Robustness of Ant Colony Optimization to Noise

Tobias Friedrich   Timo Kötzing   Martin S. Krejca   Andrew M. Sutton

Hasso Plattner Institute
University of Potsdam
Potsdam, Germany

ABSTRACT

Recently Ant Colony Optimization (ACO) algorithms have been proven to be efficient in uncertain environments, such as noisy or dynamically changing fitness functions. Most of these analyses focus on combinatorial problems, such as path finding.

We analyze an ACO algorithm in a setting where we try to optimize the simple OneMax test function, but with additive posterior noise sampled from a Gaussian distribution. Without noise the classical $(\mu + 1)$-EA outperforms any ACO algorithm, with smaller μ being better; however, with large noise, the $(\mu + 1)$-EA fails, even for high values of μ (which are known to help against small noise). In this paper we show that ACO is able to deal with *arbitrarily large* noise in a graceful manner, that is, as long as the evaporation factor ρ is small enough dependent on the parameter σ^2 of the noise and the dimension n of the search space ($\rho = o(1/(n(n+\sigma \log n)^2 \log n)))$, optimization will be successful.

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity

Keywords

Ant colony optimization; noisy fitness; theory; run time analysis

1. INTRODUCTION

Ant colony optimization (ACO) is a class of randomized general-purpose optimization algorithms inspired by the foraging behavior of ant colonies. ACO has been successfully applied as a heuristic technique for solving combinatorial optimization problems.

In real-world optimization problems, there is sometimes a large degree of uncertainty present due to the complexity

of candidate solution generation, noisy measurement processes, and rapidly changing problem environments. Empirically, ACO seems particularly well-suited to uncertain problems due to its dynamic and distributed nature, and in some cases it can outperform classical state-of-the-art approaches on dynamic network routing problems [3]. We will focus on a version of the *Max-Min Ant System* (MMAS, [15]) applied to pseudo-Boolean optimization (i.e. optimization where solutions are coded as bit strings).

Jin and Branke [10] survey a number of sources of uncertainty that randomized search heuristics must often deal with in practice, (1) noisy objective functions, (2) dynamically changing problems, (3) approximation errors in the objective function, and (4) a requirement that an optimal solution must be robust to changes in design variables and environmental parameters that occur after optimization is complete. Arguably, the two most important sources of uncertainty are (1) and (2), namely, *stochastic* problems and *dynamic* problems (see also [1] for a recent survey). In stochastic problems, the objective function value of a search point follows a random distribution, and that distribution does not change over time. In dynamic problems, the evaluation of the fitness is deterministic but changes over time.

In order to address these practical issues, the theoretical analyses of randomized search heuristics under uncertainty has recently gained momentum. For example, a number of recent papers rigorously analyzed the performance of evolutionary algorithms in stochastic environments [2, 9]. For ant colony optimization, a series of papers considered the performance of ACO on single destination shortest paths (SDSP) problems with stochastic weights. This work was initiated by Sudholt and Thyssen [16] and later followed up by Doerr, Hota, and Kötzing [4], who showed that by augmenting the ant system with a re-evaluation strategy on the best-so-far solution, many of the difficulties with noise discovered in [16] could be overcome. Feldmann and Kötzing [7] recently showed that an ant system that uses a fitness-proportional update rule (called MMAS-fp) can efficiently optimize SDSP on graphs with stochastic weights. MMAS-fp is closer to systems that are used by practitioners [15] and is the ant system variant that we analyze in the current paper.

For the optimization of functions over bit strings, analyses of ACO suggest that it often performs worse than evolutionary algorithms and simple hill-climbers in a noise-free setting [13]. On the other hand, ACO can outperform evolutionary algorithms on dynamic problems [12, 14]. So far,
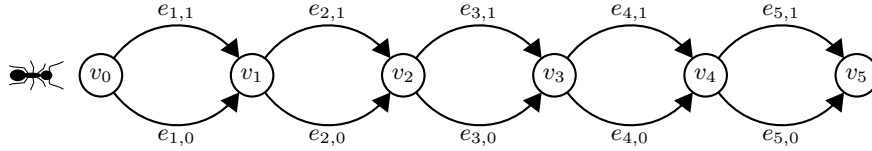
Figure 1: Construction graph for pseudo-Boolean optimization with $n = 5$ bits.

the question of how robust ACO is to noisy evaluation on pseudo-Boolean optimization remains unanswered.

The goal of this paper is to observe the robustness of ACO to noise on a simple objective function. In particular, we are interested in the *scalability* of the run time of the algorithm as a function of noise intensity (measured by variance). We study the algorithm on the well-known OneMax function, the most common starting point for rigorous analyses of randomized search heuristics. In Section 3 we show that robustness can be achieved for additive posterior noise from a Gaussian distribution: for all variances σ^2, there is a parameter setting such that MMAS-fp on OneMax $+ \mathcal{N}(0, \sigma^2)$ is successful in time $\sigma^2 \cdot \text{poly}(n)$; thus, we say that MMAS-fp handles Gaussian noise *gracefully*. Such a graceful scaling cannot be achieved by the $(\mu + 1)$ EA, as shown in [8].

A similar statement holds for an ACO algorithm which does not impose nontrivial Max-Min bounds on the pheromone values. This is achieved by not relying on the standard technique of *pheromone freezing* employed by many early works on the mathematical analysis of ACO algorithms, but instead we show that all pheromones *drift* in the right direction (in particular, they will never be close to any pheromone bound on the wrong side of the spectrum); see Corollary 6.

In Section 4 we extend our findings to other noise models and show that we achieve the same robustness also in the presence of other additive noise distributions (which fulfill a certain restriction) as well as for the model of prior noise from [6]. We discuss our findings and conclude the paper in Section 5.

2. PRELIMINARIES

We consider optimizing pseudo-Boolean functions, i.e., functions of type $f: \{0,1\}^n \to \mathbb{R}$. In the following, n always denotes the dimension of the solution space $\{0,1\}^n$. For any bit string $x \in \{0,1\}^n$, $f(x)$ is called the *fitness of x*, and we are interested in finding a solution x^* such that $f(x^*)$ is maximal. The fitness function we are going to investigate is OneMax. This function simply returns the amount of ones in a bit string, i.e., $\text{OneMax}(x) = \|x\|_1$. So the unique optimum is the all ones string.

We consider, however, a noisy version of OneMax. The noise perturbing the actual function value will be a Gaussian distributed random variable D with mean 0 and variance σ^2 that is added to the OneMax value. More formally, $\text{OM}_{\sigma^2}(x) := \text{OneMax}(x) + D$ with $D \sim \mathcal{N}(0, \sigma^2)$.

2.1 Algorithms

Our main algorithm of interest is MMAS-fp *(Min-Max Ant System)*, an ACO algorithm with a so-called *fitness-proportional* update rule.

To find an optimal solution, MMAS-fp keeps a vector $\tau \in [\ell, u]^n$, the components of which are called *pheromones*,

whereas $0 \le \ell < 1/2$ and $1 \ge u > 1/2$ are the lower and upper bound for each pheromone, respectively. $\tau^{(t)}$ with $t \in \mathbb{N}$ denotes the pheromone vector in the t-th iteration of the algorithm.

MMAS-fp starts off with every pheromone $\tau_i^{(0)} = 1/2$ and iteratively generates solutions $x^{(t)} \in \{0,1\}^n$ until $x^{(t)}$ is optimal. These solutions $x^{(t)}$ are sampled accordingly to $\tau^{(t)}$ in the manner $\forall i: \Pr(x_i^{(t)} = 0) = \tau_i^{(t)} \wedge \Pr(x_i^{(t)} = 1) = 1 - \tau_i^{(t)}$. This is equivalent to a definition of using a simulated ant walk on the construction graph as shown in Figure 1. Afterward, the pheromone vector for the next iteration is updated accordingly to the fitness of the sampled solution, i.e.,

$$\tau_i^{(t+1)} := \begin{cases} \min\left\{\tau_i^{(t)}\left(1 - \rho\frac{f(x^{(t)})}{n}\right) + \rho\frac{f(x^{(t)})}{n}, u\right\} : x_i^{(t)} = 0, \\ \max\left\{\tau_i^{(t)}\left(1 - \rho\frac{f(x^{(t)})}{n}\right), \ell\right\} \qquad\qquad : x_i^{(t)} = 1, \end{cases}$$

for all $i \in \{1, \ldots, n\}$.

The parameter $\rho \in [0, 1]$ is the so-called *evaporation factor* and a parameter of the algorithm. Intuitively, it regulates the impact of the fitness of a sampled solution on the corresponding pheromone update.

Note that MMAS-fp does not keep a best-so-far solution but, instead, always updates its pheromone vector regardless of the quality of the solution sampled. The update is proportional to the fitness of each sampled solution, hence the suffix *fp*. If an updated pheromone is not capped due to the pheromone bounds, we speak of a *normal* update.

In the following, we will denote a pheromone vector $\tau^{(t)}$ simply as τ if we are not interested in t. Likewise, x denotes the solution sampled by τ, and τ' denotes the pheromone vector after the update according to x.

The other algorithm we consider is ACO-fp. It is a special case of MMAS-fp where $\ell = 0$ and $u = 1$, i.e., only trivial pheromone bounds are enforced.

2.2 Tools Used

The proofs in this paper rely heavily on *drift theory*, a tool that allows to give upper and lower bounds on the hitting times of randomly moving processes if there is a bias in the process's motion, the so-called *drift*.

We state two drift theorems that we are going to use in our proofs.

Theorem 1 (Multiplicative Drift [5])**.** *Let $(X_t)_{t \in \mathbb{N}}$ be non-negative random variables over $\mathbb{R}$, each with finite expectation, and let $T = \min\{t \mid X_t < 1\}$.*

Suppose there exists an $\varepsilon > 0$ such that, for all t,

$$\mathrm{E}(X_t - X_{t+1} \mid X_t, t < T) \ge \varepsilon X_t \ .$$

Then

$$\mathrm{E}(T \mid X_0) \le \frac{1 + \ln X_0}{\varepsilon} \ .$$

Algorithm 1: MMAS-fp with $\rho \in (0,1]$, for optimizing $f \colon \{0,1\}^n \to \mathbb{R}$

---

1 **for** $i \in \{1, \ldots, n\}$ **do**
2 $\quad \lfloor \ \tau_i \leftarrow \frac{1}{2}$;
3 **while** optimum not found **do**
4 $\quad$ **for** $i \in \{1, \ldots, n\}$ **do**
5 $\quad\quad \lfloor \ x_i \leftarrow 0$ with probability τ_i, $x_i \leftarrow 1$ with probability $1 - \tau_i$;
6 $\quad$ **for** $i \in \{1, \ldots, n\}$ **do**
7 $\quad\quad$ **if** $x_i = 0$ **then**
8 $\quad\quad\quad \lfloor \ \tau_i \leftarrow \min\left\{\tau_i\left(1 - \rho\frac{f(x)}{n}\right) + \rho\frac{f(x)}{n}, u\right\}$;
9 $\quad\quad$ **else**
10 $\quad\quad\quad \lfloor \ \tau_i \leftarrow \max\left\{\tau_i\left(1 - \rho\frac{f(x)}{n}\right), \ell\right\}$;

Theorem 2 (Negative Drift [11])**.** *Let $(X_t)_{t\in\mathbb{N}}$ be random variables over $\mathbb{R}$, each with finite expectation, let $d > 0$, and let $T = \min\{t\colon X_t \geq d \mid X_0 \leq 0\}$. Suppose there are $s, 0 < s < d$ and $\varepsilon < 0$ such that, for all t,*

1. $\mathrm{E}(X_{t+1} - X_t \mid X_t, t < T) \leq \varepsilon$, and

2. $|X_t - X_{t+1}| < s$.

Then, for all $t \in \mathbb{N}$,

$$\Pr(T \leq t) \leq t \exp\left(-\frac{d|\varepsilon|}{16s^2}\right) \ .$$

The following proposition gives tail bounds for our noise $D \sim \mathcal{N}(0, \sigma^2)$ by using standard estimates of the complementary error function [17].

Proposition 3. *Let D be a zero-mean Gaussian random variable with variance σ^2. For all $t > 0$, we have*

$$\Pr\left(D < -t\right) = \frac{1}{2}\operatorname{erfc}\left(\frac{t}{\sigma\sqrt{2}}\right) \leq \frac{1}{2}\exp\left(\frac{-t^2}{2\sigma^2}\right) \ .$$

3. GAUSSIAN NOISE

In this section we are going to bound the run time of MMAS-fp for optimizing OM_{σ^2}, which will be denoted as f in the proofs.

First off, we bound the noise and thus the difference of two consecutive pheromones with high probability.

Lemma 4. *Consider MMAS-fp optimizing OM_{σ^2}. Then we can bound $|\tau_i - \tau_i'|$ for any pheromone τ_i with high probability as follows:*

$$|\tau_i - \tau_i'| \leq \frac{\rho}{n}(n + \mathcal{O}(\sigma \log n)) \ .$$

Proof. We make a case distinction for τ_i' and assume $\ell = 0$ and $u = 1$ since capping a pheromone would only make the difference smaller. Because of that, it does not even matter if a pheromone would exceed the bounds $[0, 1]$.

1. $\tau_i' = \tau_i\left(1 - \rho\frac{f(x)}{n}\right)$:

$$|\tau_i - \tau_i'| = \left|\tau_i - \tau_i\left(1 - \rho\frac{f(x)}{n}\right)\right|$$

$$= \frac{\rho}{n}|\overbrace{\tau_i}^{\leq 1} f(x)| \leq \frac{\rho}{n}|f(x)| \ .$$

2. $\tau_i' = \tau_i\left(1 - \rho\frac{f(x)}{n}\right) + \rho\frac{f(x)}{n}$:

$$|\tau_i - \tau_i'| = \left|\tau_i - \left(\tau_i\left(1 - \rho\frac{f(x)}{n}\right) + \rho\frac{f(x)}{n}\right)\right|$$

$$= \frac{\rho}{n}|\overbrace{(\tau_i - 1)}^{\geq -1} f(x)| \leq \frac{\rho}{n}|f(x)| \ .$$

We bound $|f(x)| = \big|\|x\|_1 + D\big|$. By Proposition 3 we get:

$$\Pr(D > t) = \Pr(D < -t) \leq \frac{1}{2}\exp\left(-\frac{t^2}{2\sigma^2}\right) \ .$$

For $t = \omega(\sigma \log n)$ this probability is superpolynomially small. So we can bound $|f(x)| \leq n + \mathcal{O}(\sigma \log n)$ and get

$$|\tau_i - \tau_i'| \leq \frac{\rho}{n}(n + \mathcal{O}(\sigma \log n)) \ . \qquad \square$$

In the following, let $s = \rho(n + \mathcal{O}(\sigma \log n))/n$, i.e., the greatest difference between two consecutive pheromones, as stated by Lemma 4. We further assume the restrictions $\ell + s < 1/2 < u - s$ for the pheromone bounds.

Also, we always condition on the event that the noise does not exceed the bounds mentioned in the proof of Lemma 4, which holds with high probability in any polynomial number of steps of MMAS-fp. Other runs are discarded as fails.

Lemma 5. *Consider MMAS-fp optimizing OM_{σ^2}. Then, for all pheromone vectors, the drift of each pheromone $\tau_i = b \in [\ell + s, u - s]$ toward 1 is $-b(1 - b)\rho/n$.*

Proof. We condition on an arbitrary pheromone vector τ where $\tau_i = b \in [\ell + s, u - s]$, thus, the pheromone is updated normally.

Let A denote the event that x_i was sampled as 1, and consider the expected value of τ_i'. Regardless of A, τ_i decreases. Only in the case of $\overline{A}$ it is additionally increased.

$$\mathrm{E}(\tau_i') = \mathrm{E}\left(\tau_i\left(1 - \rho\frac{f(x)}{n}\right)\right) + \mathrm{E}\left(\rho\frac{f(x)}{n} \,\Big|\, \overline{A}\right)\overbrace{\Pr(\overline{A})}^{=\tau_i}$$

$$= b + b\frac{\rho}{n}\left(\mathrm{E}\big(f(x) \mid \overline{A}\big) - \mathrm{E}\big(f(x)\big)\right) \ .$$

We get the negative drift:

$$\mathrm{E}(\tau_i' - \tau_i) = b + b\frac{\rho}{n}\left(\mathrm{E}\big(f(x) \mid \overline{A}\big) - \mathrm{E}\big(f(x)\big)\right) - b$$

$$= b\frac{\rho}{n}\left(\mathrm{E}(n - \|x\|_0 + D \mid \overline{A}) - \mathrm{E}(n - \|x\|_0 + D)\right)$$

$$= b\frac{\rho}{n}\left(n - \left(1 + \sum_{\substack{j \\ j \neq i}}\tau_j\right) - \left(n - \left(b + \sum_{\substack{j \\ j \neq i}}\tau_j\right)\right)\right)$$

$$= b(b - 1)\frac{\rho}{n}$$

$$= -b(1 - b)\frac{\rho}{n} \ . \qquad \square$$

Corollary 6. *Consider MMAS-fp optimizing* OM_{σ^2}, *and let* $b \in [\ell, u]$, *with* $u \geq d$ *and* $d > 1/2 + s$ *being a constant. If* $\rho = \mathrm{o}\big(nb(1-b)/\big((n + \sigma \log n)^2 \log n\big)\big)$, *then each pheromone reaches values of at least* d *in polynomial time with only superpolynomially low probability.*

Proof. First, note that the drift of Lemma 5 is an upper bound on the drift for pheromones in $(u - s, u]$, in which case the additive term turns out to be smaller than it is in a normal update case.

We can therefore apply Theorem 2 by using the bound given in Lemma 4 and the negative drift of Lemma 5.

$$\Pr(T \leq t) \leq t \exp\left(-\frac{\overbrace{\left(d - \frac{1}{2}\right)}^{>0} b(1-b)\frac{\rho}{n}}{16\frac{\rho^2}{n^2}(n + \mathcal{O}(\sigma \log n))^2}\right)$$

$$= t \exp\left(-\frac{(d - \frac{1}{2})nb(1-b)}{16\rho(n + \mathcal{O}(\sigma \log n))^2}\right) .$$

This means that the probability for a single pheromone starting from $1/2$ to reach d in any polynomial amount of steps is superpolynomially small for $\rho = \mathrm{o}\big(nb(1-b)/((n + \sigma \log n)^2 \log n)\big)$, as given, if that pheromone is in $[\ell + s, u]$.

Now consider a pheromone to be in $[\ell, \ell + s)$. Even the best update could not get that pheromone above $1/2$ because $s = \rho(n + \mathcal{O}(\sigma \log n))/n = \mathrm{o}(1)$ with ρ as given. So we can apply the above argumentation.

Note that $s = \mathrm{o}(1) < d - 1/2$, as this is what we need to apply Theorem 2. $\square$

Using a union bound argument, we can now say that none of the polynomially many pheromones reaches d as in Corollary 6 in polynomially many steps with high probability.

Lemma 7. *Consider MMAS-fp optimizing* OM_{σ^2} *with* $\rho = \mathrm{o}\big(nb(1-b)/\big((n + \sigma \log n)^2 \log n\big)\big)$, $\ell = \mathrm{o}(1/n^2)$, *and* $u > 1/2 + s$. *Consider further that* $k \in [0, n] \cap \mathbb{N}$ *pheromones dropped below* $1/n$ $(\geq \ell + s)$. *If at least one pheromone is at least* $1/n$, *then the drift of the sum of all pheromones* τ_{total} *is* $\Omega(\tau_{\mathrm{total}}\rho/n^2)$.

Proof. In this proof we pessimistically assume pheromones below $1/n$ driftwise (and only driftwise!) as having reached ℓ. Note that, thus, there cannot be any positive drift if $k = n$ since none of the pheromones could drop any lower.

We split up the overall drift of the sum of all pheromones into the (positive) drift (toward 0) of the $n - k$ pheromones that are at least $1/n$ and into the (negative) drift of the k pheromones that are below $1/n$.

Let Y denote the index set of those former $n - k$ pheromones, Z the index set of the latter k, and let K be the event that exactly k pheromones are below $1/n$. The string x^y denotes the bit string consisting only of those elements of x whose index is in Y, and x^z is defined analogously with respect to Z. The sum of pheromones with index in Y is denoted by τ_{tot}^y and the sum of those with index in Z by τ_{tot}^z.

First, we have a look at the $n - k$ pheromones. Each of these pheromones drops and all of them which got a zero sampled for their respective bit position get at most a normal

update (in the case of capped pheromones the drift would be even greater).

$$\mathrm{E}(\tau_{\mathrm{tot}}^y - \tau_{\mathrm{tot}}^{y\,\prime} \mid \overbrace{K, \tau_{\mathrm{tot}}^y = b}^{:=B})$$

$$\geq \mathrm{E}\left(\tau_{\mathrm{tot}}^y - \left(\tau_{\mathrm{tot}}^y\left(1 - \rho\frac{f(x)}{n}\right) + \|x^y\|_0\rho\frac{f(x)}{n}\right) \,\Big|\, B\right)$$

$$= \frac{\rho}{n}\mathrm{E}\left(\tau_{\mathrm{tot}}^y f(x) - \|x^y\|_0 f(x) \mid B\right)$$

$$= \frac{\rho}{n}\Big(b\mathrm{E}(f(x) \mid B) - \mathrm{E}(\|x^y\|_0 f(x) \mid B)\Big) .$$

We give bounds for the expected values needed.

$$\mathrm{E}(f(x) \mid B) = \mathrm{E}(n - \|x\|_0 + D \mid B) = n - (b + k\ell) ,$$

and

$$\mathrm{E}(\|x^y\|_0 f(x) \mid B) = \mathrm{E}(\|x^y\|_0(n - \|x\|_0 + D) \mid B)$$
$$= n\mathrm{E}(\|x^y\|_0 \mid B) - \mathrm{E}(\|x^y\|_0\|x\|_0 \mid B)$$
$$= nb - \mathrm{E}(\|x^y\|_0\|x\|_0 \mid B) .$$

Let a be a variable ranging over Y; j ranges over all pheromone indices. Let $[\cdot]$ denote the Iverson bracket (indicator function). We make use of Corollary 6.

$$\mathrm{E}(\|x^y\|_0\|x\|_0 \mid B) = \mathrm{E}\left(\sum_a \sum_j [x_a = 0][x_j = 0] \,\Big|\, B\right)$$

$$= \sum_a \sum_j \mathrm{E}\left([x_a = 0][x_j = 0] \mid B\right)$$

$$= \sum_a \sum_{\substack{j \\ j \neq a}} \mathrm{E}\left([x_a = 0] \mid B\right)\mathrm{E}\left([x_j = 0] \mid B\right) +$$

$$\quad \sum_a \mathrm{E}\left([x_a = 0][x_a = 0] \mid B\right)$$

$$= \sum_a \sum_{\substack{j \\ j \neq a}} \tau_a \tau_j + \sum_a \tau_a$$

$$= \sum_a \tau_a \sum_j \tau_j - \sum_a \tau_a^2 + \sum_a \tau_a$$

$$= \sum_a \tau_a \sum_j \tau_j + \sum_a \tau_a(1 - \overbrace{\tau_a}^{\leq d})$$

$$\geq b(b + k\ell) + (1 - d)b .$$

Now we can calculate the desired drift.

$$\mathrm{E}(\tau_{\mathrm{tot}}^y - \tau_{\mathrm{tot}}^{y\,\prime} \mid B)$$

$$\geq \frac{\rho}{n}\left(b\left(n - (b + k\ell)\right) - (nb - b(b + k\ell) - (1 - d)b)\right)$$

$$= \frac{\rho}{n}(1 - d)b .$$

For the negative drift, we have to look at τ_{tot}^z and $\tau_{\mathrm{tot}}^{z\,\prime}$. All of the k pheromones where the corresponding bit position was sampled as 1 cannot drop any lower due to our assumption. The remaining of the k pheromones get a normal update in the best case:

$$\tau_{\text{tot}}^z{}' \le \|x^z\|_1 \ell + \|x^z\|_0 \left(\ell \left(1 - \rho \frac{f(x)}{n} \right) + \rho \frac{f(x)}{n} \right)$$

$$= \overbrace{\|x^z\|_1 \ell + \|x^z\|_0 \ell}^{=k\ell} - \|x^z\|_0 \ell \rho \frac{f(x)}{n} + \|x^z\|_0 \rho \frac{f(x)}{n}$$

$$= k\ell + \frac{\rho}{n}(1-\ell)\|x^z\|_0 f(x) \ .$$

Thus we get:

$$\mathrm{E}(\tau_{\text{tot}}^z - \tau_{\text{tot}}^z{}' \mid B) \ge \mathrm{E}\left(k\ell - k\ell - \frac{\rho}{n}(1-\ell)\|x^z\|_0 f(x) \ \Big| \ B \right)$$

$$= -\frac{\rho}{n}(1-\ell)\mathrm{E}\left(\|x^z\|_0 f(x) \mid B \right) \ .$$

The remaining estimations follow analogously to the ones beforehand:

$$\mathrm{E}(\|x^z\|_0 f(x) \mid B) = \mathrm{E}(\|x^z\|_0(n - \|x\|_0 + D) \mid B)$$

$$= n\mathrm{E}(\|x^z\|_0 \mid B) - \mathrm{E}(\|x^z\|_0 \|x\|_0 \mid B)$$

$$= nk\ell - \mathrm{E}(\|x^z\|_0 \|x\|_0 \mid B) \ .$$

Note that a now refers to indices in Z.

$$\mathrm{E}(\|x^z\|_0 \|x\|_0 \mid B) = \sum_a \tau_a \sum_j \tau_j + \sum_a \tau_a (1 - \tau_a)$$

$$\ge k\ell(b + k\ell) + k\ell\,(1 - d)$$

$$= k\ell\,(b + k\ell + (1-d)) \ .$$

So we get the negative drift:

$$\mathrm{E}(\tau_{\text{tot}}^z - \tau_{\text{tot}}^z{}' \mid B) \ge -\frac{\rho}{n}(1-\ell)k\ell\,(n - b - k\ell - (1-d)) \ .$$

Now we can look at the general drift:

$$\mathrm{E}(\tau_{\text{total}} - \tau_{\text{total}}' \mid B)$$

$$\ge \frac{\rho}{n}(1-d)b - \frac{\rho}{n}(1-\ell)k\ell\,(n - b - k\ell - (1-d))$$

$$= \frac{\rho}{n}\left((1-d)b - (1-\ell)\left(nk\ell - bk\ell - k^2\ell^2 - (1-d)k\ell \right) \right) \ ,$$

which we want to be positive.

Recall that we only consider $k \in [0, n-1] \cap \mathbb{N}$. We bound b by $(n-k)/n$ since all these pheromones are at least $1/n$.

Thus, we end up having:

$$(1-d)(n-k)/n - (1-\ell)\Big(nk\ell - (n-k)k\ell/n - k^2\ell^2 -$$

$$(1-d)k\ell \Big) \ .$$

On inspection one sees that the highest order terms in the minuend are $n/n = 1$ and $-k/n$. Since we want to have a positive drift, $1 > k/n$ must hold. Because we are interested in $k \in [0, n-1] \cap \mathbb{N}$, this does hold.

The highest order term in the subtrahend is $-nk\ell$. It is absolutely largest for $k \in \Theta(n)$, i.e, $n^2\ell$. To avoid any conflict with 1, we need $1 \in \omega(n^2\ell)$, i.e.,

$$\lim_{n \to \infty} \frac{1}{n^2\ell} = \infty \ ,$$

which holds for $\ell = o(1/n^2)$, as given. This results in the drift being in $\Omega(b\rho/n)$.

Since we want the drift to be expressed in terms of τ_{total}, we need $b\rho/n \ge \varepsilon\tau_{\text{total}}$ to hold for an $\varepsilon > 0$.

We can bound $\tau_{\text{total}} < b + k/n$ by pessimistically assuming all k pheromones below $1/n$ as being $1/n$. So we want $b\rho/n \ge \varepsilon(b + k/n)$ to hold for all valid values of b. This is equivalent to $1/\varepsilon \ge (n + k/b)/\rho$. To make k/b as large as possible we bound $b \le (n-k)/n$, and we get $1/\varepsilon \ge (n + nk/(n-k))/\rho$, where $k/(n-k)$ is maximal for $k = n - 1$ (we do not consider $k = n$). This results in $\varepsilon = \mathcal{O}(\rho/n^2)$.

Overall, we finally get a drift of $\Omega(\tau_{\text{total}}\rho/n^2)$. $\qquad\square$

Theorem 8. *Consider MMAS-fp optimizing* OM_{σ^2} *with* $\rho = \mathrm{o}(1/(n(n + \sigma\log n)^2 \log n))$, $\ell = \mathrm{o}(1/n^2)$, *and* $u > 1/2 + s$. *The algorithm then finds the optimum after* $\mathcal{O}(n^2\log(n)/\rho)$ *steps with high probability.*

Proof. Let T denote the run time of the algorithm when the drift is positive, i.e., $T = \min\{t \in \mathbb{N} \mid \forall i\colon \tau_i^{(t)} < 1/n\}$.

According to Lemma 7, there is a positive drift of the pheromone sum toward 0 of at least $\tau_{\text{total}}\rho/n^2$.

With this multiplicative drift we get an upper bound on the expectation of T using Theorem 1:

$$\mathrm{E}\left(T \ \Big| \ \forall i\colon \tau_i^{(0)} = \frac{1}{2} \right) = \mathcal{O}\left(\frac{1 + \log \frac{n}{2}}{\frac{\rho}{n^2}} \right) = \mathcal{O}\left(\frac{n^2\log n}{\rho} \right) \ .$$

Now assume that all pheromones dropped below $1/n$. The probability to sample the optimum now is $(1 - 1/n)^n \ge (1 - 1/n)/e$. Thus, the optimum will be sampled in expectation in $en/(n-1) = \mathcal{O}(1)$ many tries.

The overall expected run time is bounded above by the following worst case scenario: the algorithm needs $\mathcal{O}(n^2\log(n)/\rho)$ steps until all pheromones drop below $1/n$. Then the algorithm does not sample the optimum and all pheromones are set back to $1/2$ (that does not actually happen but we are looking at a worst case). As mentioned before, in this scenario, the algorithm would need $\mathcal{O}(1)$ tries to sample the optimum, once the pheromones are low enough. Hence, the algorithm hast to be restarted $\mathcal{O}(1)$ times.

Overall, this results in an expected run time in $\mathcal{O}(n^2\log(n)/\rho)$ with high probability due to us conditioning on the noise being bounded, as stated in Lemma 4. $\qquad\square$

Corollary 9. *Consider ACO-fp optimizing* OM_{σ^2}. *If* $\rho = \mathrm{o}(1/(n(n+\sigma\log n)^2\log n))$, *the algorithm finds the optimum after* $\mathcal{O}(n^2\log(n)/\rho)$ *steps with high probability.*

Proof. Recall that the proof Lemma 4 is actually a proof for ACO-fp as well. We can thus argue analogously as in the proof of Theorem 8. The upper bound d is at most 1 because $d \le u \le 1$. The lower bound of $\ell = 0$ for ACO-fp satisfies $\ell = \mathrm{o}(1/n^2)$ as well. $\qquad\square$

4. OTHER NOISE MODELS

In this section we consider the optimization of OneMax perturbed by other noise than additive posterior Gaussian noise.

4.1 Posterior Noise

We start with a generalization of Corollary 9 by taking non-Gaussian noise into account, i.e., we optimize OM_D with $\mathrm{OM}_D(x) := \|x\|_1 + D$ and D being a random variable (possibly not Gaussian).

Corollary 10. *Consider MMAS-fp optimizing* OM_D *with* $\ell = o(1/n^2)$ *and* $u > 1/2 + s$, *where* D *is a random variable such that, for some* $t^* > 0$ *and all polynomials* p, *the following holds:*

$$\Pr(|D| > t^*) \leq 1 - o\left(\frac{1}{p(n)}\right) . \qquad (1)$$

If $\rho = o(1/(n(n + t^*)^2 \log n))$, *then the algorithm finds the optimum after* $\mathcal{O}(n^2 \log(n)/\rho)$ *steps with high probability.*

Proof. The drifts of Lemmas 5 and 7 hold here as well since the noise is additive and posterior and, thus, it cancels out in expectation.

To be able to use Corollary 6, we need to bound $|f(x)|$ with high probability. Because of Equation (1) we can do so, and it follows that any pheromone τ_i reaches values of at least a constant $d > 1/2 + s$ with superpolynomially low probability for $\rho = o(1/(n(n + t^*)^2 \log n))$.

We can now use an analogous argumentation as in Theorem 8, proving the proposition. $\qquad\square$

Corollary 11. *Consider ACO-fp optimizing* OM_D, *where* D *is a random variable such that, for some* $t^* > 0$ *and all polynomials* p, *the following holds:*

$$\Pr(|D| > t^*) \leq 1 - o\left(\frac{1}{p(n)}\right) .$$

If $\rho = o(1/(n(n + t^*)^2 \log n))$, *then the algorithm finds the optimum after* $\mathcal{O}(n^2 \log(n)/\rho)$ *steps with high probability.*

Proof. We can use the same same proof as for Corollary 10 and argue as in the proof of Corollary 9 that we do not need the special pheromone bounds. $\qquad\square$

4.2   Prior Noise

In this section we have a look at the noise model from [6], where $OM_p(x) := OneMax(\widetilde{x})$, whereas $\widetilde{x} = x$ with probability $(1 - p)$, and $\widetilde{x} = x'$ with $d_H(x, x') = 1$ for each such x' with probability p/n. That means that with probability p a single bit in x gets flipped uniformly at random.

Note that we can bound $|\tau_i - \tau_i'| \leq \rho$ in the same way as we did in the proof of Lemma 4 since $|f(x)| \leq n$. We again assume that $\ell + \rho < 1/2 < u - \rho$.

Lemma 12. *Consider MMAS-fp optimizing* OM_p. *Then, for all pheromone vectors, the drift of each pheromone* $\tau_i = b \in [\ell + \rho, u - \rho]$ *toward 1 is* $-b(1 - b)(1 - 2p/n)\rho/n$.

Proof. As the noise does not have any influence on the update of the pheromones and we have a normal update, we get the same basic equation for the desired drift as in Lemma 5. Let A denote the event that x_i was sampled as 1, and condition on any pheromone vector τ where $\tau_i = b \in [\ell + \rho, u - \rho]$.

$$E(\tau_i' - \tau_i) = b\frac{\rho}{n}\left(E\left(f(x) \mid \overline{A}\right) - E(f(x))\right) .$$

Let P denote the event that $\widetilde{x} = x'$ during the evaluation of f, i.e., a bit flipped, and let j and k range from 1 to n.

For $E\left(f(x) \mid \overline{A}\right)$ we have

$$E\left(f(x) \mid \overline{A}\right)$$
$$= E\left(f(x) \mid \overline{A}, \overline{P}\right)(1 - p) + E\left(f(x) \mid \overline{A}, P\right)p$$
$$= \left(n - \left(1 + \sum_{\substack{j \\ j \neq i}}\right)\right)(1 - p) +$$
$$\left(n - \sum_{\substack{k \\ k \neq i}}\frac{1}{n}\left(1 + (1 - \tau_k) + \sum_{\substack{j \\ j \neq k \\ j \neq i}}\tau_j\right) - \frac{1}{n}\sum_{\substack{j \\ j \neq i}}\tau_j\right)p .$$

And for $E(f(x) \mid \tau_i = b)$ we get

$$E(f(x)) = E\left(f(x) \mid \overline{P}\right)(1 - p) + E\left(f(x) \mid P\right)p$$
$$= \left(n - \left(b + \sum_{\substack{j \\ j \neq i}}\tau_j\right)\right)(1 - p) +$$
$$\left(n - \sum_{\substack{k \\ k \neq i}}\frac{1}{n}\left(b + (1 - \tau_k) + \sum_{\substack{j \\ j \neq k \\ j \neq i}}\tau_j\right) - \right.$$
$$\left.\frac{1}{n}\left((1 - b) + \sum_{\substack{j \\ j \neq i}}\tau_j\right)\right)p .$$

Thus we get

$$E\left(f(x) \mid \overline{A}\right) - E(f(x))$$
$$= -(1 - b)(1 - p) + \left(\frac{n - 1}{n}(b - 1) - \frac{1}{n}(b - 1)\right)p$$
$$= -(1 - b)(1 - p) - (1 - b)\frac{n - 2}{n}p$$
$$= -(1 - b)\left(1 - \frac{2}{n}p\right) .$$

Finally, we have a negative drift as desired:

$$E(\tau_i' - \tau_i) = -b(1 - b)\left(1 - \frac{2}{n}p\right)\frac{\rho}{n} . \qquad\square$$

Corollary 13. *Consider MMAS-fp optimizing* OM_p, *and let* $b \in [\ell, u]$, *with* $u \geq d$ *and* $d > 1/2 + \rho$ *being a constant. If* $\rho = o\left(b(1 - b)\left(1 - \frac{2}{n}p\right)/(n \log n)\right)$, *then each pheromone reaches values of at least* d *in polynomial time with only superpolynomially low probability.*

Proof. The proof is basically the same as the one of Corollary 6. Using again Theorem 2, for all $t \geq 0$, the following proposition holds for T as the hitting time of one pheromone reaching a value of at least d:

$$\Pr(T \leq t) \leq t \exp\left(-\frac{\left(d - \frac{1}{2}\right)b(1 - b)\left(1 - \frac{2}{n}p\right)\frac{\rho}{n}}{16\rho^2}\right)$$
$$= t \exp\left(-\frac{\left(d - \frac{1}{2}\right)b(1 - b)\left(1 - \frac{2}{n}p\right)}{16\rho n}\right) .$$

The probability for a single pheromone starting from $1/2$ to reach d in any polynomial amount of steps is superpolynomially small for $\rho = o\left(b(1 - b)\left(1 - \frac{2}{n}p\right)/(n \log n)\right)$. $\qquad\square$

Again, a union bound argument gives us the guarantees of Corollary 13 for all pheromones.

Lemma 14. *Consider MMAS-fp optimizing* OM_p *with* $\rho = \mathrm{o}\big(b(1-b)\big(1-\frac{2}{n}p\big)/(n\log n)\big)$, $\ell = \mathrm{o}(1/n^2)$, *and* $u > 1/2 + \rho$. *Consider further that* $k \in [0,n] \cap \mathbb{N}$ *pheromones dropped below* $1/n$ ($\geq l + \rho$). *If at least one pheromone is at least* $1/n$, *then the drift of the sum of all pheromones* τ_{total} *is* $\Omega(\tau_{\mathrm{total}}\rho/n^2)$.

Proof. We are going to use the same notation as in the proof of Lemma 7. P shall, again, denote the event that the bit flip occurred.

As mentioned before, the kind of noise has no impact on the general form of the drifts. So for the $n - k$ pheromones we have

$$\mathrm{E}(\tau_{\mathrm{tot}}^y - {\tau_{\mathrm{tot}}^y}' \mid B) \geq \frac{\rho}{n}\Big(b\mathrm{E}(f(x) \mid B) - \mathrm{E}\left(\|x^y\|_0 f(x) \mid B\right)\Big).$$

The required expected values can now be calculated:

$$\mathrm{E}(f(x) \mid B) = \mathrm{E}(f(x) \mid B, \overline{P})(1-p) + \mathrm{E}(f(x) \mid B, P)p$$
$$= (n - (b + k\ell))(1-p) + (n - \mathrm{E}(\|x\|_0 \mid B, P))\,p\,.$$

Furthermore, we get

$$\mathrm{E}(\|x\|_0 \mid B, P) = \sum_i \frac{1}{n}\Big((1 - \tau_i) + \sum_{\substack{j \\ j \neq i}} \tau_j\Big)$$

$$= \frac{1}{n}\sum_i (1 - \tau_i + b + k\ell - \tau_i)$$

$$= \frac{1}{n}\sum_i (b + k\ell + 1) - \frac{2}{n}\sum_i \tau_i$$

$$= b + k\ell + 1 - \frac{2}{n}(b + k\ell)\,.$$

This results in

$$\mathrm{E}(f(x) \mid B) = (n - b - k\ell) +$$
$$p\left(-n + b + k\ell + n - b - k\ell - 1 + \frac{2}{n}(b + k\ell)\right)$$

$$= n - b - k\ell + p\left(\frac{2}{n}(b + k\ell) - 1\right)\,.$$

Now we get for $\mathrm{E}(\|x^y\|_0 f(x) \mid B)$

$$\mathrm{E}(\|x^y\|_0 f(x) \mid B) = nb - \mathrm{E}(\|x^y\|_0 \|\widetilde{x}\|_0 \mid B)\,,$$

and by using Corollary 13

$$\mathrm{E}(\|x^y\|_0 \|\widetilde{x}\|_0 \mid B)$$
$$= \sum_a \sum_{\substack{j \\ j \neq a}} \mathrm{E}\left([x_a = 0] \mid B\right)\mathrm{E}\left([\widetilde{x}_j = 0] \mid B\right) +$$
$$\sum_a \mathrm{E}\left([x_a = 0] \mid B\right)$$

$$= \sum_a \sum_{\substack{j \\ j \neq a}} \tau_a\left(\tau_j(1-p) + \tau_j p\frac{n-1}{n} + (1 - \tau_j)\frac{p}{n}\right) + \sum_a \tau_a$$

$$= \sum_a \sum_{\substack{j \\ j \neq a}} \tau_a\left(\tau_j \frac{n - 2p}{n} + \frac{p}{n}\right) + \sum_a \tau_a$$

$$= \frac{n - 2p}{n}\sum_a \sum_{\substack{j \\ j \neq a}} \tau_a\tau_j + \frac{p}{n}\sum_a \sum_{\substack{j \\ j \neq a}} \tau_a + \sum_a \tau_a$$

$$= \frac{n - 2p}{n}\sum_a \sum_j \tau_a\tau_j - \frac{n - 2p}{n}\sum_a \tau_a^2 +$$

$$\sum_a \tau_a + \frac{p}{n}\sum_a \sum_{\substack{j \\ j \neq a}} \tau_a$$

$$\overset{b \leq d}{\geq} \frac{n - 2p}{n}b(b + k\ell) + \sum_a \tau_a\left(1 - \frac{n - 2p}{n}d\right) + b\frac{n-1}{n}p$$

$$= \frac{n - 2p}{n}b(b + k\ell) + b\left(1 - \frac{n - 2p}{n}d\right) + b\frac{n-1}{n}p\,.$$

Thus we get

$$\mathrm{E}(\|x^y\|_0 f(x) \mid B)$$
$$\leq nb - \frac{n - 2p}{n}b(b + k\ell) - b\left(1 - \frac{n - 2p}{n}d\right) - b\frac{n-1}{n}p\,.$$

Now we can calculate the desired drift as follows.

$$\mathrm{E}(\tau_{\mathrm{tot}}^y - {\tau_{\mathrm{tot}}^y}' \mid B)$$
$$\geq \frac{\rho}{n}\left(b\left(n - b - k\ell + p\left(\frac{2}{n}(b + k\ell) - 1\right)\right) - nb + \right.$$
$$\left. \frac{n - 2p}{n}b(b + k\ell) + b\left(1 - \frac{n - 2p}{n}d\right) + b\frac{n-1}{n}p\right)$$

$$= \frac{\rho}{n}\left(nb - b(b + k\ell) - nb + b(b + k\ell) + b - \right.$$
$$\left. b\frac{dn - 2dp + p}{n}\right)$$

$$\geq \frac{\rho}{n}b\,(1 - d)\,.$$

Regarding the negative drift we have

$$\mathrm{E}(\tau_{\mathrm{tot}}^z - {\tau_{\mathrm{tot}}^z}' \mid B) = -\frac{\rho}{n}(1 - \ell)\mathrm{E}\left(\|x^z\|_0 f(x) \mid B\right)\,.$$

We also get

$$\mathrm{E}(\|x^z\|_0 f(x) \mid B) = nk\ell - \mathrm{E}(\|x^z\|_0 \|\widetilde{x}\|_0 \mid B)$$
$$\leq nk\ell - \frac{n - 2p}{n}k\ell(b + k\ell) + k\ell\left(1 - \frac{n - 2p}{n}d\right) +$$
$$k\ell\frac{n-1}{n}p$$

$$= k\ell\left(n - \frac{n - 2p}{n}(b + k\ell + d) + 1 + \frac{n-1}{n}p\right)\,.$$

Now we have the negative drift as follows

$$\mathrm{E}(\tau_{\mathrm{tot}}^z - {\tau_{\mathrm{tot}}^z}' \mid B)$$
$$\geq -\frac{\rho}{n}(1 - \ell)k\ell\left(n - \frac{n - 2p}{n}(b + k\ell + d) + 1 + \frac{n-1}{n}p\right)\,.$$

Taking all the parts together, we get the total drift

$$\mathrm{E}(\tau_{\mathrm{total}} - \tau_{\mathrm{total}}' \mid B)$$
$$\geq \frac{\rho}{n}b(1 - d) - \frac{\rho}{n}(1 - \ell)k\ell\left(n - \frac{n - 2p}{n}(b + k\ell + d) + \right.$$
$$\left. 1 + \frac{n-1}{n}p\right)$$

$$= \frac{\rho}{n}\left(b(1 - d) - (1 - \ell)k\ell\left(n - \frac{n - 2p}{n}(b + k\ell + d) + \right.\right.$$
$$\left.\left. 1 + \frac{n-1}{n}p\right)\right)\,,$$

which we want to be positive.

23

Again, we only consider $k \in [0, n-1] \cap \mathbb{N}$, and we use $b \geq (n-k)/n$. We thus end up having

$$(n-k)(1-d)/n - (1-\ell)k\ell \left(n - \frac{n-2p}{n}(b+k\ell+d) \right.$$
$$\left. + 1 + \frac{n-1}{n}p \right).$$

The highest order terms are the same as in the proof of Lemma 7, and the drift is again $\Omega(b\rho/n)$. So we can conclude analogously. □

Theorem 15. *Consider MMAS-fp with $\rho = \mathrm{o}(1/(n^3 \log n))$, $\ell = \mathrm{o}(1/n^2)$, and $u > 1/2 + \rho$ or ACO-fp optimizing* OM_p. *Both algorithms then find the optimum after* $\mathcal{O}(n^2 \log(n)/\rho)$ *steps with high probability.*

Proof. The argumentation is exactly as in the proof of Theorem 8 or Corollary 9, respectively, but now we use Lemmas 12, 14, and Corollary 13. The value of ρ has to be changed accordingly using Corollary 13. □

5. DISCUSSION AND SUMMARY

In this work we saw that two simple ACO algorithms on OneMax scale gracefully with noise for Gaussian distributions, i.e., the run time depends only linearly on the variance of the noise. We get similar results for many other noise models and for different ACO algorithms, suggesting that ACO algorithms are generally good for dealing with noise (at least in settings where the underlying fitness function is simple enough, as in the case of OneMax). Many of these settings are not solvable by simple hill climbers [6].

The analysis of meta-heuristics such as ACO on noisy fitness functions is of particular interest because this is one particular area where very few tailored approaches exist for designing efficient algorithms.

One drawback of the ACO algorithms analyzed in this paper is that the variance of the noise must be known in order to correctly set the evaporation factor ρ. This problem can be bypassed as follows: guess a variance of 1 and run the algorithm until it has a constant success probability if the guess was correct. If the optimum was not found so far, double the guess and repeat. This standard doubling scheme leads to a *noise-oblivious* ACO algorithm with an expected run time of at most a constant factor away from the ACO which knows the noise in advance.

Acknowledgements

6. REFERENCES

[1] L. Bianchi, M. Dorigo, L. Gambardella, and W. Gutjahr. A survey on metaheuristics for stochastic combinatorial optimization. *Natural Computing*, 8:239–287, 2009.

[2] D.-C. Dang and P. K. Lehre. Evolution under partial information. In *Proc. of GECCO'14*, pages 1359–1366, 2014.

[3] G. A. Di Caro, F. Ducatelle, and L. M. Gambardella. Theory and practice of Ant Colony Optimization for routing in dynamic telecommunications networks. In N. Sala and F. Orsucci, editors, *Reflecting interfaces: the complex coevolution of information technology ecosystems*, pages 185–216. Idea Group, 2008.

[4] B. Doerr, A. Hota, and T. Kötzing. Ants easily solve stochastic shortest path problems. In *Proc. of GECCO'12*, pages 17–24, 2012.

[5] B. Doerr, D. Johannsen, and C. Winzen. Multiplicative drift analysis. *Algorithmica*, 64(4):673–697, 2012.

[6] S. Droste. Analysis of the (1+1) EA for a Noisy OneMax. In *Proc. of GECCO'04*, pages 1088–1099, 2004.

[7] M. Feldmann and T. Kötzing. Optimizing expected path lengths with ant colony optimization using fitness proportional update. In *Proc. of FOGA'13*, pages 65–74, 2013.

[8] T. Friedrich, T. Kötzing, M. S. Krejca, and A. M. Sutton. The benefit of sex in noisy evolutionary search. `arXiv:1502.02793 [cs.NE]`, 2015.

[9] C. Gießen and T. Kötzing. Robustness of populations in stochastic environments. In *Proc. of GECCO'14*, pages 1383–1390, 2014.

[10] Y. Jin and J. Branke. Evolutionary optimization in uncertain environments—a survey. *IEEE Trans. on Evol. Comp.*, 9:303–317, 2005.

[11] T. Kötzing. Concentration of first hitting times under additive drift. In *Proc. of GECCO'14*, pages 1391–1397, 2014.

[12] T. Kötzing and H. Molter. ACO beats EA on a dynamic pseudo-boolean function. In *Proc. of PPSN'12*, pages 113–122, 2012.

[13] T. Kötzing, F. Neumann, D. Sudholt, and M. Wagner. Simple max-min ant systems and the optimization of linear pseudo-boolean functions. In *Proc. of FOGA'11*, pages 209–218. ACM, 2011.

[14] A. Lissovoi and C. Witt. MMAS versus population-based EA on a family of dynamic fitness functions. *Algorithmica*, 2015. (To appear).

[15] T. Stützle and H. H. Hoos. MAX-MIN Ant System. *Future Generation Computer Systems*, 16:889–914, 2000.

[16] D. Sudholt and C. Thyssen. A simple ant colony optimizer for stochastic shortest path problems. *Algorithmica*, 64:643–672, 2012.

[17] E. W. Weisstein. Erfc, 2015. From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Erfc.html.