



HAL
open science

Exploring Self-Assembling Behaviors in a Swarm of Bio-micro-robots using Surrogate-Assisted MAP-Elites

Leo Cazenille, Nicolas Bredeche, Nathanael Aubert-Kato

► **To cite this version:**

Leo Cazenille, Nicolas Bredeche, Nathanael Aubert-Kato. Exploring Self-Assembling Behaviors in a Swarm of Bio-micro-robots using Surrogate-Assisted MAP-Elites. IEEE ALIFE, 2020, Xiamen, China. hal-03175249

HAL Id: hal-03175249

<https://hal.sorbonne-universite.fr/hal-03175249v1>

Submitted on 22 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exploring Self-Assembling Behaviors in a Swarm of Bio-micro-robots using Surrogate-Assisted MAP-Elites

Leo Cazenille

Department of Information Sciences
Ochanomizu University
Tokyo, Japan
cazenille.leo@ocha.ac.jp

Nicolas Bredeche

Sorbonne Université, CNRS
Institut des Systèmes Intelligents
et de Robotique (ISIR)
F-75005 Paris, France
nicolas.bredeche@sorbonne-universite.fr

Nathanael Aubert-Kato

Department of Information Sciences
Ochanomizu University
Tokyo, Japan
naubertkato@is.ocha.ac.jp

Abstract—Swarms of molecular robots are a promising approach to create specific shapes at the microscopic scale through self-assembly. However, controlling their behavior is a challenging problem as it involves complex non-linear dynamics and high experimental variability. Hand-crafting a molecular controller will often be time-consuming and give sub-optimal results. Optimization methods, like the BIONEAT algorithm, were previously employed to partially overcome these difficulties, but they still had to cope with deceptive high-dimensional search spaces and computationally expensive simulations.

Here, we describe a novel approach to solve this problem by using MAP-Elites, an algorithm that searches for both high-performing and diverse solutions. We then apply it to a molecular robotic framework we recently introduced that allows sensing, signaling and self-assembly at the micro-scale and show that MAP-Elites outperforms previous approaches. Additionally, we propose a surrogate model of micro-robots physics and chemical reaction dynamics to reduce the computational costs of simulation. We show that the resulting methodology is capable of optimizing controllers with similar accuracy as when using only a full-fledged realistic model, with half the computational budget.

Keywords—Bio-micro-robots, swarm robotics, molecular programming, evolutionary robotics, surrogate models, quality-diversity algorithms, MAP-Elites

I. INTRODUCTION

Swarms are impressive examples of collective behavior leading to the emergence of new properties [4], [25], [7]. One of the most visible characteristics of swarms is their display of complex, dynamic structures [13], [1], [6]. Here, we are specifically interested in the creation of structure through cooperative assembly. Such behavior has been investigated both in the natural world [1], [6] as well as in robotic swarms [14], [26], [19], [23].

In this paper, we are interested in the automated programming of a swarm of micro-robots ($\gg 1000$). The micro-robots are microscopic agarose beads functionalized with biomolecules (single strand DNA of 12-24 base pairs) [12]. The beads forming the body of the robots are small enough to move through Brownian motion. The DNA functionalization allows them to produce (a) chemical signals that may impact

the behavior of nearby beads and (b) an anchoring signal that will attach them to their neighbors. Clusters created from aggregated beads move slower or even stop, based on their size. That property allows us to control where the beads should go, which we use to create self-assembled structures. Thanks to that scale and the low price and availability of the molecular components, we previously designed and validated *in vitro* a system of a million micro-robots that self-assemble [3], [28].

We previously introduced the BIONEAT algorithm [3], inspired from state-of-the-art NEAT algorithm [24]. BIONEAT aims at producing chemical reaction networks (CRN) which represent bio-molecules to be attached to beads or left floating in the 2-dimensional substrate. Given a target for self-assembly (*e.g.*, in Fig. 2), the objective is that micro-robots assemble to one another in the target area.

Though initial results are promising, this previous work raised several important issues:

- **target complexity**: optimization was limited to simple targets, with one single rectangular region for self-assembly.
- **computational cost**: simulating chemical reactions in a realistic fashion is extremely expensive. Several days of computation were necessary for running the optimization algorithm.

In this paper, we address both issues. Firstly, we introduce the use of MAP-Elites [18], an illumination algorithm that favors exploration over pure optimization, thus reducing the risk of premature convergence during the search. We also implement a new mechanism, termed *topological initialization* that helps MAP-Elites to bootstrap exploration when confronted with search spaces with large neutral regions.

Secondly, we introduce a surrogate model [22] to approximate the (costly) pseudo-realistic simulation of real chemical reactions. We also introduce transferability [16] in combination with the proposed surrogate model, *i.e.*, bootstrapping optimization using the surrogate model and then reverting to the costly pseudo-realistic simulator.

In the following, we first describe the Methods, both in terms of experimental setup, objective function and optimization algorithms. Then, we present the results obtained using a target function that combines several regions of interest. Finally, we discuss the remaining challenges of the approach and conclude.

II. METHODS

A. Molecular robotic swarm

We consider a swarm comprised of molecular robots made from spherical beads grafted with specific strands of DNA (Fig. 1). Those strands are of two types: (a) templates used to capture signal DNA from the environment, process it and produce signal back; (b) anchoring points, which are used for aggregation when the anchoring signal is present.

Processing and production of signals are based on the PEN DNA toolbox [17] (Fig. 1, c.): a signal strand can attach to a compatible template and either triggers the production of a different DNA strand (activation) or prevents activation from other strands (inhibition). The activation mechanism works as follows: once a compatible signal strand is double-stranded to a template, it gets extended into a fully complementary strand by a first enzyme called polymerase. The extend strand is designed to contain the recognition site of a second enzyme, called nickase, which splits it in half. That activity recovers the original signal strand (green in Fig. 1, c.) as well as produces a new signal strand (orange in Fig. 1, c.). That structure is unstable by design and both strands eventually fall off, allowing them to interact with other templates in the system. The inhibition mechanism relies on a signal strand attaching to a target template in a way that does not trigger the activity of either enzymes: a 3'-end mismatch prevents polymerase extension, while missing bases on the 5'-end of the strand prevents recognition by the nickase. At the same time, the inhibiting strand physically prevents activation signals to interact with the template, effectively inactivating it. As with activation signal, this structure is designed to be unstable, allowing the template to be eventually freed.

Signal strands are spread in the environment through chemical diffusion and are degraded over time by a third enzyme called exonuclease. Templates are chemically protected against that activity, thus remaining stable.

Additionally to templates, anchor strands are grafted on the beads. A specific signal, called anchoring signal, can link together anchor strands from two separate beads to link them together (Fig. 1, b). Past experimental results showed that this mechanism is strong enough to produce stable aggregation [3].

Robots move through Brownian motion: individual robots move much faster than aggregated clusters, which in turn may become completely motionless when they get large enough.

B. Simulation models

We use two distinct models to simulate our molecular systems: a “full” model that explicitly computes beads positions and aggregation, and a “surrogate” model that considers a

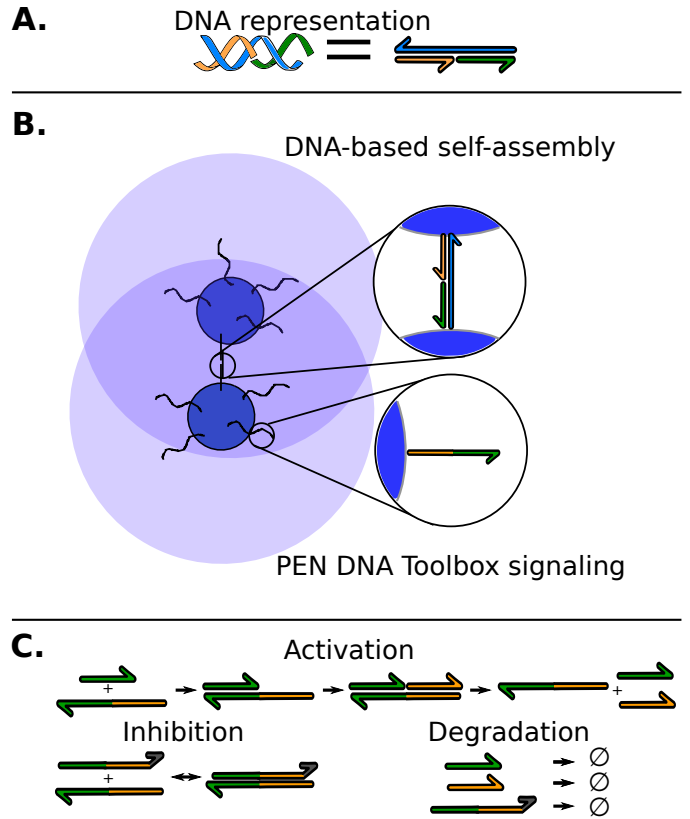


Fig. 1. **A.** Model of DNA: the 3D structures is abstracted, keeping only the orientation of the molecule. Here, two strands (orange and green) are double-stranded to a longer blue strand. **B.** Model of the robots. Two robots are shown here, both producing and emitting a DNA signal (blue diffusion area). The DNA strands grafted on their surface are used both for aggregation and signalling. **C.** Signalling mechanism. DNA strands diffusing in the environment interact with those grafted on the body of the robots, either producing other DNA strands (activation) or preventing production (inhibition). Signal strands are slowly degraded over time by an enzyme present in the environment. Strands grafted on the surface of the robot are chemically protected against this enzymatic activity and remain stable over time.

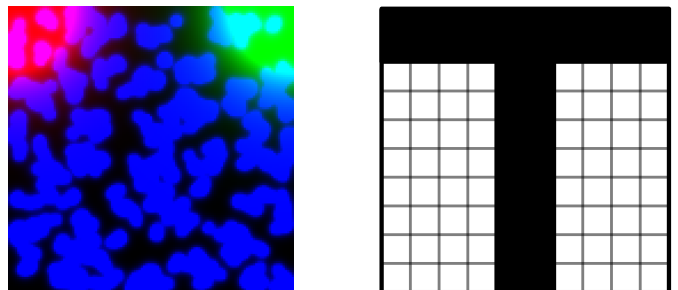


Fig. 2. **Left:** Simulation of an experiment where a swarm of micro-beads (in blue) interact with their environment and self-aggregate. Two gradient sources, in the top-left and top-right positions, are continuously diffused isotropically in the environment. They are shown respectively in red and green. **Right:** Target for self-assembly, with the shape of the letter T. Beads should aggregate in the black area.

simplified scenario where beads are uniformly spread in the environment.

In both cases, the production, diffusion, and degradation of signal species are modeled through reaction-diffusion. For a given signal S , we have:

$$\frac{\partial[S]}{\partial t}(t, x, y) = R_S(t, x, y) + D_S \cdot \Delta[S](t, x, y)$$

where $[S]$ is the concentration of S , R_S is the contribution from reactions (*e.g.*, production or degradation, see Figure 1.C. for a graphical representation and [2] for an explicit formula), D_S is the diffusion coefficient of S and Δ is the Laplacian operator. That approach generates a set of Partial Differential Equations (PDEs) that we solve numerically.

In the full model, beads are modelled as disks moving in the environment through Brownian motion. We implement it by adding Gaussian noise to their position over time, with mean 0 and variance $\sigma^2 = 2D$ where:

$$D = \frac{k_B T}{6\pi d \eta_{\text{water}}}$$

with k_B the Boltzmann constant, T the temperature, d the bead diameter and η_{water} the viscosity of water. To integrate with the reaction-diffusion part of the system, that Gaussian noise is scaled by the length of a time step from the PDE solver.

While the simulation is 2-dimensional, we consider that the beads are moving in a 3-dimensional environment, which allows us to ignore collisions. In presence of the anchoring signal, beads have a probability to aggregate. That probability is computed by implementing a Gillespie-like step: considering the current concentration of signal, we predict when the next aggregation event will occur [11]. If that event happens before the next time step of the PDE solver, aggregation is considered successful. The reverse reaction, a bead separating from an aggregate, is handled the same way. Since reactions probability distributions are memoryless, that approach does not change the overall behavior of the system. Once aggregated, we consider a cluster of N moves in a similar fashion to a single bead with a radius N times larger.

Note that keeping track of the behavior of each bead is computationally intensive. To minimize costs, we use far fewer beads in simulations than there would be in wetlab experiments. To still cover enough of the environment, a larger radius is used both to compute signal production and aggregation range. However, that radius would make the beads extremely slow; to keep the system dynamic, we consider that the beads move as if they had a smaller radius. Those two radii are indicated as *bead size (aggregation)* and *bead size (Brownian motion)* respectively in Table I.

The local concentration of each signal-producing species is directly proportional to the number of beads at a given point of space. That is, we sum the concentrations of DNA molecules grafted on all the beads that are present. Note that, due to the non-linearity of the enzymatic reactions involved, a linear increase in signal-producing species does not necessarily mean a linear increase in the production of signal species.

In the surrogate model, we use the same reaction-diffusion approach, however the concentrations of signal-producing strands are considered spatially uniform and are equal to the respective concentrations of species present on the surface of a single bead. In practice, this model is equivalent to having motionless beads uniformly spread out in the environment or having the signal-producing species grafted on the surface of the environment rather than on beads. Note that such reaction-diffusion system can be experimentally implemented as well [20].

One of the main difference between these two models is that the full model is stochastic, due to the random motion of beads and probabilistic aggregations, while the surrogate model is deterministic. This characteristic makes the surrogate model much cheaper to compute as multiple repeats are not required. At the same time, not having to simulate the Brownian motion and aggregation of beads gives us an additional cut in computational cost. On the negative side, the surrogate model may fail to capture the lack of robustness of a given system. It is also incapable of taking into account possible designs relying on the increase in local concentrations when beads aggregate. However, we have found that, for the task presented in this paper, performances achieved with the simulated model or with the full model and with the surrogate model are correlated (Fig. 3 and Results section).

C. Target Objective

We extend our work from [3] with a more challenging problem, described in this paragraph. We previously devised a methodology to optimize CRN controllers driving a swarm of micro-beads to self-assemble into simple shapes composed of only one straight line. Here, we extend this approach further by focusing on a more complex type of pattern, composed of two lines. This requires that micro-robots to consider several chemical signals in order to perform a collection of attraction and repulsion behaviors to navigate to the target area.

We define a target for self-assembly corresponding to the shape of the letter **T** (Fig. 2 right). Our objective is for the micro-beads to self-assemble into the target area, shown in black, starting from their initial random positions. Two fixed gradient sources are arranged in the top-left and top-right positions in the arena (Fig. 2 left). They each emit a respective type of signal strand that diffuse throughout the environment. They provide the micro-beads information about their localization, potentially inducing self-assembly. As the gradients diffuse isotropically, self-assembling into straight patterns can be difficult.

We quantify the performance of a simulation by using the method described in [3]. The experimental arena is discretized into a $N \times N$ matrix of cells, with $N = 160$. This allows to compute the following *match-nomatch* score:

$$f = \underbrace{\sum_{(\tilde{x}, \tilde{y}) \in T} r * B(\tilde{x}, \tilde{y})}_{\text{reward term}} - \underbrace{\sum_{(\tilde{x}, \tilde{y}) \notin T} p * e^{a*d(\tilde{x}, \tilde{y})} * B(\tilde{x}, \tilde{y})}_{\text{penalization term}}$$

with T the set of (\tilde{x}, \tilde{y}) cells in the target area, r a reward parameter, $B(\tilde{x}, \tilde{y})$ a Bool function of the presence of beads at position (\tilde{x}, \tilde{y}) , d the distance of a cell towards the closest position in T and a a scaling parameter. Individuals are rewarded according to the number of cells within T and penalized for cells outside of T , with penalization increasing with distance to the target area. Table I lists the parameters used for simulation and fitness computation.

We define two fitness functions catering respectively to the full and surrogate models. The surrogate model is deterministic, so the fitness is equal to the *match-nomatch* score over one simulation. The full model is stochastic. To alleviate simulation stochasticity, we assess the performance of an individual CRN by reevaluating 5 simulations, computing their respective *match-nomatch* scores, and defining the resulting fitness value as the median of these scores. The selection of the number of reevaluations corresponds to a trade-off between precision and computational costs. Our choice of 5 reevaluations was motivated by the results presented in Fig. 4 assessing fitness variability depending on the number of reevaluations.

Individuals are deemed valid only if their respective topology matches the requirements of Table II in term of number of signal species and number of activation and inhibition templates.

D. Optimization

We use a two-step experimental protocol for optimization. As a first step, we perform both structural and parametric optimization to obtain chemical reaction networks (CRNs) through two possible methods: BIONEAT and Map-Elites. Structural optimization modifies which templates are added to the beads, thus changing the set of reactions. Parametric optimization focuses on optimizing continuous values, namely DNA sequences stabilities and template concentrations. After convergence or exhaustion of the evaluation budget, we refine the designs obtained by performing a parametric-only opti-

TABLE I
SIMULATIONS AND FITNESS PARAMETERS.

Simulation Parameter	Value
Arena size	$1mm \times 1mm$
Beads	500
Bead size (aggregation)	$50\mu m$
Bead size (Brownian motion)	$5\mu m$
Temperature	$43^\circ C$
Grid size	160×160
Time discretization	0.1 min per step
Simulation duration	4000 steps (<i>i.e.</i> , 400 min)
Target width	$0.20mm$ (20% of arena width)
Fitness Parameter	Value
r (reward)	1.0
p (penalty)	0.2
a (scaling)	0.1

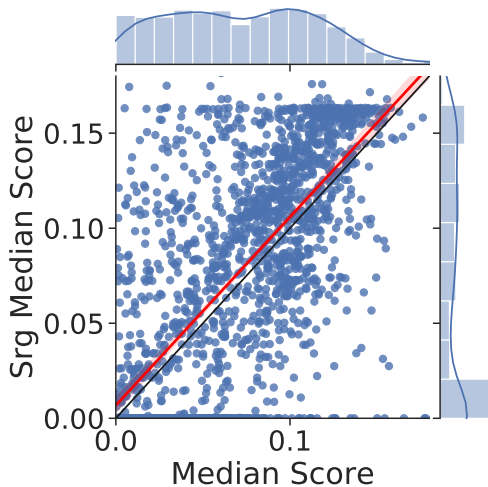


Fig. 3. Comparison of fitness evaluations results between the full model and the surrogate model across 5000 randomly generated individuals. A linear regression of the data is shown in red. The curve in black corresponds to the reference $x = y$. The surrogate model is shown to be an approximate of the basic model ($RMSE = 0.0550$), with a weak (noisy) conservation of score ordering.

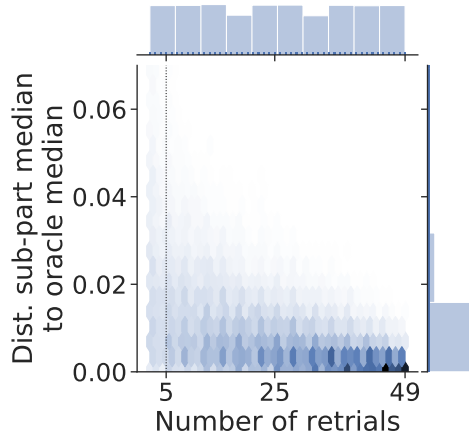


Fig. 4. Variability of evaluation scores on 5000 randomly generated CRN individuals. Each individual is reevaluated 50 times (*oracle*). For each individual, we generate 1000 random subsets, of size 2 to 49 from the *oracle* sets of 50 reevaluations. Medians of all subsets are compared to the respective medians of the *oracle*. We postulate that 5 reevaluations per individuals is a good trade-off between precision and computational cost, and define our fitness metric as the median score across $N = 5$ reevaluations of the same individual.

mization using CMA-ES[15], a state-of-the-art evolutionary optimization method for continuous values.

BIONEAT is an evolutionary algorithm we first introduced in [9]. It takes inspiration from the famous state-of-the-art NEAT algorithm [24], which was originally designed to optimize artificial neural networks. BIONEAT uses specific variation operators to navigate the search space of chemical reaction networks. It is also capable to protect innovation, that is, to explore several regions of the search space simultaneously, balancing between novelty of a particular design and

the quality of solutions. We previously showed that BIONEAT provides efficient solutions for targets comprised of a single regions (horizontal or vertical lines, see [3] for details). The main limit of BIONEAT is that while it can protect innovation for some time, there is not guarantee that it is capable of escaping completely the curse of premature convergence due to the pressure for optimizing towards (possibly only temporary) better solutions.

BIONEAT searches through CRN topologies iteratively. This behavior is inherited from NEAT that postulated that iterated small changes in topologies would often only result in a mild effect on the fitness values. It may not be the case with CRN [27], where small changes in topology can have severe effects in fitness values, a particular trait of deceptive and hard-explore problems. This may explain why BIONEAT is prone to premature convergence during optimization, as it does not possess a way to explore a totally unexplored niche in the space of topologies. While this aspect is partially mitigated with BIONEAT mechanism of speciation, which allows the parallel optimization of several topological niches, it does not enforce the discovery of totally novel niches. Worst, as new species are only created through atomic mutations (*i.e.*, change only a small part of the topology), the surviving species may contains individuals with very similar topologies, possibly in the same topological niche.

In order to favor exploration over pure optimization, we rely on MAP-Elites [18], a Quality-Diversity algorithm [21], that decomposes the search space into regions based on feature descriptions. It considers how does a candidate solution *looks like* in the phenotypic space instead of considering how it is coded in the genotypic space. This method is particularly suited to cope with multi-modal, deceptive, hard exploration problems where traditional optimization algorithms would be prone to premature convergence, as in evolutionary robotics [18], [8], [10].

TABLE II
PARAMETERS OF BIONEAT AND MAP-ELITES.

Optimization Parameter	Value
Evaluation Budget	3000
Number of retrials per individuals	5
Maximal number of activation signals	6
Maximal number of activation templates	7
Maximal number of inhibition templates	7
BIONEAT Parameter	Value
Target number of BIONEAT species	20
Population Size	50
Number of templates	1 – 13
MAP-Elites Parameter	Value
Number of bins	7
Batch size	50
Number of elites per grid bin	1
Number of templates	7 – 13
Mutation operator	Probability
Parameter mutation	0.80
Add template strand	0.05
Remove template strand	0.05
Add signal species	0.05
Add inhibition species	0.05

Algorithm 1: Topological initialization of a CRN

```

input:  $T_{min}, T_{max}$ : min and max nr of templates
input:  $pmut_{min}, pmut_{max}$ : min and max nr of
parameter mutations
1  $ind_{base} \leftarrow$  "3 signals (2 gradients + 1 anchoring signal),
one autocatalytic activation template applied to the
anchoring signal."
2 while True do
| /* Initialize random topology */
3  $ind \leftarrow ind_{base}$ 
4  $T_{target} \leftarrow \text{randint}(T_{min}, T_{max})$ 
5 while True do
6 | if  $ind.nb\_templates = T_{target}$  then
7 | | break
8 | else if  $ind.nb\_templates > T_{target}$  then
9 | |  $\text{disable\_template}(ind)$ 
10 | else
11 | |  $\text{apply\_topological\_mutations}(ind)$ 
| /* Apply parameter mutations */
12  $pmut_{target} \leftarrow \text{randint}(pmut_{min}, pmut_{max})$ 
13 for  $\_$  in  $1..pmut_{target}$  do
14 | |  $\text{apply\_param\_mutation}(ind)$ 
| /* Verify if ind is valid */
15 if  $ind.valid$  then
16 | | break
17 return  $ind$ 

```

MAP-Elites iteratively regroups the explored solutions in a grid of elites. This results in a collection of high-performing individuals across a number of features selected by the user, corresponding to the axes of the grid. Here, we only consider a single feature corresponding to the total number of templates in the topology of an individual. CRN with smaller number of templates are easier to test experimentally but lose expressivity. Conversely, large-sized CRN can describe more complex behaviors, which may be necessary for the beads to successfully self-aggregate into the target shape. As such, a trade-off in term of topology complexity has to be considered, possibly after the optimization process. This substantiates methodologies that concurrently search for topologies of differing sizes.

MAP-Elites is equipped with the same set of mutation operators as BIONEAT, and also retains its capability to optimize iteratively the topologies. We introduce a novel methodology to bootstrap MAP-Elites exploration by initializing a collection of individuals with random topologies. This approach, described in Algorithm 1, allows MAP-Elites to consider a large number of differing topologies. It makes use of the BIONEAT mutation operators to generate individuals of varying topologies across a range of number of templates. In one optimization run, 10% of the individuals are initialized with a random topology (*i.e.*, 300 individuals for an evaluation budget of 3000). This contrasts with the BIONEAT approach, where only small iterative changes in topologies are possi-

ble from mutations, and where individuals with totally new topologies are not initialized.

Table II lists the chosen parameters for the MAP-Elites algorithm. We use our own implementation of BIONEAT and MAP-Elites. BIONEAT is open source and available from <https://bitbucket.org/AubertKato/bioneat/>. MAP-Elites is coded in Python using the QDpy library [5] and freely available as open source software at <https://gitlab.com/leo.cazenille/qdpy>. All additional scripts used in this paper are available at <https://bitbucket.org/leo-cazenille/daccad-qd>.

III. RESULTS

The BIONEAT and MAP-Elites algorithms (shorten as BN and ME for brevity) are used to optimize CRNs on the **T** target in three settings. The **full** and **srg** settings involve, respectively, the full and the surrogate simulations on a budget of 3000 evaluations. The **srg+tsf** setting tests the transferability of the surrogate model: individuals are first optimized using the surrogate model during 1500 evaluations. Then, all individuals of the last iteration (last population for BIONEAT, final grid of MAP-Elites) are reevaluated and optimized using the full model over 1500 evaluations.

All results are compared to a CRN designed by an expert, corresponding to the **Expert** setting.

The best-performing individuals of each method are reevaluated using the full model, with scores presented in Table III. For each method, CRN parameters of the best reevaluated individuals are optimized by CMA-ES [15] over 2000 evaluations, as in [3]: the topology of these individuals are fixed, but templates concentration and activation signals stability parameters are optimized.

Full runs are computed in 15 hours per run, **srg** runs take 90 minutes per run and **srg+tsf** runs take 9 hours per run.

For each setting, the evolution of fitness values during 16 optimization runs are presented in Fig. 5.

Figure 6 shows the final state of a full simulation of the best-performing solutions for each setting, before and after optimization by CMA-ES. Only the anchoring signal is represented, corresponding to self-assembled clusters of beads.

To compare the performance of the different settings, we collected the best individuals from each run (16 individuals per setting). For any two given settings, we used a Mann-Whitney U test on the distribution of fitness of their respective best individuals (Table IV). The algorithms can be ordered in terms of mean performance: $ME \simeq ME+srg+tsf \gg BN+srg+tsf \simeq BN \gg ME+srg \simeq BN+srg$, where \gg indicates a statistically significant difference ($p < 0.05$). The remaining cases, shown by a \simeq have $p = 0.11$, $p = 0.12$ and $p = 0.19$ respectively.

In order to evaluate the impact of further parametric optimization refinement, we picked the best performing individual for each method, and compare performance *before* and *after* an extra parametric optimization step (*i.e.*, the topology is fixed) using the CMA-ES optimization algorithm. All individuals were reevaluated 100 times with the full model and the distribution of their evaluations were compared with a Mann-Whitney U test).

Ordering the best individuals from each method *before* parametric optimization gives the following: $ME \simeq ME+srg+tsf \gg BN \gg Expert \gg BN+srg+tsf \gg ME+srg \gg BN+srg$.

After parametric optimization, the order of the best individuals becomes: $Expert+CMAES \simeq ME+CMAES \gg BN+srg+CMAES \gg ME+srg+tsf+CMAES \gg BN+CMAES \gg ME+srg+CMAES \gg BN+srg+tsf+CMAES$.

Finally, we also measured the impact of the optimization by CMA-ES by comparing the distribution of evaluations of a given individual and that of its optimized version. We used once again 100 retrials with a Mann-Whitney U test. The p -values were added to Table III. All optimized individuals were significantly better than their original ($p < 0.05$) except for the best individual of the ME with Full model setting ($p = 0.08$).

In accordance with the specifications of Table II, optimized individuals are noticeably smaller than the expert-designed CRN (19 templates, Fig. 7): 13 for BN, 12 for BN+srg and BN+srg+tsf, 11 for ME, 10 for ME+srg and ME+srg+tsf. Fewer templates is actually beneficial with respect to potential *in vitro* experimentation, as the models used loose accuracy when the number of chemical interactions grows.

IV. DISCUSSION AND CONCLUSION

We presented an approach to optimize CRNs driving the behavior of a swarm of bio-micro-robots to self-assemble into a **T**-shaped target. Our approach improves upon the methodology in [3], based on BIONEAT. We then described a method to optimize CRNs using the MAP-Elites [18]. It is a Quality-Diversity algorithm [21], allowing us to explore solutions that are both high-performing and diverse. It is particularly suitable to cope with the deceptive nature of our target problem. We also introduced the notion of topological initializations to bootstrap MAP-Elites. We showed that MAP-Elites outperforms BIONEAT both in term of performance and convergence speed. While the best-performing individuals optimized by BIONEAT still require an additional optimization stage using CMA-ES to provide competitive results, MAP-Elites does not necessitate this procedure.

We presented a way to reduce the computational costs of simulations through the use of a surrogate model of the robot swarm dynamics. Our optimization approach required far less computational power, while still retaining sufficient accuracy. We described how individuals optimized through the surrogate model can be transferred into optimization processes implementing the full model. We showed that such results are competitive with results obtained using only the full model, nearly dividing by two the optimization effort (60%).

We limited the optimization runs to small CRN (≤ 13 templates) to ease potential experimental validations. Indeed, smaller CRNs are easier to test experimentally, but lose expressivity. We showed that all methods are capable of optimizing small CRN while still being competitive with results obtained by the expert-designed CRN, that involves a far larger number of templates (19). Additionally, MAP-Elites allows for the concurrent optimization of CRN of differing size (number of templates), resulting in several CRNs that could be selected

TABLE III

FITNESS SCORES OF THE BEST-PERFORMING INDIVIDUALS FOR ALL METHODS, ACROSS 100 RETRIALS. THOSE INDIVIDUALS ARE FURTHER OPTIMIZED BY CMA-ES AND COMPARED TO THEIR ORIGINAL WITH THE MANN-WHITNEY U TEST. THE p VALUE IS SHOWN IN PARENTHESIS.

Method	Full	Full+CMAES	srg	srg+CMAES	srg+trf	srg+tsf+CMAES
Expert	0.52 ± 0.05	0.63 ± 0.04 ($p < 0.01$)	-	-	-	-
BN	0.54 ± 0.02	0.60 ± 0.02 ($p < 0.01$)	0.42 ± 0.04	0.62 ± 0.03 ($p < 0.01$)	0.49 ± 0.03	0.51 ± 0.04 ($p < 0.01$)
ME	0.60 ± 0.04	0.63 ± 0.04 ($p < 0.01$)	0.43 ± 0.06	0.55 ± 0.03 ($p < 0.01$)	0.60 ± 0.03	0.61 ± 0.03 ($p < 0.01$)

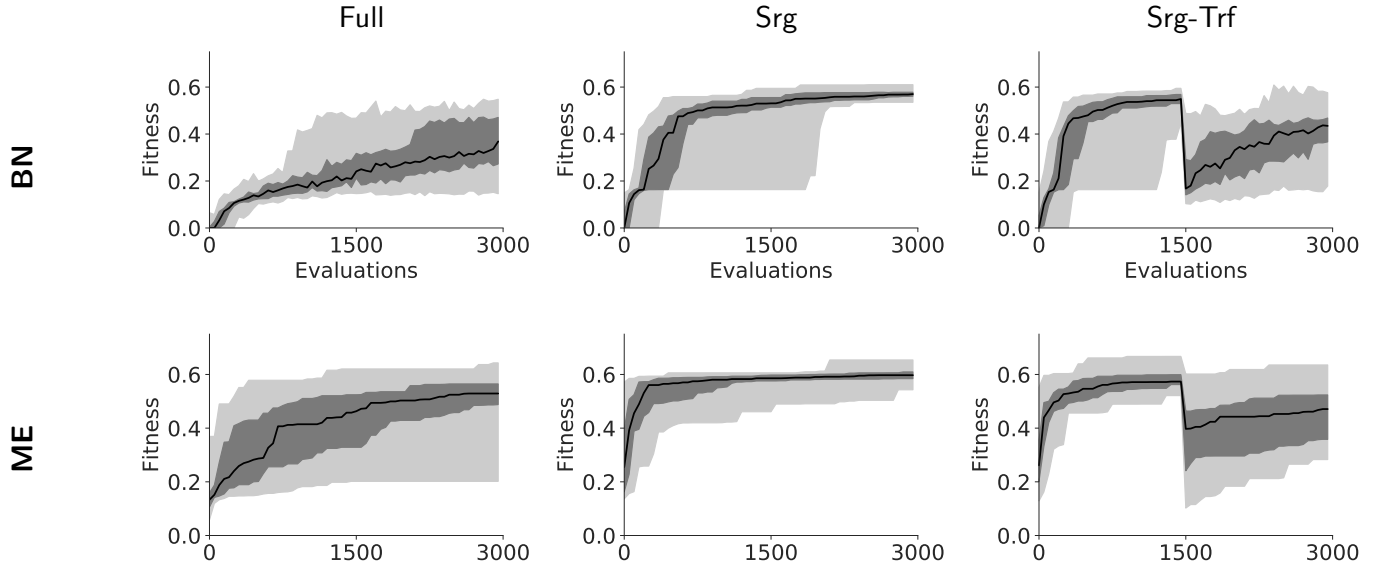


Fig. 5. Evolution of the median quality of the best-performing individuals for each target. Optimization methods are tested across 16 different runs. A fitness of 1.0 corresponds to the best performance. The darker shade represents the 25 to 75 percentiles. The lighter shade encompasses the minimal and maximal values.

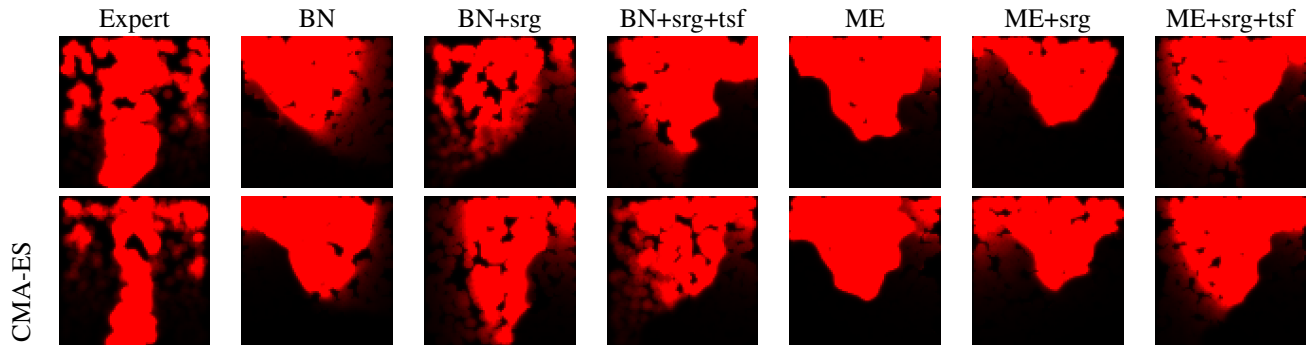


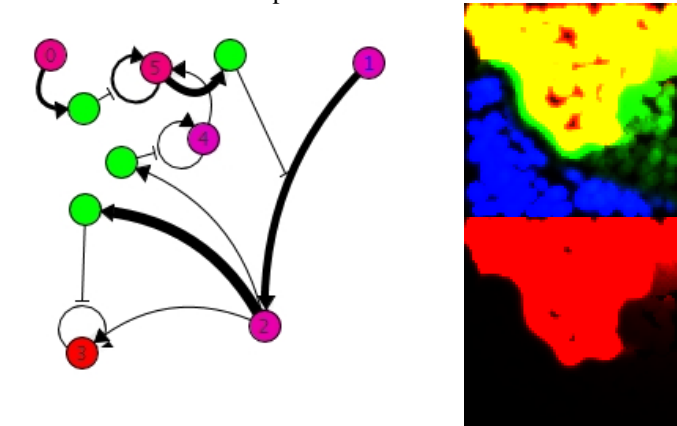
Fig. 6. Examples of final states obtained by **Full** simulations of the best-performing individuals optimized by each method, before and after calibration by CMA-ES.

TABLE IV

MANN-WHITNEY U TEST FOR COMPARING THE DISTRIBUTION OF THE BEST INDIVIDUALS FOUND BY THE DIFFERENT ALGORITHMS. THE TABLE SHOWS WHICH ALGORITHM HAS THE BETTER MEAN BETWEEN TWO GIVEN PAIRS AND THE p -VALUE CALCULATED BY THE TEST.

	BN+srg	BN+srg+tsf	ME	ME+srg	ME+srg+tsf
BN	BN ($p = 7 \times 10^{-6}$)	BN+srg+tsf ($p = 0.12$)	ME ($p = 8 \times 10^{-4}$)	BN ($p = 10^{-5}$)	ME+srg+tsf ($p = 5 \times 10^{-4}$)
BN+srg	-	BN+srg+tsf ($p = 9 \times 10^{-7}$)	ME ($p = 10^{-8}$)	ME+srg ($p = 0.19$)	ME+srg+tsf ($p = 6 \times 10^{-8}$)
BN+srg+tsf	-	-	ME ($p = 2 \times 10^{-4}$)	BN+srg+tsf ($p = 2 \times 10^{-5}$)	ME+srg+tsf ($p = 9 \times 10^{-3}$)
ME	-	-	-	ME ($p = 8 \times 10^{-8}$)	ME ($p = 0.10$)
ME+srg	-	-	-	-	ME+srg+tsf ($p = 10^{-6}$)

Best individual before optimization with CMA-ES: ME



Best individual after optimization with CMA-ES: Expert

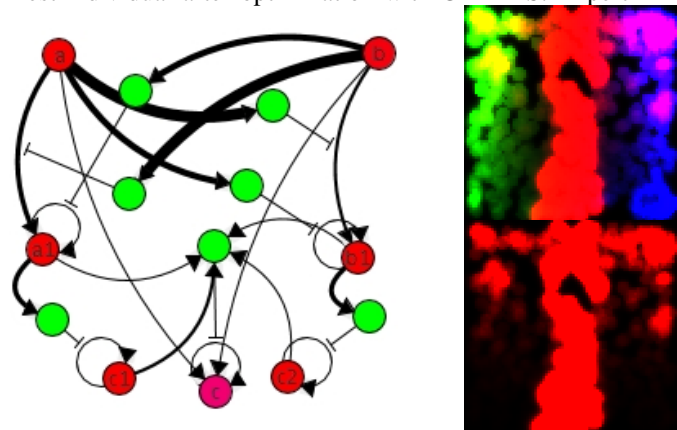


Fig. 7. Best-performing solutions before (top) and after (bottom) parameter tuning done by CMA-ES. Left: chemical reaction networks, vertices are signal strands (red: activating, green: inhibiting), edges are templates, inhibition is shown by a bar-headed arrow; top-right: chemical concentrations; bottom-right: production of anchoring signals.

after the optimization process for experimentation, depending on the choices and requirements of the experimenter.

Our approach with MAP-Elites could be improved by testing additional features rather than only the number of templates, including topological features (e.g., number of vertices, properties of the graph) or behavioral features (e.g., occupation of various zones of the arena, speed of reaching a stable state). The interplay between surrogate and full model could be improved further by finding automatically when to perform a transfer. The optimization process could also switch several times between surrogate and full model evaluations, either periodically or under some predetermined conditions. Finally, our approach could be applied to more complex target shapes, possibly requiring the implementation of mechanisms to optimized CRN iteratively, or by automatically decomposing the target shapes into sub-shapes.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number JP17K00399 and by Grant-in-Aid for JSPS Fellows JP19F19722. This work was also supported by the Agence Nationale pour la Recherche under Grant ANR-18-CE33-0006 (MSR project).

REFERENCES

- [1] C Anderson, G Theraulaz, and JL Deneubourg. Self-assemblages in insect societies. *Insectes sociaux*, 49(2):99–110, 2002.
- [2] N Aubert, C Mosca, T Fujii, M Hagiya, and Y Rondelez. Computer-assisted design for scaling up systems based on dna reaction networks. *J. R. Soc. Interface*, 11(93):20131167, 2014.
- [3] N Aubert-Kato, C Fosseprez, G Gines, I Kawamata, H Dinh, L Cazenille, A Estevez-Tores, M Hagiya, Y Rondelez, and N Bredeche. Evolutionary optimization of self-assembly in a swarm of bio-micro-robots. In *GECCO*, pages 59–66. ACM, 2017.
- [4] S Camazine. *Self-organization in biological systems*. Princeton University Press, 2003.
- [5] L Cazenille. Qdpy: A python framework for quality-diversity. <https://gitlab.com/leo.cazenille/qdpy>, 2018.
- [6] MF Copeland and DB Weibel. Bacterial swarming: a model system for studying dynamic self-assembly. *Soft matter*, 5(6):1174–1187, 2009.
- [7] ID Couzin. Collective cognition in animal groups. *Trends in cognitive sciences*, 13(1):36–43, 2009.
- [8] A Cully, J Clune, D Tarapore, and JB Mouret. Robots that can adapt like animals. *Nature*, 521(7553):503, 2015.
- [9] HQ Dinh, N Aubert, N Noman, T Fujii, Y Rondelez, and H Iba. An effective method for evolving reaction networks in synthetic biochemical systems. *Trans. Evol. Comp.*, 19(3):374–386, 2014.
- [10] M Duarte, J Gomes, SM Oliveira, and AL Christensen. Evolution of repertoire-based control for robots with complex locomotor systems. *Trans. Evol. Comp.*, 22(2):314–328, 2018.
- [11] DT Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.
- [12] G Gines, AS Zadorin, JC Galas, T Fujii, A Estevez-Torres, and Y Rondelez. Microscopic agents programmed by dna circuits. *Nat. Nano.*, 12(4):351, 2017.
- [13] AE Goodenough, N Little, WS Carpenter, and AG Hart. Birds of a feather flock together: Insights into starling murmuration behaviour revealed using citizen science. *PLoS one*, 12(6):e0179277, 2017.
- [14] R Groß, M Bonani, F Mondada, and M Dorigo. Autonomous self-assembly in swarm-bots. *Trans. robot.*, 22(6):1115–1130, 2006.
- [15] N Hansen. The CMA evolution strategy: a comparing review. In *Towards a new evolutionary computation*, pages 75–102. Springer, 2006.
- [16] S Koos, JB Mouret, and S Doncieux. The transferability approach: Crossing the reality gap in evolutionary robotics. *Trans. Evol. Comp.*, 17(1):122–145, February 2013.
- [17] K Montagne, R Plasson, Y Sakai, T Fujii, and Y Rondelez. Programming an in vitro dna oscillator using a molecular networking strategy. *Molecular systems biology*, 7(1):466, 2011.
- [18] JB Mouret and J Clune. Illuminating search spaces by mapping elites. *arXiv preprint arXiv:1504.04909*, 2015.
- [19] R O’Grady, AL Christensen, and M Dorigo. Swarmorph: multi-robot morphogenesis using directional self-assembly. *Trans. Robot.*, 25(3):738–743, 2009.
- [20] A Padirac, T Fujii, A Estevez-Torres, and Y Rondelez. Spatial waves in synthetic biochemical networks. *JACS*, 135(39):14586–14592, 2013.
- [21] JK Pugh, LB Soros, and KO Stanley. Quality diversity: A new frontier for evolutionary computation. *Frontiers in Robotics and AI*, 3:40, 2016.
- [22] NV Queipo, RT Haftka, W Shyy, T Goel, R Vaidyanathan, and PK Tucker. Surrogate-based analysis and optimization. *Progress in aerospace sciences*, 41(1):1–28, 2005.
- [23] M Rubenstein, A Cornejo, and R Nagpal. Programmable self-assembly in a thousand-robot swarm. *Science*, 345(6198):795–799, 2014.
- [24] KO Stanley and R Miikkulainen. Evolving neural networks through augmenting topologies. *Evolutionary computation*, 10(2):99–127, 2002.
- [25] David JT Sumpter. The principles of collective animal behaviour. *Proc. Royal Soc. B*, 361(1465):5–22, 2006.
- [26] E Tuci, R Groß, V Trianni, F Mondada, M Bonani, and M Dorigo. Cooperation through self-assembly in multi-robot systems. *TAAS*, 1(2):115–150, 2006.
- [27] W Yahiro, N Aubert-Kato, and M Hagiya. A reservoir computing approach for molecular computing. In *ALIFE*, pages 31–38. MIT Press, 2018.
- [28] AS Zadorin, Y Rondelez, G Gines, V Dilhas, G Urtel, A Zambrano, JC Galas, and A Estévez-Torres. Synthesis and materialization of a reaction-diffusion french flag pattern. *Nat. Chem.*, 9(10):990, 2017.