Control of stochastic boundary coverage by multi-robot systems

Theodore P. Pavlic*
Associate Research Scientist
School of Life Sciences
Arizona State University
Tempe, AZ 85281
Email: tpavlic@asu.edu

Sean Wilson
Graduate Research Associate
School for Engineering
of Matter, Transport and Energy
Arizona State University
Tempe, AZ 85281
Email: Sean.T.Wilson@asu.edu

Ganesh P. Kumar
Graduate Research Associate
School for Computing, Informatics,
and Decision Systems Engineering
Arizona State University
Tempe, AZ 85281
Email: Ganesh.P.Kumar@asu.edu

Spring Berman
Assistant Professor
School for Engineering
of Matter, Transport and Energy
Arizona State University
Tempe, AZ 85281
Email: Spring.Berman@asu.edu

This technical brief summarizes and extends our recently introduced control framework for stochastically allocating a swarm of robots among boundaries of circular regions. As in the previous work, a macroscopic model of the swarm population dynamics is used to synthesize robot control policies that establish and maintain stable predictable team sizes around region boundaries. However, this extension shows that the control strategy can be implemented with no robot-to-robot communication. Moreover, target team sizes can vary across different types of regions, where a region's type is a subjective characteristic that only needs to be detectable by each individual robot. Thus, regions of one type may have a higher equilibrium team size than regions of another type. In other work that predicts and controls stochastic swarm behaviors using macroscopic models, the equilibrium allocations of the swarm are sensitive to changes in the mean robot encounter rates with objects in the environment. Thus, in those works, as the swarm density or number of objects changes, the control policies on each robot must be re-tuned to achieve the desired allocations. However, our approach is insensitive to changes in encounter rate and therefore requires no re-tuning as the environment changes. In this extension, we validate these claims and show how the convergence rate to the target equilibrium allocations can be controlled in swarms with a sufficiently large free-robot population. Furthermore, we demonstrate how our framework can be used to experimentally measure the rates of robot encounters with occupied and unoccupied sections of region boundaries. Thus, our method can be viewed both as an encounter-rate-independent allocation strategy as well as a tool for accurately measuring encounter rates when using other swarm control strategies that depend on them.

Keywords: modeling of dynamic systems; robotics; uncertain systems and robust control

1 Introduction

In this technical brief, we extend our prior work [1] on developing a scalable, robust control framework for the problem of allocating a robotic swarm in target group sizes around the boundaries of disjoint, stationary circular regions. We use a *stochastic* approach where agents behave randomly, and the designer chooses probabilistic parameters that shape the distribution of possible behaviors by each agent. Although deterministic strategies usually exist that perform better than stochastic strategies for the same problem, the deterministic approaches often require sophisticated communication and navigation capabilities and are accordingly more sensitive to communication and navigation errors [2]. Consequently, there are already a variety of existing stochastic strategies for achieving task allocation, assembly and self-assembly, and formation control in swarm robotic systems [e.g., 3–11]. In most cases, the stochastic control policies are either inspired by or can be translated into an abstract chemical reaction network (CRN). That is, each robot *reacts* probabilistically to events generated by its environment much like two entities of a gas probabilistically reacting after coming into close proximity. These artificial reactions

*Corresponding author.

are engineered to be reversible – state transitions triggered by external events will eventually be undone by an artificial decay process implemented by stochastic timer within each robot. After a transient period, the total flux of event-triggered transitions will be precisely balanced by the flux of decay events, and the system reaches an equilibrium state allocation. Unfortunately, the parameters of the internal decay events must be tuned to match empirical measurements of the environment. Consequently, the equilibrium allocations are impractically sensitive to minor changes in the environment. For example, increasing the number of robots in a swarm within a fixed volume leads to an increased encounter rate between robots, and so new control policies must be broadcast to all robots to update their delay parameters so that the flux-balancing equilibrium is not changed. To address this issue, our prior stochastic task-allocation work [1] replaced delay-driven reverse reactions with novel event-driven catalytic reverse reactions. This modification allows the system to compensate for changes in encounter rate, and thus the equilibrium allocations are robust to changes in swarm density, robot speed, and environmental properties such as number and size of allocation regions.

Here, we further validate our claims that catalytic reverse reactions provide significant increases in robustness and modularity. Our previously proposed robotic implementation required each robot to communicate one bit of information to other robots randomly encountered while moving around the arena. Since then, we have modified our implementation to be communication free. We have also generated additional data that demonstrates that our approach is effective at decomposing a relatively complicated task allocation problem over multiple task types into several simple and independent single-type problems. Furthermore, we have found that our approach is a generalized case of a classic problem in probability theory, the sequential parking problem [12, 13], which explains why the so-called *parking constant* [14] emerges from our results for special cases.

The stochastic control approach that we develop for multi-robot boundary coverage is related to other recent work of ours on developing models and control strategies for multi-robot collective transport of payloads that are too heavy to be transported by a single robot. In one case [15], we fit a stochastic hybrid system model to experimental data on group retrieval of a standardized circular load by self-organized teams of ants. We modeled the ants as spontaneously detaching from the load at a constant rate. Here, the analogous detachment process is driven by encounters between bound and unbound robots. In our more recent work [16], we show how ant-like spontaneous detachment can be emulated by this catalytic detachment process, allowing us to define robot control policies that reproduce the robust transport behaviors exhibited by ants in a way that is independent of the robots' environment.

The remaining sections of this technical brief are organized as follows. In Section 2, we present an overview of the boundary coverage problem and the salient features of the catalytic stochastic strategy that we use to solve it. We also discuss how this approach improves upon others from the literature. In Section 3, we summarize the key theoretical results from our prior work on this problem. Then, in Section 4, we present new data that demonstrates that this approach can decompose a coverage problem with multiple region types into several independent single-type problems. Furthermore, in Section 5, we show how a reduced-order linear model closely approximates the dynamics of robotic swarms that are either sufficiently large or sufficiently close to their equilibrium. Using this reduced-order model, we incorporate optimization results from other literature on stochastic task allocation. Finally, we give some concluding remarks in Section 6.

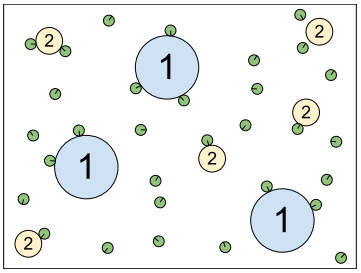


Fig. 1 Example scenario with two types of disk-shaped regions, labeled 1 and 2. The unlabeled circles are robots that are allocating themselves to the region boundaries.

2 Overview and Contribution

We consider the problem of allocating a swarm of identical robots around the boundaries of a set of circular regions. Regions are classified by *type*, which reflects their physical or subjective properties. For instance, in the case where the regions are payloads to be transported, *type* can signify the effort required to move the corresponding payload, based on its size or weight. Thus, different payload types will require different robot team sizes. In surveillance applications, regions may require different levels of coverage based on a subjective measure. For example, although all regions may be the same size, some regions may be known to have a higher occurrence rate of interesting events. We discuss other possible applications in our prior work [1]. In general, we view the boundary-allocation problem as a prototypical example of encounter-based task allocation in robot collectives. Our approach can be used as a template for introducing minor changes to existing stochastic task allocation strategies in order to significantly reduce their sensitivity to changes in environmental parameters.

An example of the focal scenario is depicted in Fig. 1, where the objective is for three robots to be allocated to each type-1 disk and one robot to each type-2 disk on average. As in our prior work [1], we consider only circular regions here for simplicity; we refer to these regions as *disks*. We assume that each robot has a small sensing radius and is capable of detecting other robots and disk boundaries. If there are multiple disk types, robots must have the ability to differentiate between them (e.g., by color or surface texture or some other physical variable). In our prior work [1], we assumed that

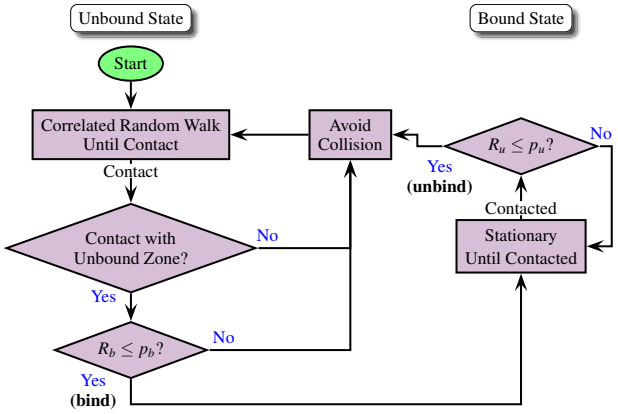


Fig. 2 Control flow chart. Unbound robots move randomly until encountering other robots, which are avoided, or unbound zones of disk boundary. On encountering an unbound zone, a robot compares a generated pseudo-random number $R_b \sim \text{unif}(0,1)$ to p_b to determine whether to bind to that zone. Once bound, the robot waits to detect the close proximity of another free robot. On that event, the robot generates a second pseudo-random number $R_u \sim \text{unif}(0,1)$ to compare to p_u to determine whether to unbind from the disk and begin the random walk again.

robots could also send a single bit of information to other encountered robots. However, we have now reformulated the control strategy so that no communication is necessary. The controller for each robot is shown in the flow chart in Fig. 2.

In order to disperse throughout an area, robots move according to a *correlated random walk* (CRW). That is, each robot's CRW motion iterates through two phases – it moves straight ahead for a short distance and then turns to a random angle before repeating this iteration. The *turning angle* of a CRW is a measure of the spread in the angular distribution of each turn. Thus, having a *low turning angle* means that robots will primarily drift in one direction because turns will be very subtle. Furthermore, having a *high turning angle* means that no one robot will travel in a straight line for very long, and groups of robots will slowly diffuse in all directions. Although we use a CRW in our simulations to validate our model, other motion programs are applicable.

Stochastic Allocation Strategy: In a deterministic allocation strategy, a robot might survey an encountered disk boundary before determining whether to bind to the disk. In that deterministic case, it would be counterintuitive, counterproductive, and inefficient for bound robots to unbind and become free after deliberately choosing to bind. However, this reverse transition occurs regularly in stochastic allocation strategies in lieu of a complicated deterministic site survey before binding. In the stochastic strategy described here, after a robot encounters a disk boundary, it gathers no additional information about other robots on the boundary and instead chooses to bind to the boundary with a probability p_b that depends on disk type but was fixed *a priori*. When a free robot encounters any other robot, bound or unbound, it executes maneuvers to avoid collisions. However, if a bound robot detects that it was encountered by a free robot, then the bound robot chooses to unbind from the disk with probability p_u that depends on the disk type. As will be described in detail in Section 3, the equilibrium concentration of robots on each disk type is set by the ratio of the type's binding and unbinding rates. Thus, reverse reactions in stochastic allocation strategies are used as a substitute for complicated detection and estimation strategies used in deterministic allocation strategies.

Comparison to Other Stochastic Strategies: A critical difference between our work and similar stochastic task-allocation approaches [e.g., 4, 6–11, 17, 18] is that the only way a bound robot can become free is by being encountered by another free robot. Using those other approaches, the bound robot would spontaneously become unbound according to a delay-triggered event. Consequently, the programmed mean delays have to be closely tuned to match environmental parameters that non-trivially determine the rate of encounters between robots and disks. In our case, because forward and reverse reactions are both triggered exclusively by encounter events, they self-compensate for environmental changes in those rates. For example, if the number of robots in a swarm is increased, the rate of robot–disk encounters will increase in the system. In other approaches, the equilibrium population of robots on disks would also increase unless each robot was re-programmed to reduce the time it waits to unbind. Thus, the traditional approaches behave like classical Langmuir chemical adsorption processes [19] – the equilibrium concentration of adsorbed particles on a surface increases with the density of particles in the solution. However, in our case, the increase in robots in the swarm will similarly increase both the robot–disk encounter rate, which determines the rate of binding, and the encounter rate of free robots with bound robots, which determines the rate of unbinding. Thus, the ratio of the binding and unbinding rates will be unchanged, and the equilibrium concentration will not vary with the number of robots in the swarm or their density. We are able to avoid the Langmuir phenomenon because our forward and reverse reactions are implemented as two *separate irreversible* reactions. Consequently, rather than simply simulating gas-like behavior on physical robots, we capitalize on the agency of each robot to implement a thermodynamic equilibrium that differs from those traditionally studied in statistical mechanics.

There is an alternative approach in the stochastic-robotics literature that might be used in this scenario to achieve task allocations that are independent of swarm size [7]. In that case, the forward and reverse reaction rates themselves are treated as control laws which can vary in time based on the current population of robots allocated to the focal and adjacent tasks. Like many deterministic strategies, that stochastic approach requires robots to be able to estimate and communicate the populations of robots around them. In contrast, our approach achieves desired target allocations without requiring these robot capabilities.

3 Analysis and Control

For each disk type, the design problem is to find a corresponding (p_b, p_u) pair that results in the desired equilibrium concentration of robots around disks of that type. As we will show, the equilibrium concentration for any one disk type is independent of the equilibrium concentration of any other disk type. Consequently, the design problem for each type is identical to the design problem for that type in isolation. So, without loss of generality, we initially consider the scenario of a single disk type, and then we conclude this section with a discussion of how to extend these results to the multiple-type case.

The robot–disk system described in Section 1 is a generalization of the classic linear sequential parking problem [12, 13]. There, given an arbitrarily long but finite length of street and cars that arrive and park at random locations, the problem is to find the mean occupancy when the street becomes saturated. The result is the so-called *parking constant* [14], which is approximately 75%. In our case, robots are parking along a finite line segment wrapped around the circumference of each region. However, parked robots may also vacate. Thus, our goal is to design egress and ingress policies to allow for any allocation from 0% to the parking constant.

In order to model this generalization of the classic parking problem, we begin in Section 3.1 by modeling the robot–disk system as a gas composed of three species: *free robots*, *bound* disk zones, and *unbound* disk zones. This description is motivated by the fact that the sum of bound and unbound space is invariant of the allocation dynamics. For theoretical convenience, we will begin by assuming that this gas is well mixed. This assumption is equivalent to assuming that the availability of bound and unbound space is proportional to their relative proportions. However, because two robots can bind to the same disk without allowing sufficient room for a third robot to bind between them, the assumption of a well-mixed system is not valid. In fact, this phenomenon is the reason why the solution to the parking problem is less than unity. Below, in Section 3.2, we will show how to convert the equilibrium of this well-mixed model into the equilibrium of a system that accurately models these spatial crowding effects. Then, in Section 3.3, we will describe how to use these results to synthesize a (p_b, p_u) control strategy for each disk type in a multi-disk allocation scenario.

3.1 ODE Dynamics of a Well-mixed CRN Model

The well-mixed chemical reaction network (CRN) that corresponds to the single-type case consists of the two irreversible bimolecular reactions:

$$r + U \xrightarrow{p_b e_u} B \tag{1a}$$

$$r + B \xrightarrow{p_u e_b} U + 2r \tag{1b}$$

where r represents the free-robot species, B represents bound zones, and U represents unbound zones. The mass-action rates $p_b e_u$ and $p_u e_b$ will be explained below. The $r + U \to B$ reaction represents the possibility of a free robot r encountering an unbound zone U and binding to it, thus converting the combination of the two into a single bound zone B. Similarly, the $r + B \to U + 2r$ reaction represents the possibility of a free robot r encountering a bound zone B and triggering the bound robot to unbind. In the unbinding event, not only does the bound zone convert to an unbound zone U, but both free robots $2r$ emerge. Thus, Eq. (1b) may be viewed as the reversal of Eq. (1a); however, it depends on the presence of an additional "catalytic" robot that is not consumed by the reaction.

The mass-action rates $p_b e_u$ and $p_u e_b$ represent the probability per unit of time that the corresponding reaction will occur when the two reactants are available. That is, if there exists a free robot r and an unbound zone U, then the two will encounter each other at rate e_u. In other words, on average, the two entities will wait $1/e_u$ time before meeting each other. On an encounter between r and U, the binding reaction will occur with probability p_b. Thus, each $r + U$ combination in the system will wait an average $1/(p_b e_u)$ time before being converted into a bound zone B. Moreover, if we let $[r]$ and $[U]$ be the number of free robots and unbound zones at any time, then the probability per unit time of any $r + U \to B$ transition is $[r][U] p_b e_u$. Similarly, the so-called *transition intensity* [20] of the $r + B \to U + 2r$ reaction is $[r][B] p_u e_b$. As we discuss in our prior work [1], these transition intensities can be used to define an ordinary differential equation (ODE) model of the mean-field dynamics of the CRN. The asymptotically stable equilibrium (B^*, U^*, r^*) of the resulting ODE is such that

$$\frac{B^*}{U^*} = \frac{p_b e_u}{p_u e_b} \tag{2a}$$

or, equivalently,

$$\frac{B^*}{B^* + U^*} = \frac{p_b e_u}{p_b e_u + p_u e_b} = \frac{\frac{p_b}{p_u}}{\frac{p_b}{p_u} + \frac{e_b}{e_u}} \tag{2b}$$

so long as the initial number of free robots r_0 and the initial number of bound robots B_0 are such that $r_0 + B_0 > B^*$. That is, at thermodynamic equilibrium, the average flux $r^* U^* p_b e_u$ of free robots into bound zones is precisely balanced by the average flux $r^* B^* p_u e_b$ of bound zones into free robots. So although the actual number of bound zones, unbound zones, and free robots fluctuates randomly for all time, the mean demographics stabilize to levels parameterized by the ratios p_b/p_u and e_u/e_b.

As shown in Eq. (2), so long as the parameter ratio e_b/e_u is a known constant or relatively fixed, the equilibrium concentration of bound and unbound zones will be determined entirely by the probabilities p_b and p_u which are under designer control. Because both encounter rates, e_b and e_u, are driven by robot motion, then environmental parameters that affect robot motion will similarly reduce or enhance both encounter rates. Consequently, the ratio e_b/e_u is expected to

be robust to most environmental variations and significantly simpler to estimate than either encounter rate alone. Alternatively, the catalytic reverse reaction in Eq. (1b) could be removed and the reaction in Eq. (1a) could be made to be reversible with a reverse mass-action rate ρ that represents that a bound robot waits $1/\rho$ time, on average, to spontaneously unbind and become free again. In this implementation, the delay $1/\rho$ would have to be adjusted every time a change in the environment caused a change in the either encounter rate (e.g., if the density or speed of the robots change). However, because both reactions in Eq. (1) are encounter driven, then these environmental changes only affect the convergence rate of the system; its equilibrium allocation is unchanged. Thus, using pairs of irreversible reactions allows for greater robustness in stochastic robot swarms. Other stochastic task-allocation strategies may similarly benefit from using catalytic encounter-driven reverse reactions as opposed to delay-based strategies.

3.2 Corrections for Spatial Effects on Boundaries

As discussed above, the robot–disk system is not well approximated by a well-mixed gas. In practice, it is likely that two robots interacting stochastically with a disk will bind with non-zero inter-robot space between them that is nevertheless too small for another robot to encounter. This intermediate-sized space will be wasted. It is as if a reactant in a gas becomes spatially isolated from other reactants, and this spatial isolation prevents corresponding reactions from occurring. There is a similar effect in the classical sequential parking problem [12, 13], and consequently the average allocation on a street at saturation is roughly 75% [14]; the other 25% of the street is wasted as space between cars that is too small to be used. However, our robots also have the ability to unbind. Thus, we must model this non-well-mixed effect at allocation levels well under saturation. Here, we describe how insights from the parking problem can be used to modify the equilibrium from Eq. (2) to accommodate non-well-mixed effects at allocation levels from 0% to saturation.

In our prior work [1], we show that at equilibrium, the non-well-mixed spatial effects lead to a magnification of each bound zone by a factor $(1+\delta)$ where δ is a measure of the mean distance between robots. First, we define the *avoidance distance* $a \geq 0$ be the minimum distance that each robot allows between itself and another robot. Then, we take $\delta : [0,1] \to [a, 1+2a]$ to be a function that maps the equilibrium allocation $B^*/(U^*+B^*)$ to the corresponding magnification of each bound zone. For convenience, we abbreviate $\delta(B^*/(U^*+B^*))$ as simply δ. Thus, using the substitution $B^* \mapsto B^* + \delta B^*$ and the complementary $U^* \mapsto U^* - \delta B^*$, the well-mixed equilibrium of Eq. (2) can be transformed into the spatially corrected equilibrium

$$\underbrace{\frac{(1+\delta)}{1-\delta\frac{B^*}{U^*}}}_{\text{Correction factor}} \frac{B^*}{U^*} = \frac{e_u p_b}{e_b p_u} \tag{3a}$$

where the underbraced expression is a correction factor for the spatial effects in a physical robot scenario. For comparison, the corrected and idealized allocations are related by

$$\begin{aligned} \frac{B^*}{U^*+B^*} &= \frac{B^*}{(U^*-\delta B^*)+(1+\delta)B^*} \\ &= \frac{\frac{B^*}{(1+\delta)B^*}}{\frac{U^*-\delta B^*}{(1+\delta)B^*}+1} = \frac{1}{1+\delta}\underbrace{\frac{\frac{p_b}{p_u}}{\frac{p_b}{p_u}+\frac{e_b}{e_u}}}_{\text{Idealized allocation}} \end{aligned} \tag{3b}$$

where the underbraced expression matches the idealized allocation ratio in Eq. (2b). So the ideal and actual allocations are predicted to be related by a $1/(1+\delta)$ gain. For low allocations, δ takes its maximum value of approximately $1+2a$, and the gain is predicted to approximately $1/(2+2a)$. For high allocations near saturation, δ will approach its minimum value of avoidance distance a, but it will be significantly larger than a due to non-well-mixed effects. For example, in the special case where $a=0$, δ will saturate such that the gain $1/(1+\delta)$ will be equal to the linear parking constant [14] of approximately 75%; that is, δ saturates at approximately $1/3$ when $a=0$.

An important future direction is to develop theory to predict the precise shape of the δ function from first principles. This problem is tantamount to solving the generalized linear parking problem for parked cars that can also leave. Consequently, it is a significant theoretical challenge. However, as we showed in our prior work [1], we have strong experimental evidence that

$$\delta(r) = A\cos\left(2\pi\frac{r}{T}+c\right)$$

subject to the constraints that $A \geq (1+a)/2$, $T \geq 2/(1+a)$, $c \geq 0$, $\delta(1/1+a) = a$, and $\delta(0) = L/\lceil (L/(1+2a)) \rceil$ where L is the linear distance of the disk circumference in units of robot width. In the special case when $a=0$, the additional constraint that $\delta(\gamma) = (1-\gamma)/\gamma$ can be used for the parking constant $\gamma \approx 0.7476$. Thus, rather than solving the generalized parking problem directly, it may be simpler to explain how this inter-robot δ spacing relationship emerges.

3.3 Control Synthesis for Multiple Disk Types

For initial number of free robots r_0 (i.e., roughly the size of the swarm), if $r_0 > B^*$, then the swarm allocation system for one disk type will converge to the equilibrium described by Eq. (3a). Thus, a (p_b, p_u) control policy for the single-type case can be synthesized using the rule

$$\frac{p_b}{p_u} = \frac{e_b}{e_u}\frac{B^*}{U^*}\frac{(1+\delta)}{1-\delta\frac{B^*}{U^*}} \tag{4}$$

where B^*/U^* is the desired bound–unbound ratio of zones at equilibrium. For example, if robots at equilibrium are to occupy 60% of a disk boundary on average, then the desired ratio $B^*/U^* = 0.6/(1-0.6) = 1.5$, and Eq. (4) can be applied

to generate the required p_b/p_u ratio for any given e_b/e_u ratio. For any given desired allocation, there is a continuum of corresponding (p_b, p_u) pairs. Consequently, as we demonstrate in Section 5.2, the absolute values of p_b and p_u can be chosen to optimize some other criteria so long as the p_b/p_u ratio satisfies Eq. (4).

Under the assumption that a robot can detect the difference between disk types, the robotic swarm can be configured to achieve a different equilibrium allocation for each disk type. If one additional disk type is added to the system, the CRN from Eq. (1) can simply be augmented with the two additional reactions

$$r + U' \xrightarrow{p'_b e'_u} B' \tag{5a}$$

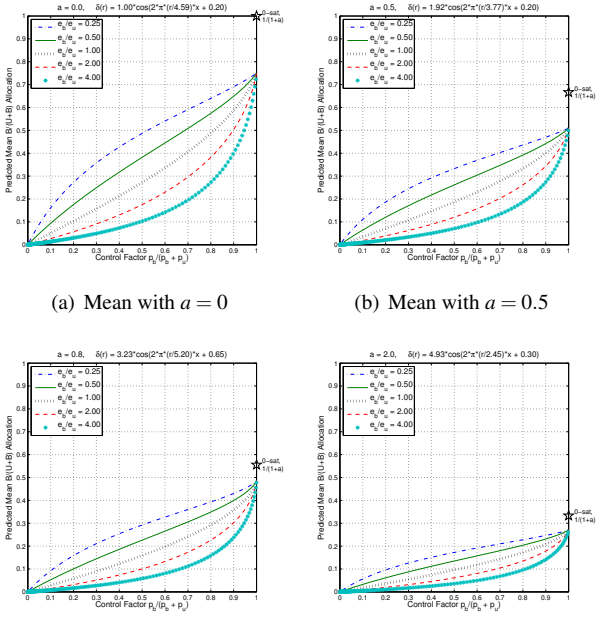
$$r + B' \xrightarrow{p'_u e'_b} U' + 2r \tag{5b}$$

where U' and B' represent the unbound and bound zones on the new disk type, and $p'_b e'_u$ and $p'_u e'_b$ represent the mass-action rates of the new reactions. So long as $r_0 > B^* + B'^*$, then both pairs of reactions in Eqs. (1) and (5) will meet the sufficient conditions for convergence to equilibria described by Eq. (3a). Thus, the (p_b, p_u) pair for the first type will be independent of the introduction of the second type, and the (p'_b, p'_u) pair for the second type can be designed independently of the control pair for the first type. Moreover, so long as the robotic swarm has a sufficiently large size r_0, any number of disk types can be considered, and the design for each type will be independent of the design for every other type.

The catalytic reverse reactions in Eq. (1b) and (5b) allow for the multi-type decomposition described above. In general, regardless of the number of disk types, adding more disks to a finite volume reduces the free space available for movement. Consequently, the introduction of new disks increases the encounter rates between robots and disks. If, instead of a catalytic reverse reaction, a traditional timer-based reverse reaction was implemented, the mean binding time would have to be reduced on every robot to compensate for the increased robot density. That is, the increased density would lead to a higher binding event rate, and so the unbinding event rate would have to be similarly increased to maintain the desired equilibrium allocations. As discussed in Section 3.1, the catalytic reverse reactions allow for the control policy in Eq. (4) to be independent of density. Therefore, so long as the robotic swarm is sufficiently large, additional disk types can be added with no impact on existing control designs.

4 Model Predictions and Validation

The predicted equilibrium allocations for several e_b/e_u ratios and several values of avoidance distance a are shown in Fig. 3. In each graph, the horizontal axis is the so-called *control factor* $p_b/(p_b + p_u)$, which varies from 0 to 1. Following a convention discussed later in Section 5.2, $p_u \equiv 1$ for all control factors less than 0.5, and $p_b \equiv 1$ for all control factors greater than 0.5. Under this convention, there is a one-to-one mapping between control factors and (p_b, p_u) control policies. For many robot motion behaviors, the encounter-rate ratio e_b/e_u may be approximated by dividing the sum of the areas of a robot and an unbound zone sector by the area of an unbound zone sector alone. However, it can be estimated in general from fitting the predicted mean allocation curves in Fig. 3 to measured allocation data. Inferring this encounter-rate *ratio* is empirically much simpler than inferring actual encounter rates [21, 22].

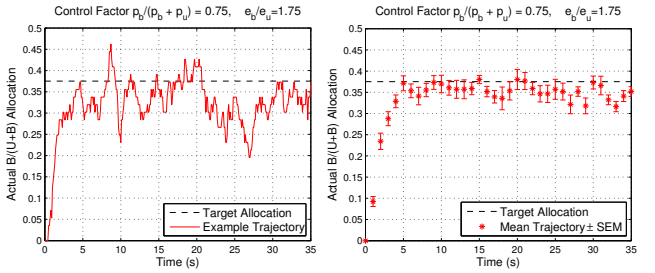


(a) Mean with $a = 0$ (b) Mean with $a = 0.5$

(c) Mean with $a = 0.8$ (d) Mean with $a = 2.0$

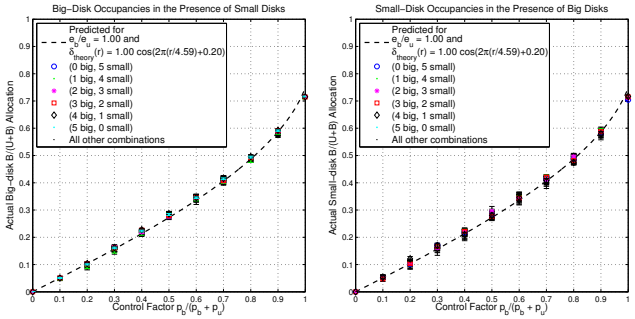
Fig. 3 Effect of encounter ratio and avoidance. For each e_b/e_u ratio and avoidance distance a, the plot shows the predicted relationship between the control factor $p_b/(p_b + p_u)$ and the predicted equilibrium mean allocation. The predictions of the mean come from Eq. (3b). The star in each plot represents the levels corresponding to the $(p_b, p_u) = (1, 0)$ case if robots could re-assort on boundaries to to optimally pack and eliminate wasted space. Consistent with expectations from the classical linear sequential parking problem, the curves in (a) all predict the parking constant of approximately 75% in the saturated case.

In our prior work [1], we used simulated experimental trials to show that this control strategy leads to equilibrium allocations for the single-type case that match the predictions in Fig. 3 across a wide range of environmental conditions. In particular, despite varying robot swarm size, robot radius, number of disks, and disk radius, the resulting control-factor–allocation plots fell on a single curve from Fig. 3 corresponding to a particular e_b/e_u ratio that varied only with individual robot parameters and not externally generated environmental parameters. Furthermore, we showed how robot motion primitives could be changed to skew the e_b/e_u ratio in predictable ways. In this extension, we validate our claims



(a) Single trajectory (b) 10-simulation average trajecotry

Fig. 4 Example simulated trajectories. In (a), a single execution of a simulation of 300 robots allocating to 3 disks is shown. In (b), the average trajectory is shown across 10 simulations. The (p_b, p_u) policy chosen was picked to achieve a mean $[B]/([B]+[U])$ allocation of 37.5%, shown as a dashed horizontal line.



(a) Big-disk allocations (b) Small-disk allocations

Fig. 5 Effect of varying disk mixture. Ten trials were generated for each experimental treatment (i.e., number of big disks, number of small disks, big allocation ratio, and small allocation ratio); the big-disk type is twice as large as the small-disk type. The averages shown for each big-disk allocation in (a) are taken across the pool of 110 trials that include all eleven small-disk allocation ratios for the corresponding big-disk allocation ratio and big–small mixture (and similar for the statistics in (b)). Error bars show ± 1 standard error of the mean (SEM). The broken line in the plot shows the expected allocation curve from theory.

from Section 3.3 that control policies for different boundary types can be designed in isolation and independently composed with no additional modifications necessary.

Following our earlier experimental protocol [1], simulations were conducted using NetLogo[1] [23], which allowed for simulating hundreds of mobile robots randomly interacting with each other and with disks of different types. Example simulated trajectories are shown in Fig. 4, which are effectively macroscopic step responses of a system of initially free robots allocating in order to achieve a desired bound–unbound ratio.

Simulations of 500 robots were conducted in arenas with varying mixtures of two different types of disk – "big" and "small" – at random locations throughout an arena. Each experimental arena has up to five of each of the two types of disk, and all (6^2-1) combinations of disk-type mixtures are explored. For each mixture and each disk type, experimental trials were generated for eleven different target allocation ratios. All eleven small-disk allocation ratios were paired with all eleven big-disk allocation ratios, and each pairing received ten simulation runs. Consequently, there are 110 experimental trials for each small-disk control policy and 110 experimental trials for each big-disk control policy. The resulting data are summarized in Fig. 5, which shows the control-factor–allocation curves for the big type (Fig. 5(a)) and the small type (Fig. 5(b)). Six disk mixtures are highlighted in the legend, but all mixtures are plotted together along with standard-error bars over the 110 trials for each control policy. Thus, because there are always sufficiently many free robots available to populate the desired occupancy levels, there is no effect of the number of disks or presence of additional types of disks. Moreover, as shown by the broken line, the data fall along the predicted curve for a unity e_b/e_u ratio.

In this experiment, the disk types are physically different – one type has a circumference of 28.27 robot arc lengths, and the other type has a circumference of 14.14 robot arc lengths. In practice, disk types need not be physically different; a disk's type is generally a detectable feature that allows for a robot to choose a different control policy when encountering the disk. However, because these disks differ by size, it is possible to see the effect of total number of zones on allocation variance. In particular, the standard-error bars in Fig. 5 are proportional to the sample variance across the 110 trials. The error bars in Fig. 5(b) are larger than the error bars in Fig. 5(a) because there are half as many zones on a small disk as there are on a large disk. As we will discuss in Section 5, although the CRN we use to describe this system is bimolecular, the relationship between number of zones and variance matches what is expected in a unimolecular CRN [6]. Moreover, catalytic reverse reactions provide a mechanism for implementing unimolecular CRN designs with increased robustness to environmental variations.

5 Reduced-Order Model and Its Applications

Let r_0 be the initial number of free robots. Assuming that no robots are initially bound, then r_0 is the total size of the swarm. For a large robotic swarm with $r_0 \gg B^*$, the concentration $[r]$ will essentially be constant. That is, while the concentrations $[B]$ and $[U]$ will move opposite to each other in relatively large swings, the fractional change in $[r]$ will be negligible [1]. Consequently, the bimolecular reactions in Eq. (1) can be viewed as the reversible unimolecular reaction

$$U \underset{p_u e_b r_0}{\overset{p_b e_u r_0}{\rightleftharpoons}} B, \tag{6}$$

which introduces the swarm size r_0 as a scalar on both reaction rate constants. This unimolecular approximation will also hold close to the equilibrium of Eq. (1). Thus, it also represents how the system returns to equilibrium after small perturbations, and its equilibrium matches the one predicted

[1] To obtain the code, contact Dr. Theodore Pavlic (tpavlic@asu.edu).

for the full bimolecular model. Moreover, the mean-field dynamics of this unimolecular CRN is a second-order linear time-invariant ODE, which allows for exponential characterizations of its convergence rate.

5.1 Convergence Rate and Variance

Unimolecular reversible reactions have been used in much of the stochastic task-allocation literature. For these systems [6, Section 3.2], the convergence rate is independent of the equilibrium distribution. For this system in particular, the time constant

$$\tau \triangleq \frac{1}{(p_b e_u + p_u e_b) r_0}. \tag{7}$$

So increasing the swarm size r_0 has no effect on the equilibrium distribution; however, it increases the speed of the system. Moreover, by similar reasoning for unimolecular reversible reactions [6, Section 3.2], the allocation variance is independent of the convergence rate. So maximizing the convergence rate of the system has no cost in terms of variance. Consequently, arbitrarily large swarms can be used to quickly achieve and maintain the equilibrium allocation with very little performance cost. The larger the swarm, the faster and more linear the system behaves. As shown in Fig. 5, variance decreases with the number of zones, which is consistent with unimolecular arguments about control of variance [6]. So the variance of the system will be improved by using robots which are relatively small with respect to the boundaries they allocate to. By using small robots, the number of zones necessarily increases, and the variance in allocation decreases accordingly.

5.2 Optimal Choice of Control Policy

We have discussed how increasing swarm size and number of zones has an effect on convergence rate and allocation variance without affecting the equilibrium mean allocation ratio. Furthermore, the control policy in Eq. (4) includes one degree of freedom over which (p_b, p_u) pairs can be chosen to optimize some feature of the system with no effect on mean allocation ratio. By Eq. (7), the convergence rate of the system can be maximized by making the sum $p_b + p_u$ as large as possible. Thus, for the fastest convergence to a desired $B^*/(U^*+B^*)$ equilibrium occupancy, the controls p_b and p_u should be chosen so that

$$(p_b, p_u) = \begin{cases} \left(\frac{e_b}{e_u} \frac{B^*}{U^*} \frac{(1+\delta)}{1-\delta B^*/U^*}, 1 \right) & \text{if } e_b B^* < e_u U^* \frac{1-\delta B^*/U^*}{1+\delta}, \\ \left(1, \frac{e_u}{e_b} \frac{U^*}{B^*} \frac{1-\delta B^*/U^*}{(1+\delta)} \right) & \text{otherwise.} \end{cases} \tag{8}$$

However, in an implementation of these strategies on real robots subject to mechanical fatigue or appreciable task-switching times, there may be some constraint which limits the $p_b + p_u$ sum.

Other Optimization Approaches: In the optimization approaches discussed above, we have shown how certain parameter choices of the system affect convergence rate and allocation variance without having an effect on mean occupancy itself. Nevertheless, some mean occupancies will have higher variances than others [6]. For example, the target allocations that correspond to extreme control policies $(p_b, p_u) \in \{(1,0),(0,1)\}$ will have relatively little variance at equilibrium and a relatively low convergence rate. However, the spatial avoidance parameter a has a predictable effect on the mean occupancy for each (p_b, p_u) pair (Fig. 3), and thus it can be used to alter the convergence rate and variance characteristics of each target allocation ratio. For example, for $(p_b, p_u) = (1,0)$, the parameter a can be chosen to compress the saturation occupancy well below the parking constant. Thus, with the appropriate choice of a, the variance of target allocations can be reduced at the possible cost of decreasing the convergence rate of the system. In our prior work, we described how the e_b/e_u ratio can be manipulated by changing sensor characteristics or motion primitives of each robot. Thus, this ratio also expands the optimization parameter space to achieve different target allocation ratios with improved variance or convergence rate.

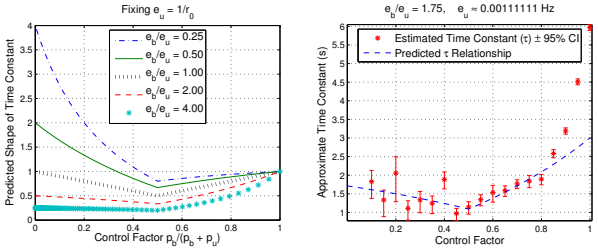
5.3 Absolute Encounter-Rate Estimation

Although we believe catalytic reverse reactions can greatly improve the design flow of a variety of stochastic task-allocation control strategies, there may still be problems for which control policies require an accurate estimation of encounter rates. In these cases, robots can be temporarily configured to implement a toy task-allocation problem like the one described in this paper. The resulting experimentally determined control-factor–allocation curve can be used to estimate encounter rates. This method is significantly simpler than estimating rates from sampling encounter times [21, 22].

The encounter-rate ratio e_b/e_u can be estimated geometrically or numerically using the curves in Fig. 3. Although τ described in Eq. (7) varies with e_b and e_u, its shape is set by the ratio e_b/e_u. These shapes are shown in Fig. 6(a). Once the time-constant shape is fixed, there is one additional degree of freedom to fit the actual time-constant curve to data. Thus, as shown in Fig. 6(b), the expected time-constant shape can be fit to measured convergence-rate data to estimate the actual values of e_b (or, equivalently, e_u).

6 Conclusions and Future Work

We have extended our prior work to the synthesis of communication-free stochastic control policies for robotic swarms that establish and maintain desired task allocations which are robust to changes in the environment. Moreover, our simulation results validate that design for multiple task types can be decomposed into isolated simpler single-type problems whose solutions are independent of each other. Finally, we have explored how the system provides several degrees of freedom for optimization of convergence rate and task-allocation variance. More generally, we have used a simple boundary-allocation problem to demonstrate



(a) Predicted τ shapes (b) Simulated τ Data

Fig. 6 Effect of encounter ratio on time constant. The time constant τ described in Eq. (7) cannot be predicted without knowledge of encounter rates e_u and e_b. However, the shape of τ is set by the ratio only e_b/e_u. In (a), e_u is arbitrarily assumed to be $1/r_0$ and the ratio e_b/e_u is varied. In (b), time-constant data are shown that were generated from simulations of 300 robots across a range of control factors. For each factor, 10 trials were combined into an average step response, as in Fig. 4(b), which was then used to estimate a corresponding time constant. Using an e_b/e_u ratio estimated from allocation data, a corresponding e_u was fit to the time-constant data.

how specially designed irreversible processes can be used to implement higher level reversible processes that are less sensitive to environmental variations than the reversible processes typically studied in traditional statistical mechanics.

In our stochastic control approach, robots continue to bind and unbind randomly even after the equilibrium allocation has been reached, which apparently wastes a significant amount of power. This task switching at equilibrium, and its attendant power consumption, is a property of most stochastic strategies that already exist in the literature [e.g., 3–11]. In principle, a deterministic strategy could achieve the same distributions of robots to boundaries (possibly at allocations above the 75% parking constant) and then trigger robots to switch out of the costly allocation mode. However, as has been discussed elsewhere for a particular deterministic–stochastic comparison [2], the deterministic approach would likely require significantly more communication between robots and additional sensing and navigation capabilities. Our stochastic approach allows for simpler robots that actually require *less* power and hardware for communication and sensing. In fact, the weight and physical space that is not required for communication and sensing may be used for additional battery capabilities. Additionally, our stochastic approach is insensitive to the communication and navigational errors that would be a significant concern for a deterministic approach. Furthermore, the continued operation of the stochastic allocation process allows the system to detect newly added boundaries and allocate to them automatically. Consequently, although the stochastic strategy we employ may have some performance weaknesses, it compensates for them by being less costly to implement, less sensitive to sensing and actuation errors, and adaptive to disturbances and dynamic changes in the environment.

In addition, our stochastic strategy may have certain advantages over deterministic strategies in reducing steady-state power consumption. A deterministic approach will not necessarily come to a static equilibrium, either due to operational constraints or due to the design of the algorithm. For example, many autonomous air vehicles must stay in constant motion to remain in flight. Thus, surveillance drones flying through different regions will naturally "bind" and "unbind" to their assignments as part of an operational constraint. In fact, the frequent movement of vehicles from one region to another may help to reduce their likelihood of being detected. Additionally, there are many deterministic algorithms that do not lead to static equilibria. For example, primal-space parallel distributed optimization algorithms [24] are both deterministic and widely used, but they usually lead to a cyclic steady-state behavior that oscillates around an optimal solution. Like the continued task switching at equilibrium in the stochastic approaches, these steady-state oscillations entail power consumption that is not directly applied to task execution. Moreover, not all such deterministic strategies have convergence rates that are as well characterized as those in our stochastic strategy. Because our system can be shown to converge to equilibrium at an exponential rate, robots can be safely switched out of their energetically costly allocation mode after a time when it is very likely that their allocations have reached a specified distance from the target equilibrium (e.g., after three time constants). This timed event could be internally generated or externally generated by a signal broadcast to all robots from a central supervisor.

In future work, we will better characterize the mean inter-robot distance and thereby move closer to solving the generalized sequential parking problem. Additionally, we will show how irreversible catalytic implementations of reversible processes can improve the robustness of other applications of stochastic robotics, such as stochastic self-assembly. Finally, we intend to validate the results of this work on a physical multi-robot testbed in the near future.

Acknowledgements

Comments by Simon DeDeo regarding the models used here have been helpful in framing our thoughts about the implications of this control design approach. Interaction with him was possible due to a workshop on Information, Complexity, and Life organized by the BEYOND Center for Fundamental Concepts in Science at Arizona State University. We are also grateful for the helpful comments of two anonymous referees. This work was supported in part by the National Science Foundation (grant no. CCF-1012029).

References

[1] Pavlic, T. P., Wilson, S., Kumar, G. P., and Berman, S., 2013. "An enzyme-inspired approach to stochastic allocation of robotic swarms around boundaries". In Proceedings of the 16th International Symposium on Robotics Research (ISRR 2013).

[2] Dantu, K., Berman, S., Kate, B., and Nagpal, R., 2012. "A comparison of deterministic and stochastic approaches for allocating spatially dependent tasks in micro-aerial vehicle collectives". In Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 793–800.

[3] Correll, N., and Martinoli, A., 2004. "Modeling and optimization of a swarm-intelligent inspection system". In Proceedings of the Seventh International Symposium on Distributed Autonomous Robotics Systems (DARS 2004), pp. 369–378.

[4] Berman, S., Halász, Á., Hsieh, M. A., and Kumar, V., 2009. "Optimized stochastic policies for task allocation in swarms of robots". *IEEE Trans. Robot.,* **25**(4), August, pp. 927–937.

[5] Odhner, L. U., Ueda, J., and Asada, H. H., 2008. "Feedback control of stochastic cellular actuators". In Proceedings of the 10th International Symposium on Experimental Robotics, O. Khatib, V. Kumar, and D. Rus, eds., Vol. 39 of *Springer Tracts in Advanced Robotics*, pp. 481–490.

[6] Odhner, L. U., and Asada, H., 2010. "Stochastic recruitment control of large ensemble systems with limited feedback". *J. Dyn. Syst., Meas., Control.,* **132**(4), July.

[7] Mather, T. W., and Hsieh, M. A., 2011. "Distributed robot ensemble control for deployment to multiple sites". In Proceedings of Robotics: Science and Systems VII.

[8] Matthey, L., Berman, S., and Kumar, V., 2009. "Stochastic strategies for a swarm robotic assembly system". In Proceedings of the 2009 IEEE International Conference on Robotics and Automation, pp. 1953–1958.

[9] Napp, N., Burden, S., and Klavins, E., 2009. "Setpoint regulation for stochastically interacting robots". In Proceedings of Robotics: Science and Systems V.

[10] Napp, N., and Klavins, E., 2011. "A compositional framework for programming stochastically interacting robots". *Int. J. Robot. Res. [Special Issue Stochasticity in Robot. Bio-Systems Part 2],* **30**(6), May, pp. 713–729.

[11] Berman, S., Nagpal, R., and Halász, Á., 2011. "Optimization of stochastic strategies for spatially inhomogeneous robot swarms: a case study in commercial pollination". In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3923–3930.

[12] Rényi, A., 1958. "On a one-dimensional problem concerning random space-filling". *Publ. Math. Inst. Hung. Acad. Sci.,* **3**, pp. 109–127.

[13] Solomon, H., and Weiner, H., 1986. "A review of the packing problem". *Commun. in Stat. - Theory Methods,* **15**(9), pp. 2571–2607.

[14] Finch, S. R., 2003. *Mathematical Constants*, Vol. 94 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press.

[15] Kumar, G. P., Buffin, A., Pavlic, T. P., Pratt, S. C., and Berman, S. M., 2013. "A stochastic hybrid system model of collective transport in the desert ant *Aphaenogaster cockerelli*". In Proceedings of the 16th ACM International Conference on Hybrid Systems: Computation and Control, pp. 119–124.

[16] Wilson, S., Pavlic, T. P., Kumar, G. P., Buffin, A., Pratt, S. C., and Berman, S., 2014. "Design of ant-inspired stochastic control policies for collective transport by robotic swarms". *Swarm Intell.* Submitted.

[17] Klavins, E., Burden, S., and Napp, N., 2006. "Optimal rules for programmed stochastic self-assembly". In Proceedings of Robotics: Science and Systems II.

[18] Berman, S., Kumar, V., and Nagpal, R., 2011. "Design of control policies for spatially inhomogeneous robot swarms with application to commercial pollination". In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, pp. 378–385.

[19] Langmuir, I., 1918. "The adsorption of gases on plane surfaces of glass, mica and platinum". *J. Am. Chem. Soc.,* **40**(9), pp. 1361–1403.

[20] Hespanha, J. P., and Singh, A., 2005. "Stochastic models for chemically reacting systems using polynomial stochastic hybrid systems". *Int. J. Robust Nonlinear Control.,* **15**(15), October, pp. 669–689.

[21] Hutchinson, J. M. C., and Waser, P. M., 2007. "Use, misuse and extensions of "ideal gas" models of animal encounter". *Biol. Rev.,* **82**(3), August, pp. 335–359.

[22] Gurarie, E., 2008. "Models and analysis of animal movements: From individual tracks to mass dispersal". PhD thesis, University of Washington.

[23] Wilensky, U., 1999. *NetLogo*. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, USA.

[24] Bertsekas, D. P., and Tsitsiklis, J. N., 1997. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, Belmont, Massachusetts.