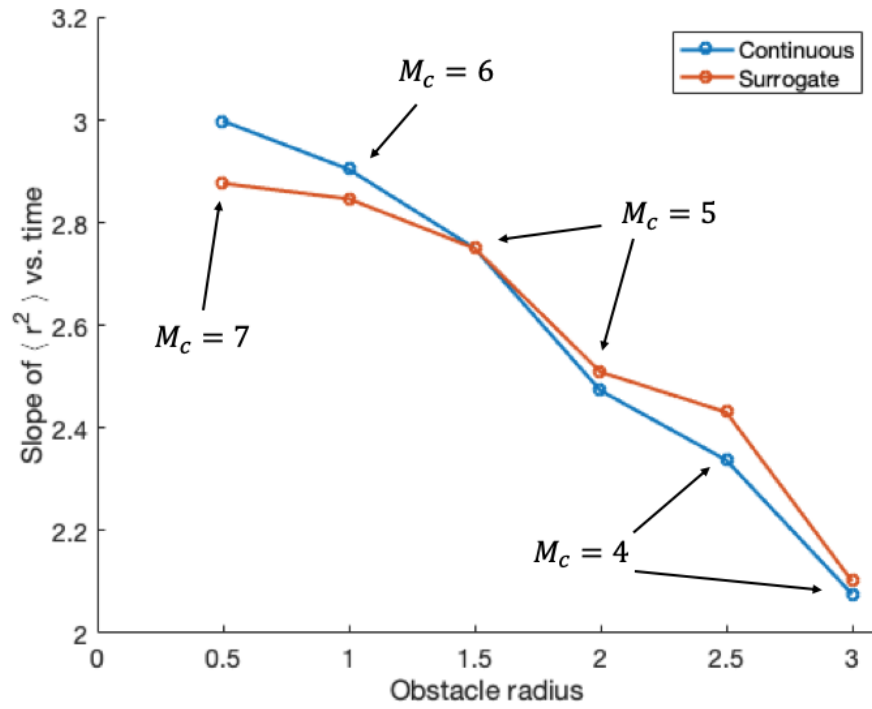
**Figure 5.18** Equally spaced obstacles. Slope of  $\langle r^2 \rangle$  versus time for an increasing number of realizations, along with a cumulative average of the slope value. Shown for particle radius 1.0 and obstacle radius 2.0. (a)  $M_c = 3$ , (b)  $M_c = 4$ , (c)  $M_c = 5$ , (d)  $M_c = 6$

**Table 5.11** Equally spaced obstacles. Slope values for  $\langle r^2 \rangle$  versus time for continuous model and surrogate model with the best  $M_c$  values.

$R_{\text{obs}}$	$M_c$	Continuous model slope	Surrogate model slope	Percent difference
0.5	7	2.9984	2.8770	4.049%
1.0	6	2.9044	2.8468	1.983%
1.5	5	2.7507	2.7506	0.004%
2.0	5	2.4731	2.5094	1.468%
2.5	4	2.3376	2.4310	3.996%
3.0	4	2.0765	2.1016	1.209%



**Figure 5.19** Equally spaced obstacles. Slope of  $\langle r^2 \rangle$  vs. time for increasing number of realizations, along with a cumulative average of the slope value. For  $R_{\text{part}} = 1.0$ . (a)  $R_{\text{obs}} = 0.5, M_c = 7$ , (b)  $R_{\text{obs}} = 1.0, M_c = 6$ , (c)  $R_{\text{obs}} = 1.5, M_c = 5$ , (d)  $R_{\text{obs}} = 2.0, M_c = 5$ , (e)  $R_{\text{obs}} = 2.5, M_c = 4$ , and (f)  $R_{\text{obs}} = 3.0, M_c = 4$ .



**Figure 5.20** Equally spaced obstacles. Slope of  $\langle r^2 \rangle$  versus time for continuous model and resulting surrogate model with best  $M_c$  values. Results for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$ .

The rows in Table 5.12 correspond with the transition probability vector  $\vec{p}^E$  (section 3.3.1.1).

**Table 5.12** Equally spaced obstacles model. Transition probabilities by type for optimal  $M_c$  for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  after 2000 realizations.

Obstacle radius	$M_c$	Probability by transition type	
		1	2
0.5	7	0.99276	$1.8099 \times 10^{-3}$
1.0	6	0.99291	$1.7724 \times 10^{-3}$
1.5	5	0.99297	$1.7573 \times 10^{-3}$
2.0	5	0.99371	$1.5734 \times 10^{-3}$
2.5	4	0.99384	$1.5401 \times 10^{-3}$
3.0	4	0.99493	$1.2676 \times 10^{-3}$

### 5.2.1.1 Diffusivity for equally spaced obstacles

#### 5.2.1.1.1 Simulated diffusivities

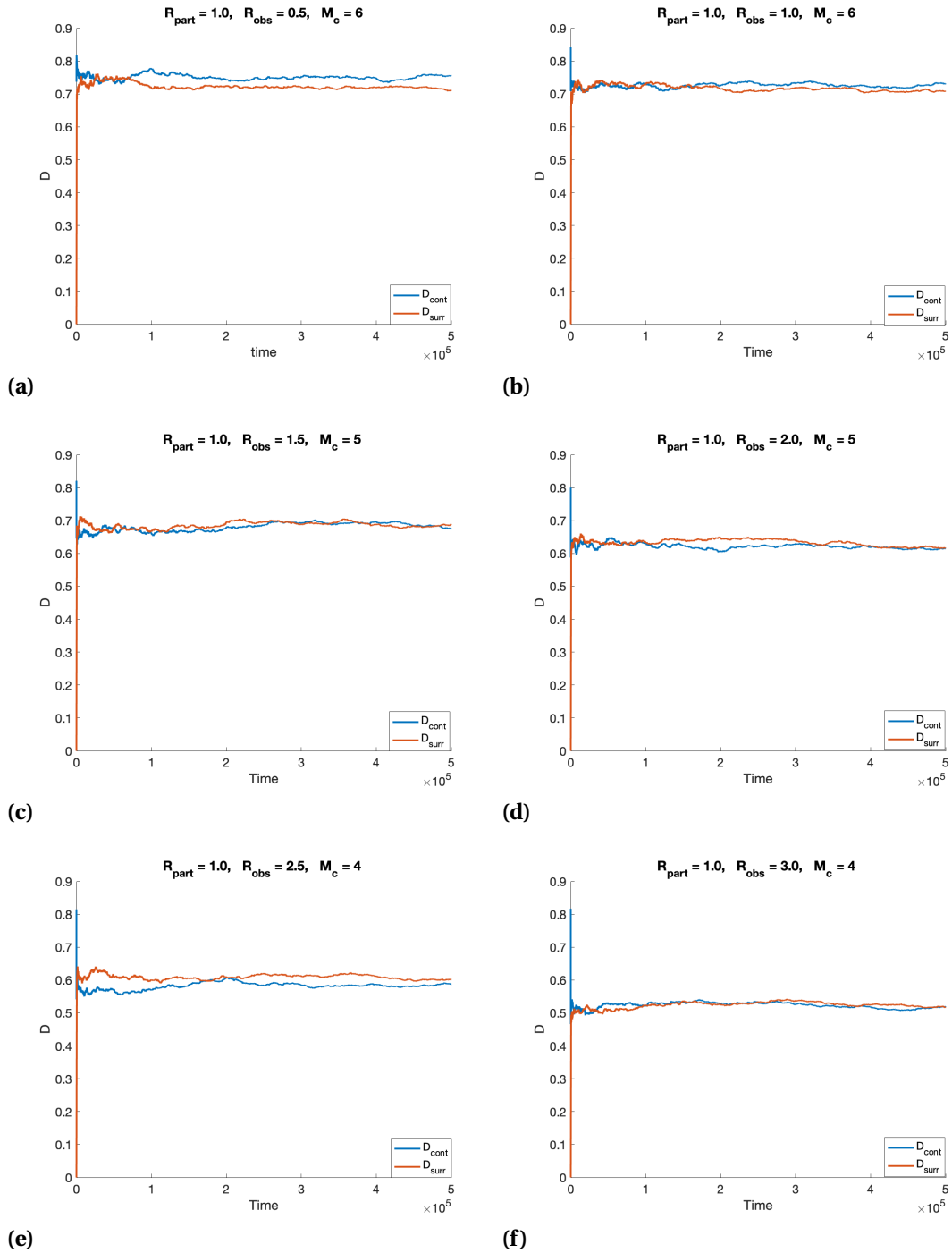
Having obtained results for  $\langle r^2 \rangle$  for both the continuous and surrogate models, we can estimate the diffusivity  $D$ , as discussed in section 4.2.1. We first estimate the simulated diffusivity via the equation

$$D = \frac{\langle r^2 \rangle}{4t} = \frac{\langle r^2 \rangle}{4N(\Delta t)}, \quad (5.8)$$

where  $N$  is the number of timesteps. Equation 5.8 allows us to estimate diffusion constants from simulated results for the continuous and surrogate models ( $D_{\text{cont}}$  and  $D_{\text{surr}}$ , respectively), where

$$\begin{aligned} D_{\text{cont}} &= \frac{\langle r_{\text{cont}}^2 \rangle}{4N(\Delta t)}, \\ D_{\text{surr}} &= \frac{\langle r_{\text{surr}}^2 \rangle}{4N(\Delta t)}. \end{aligned} \quad (5.9)$$

Figure 5.21 shows the resulting estimated diffusivity versus time for each obstacle radius value. For a final comparison we average the last 100,000 values (last 10%) of  $D_{\text{cont}}$  and  $D_{\text{surr}}$ , as summarized in Table 5.13. The continuous and surrogate diffusivities have a correlation coefficient of  $R = 0.9784$ .



**Figure 5.21** Equally spaced obstacles model. Estimated diffusivity for continuous and surrogate models for  $R_{\text{part}} = 1.0$ . (a)  $R_{\text{obs}} = 0.5$ ,  $M_c = 7$ , (b)  $R_{\text{obs}} = 1.0$ ,  $M_c = 6$ , (c)  $R_{\text{obs}} = 1.5$ ,  $M_c = 5$ , (d)  $R_{\text{obs}} = 2.0$ ,  $M_c = 5$ , (e)  $R_{\text{obs}} = 2.5$ ,  $M_c = 4$ , (f)  $R_{\text{obs}} = 3.0$ ,  $M_c = 4$ .

**Table 5.13** Equally spaced obstacles model. Table showing commitment index  $M_c$ , probability of staying in current subdomain  $q$ , and diffusivities obtained via simulation for the continuous and surrogate models ( $D_{\text{cont}}$  and  $D_{\text{surr}}$ , respectively) for each obstacle radius value.

Obstacle radius	$M_c$	$q$	$D_{\text{cont}}$	$D_{\text{surr}}$	Percent difference
0.5	7	0.99276	0.7550	0.7100	5.960%
1.0	6	0.99291	0.7314	0.7084	3.145%
1.5	5	0.99297	0.6759	0.6862	1.524%
2.0	5	0.99371	0.6146	0.6165	0.309%
2.5	4	0.99384	0.5885	0.6016	2.226%
3.0	4	0.99493	0.5188	0.5187	0.019%

#### 5.2.1.1.2 Using the ratio of $D_{\text{surr}}$ to $D_{\text{TS}}$ to find $\Delta x$

In the previous section we used simulation results to estimate diffusivities for the continuous and surrogate models using the equations:

$$D_{\text{cont}} = \frac{\langle r_{\text{cont}}^2 \rangle}{4N(\Delta t)},$$

$$D_{\text{surr}} = \frac{\langle r_{\text{surr}}^2 \rangle}{4N(\Delta t)}.$$
(5.10)

In section 4.2 we discussed comparing our simulated surrogate estimate for diffusivity  $D_{\text{surr}}$  to a theoretical estimate using a Taylor Series expansion. In section 4.2.3 we found the Taylor Series Expansion estimate for diffusivity of a random walk on a lattice to be:

$$D_{\text{TS}} = \frac{(1-q)(\Delta x)^2}{4(\Delta t)},$$
(5.11)

where  $q$  is the probability of staying in the current subdomain,  $\Delta t$  is the timestep, and  $\Delta x$  is the effective spatial stepsize, which needs to be estimated. In section 4.2.6 we related  $D_{\text{surr}}$  and  $D_{\text{TS}}$  using the relation

$$D_{\text{surr}} \approx D_{\text{TS}}$$

$$\approx \frac{1}{\alpha^2} D_{\text{TS}} | \Delta x = 10,$$
(5.12)



where  $D_{\text{TS}|\Delta x=10}$  is calculated according to equation (5.11) using an initial estimate  $\Delta x = L = 10$ , and  $\alpha$  is a constant such that

$$\Delta x = \frac{L}{\alpha}. \quad (5.13)$$

Thus, determining the ratio

$$\frac{D_{\text{TS}|\Delta x=10}}{D_{\text{surr}}} \approx \alpha^2, \quad (5.14)$$

allows us to estimate the appropriate  $\Delta x$  for  $D_{\text{TS}}$  using equation (5.13).

Table 5.14 shows the resulting estimates for the diffusion constant for the surrogate model  $D_{\text{surr}}$ , the diffusion constant using  $\Delta x = L = 10$  for the Taylor Series result  $D_{\text{TS}|\Delta x=10}$ ,  $\alpha^2$ , and the final resulting  $D_{\text{TS}}$  using equation (5.13). The estimated  $\alpha^2$  has mean 5.0466 and standard deviation 0.0888, and  $\Delta x$  has mean 4.4519 and standard deviation 0.0397. We calculated the correlation coefficient for the obstacle radius value and  $\alpha^2$ , and found  $R = 0.5532$ , indicating that the value of  $\alpha^2$  is not correlated with the obstacle radius value, and suggesting that it could be considered constant across obstacle radius values.

**Table 5.14** Equally spaced obstacles model. Table showing diffusivity and effective  $\Delta x$ .

Obstacle radius	$q$	$\alpha^2$	$\Delta x$	$D_{\text{cont}}$	$D_{\text{surr}}$	$D_{\text{TS}}$	$D_{\text{TS} \Delta x=10}$
0.5	0.99276	5.1048	4.4260	0.7550	0.7100	0.7100	3.6243
1.0	0.99291	5.0027	4.4709	0.7314	0.7084	0.7084	3.5440
1.5	0.99297	5.1224	4.4184	0.6759	0.6862	0.6862	3.5149
2.0	0.99371	5.0907	4.4321	0.6146	0.6165	0.6165	3.1384
2.5	0.99384	5.0728	4.4399	0.5885	0.6016	0.6016	3.0517
3.0	0.99493	4.8863	4.5239	0.5188	0.5187	0.5187	2.5343

### 5.2.1.1.3 Integral distance derivation

In section 4.2.7.2, we discussed using integrals to evaluate a particle's average distance to neighboring subdomains,  $d_{\text{approx}}$ . Table 5.15 shows the integral approximations of distances for each value of  $R_{\text{obs}}$  (as shown previously in section 4.2.7.2).

While values for  $d_{\text{approx}}$  in Table 5.15 are larger than the values for  $\Delta x$  shown in Table 5.14, the values for  $d_{\text{approx}}$  are closer to the values for  $\Delta x$  in Table 5.14 than to  $L = 10$  (the subdomain length in the continuous model). In Table 5.14 the surrogate model results

**Table 5.15** Equally spaced obstacles model. Resulting average distance  $d_{\text{approx}}$  to a neighboring subdomain based on different values of the obstacle radius  $R_{\text{obs}}$ . The value for  $R_{\text{obs}}$  refers to the no obstacle case.

Obstacle radius $R_{\text{obs}}$	0	0.5	1.0	1.5	2.0	2.5	3.0
$r = R_{\text{obs}} + R_{\text{part}}$	-	1.5	2.0	2.5	3.0	3.5	4.0
$d_{\text{approx}}$	6.5176	6.1108	5.9658	5.8257	5.6952	5.5792	5.4822

for diffusivity are based on probabilities generated from the continuous model. In initial simulations,  $M_c = 0$ , and the  $\langle r^2 \rangle$  value in the surrogate model was significantly greater than in the continuous model (e.g. the slope of that curve was greater in magnitude). Increasing the value of  $M_c$  led to a *decreasing* value for the slope, until we found a value for  $M_c$  for which continuous and surrogate model results were in much better agreement.

The need for  $M_c$  was due to the inherent randomness in particles engaging in random walks (as in the continuous model) – for an arbitrarily large number of realizations, there will be *some* realization(s) in which the particle moves back-and-forth repeatedly. This skews the probability of a transition higher, and leads to a surrogate model in which the particle travels much further outward than in the continuous model. The integral approximation  $d_{\text{approx}}$  assumes: (1) the particle is equally likely to transition to any neighboring subdomain, and (2) the particle is equally likely to be located anywhere in a particular subdomain at any given time. However, in the continuous model, moving back-and-forth near subdomain boundaries is more common, and the particle tends to be located near subdomain edges for many timesteps near when a transition occurs. This bias towards boundary re-crossing is more pronounced when the obstacle radius is smaller or nonexistent (e.g.  $R_{\text{obs}} = 0.5$ ), and for these values we also see greater discrepancy with the estimate of  $\Delta x$  from Table 5.14 .

## 5.2.2 Four internal obstacles

As in the equally spaced obstacles case, for the four internal obstacles case we began by identifying the optimal values of  $M_c$  for  $R_{\text{part}} = 1.0$  and varying obstacle radius. For each obstacle radius  $R_{\text{obs}}$  we ran the surrogate model multiple times using probabilities obtained using the continuous model as the commitment index  $M_c$  was varied from 3 to 8. We determined the value for  $M_c$  by comparing the continuous and surrogate plots of mean squared displacement  $\langle r^2 \rangle$  versus time. We consider the resulting slope of  $\langle r^2 \rangle$  versus time to be a quantity of interest.

Figure 5.22 visualizes these resulting slopes versus increasing number of realizations

for the best values was  $M_c = 4$  in all cases. Based on these results, it appears that our slope values for both the continuous and surrogate cases have reached stationary values.

In Figure 5.23 we show the resulting slope of  $\langle r^2 \rangle$  versus time for increasing obstacle radius value, showing results for the continuous model and the surrogate model with the best  $M_c$  value. We summarize these slope results and percent differences in Table 5.16. The correlation coefficient for the continuous slope values and surrogate slope values was  $R = 0.9689$ . Although we had good agreement between the continuous and surrogate models in the four internal obstacles case, the agreement was not as good as in the equally spaced obstacles case.

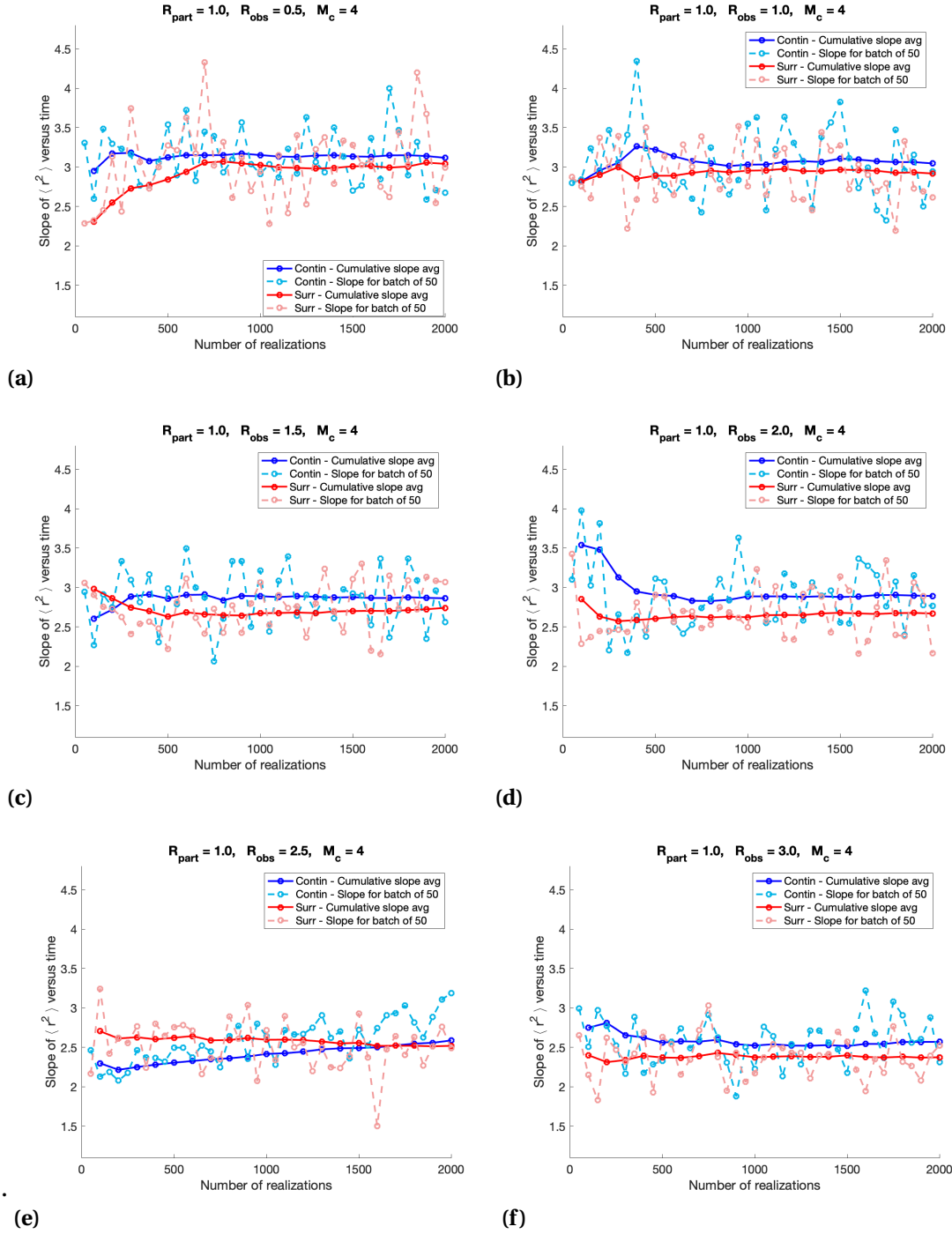
**Table 5.16** Four internal obstacles model. Slope values for  $\langle r^2 \rangle$  versus time for continuous model and surrogate model with the best  $M_c$  values.

$R_{\text{obs}}$	$M_c$	Continuous model slope	Surrogate model slope	Percent difference
0.5	4	3.1170	3.0410	2.438%
1.0	4	3.0459	2.9181	4.196%
1.5	4	2.8631	2.7416	4.244%
2.0	4	2.8914	2.6702	7.650%
2.5	4	2.5876	2.5195	2.632%
3.0	4	2.5700	2.3744	7.611%

After finding the appropriate  $M_c$  values, we are able to update the transition probabilities accordingly. Our resulting  $M_c$  values and probabilities are summarized in Table 5.17 -5.19. The probability of staying in a particular subdomain is dependent on the type of subdomain (center, onewall, corner). The rows in Table 5.17 correspond with the transition probability vector  $\vec{p}_C^F$ , the rows in Table 5.18 correspond with the transition probability vector  $\vec{p}_I^F$ , and the rows in Table 5.19 correspond with the transition probability vector  $\vec{p}_L^F$ , as detailed in section 3.3.2.1. For all subdomain types, the particle is most likely to stay in its current subdomain, especially when in the center subdomain, which has the most obstacles along its boundary.

### 5.2.2.1 Diffusivity for four internal obstacles

For the four internal obstacles case, we only considered the simulated diffusivity. Since this model has much fewer symmetries, we cannot relate the simulated diffusivities to a theoretical diffusivity from a Taylor Series expansion. The Taylor Series expansion in



**Figure 5.22** Four internal obstacles model. Slope of  $\langle r^2 \rangle$  vs. time over differing number of realizations, along with a cumulative average of the slope value. For  $R_{\text{part}} = 1.0$ . (a)  $R_{\text{obs}} = 0.5, M_c = 4$ , (b)  $R_{\text{obs}} = 1.0, M_c = 4$ , (c)  $R_{\text{obs}} = 1.5, M_c = 4$ , (d)  $R_{\text{obs}} = 2.0, M_c = 4$ , (e)  $R_{\text{obs}} = 2.5, M_c = 4$ , and (f)  $R_{\text{obs}} = 3.0, M_c = 4$ .

**Table 5.17** Four internal obstacles model. Transition probabilities for center subdomain (C) by transition type for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  for after 2000 realizations (for  $M_c = 4$ ).

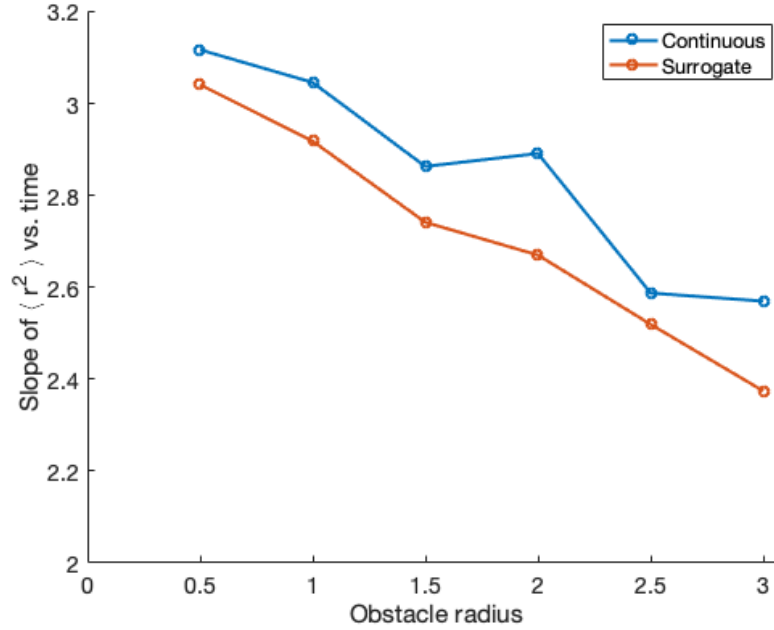
Obstacle radius	Probability by transition type	
	1	2
0.5	0.96160	$9.5959 \times 10^{-3}$
1.0	0.96492	$8.7682 \times 10^{-3}$
1.5	0.96826	$7.9345 \times 10^{-3}$
2.0	0.97167	$7.0811 \times 10^{-3}$
2.5	0.97520	$6.1998 \times 10^{-3}$
3.0	0.97971	$5.0736 \times 10^{-3}$

**Table 5.18** Four internal obstacles model. Transition probabilities for onewall subdomain (I) by transition type for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  for after 2000 realizations (for  $M_c = 4$ ).

Obstacle radius	Probability by transition type				
	1	2	3	4	5
0.5	0.95538	$1.1107 \times 10^{-2}$	$1.3266 \times 10^{-4}$	$1.2929 \times 10^{-2}$	$9.2035 \times 10^{-3}$
1.0	0.95671	$1.0741 \times 10^{-2}$	$1.3599 \times 10^{-4}$	$1.3365 \times 10^{-2}$	$8.1691 \times 10^{-3}$
1.5	0.95790	$1.0446 \times 10^{-2}$	$1.4010 \times 10^{-4}$	$1.3862 \times 10^{-2}$	$7.0601 \times 10^{-3}$
2.0	0.95877	$1.0249 \times 10^{-2}$	$1.4911 \times 10^{-4}$	$1.4554 \times 10^{-2}$	$5.8756 \times 10^{-3}$
2.5	0.95930	$1.0093 \times 10^{-2}$	$1.5779 \times 10^{-4}$	$1.5536 \times 10^{-2}$	$4.6658 \times 10^{-3}$
3.0	0.95941	$1.0060 \times 10^{-2}$	$1.7029 \times 10^{-4}$	$1.6809 \times 10^{-2}$	$3.3223 \times 10^{-3}$

**Table 5.19** Four internal obstacles model. Transition probabilities for corner subdomain (L) by transition type for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$  for after 2000 realizations (for  $M_c = 4$ ).

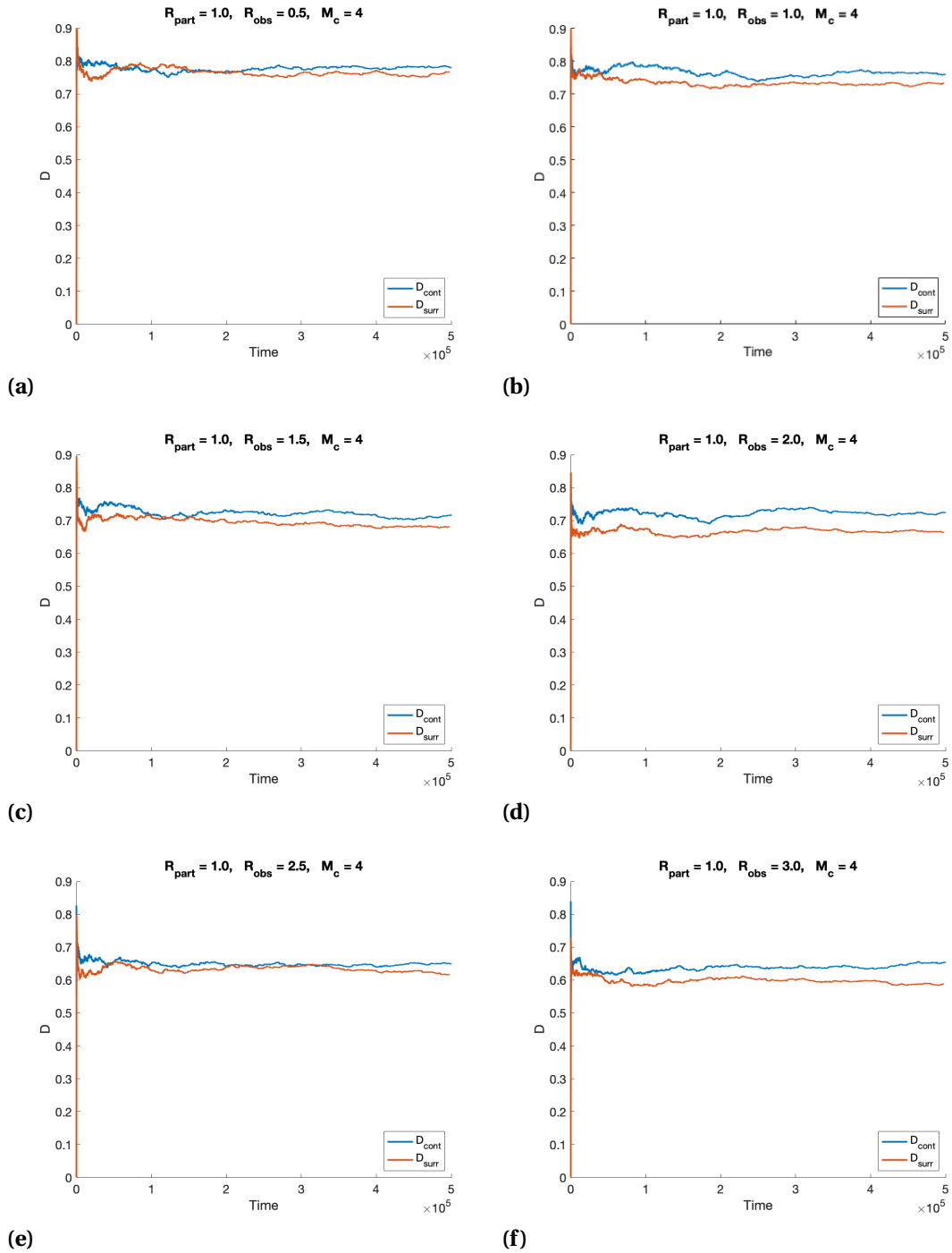
Obstacle radius	Probability by transition type				
	1	2	3	4	5
0.5	0.95236	$1.0912 \times 10^{-2}$	$1.3042 \times 10^{-4}$	$1.2712 \times 10^{-2}$	$1.2923 \times 10^{-4}$
1.0	0.95314	$1.0390 \times 10^{-2}$	$1.3129 \times 10^{-4}$	$1.2844 \times 10^{-2}$	$1.3243 \times 10^{-4}$
1.5	0.95352	$9.8880 \times 10^{-3}$	$1.3399 \times 10^{-4}$	$1.3149 \times 10^{-2}$	$1.3379 \times 10^{-4}$
2.0	0.95382	$9.4388 \times 10^{-3}$	$1.3689 \times 10^{-4}$	$1.3448 \times 10^{-2}$	$1.3630 \times 10^{-4}$
2.5	0.95400	$8.9760 \times 10^{-3}$	$1.4082 \times 10^{-4}$	$1.3811 \times 10^{-2}$	$1.4138 \times 10^{-4}$
3.0	0.95379	$8.5664 \times 10^{-3}$	$1.4594 \times 10^{-4}$	$1.4322 \times 10^{-2}$	$1.4506 \times 10^{-4}$



**Figure 5.23** Four internal obstacles model. Slope of  $\langle r^2 \rangle$  versus time for continuous model and resulting surrogate model with best  $M_c$  values ( $M_c = 4$  for all cases). Results for particle radius  $R_{\text{part}} = 1.0$  and varying obstacle radius  $R_{\text{obs}}$ .

the equally spaced obstacle case and in the multisize obstacle case relied on heavily on symmetric RVEs and symmetric transition probabilities. In the equally spaced obstacles case, we had only cardinal direction transitions, each of which had an equally probability of occurring, thus, we only had two total transition probabilities (stay versus leave), only one of which was an independent transition probability. In the multisize obstacles case, we also only had cardinal direction transitions, and four total probabilities (stay, leave vertically, leave horizontally in RVE, leave horizontally to extended domain), three of which were independent transition probabilities. However, in the four internal obstacles case, we had both cardinal direction transitions and diagonal transitions, and, additionally, the transition types and probabilities were highly depended on the subdomain type. We had twelve total transition probabilities, with two for the center subdomain, five for the onewall subdomain, and five for the corner subdomain. Of the twelve transition probabilities, nine were independent of the other subdomain-specific transition probabilities. Thus, when considering a Taylor Series expansion, we were unable to make many of the necessary assumptions needed to carry out such an investigation.

We estimated the effective diffusivities for the continuous and surrogate models ( $D_{\text{cont}}$  and  $D_{\text{surr}}$ , respectively) using  $\langle r^2 \rangle$  simulation results according to equation (5.9). Figure



**Figure 5.24** Four internal obstacles model. Estimated diffusivity for continuous and surrogate models for  $R_{\text{part}} = 1.0$  when  $M_c = 4$ . (a)  $R_{\text{obs}} = 0.5$ , (b)  $R_{\text{obs}} = 1.0$ , (c)  $R_{\text{obs}} = 1.5$ , (d)  $R_{\text{obs}} = 2.0$ , (e)  $R_{\text{obs}} = 2.5$ , (f)  $R_{\text{obs}} = 3.0$ .

5.24 shows the resulting estimated diffusivity versus time for each obstacle radius value. For a final comparison we average the last 100,000 values (last 10%) of  $D_{\text{cont}}$  and  $D_{\text{surr}}$ , as summarized in Table 5.20. The correlation coefficient for  $D_{\text{cont}}$  and  $D_{\text{surr}}$  was  $R = 0.9722$ .

**Table 5.20** Four internal obstacles model. Table showing the diffusivities obtained via simulation for the continuous and surrogate models ( $D_{\text{cont}}$  and  $D_{\text{surr}}$ , respectively) for each obstacle radius value (commitment index  $M_c = 4$  for all cases).

Obstacle radius	$D_{\text{cont}}$	$D_{\text{surr}}$	Percent difference
0.5	0.7815	0.7667	1.894%
1.0	0.7587	0.7322	3.493%
1.5	0.7159	0.6804	4.959%
2.0	0.7245	0.6634	8.433%
2.5	0.6499	0.6165	5.139%
3.0	0.6531	0.5878	9.998%

### 5.2.3 Multisize obstacles

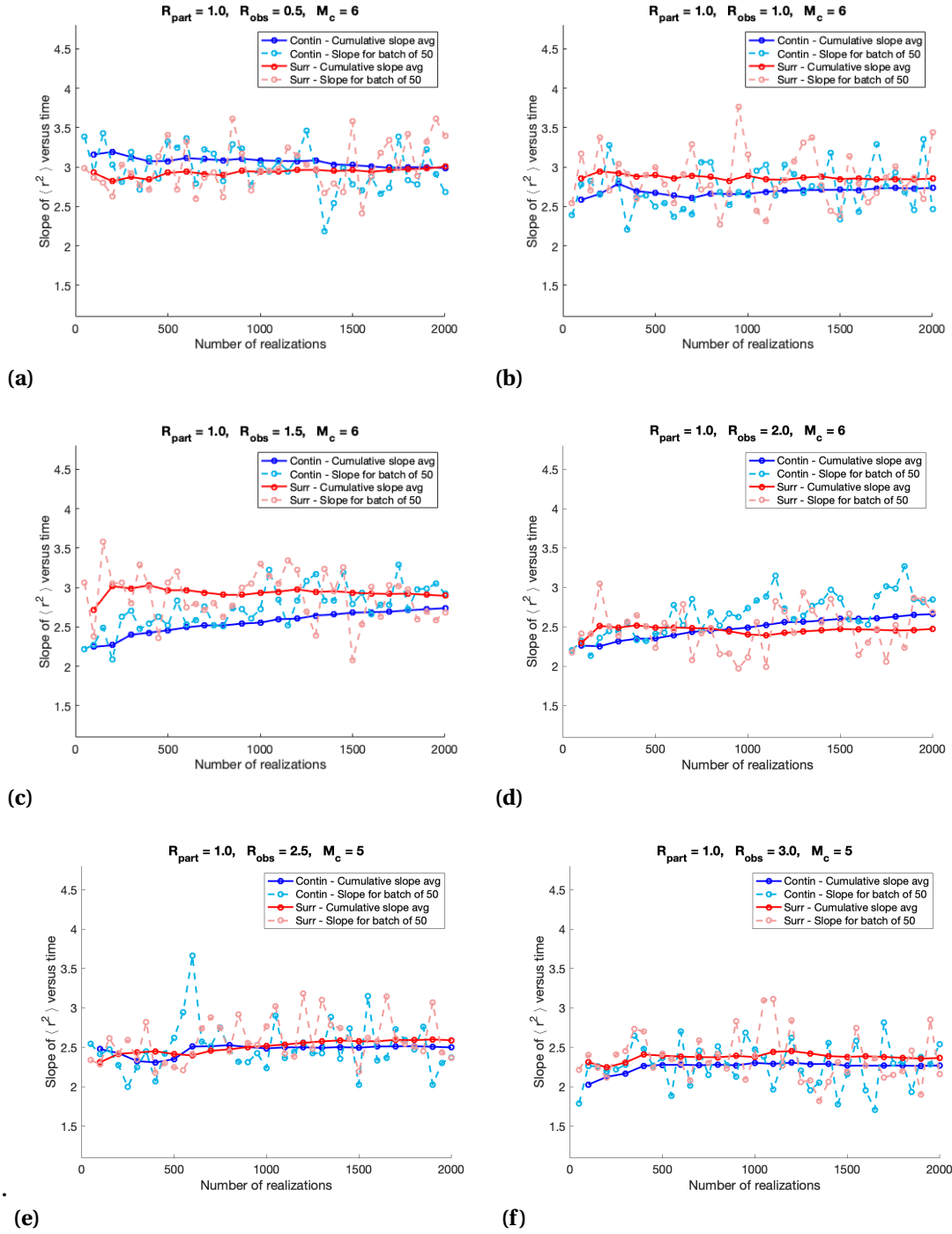
As in the cases of equally spaced obstacles and four internal obstacles, for the multisize obstacles model we began by identifying the optimal values of  $M_c$ . We fixed  $R_{\text{part}} = 1.0$  and  $R_{\text{obs}_1} = 1.0$  and varied  $R_{\text{obs}_2}$ . For each obstacle radius  $R_{\text{obs}}$  we ran the surrogate model multiple times using probabilities obtained using the continuous model as the commitment index  $M_c$  was varied from 3 to 8. We determined the value for  $M_c$  by comparing the continuous and surrogate plots of mean squared displacement  $\langle r^2 \rangle$  versus time, again considering the resulting slope of  $\langle r^2 \rangle$  versus time to be a quantity of interest.

Figure 5.25 visualizes these resulting slopes versus increasing realizations for the best values of  $M_c$  for all cases in which the obstacle radius was varied. Note that the best value for  $M_c$  ranges from  $M_c = 5$  ( $R_{\text{obs}} = 3$ ) up to  $M_c = 6$  ( $R_{\text{obs}} = 0.5$ ). Based on these results, it appears that our slope values for both the continuous and surrogate cases have reached stationary values. We also observe agreement between the continuous and surrogate models.

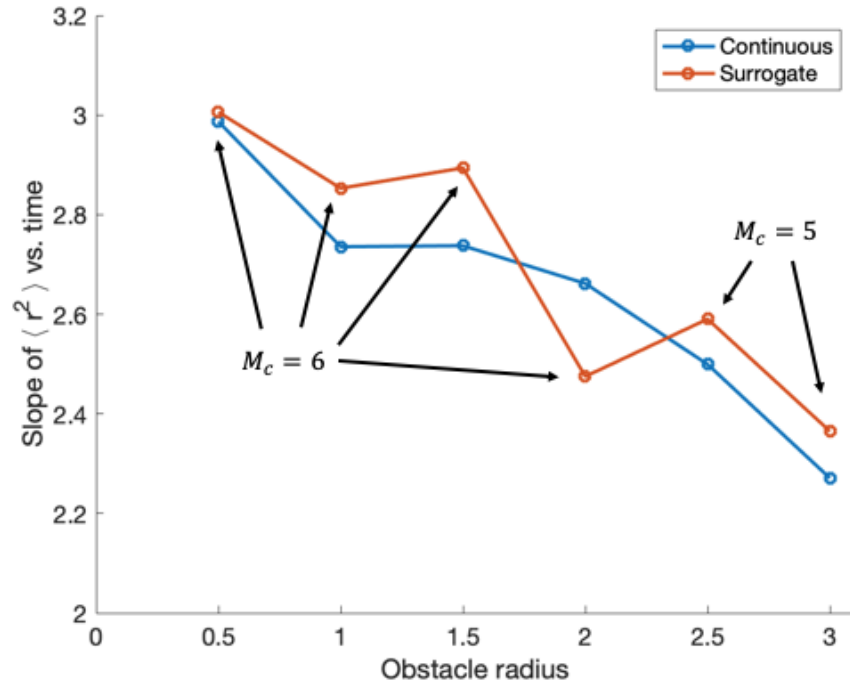
In Figure 5.26 we show the resulting slope of  $\langle r^2 \rangle$  versus time for increasing obstacle radius values, showing results for the continuous model and the surrogate model with the best  $M_c$  value. We summarize these slope results and their percent difference in Table 5.21. The correlation coefficient for continuous and surrogate model results was  $R = 0.8789$ .

After finding the appropriate  $M_c$  values, we are able to update the transition probabilities accordingly. Our resulting  $M_c$  values and probabilities are summarized in Table 5.22. The





**Figure 5.25** Multisize obstacles model. Slope of  $\langle r^2 \rangle$  vs. time over differing number of realizations, along with a cumulative average of the slope value. For  $R_{\text{part}} = 1.0$ . (a)  $R_{\text{obs}} = 0.5, M_c = 7$ , (b)  $R_{\text{obs}} = 1.0, M_c = 6$ , (c)  $R_{\text{obs}} = 1.5, M_c = 5$ , (d)  $R_{\text{obs}} = 2.0, M_c = 5$ , (e)  $R_{\text{obs}} = 2.5, M_c = 4$ , and (f)  $R_{\text{obs}} = 3.0, M_c = 4$ .



**Figure 5.26** Multisize obstacles model. Slope of  $\langle r^2 \rangle$  versus time for continuous model and resulting surrogate model with best  $M_c$  values. Results for  $R_{\text{part}} = 1.0$ ,  $R_{\text{obs}_1} = 1.0$ , and varying  $R_{\text{obs}_2}$ .

**Table 5.21** Multisize obstacles model. Slope values for  $\langle r^2 \rangle$  versus time for continuous model and surrogate model with the best  $M_c$  values.  $R_{\text{part}} = 1.0$ ,  $R_{\text{obs}_1} = 1.0$ .

$R_{\text{obs}_2}$	$M_c$	Continuous model slope	Surrogate model slope	Percent difference
0.5	6	2.9870	3.0059	0.632%
1.0	6	2.7357	2.8531	4.291%
1.5	6	2.7376	2.8939	5.709%
2.0	6	2.6618	2.4751	7.014%
2.5	5	2.4995	2.5904	3.637%
3.0	5	2.2710	2.3651	4.144%

rows in Table 5.22 correspond with the transition probability vector  $\vec{p}^M$  (section 3.3.3.1).

**Table 5.22** Multisize obstacles model. Transition probabilities for center subdomain by transition type for particle radius  $R_{\text{part}} = 1.0$ ,  $R_{\text{obs}_1} = 1.0$  and varying obstacle radius  $R_{\text{obs}_2}$  for after 2000 realizations (for  $M_c = 0$ ).

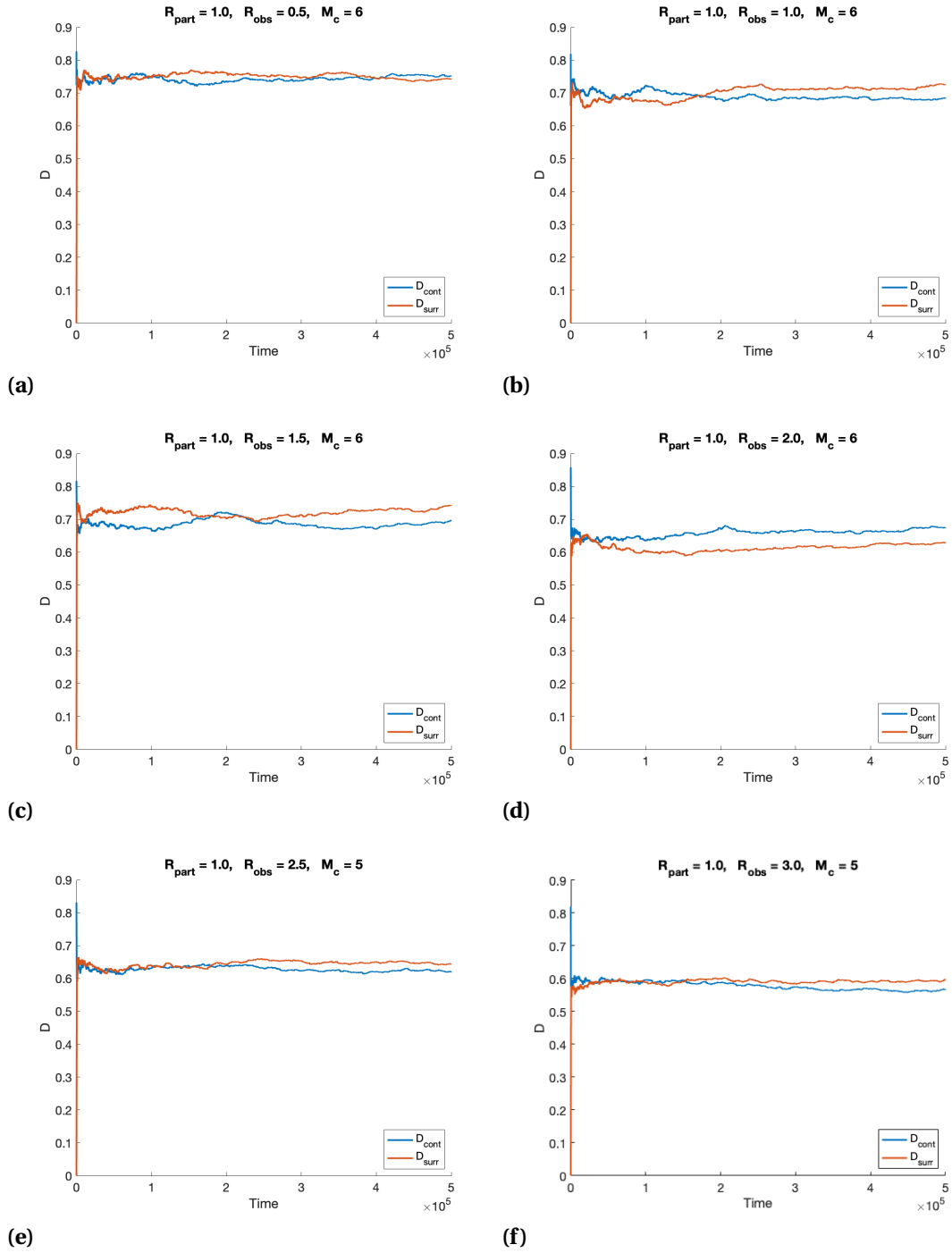
$R_{\text{obs}_2}$	$M_c$	Probability by transition type			
		1	2	3	4
0.5	6	0.99253	$1.8680 \times 10^{-3}$	$2.0159 \times 10^{-3}$	$1.7147 \times 10^{-3}$
1.0	6	0.99291	$1.7715 \times 10^{-3}$	$1.7764 \times 10^{-3}$	$1.7725 \times 10^{-3}$
1.5	6	0.99324	$1.6894 \times 10^{-3}$	$1.5339 \times 10^{-3}$	$1.8510 \times 10^{-3}$
2.0	6	0.99352	$1.6203 \times 10^{-3}$	$1.2909 \times 10^{-3}$	$1.9522 \times 10^{-3}$
2.5	5	0.99316	$1.7098 \times 10^{-3}$	$1.1322 \times 10^{-3}$	$2.2850 \times 10^{-3}$
3.0	5	0.99339	$1.6502 \times 10^{-3}$	$8.1439 \times 10^{-4}$	$2.4950 \times 10^{-3}$

### 5.2.3.1 Diffusivity for multisize obstacles

For the multisize obstacles case, we again estimate the effective diffusivities from simulation results for the continuous and surrogate models ( $D_{\text{cont}}$  and  $D_{\text{surr}}$ , respectively) according to equation (5.9). Figure 5.27 shows the resulting estimated diffusivity versus time for each obstacle radius value. For a final comparison we average the last 100,000 values (last 10%) of  $D_{\text{cont}}$  and  $D_{\text{surr}}$ , as summarized in Table 5.23. The diffusivities for the continuous and surrogate models are correlated, with a correlation coefficient of  $R = 0.8503$ .

**Table 5.23** Multisize obstacles model. Table showing commitment index  $M_c$ , probability of staying in current subdomain  $q$ , and diffusivities obtained via simulation for the continuous and surrogate models ( $D_{\text{cont}}$  and  $D_{\text{surr}}$ , respectively) for each obstacle radius value.

Obstacle radius	$M_c$	$q$	$D_{\text{cont}}$	$D_{\text{surr}}$	Percent difference
0.5	6	0.9925	0.7509	0.7428	1.084%
1.0	6	0.9929	0.6842	0.7261	6.133 %
1.5	6	0.9932	0.6943	0.7418	6.845 %
2.0	6	0.9935	0.6746	0.6290	6.755%
2.5	5	0.9932	0.6205	0.6440	3.785%
3.0	5	0.9939	0.5667	0.5942	4.859 %



**Figure 5.27** Multisize obstacles model. Estimated effective diffusivity for continuous and surrogate models for  $R_{\text{part}} = 1.0$ ,  $R_{\text{obs}_1} = 1.0$  (based on simulation results). (a)  $R_{\text{obs}_2} = 0.5$ ,  $M_c = 6$ . (b)  $R_{\text{obs}_2} = 1.0$ ,  $M_c = 6$ . (c)  $R_{\text{obs}_2} = 1.5$ ,  $M_c = 6$ . (d)  $R_{\text{obs}_2} = 2.0$ ,  $M_c = 6$ . (e)  $R_{\text{obs}_2} = 2.5$ ,  $M_c = 6$ . (f)  $R_{\text{obs}_2} = 3.0$ ,  $M_c = 6$ .

Based on the structure in the multisize obstacles model, we expect to have different inherent diffusivities in the horizontal and vertical directions. In section 4.2.4, we used the following probabilities:

- (i)  $q_1$  is the probability that the particle stays in its current subdomain;
- (ii)  $q_2$  is the probability that the particle moves to a subdomain in the horizontal direction;
- (iii)  $q_3$  is the probability that the particle moves to a subdomain in the vertical direction;

and found the diffusion equation to be:

$$u_t = \frac{(\Delta x_M)^2}{2(\Delta t)} (q_2 u_{xx} + q_3 u_{yy}). \quad (5.15)$$

We can then consider the horizontal diffusivity,  $D_H$ , to be

$$D_H = \frac{q_2 (\Delta x_M)^2}{2(\Delta t)}, \quad (5.16)$$

and the vertical diffusivity,  $D_V$ , to be

$$D_V = \frac{q_3 (\Delta x_M)^2}{2(\Delta t)}, \quad (5.17)$$

where  $\Delta x_M$  is the effective spatial step size, and  $D_H$  and  $D_V$  are horizontal and vertical diffusivities, respectively. However, due to one column of obstacles having radius  $R_{\text{obs}_1} = 1.0$  and the other column of obstacles having varying radius  $R_{\text{obs}_2}$ , the particle has a bias in either the right or left direction, depending on where it is currently located. Thus, we cannot use these equations to calculate  $\Delta x_M$  as in the equally spaced obstacles model.

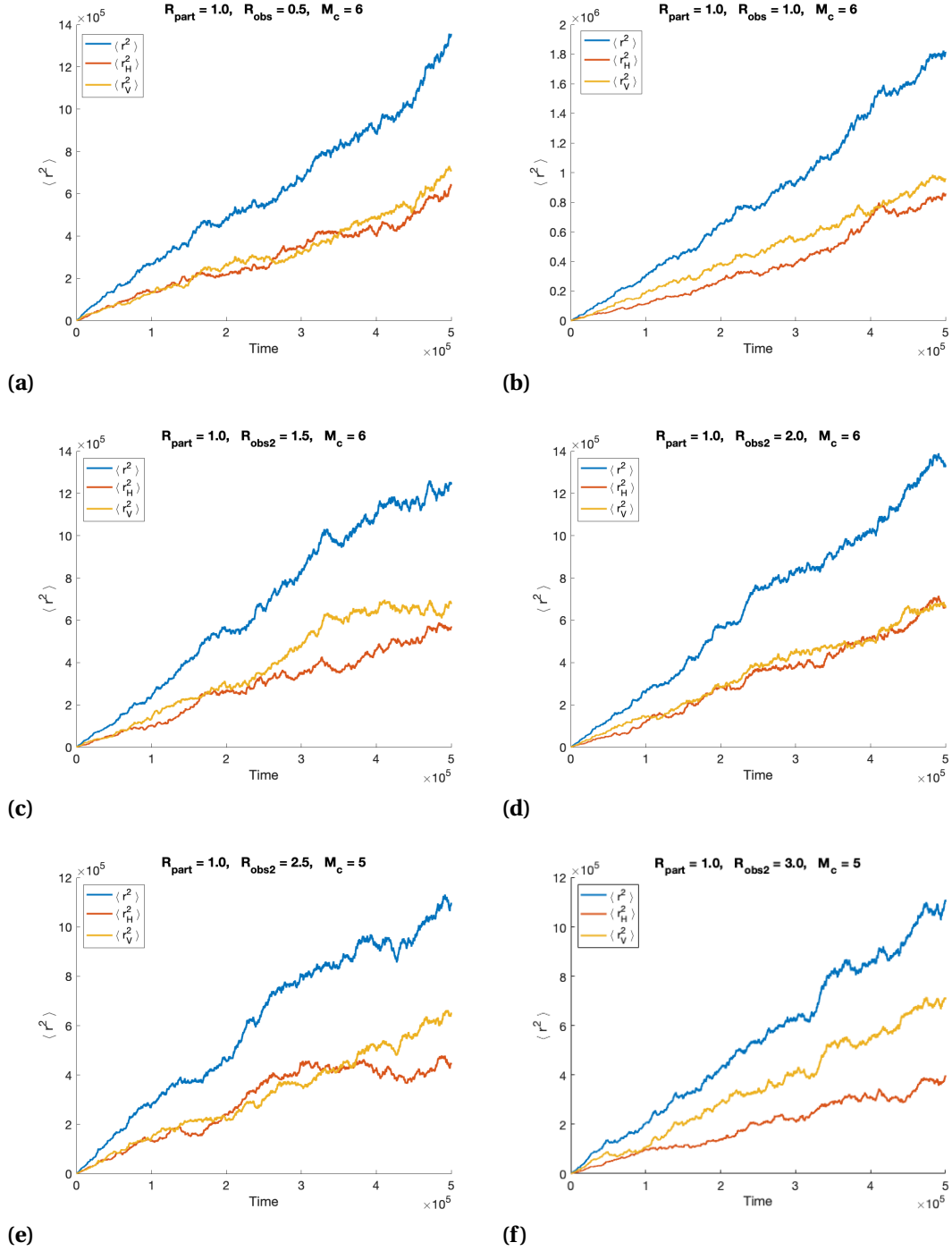
We consider simulation results in light of horizontal versus vertical diffusion. In the surrogate model we record the horizontal and vertical mean square displacements,  $\langle r_H^2 \rangle$  and  $\langle r_V^2 \rangle$ , respectively, by using the center coordinate of the corresponding subdomain in the continuous model. Figure 5.28 shows these displacements.

Although we cannot calculate the horizontal and vertical diffusivities directly, we can calculate the ratio

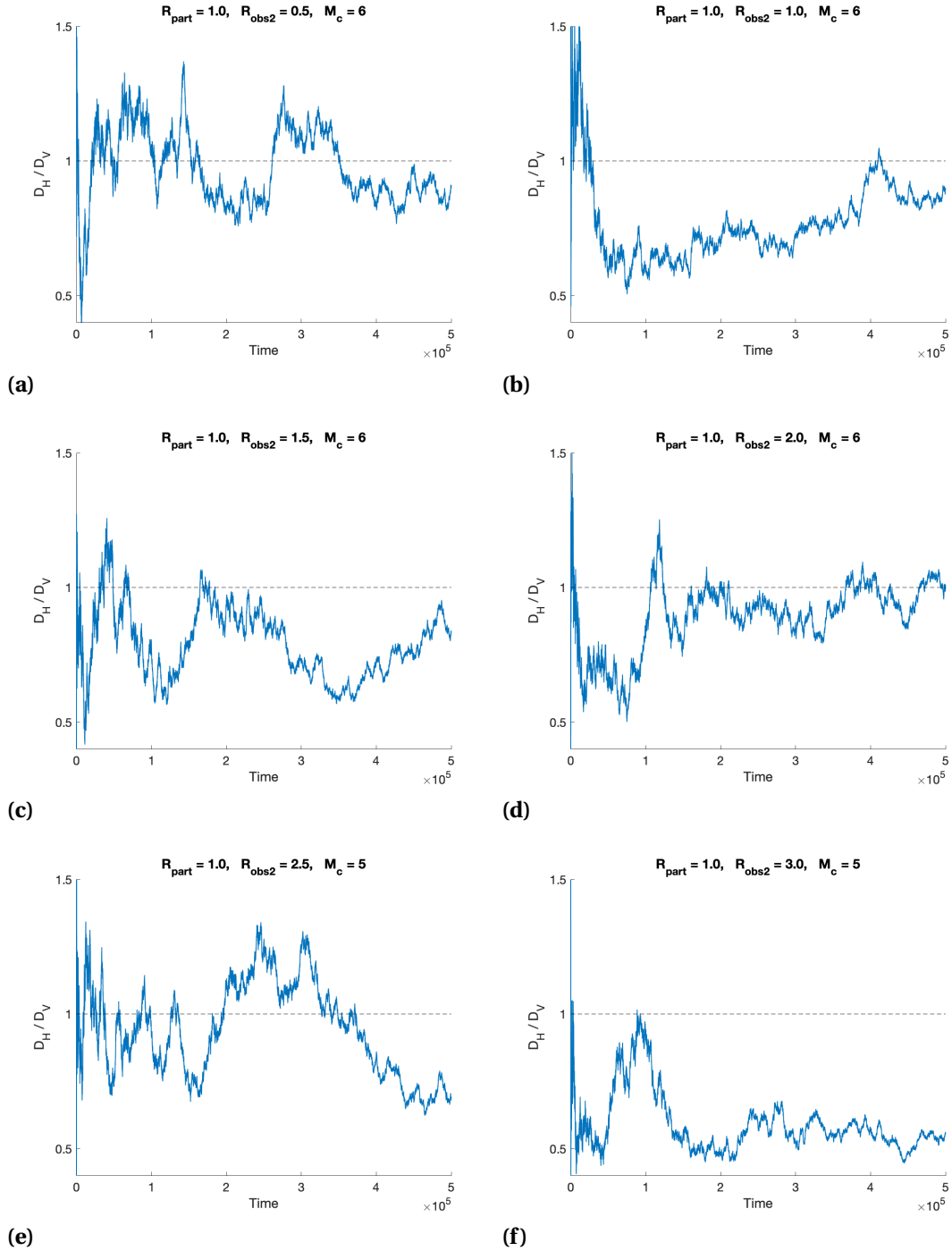
$$\frac{D_H}{D_V}, \quad (5.18)$$

using the ratio

$$\frac{\langle r_H^2 \rangle}{\langle r_V^2 \rangle}. \quad (5.19)$$

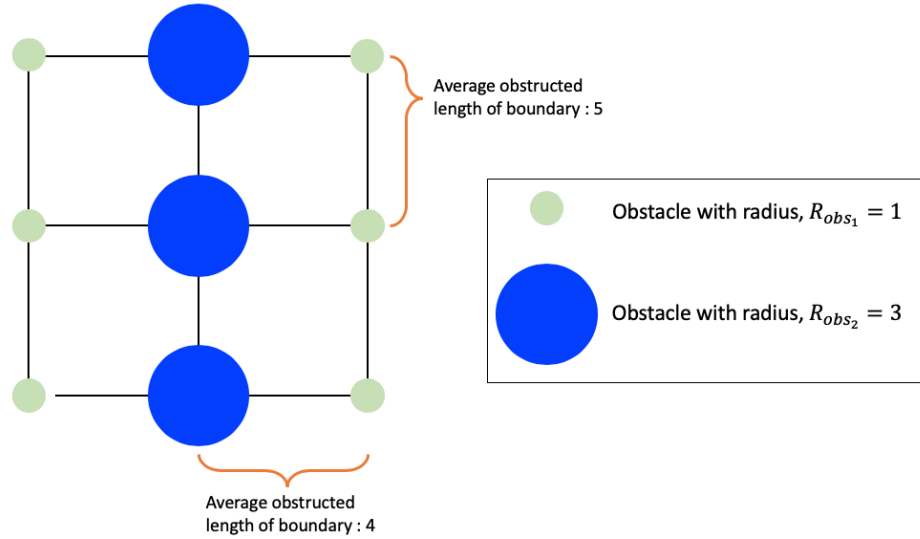


**Figure 5.28** Multisize obstacles model. Horizontal and vertical mean squared displacements for the surrogate model.  $R_{\text{part}} = 1.0, R_{\text{obs}_1} = 1.0$ . (a)  $R_{\text{obs}_2} = 0.5, M_c = 6$ . (b)  $R_{\text{obs}_2} = 1.0, M_c = 6$ . (c)  $R_{\text{obs}_2} = 1.5, M_c = 6$ . (d)  $R_{\text{obs}_2} = 2.0, M_c = 6$ . (e)  $R_{\text{obs}_2} = 2.5, M_c = 6$ . (f)  $R_{\text{obs}_2} = 3.0, M_c = 6$ .



**Figure 5.29** Multisize obstacles model. Ratio of horizontal diffusivity to vertical diffusivity  $R_{\text{part}} = 1.0, R_{\text{obs1}} = 1.0$ . (a)  $R_{\text{obs2}} = 0.5, M_c = 6$ . (b)  $R_{\text{obs2}} = 1.0, M_c = 6$ . (c)  $R_{\text{obs2}} = 1.5, M_c = 6$ . (d)  $R_{\text{obs2}} = 2.0, M_c = 6$ . (e)  $R_{\text{obs2}} = 2.5, M_c = 6$ . (f)  $R_{\text{obs2}} = 3.0, M_c = 6$ .

We illustrate the ratio of diffusivities in Figure 5.29. In the multisize obstacles case, depending on the value of  $R_{\text{obs}_2}$ , the RVE is more open in either the horizontal or vertical direction. This effect on the diffusivity is subtle except for in our most extreme case,  $R_{\text{obs}_2} = 3.0$  (depicted in Figure 5.30), where the average boundary length obstructed is greater in the *horizontal* direction than in the vertical direction, and so our ratio  $D_H/D_V < 1$ .



**Figure 5.30** Multisize obstacles model. RVE illustrating case where  $R_{\text{obs}_1} = 1.0$  and  $R_{\text{obs}_2} = 3.0$  (to scale).

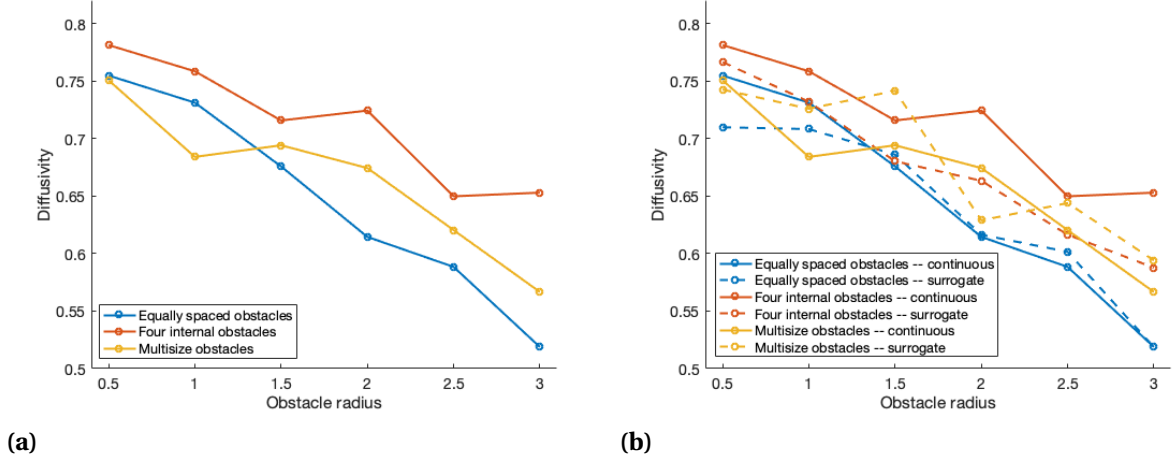
### 5.3 Model comparison

The models for 2D diffusion via random walks in a RVE with obstacles, developed in this chapter, can help identify the relationships between geometric features of the material being modeled and diffusivity of the particle undergoing diffusive transport. To this end, we compare the simulated diffusivity versus the varied obstacle radius ( $R_{\text{obs}}$  or  $R_{\text{obs}_2}$ ) for the equally spaced obstacles model, the four internal obstacles model, and the multisize obstacles model in Figure 5.31(a). We also summarize the corresponding diffusivities in Table 5.24.

We observe that the continuous models were most similar to each other for small obstacle radii and that this difference tends to increase as the obstacles increase in size. The continuous model and corresponding surrogate model were in best agreement for the equally spaced obstacles model. This finding suggests that the accuracy of our approach



for developing a surrogate lattice model may decrease as the complexity of the material being modeled increases.



**Figure 5.31** Comparison of diffusivity versus obstacle radius for all models considered. (a) Continuous models only. (b) Continuous and surrogate models.

**Table 5.24** Diffusivity across obstacle radii for all models considered.

		Obstacle radius					
		0.5	1.0	1.5	2.0	2.5	3.0
Equally spaced obstacles model	$D_{\text{cont}}$	0.7550	0.7314	0.6759	0.6146	0.5885	0.5188
	$D_{\text{surr}}$	0.7100	0.7084	0.6862	0.6165	0.6016	0.5187
Four internal obstacles model	$D_{\text{cont}}$	0.7509	0.6842	0.6943	0.6745	0.6205	0.5667
	$D_{\text{surr}}$	0.7428	0.7261	0.7418	0.6290	0.6440	0.5942
Multisize obstacles model	$D_{\text{cont}}$	0.7815	0.7587	0.7159	0.7245	0.6499	0.6531
	$D_{\text{surr}}$	0.7667	0.7322	0.6804	0.6634	0.6165	0.5878

For comparison purposes, we also ran a continuous model simulation with obstacle collisions turned off, i.e.  $R_{obs} = 0$ . For this special case, we again ran the continuous model for  $10^7$  time steps over 2000 realizations. For the obstacle free model, the slope of  $\langle r^2 \rangle$  versus time was 2.1026, and the diffusivity from simulation results (averaging the last 100,000 timesteps) was  $D_{\text{cont}} = 0.8733$ .

With the addition of the obstacle-free simulation, we can also compare our diffusivity

estimates to the classical Mackie-Mears relation [22] used in many continuum porous media models [3].

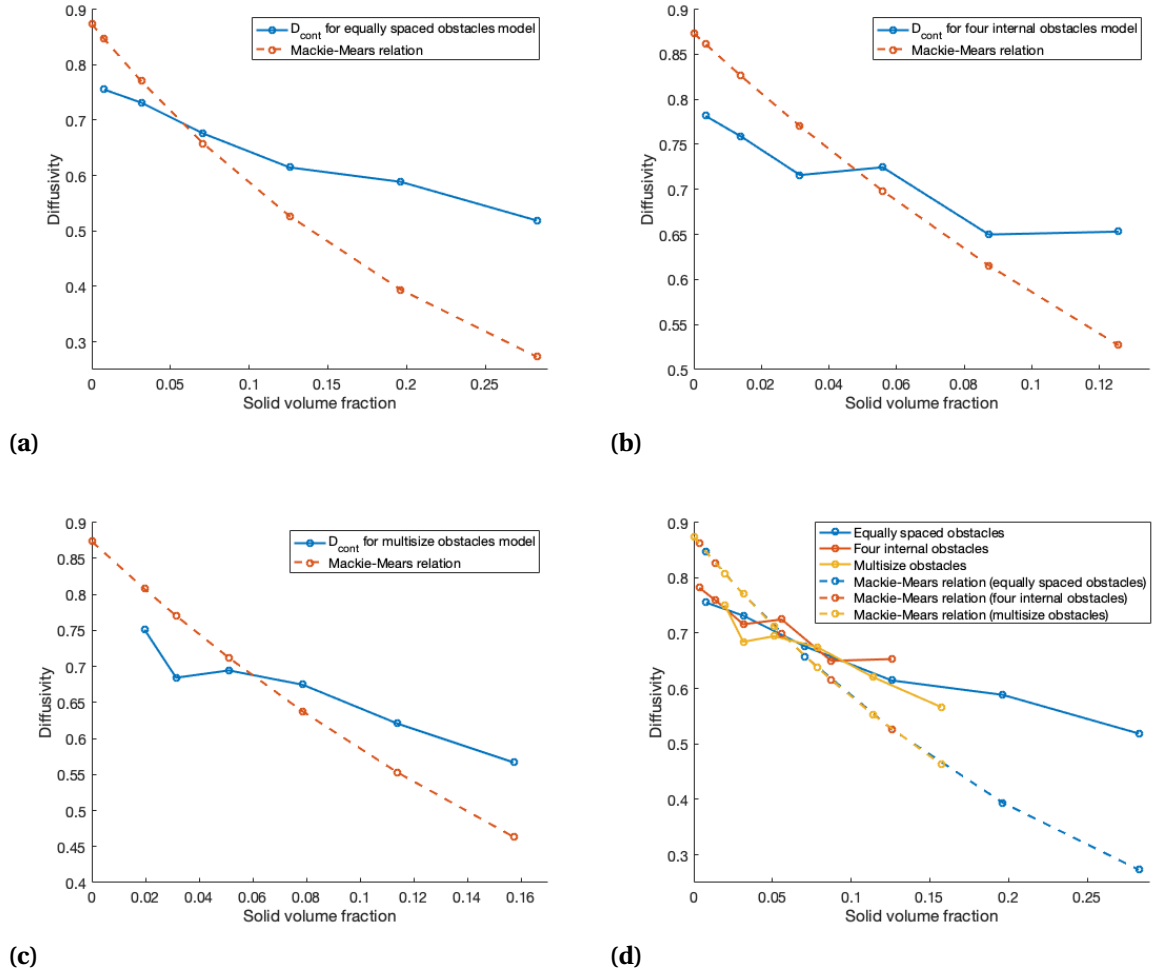
$$D^N[\phi^S] = D_0^N \left( \frac{1 - \phi^2}{1 + \phi^S} \right)^2. \quad (5.20)$$

In equation (5.20),  $\phi^S$  is the solid volume fraction in a mixture,  $D_0^N$  is the diffusivity in a free solution (one without obstacles) and, thus,  $D^N[\phi^S]$  describes the diffusivity with varying porosity [22]. For our three models of a porous medium, we calculate the corresponding solid volume fractions as follows:

$$\begin{aligned} \phi_E^S &= \frac{9\pi R_{\text{obs}}^2}{900}, \\ \phi_F^S &= \frac{4\pi R_{\text{obs}}^2}{900}, \\ \phi_M^S &= \frac{2\pi R_{\text{obs}_1}^2 + 2\pi R_{\text{obs}_2}^2}{400}, \end{aligned} \quad (5.21)$$

where  $\phi_E^S$ ,  $\phi_F^S$ , and  $\phi_M^S$  are the solid volume fractions for the equally spaced obstacles model, the four internal obstacles model, and the multisize obstacles model, respectively. Note that we assume a constant thickness in the RVE and obstacles and recall that the RVE is of size  $30 \times 30$  for the first two models and  $20 \times 20$  for the multisize obstacles model. We plot model diffusivities versus solid volume fraction along with the Mackie-Mears relations in Figure 5.32. The estimated diffusivities from our models compared to those from the Mackie-Mears relation are shown in Tables 5.25 - 5.27.

The Mackie-Mears relation applies to dilute solutions and assumes isotropy in the material through which the particle is diffusing. Based on Fig. 5.32, we first observe that the relations predicted for our three models are much closer to each other than to the (classical) Mackie-Mears relation. Second, we observe that relationships between diffusivity and the RVE solid volume fraction predicted by our three models are significantly different from those observed by the classical Mackie-Mears model. The first observation suggests that while each of our three models produces differing relations, the nature of the anisotropy may be similar. This is likely due to the property that all three models are comprised of circular obstacles of similar sizes arranged along principal (coordinate) directions in the RVE. Lastly, the second observation suggests that the accuracy of relations between diffusivity and solid volume fraction in porous material models could be improved using the approach developed herein.



**Figure 5.32** Comparison of effective diffusivity for continuous simulations and corresponding Mackie-Mears relation. (a) Equally spaced obstacles model. (b) Four internal obstacles model. (c) Multisize obstacles. (d) All models considered and Mackie-Mears relation. (a)-(c) plotted at a different scale.

**Table 5.25** Diffusivity across obstacle radii and corresponding solid volume fraction for equally spaced obstacles model and Mackie-Mears relation.

Obstacle radius	0.5	1.0	1.5	2.0	2.5	3.0
Solid volume fraction	0.0079	0.0314	0.0707	0.1257	0.1963	0.2827
Diffusivity $D_{\text{cont}}$	0.7550	0.7314	0.6759	0.6146	0.5885	0.5188
Mackie-Mears Relation	0.8463	0.7702	0.6579	0.5269	0.3941	0.2730

**Table 5.26** Diffusivity across obstacle radii and corresponding solid volume fraction for four internal obstacles model and Mackie-Mears relation.

Obstacle radius	0.5	1.0	1.5	2.0	2.5	3.0
Solid volume fraction	0.0035	0.0140	0.0314	0.0559	0.0873	0.1257
Diffusivity $D_{\text{cont}}$	0.7815	0.7587	0.7159	0.7245	0.6499	0.6531
Mackie-Mears Relation	0.8612	0.8259	0.7702	0.6983	0.6154	0.5269

**Table 5.27** Diffusivity across obstacle radii and corresponding solid volume fraction for multi-sized obstacles model and Mackie-Mears relation.

Obstacle radius	0.5	1.0	1.5	2.0	2.5	3.0
Solid volume fraction	0.0196	0.0314	0.0511	0.0785	0.1139	0.1571
Diffusivity $D_{\text{cont}}$	0.7509	0.6842	0.6943	0.6745	0.6205	0.5667
Mackie-Mears Relation	0.8073	0.7702	0.7119	0.6375	0.5527	0.4635

## 5.4 Conclusions

Direct numerical simulation is a useful tool in modeling systems of particles. Direct simulation approaches using representative volume elements or agent-based models have both been used to analyze system interactions in various applications [13, 14, 2, 25, 3]. When carried out over many realizations, such simulations can yield information used to develop probabilistic surrogate models that are able to reproduce essential features of the full direct simulation [4]. Lattice-based models are commonly used in biological applications such as tissue modeling or cancer modeling [3, 14]. Previous studies have used lattice models of single particles to analyze diffusion in crowded media [32, 31], and local, microscopic properties of the material strongly affect the overall diffusivity.

In this dissertation, we developed continuous model simulations of particle movement and then used these results to develop probabilistic surrogate models. We developed two different models: Model 1, which simulated a two-dimensional particle system, and Model 2, which simulated 2D single particle random walks for different obstacle configurations. In Model 1 we modeled twenty-seven particles which had perfectly elastic collisions with other particles and with domain boundaries. We considered varying values of particle

radii, tracking state changes for the different subdomains and used these results to develop a surrogate Markov chain (MC) model. Continuous and MC model results were in good agreement, and we observed a linear relationship between subdomain states and the particle radius size for each subdomain type. The approach used was also able to provide estimates of uncertainty for the quantities of interest.

In Model 2 we simulated single particle random walks in a 2D RVE with obstacles. We considered varying obstacle configurations (equally spaced obstacles, four internal obstacles, and multisize obstacles). For each obstacle configuration we considered a fixed particle radius with varying obstacle radii, and, in the equally spaced obstacles model, we also considered a fixed obstacle radius and varying particle radius. We divided the RVE into subdomains based on obstacle location. Subdomain transition probabilities were developed and used to construct the surrogate lattice model. As described in previous studies, we observed a linear relationship between the mean squared displacement and time. The mean squared displacement was used to estimate diffusivities for both the continuous models and surrogate models. Additionally, we related the equally spaced obstacle model's effective diffusivity from simulation results to a theoretical estimate of diffusivity, which we obtained by (1) using a Taylor Series expansion of random walks on a 2D lattice and (2) approximating an effective  $\Delta x$  for the theoretical diffusivity estimate on the lattice. We evaluated the effect of obstacle configuration and obstacle radius size on the slope of mean squared displacement versus time and on the estimated diffusivity.

The continuous and surrogate models were in good agreement for all obstacle configurations, and were in best agreement for the less anisotropic models, in particular the equally spaced obstacle model. This study showed that direct simulation of diffusion in continuous media with obstacles can be represented via lattice-based surrogate models. Our surrogate models efficiently preserved key features of the continuous model. In future studies, this approach can be applied to a variety of materials with different obstacle arrangements, shapes, and symmetries.

In Chapter 2 we introduced Model 1, which models a two-dimensional interacting particle system and serves as an initial motivating problem. We directly simulated the continuous model, which consisted of twenty-seven particles that exhibited perfectly elastic collisions with other particles and with obstacle boundaries. We partitioned the domain into nine subdomains, enabling us to track state-changes (number of particles in subdomain) for the different subdomain types. We then used these results to develop a surrogate Markov chain (MC) model. We observed excellent agreement between the continuous model and surrogate model results. A truncated normal distribution was well-suited to fitting a probability density function for each of the three subdomain types. In addition

to good agreement between continuous model results and expected values from the MC model, using a probability density function also provided an estimate of uncertainty. We observed a linear relationship between average number of particles and the particle radius size for each subdomain type, with the average number of particles decreasing as the particle radius increased in the corner subdomain, and the average number of particles increasing as the particle radius increased in the center subdomain.

In Chapter 3 we described Model 2, which models single particle random walks in a 2D representative volume element. We described the three different continuous model obstacle configurations – equally spaced obstacles, four internal obstacles, and multisize obstacles. Subdomains within the RVE were identified based on obstacle positions, and, in the four internal obstacles case, we had three distinct subdomain types (center, onewall, and corner). For all models we fixed the particle radius at  $R_{\text{part}} = 1.0$  and varied the obstacle radius  $R_{\text{obs}}$  (or, in the multisize obstacles case, we also fixed  $R_{\text{obs}_1} = 1.0$  and varied  $R_{\text{obs}_2}$ ). In the equally spaced obstacles model, we also considered fixed obstacle radius  $R_{\text{obs}} = 2.0$  and varying particle radius  $R_{\text{part}}$ . We observed similar sensitivity to varying particle radius as to varying obstacle radius, so we chose to focus on varying obstacle radius for the other two models.

In the second half of Chapter 3 we discussed development of the surrogate model. We used results from the continuous model to develop transition probabilities for the surrogate model. Each model had a different number of transition probabilities: the equally spaced obstacles model had two transition probabilities, the four internal obstacles model had twelve transition probabilities, and the multisize obstacles model had three transition probabilities. The equally spaced obstacles model and multisize obstacles model only had cardinal direction transitions, but in the four internal obstacles model, there were both cardinal direction transitions and diagonal transitions.

In Chapter 4 we discussed model calibration and diffusivity estimation. We introduced the commitment index  $M_c$  for identifying when a particle commits to a new subdomain (due to the particle's random walk, it has a tendency to cross and recross a boundary between subdomains multiple times). In order to estimate the diffusivity for the continuous and surrogate models, we used results from previous studies relating a particle's mean squared displacement  $\langle r^2 \rangle$  and diffusivity [32, 31, 35]. We also developed a theoretical estimate for diffusivity  $D_{\text{TS}}$  using a 2D Taylor Series expansion for a random walk on a lattice. Limits for a random walk on a 2D lattice were used in order to relate and compare the effective diffusivity from surrogate result  $D_{\text{surr}}$  with the theoretical diffusivity  $D_{\text{TS}}$ , which then allowed us to identify an effective  $\Delta x$  for  $D_{\text{TS}}$ .

In Chapter 5 we presented results for Model 2. We first discussed continuous model

results for each obstacle configuration. Transition probabilities from the continuous model were developed for increasing values of  $M_c$ , comparing the resulting surrogate model results to identify the best value for  $M_c$  for each obstacle configuration and radius. The resulting transition probabilities were dependent on the obstacle configuration and obstacle radius value. For all obstacle configurations, the probability of staying in the current subdomain was overwhelmingly more likely than any type of transition.

For all three obstacle configurations, we had good agreement between the continuous and surrogate models. We used the mean squared displacement  $\langle r^2 \rangle$  to compute an effective diffusivity for all models, which we also compared between continuous and surrogate models ( $D_{\text{cont}}$  and  $D_{\text{surr}}$ , respectively). For the equally spaced obstacles configuration, we also computed a theoretical estimate for diffusivity,  $D_{\text{TS}}$ , using a Taylor Series expansion of a random walk on a 2D lattice. We then related the effective diffusivity for the surrogate model  $D_{\text{surr}}$  to  $D_{\text{TS}}$  by finding an effective  $\Delta x$ . For the multisize obstacle case, we also estimated horizontal versus vertical diffusivity via the ratio of  $D_H/D_V$ .

Lastly, the effects of the three different obstacle configurations on diffusivity were compared. We also evaluated the diffusivity versus solid volume fraction for all models and compared these results to the classical Mackie-Mears relationship for diffusion in porous media. Our models were in better agreement with each other than with the Mackie-Mears relationship.

Overall, the techniques developed in this dissertation allow for internal geometric properties to be directly incorporated into the modeling and simulation process, enabling the calculation of diffusivity and the identification of an associated surrogate model that can be incorporated into lattice-based modeling frameworks, where diffusive transport is one of many important interacting mechanisms. In Model 1 we showed that a Markov chain model can be developed when the quantities of interest can be uniquely identified as states with transitions among the states (e.g. number of particles per subdomain). However, in Model 2, the quantity of interest was the mean squared displacement, so a Markov chain model was not possible, but a surrogate lattice model efficiently reproduced characteristics of the continuous model. Since this approach is versatile, the RVE geometry can be set up in many ways in the continuous model and can be tailored to the specific material being investigated. However, continuous model simulations are time-intensive and require the use of HPC. This approach could also be extended to a 3D RVE, but one would need to balance the complexity of the model under consideration with the specific HPC resources available to ensure that results can be obtained in a practical amount of computation time.

## BIBLIOGRAPHY

- [1] Alarcón, T., Bryne, H. & Maini, P. “A multiple scale model for tumor growth”. *Multiscale Modeling & Simulation* **3.2** (2005), pp. 440–475.
- [2] An, G, Fitzpatrick, B., Christley, S, Federico, P, Kanarek, A, Neilan, R. M., Oremland, M, Salinas, R, Laubenbacher, R. & Lenhart, S. “Optimization and control of agent-based models in biology: a perspective”. *Bulletin of Mathematical Biology* **79.1** (2017), pp. 63–87.
- [3] Anderson, A. & Rejniak, K. *Single-Cell-Based Models in Biology and Medicine*. Springer Science & Business Media, 2007.
- [4] Ben Salem, M., Roustant, O., Gamboa, F. & Tomaso, L. “Universal prediction distribution for surrogate models”. *SIAM/ASA Journal on Uncertainty Quantification* **5.1** (2017), pp. 1086–1109.
- [5] Dallon, J., Sherratt, J., Maini, P. & Ferguson, M. “Biological implications of a discrete mathematical model for collagen deposition and alignment in dermal wound repair”. *Mathematical Medicine and Biology: A Journal of the IMA* **17.4** (2000), pp. 379–393.
- [6] Drasdo, D., Kree, R & McCaskill, J. “Monte Carlo approach to tissue-cell populations”. *Physical Review E* **52.6** (1995), p. 6635.
- [7] Erickson, I. E., Huang, A. H., Sengupta, S., Kestle, S., Burdick, J. A. & Mauck, R. L. “Macromer density influences mesenchymal stem cell chondrogenesis and maturation in photocrosslinked hyaluronic acid hydrogels”. *Osteoarthritis and Cartilage* **17.12** (2009), pp. 1639–1648.
- [8] Feder, T. J., Brust-Mascher, I., Slattery, J. P., Baird, B. & Webb, W. W. “Constrained diffusion or immobile fraction on cell surfaces: a new interpretation”. *Biophysical journal* **70.6** (1996), pp. 2767–2773.
- [9] Fernandes, H., Moroni, L., Van Blitterswijk, C. & De Boer, J. “Extracellular matrix and tissue engineering applications”. *Journal of Materials Chemistry* **19.31** (2009), pp. 5474–5484.
- [10] Fish, J. *Multiscale Methods: Bridging the Scales in Science and Engineering*. Oxford University Press on Demand, 2010.
- [11] Freyria, A.-M. & Mallein-Gerin, F. “Chondrocytes or adult stem cells for cartilage repair: the indisputable role of growth factors”. *Injury* **43.3** (2012), pp. 259–265.
- [12] Glazier, J. A. & Graner, F. “Simulation of the differential adhesion driven rearrangement of biological cells”. *Physical Review E* **47.3** (1993), p. 2128.



- [13] Guo, Z., Shi, X., Chen, Y., Chen, H., Peng, X. & Harrison, P. "Mechanical modeling of incompressible particle-reinforced neo-Hookean composites based on numerical homogenization". *Mechanics of Materials* **70** (2014), pp. 1–17.
- [14] Halloran, J. P., Sibole, S. C. & Erdemir, A. "The potential for intercellular mechanical interaction: simulations of single chondrocyte versus anatomically based distribution". *Biomechanics and Modeling in Mechanobiology* **17.1** (2018), pp. 159–168.
- [15] Holmes, M. H. *Introduction to the Foundations of Applied Mathematics*. Springer, 2009.
- [16] Jamali, Y., Jamali, T. & Mofrad, M. R. K. "An agent based model of integrin clustering: Exploring the role of ligand clustering, integrin homo-oligomerization, integrin-ligand affinity, membrane crowdedness and ligand mobility." *Journal of Computational Physics* **244** (2013), pp. 264–278.
- [17] Jiang, Y., Pjesivac-Grbovic, J., Cantrell, C. & Freyer, J. P. "A multiscale model for avascular tumor growth". *Biophysical journal* **89.6** (2005), pp. 3884–3894.
- [18] Johnson, N. L., Kotz, S. & Balakrishnan, N. *Continuous Univariate Distributions*. Wiley New York, 1994.
- [19] Kanit, T., Forest, S., Galliet, I., Mounoury, V. & Jeulin, D. "Determination of the size of the representative volume element for random composites: statistical and numerical approach". *International Journal of Solids and Structure* **40** (2003), pp. 3647–3679.
- [20] Keeney, M., Lai, J. H. & Yang, F. "Recent progress in cartilage tissue engineering". *Current Opinion in Biotechnology* **22.5** (2011), pp. 734–740.
- [21] Kock, L., Donkelaar, C. C. van & Ito, K. "Tissue engineering of functional articular cartilage: the current status". *Cell and Tissue Research* **347.3** (2012), pp. 613–627.
- [22] Masaro, L & Zhu, X. "Physical models of diffusion for polymer solutions, gels and solids". *Progress in Polymer Science* **24.5** (1999), pp. 731–775.
- [23] Matsiko, A., Gleeson, J. P. & O'Brien, F. J. "Scaffold mean pore size influences mesenchymal stem cell chondrogenic differentiation and matrix deposition". *Tissue Engineering Part A* **21.3-4** (2015), pp. 486–497.
- [24] Olson, S. D. & Haider, M. A. "A computational reaction–diffusion model for biosynthesis and linking of cartilage extracellular matrix in cell-seeded scaffolds with varying porosity". *Biomechanics and Modeling in Mechanobiology* **18.3** (2019), pp. 701–716.
- [25] Poleszczuk, J., Macklin, P. & Enderling, H. "Agent-based modeling of cancer stem cell driven solid tumor growth." *Methods in Molecular Biology* **1516** (2016), pp. 335–346.
- [26] Privault, N. *Understanding Markov Chains: Examples and Applications*. Springer, NY, 2013.

- [27] Reddi, A. H., Becerra, J. & Andrades, J. A. "Nanomaterials and hydrogel scaffolds for articular cartilage regeneration". *Tissue Engineering Part B: Reviews* **17.5** (2011), pp. 301–305.
- [28] Rejniak, K. A. "A single-cell approach in modeling the dynamics of tumor microregions". *Mathematical Biosciences & Engineering* **2.3** (2005), p. 643.
- [29] Savill, N. J. & Hogeweg, P. "Modelling morphogenesis: from single cells to crawling slugs". *Journal of Theoretical Biology* **184.3** (1997), pp. 229–235.
- [30] Savill, N. J. & Sherratt, J. A. "Control of epidermal stem cell clusters by Notch-mediated lateral induction". *Developmental Biology* **258.1** (2003), pp. 141–153.
- [31] Saxton, M. J. "Single-particle tracking: the distribution of diffusion coefficients". *Biophysical journal* **72.4** (1997), pp. 1744–1753.
- [32] Saxton, M. "Anomalous Diffusion Due to Obstacles: A Monte Carlo Study". *Biophysical Journal* **66** (1994), pp. 394–401.
- [33] Smith, R. C. *Uncertainty Quantification: Theory, Implementation, and Applications*. Vol. 12. SIAM, 2013.
- [34] Stott, E. L., Britton, N. F., Glazier, J. A. & Zajac, M. "Stochastic simulation of benign avascular tumour growth using the Potts model". *Mathematical and Computer Modelling* **30.5-6** (1999), pp. 183–198.
- [35] Vilaseca, E., Isvoran, A., Madurga, S., Pastor, I., Garcés, J. L. & Mas, F. "New insights into diffusion in 3D crowded media by Monte Carlo simulations: effect of size, mobility and spatial distribution of obstacles". *Physical Chemistry Chemical Physics* **13.16** (2011), pp. 7396–7407.
- [36] Wong, M & Carter, D. "Articular cartilage functional histomorphology and mechanobiology: a research perspective". *Bone* **33.1** (2003), pp. 1–13.