Article

Bridging Continuous and Lattice-Based Models of Two-Dimensional Diffusion: A Systematic Approach for Estimating Transition Probabilities, Grid Size and Diffusivity

Tricity M. Andrew †,‡ and Mansoor A. Haider *‡

Department of Mathematics, North Carolina State University, Raleigh, NC 27695, USA; tricity.andrew@seattlechildrens.org
* Correspondence: mahaider@ncsu.edu
† Current address: Center for Developmental Biology and Regenerative Medicine, Seattle Children’s Research Institute, Seattle, WA 98101, USA.
‡ These authors contributed equally to this work.

Abstract: Lattice-based models have been broadly applied in mathematical and computational modeling of biological and biomedical systems for which spatial effects are important. These discrete models commonly include diffusion of mobile constituents as a key underlying mechanism. While the direct simulation of diffusion in continuous (off-lattice) domains is possible, it is computationally intensive, particularly when multiple coupled mechanisms are involved. This study presents a systematic approach for connecting continuous models of two-dimensional diffusion with internal obstacles to discrete, lattice-based (surrogate) models of diffusion. Results from continuous model simulations on a representative domain, and over many realizations, are used to develop accurate lattice-based surrogate models by exploiting internal symmetries. Probabilities determined for the lattice-based surrogate models are also connected to theoretical diffusivities for 2D random walks on a square lattice, necessitating the calibration of a spatial grid size. This approach can facilitate the inclusion of more accurate diffusive transport models of complex media within the general framework of lattice-based models that incorporate multiple coupled mechanisms.

Keywords: diffusion; mathematical model; random walk; lattice model; surrogate model; simulation

1. Introduction

Lattice-based models have seen broad application in the study of a variety of mechanisms and dynamics through simulation in environments where incorporation of spatial effects is important. For example, they have been commonly used for modeling biotransport, cell dynamics and tissue dynamics with applications in many biological and biomedical areas [1–6]. One example is the cellular Potts model, in which a biological cell is defined over a region comprising multiple lattice sites [2,5,6]. This lattice-based modeling framework enables the study of coupled mechanisms in biological cell dynamics and has been applied to various biomedical problems, including epidermal biology, cancer, angiogenesis and vasculogenesis [1,4]. Such discrete, lattice-based models commonly involve diffusion (e.g., of mobile solutes) as a key transport mechanism. Alternatively, agent-based models [7–9] often restrict movement to a lattice, with random walks commonly used as a simple or convenient approach to advance individual agent locations at each time increment. In such cases, diffusion of nutrients and cell-synthesized solutes is an important coupled mechanism.

As computing power has advanced, direct simulation in continuous-domain (off-lattice) models is a useful tool for investigating systems where particles or agents interact with a background medium or with each other. However, the direct approach can become tedious or computationally intensive in situations involving several coupled mechanisms or complex geometry or in the presence of many particles or agents. The ability to systematically connect off-lattice (continuous) models to lattice-based (discrete) models, accounting
for a few key underlying model features or mechanisms, has the potential to improve the overall accuracy of discrete models. Moreover, such discrete models can also be connected to estimates of theoretical diffusivities for random walks on a two-dimensional (2D) square lattice. An appropriate spatial length scale (grid size) must be determined for the discrete model, especially when diffusion is one of several coupled mechanisms in a 2D lattice-based model.

In this study, we evaluate the extent to which 2D lattice-based surrogate models can preserve measures of particle diffusivity observed in direct simulation of continuous models based on 2D random walks with internal obstacles. Our overall approach is illustrated in Figure 1. Results from continuous model simulations (Figure 1A) on a 2D representative volume element (RVE), and over many realizations, are used to develop lattice-based surrogate models (Figure 1B) that reproduce key features of the continuous model. This is accomplished by exploiting internal symmetries in the continuous model representation. Probabilities in the lattice-based surrogate models are also connected to theoretical diffusivities for 2D random walks on a square lattice (Figure 1C). This final step requires calibration of the discrete length scale (grid size) in the square lattice surrogate model.

**Figure 1.** Illustration of the approach used for bridging continuous and lattice-based models in this study.

The random walker is prescribed a fixed particle radius, and obstacle radii are varied in the context of two different obstacle configurations, each having distinct internal symmetries. Particle dynamics observed for the continuous model simulations are used to develop probabilistic lattice-based surrogate models for transitions across RVE subdomains. The subdomains are chosen to preserve internal symmetries for each of the two obstacle arrangements. A primary advantage of this approach is the capability of the surrogate models to bypass the need to directly represent the obstacles in the simulation domain.

The well-known linear relationship between the mean squared particle displacement and time for the continuous model \([10–12]\) is used to estimate and compare diffusivities for both the continuous and lattice-based surrogate models. This comparison motivates the introduction of a simulation parameter (commitment index) that is calibrated to minimize adverse effects of noise on achieving stationarity of transition probabilities in the lattice-based surrogate model. Theoretical derivations for isotropic and anisotropic diffusion on a 2D lattice are used to connect the transition probabilities in the lattice-based surrogate models to theoretical estimates of the diffusivity for a square lattice surrogate model. Relations enforcing consistency between diffusivity estimates, based on continuous model simulations and analytical results on the lattice, are used to estimate the effective spacing (grid size) in a (square) lattice-based model for both obstacle configurations.

This paper is organized as follows. In Section 2, effective diffusivity is first outlined in terms of the mean squared displacement of a single particle undergoing a 2D random walk in a continuous domain or on a 2D lattice. Section 3 presents our equally spaced obstacles model (Section 3.2) and our multisize obstacles model (Section 3.3). Within each model, quantities are defined to determine probabilities connecting the continuous model with its lattice-based surrogate model. In Section 4, methods for calibrating the lattice-based surrogate models based on the continuous model simulation results are first presented...
(Section 4.1). A framework is then presented to connect the lattice-based surrogate models to square lattice surrogate models (Section 4.2), necessitating the calibration of the square lattice grid size parameter. Our results are presented for both models in Section 5, including the calculation of all model probabilities and calibration of other model parameters, followed by a brief discussion and conclusions in Section 6. Supplementary derivations are presented in Appendices A and B.

2. Effective Diffusivity

We first summarize approaches for estimating the effective diffusivity of a particle undergoing 2D random walks based on observations of its mean squared displacement.

2.1. Effective Diffusivity in a Continuous Domain

Normal diffusion describes diffusion via random walks of a particle in a continuous domain. The mean-squared displacement $\langle r^2 \rangle$, relative to a fixed starting location at $t = 0$, is calculated by averaging displacement trajectories over a large number of realizations. In prior studies of 2D single particle random walks, it was shown that $\langle r^2 \rangle$ grows linearly with time ($t$) after a sufficiently large number of time steps have passed [10–12]. The general formula for normal diffusion is written in terms of an effective diffusivity ($\kappa$) as [11–13]

$$\langle r^2 \rangle = (2d) \kappa t,$$

where $d$ is the dimension of the medium. In our study, all simulation results are sampled in a stationary time regime where this observed relationship is linear; the effective diffusivity is estimated by rearranging (1) to obtain

$$\kappa = \frac{\langle r^2 \rangle}{(2d) N \Delta t},$$

where $N$ is the number of time steps (of size $\Delta t$) in the simulation.

2.2. Effective Diffusivity on a 2D Lattice

To develop our lattice-based surrogate models, we estimate the diffusivity analogous to (2) but based on analyzing $\langle r^2 \rangle$ on the lattice. In this section, we demonstrate, via an analytical derivation, that a particle undergoing random walks on a 2D lattice exhibits normal diffusion, to leading order. Overall, our derivations of mean squared displacements and theoretical diffusivities for random walks on 2D lattices are based on analytical approaches outlined in Holmes [14].

We can track $\langle r^2 \rangle$ systematically through a set of transition probabilities (Figure 2). Here, $p_x$ and $p_y$ denote the probability of moving, at the next time step, to adjacent horizontal (by $\Delta x$) or vertical (by $\Delta y$) lattice sites, respectively. The particle remains in its current location with probability $1 - p_x - p_y$. In this section, we denote the mean squared displacement as $\langle r^2_k(N) \rangle$, with $K$ referring to the number of realizations of a single particle random walk on the 2D lattice and $N$ denoting a particular time step.
Figure 2. Illustration of a particle undergoing a random walk on a 2D lattice. The particle moves one step (size $\Delta x$) in the horizontal direction with probability $p_x$ (left) and one step (size $\Delta y$) in the vertical direction with probability $p_y$ (right). The probability of staying in place is $1 - p_x - p_y$.

2.2.1. Lattice Representation of Mean Squared Displacement

The mean squared displacement after $N$ time steps and for $K$ realizations is

$$\langle r^2 \rangle_K(N) = \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(N) + y_i^2(N), \quad (3)$$

where $x_i(N)$ and $y_i(N)$ are the horizontal and vertical particle locations, respectively. The mean squared displacement is then $\langle r^2 \rangle = \lim_{K \to \infty} \langle r^2 \rangle_K(N)$.

For brevity, let $x_i^N$ and $y_i^N$ denote the horizontal and vertical positions of the $i^{th}$ particle at time step $N$, respectively, (the full notation is utilized in the Appendices). We write these positions as

$$x_i^N = x_i^{N-1} + h_{xi}^N, \quad y_i^N = y_i^{N-1} + h_{yi}^N, \quad i = 1, \ldots, K, \quad (4)$$

where $h_{xi}^N$ and $h_{yi}^N$ are the possible displacements with associated step sizes of $\Delta x$ and $\Delta y$, respectively. In any one time step $\Delta t$, the five possible transitions are

$$\begin{align*}
(i) & \quad h_{xi}^N = -\Delta x, h_{yi}^N = 0, \\
(ii) & \quad h_{xi}^N = 0, h_{yi}^N = \pm \Delta y, \\
(iii) & \quad h_{xi}^N = h_{yi}^N = 0,
\end{align*} \quad (5)$$

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

Substituting (4) into (3), $\langle r^2 \rangle$ can be written in recursive form as

$$\langle r^2 \rangle = \lim_{K \to \infty} \langle r^2 \rangle_K(N) = \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} \left( x_i^{N-1} + h_{xi}^N \right)^2 + \left( y_i^{N-1} + h_{yi}^N \right)^2,$$

$$= \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} \left[ \left( x_i^{N-1} \right)^2 + \left( y_i^{N-1} \right)^2 + h_{xi}^N \right]^2 + \left[ \left( y_i^{N-1} \right)^2 + \left( y_i^{N-1} \right)^2 + h_{yi}^N \right]^2,$$

$$= \lim_{K \to \infty} \langle r^2 \rangle_{K-1}(N) + \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} \left[ x_i^{N-1} h_{xi}^N + \left( h_{xi}^N \right)^2 + y_i^{N-1} h_{yi}^N + \left( h_{yi}^N \right)^2 \right]. \quad (6)$$

In the next subsection, we briefly discuss limits used to simplify (6).
2.2.2. Useful Limits for Simplifying Equation (6)

We use two main limit results in simplifying Equation (6). Direct evaluation of these limits is summarized in Appendices A and B, resulting in the following values associated with case (i) (horizontal transition) above:

\[
\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} (h_{xi}^N)^2 = p_x (\Delta x)^2, \quad \lim_{K \to \infty} \frac{2}{K} \sum_{i=1}^{K} x_i^N h_{xi}^M = 0 \text{ when } M > N. \quad (7)
\]

Analogous results also hold for the vertical transition case (ii).

2.2.3. Reduced Form of Mean Squared Displacement

Using (7), we rewrite (6) as

\[
\lim_{K \to \infty} \langle r_{K}^2(N) \rangle = \lim_{K \to \infty} \langle r_{K-1}^2(N) \rangle + \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^{N-1} h_{xi}^N + \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} y_i^{N-1} h_{yi}^N
\]

\[
+ \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} (h_{xi}^N)^2 + \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} (h_{yi}^N)^2, \quad (8)
\]

to simplify (8) to the relation,

\[
\lim_{K \to \infty} \langle r_{K}^2(N) \rangle = \lim_{K \to \infty} \langle r_{K-1}^2(N) \rangle + p_x (\Delta x)^2 + p_y (\Delta y)^2. \quad (9)
\]

Lastly, the recurrence property enables (9) to be simplified to

\[
\langle r^2 \rangle = \lim_{K \to \infty} \langle r_{K}^2(N) \rangle = N p_x (\Delta x)^2 + N p_y (\Delta y)^2. \quad (10)
\]

On a square lattice ($\Delta x = \Delta y$), (10) can be reduced to the linear relation

\[
\langle r^2 \rangle = \lim_{K \to \infty} \langle r_{K}^2(N) \rangle = N(p_x + p_y)(\Delta x)^2 = N(1 - q)(\Delta x)^2, \quad (11)
\]

where we denote $q$ as the probability of staying in the current subdomain, i.e., since $q + p_x + p_y = 1$. Equation (11) will serve as the basis for connecting the square lattice surrogate models with the lattice-based surrogate models obtained from the continuous model simulations (see Section 4.3).

3. Models

We consider direct simulation of 2D random walks on a continuous domain. Results from these simulations are used to determine transition probabilities in the lattice-based surrogate model. The number of such independent probabilities, as well as their values, change for different obstacle arrangements. We consider a single particle undergoing random walks in a 2D representative volume element (RVE). Within the RVE, we place a small number of rigid obstacles with which the particle collides; the collisions are idealized to be perfectly elastic with the angle of reflection equal to the angle of incidence. The RVE is assumed to have periodic boundary conditions, further capturing internal symmetries in the background medium. Two different obstacle arrangements are considered (Figure 3): (i) equally spaced obstacles (Figure 3a) and (ii) multisize obstacles (Figure 3b). While both cases model inhomogeneous, anisotropic materials, case (i) exhibits a greater degree of (rotational) symmetry as compared to case (ii). The second case also serves as a simple analogue of a composite with two different obstacle sizes.
Figure 3. Sample domains (RVEs) for continuous model simulations (\( R_{\text{part}} = 1.0 \)), with numbered subdomains: (a) equally spaced obstacles model (with \( R_{\text{obs}} = 2.0 \)); (b) multisize obstacles model (with \( R_{\text{obs,1}} = 1.0 \) and \( R_{\text{obs,2}} = 2.0 \)). Obstacles are shown as black circles and the particle undergoing random walks is shown as a red circle.

3.1. Continuous Model Simulation Details

We consider symmetries inherent to each case that are evident when the RVE is partitioned into nine equally sized subdomains for case (i) (Figure 3a) and four equally sized subdomains in case (ii) (Figure 3b). Each (square) subdomain is prescribed a length \( L = 10 \); thus, to capture internal symmetries, the RVE is \( 30 \times 30 \) for case (i) and \( 20 \times 20 \) for case (ii). The state of the system at time \( t \) is the subdomain in which the particle is located (1–9 in case (i) and 1–4 in case (ii), Figure 3). State changes, which are vertical or horizontal transitions between neighboring subdomains, are tracked to determine transition probabilities for this surrogate model (see Sections 3.2.2 and 3.3.2).

Additional assumptions are as follows: (a) our single particle has prescribed radius \( R_{\text{part}} \) and moves with a speed \( v = 8 \); (b) at \( t = 0 \), the particle is placed at the center of subdomain 5 for the equally spaced obstacles model and at the center of subdomain 3 for the multisize obstacles model; (c) at each time step of prescribed temporal step size \( \Delta t = 0.05 \), a new direction is chosen, via uniform sampling, from \( \theta \in [0, 2\pi] \), and the particle is advanced a distance of \( \Delta x = v\Delta t \) in this direction; (d) collisions with the obstacles are perfectly elastic, i.e., the particle bounces off the obstacle at the (reflected) incidence angle relative to the local tangent line at the point of collision.

We simulate many realizations over \( N_t = 1 \times 10^7 \) time steps to generate plots of the mean-squared displacement \( \langle r^2 \rangle \) versus time. For all simulations considered in this study, 2000 realizations on a high-performance cluster computer were sufficient for the slopes of these linear curves to reach stationary values.

3.2. Equally Spaced Obstacles Model

3.2.1. Continuous Model

For the equally spaced obstacles model, the circular obstacles (radius \( R_{\text{obs}} \)) are centered on the subdomain intersection points. The particle radius is fixed at \( R_{\text{part}} = 1.0 \), and the obstacle radius is varied from 0.5 to 3.0 by increments of 0.5. This range was chosen to enable particle movement between obstacles; an obstacle radius greater than 3.0 inhibits transitions between subdomains.

3.2.2. Lattice-Based Surrogate Model

Figure 4 illustrates the construction of the 2D lattice-based surrogate model for the equally spaced obstacles case (Figure 3a). A periodic extension of the RVE is shown, with a primary transition from the center subdomain illustrated. Due to extensive symmetries in this case, all internal transitions across boundary walls between subdomains are equivalent. The probability of diagonal transitions is zero due to the presence of obstacles at the subdomain corners. Hence, the lattice model has only two distinct transition probabilities, that the particle (1) stays in its current subdomain (\( p^1_1 \)) or (2) moves to a neighboring subdomain in a cardinal direction (\( p^1_2 \)). The corresponding state transitions are illustrated.
for subdomain 5 in Figure 4b. Since transition probabilities within a subdomain must sum to 1, only one surrogate model probability needs to be calculated from the results of continuous model simulations.

**Figure 4.** Building a lattice-based surrogate model for the continuous model in the equally spaced obstacles case: (a) nine subdomains are used to exploit symmetry of the obstacle arrangement and transitions between subdomains are tracked. (b) Lattice-based surrogate representation is shown, with red circles denoting corresponding subdomains in the continuous model.

Results of the continuous model simulations (Figure 4a) are tracked via the time course \( \eta(t_k) \) \( (t_k = k\Delta t, \ k = 0, \ldots, N) \), where \( N \) is the total number of time steps. \( \eta(t_k) \) records the state of a particle as the index of the \( i \)th subdomain \( (i = 1, \ldots, 9) \) at time \( t_k \) if its center is located in that subdomain at time \( t_k \). A \( 9 \times 9 \) matrix \( S^{(k)} \) is used to track all transitions (including remaining in the same subdomain) up to the current time step \( (S^{(0)} \) is the zero matrix). \( S^{(k)} \) is updated according to the formula

\[
S^{(k)}(\eta(t_{k-1}), \eta(t_k)) = S^{(k-1)}(\eta(t_{k-1}), \eta(t_k)) + 1, \ k = 1, \ldots, N. \tag{12}
\]

\( S^{(k)} \) is then used to build a vector \( \vec{p}^{E} \) of transition probabilities. Within each time step, a vector \( \vec{T}^{E}_i \) \( (i = 1, \ldots, 9) \) is created to count transitions that occurred for each subdomain. In particular, the rows of \( S^{(k)} \) are used to create vector \( \vec{T}^{E}_i \) by grouping transition sums by transition type; since the equally spaced obstacles model has two transition types, \( \vec{T}^{E}_i \) has dimensions of \( 1 \times 2 \). At the current time \( t_k \), we construct \( \vec{T}^{E}_i \) from \( S^{(k)} \) as follows,

\[
\vec{T}^{E}_i = \left[ S^{(k)}(i,i), \sum_{\{j \neq i\}} S^{(k)}(i,j) \right], \ i = 1, \ldots, 9. \tag{13}
\]

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

\[
\vec{p}^{E} = \frac{1}{N} \sum_{i=1}^{N} \vec{T}^{E}_i. \tag{14}
\]

### 3.3. Multisize Obstacles Model

#### 3.3.1. Continuous Model

For the multisize obstacles model, the RVE has one set of obstacles with a fixed obstacle radius \( R_{\text{obs}1} \) and a second set of obstacles of radius \( R_{\text{obs}2} \) that is varied across cases. Symmetry for this model is preserved by partitioning the RVE using a \( 2 \times 2 \) grid of subdomains (Figure 3b). We view the domain as having “columns” of obstacles; obstacles in the same column have identical radii. The particle radius and first obstacle radius are
fixed at values of \( R_{\text{part}} = 1.0 \) and \( R_{\text{obs}1} = 1.0 \), respectively. The second obstacle radius \( R_{\text{obs}2} \) is then varied between 0.5 to 3.0 in increments of 0.5.

### 3.3.2. Lattice-Based Surrogate Model

Figure 5 illustrates the construction of the 2D lattice-based surrogate model for the multisize obstacles case. A periodic extension of the RVE is shown, with primary transitions from subdomain 1 (top-left corner) illustrated. Due to inherent symmetries, all subdomains have the same number of distinct transition types. There are four types of transition with corresponding probabilities (Figure 5b). In type 1, the particle stays in its current subdomain \( (p_{M1}) \). In types 2–4, it transitions to a neighboring subdomain (2,3 or 4) having two obstacles. We delineate the three latter types into probabilities as follows: the two obstacles are of different sizes \( (R_{\text{obs}1}, R_{\text{obs}2}) \) along the boundary \( (p_{M2}) \); the two obstacles are the same size \( R_{\text{obs}2} \) along the boundary \( (p_{M3}) \); or the two obstacles are the same size \( R_{\text{obs}1} \) along the boundary \( (p_{M4}) \). Again, the probabilities must sum to one; the three surrogate model probabilities need to be calculated from the results of the continuous model simulations.

![Figure 5](image)

**Figure 5.** Building a lattice-based surrogate model for the continuous model in the multisize obstacles case: (a) four subdomains are used to exploit symmetries of the obstacle arrangement, and transitions between subdomains are tracked. Smaller green circles indicate obstacles of fixed radius \( R_{\text{obs}1} = 1.0 \), and larger blue circles indicate obstacles with varying obstacle radius \( R_{\text{obs}2} \); (b) lattice-based surrogate representation with red circles denoting corresponding subdomains in the continuous model.

Again, the time course \( \eta(t_k) \) records the state of the particle as the index of the \( i \)th subdomain \( (i = 1, ..., 4) \) at time \( t_k \) if its center is located in that subdomain at time \( t_k \). For the multisize obstacles case, we differentiate between a horizontal transition within the RVE versus a horizontal transition into an extended domain by introducing an additional time course \( \zeta(t_k) \) \( (t_k = k\Delta t, k = 0, ..., N) \), where

\[
\zeta(t_k) = \begin{cases} 
1 & \text{if a horizontal transition occurred into an extended domain} \\
0 & \text{otherwise (includes all vertical transitions).}
\end{cases}
\]  

(15)

In this case, the transition matrix \( S^{(k)} \) (for time steps \( k = 1, ..., N \)) is a \( 4 \times 5 \) (augmented) matrix; the fifth column tracks horizontal transitions occurring when the particle moves into an extended domain. For a large majority of time steps, \( \zeta(t_k) = 0 \) and \( S^{(k)} \) are updated in a similar manner as in the equally spaced obstacles case via Equation (12) (Section 3.2.2). Alternatively, when \( \zeta(t_k) = 1 \), only the new column in the augmented matrix is updated via the formula

\[
S^{(k)}(\eta(t_{k-1}), 5) = S^{(k)}(\eta(t_{k-1}), 5) + 1, \quad k = 1, ..., N.
\]  

(16)

and \( S^{(k)} \) is then used to build the vector \( p^M \) of five transition probabilities.

Within each time step and for each of the four subdomains, a vector \( \vec{T}^M_i \) \( (i = 1, 2, 3, 4) \) is created to count transitions, according to the following rules: (i) the first component counts
the number of times a particle stayed in its current subdomain; (ii) the second component counts the number of times a particle moved vertically; (iii) the third component counts the number of times a particle moved horizontally but remained in the original RVE; (iv) the fourth component counts the number of times a particle moved horizontally but moved to the extended domain.

Due to the symmetries in this case, we create the transition probability vector $\vec{p}^M$ via the average:

$$\vec{p}^M = \frac{1}{N} \sum_{j=1}^{N} \vec{T}^M_j.$$

(17)

4. Model Calibration and Square Lattice Surrogate Models

4.1. Commitment Index

For random walks in our continuous model, a particle may cross, return and then recross a subdomain boundary several times before “committing” to a new subdomain. This leads to inaccuracy in the calculation of transition probabilities as stationary quantities in the lattice-based surrogate models. Consequently, we introduce a commitment index, $M_C$, to indicate the number of successive time steps after which a particle is deemed to commit to a new subdomain. The role of this parameter is illustrated in Figure 6 for the case $M_C = 4$.

While the value of $M_C$ requires calibration via direct comparison of simulated results for mean-squared displacement in the continuous and lattice-based surrogate models, we now explain its narrow range of variation.

Figure 6. Illustration of the role of the commitment index (calibration) parameter $M_C$: (a) The particle (solid black circle) has just crossed the boundary between subdomains (dashed gray line) to reach its 1st-step location. Possible 2nd-step locations are indicated by the solid black ring. (b) Three possible 2nd-step locations are indicated by solid red circles. Some possible next steps (3rd-step locations) are shown with red rings. (c) A red ring illustrates the particle’s remaining possible 3rd-step locations when it has remained in the current subdomain at the prior step. (d) Possible 4th-step locations are indicated with blue rings.

(i) Consider a particle, committed to a subdomain, that has just crossed the boundary into a neighboring subdomain (first-step location). A (black) ring indicates all possible particle locations at the next step, including some locations that result in recrossing the same boundary (Figure 6a).

(ii) We next illustrate a few possible steps relative to the particle’s first-step location, thus indicating possible second-step locations. A (red) ring of possible 3rd-step locations is shown around each such second-step location (Figure 6b).

(iii) Choosing one of the possible second-step locations that is in the new subdomain (gray ring), the resulting set of possible third-step locations is shown in red (Figure 6c).

(iv) Lastly, the larger the number of successive steps for which a particle remains in the new subdomain, the smaller the probability that it will recross back into the
subdomain to which it was previously committed; in Figure 6d, blue rings depict fourth-step locations.

For each of our two obstacle arrangements, \( M_C \) is calibrated by constructing transition probabilities from continuous model simulations over a range of \( M_C \) values and by comparing mean-squared displacement for the continuous and lattice-based surrogate models to identify the case with the best agreement (see Section 5.1).

4.2. Diffusivities for the Square Lattice Surrogate Models

Recall from Section 2 the following estimates of the effective diffusivity (e.g., (2) with \( d = 2 \)) for the continuous model and the surrogate model, respectively,

\[
\kappa_{\text{cont}} = \frac{\langle r_{\text{cont}}^2 \rangle}{4N\Delta t}, \quad \text{and} \quad \kappa_{\text{surr}} = \frac{\langle r_{\text{surr}}^2 \rangle}{4N\Delta t},
\]  

(18)

In Section 4.2, we use 2D Taylor series expansions to derive the (continuum) diffusion equation and the associated theoretical diffusivities on a lattice for both obstacle arrangements.

4.2.1. Equally Spaced Obstacles Model

Consider a 2D rectangular lattice with horizontal spacing \( h(=\Delta x) \) and vertical spacing \( k(=\Delta y) \). Assuming a fixed volume and a uniform particle mass, let \( u(x,y,t) \) be the density of particles and let \( q \) be the probability that a particle stays in its current subdomain; \( (1-q) \) is the probability of transitioning to a neighboring lattice point. The resulting particle density is given by

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

(19)

Consider the Taylor series expansions about the point \( P(x,y,t) \),

\[
u(x,y,t+\Delta t) = u + u_t(\Delta t) + \frac{1}{2}u_{tt}(\Delta t)^2 + \frac{1}{6}u_{ttt}(\Delta t)^3 + \ldots,
\]  

(20)

\[
u(x \pm h,y,t) = u \pm ux(\Delta x) + \frac{1}{2}u_{xx}(\Delta x)^2 \pm \frac{1}{6}u_{xxx}(\Delta x)^3 + \ldots,
\]  

(21)

\[
u(x,y \pm k,t) = u \pm uy(\Delta y) + \frac{1}{2}u_{yy}(\Delta y)^2 \pm \frac{1}{6}u_{yyy}(\Delta y)^3 + \ldots.
\]  

(22)

Using these series expansions, (19) becomes

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

(23)

Assuming a 2D square lattice \((\Delta x = \Delta y)\) and dividing by \( \Delta t \), we can rewrite (23) as

\[
u_t = O(\Delta t) = \frac{(1-q)}{4}\left(\frac{u_{xx}}{\Delta t} + \frac{u_{yy}}{(\Delta t)} + O\left(\frac{(\Delta x)^4}{(\Delta t)}\right)\right),
\]  

(24)

resulting in the approximate equation (assuming \((1-q)(\Delta x)^4 \ll \Delta t\),

\[
u_t \approx \frac{(1-q)(\Delta x)^2}{4(\Delta t)}(u_{xx} + u_{yy}).
\]  

(25)
Recognizing (25) as the 2D diffusion equation, the theoretical diffusivity for the equal obstacles model \(D\) is given in terms of square lattice model properties by

\[
D = \frac{(1 - q)(\Delta x)^2}{4(\Delta t)}.
\]  

(26)

4.2.2. Multisize Obstacles Model

In the multisize obstacles model, the particle has different diffusivities in the vertical and horizontal directions. Yet, the large majority of lattice-based models in computational biology employ a square lattice. Thus, while our derivation below allows for distinct values of \(\Delta x\) and \(\Delta y\), we ultimately set these quantities to be equal in order to reconcile the two viewpoints.

We follow the same procedure as in Section 4.2.1, but for this case, note the following: (i) \(q_1\) is the probability that the particle stays at its current lattice point; (ii) \(q_x\) is the probability that the particle moves to a neighboring lattice point in the horizontal direction; (iii) \(q_y\) is the probability that the particle moves to a neighboring lattice point in the vertical direction. Note that \(q_1 + q_x + q_y = 1\). We can then write the following relation for the particle density:

\[
u(x,y,t + \Delta t) = q_1 u(x,y,t) + \frac{q_x}{2} u(x + h,y,t) + \frac{q_y}{2} u(x - h,y,t)
\]

\[+ \frac{q_y}{2} u(x,y + k,t) + \frac{q_y}{2} u(x,y - k,t).
\]  

(27)

Via the Taylor series expansions in (20)–(22), the above equation can be expanded and simplified to

\[
u + \nu_t(\Delta t) + O((\Delta t)^2) = q_1 u + q_x u + q_y u + \frac{q_x}{2} \left( u_{xx}(\Delta x)^2 + O((\Delta x)^4) \right)
\]

\[+ \frac{q_y}{2} \left( u_{yy}(\Delta y)^2 + O((\Delta y)^4) \right).
\]  

(28)

Since \(q_1 + q_x + q_y = 1\), upon dividing by \(\Delta t\) we can rewrite (28) as

\[
u_t + O(\Delta t) = \frac{q_x}{2} u_{xx}(\Delta x)^2 + \frac{q_y}{2} u_{yy}(\Delta y)^2 + \frac{q_x}{2} O \left( \frac{(\Delta x)^4}{(\Delta t)} \right) + \frac{q_y}{2} O \left( \frac{(\Delta y)^4}{(\Delta t)} \right).
\]  

(29)

Imposing our restriction to a uniform 2D lattice (\(\Delta x = \Delta y\)), the simplified diffusion equation for the multisize obstacles case is thus (assuming \(q_x(\Delta x)^4 \ll \Delta t\) and \(q_y(\Delta y)^4 \ll \Delta t\))

\[
u_t \approx \left( \frac{(\Delta x)^2}{2(\Delta t)} \right) (q_x u_{xx} + q_y u_{yy}).
\]  

(30)

Hence, when a 2D square lattice is used as a surrogate model for the multisize obstacles case, there are two theoretical diffusivities that we refer to as the horizontal diffusivity \((D_H)\) and the vertical diffusivity \((D_V)\), which are given by

\[
D_H = \frac{q_x(\Delta x)^2}{2(\Delta t)}, \quad D_V = \frac{q_y(\Delta y)^2}{2(\Delta t)}.
\]  

(31)

We note that in the case \(q_x = q_y\), the diffusion equation for the multisize obstacles model (30) reduces to the diffusion equation for the equally spaced obstacles model (25).

4.3. Relating the Diffusivities

We now relate our effective diffusivities \((\kappa_{\text{cont}}, \kappa_{\text{surr}})\) in (18) to the theoretical diffusivities for each of the two obstacle models, i.e., (26) and (31).
4.3.1. Equally Spaced Obstacles Model

The effective diffusivity for the surrogate model $\kappa_{\text{surr}}$ is first used to approximate the effective diffusivity for the continuous model $\kappa_{\text{cont}}$. Specifically, slopes of the mean squared displacement response are matched by calibrating the commitment index $M_C$ in the surrogate model (see Section 4.1).

Substituting $\langle r^2 \rangle$ from (11) for the surrogate model into the second relation of (18), we obtain the relations,

$$\kappa_{\text{surr}} = \frac{\langle r^2_{\text{surr}} \rangle}{4N\Delta t} \approx \frac{N(1-q)(\Delta x)^2}{4N\Delta t} = \frac{(1-q)(\Delta x)^2}{4\Delta t} = D. \quad (32)$$

Since $\Delta t$ is the same across all representations, it remains to calibrate $\Delta x$, i.e., the effective spatial grid size for the square lattice model with theoretical diffusivity $D$.

Although $L = 10$ is a characteristic length scale for both the continuous and (surrogate) lattice models, it is not an appropriate choice for $\Delta x$ in (32). A particle at an arbitrary location in a subdomain has many different net distances it can travel before crossing into and committing to a neighboring subdomain. We determine the spatial step size via the representation $\Delta x = L^\alpha$, where $\alpha > 1$ will be calibrated based on our numerical simulations.

Substituting $\Delta x = L^\alpha$ into (32), we obtain

$$\kappa_{\text{surr}} \approx D|_{\Delta x=\frac{L}{\alpha}} = \frac{(1-q)(\frac{L}{\alpha})^2}{4\Delta t} = \frac{1}{\alpha^2}D|_{\Delta x=L}. \quad (33)$$

Hence, the ratio of the two diffusivities can be used to approximate $\alpha$ via the formula

$$\alpha \approx \sqrt{\frac{D|_{\Delta x=L}}{\kappa_{\text{surr}}}}. \quad (34)$$

Using this coordinated approach, we ensure that all diffusivity estimates obtained via simulations using the continuous model and the surrogate model are consistent with theoretical estimates for 2D random walks on a lattice based on the choice $\Delta x = \frac{L}{\alpha}$.

4.3.2. Multisize Obstacles Model

Considering the horizontal and vertical effective diffusivities independently, and using (2) with $d = 1$, we can define $\kappa_{H}$ and $\kappa_{V}$ as

$$\kappa_{H} = \frac{\langle r_{H}^2 \rangle}{2N\Delta t}, \quad \kappa_{V} = \frac{\langle r_{V}^2 \rangle}{2N\Delta t}, \quad (35)$$

where $\langle r_{H}^2 \rangle$ and $\langle r_{V}^2 \rangle$ are the horizontal and vertical mean squared displacements. In the multisize obstacles model, we identified horizontal and vertical components of diffusivity in terms of the corresponding probabilities (see (31)).

In contrast to the equally spaced obstacles model, here we do not directly identify the probabilities $q_x$ (of a horizontal transition) and $q_y$ (of a vertical transition) in the square lattice surrogate model. In particular, $q_x$ is not explicitly identified from our simulations, since our results delineate (1) a horizontal transition in the original RVE and (2) a horizontal transition to the extended domain. We, instead, relate the following ratios:

$$\frac{D_{H}}{D_{V}} \approx \frac{\kappa_{H}}{\kappa_{V}} \approx \frac{\langle r_{H}^2 \rangle}{\langle r_{V}^2 \rangle} \approx \frac{q_x}{q_y}. \quad (36)$$

Note that in the case where all obstacles are the same size, this yields $D_{H} \approx D_{V}$.

For the lattice-based surrogate model, the four probabilities $p_1, p_2, p_3, p_4$ correspond to the transitions depicted in Figure 5b. Recall that $q_1 + q_x + q_y = 1$ in the square lattice surrogate model and $p_1 + 2p_2 + p_3 + p_4 = 1$ in the lattice-based surrogate model. Equating
these two relations and assuming that $q_1 = p_1$ and that $2p_2 = q_y$, we obtain the reduced equation

$$q_x = p_3 + p_4.$$  \hfill (37)

Assuming $\kappa_H \approx D_H$ and $\kappa_V \approx D_V$, we can use simulation results to calibrate $\Delta x$ using each of the two relations in (31) (see Section 5.2.2). Moreover, we note that the ratio of diffusivities in (36) is independent of the square lattice grid size.

5. Results

For both the equally spaced obstacles model (Section 3.2) and the multisize obstacles model (Section 3.3), we use results of the continuous model simulations to determine transition probabilities for the two surrogate models, as described in Section 3.2.2 and Section 3.3.2, respectively. As outlined in Section 3.1, we simulate a single particle (radius $R_{\text{part}}$) undergoing a 2D random walk (speed $v = 8$) with time step $\Delta t = 0.05$. At each step, the direction is sampled from a uniform distribution and all collisions with internal obstacles are perfectly elastic. All continuous model simulations were carried out on a high-performance cluster computer for $N = 2000$ realizations. Once transition probabilities were determined, the corresponding lattice-based surrogate model simulations were carried out on a laptop computer (MacBook Pro) in under 2.5 h. By contrast, the continuous model simulations were completed at a rate of 400 realizations per 27 h, distributed over eight processors (50 realizations per processor), for a specified set of model parameters. A total of $N_t = 1 \times 10^7$ time steps were sufficient for all quantities of interest to exhibit stationary responses.

5.1. Building the Lattice-Based Surrogate Models

5.1.1. Equally Spaced Obstacles Case

Obstacle radii: Continuous model simulations were performed on the RVE for the equally spaced obstacles model (see Section 3.2.1 and Figure 3a). The particle radius was first fixed at $R_{\text{part}} = 1.0$, and the obstacle radius ($R_{\text{obs}}$) was varied from 0.5 to 3.0 by increments of 0.5. Stationarity and linearity of $\langle r^2 \rangle$ versus time (Section 2.1) for an increasing number of realizations is illustrated, for a representative case ($R_{\text{part}} = 1.0$, $R_{\text{obs}} = 1.0$), in Figure 7.

![Figure 7](image)

**Figure 7.** Mean squared displacement versus time for the equally spaced obstacles model. Effects of an increasing number of realizations are shown for the case $R_{\text{part}} = 1.0$ and $R_{\text{obs}} = 1.0$: (a) curves; (b) linear fits to the curves.

Determining transition probabilities: Based on the approach illustrated in Figure 4, continuous model results were used to determine transition probabilities for the equally spaced obstacles model via (14). These probabilities are illustrated for transitions from subdomain 6 over the time course of one realization in Figure 8b–d; rescaling the range more effectively illustrates the stay and leave probabilities (Figure 8c,d, respectively). This figure also demonstrates the stationary nature of these probabilities at later times. In
determining the final stay ($p_E^F(1)$) and leave ($p_E^F(2)$) transition probabilities for the lattice-based surrogate model, values of each type were pooled together (see (14)) and then averaged across all realizations. Note that our results are consistent with the improbability of diagonal transitions (Figure 8d, blue curve).

Figure 8. Transition probabilities for building the lattice-based surrogate model from the equally spaced obstacles (continuous) model. Transitions from subdomain 6 are illustrated: (a) simulation domain showing the 9 subdomains; (b) plot of all transition probabilities (here, only the stay probability is discernible); (c) magnified plot to illustrate the stay probability; (d) magnified plot to illustrate the leave probabilities.

Commitment index calibration: Lattice-based surrogate models were built using the transition probabilities obtained from the continuous model (see (14), Section 3.2.2) by first calibrating the commitment index $M_C$ (Section 4.1). $M_C$ was varied from 3 to 8 and calibrated by comparing the continuous and surrogate plots of $\langle r^2 \rangle$ versus time. This procedure was repeated for varying obstacle radii, with the particle radius fixed at $R_{\text{part}}=1.0$, illustrated in Figure 9 for $R_{\text{part}}=1.0$ and $R_{\text{obs}}=2.0$. Here, $M_C = 5$ yields the best result; 2000 realizations were sufficient to observe stationary values sufficient to calibrate $M_C$ (Figure 9—right). Probabilities and commitment index values are shown in Table 1. We observe that the sum of the first probability and four times the second probability is very close to 1 (based on symmetries in the RVE); the largest discrepancy was observed to be $3.6 \times 10^{-6}$. We also observe that the probability of staying in the current subdomain increases with the obstacle radius. Comparisons of slopes for $\langle r^2 \rangle$ versus time for the continuous and surrogate models are compared in Table 2 and Figure 10. We observe differences in slopes ranging from 0.004% to roughly 4%.
Figure 9. Mean squared displacement and its slope for the equally spaced obstacles model. Results are shown for particle radius \( R_{\text{part}} = 1.0 \) and obstacle radius \( R_{\text{obs}} = 2.0 \): (left) \( \langle r^2 \rangle \) versus time for continuous model and resulting surrogate model as the commitment index \( M_c \) is varied; (right) slope as a function of obstacle radius for an increasing number of realizations.

Table 1. Transition probabilities in the lattice-based surrogate model for the equally spaced obstacles (continuous) model for the best values of \( M_c \) with particle radius \( R_{\text{part}} = 1.0 \).

<table>
<thead>
<tr>
<th>( R_{\text{obs}} )</th>
<th>( M_c )</th>
<th>( p_1^E )</th>
<th>( p_2^E )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>7</td>
<td>0.99276</td>
<td>( 1.8099 \times 10^{-3} )</td>
</tr>
<tr>
<td>1.0</td>
<td>6</td>
<td>0.99291</td>
<td>( 1.7724 \times 10^{-3} )</td>
</tr>
<tr>
<td>1.5</td>
<td>5</td>
<td>0.99297</td>
<td>( 1.7573 \times 10^{-3} )</td>
</tr>
<tr>
<td>2.0</td>
<td>5</td>
<td>0.99371</td>
<td>( 1.5734 \times 10^{-3} )</td>
</tr>
<tr>
<td>2.5</td>
<td>4</td>
<td>0.99384</td>
<td>( 1.5401 \times 10^{-3} )</td>
</tr>
<tr>
<td>3.0</td>
<td>4</td>
<td>0.99493</td>
<td>( 1.2676 \times 10^{-3} )</td>
</tr>
</tbody>
</table>

Table 2. Slope values for \( \langle r^2 \rangle \) vs. \( t \) in the equally spaced obstacles (continuous) model \( (R_{\text{part}} = 1.0) \) and the lattice-based surrogate model. The best values of \( M_c \) are also shown.

<table>
<thead>
<tr>
<th>( R_{\text{obs}} )</th>
<th>( M_c )</th>
<th>Slope (Continuous)</th>
<th>Slope (Surrogate)</th>
<th>Percent Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>7</td>
<td>2.9984</td>
<td>2.8770</td>
<td>4.049%</td>
</tr>
<tr>
<td>1.0</td>
<td>6</td>
<td>2.9044</td>
<td>2.8468</td>
<td>1.983%</td>
</tr>
<tr>
<td>1.5</td>
<td>5</td>
<td>2.7507</td>
<td>2.7506</td>
<td>0.004%</td>
</tr>
<tr>
<td>2.0</td>
<td>5</td>
<td>2.4731</td>
<td>2.5094</td>
<td>1.468%</td>
</tr>
<tr>
<td>2.5</td>
<td>4</td>
<td>2.3376</td>
<td>2.4310</td>
<td>3.996%</td>
</tr>
<tr>
<td>3.0</td>
<td>4</td>
<td>2.0765</td>
<td>2.1016</td>
<td>1.209%</td>
</tr>
</tbody>
</table>

Figure 10. Slope of \( \langle r^2 \rangle \) versus time for equally spaced obstacles (continuous) model and the lattice-based surrogate model with best \( M_c \) values. Results are shown for particle radius \( R_{\text{part}} = 1.0 \) and varying obstacle radius \( R_{\text{obs}} \).
5.1.2. Multisize Obstacles Case

Obstacle radii: Continuous model simulations were also performed for the multisize obstacles model (Figure 5) on an RVE partitioned into four subdomains as described in Section 3.3.1. The particle radius was fixed at $R_{\text{part}} = 1.0$, with one obstacle radius fixed at $R_{\text{obs1}} = 1.0$ and the other obstacle radius ($R_{\text{obs2}}$) varied from 0.5 to 3.0 in increments of 0.5. Stationarity and linearity of $\langle r^2 \rangle$ versus time is illustrated, for a representative case ($R_{\text{part}} = 1.0$, $R_{\text{obs1}} = 1.0$ and $R_{\text{obs2}} = 1.0$), in Figure 11.

**Figure 11.** Mean squared displacement versus time for the multisize obstacles model. Effects of an increasing number of realizations are shown for the case $R_{\text{part}} = 1.0$, $R_{\text{obs1}} = 1.0$ and $R_{\text{obs2}} = 1.5$: (a) curves; (b) linear fits to the curves.

Determining transition probabilities: Based on the approach illustrated in Figure 5, continuous model results were used to determine transition probabilities for the multisize obstacles model using (17).

Commitment index calibration: Lattice-based surrogate models were built using the transition probabilities obtained from the continuous model (Section 3.3.2) by first calibrating the commitment index $M_C$ (Section 4.1). For each case considered, $M_C$ was varied from 3 to 8 and calibrated by comparing the continuous and surrogate plots of $\langle r^2 \rangle$ versus time. A sample plot illustrates increasing stationarity of the slope of $\langle r^2 \rangle$ versus obstacle radius as the number of realizations increases (Figure 12—left).

**Figure 12.** Slope of the mean squared displacement for the multisize obstacles model: (left) variation with obstacle radius as the number of realizations is increased; (right) comparison of the continuous and lattice-based surrogate models for the best $M_C$ values.

We observe the manner in which the slope decreases with increasing obstacle radius $R_{\text{obs2}}$, which differs from that observed in the equally spaced obstacles case (Figure 10). A summary of the transition probabilities and values of $M_C$ needed to build the lattice-based surrogate model for all cases considered is provided in Table 3. Note that the sum of the probabilities in each row is 1 when the second probability is weighted by a factor of 2 due to symmetries in the RVE.
Comparisons of slopes for the relation between \( \langle r^2 \rangle \), including the best values for \( M_C \), are shown for the continuous and surrogate models in Table 4 and Figure 12—right. These results exhibit a difference in slopes ranging from 0.6% to roughly 7%.

### Table 3. Transition probabilities in the lattice-based surrogate model for the multisize obstacles (continuous) model for the best values of \( M_c \) with particle radius \( R_{\text{part}} = 1.0 \) and \( R_{\text{obs}} = 1.0 \).

<table>
<thead>
<tr>
<th>( R_{\text{obs}} )</th>
<th>( M_c )</th>
<th>( p_1^M )</th>
<th>( p_2^M )</th>
<th>( p_3^M )</th>
<th>( p_4^M )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>6</td>
<td>0.99253</td>
<td>1.8680 ( \times 10^{-3} )</td>
<td>2.0159 ( \times 10^{-3} )</td>
<td>1.7147 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>1.0</td>
<td>6</td>
<td>0.99291</td>
<td>1.7715 ( \times 10^{-3} )</td>
<td>1.7764 ( \times 10^{-3} )</td>
<td>1.7725 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>1.5</td>
<td>6</td>
<td>0.99324</td>
<td>1.6894 ( \times 10^{-3} )</td>
<td>1.5339 ( \times 10^{-3} )</td>
<td>1.8510 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>2.0</td>
<td>6</td>
<td>0.99352</td>
<td>1.6203 ( \times 10^{-3} )</td>
<td>1.2909 ( \times 10^{-3} )</td>
<td>1.9522 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>2.5</td>
<td>5</td>
<td>0.99316</td>
<td>1.7098 ( \times 10^{-3} )</td>
<td>1.1322 ( \times 10^{-3} )</td>
<td>2.2850 ( \times 10^{-3} )</td>
</tr>
<tr>
<td>3.0</td>
<td>5</td>
<td>0.99339</td>
<td>1.6502 ( \times 10^{-3} )</td>
<td>8.1439 ( \times 10^{-4} )</td>
<td>2.4950 ( \times 10^{-3} )</td>
</tr>
</tbody>
</table>

### Table 4. Slope values for \( \langle r^2 \rangle \) vs. \( t \) in the multisize obstacles (continuous) model \( (R_{\text{part}} = 1.0, R_{\text{obs}} = 1.0) \) and the lattice-based surrogate model. The best values of \( M_C \) are also shown.

<table>
<thead>
<tr>
<th>( R_{\text{obs}} )</th>
<th>( M_c )</th>
<th>Slope (Continuous)</th>
<th>Slope (Surrogate)</th>
<th>Percent Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>6</td>
<td>2.9870</td>
<td>3.0059</td>
<td>0.632%</td>
</tr>
<tr>
<td>1.0</td>
<td>6</td>
<td>2.7357</td>
<td>2.8531</td>
<td>4.291%</td>
</tr>
<tr>
<td>1.5</td>
<td>6</td>
<td>2.7376</td>
<td>2.8939</td>
<td>5.709%</td>
</tr>
<tr>
<td>2.0</td>
<td>6</td>
<td>2.6618</td>
<td>2.4751</td>
<td>7.014%</td>
</tr>
<tr>
<td>2.5</td>
<td>5</td>
<td>2.4995</td>
<td>2.5904</td>
<td>3.637%</td>
</tr>
<tr>
<td>3.0</td>
<td>5</td>
<td>2.2710</td>
<td>2.3651</td>
<td>4.144%</td>
</tr>
</tbody>
</table>

### 5.2. Estimating Diffusivities

We use the continuous and lattice-based surrogate model results to estimate effective and theoretical diffusivities (see Figure 1).

#### 5.2.1. Equally Spaced Obstacles Case

We estimate the effective diffusivities from simulated results for the mean squared displacement in both the continuous \( (\kappa_{\text{cont}}) \) and surrogate \( (\kappa_{\text{surr}}) \) models via (18) and for each obstacle radius (Figure 13).

Time averages for both effective diffusivities over the last 100,000 values are shown in Table 5, along with the commitment index \( M_C \) and the stay probability \( q \). The continuous and surrogate diffusivities exhibit a correlation coefficient of \( R = 0.9784 \) and percent differences between 0.02% and roughly 6%.

### Table 5. Effective diffusivities in the equally spaced obstacles case for the continuous model \( (\kappa_{\text{cont}}) \) and the lattice-based surrogate model \( (\kappa_{\text{surr}}) \). The stay probability \( q \) and best commitment index \( (M_C) \) are also shown for each obstacle radius \( (R_{\text{obs}}) \).

<table>
<thead>
<tr>
<th>( R_{\text{obs}} )</th>
<th>( M_C )</th>
<th>( q )</th>
<th>( \kappa_{\text{cont}} )</th>
<th>( \kappa_{\text{surr}} )</th>
<th>Percent Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>7</td>
<td>0.99276</td>
<td>0.7550</td>
<td>0.7100</td>
<td>5.960%</td>
</tr>
<tr>
<td>1.0</td>
<td>6</td>
<td>0.99291</td>
<td>0.7314</td>
<td>0.7084</td>
<td>3.145%</td>
</tr>
<tr>
<td>1.5</td>
<td>5</td>
<td>0.99297</td>
<td>0.6759</td>
<td>0.6862</td>
<td>1.524%</td>
</tr>
<tr>
<td>2.0</td>
<td>5</td>
<td>0.99371</td>
<td>0.6146</td>
<td>0.6165</td>
<td>0.309%</td>
</tr>
<tr>
<td>2.5</td>
<td>4</td>
<td>0.99384</td>
<td>0.5885</td>
<td>0.6016</td>
<td>2.226%</td>
</tr>
<tr>
<td>3.0</td>
<td>4</td>
<td>0.99493</td>
<td>0.5188</td>
<td>0.5187</td>
<td>0.019%</td>
</tr>
</tbody>
</table>
Figure 13. Estimated effective diffusivities for the continuous and lattice-based surrogate models in the equally spaced obstacles case with $R_{\text{part}} = 1.0$: (a) $R_{\text{obs}} = 0.5$, $M_C = 6$; (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) $R_{\text{obs}} = 3.0$, $M_C = 4$.

Estimated diffusivities for the surrogate models were used to compute values of $\alpha^2$ and $\Delta x (= \frac{L}{\alpha})$, for each case. Comparisons of the effective diffusivities to the resulting theoretical diffusivities, determined via (33), are shown in Table 6. For comparison, the theoretical estimate of the diffusivity for the case $L = 10$, corresponding to $\alpha^2 = 1$, is also shown in the table. The estimated values of both $\alpha^2$ (mean 5.0466, st. dev. 0.0888) and $\Delta x$ (mean 4.4519, st. dev. 0.0397) exhibited little variation or correlation ($R = 0.5532$) across the range of obstacle sizes considered. This finding suggests that the coefficient $\alpha^2$ can be approximated as constant across the range of obstacle radii considered.

The 2D lattice-based surrogate model for the equally spaced obstacles case is thus fully represented by the set of parameters $q, p (= 1 - q), \Delta x$ and $\alpha$ shown in Table 6 for each obstacle radius.
Table 6. Theoretical diffusivity ($D$) and grid size ($\Delta x$) in the equally spaced obstacles case. The stay probability ($q$), best commitment index ($M_C$) and theoretical diffusivity with $L = 10$ ($D|\Delta x=10$) are also shown for each obstacle radius ($R_{\text{obs}}$).

| $R_{\text{obs}}$ | $q$       | $a^2$  | $\Delta x$ | $\kappa_{\text{cont}}$ | $\kappa_{\text{surr}}$ | $D$         | $D|\Delta x=10$ |
|---------------|-----------|--------|------------|-------------------------|-------------------------|-------------|-----------------|
| 0.5           | 0.99276   | 5.1048 | 4.260      | 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          |

5.2.2. Multisize Obstacles Case

We also estimated effective diffusivities from simulated results for the continuous and surrogate models via (18) and for each obstacle radius combination (Figure 14). Time averages for each of these effective diffusivities over the last 100,000 values are shown in Table 7 along with the commitment index $M_C$ and the stay probability $q$. The continuous and surrogate diffusivities exhibit a correlation coefficient of $R = 0.8503$ and differences ranging between roughly 1.1% and 6.8%.

Figure 14. Estimated effective diffusivities for the continuous and lattice-based surrogate models in the multisize obstacles case with $R_{\text{part}} = 1.0$, $R_{\text{obs1}} = 1.0$: (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) $R_{\text{obs2}} = 3.0$, $M_C = 6$. 
Table 7. Effective diffusivities in the multisize obstacles case for the continuous model ($\kappa_{\text{cont}}$) and the lattice-based surrogate model ($\kappa_{\text{surr}}$). The resulting stay probability ($q_1$) in the square lattice surrogate model and the best commitment index ($M_C$) are also shown for each obstacle radius ($R_{\text{obs}}$).

<table>
<thead>
<tr>
<th>$R_{\text{obs}}$</th>
<th>$M_c$</th>
<th>$q_1$</th>
<th>$\kappa_{\text{cont}}$</th>
<th>$\kappa_{\text{surr}}$</th>
<th>Percent Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>6</td>
<td>0.9925</td>
<td>0.7509</td>
<td>0.7428</td>
<td>1.084%</td>
</tr>
<tr>
<td>1.0</td>
<td>6</td>
<td>0.9929</td>
<td>0.6842</td>
<td>0.7261</td>
<td>6.133%</td>
</tr>
<tr>
<td>1.5</td>
<td>6</td>
<td>0.9932</td>
<td>0.6943</td>
<td>0.7418</td>
<td>6.845%</td>
</tr>
<tr>
<td>2.0</td>
<td>6</td>
<td>0.9935</td>
<td>0.6746</td>
<td>0.6290</td>
<td>6.755%</td>
</tr>
<tr>
<td>2.5</td>
<td>5</td>
<td>0.9932</td>
<td>0.6205</td>
<td>0.6440</td>
<td>3.785%</td>
</tr>
<tr>
<td>3.0</td>
<td>5</td>
<td>0.9939</td>
<td>0.5667</td>
<td>0.5942</td>
<td>4.859%</td>
</tr>
</tbody>
</table>

Recalling Equation (31) (Section 4.2.2), for the theoretical diffusivities in the multisize obstacles case, we write

$$D_H = \frac{q_x (\Delta x_M)^2}{2(\Delta t)}, \quad D_V = \frac{q_y (\Delta x_M)^2}{2(\Delta t)}.$$  (38)

Here, $\Delta x_M$ denotes the effective grid size for the multisize obstacles case, and its value remains to be calibrated. Note, however, that the ratio of these two diffusivities is independent of this parameter. Consequently, we first examined this ratio using (36) by evaluating the relevant mean squared displacements (Figure 15) and their ratio (Figure 16).

Figure 15. Horizontal, vertical and (regular) mean squared displacements for the lattice-based surrogate model in the multisize obstacles case with $R_{\text{part}} = 1.0$, $R_{\text{obs}} = 1.0$: (a) $R_{\text{obs}} = 0.5$, $M_c = 6$; (b) $R_{\text{obs}} = 1.0$, $M_c = 6$; (c) $R_{\text{obs}} = 1.5$, $M_c = 6$; (d) $R_{\text{obs}} = 2.0$, $M_c = 6$; (e) $R_{\text{obs}} = 2.5$, $M_c = 6$; (f) $R_{\text{obs}} = 3.0$, $M_c = 6$. 
The multisize obstacles model has the property that the RVE is more open in either the horizontal ($R_{\text{obs}} < 1$) or vertical ($R_{\text{obs}} > 1$) direction, depending on the value of $R_{\text{obs}}$. This effect is subtle for $R_{\text{obs}} < 1$ due to the large amount of available space for particle diffusion (Figures 15a and 16a). Significant effects on the ratio of diffusivities are observed for larger values in the regime $R_{\text{obs}} > 1$, where available space in the horizontal direction becomes more limited (Figures 15e,f and 16e,f). An increased time to achieve a stationary response is also observed for these cases.

The effective grid size $\Delta x_M$ can be computed in one of two ways, by approximating $D_H$ and $D_V$ by $\kappa_H$ and $\kappa_V$, respectively, and solving the corresponding relation in (38) for $\Delta x_M$ (Table 8). The values of $\Delta x_M$ computed in this manner are 4.2268 ± 0.1444 (mean ± SD) using $\kappa_H$ and 4.0684 ± 0.2099 using $\kappa_V$. A single-factor ANOVA analysis results in p-values of 0.159 for $R_{\text{obs}}$ between 0.5 and 3.0, 0.226 for $R_{\text{obs}}$ between 0.5 and 2.5 and 0.394 for $R_{\text{obs}}$ between 0.5 and 2.0. These findings suggest that $\Delta x_M$ can be approximated as constant, with some caution needed for the case with the largest obstacle sizes.
Table 8. Estimated values of grid size $\Delta x_M$ based on the lattice-based surrogate model simulation results for the multisize case using the effective horizontal diffusivity ($\kappa_H$) and the effective vertical diffusivity ($\kappa_V$). The resulting transition (leave) probabilities in the square lattice surrogate model are also shown.

<table>
<thead>
<tr>
<th>$R_{obs}$</th>
<th>$q_y = 2p_2$</th>
<th>$q_z = p_3 + p_4$</th>
<th>$\Delta x_M$ ($\kappa_V$ est.)</th>
<th>$\Delta x_M$ ($\kappa_H$ est.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.0037360</td>
<td>0.0037629</td>
<td>4.2097</td>
<td>4.1652</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0035430</td>
<td>0.0035489</td>
<td>4.0917</td>
<td>4.1101</td>
</tr>
<tr>
<td>1.5</td>
<td>0.003788</td>
<td>0.0033849</td>
<td>4.3173</td>
<td>4.1280</td>
</tr>
<tr>
<td>2.0</td>
<td>0.0032406</td>
<td>0.0032431</td>
<td>4.4659</td>
<td>4.3340</td>
</tr>
<tr>
<td>2.5</td>
<td>0.0034196</td>
<td>0.0034172</td>
<td>4.1858</td>
<td>3.9538</td>
</tr>
<tr>
<td>3.0</td>
<td>0.0033004</td>
<td>0.0033094</td>
<td>4.0902</td>
<td>3.7193</td>
</tr>
</tbody>
</table>

The 2D lattice-based surrogate model for the multisize obstacles case is thus fully represented by the set of parameters $q_x$, $q_y$, $q_1 (= 1 - q_x - q_y)$ and $\Delta x_M$, shown for each obstacle radius in Table 8. In reality, using an average of the two calculated values of $\Delta x_M$ is recommended.

6. Discussion and Conclusions

In this study, we developed a systematic approach (outlined in Figure 1) for bridging two-dimensional continuous models of diffusive transport in the presence of rigid internal obstacles with perfectly elastic collisions with two types of surrogate models. The first type is a lattice-based surrogate model that exploits symmetries in the RVE of the continuous domain used for direct numerical simulation. The second type is a square lattice surrogate model that can enable the accurate estimation of diffusivities to be subsequently used in more general lattice-based modeling frameworks where diffusion is one of several important coupled mechanisms. Here, diffusion is often the fastest spatiotemporal mechanism. Our systematic approach can be used, first, to establish the discrete model probabilities and grid size for building 2D lattice models in applications that couple diffusion with additional mechanisms in a background medium having specific geometric properties. Conversely, in other applications, the ability to calibrate the grid size such that the associated probabilities are consistent with an observed or prescribed diffusivity may also be beneficial.

While the level of noise and error increases in the more complex (multisize obstacles) model considered in this study, it remains in a range that is comparable to the level of noise or variation typical for 2D lattice-based simulations with underlying stochastic or probabilistic mechanisms. Our approach is versatile in the sense that it can be extended to other geometric configurations of obstacles in an RVE exhibiting several degrees of symmetry. Similarly, the techniques used in our overall approach can be extended to three-dimensional lattice-based models for materials with internal geometric symmetries. Nevertheless, it should be noted that the computational time needed to achieve stationarity of the underlying continuous and lattice-based surrogate model probabilities can increase substantially with increasing model complexity. Consequently, the extent to which computation time is practical (versus prohibitive) will depend on the degree of internal symmetry in the background medium as well as availability of ultra-high-performance computing resources.

Author Contributions: Conceptualization, T.M.A. and M.A.H.; Methodology, T.M.A. and M.A.H.; Validation, T.M.A. and M.A.H.; Writing—original draft, T.M.A. and M.A.H.; Writing—review & editing, T.M.A. and M.A.H.; Visualization, T.M.A. and M.A.H. All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded and the work was supported by the US National Science Foundation via grant DMS-1638521 and an NSF Graduate Research Fellowship.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.
Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author. The data presented in this study are available on request from the corresponding author.

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

Appendix A. Overview and Derivation of Type 1 Limits

We summarize several useful limits in evaluating Equation (6) (Section 2.2.2). We delineate these limits as those involving only the relative displacements (Type 1) and as those involving the position of the particle (Type 2). Direct evaluation of the Type 1 limits yields the following results associated with case (i):

\[
\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}^N = 0, \quad (A1)
\]

\[
\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}^N h_{xi}^M = 0 \quad \text{when } N \neq M, \quad (A2)
\]

\[
\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} (h_{xi}^N)^2 = p_x(\Delta x)^2 \quad \text{i.e., when } N = M. \quad (A3)
\]

Analogous results also hold for the vertical transition case (ii) and are derived using the same steps shown in this appendix.

The Type 2 limits can also be evaluated directly (see Appendix B) and yield the following results:

\[
\lim_{K \to \infty} \frac{2}{K} \sum_{i=1}^{K} x_i^N h_{xi}^M = 0 \quad \text{when } M > N, \quad (A4)
\]

\[
\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} (x_i^N)^2 = N p_x(\Delta x)^2. \quad (A5)
\]

The above limits are for the horizontal transition case \((x_i)\), but there are similar analogous results that can be derived using the same steps for the vertical transition case.

Derivation of Type 1 Limits

Evaluating Equation (A1). Consider the average displacement in the horizontal direction, over all realizations, as \(K \to \infty\),

\[
\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}(N). \quad (A6)
\]

Let \(k_{x1}, k_{x2}\) and \(k_{x3}\) be defined as follows:

(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 then write

\[
\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}(N) = \lim_{K \to \infty} \frac{1}{K} \left[ k_{x2}(\Delta x) + k_{x3}(-\Delta x) \right] = \lim_{K \to \infty} (\Delta x) \left( \frac{k_{x2}}{K} - \frac{k_{x3}}{K} \right). \quad (A7)
\]
For the equally spaced obstacles problem, moving left or right is equally likely. Hence, for very large values of $K$, it follows that $\frac{k_x}{K} = \frac{p_x}{2}$ and that $\frac{k_y}{K} = \frac{p_y}{2}$. We then write

$$\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}(N) = \lim_{K \to \infty} (\Delta x) \left( \frac{k_x}{K} - \frac{k_x}{K} \right) = (\Delta x) \left( \frac{p_x}{2} - \frac{p_x}{2} \right) = 0. \quad (A8)$$

Similarly, in the vertical direction, by analogous steps we obtain the result

$$\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{yi}(N) = 0. \quad (A9)$$

Evaluating Equation (A2). We have three cases for $h_{xi}(N)$ and $h_{yi}(M)$:

(i) $h_{xi}(N) = \Delta x$ and $h_{yi}(M) = -\Delta x$ (or vice versa),

(ii) $h_{xi}(N) = h_{yi}(M) = \Delta x$ or $h_{xi}(N) = h_{yi}(M) = -\Delta x$,

(iii) $h_{xi}(N) = 0$ and/or $h_{yi}(M) = 0$.

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

(i) $k_{x4}$: the number of times that $h_{xi}(N) = \Delta x$ and $h_{yi}(M) = -\Delta x$ (or vice versa),

(ii) $k_{x5}$: the number of times that $h_{xi}(N) = h_{yi}(M) = \Delta x$ or $h_{xi}(N) = h_{yi}(M) = -\Delta x$,

(iii) $k_{x6}$: the number of times that $h_{xi}(N) = 0$ and/or $h_{yi}(M) = 0$.

Consider the limit

$$\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}(N) h_{yi}(M) = \lim_{K \to \infty} (\Delta x)^2 \left[ \frac{k_{x5}}{K} - \frac{k_{x4}}{K} \right]. \quad (A10)$$

We have equal probabilities for events corresponding with $k_{x4}$ and $k_{x5}$, and we expect for very large values of $K$ that $\frac{k_{x4}}{K} = \frac{k_{x5}}{K}$. Hence,

$$\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}(N) h_{yi}(M) = \lim_{K \to \infty} (\Delta x)^2 \left[ \frac{k_{x5}}{K} - \frac{k_{x4}}{K} \right] = 0. \quad (A11)$$

Similarly, in the vertical direction, by analogous steps we obtain the result

$$\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{yi}(N) h_{yi}(M) = 0, \quad (A12)$$

when $N \neq M$.

Evaluating Equation (7). In the case $N = M$, and with $k_{x1}$, $k_{x2}$ and $k_{x3}$ defined as above,

$$\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}(N) h_{xi}(M) = \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{yi}^2(N) = \lim_{K \to \infty} \left( \frac{k_{x2} + k_{x3}}{K} \right) (\Delta x)^2, \quad (A13)$$

yielding the result

$$\lim_{K \to \infty} \left( \frac{k_{x2} + k_{x3}}{K} \right) = p_x. \quad (A14)$$

Hence, for $N = M$,

$$\lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_{xi}(N) h_{xi}(M) = p_x (\Delta x)^2, \quad (A15)$$

which corresponds to the first relation in (7). We can again follow a similar line of reasoning to obtain a similar result in the vertical direction.
Appendix B. Derivation of Type 2 Limits

Consider a particle’s possible horizontal locations after one time step,

\[ x_i(N) = x_i(N - 1) + h_{xi}(N), \quad (A16) \]

We use induction to derive the second equation in (7).

When \( N = 0 \), we obtain

\[ \lim_{K \to \infty} 2 K \sum_{i=1}^{K} x_i(0) h_{xi}(M) = 0, \quad M > 0. \]  

(A17)

When \( N = 1 \), we have,

\[ \lim_{K \to \infty} 2 K \sum_{i=1}^{K} x_i(1) h_{xi}(M), \quad M > 1 \]
\[ = \lim_{K \to \infty} 2 K \sum_{i=1}^{K} (x_i(0) + h_{xi}(1)) h_{xi}(M) \]
\[ = \lim_{K \to \infty} 2 K \sum_{i=1}^{K} h_{xi}(1) h_{xi}(M) = 0. \]  

(A18)

For \( N > 1 \), we employ mathematical induction, first assuming that

\[ \lim_{K \to \infty} 2 K \sum_{i=1}^{K} x_i(N) h_{xi}(M) = 0 \quad \text{when} \quad M > N, \]  

(A19)

holds for \( N > 1 \). We now show that it also holds for \( N + 1 > 2 \) via the following steps:

\[ \lim_{K \to \infty} 2 K \sum_{i=1}^{K} x_i(N+1) h_{xi}(M) \]
\[ = 2 K \sum_{i=1}^{K} (x_i(N) + h_{xi}(N+1)) h_{xi}(M) \]
\[ = 2 K \sum_{i=1}^{K} x_i(N) h_{xi}(M) + 2 K \sum_{i=1}^{K} h_{xi}(N+1) h_{xi}(M). \]  

(A20)

Using the inductive hypothesis in Equation (A19), we obtain the result

\[ \lim_{K \to \infty} 2 K \sum_{i=1}^{K} x_i(N+1) h_{xi}(M) = 0 \quad \text{for} \quad M > N, \]  

(A21)

which corresponds to the second relation in (7).

Lastly, we use mathematical induction to show that

\[ \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(N) = N p_x(\Delta x)^2. \]  

(A22)

When \( N = 1 \), Equation (A22) becomes

\[ \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(1) = \lim_{K \to \infty} \frac{1}{K} \sum_{i=1}^{K} (x_i(0) + h_i(1))^2 = p_x(\Delta x)^2. \]  

(A23)
When $N = 2$, Equation (A22) becomes

$$\lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(2) = \lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} (x_i(1) + h_i(2))^2$$

$$= \lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} (x_i^2(1) + 2x_i(1)h_i(2) + h_i^2(2))$$

$$= \lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(1) + \lim_{k \to \infty} \frac{2}{K} \sum_{i=1}^{K} x_i(1)h_i(2)$$

$$+ \lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_i^2(2).$$

(A24)

We can rewrite this as

$$\lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(2) = 2p_x(\Delta x)^2 + \lim_{k \to \infty} \frac{2}{K} \sum_{i=1}^{K} x_i(1)h_i(2)$$

$$= 2p_x(\Delta x)^2 + 0.$$

(A25)

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

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

(A26)

holds for $N$. We now show that it also holds for $N + 1$ via the following steps:

$$\lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(N + 1) = \lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} (x_i(N) + h_i(N + 1))^2$$

$$= \lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} \left( x_i^2(N) + 2x_i(N)h_i(N + 1) + h_i^2(N + 1) \right)$$

(A27)

$$= \lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(N) + \lim_{k \to \infty} \frac{2}{K} \sum_{i=1}^{K} x_i(N)h_i(N + 1)$$

$$+ \lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} h_i^2(N + 1).$$

Using the inductive hypothesis in Equation (A26), we obtain,

$$\lim_{k \to \infty} \frac{1}{K} \sum_{i=1}^{K} x_i^2(N + 1)$$

$$= Np_x(\Delta x)^2 + \lim_{k \to \infty} \frac{2}{K} \sum_{i=1}^{K} x_i(N)h_i(N + 1) + p_x(\Delta x)^2$$

(A28)

$$= (N + 1)p_x(\Delta x)^2 + \lim_{k \to \infty} \frac{2}{K} \sum_{i=1}^{K} x_i(N)h_i(N + 1).$$

The last term above is zero, so that

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

(A29)

Hence,

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

(A30)
References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.