

# The Impact of Migration Topology on the Runtime of Island Models in Dynamic Optimization

Andrei Lissovoi  
Department of Computer Science  
University of Sheffield  
Sheffield, S1 4DP, United Kingdom

Carsten Witt  
DTU Compute  
Technical University of Denmark  
Kongens Lyngby, Denmark

## ABSTRACT

We introduce a simplified island model with behavior similar to the  $\lambda(1+1)$  islands optimizing the MAZE fitness function, and investigate the effects of the migration topology on the ability of the simplified island model to track the optimum of a dynamic fitness function. More specifically, we prove that there exist choices of model parameters for which using a unidirectional ring as the migration topology allows the model to track the oscillating optimum through  $n$  MAZE-like phases with high probability, while using a complete graph as the migration topology results in the island model losing track of the optimum with overwhelming probability. Additionally, we prove that if migration occurs only rarely, denser migration topologies may be advantageous. This serves to illustrate that while a less-dense migration topology may be useful when optimizing dynamic functions with oscillating behavior, and requires less problem-specific knowledge to determine when migration may be allowed to occur, care must be taken to ensure that a sufficient amount of migration occurs during the optimization process.

## 1. INTRODUCTION

Optimization problems are often dynamic in nature, as the environment in which they have to be solved often changes with the passing of time. Nature-inspired algorithms are based on approaches to solving optimization problems observed in nature, and we might therefore hope that they would also provide a reasonable solution to coping with dynamic changes in optimization problems. The performance of nature-inspired algorithms on dynamic problems has been considered in the literature [17, 1], including a number of runtime analyses of evolutionary algorithms on dynamic problems [5, 8, 18, 9, 10, 3].

In a dynamic optimization problem, the optimum is allowed to move in the search space over time, as conditions of the problem change. The goal of the evolutionary algorithm is then not only to locate the optimum once, as in the case of static optimization problems, but also be able to

track the optimum as it moves, maintaining good solutions over time.

The MAZE fitness function, introduced in [11], is an artificial fitness function consisting of  $n + 1$  long oscillating phases, over the course of which the optimum slowly shifts from the all-ones bit string to the all-zeros bit string, while oscillating between two specific solutions during each phase. In [11], it is shown that a simple  $(1+1)$  EA is not able to track the oscillating optimum through all  $n$  phases. Subsequent work [14, 15] has considered how various diversity mechanisms impact the ability of evolutionary algorithms to track the optimum of this function, observing that an island model can provide the necessary diversity as long as migration on a complete migration topology is carefully managed, and does not occur too close to a MAZE phase transition – conditions which require somewhat specific knowledge of the fitness function, which may not be available for other problems.

In addition to literature considering the impact of migration intervals on the performance of parallel algorithms [16, 15], there is also previous work highlighting the impact of the migration topology [2, 12, 19] on the optimization process.

In this paper, we investigate whether using a less-dense migration topology, such as a unidirectional ring, can be beneficial on a dynamic problem like MAZE, allowing some of the requirements on when migration is allowed to occur to be relaxed. Intuitively, decreasing the density of the migration topology weakens the negative effect of migration on population diversity, and may allow good solutions to survive migration occurring at inopportune times.

We base our analysis on a simplified version of the island model studied in [15], which incorporates the major elements of the original setting: an oscillating fitness function, islands performing independent mutation/selection steps, and the effect of MAZE phase transitions on the islands' ability to track the optimum based on their current-best individuals at the time of the transition. The simplified model incorporates more randomization, as both the oscillating pattern and migration are randomized, which both simplifies the analysis, and disallows some of the more artificial solutions possible in the original model, such as only performing migrations on iterations that assign a higher fitness value to the desirable solution.

Using this simplified model, we prove, through rigorous runtime analysis, that the unidirectional ring migration topology allows the island model to track the optimum of the dynamic fitness functions in some settings where the complete

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

GECCO '16, July 20 - 24, 2016, Denver, CO, USA

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-4206-3/16/07...\$15.00

DOI: <http://dx.doi.org/10.1145/2908812.2908843>

| Phase          | 0       | 1             | 2             | 3             | ... | $n$           |
|----------------|---------|---------------|---------------|---------------|-----|---------------|
| $\text{OPT}_p$ | $(1^n)$ | $0^1 1^{n-1}$ | $0^2 1^{n-2}$ | $0^3 1^{n-3}$ | ... | $0^n$         |
| $\text{ALT}_p$ |         | $1^n$         | $0^1 1^{n-1}$ | $0^2 1^{n-2}$ | ... | $0^{n-1} 1^1$ |

**Figure 1: The Maze dynamic fitness function: in  $n$  oscillating phases after an initial OneMax phase, two bit strings,  $\text{OPT}_p$  and  $\text{ALT}_p$  are assigned higher-than-OneMax fitness values. The optimum oscillates in an  $\text{OPT}_p$ - $\text{OPT}_p$ - $\text{ALT}_p$  pattern, with  $f(\text{OPT}_p) > f(\text{ALT}_p)$  two iterations out of three.**

migration topology does not, as well as a converse result which applies if migration does not occur frequently enough.

This paper is structured as follows. In the next section, we introduce the simplified island model, and highlight its key differences by comparing to the setting of [15], as well as some of the tools used in subsequent proofs. Sections 3.1 and 3.2 consider the case of migration occurring in every iteration, the former proving that a complete migration topology leads to a failure to track the optimum, while the latter proves that switching to the unidirectional ring topology allows tracking the optimum with high probability. Sections 4.1 and 4.2 consider the effects of very infrequent migrations, proving that in such settings, a denser migration topology may aid in tracking the oscillating optimum. We finish with some conclusions, as well as a discussion of further possibilities for analysis.

## 2. PRELIMINARIES

In order to analyze the impact of migration topology on the island model behavior, and remove some of the artifacts arising from the MAZE fitness function (summarized in Figure 1), we will construct a somewhat simplified model of the optimization algorithm, while maintaining similarities to  $\lambda$  islands using (1+1) EAs to optimize MAZE. The simplified model is shown as Algorithm 1 below; this section discusses the key differences. We note that the original MAZE fitness function has been used to show that there are advantages for the dynamic fitness functions that are easier for ant colony-based algorithms to optimize (compared to the (1+1) EA), and as its rapid oscillations favouring a specific individual can more easily be represented in a pheromone memory (versus the single ancestor individual of a (1+1) EA).

Some changes have been made to the model of the MAZE fitness function: instead of a fully deterministic oscillation defined on bit strings, Algorithm 1 distinguishes between three states, OPT, ALT, and LOST, and randomly selects which of OPT and ALT has a higher fitness value, independently favoring OPT over ALT in each iteration with probability  $p_{\text{OPT}}$ . When a MAZE phase transition occurs, all islands in the OPT state transition to the ALT state, while all other islands transition to the LOST state. The OPT, ALT and LOST states thus correspond to having OPT, ALT, and OneMax-valued individuals in the original MAZE, where the OPT individual in each phase becomes the ALT individual of the next phase, while the ALT individual becomes a OneMax-valued individual following a phase transition.

Each island behaves like a simplified (1+1) EA, maintaining a current-best solution  $x_i^*(t)$  by applying mutation and selection. The mutation operator is simplified to allow construction of OPT from ALT (or vice versa) with proba-

bility  $p_{\text{mut}}$ , while preventing construction of either OPT or ALT from LOST. With an appropriate choice of  $p_{\text{mut}}$  based on a probability of a specific single-bit mutation occurring, this is a pessimistic model of (1+1) EA's behavior on MAZE, where, in the later phases, beginning a phase with a OneMax valued individual (i.e., in the LOST state in the simplified model) would cause the (1+1) EA to revert to optimizing OneMax with at least constant probability, leaving it with an overwhelmingly small probability of finding the oscillating optimum again.

Additionally, migration is randomized by allowing it to occur in each iteration independently at random with probability  $p_{\text{mig}}$ . This, as well as the randomization of which of ALT and OPT is favored in any given iteration, prevents the algorithm from ensuring that it performs migration only on OPT-favoring iterations.

---

**Algorithm 1** Simplified island model for Maze with a directed graph  $G = (V, \mathcal{A})$  used as the migration topology.

---

Initialize  $x_i^*(1) \leftarrow \text{OPT}$  for all  $i \in V$ .

**for**  $t \leftarrow 1, 2, \dots$  **do**

    With probability  $p_{\text{OPT}}$ ,  $y^+ \leftarrow \text{OPT}$ ,  $y^- \leftarrow \text{ALT}$ ; otherwise, vice versa.

$M \leftarrow \text{bernoulli}(p_{\text{mig}})$      $\triangleright$  Does migration occur now?

**for all**  $i \in V$  **in parallel do**

        Let  $N = \{x_i^*(t)\} \cup \{x_j^*(t) \mid M = 1 \wedge (j, i) \in \mathcal{A}\}$

$x'_i \leftarrow \begin{cases} y^+ & \text{if } y^+ \in N \\ y^- & \text{if } y^+ \notin N \wedge y^- \in N \\ \text{LOST} & \text{otherwise} \end{cases}$

$x_i^*(t+1) \leftarrow \begin{cases} y^+ & \text{with probability } p_{\text{mut}} \text{ if } x'_i = y^- \\ x'_i & \text{otherwise} \end{cases}$

**if**  $t \bmod t_0 = 0$  **then**     $\triangleright$  A phase transition occurs

**for all**  $i \in V$  **do**

$x_i^*(t+1) \leftarrow \begin{cases} \text{ALT} & \text{if } x_i^*(t+1) = \text{OPT} \\ \text{LOST} & \text{otherwise} \end{cases}$

---

Summarized concisely, Algorithm 1 models  $\lambda$  islands performing mutation in every iteration, each one updating its current-best individual  $x_i^*(t+1)$  while the dynamic fitness function oscillates randomly assigning higher fitness values to OPT and ALT. When migration occurs, each island receives the previous iteration's current-best individuals from all of the islands that have are connected to it by a directed edge in the migration topology  $G = (V, \mathcal{A})$ , and uses the best of these individuals and its own  $x_i^*(t)$  (evaluated in the current iteration) as the ancestor for the current iteration. Thus, the parameters of the simplified model are:

- $n$ , the number of phases being considered,
- $t_0$ , the number of iterations in each phase,
- $p_{\text{OPT}}$ , the probability of OPT having a higher fitness value than ALT in an iteration,
- $p_{\text{mut}}$ , the probability of constructing OPT from ALT and vice versa,
- $p_{\text{mig}}$ , the probability of migration occurring in an iteration,
- $\lambda$ , the number of islands,
- $G = (V, \mathcal{A})$ , a directed graph specifying the migration topology, with  $V$  being the set of vertices (islands, and therefore  $|V| = \lambda$ ), and  $\mathcal{A}$  a set of directed arcs specifying how migration transfers current-best individuals.

In this paper, we consider the impact of the migration topology  $G$  on the algorithm's ability to track the oscillating OPT/ALT optimum (more specifically, whether at least one island remains in a non-LOST state after  $n$  oscillating phases).

The following choice of parameters yields a setting similar to the original MAZE considered in [11, 15]:  $t_0 = n^3$ ,  $p_{\text{OPT}} = 2/3$ ,  $\lambda = \Omega(\log n)$ ,  $p_{\text{mut}} = \Theta(1/n)$ ,  $p_{\text{mig}} = 1/\tau$  (where  $\tau$  is the deterministic migration interval), and  $G = K_\lambda$ . In [11],  $t_0 = kn^3$  for a constant  $k > 0$  is used to provide an ant colony-based algorithm with sufficient time to adjust its pheromone memory; we somewhat relax the choices for this parameter in our results, requiring most often that it is in  $\Omega(n^2)$  and polynomial with respect to  $n$ . Generally, a constant  $1/2 < p_{\text{OPT}} < 1$  is required to allow the ant colony to adjust to the solution useful in the next iteration.

It is worth noting that in the original MAZE setting,  $n$  serves as both the number of bits the individuals are composed of, and the number of oscillating phases in the MAZE function. This motivates the relationship between  $n$ ,  $p_{\text{mut}}$ , and  $\lambda$  which persists even in the simplified setting: although the simplified setting no longer deals with  $n$ -bit strings directly, it is still serving as a model for islands using actual  $(1+1)$  EAs on  $n$ -bit strings.

To derive our theoretical results, we use the following drift theorem, describing the expectation of first-hitting time in the presence of additive drift.

**THEOREM 1** (ADDITIVE DRIFT, EXPECTED TIME, [13, 7]).

Let  $(X^t)_{t \geq 0}$ , be a stochastic process over a bounded state space  $S \subseteq \mathbb{R}_0^+$ , and  $\mathcal{F}_t$  a filtration to which the stochastic process is adapted (e.g., the natural filtration) and let  $T_0 := \min\{t \geq 0 : X^t = 0\}$  denote the first hitting time of 0 and assume that both  $\mathbb{E}(X^0)$  and  $\mathbb{E}(T_0 | X^0)$  are finite. Then, if

$$\mathbb{E}(X^t - X^{t+1} | \mathcal{F}_t; X^t > 0) \geq \epsilon,$$

it holds that  $\mathbb{E}(T_0 | X^0) \leq X^0/\epsilon$ .

Additionally, the classical gambler's ruin problem [6] is used to bound the probability that a process that shrinks in expectation grows to a particular size in Lemma 5. In the canonical setting, this would be equivalent to determining the probability that a gambler who starts with a single coin is able to collect a certain number of coins in an unfair coin flipping game, where he is more likely to lose a coin than win a coin in each round.

### 3. FREQUENT MIGRATION

As a simple case, consider setting  $p_{\text{mig}} = 1$ , i.e., requiring migration to occur in every iteration. We consider the impact of two extreme choices of  $G$  in this setting: a complete graph  $K_\lambda$ , and a  $\lambda$ -vertex unidirectional ring.

#### 3.1 Complete migration topology

We first prove that using the complete migration topology with migration occurring in every iteration results in the simplified model being unable to track the optimum of the MAZE through all  $n$  phases.

**THEOREM 2.** When  $t_0 \in \Omega(n) \cap O(\text{poly}(n))$ ,  $0 < p_{\text{OPT}} < 1$  is a constant,  $p_{\text{mut}} = 1/(en)$ ,  $p_{\text{mig}} = 1$ ,  $\lambda = O(n)$ , and  $G$

is a  $\lambda$ -vertex complete graph, i.e., in the case of using the complete migration topology with migration occurring in every iteration, the probability that all islands are in the LOST state after  $n \cdot t_0$  iterations is  $1 - 2^{-\Omega(n)}$ .

**PROOF.** We note that at least one mutation occurs during a phase with at least a constant probability:

$$1 - (1 - p_{\text{mut}})^{\lambda t_0} \geq 1 - e^{-c}$$

and the probability that no mutation occurs in a single iteration is also at least a constant:

$$(1 - p_{\text{mut}})^\lambda \geq e^{-c'}$$

where  $c > 0$  and  $c' > 0$  are constants.

Thus, with at least a constant probability, the last mutation in a phase occurs at least one iteration before the phase transition. With probability  $(1 - p_{\text{OPT}})^2$ , i.e., at least a constant probability, both the iteration when the last mutation occurs, and the iteration immediately following it favor ALT over OPT; thus, if all islands were in the OPT state, the mutation would produce an ALT individual which would migrate to all islands, while if at least one island was in the ALT state, its original individual would migrate to all other islands. As no further mutation occurs before the phase transition, we conclude that each phase has at least a constant probability of ending with all islands having the ALT individual, and thus losing track of the oscillating optimum following the next phase transition.

Thus, if each of  $n$  phases has at least a constant probability of failing  $p_f > 0$ , the probability that at least one of  $n$  phases ends with all islands in the LOST state is at least  $1 - (1 - p_f)^n = 1 - 2^{-\Omega(n)}$ .  $\square$

It is worth noting that this proof approach is flexible enough to be adapted to settings where migration occurs less often, such as once in every constant number of iterations. The proof of Theorem 2 essentially relies on mutations occurring in at most a constant number of iterations following the final migration. When  $p_{\text{OPT}}$  is at most a constant smaller than 1, and this condition on the number of iterations with mutation holds, all mutations that occur following the final migration are rejected with at least a constant probability, which allows all islands to remain in the ALT state, and therefore become LOST following the phase transition.

#### 3.2 Unidirectional ring topology

For the other extreme, suppose that  $G$  is minimally connected, i.e., a unidirectional ring of  $\lambda$  vertices and  $\lambda$  arcs. This reduces the effect of migration on the island memory, making it impossible to propagate an undesirable individual to all islands in a single migration. In this section, we will prove that the simplified island model is able to track the oscillating optimum for the full  $n$  phases.

**THEOREM 3.** When  $t_0 \in \Omega(n^2) \cap O(\text{poly}(n))$ ,  $p_{\text{OPT}} = 2/3$ ,  $p_{\text{mut}} = 1/(en)$ ,  $p_{\text{mig}} = 1$ ,  $\lambda = c \log n$ , where  $c > 0$  is a sufficiently large constant, and  $G$  is a  $\lambda$ -vertex unidirectional ring, the simplified island model is able to track the oscillating optimum for at least  $n$  phases with high probability.

We will prove this by showing that as long as each phase begins with at least one island still tracking the optimum, the phase will end with at least one island having  $x_i^*(t) = \text{OPT}$ . Notably, any constant  $p_{\text{OPT}} > 1/2$  is sufficient, and

$p_{\text{OPT}} = 2/3$  was chosen to correspond to the oscillation pattern of the original MAZE.

LEMMA 4. Let  $t_0 \in \Omega(n^2) \cap O(\text{poly}(n))$ ,  $p_{\text{OPT}} = 2/3$ ,  $p_{\text{mut}} = 1/(en)$ ,  $p_{\text{mig}} = 1$ ,  $\lambda = c \log n$ , where  $c > 0$  is a sufficiently large constant, and  $G$  be a  $\lambda$ -vertex unidirectional ring. If a phase begins with at least one island having  $x_i^*(t) \neq \text{LOST}$ , there will with high probability exist an iteration  $t'$  before the phase ends such that all islands will have  $x_i^*(t') = \text{OPT}$ .

PROOF. We note that after at most  $\lambda$  iterations, no islands will be in the LOST state – in the worst case,  $\lambda$  iterations are enough to migrate the non-LOST individual from the single surviving island to all other islands, with fewer iterations being required if there is more than one surviving island.

Let  $t'$  be the iteration during which no islands are in the LOST state, and consider the drift in  $X_t$ , the number of islands having  $x_i^*(t + t') = \text{ALT}$ . Let  $S_t$  be the number of arcs  $(u, v) \in \mathcal{A}$  in the migration topology for which it holds that  $x_u^*(t + t') = \text{OPT}$  and  $x_v^*(t + t') = \text{ALT}$ , i.e., the number of segments in the unidirectional ring of the migration topology which are composed of islands having OPT as their current-best solution. We note that when migration occurs,  $S_t$  islands change their current-best solution because of migration – there are  $S_t$  segments of OPT islands in the migration topology; the first island in each segment will receive ALT through migration, while the islands following each segment will receive OPT through migration, so regardless of which solution is favored in an iteration, migration will cause  $S_t$  islands to switch state. The remaining islands, if they are in the non-favored state, may change their current-best solution through mutation. The expected change in  $X_t$  is a combination of these effects:

$$\begin{aligned} E(X_t - X_{t+1} \mid X_t < \lambda, S_t) &\geq p_{\text{OPT}}(p_{\text{mig}}S_t + p_{\text{mut}}(\lambda - X_t - S_t)) \\ &\quad - (1 - p_{\text{OPT}})(p_{\text{mig}}S_t + p_{\text{mut}}(X_t - S_t)) \\ &> 2p_{\text{OPT}}p_{\text{mig}}S_t - p_{\text{mig}}S_t - p_{\text{mut}}\lambda \\ &> \Omega(1) - p_{\text{mut}}\lambda = \Omega(1) \end{aligned}$$

where the inequalities reflect lower-bounding the positive contribution of mutation as 0, upper-bounding  $(X_t - S_t) < \lambda$  in the negative contribution of mutation, and recalling that  $p_{\text{mig}} = 1$  and  $2p_{\text{OPT}} > 1$ .

When the iteration begins with no islands holding OPT, a drift toward  $X_t = 0$  exists,

$$E(X_t - X_{t+1} \mid X_t = \lambda) = p_{\text{OPT}}p_{\text{mut}}\lambda = \frac{2c \log n}{3en}$$

and can be used as a lower bound on the overall drift throughout the whole process.

Applying the additive drift theorem, the expected first hitting time  $T = \min\{t : X_t = 0\} = O(\lambda / \frac{2c \log n}{3en}) = O(n)$ . As this is much shorter than the phase length  $t_0 \in \Omega(n^2)$ , we can conclude that  $X_t = 0$  is hit during the phase with high probability (by applying a Markov bound on the probability that the first hitting time exceeds twice the expectation, and repeating the argument  $n$  times), and hence at least at some point during the phase, all islands have OPT as their current-best solution.  $\square$

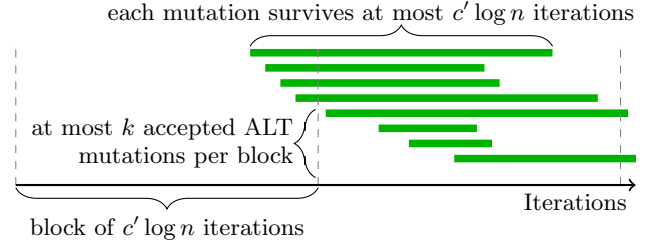


Figure 2: Lemma 5: the oscillating phase is divided into blocks of  $c' \log n$  iterations; with high probability, no more than  $k = c' \ln n$  OPT-to-ALT mutations are accepted within each block, each one creating an ALT segment that disappears after at most  $c' \log n$  iterations with high probability. Thus, with high probability, at most  $2k$  ALT segments may be active at the same time.

We then need to show that it is not likely that the island model will manage to replace OPT with ALT on all islands during the remainder of the current phase.

LEMMA 5. Let  $t_0 \in \Omega(n^2) \cap O(\text{poly}(n))$ ,  $p_{\text{OPT}} = 2/3$ ,  $p_{\text{mut}} = 1/(en)$ ,  $p_{\text{mig}} = 1$ ,  $\lambda = c \log n$ , where  $c > 0$  is a sufficiently large constant, and  $G$  be a  $\lambda$ -vertex unidirectional ring. If there occurs an iteration where  $x_i^*(t) = \text{OPT}$  for all islands  $i$ , with high probability, an iteration where for all islands  $i$ ,  $x_i^*(t) \neq \text{OPT}$  does not occur before the next phase transition.

PROOF. We note that it is difficult to apply a negative drift theorem directly in this setting, as the drift would depend on  $S_t$ : if there are many OPT/ALT boundaries in the migration topology, migration may cause drastic changes in the number of islands having OPT as their current-best individual. Instead, our strategy is to bound the number of islands having ALT as their current-best individual by considering the effects of each OPT-to-ALT mutation that occurs in isolation, i.e., as if it created the only ALT segment around at any specific time. An upper bound on the total number of islands having ALT as their current-best solution at any one time can then be derived from bounds on the maximum length each isolated ALT segment may reach, the number of iterations isolated ALT segments survive, and the rate at which such segments are created.

When considered in isolation, an OPT-to-ALT mutation creates an ALT segment with initial length 1 in the migration topology. We only consider its length to be modified by migration: it increases by 1 if migration occurs during an ALT-favoring iteration, and decreases by 1 if migration occurs during an OPT-favoring iteration; any further OPT-to-ALT mutations would be treated as separate isolated segments, and pessimistically, no ALT-to-OPT mutations occur within the ALT segment. There is a tendency towards decreasing the length. Instead of a drift theorem, here we even can directly apply results on the classical gambler's ruin problem [6] to bound the maximum length of such an isolated segment: it decreases with probability at least  $2/3$ , as  $p_{\text{OPT}} = 2/3$ , the maximal possible change is by 1 in either direction, and the probability of increasing is at most  $1/3$ . Thus, using the ruin problem on  $\{0, \dots, \ell\}$  with starting state 1, the probability that the length of an ALT segment

exceeds  $\ell = c' \log n$ , where  $c' > 0$  is constant is no more than

$$\frac{(1/2)^\ell}{1 - (1/2)^{\ell+1}} \leq 2^{-\ell+1} = O(n^{-c'}). \quad (1)$$

Next, consider the number of iterations before migration reduces a freshly-created ALT segment to length 0. This can be upper-bounded using a tail bound on the binomial distribution: the segment is guaranteed to be reduced to length 0 if, in  $2k$  iterations, more than  $k$  favor OPT. Let  $X_{2k}$  be the number of iterations that favor OPT of  $2k$  iterations:

$$P(X_{2k} \leq k) \leq \exp\left(-2 \frac{(2kp_{\text{OPT}} - k)^2}{k}\right) = e^{-2k/9}$$

using Hoeffding's inequality and recalling  $p_{\text{OPT}} = 2/3$ . Setting also  $k = c' \log n$  for large enough but constant  $c' > 0$ , we conclude that with probability  $1 - n^{-\Omega(c')}$ , an OPT-to-ALT mutation disappears after  $O(\log n)$  iterations. We note that in total, the expected number of OPT-to-ALT mutations within a phase is at most  $(1 - p_{\text{OPT}})p_{\text{mut}}\lambda t_0 = O(t_0 \log(n)/n)$  since  $p_{\text{mut}} = 1/(ne)$ , so by a straightforward union bound on the probabilities of an ALT iteration surviving more than  $O(\log n)$  iterations (if  $c'$  large enough), none of the OPT-to-ALT mutations that occur in the considered interval survive for more than the desired number of iterations with high probability.

Finally, we need to show that the rate at which OPT-to-ALT mutations are accepted is low enough to allow any accepted mutations to dissolve through migration without overrunning the island model. To that end, we can bound  $Y_k$ , the number of OPT-to-ALT mutations that are accepted within  $k = c' \log n$  iterations using a Chernoff bound:

$$E(Y_k) < k\lambda p_{\text{mut}} = O((\log^2 n)/n) = o(1)$$

$$P(Y_k \geq c' \ln n) \leq e^{-c' \ln n} = n^{-c'}$$

recalling that  $\lambda = O(\log n)$ ,  $p_{\text{mut}} = 1/(ne)$ , and ignoring the possibility that some of these mutations occur during iterations which assign a higher fitness value to OPT, and therefore would not be accepted.

Thus, no more than  $c' \ln n$  OPT-to-ALT mutations are accepted during a  $c' \log n$  iteration period with high probability, and all accepted mutations disappear after  $c' \log n$  iterations with high probability. By dividing the MAZE phase into blocks of  $c' \log n$  iterations each, as illustrated in Figure 2, we can conclude that with high probability, at most  $2 \cdot c' \log n = O(\log n)$  OPT-to-ALT segments can be active at the same time: with high probability, no more than  $c' \log n$  appear at the exact end of an  $c' \log n$  iteration block, and no more than  $c' \log n$  appear during the next block, with the former group all being reduced to length 0 before the next-next block begins.

We are finally ready to bound the total number of islands that can have ALT as their best-so-far individual at the same time: denoting by  $k$  the number of segments consisting of ALT-individuals, we define  $L_i$  as the length of the  $i$ -th segment. We are interested in  $S := \sum_{i=1}^k L_i$ , which is the total number of ALT-islands. By (1), we have  $P(L_i \geq j) \leq 2^{-j+1}$ , independently from the other segments. Hence,  $L_i - 1$  is stochastically dominated by a geometric distribution with parameter  $1/2$  and  $S - k$  is dominated by the sum of  $k$  such random variables. We assume  $k \leq 2c' \log n$ , which, as argued before, holds with high probability. Now we can apply Theorem 1.14 from [4] on the sum of geometric random vari-

ables, choosing  $\delta = 3$ , and get that  $P(S \geq k + 6c' \log n) \leq e^{-\frac{9(2c' \log n - 1)}{8}} \leq n^{-c'}$ . Altogether, for a sufficiently large  $n$  and sufficiently large constant  $c$  from the lemma, there will with high probability still be an island with  $x_i^*(t) = \text{OPT}$  at the end of the phase.  $\square$

We note that the bounds used in Lemma 5 take a very dim view of the situation, and could probably be improved significantly. In practical simulations, we observe that the simplified island model converges to a larger-than- $p_{\text{OPT}}$  majority of islands having OPT as their current-best solution, and any OPT-to-ALT mutations disappear quickly.

Applying Lemmas 4 and 5 inductively over  $n$  phases yields a proof of Theorem 3.

**PROOF OF THEOREM 3.** For the first iteration, Lemma 5 may be applied immediately, as all islands are initialized with the OPT individual. Per the lemma, at least one island ends the phase with  $x_i^*(t) = \text{OPT}$  with high probability, allowing Lemma 4 to be applied at the beginning of the next phase. Per that lemma, there is with high probability an iteration within the phase when OPT is the current-best individual on all islands, allowing Lemma 5 to be applied again.

As the events described in both of these lemmas occur with high probability, and we only require  $n$  repeated applications of each lemma to cover the whole optimization process, a simple union bound on the failure probabilities can be used to conclude that with high probability, at least one island is still tracking the oscillating optimum after the  $n$  phases are over.  $\square$

Thus, we have proven that using a unidirectional ring as the migration topology can allow the simplified island model to track the oscillating optimum of the MAZE in settings where this is not possible for the complete migration topology. Intuitively, this is achieved by removing the ability of a single ill-timed migration to propagate an undesirable individual to all islands.

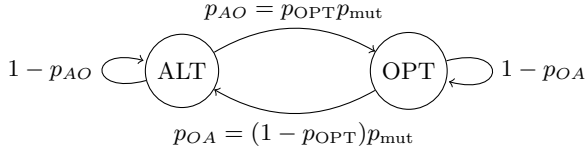
## 4. OCCASIONAL MIGRATION

In this section, we consider the behavior of the island model when migration occurs less frequently. In particular, we demonstrate that with  $p_{\text{mig}} = O(1/t_0)$ , the ring topology is not able to track the optimum through  $n$  phases.

The following lemma provides a useful bound on the distribution of the non-LOST island states immediately prior to a phase transition in cases where migration does not occur close to the phase transition. Its proof follows the approach used in [15] to analyze the behavior of a single  $(1+1)$  EA island on MAZE.

**LEMMA 6.** *Let  $0 < p_{\text{OPT}} < 1$  be a constant. Assuming no migration or phase transitions have occurred for at least  $t = 2k/p_{\text{mut}}$  iterations, where  $k$  is a large-enough constant, the probability  $p_A$  that a non-LOST is in an ALT state can be bounded by constants  $a \leq p_A \leq b$  such that  $a > 0$  and  $b < 1$ .*

**PROOF.** In the absence of migration, the behavior of a non-LOST island can be modeled by a Markov chain over the two states ALT and OPT (corresponding to the current-best individuals), illustrated in Figure 3. The steps of the Markov chain correspond to iterations of the simplified island model,



**Figure 3: Island behavior in the absence of migration and phase transitions, modeled as a two-state Markov chain in the proof of Lemma 6.**

and the probabilities of transitioning between the states are  $p_{AO} = p_{\text{OPT}} p_{\text{mut}}$  and  $p_{OA} = (1 - p_{\text{OPT}}) p_{\text{mut}}$  respectively.

Consider the steady-state probability  $\pi_{\text{ALT}}$  of the chain being in the ALT state:

$$\begin{aligned} \pi_{\text{ALT}} &= \pi_{\text{OPT}} p_{OA} + \pi_{\text{ALT}} (1 - p_{AO}) \\ &= \pi_{\text{OPT}} (1 - p_{\text{OPT}}) p_{\text{mut}} + \pi_{\text{ALT}} (1 - p_{\text{OPT}} p_{\text{mut}}) \\ &= (1 - \pi_{\text{ALT}}) (1 - p_{\text{OPT}}) p_{\text{mut}} + \pi_{\text{ALT}} (1 - p_{\text{OPT}} p_{\text{mut}}) \\ &= (1 - p_{\text{OPT}}) p_{\text{mut}} + \pi_{\text{ALT}} (1 - p_{\text{mut}}) \\ &= 1 - p_{\text{OPT}} \end{aligned}$$

as  $\pi_{\text{OPT}} + \pi_{\text{ALT}} = 1$ .

Over time, the distribution of the island's state approaches the steady-state distribution of this Markov chain. To bound the total variation distance (i.e., deviation from the steady-state distribution at a particular time), we use a coupling time argument: the total variation distance at time  $t$  is at most the probability that two independent instances of the Markov chain, started in different states, have not ever been in the same state by time  $t$ . Using  $p_{OA}$  and  $p_{AO}$  to denote the transition probabilities, the total variation distance at time  $t$  is thus at most

$$\begin{aligned} (1 - p_{OA}(1 - p_{AO}) - (1 - p_{OA})p_{AO})^t &= \\ (1 - p_{OA} - p_{AO} + 2p_{OA}p_{AO})^t &= \\ (1 - p_{\text{mut}} + 2p_{\text{mut}}^2(p_{\text{OPT}} - p_{\text{OPT}}^2))^t &\leq \\ (1 - p_{\text{mut}}/2)^t & \end{aligned}$$

using that  $(p_{\text{OPT}} - p_{\text{OPT}}^2) < 1$ , and  $2p_{\text{mut}}^2 \leq p_{\text{mut}}/2$  for  $p_{\text{mut}} \leq 1/4$ . Then, picking  $t = 2k/p_{\text{mut}}$ ,

$$(1 - p_{\text{mut}}/2)^{2k/p_{\text{mut}}} \leq e^{-k}$$

which yields bounds on  $p_A$ , the probability that the island ends the phase with ALT as its current-best individual:

$$\begin{aligned} \pi_{\text{ALT}} - e^{-k} &\leq p_A \leq \pi_{\text{ALT}} + e^{-k} \\ (1 - p_{\text{OPT}}) - e^{-k} &\leq p_A \leq (1 - p_{\text{OPT}}) + e^{-k} \end{aligned}$$

which are constant when  $p_{\text{OPT}}$  is a constant and  $k$  is a large-enough constant.  $\square$

For standard choices of model parameters, the lemma states that there exists a constant  $k > 0$  such that if no migration occurs  $kn$  iterations before a phase transition, the probability that any non-LOST island has ALT as its current-best solution at the phase transition can be bounded by constants from both directions.

**COROLLARY 7.** *When migration does not occur significantly more often than mutation, i.e.,  $p_{\text{mig}} \in O(p_{\text{mut}})$ , and  $0 < p_{\text{OPT}} < 1$  is a constant, the probability  $p_A$  that a non-LOST island has ALT as its current-best individual  $\Omega(1/p_{\text{mut}})$*

iterations after a phase transition (or after the island becoming non-LOST), can be bounded by constants  $a \leq p_A \leq b$ , where  $a > 0, b < 1$ .

**PROOF.** Proof of Lemma 6 can be adapted to this setting.

For the lower bound on  $p_A$ , we pessimistically assume that migration, when it occurs, always causes a transition from the ALT state to the OPT state; as  $p_{\text{mig}} \in O(p_{\text{mut}})$ , this increases  $p_{AO}$  by at most a constant factor, and hence increases  $\pi_{\text{ALT}}$  by at most a constant.

For the upper bound on  $p_A$ , we similarly assume that migration always causes a transition from the OPT state to the ALT state, increasing  $p_{OA}$  by at most a constant factor, and hence decreasing  $\pi_{\text{ALT}}$  by at most a constant.

Increasing the transition probabilities between states can only shorten the time required to reduce the total variation distance down to the desired level, so the  $e^{-k}$  bound on total variation distance from the Markov chain steady-state distribution can be applied without further modifications.  $\square$

## 4.1 Ring topology

With migration occurring an expected constant number of times in each phase, the using the unidirectional ring as the migration topology results in all islands being in the LOST state at the end of  $n$  phases.

**THEOREM 8.** *When  $t_0 \in \Omega(n^2) \cap O(\text{poly}(n))$ ,  $p_{\text{OPT}} = 2/3$ ,  $p_{\text{mut}} = 1/(en)$ ,  $p_{\text{mig}} = 1/(kt_0)$ , where  $k > 1$  is a large-enough constant,  $\lambda = O(n^{1-\epsilon})$ , and  $G$  is a  $\lambda$ -vertex unidirectional ring, the simplified island model will with high probability have all islands in the LOST state by the end of phase  $n$ .*

**PROOF.** Let  $X_p$  be the number of islands *not* in the longest LOST segment in the migration topology at the start of phase  $p$ . We would like to apply the additive drift theorem to  $X_p$ , showing that there exists a drift toward 0, and, as  $\lambda = O(n^{1-\epsilon})$ ,  $X_p = 0$  is hit before the  $n$  phases are over.

We bound the drift on  $X_p$  in terms of two components: the negative contribution  $\delta^-$  from migration recovering LOST islands at the start of the segment, and the positive contribution  $\delta^+$  from the phase transition being able to change the island immediately following the segment into the LOST state. Thus, we would like to show that:

$$E(X_p - X_{p+1} \mid 0 < X_p < \lambda) = \delta^+ - \delta^- > c$$

where  $c > 0$  is constant, as well as a similar statement for the  $X_p = \lambda$  case.

The negative contribution  $\delta^-$  can be upper-bounded as the number of migrations that occur during the  $t_0$  iterations in a phase, i.e., a binomially-distributed variable:

$$E(\delta^-) \leq t_0 p_{\text{mig}} = 1/k.$$

For the positive contribution  $\delta^+$ , we can derive a lower bound by focusing on the non-LOST island immediately following the  $(\lambda - X_p)$ -island LOST segment.

We note that for as long as the LOST segment preceding it does not disappear entirely, this island is not affected by migration, which allows the strategy used in the proof of Lemma 6 to be applied. For the island in question, the true probability of having ALT as the current-best individual approaches  $\pi_{\text{ALT}} = (1 - p_{\text{OPT}})$  from above, as the island begins phase  $p$  with ALT as its current-best solution (due

to the phase transition preceding phase  $p$ ). This allows us to use  $(1 - p_{\text{OPT}})$  as a lower bound on  $\delta^+$  when this island is not affected by migration.

When the island *is* affected by migration, Corollary 7 can be applied: even in the presence of migration to the considered island, the probability that it ends the phase in an ALT state, and hence  $\delta^+$ , can be lower-bounded by a positive constant.

Returning to the overall drift,

$$E(X_p - X_{p+1} \mid 0 < X_p < \lambda) = \delta^+ - \delta^- \geq \Omega(1) - 1/k$$

i. e., for a large-enough constant  $k$ , there is a constant drift toward  $X_p = 0$ .

For  $X_p = \lambda$ , we can show that a constant drift exists by referring to Lemma 6: with probability greater than  $1 - p_{\text{mig}} \cdot 2k/p_{\text{mut}} = 1 - O(n^{-2})$ , no migration occurs too close to the phase transition, and thus all islands are within a constant variation distance of the steady-state distribution. The drift can be bounded by considering the contribution of single island:

$$\begin{aligned} E(X_p - X_{p+1} \mid X_p = \lambda) &\geq (1 - p_{\text{OPT}} - e^{-k})(1 - O(n^{-2})) \\ &= \Theta(1) \end{aligned}$$

Applying the additive drift theorem, the expected first hitting time of  $X_p = 0$  is  $O(\lambda) = O(n^{1-\epsilon})$  phases. We note that the probability that this does not happen in twice the expected number of phases is, by applying Markov's inequality, at most 0.5; and after  $\Omega(n^\epsilon)$  repetitions, at most  $2^{-\Omega(n^\epsilon)}$ . Therefore, with high probability, the ring topology loses track of the optimum on all islands before the  $n$  phases are over.  $\square$

This serves as an illustration that with  $p_{\text{mig}} < 1/(kt_0)$ , where  $k > 1$  is a sufficiently large constant, migration on a ring topology is not able to recover islands lost in phase transitions sufficiently quickly. In such circumstances, denser migration topologies may have an advantage, as they are able to repopulate more islands per migration, and therefore also track the optimum through a greater number of phases.

## 4.2 Complete topology

In [15], it was proved that a complete migration topology loses track of the MAZE optimum if migrations occurred less frequently than once in every  $O(\log(\lambda)t_0)$  iterations. This result also points to a negative result for the complete topology with  $p_{\text{mig}} = \Omega(1/t_0)$  in the simplified model, as the time between migrations, which is approximately geometrically distributed, may exceed  $c \log(n)t_0$  iterations with probability  $n^{-c/k}$ , where  $k > 0$  is a constant; allowing for  $n/\log n$  repetitions, this would mean that with migration rate  $p_{\text{mig}} = 1/(kt_0)$ , the complete topology fails to track the optimum through all  $n$  phases with high probability.

We note that Theorem 8 would also apply to any migration schedule with the same expected number of migrations. On the other hand, there is a randomized migration schedule, with the same expected number of migrations, for which a complete migration topology is able to track the optimum through all  $n$  phases even with  $\lambda = \log(n)$  islands.

**THEOREM 9.** *Let  $t_0 \in \Omega(n^2) \cap O(\text{poly}(n))$ ,  $p_{\text{OPT}} = 2/3$ ,  $p_{\text{mut}} = 1/(en)$ ,  $\lambda = \Omega(\log n)$ ,  $G$  be a complete  $\lambda$ -vertex*

*graph, and let migration occur during one of every  $kt_0$  iterations (where  $k > 1$  is a constant), chosen uniformly at random. The simplified island model is able to track the optimum through  $n$  phases of  $t_0$  iterations each with polynomially high probability.*

**PROOF.** We note that the maximum number of iterations between any two migrations in this schedule is  $2kt_0$ , corresponding to migration occurring on the first and last iterations of two adjacent  $kt_0$  iteration blocks; thus, at most  $2k$  phases can elapse without migration.

Consider the probability that a single island loses track of the oscillating optimum in  $2k$  phase transitions: in the absence of migration, Lemma 6 applies, and the probability of a non-LOST island ending a phase with an ALT current-best individual is at most a constant smaller than 1. Thus, the probability that the island survives through  $2k$  phase transitions, where  $k$  is a constant, is also a constant; and therefore, the probability that at least one of  $\lambda = \Omega(\log n)$  islands survives is at least  $1 - n^{-c}$ , where  $c > 0$  is a constant.

Thus, as long as at least one island survives a migration-less period, the complete migration topology will allow all islands to recover from the LOST state. With a sufficiently large  $\lambda$ , the probability that at least one island survives through each of the at most  $O(n)$  migration-less periods can be made polynomially high, and hence the complete migration topology will be able to track the oscillating optimum through all  $n$  phases with high probability.

We note that this process relies on no migration occurring too close to a phase transition, as, in the worst case, this could migrate the ALT individual to all islands, resulting in all islands losing track of the oscillating optimum when the phase transition occurs. Per Lemma 6, this is not a problem as long as no migration occurs within  $O(1/p_{\text{mut}}) = O(n)$  iterations of each phase transitions; and so we note that there are at most  $O(n^2)$  iterations during which migration should not occur, and this constraint is respected with probability at least  $1 - O(n^2 p_{\text{mig}}) = 1 - O(n^{-1})$ . Thus, with polynomially high probability, this problematic situation does not occur.  $\square$

## 5. CONCLUSION

We have demonstrated using rigorous runtime analysis that there exist choices of parameters for the simplified island model for which a complete migration topology with high probability results in a failure to track the oscillating optimum through all  $n$  phases, while a unidirectional ring migration topology allows the optimum to be tracked through all  $n$  phases with high probability. This example illustrates that a less-dense migration topology can mitigate the effects of migration occurring during unfavorable iterations of an oscillating fitness function, reducing the need to rely on problem-specific knowledge as in [15]. At the other extreme, we have also proven that denser migration topologies may be advantageous if migration occurs only rarely, as in this setting the ring topology may not allow lost islands to be recovered quickly enough to replenish those which lose track of the oscillating optimum during phase transitions.

While this paper introduced and derived results based on the simplified island model, we believe that the presented results could be transferred to the original setting of  $(1+1)$  EA islands tracking the original MAZE function.

In future work, it would be useful to extend these results beyond the extreme cases of  $p_{\text{mig}} = 1$  and  $p_{\text{mig}} < 1/(kt_0)$ , i. e., for migration occurring just often enough for the ring topology to recover LOST islands. We note that while our theoretical analysis here does not prove this directly,  $p_{\text{mig}} = 1$  combined with a low  $\lambda p_{\text{mut}}$  actually leads to a reduction in population diversity, with the majority of the islands settling on OPT as their current-best solution, rather than achieving a  $p_{\text{OPT}}$ -like balance between OPT and ALT islands. We conjecture that with slower migration, such a balance could be achieved.

## 6. REFERENCES

- [1] Enrique Alba, Amir Nakib, and Patrick Siarry. *Metaheuristics for Dynamic Optimization*. Studies in Computational Intelligence. Springer, 2013.
- [2] Enrique Alba and José M Troya. A survey of parallel distributed genetic algorithms. *Complexity*, 4(4):31–52, 1999.
- [3] Duc-Cuong Dang, Thomas Jansen, and Per Kristian Lehre. Populations can be essential in dynamic optimisation. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '15)*, pages 1407–1414, 2015.
- [4] Benjamin Doerr. Analyzing randomized search heuristics: Tools from probability theory. In Anne Auger and Benjamin Doerr, editors, *Theory of randomized search heuristics*. World Scientific, 2011.
- [5] Stefan Droste. Analysis of the (1+1) EA for a dynamically bitwise changing OneMax. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '03)*, pages 909–921. Springer, 2003.
- [6] W. Feller. *An introduction to probability theory and its applications*. Wiley, 3rd edition, 1968.
- [7] Jun He and Xin Yao. Drift analysis and average time complexity of evolutionary algorithms. *Artificial Intelligence*, 127:57–85, 2001. Erratum in *Artif. Intell.* 140(1/2): 245–248 (2002).
- [8] Thomas Jansen and Ulf Schellbach. Theoretical analysis of a mutation-based evolutionary algorithm for a tracking problem in the lattice. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '05)*, pages 841–848. ACM Press, 2005.
- [9] Thomas Jansen and Christine Zarges. Evolutionary algorithms and artificial immune systems on a bi-stable dynamic optimisation problem. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '14)*, pages 975–982. ACM Press, 2014.
- [10] Timo Kötzing, Andrei Lissovoi, and Carsten Witt. (1+1) EA on generalized dynamic OneMax. In *Proceedings of Foundations of Genetic Algorithms Workshop (FOGA '15)*, pages 40–51. ACM Press, 2015.
- [11] Timo Kötzing and Hendrik Molter. ACO beats EA on a dynamic pseudo-boolean function. In *Proceedings of Parallel Problem Solving from Nature (PPSN XII)*, pages 113–122. Springer, 2012.
- [12] Jörg Lässig and Dirk Sudholt. Design and analysis of migration in parallel evolutionary algorithms. *Soft Comput.*, 17(7):1121–1144, 2013.
- [13] Per Kristian Lehre and Carsten Witt. Concentrated hitting times of randomized search heuristics with variable drift. In *Proceedings of the 25th International Symposium on Algorithms and Computation (ISAAC'14)*, volume 8889 of *Lecture Notes in Computer Science*, pages 686–697. Springer, 2014. extended version at <http://arxiv.org/abs/1307.2559>.
- [14] Andrei Lissovoi and Carsten Witt. MMAS versus population-based EA on a family of dynamic fitness functions. *Algorithmica*, 2015. in press, final version online at <http://dx.doi.org/10.1007/s00453-015-9975-z>.
- [15] Andrei Lissovoi and Carsten Witt. On the utility of island models in dynamic optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '15)*, pages 1447–1454, 2015.
- [16] Andrea Mambrini and Dirk Sudholt. Design and analysis of adaptive migration intervals in parallel evolutionary algorithms. In *Genetic and Evolutionary Computation Conference, GECCO '14, Vancouver, BC, Canada, July 12–16, 2014*, pages 1047–1054, 2014.
- [17] Trung Thanh Nguyen, Shengxiang Yang, and Jürgen Branke. Evolutionary dynamic optimization: A survey of the state of the art. *Swarm and Evolutionary Computation*, 6:1–24, 2012.
- [18] Philipp Rohlfshagen, Per Kristian Lehre, and Xin Yao. Dynamic evolutionary optimisation: An analysis of frequency and magnitude of change. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '09)*, pages 1713–1720. ACM Press, 2009.
- [19] M. Ruciński, D. Izzo, and F. Biscani. On the impact of the migration topology on the island model. *Parallel Computing*, 36(10-11):555–571, 2010. Parallel Architectures and Bioinspired Algorithms.