The Runtime of an Ant Colony Optimization algorithm on Jump Functions

Riade Benbaki and Ziyad Benomar
supervised by Benjamin Doerr
École Polytechnique
Institut Polytechnique de Paris
Palaiseau
France

December 8, 2020

1 Introduction

Ant Colony Optimization (ACO) algorithms are a class of stochastic search heuristics, where solutions are constructed by randomly traversing a certain graph. The graph traversal is influenced by pheromone values associated to every edge, which are updated after every iteration depending on the constructed solution. This heuristic have been applied to various hard combinatorial optimization problems and achieved good results in both static and dynamic problems [DS03]. In this project, we study a particular variant of ACO, with an iteration-best update scheme, where pheromones are updated at each iteration using the best generated solution at that iteration.

This variant can also be seen as a particular case of an Estimation of Density Algorithms (EDA), a particular case of evolutionary algorithms where a probability density over the search space is refined iteratively. The rigorous understanding of EDAs is much less developed than that of classical evolutionary algorithms. A recent result by Doerr [Doe20] proved an $O(\mu n)$ runtime for the cGA on Jump functions, a class of multimodal functions. In this project, we explore the runtime of an ACO on the same problem.

2 Problem Description

2.1 Description of the Algorithm

The ACO algorithm we study is the 2-MMAS$_{ib}$, a specialization of the Max Min Ant System with iteration-best update, with a population size of 2. The aim of algorithm is to find the optimum of a given pseudo-boolean function $\mathcal{F}$, and it does so by updating a frequency vector $f \in [0,1]^n$ at each iteration. We will write $X = (X_1, \dots, X_n) \sim \text{Sample}(f)$ to indicate that the point X is sampled according to the distribution where each X_i is an independent Bernoulli variable of parameter f_i.
The description of the algorithm is given below:

Algorithm 1: The MMAS$_{ib}$ with population size λ to maximize a function $\mathcal{F} : \{0,1\}^n \to \mathbb{R}$.

$t \leftarrow 0$;
$f_t = (\frac{1}{2}, \dots, \frac{1}{2}) \in [0,1]^n$;
repeat
  $S \leftarrow \emptyset$;
  for $i \in [1..\lambda]$ **do**
    $x \leftarrow$Sample in $\{0,1\}^n$ using f_t
    $S \leftarrow S \cup \{x\}$
  $y \leftarrow \operatorname{argmax}_{x \in S} \mathcal{F}(x)$
  $f_{t+1} \leftarrow (1-\rho) f_t + \rho y$;
  $t \leftarrow t+1$;
until *forever*;

We study the runtime of the 2-MMAS$_{ib}$ algorithm until an optimum is sampled.

The update scheme is a convex combination between the previous frequency vector and the fittest sample. This is somewhat similar to the update scheme in the cGA algorithm, where two samples x^1 and x^2 are taken, sorted by fitness into y^1, y^2 ($\mathcal{F}(y^1) \geq \mathcal{F}(y^2)$), so that $f_{t+1} = f_t + \frac{1}{\mu}(y^1 - y^2)$, which suggests that the runtime of the 2-MMAS$_{ib}$ on Jump functions could be similar to the cGA's.

2.2 Jump functions

The simplest pseudo-Boolean benchmark function for evolutionary and stochastic optimization is the OneMax function, counting the number of ones in a string of n bits. In [NSW10], it is proved that the 2-MMAS$_{ib}$ on the OneMax function has an expected runtime of $O(\frac{\sqrt{n}}{\rho})$ for $\rho = O(\frac{1}{\sqrt{n}\log n})$.

A much harder type of functions are Jump functions, defined by

$$\textsc{Jump}_{nk}(x) = \begin{cases} \|x\|_1 + k & \text{if } \|x\|_1 \in [0..n-k] \cup \{n\} \\ n - \|x\|_1 & \text{if } \|x\|_1 \in [n-k+1..n-1] \end{cases}$$

Jump functions have a fitness landscape similar to the OneMax, except on a fitness valley or gap G_{nk}:

$$G_{nk} := \{x \in \{0,1\}^n \mid n-k < \|x\|_1 < n\}$$

which size grows with k. They are therefore much harder to optimize for classical evolutionary algorithms [JW99].

3 Optimization phases

We follow roughly the same steps used in [Doe20] to analyze the runtime of the cGA on Jump functions, where the optimization process is split into 3 steps. The first step we describe in this report starts at the beginning and last until $D_t := n - \|f_t\|_1 = O(\log n)$. During this phase, we have a strong decrease of D_t and we can prove that with high probability, we do not sample in the gap, which makes the process similar to optimizing OneMax. The next step of the project is to study the behavior of the algorithm afterwards, where we can no lounger ignore the gap, but, since D_t is small enough, we sample the optimum with high probability.

The runtime we are aiming at is most likely true with high probability but not in expectation, which means that we need to find a result with high probability for the OneMax-style first phase, which is difficult and uses other tools those used in [NSW10] to prove a runtime of the OneMax in expectation.

4 Runtime of the First Phase

Let us assume that the jump size is $k = \Theta(\log n)$. Let $\beta > 0$ be a constant, and $K = \beta \log n$. In this section, we will focus on the first phase, that begins with $f_0 = (\frac{1}{2} \ldots, \frac{1}{2})$ - which means that $D_0 = n/2$ - and lasts until $D_t \leq K$ or some frequency $f_{j,t}$ drops below 1/3. We will prove that if β is large enough and $\rho = O((\sqrt{n}\log n)^{-1})$, then the duration of this phase is $T = O(\sqrt{n}/\rho)$.

Our objective is to have $D_t \leq K$. But for our proof, we will need the frequencies to be lower bounded (by 1/3 for example). Taking the case where the frequencies are not lower bounded will make our proof much more complicated. This is why it easier to assume that this condition is true and restart our algorithm since the beginning whenever it is violated.
The theorem 2-a from [DZ20] guarantees that if $T = O(\sqrt{n}/\rho)$, then for each $j \in [1..n]$, we have

$$\Pr\left[\forall t \in [0..T] : f_{j,t} > \frac{1}{3}\right] \geq 1 - 2\exp\left(-\Omega\left(\frac{1}{\rho\sqrt{n}}\right)\right).$$

Hence

$$\Pr\left[\forall t \in [0..T], \forall j \in [1..n] : f_{j,t} > \frac{1}{3}\right] \geq 1 - 2n\exp\left(-\Omega\left(\frac{1}{\rho\sqrt{n}}\right)\right).$$

If $\rho \leq c(\sqrt{n}\log n)^{-1}$ and c is sufficiently small, then $n\exp\left(-\Omega\left(\frac{1}{\rho\sqrt{n}}\right)\right) = o(1)$, and restarting the algorithm whenever a frequency goes below 1/3 will not reduce its efficiency.

4.1 Martingale Concentration Results

Assume that we are given a sequence of real-valued supermartingale differences $(\xi_j, \mathcal{F}_j)_{j=0,\ldots,n}$ defined on some probability space $(\Omega, \mathcal{F}, \Pr)$, where $\xi_0 = 0$ and $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq \ldots \subseteq \mathcal{F}_m \subseteq \mathcal{F}$ are increasing σ -fields. So we have $\mathbf{E}\left(\xi_j \mid \mathcal{F}_{j-1}\right) \leq 0, j = 1, \ldots, m$, by definition. Set

$$S_k = \sum_{j=1}^{k} \xi_j, \quad k = 1, \ldots, m.$$

Then $S = (S_k, \mathcal{F}_k)_{k=1,\ldots,m}$ is a supermartingale. Let $[S]$ be the squared variation of the supermartingale S :

$$[S]_k = \sum_{j=1}^{k} \xi_j^2.$$

The following result corresponds to the theorem 2.1 in [FGL15].

Theorem 1 ([FGL15]). *Assume that for each $j \in [1..t]$, V_{j-1} is a non-negative and $\mathcal{F}_{j-1}$-measureable random variable. Suppose that*

$$E\left[\exp\left(\lambda\xi_j - g(\lambda)\xi_j^2\right) \mid \mathcal{F}_{j-1}\right] \leq 1 + f(\lambda)V_{j-1}$$

for some $\lambda \in (0, \infty)$, for two non-negative functions $f(\lambda)$ and $g(\lambda)$, and for all $j \in [1..t]$. Then, for all $x, v, w > 0$,

$$\begin{aligned}
&\Pr\left[S_k \geq x, [S]_k \leq v^2 \textit{ and } \sum_{j=1}^{k} V_{j-1} \leq w \textit{ for some } k \in [1..t]\right] \\
&\leq \exp\left(-\lambda x + g(\lambda)v^2 + t\log\left(1 + \frac{f(\lambda)}{t}w\right)\right) \\
&\leq \exp\left(-\lambda x + g(\lambda)v^2 + f(\lambda)w\right).
\end{aligned}$$

For our proof, we will only need a weaker version of this thorem, where g is the null function, and the random variables V_j are just constants.

Corollary 2. *Suppose that for any $t \in \mathbb{N}^*$, $E\left[\exp\left(\lambda\xi_t\right) \mid \mathcal{F}_{t-1}\right] \leq 1 + f(\lambda)$ for some $\lambda \in (0, \infty)$, for a non-negative function $f(\lambda)$. Then, for all $x > 0$,*

$$\Pr\left[S_t \geq x\right] \leq \exp\left(-\lambda x + tf(\lambda)\right).$$

Proof. Let $t \in \mathbb{N}^*$. The hypothesis of the previous theorem are verified with g the null function and $\forall j \in \mathbb{N},\ V_j = 1$.
Hence, for each $x, v, w > 0$, we have

$$\begin{aligned}
&\Pr\left[S_t \geq x, [S]_t \leq v^2 \text{ and } \sum_{j=1}^{t} V_{j-1} \leq w\right] \\
&\leq \Pr\left[S_k \geq x, [S]_k \leq v^2 \text{ and } \sum_{j=1}^{k} V_{j-1} \leq w \text{ for some } k \in [1..t]\right] \\
&\leq \exp\left(-\lambda x + wf(\lambda)\right).
\end{aligned}$$

In our case, $V_{j-1} = 1$ for all values of $j \in [1, t]$, so by choosing $w = t$, $\sum_{i=1}^{k} V_{i-1} \leq w$ is always true for $k \leq t$. By choosing $v = m \in \mathbb{N}$, we have

$$\Pr\left[S_t \geq x, [S]_t \leq m^2\right] \leq \exp\left(-\lambda x + tf(\lambda)\right).$$

Therefore,

$$\Pr\left[S_t \geq x\right] = \lim_{m \to +\infty} \Pr\left[S_t \geq x, [S]_t \leq m^2\right] \leq \exp\left(-\lambda x + tf(\lambda)\right).$$

□

Corollary 3. *If $(Y_t)_t$ is a supermartingale such that $\forall t \in \mathbb{N}^*$*

$$E\left[\exp\left(\lambda(Y_t - Y_{t-1}\right) \mid Y_0, ..., Y_{t-1}\right] \leq 1 + f(\lambda)$$

for some $\lambda > 0$, and a non-negative function $f(\lambda)$. Then, $\forall t \in \mathbb{N}^*, x > 0$,

$$\Pr\left[Y_t - Y_0 \geq x\right] \leq \exp\left(-\lambda x + f(\lambda)t\right).$$

Proof. Take $\xi_0 = 0$ and $\xi_t = Y_t - Y_{t-1}$ for $t \geq 1$ in the previous corollary □

4.2 Technical Tools

The following result can be found in [Doe20c], Theorems 1.10.9 and 1.10.31. It is taken from [Hoe63], Theorem 2 together with (2.17).

Theorem 4. *Let $Z_1, \ldots, Z_n$ be independent random variables such that for all $i \in [1..n]$, the variable Z_i takes values in some interval $[a_i, b_i]$ and has expectation $E[Z_i] = 0$. Then for all $\lambda \geq 0$, we have*

$$\Pr\left[\exists j \in [1..n] : \sum_{i=1}^{j} Z_i \geq \lambda\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right),$$
$$\Pr\left[\exists j \in [1..n] : \sum_{i=1}^{j} Z_i \leq -\lambda\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

Corollary 5. *Under the same hypothesis as in the previous theorem, we have*

$$\Pr\left[\sum_{i=1}^{n} Z_i \geq \lambda\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right),$$
$$\Pr\left[\sum_{i=1}^{n} Z_i \leq -\lambda\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

Lemma 6. *Let Z be a discrete non-negative random variable, defined on a probability space $(\Omega, \mathcal{F}, \Pr)$, and $\varphi : \mathbb{R}_+ \to \mathbb{R}$ a non decreasing function, continously differentiable such that $\varphi(0) = 0$. Then*

$$E[\varphi(Z)] = \int_0^{+\infty} \varphi'(x) \Pr[Z \geq x] dx.$$

Proof.

$$\begin{aligned}
E[\varphi(Z)] &= \sum_{z \in Z(\Omega)} \varphi(z) \Pr[Z = z] = \sum_{z \in Z(\Omega)} \int_0^z \varphi'(x) dx \Pr[Z = z] \\
&= \sum_{z \in Z(\Omega)} \int_0^{+\infty} \varphi'(x) \mathbf{1}_{z \geq x} dx \Pr[Z = z] \qquad (\varphi(0) = 0) \\
&= \int_0^{+\infty} \varphi'(x) \sum_{z \in Z(\Omega)} \left(\mathbf{1}_{z \geq x} \Pr[Z = z]\right) dx \\
&= \int_0^{+\infty} \varphi'(x) \Pr[Z \geq x] dx.
\end{aligned}$$

The permutation of the sum and the integral is correct because φ is non decreasing, and hence all the terms in the integral are positive. □

Lemma 7. *For all reals $a, b, \lambda > 0$, we have*

(i) $\int_0^{+\infty} x^2 e^{-a(x-b)^2} dx \leq \sqrt{\frac{\pi}{a}}(b^2 + \frac{1}{2a})$;

(ii) $\int_0^{+\infty} 2x e^{-a(x-b)^2} dx \leq \frac{1}{a} + 2b\sqrt{\frac{\pi}{a}}$;

(iii) $\int_0^{+\infty} (2x + \lambda x^2) e^{-a(x-b)^2} dx \leq \frac{1}{a} + \sqrt{\frac{\pi}{a}} \left(2b + \lambda b^2 + \frac{\lambda}{2a}\right)$.

Proof. Let $a, b, \lambda > 0$, and $Z \sim \mathcal{N}\left(b, \frac{1}{\sqrt{2a}}\right)$. Then

$$\begin{aligned}
\int_0^{+\infty} x^2 e^{-a(x-b)^2} dx &\leq \int_{-\infty}^{+\infty} x^2 e^{-a(x-b)^2} dx \\
&= \sqrt{\frac{\pi}{a}} E\left[Z^2\right] = \sqrt{\frac{\pi}{a}}(\mathrm{Var}[Z] + E[Z]^2) \\
&= \sqrt{\frac{\pi}{a}}(b^2 + \frac{1}{2a}); \\
\int_0^{+\infty} 2x e^{-a(x-b)^2} dx &= \frac{1}{a}\int_0^{+\infty} 2a(x-b) e^{-a(x-b)^2} dx + 2b \int_0^{+\infty} e^{-a(x-b)^2} dx \\
&\leq \frac{1}{a}\left[-e^{-a(x-b)^2}\right]_0^{+\infty} + 2b\int_{-\infty}^{+\infty} e^{-a(x-b)^2} dx \\
&= \frac{e^{-ab^2}}{a} + 2b\sqrt{\frac{\pi}{a}} \leq \frac{1}{a} + 2b\sqrt{\frac{\pi}{a}}; \\
\int_0^{+\infty} \left(2x + \lambda x^2\right) e^{-a(x-b)^2} dx &\leq \frac{1}{a} + 2b\sqrt{\frac{\pi}{a}} + \lambda\sqrt{\frac{\pi}{a}}\left(\frac{1}{2a} + b^2\right) \\
&= \frac{1}{a} + \sqrt{\frac{\pi}{a}}\left(2b + \lambda b^2 + \frac{\lambda}{2a}\right).
\end{aligned}$$

□

4.3 Proof of the Runtime of the First Phase

We start by establishing a runtime with high probability for a run of 2-MMAS$_{ib}$ the OneMax. Define $D_t := n - \|f_t\|_1$ for all $t \geq 1$, and T the first time when D_t is below $K = \beta \log n$.

Inspired from Doerr's proof of the runtime of the first phase in the [Doe20], we divide the process in this phase into several smaller phases.

For $i = 1, 2, \ldots$, let $d_i = 2^{-i}n$. Without loss of generality, we may assume that $K = 2^{-\ell-1}n$ for some $\ell \in \mathbb{N}$. Note that $\ell \leq \log_2 n$. We say that the optimization process enters Phase i (and thus leaves its current phase) when for the first time $D_t \leq d_i$. Note that we stay in Phase i even when after entering this phase D_t increases beyond d_i. Note further that, by definition, the process starts in Phase 1. We also say that the current phase ends when a frequency reaches a value below $\frac{1}{3}$.

Let $i \in [1..n]$. Let t'_i be the time when the optimization process enters phase i, and let T_i be the duration of this phase.
Let $X_0^{(i)} = d_i - D_{t'_i} \geq 0$, and for each $t \in \mathbb{N}^*$,

$$X_t^{(i)} := \begin{cases} d_i - D_{t'_i+t} & \text{if } X_{t-1}^{(i)} < d_{i+1}, \\ \\ X_{t-1}^{(i)} + \dfrac{\rho}{33}\sqrt{d_{i+1}} & \text{if } X_{t-1}^{(i)} \geq d_{i+1}. \end{cases}$$

Then we can write

$$\begin{aligned} T_i &= \min\{t \in \mathbb{N} \ : \ D_{t'_i+t} \leq d_{i+1} \mid D_{t'_i} \leq d_i\} \\ &= \min\{t \in \mathbb{N} \ : \ X_t^{(i)} \geq d_{i+1} \mid X_0^{(i)} \geq 0\}. \end{aligned}$$

Note that $X_t^{(i)} = d_i - D_{t'_i+t} = d_i - n + \|f_{t'_i+t}\|$ for $t \in [1..T_i]$, and $X_t^{(i)} = X_{T_i}^{(i)} + (t - T_i)\dfrac{\rho}{33}\sqrt{d_{i+1}}$ for $t > T_i$.
We prove below that with $\varepsilon_i = \frac{\rho}{33}\sqrt{d_{i+1}}$, $(\varepsilon_i t - X_t^{(i)})_t$ is a super-martingale, and that it verifies the the hypothesis of the Corollary 3 with a function $f(\lambda)$ we will explicit later. Then, a suitable choice of the parameters t, x, λ will give the targeted runtime.

Lemma 8. *In an iteration t, if all frequencies are above $\frac{1}{3}$, and if n is large enough, then for any $j \in [1..n]$*

$$E\left[f_{j,t+1} \mid f_t\right] \geq f_{j,t} + \frac{\rho}{11} f_{j,t}\left(1 - f_{j,t}\right)\left(\sum_{r \neq j} f_{r,t}\left(1 - f_{r,t}\right)\right)^{-1/2}.$$

Proof. To prove this, we adapt the proof in [NSW10] for Lemma 1. From there we have

$$E\left[f_{j,t+1} \mid f_t\right] \geq f_{j,t} + \rho f_{j,t}\left(1 - f_{j,t}\right)\frac{4}{9(1 + 2\sqrt{3}\sigma)},$$

With $\sigma^2 = \sum\limits_{i \neq j} f_{i,t}(1 - f_{i,t}) \geq \frac{1}{3}D_t \geq \frac{1}{3}K \geq 1$ for n large enough.

And therefore $\frac{9}{4}(1+2\sqrt{3}\sigma) \leq \frac{9}{4}(1+2\sqrt{3})\sigma \leq 11\sigma$, which gives the required result.

$\square$

Lemma 9. *Let* $\varepsilon_i := \dfrac{\rho}{33}\sqrt{d_{i+1}}$. *If n is large enough, then for any $t \in \mathbb{N}$*

$$E\left[X_{t+1}^{(i)} - X_t^{(i)} \mid X_t^{(i)}\right] \geq \varepsilon_i.$$

Proof. Let t be a positive integer. From lemma 8 we have, for all $j \in [1..n]$

$$E\left[f_{j,t+1} \mid f_t\right] \geq f_{j,t} + \frac{\rho}{11} f_{j,t}\left(1 - f_{j,t}\right)\left(\sum_{r \neq j} f_{r,t}\left(1 - f_{r,t}\right)\right)^{-1/2}.$$

But $\quad \sum\limits_{r \neq j} f_{r,t}\left(1 - f_{r,t}\right) \leqslant \sum\limits_{r \neq j}\left(1 - f_{r,t}\right) \leqslant \sum\limits_{r=1}^{n}\left(1 - f_{r,t}\right) = n - \|f_t\|_1 = D_t$.

Hence for each bit j, we have

$$\begin{aligned}
E\left[f_{j,t+1} \mid f_t\right] &\geqslant f_{j,t} + \frac{\rho}{11} f_{j,t}\left(1 - f_{j,t}\right) D_t^{-1/2} \geqslant f_{j,t} + \frac{\rho}{33}\left(1 - f_{j,t}\right) D_t^{-1/2}; \\
E\left[D_{t+1} \mid f_t\right] &= n - \sum_{j=1}^{n} E\left[f_{j,t+1} \mid f_t\right] \\
&\leq n - \sum_{j=1}^{n}\left(f_{j,t} - \frac{\rho}{33} D_t^{-1/2}\left(1 - \rho_{j,t}\right)\right) \\
&= n - \|f_{i,t}\|_1 - \frac{\rho}{33} D_t^{-1/2} \sum_{j=1}^{n}\left(1 - f_{j,t}\right) \\
&= D_t - \frac{\rho}{33} D_t^{-1/2} D_t \\
&= D_t - \frac{\rho}{33}\sqrt{D_t}.
\end{aligned}$$

We deduce that $\quad E\left[D_t - D_{t+1} \mid D_t\right] \geq \dfrac{\rho}{33}\sqrt{D_t}$.

In particular, if $X_t^{(i)} < d_{i+1}$, then $D_{t_i'+t} = d_i - X_t^{(i)} > d_i - d_{i+1} = d_{i+1}$, and

$$\begin{aligned}
E\left[X_{t+1}^{(i)} - X_t^{(i)} \mid X_t^{(i)}\right] &= E\left[D_{t_i'+t} - D_{t_i'+t+1} \mid D_t\right] \\
&\geq \frac{\rho}{33}\sqrt{D_{t_i'+t}} \geq \frac{\rho}{33}\sqrt{d_{i+1}} = \varepsilon_i.
\end{aligned}$$

And if $X_t^{(i)} \geq d_{i+1}$, then $X_{t+1}^{(i)} = X_t^{(i)} + \varepsilon_i$, and $E\left[X_{t+1}^{(i)} - X_t^{(i)} \mid X_t^{(i)}\right] = \varepsilon_i$.

$\square$

Lemma 10. *Let $Y_t^{(i)} := \varepsilon_i t - X_t^{(i)}$ for each $t \in \mathbb{N}$. If n is large enough, then $(Y_t^{(i)})_t$ is a supermartingale.*

Lemma 11. *For any $t \in \mathbb{N}, \alpha > 0$,*

$$\Pr\left[|X_{t+1}^{(i)} - X_t^{(i)}| \geq \alpha \mid X_t^{(i)}\right] \leq 4\exp\left(-\frac{2\alpha^2}{n\rho^2}\right).$$

Proof. Let $t \in \mathbb{N}$ and $\alpha > 0$. if $X_t^{(i)} < d_{i+1}$, we have

$$|X_{t+1}^{(i)} - X_t^{(i)}| = |\|f_{t'_i+t+1}\|_1 - \|f_{t'_i+t}\|_1| = \rho|\|y^{(1)}\|_1 - \|f_{t'_i+t}\|_1|.$$

The second equality is because $f_{t'_i+t+1} = (1-\rho)f_{t'_i+t} + \rho y^{(1)}$.
We remind the computation of $y^{(1)}$: we compute first two vectors $x^{(1)}, x^{(2)} \sim \text{Sample}(f_{t'_i+t})$, and $y^{(1)} := x^{(r)}$ for the $r \in \{1,2\}$ that maximizes $\textsc{Jump}_{n,k}(x^{(r)})$. Therefore

$$\begin{aligned}
\Pr\left[|X_{t+1}^{(i)} - X_t^{(i)}| \geq \alpha \mid X_t^{(i)}\right] &= \Pr\left[\left|\|y^{(1)}\|_1 - \|f_{t'+t}\|_1\right| \geq \frac{\alpha}{\rho} \mid X_t^{(i)}\right] \\
&\leq \Pr\left[\left|\|x^{(1)}\|_1 - \|f_{t'+t}\|_1\right| \geq \frac{\alpha}{\rho} \mid X_t^{(i)}\right] + \Pr\left[\left|\|x^{(2)}\|_1 - \|f_{t'+t}\|_1\right| \geq \frac{\alpha}{\rho} \mid X_t^{(i)}\right] \\
&= 2\Pr\left[\left|\|x^{(1)}\|_1 - \|f_{t'+t}\|_1\right| \geq \frac{\alpha}{\rho} \mid X_t^{(i)}\right].
\end{aligned}$$

We can write $\|x^{(1)}\|_1 - \left\|f_{t'_i+t}\right\|_1 = \sum_{j=1}^n \left(x_j^{(1)} - f_{j,t'_i+t}\right) = \sum_{j=1}^n Z_j$, with $Z_j = x_j^{(1)} - f_{j,t'_i+t}$ for each $j \in [1..n]$. We have

- $Z_1, \dots, Z_n$ are independant,
- $\forall j \in [1..n]$, $Z_j \in \{-f_{j,t'_i+t}, 1 - f_{j,t'_i+t}\}$,
- $\forall j \in [1..n]$, $E[Z_j] = 0$.

Using the Corollary 5 we have $\forall \lambda > 0$

$$\Pr\left[\sum_{j=1}^n Z_j \geq \lambda \mid X_t^{(i)}\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{r=1}^n 1}\right) = \exp\left(-\frac{2\lambda^2}{n}\right),$$
$$\Pr\left[\sum_{j=1}^n Z_j \leq -\lambda \mid X_t^{(i)}\right] \leq \exp\left(-\frac{2\lambda^2}{\sum_{r=1}^n 1}\right) = \exp\left(-\frac{2\lambda^2}{n}\right).$$

Hence

$$\begin{aligned}
\Pr\left[|X_{t+1}^{(i)} - X_t^{(i)}| \geq \alpha \mid X_t^{(i)}\right] &\leq 2\Pr\left[\sum_{j=1}^{n} |Z_j| \geq \frac{\alpha}{\rho} \mid X_t^{(i)}\right] \\
&\leq 2\Pr\left[\sum_{j=1}^{n} Z_j \geq \frac{\alpha}{\rho} \mid X_t^{(i)}\right] + 2\Pr\left[\sum_{j=1}^{n} Z_j \leq -\frac{\alpha}{\rho} \mid X_t^{(i)}\right] \\
&\leq 4\exp\left(-\frac{2\alpha^2}{\rho^2 n}\right).
\end{aligned}$$

And now, if $X_t^{(i)} \geq d_{i+1}$, then

$$\Pr\left[|X_{t+1}^{(i)} - X_t^{(i)}| \geq \alpha \mid X_t^{(i)}\right] = \Pr\left[\varepsilon_i \geq \alpha\right] = \mathbf{1}_{\{\alpha \leq \varepsilon_i\}}.$$

For $\alpha > \varepsilon_i$: $\Pr\left[|X_{t+1}^{(i)} - X_t^{(i)}| \geq \alpha \mid X_t^{(i)}\right] = 0 < 4\exp\left(-\frac{2\alpha^2}{\rho^2 n}\right)$.
And for $\alpha \leq \varepsilon_i$:

$$\begin{aligned}
4\exp\left(-\frac{2\alpha^2}{\rho^2 n}\right) \geq 4\exp\left(-\frac{2\varepsilon_i^2}{\rho^2 n}\right) &= 4\exp\left(-\frac{2}{\rho^2 n}\frac{\rho^2 d_{i+1}}{33^2}\right) = 4\exp\left(-\frac{2d_{i+1}}{33^2 n}\right) \\
&\geq 4\exp\left(-\frac{2}{33^2}\right) \simeq 3.99\ldots \\
&> \Pr\left[\left|X_{t+1}^{(i)} - X_t^{(i)}\right| \geq \alpha \mid X_t^{(i)}\right].
\end{aligned}$$

□

Lemma 12. *For any $t \in \mathbb{N}$, $\alpha > 0$,*

$$\Pr\left[\left|Y_{t+1}^{(i)} - Y_t^{(i)}\right| \geq \alpha \mid X_t^{(i)}\right] \leq 4\exp\left(-2\frac{(x-\varepsilon_i)^2}{\rho^2 n}\right).$$

Proof. Let $t \in \mathbb{N}$ and $\alpha > 0$. We have

$$\left|Y_{t+1}^{(i)} - Y_t^{(i)}\right| = \left|X_{t+1}^{(i)} - X_t^{(i)} - \varepsilon_i\right| \leq \left|X_{t+1}^{(i)} - X_t^{(i)}\right| + \varepsilon_i.$$

Hence $\left|Y_{t+1}^{(i)} - Y_t^{(i)}\right| \geq \alpha \Rightarrow \left|X_{t+1}^{(i)} - X_t^{(i)}\right| \geq \alpha - \varepsilon_i$.

If $\alpha > \varepsilon_i$, then we can apply the result of the previous lemma

$$\begin{aligned}
\Pr\left[\left|Y_{t+1}^{(i)} - Y_t^{(i)}\right| \geq \alpha \mid X_t^{(i)}\right] &\leq \Pr\left[\left|X_{t+1}^{(i)} - X_t^{(i)}\right| \geq \alpha - \varepsilon_i \mid X_t^{(i)}\right] \\
&\leq 4\exp\left(-2\frac{(\alpha-\varepsilon_i)^2}{\rho^2 n}\right).
\end{aligned}$$

If $0 < \alpha \leq \varepsilon_i$, then $|\alpha - \varepsilon_i| \leq \varepsilon_i$, and (as in the proof of the previous lemma)

$$4 \exp\left(-2\frac{(x-\varepsilon_i)^2}{\rho^2 n}\right) \geq 4 \exp\left(-\frac{2\varepsilon_i^2}{\rho^2 n}\right) > 1 \geq \Pr\left[\left|Y_{t+1}^{(i)} - Y_t^{(i)}\right| \geq \alpha \mid X_t^{(i)}\right].$$

□

Lemma 13. *If n is large enough, then for all $t \in \mathbb{N}^*, \lambda > 0$, we have*

$$E\left[\exp(\lambda(Y_t - Y_{t-1}) \mid X_{t-1}^{(i)}\right] \leq 1 + f(\lambda).$$

with

$$f(\lambda) := 2(\lambda\rho\sqrt{n})^2 \left[\frac{1}{2} + \sqrt{\frac{\pi}{2}}\left(\frac{2}{33} + \lambda\rho\sqrt{n}\left(\frac{3}{2} + \left(\frac{1}{33} + \frac{\lambda\rho\sqrt{n}}{4}\right)^2\right)\right)\right] e^{\frac{\lambda\rho\sqrt{n}}{33} + \frac{\lambda^2\rho^2 n}{8}}.$$

Proof. Let $t \in \mathbb{N}^*, \lambda > 0$, and let $\xi_t = Y_t^{(i)} - Y_{t-1}^{(i)}$. It is easy to see that $\forall x \in \mathbb{R},\ e^x \leq 1 + x + \frac{x^2}{2}e^{|x|}$. Therefore, we have

$$\begin{aligned}
&E\left[\exp(\lambda\xi_t) \mid X_{t-1}^{(i)}\right] \\
&\quad \leq 1 + \lambda E\left[\xi_t \mid X_{t-1}^{(i)}\right] + \frac{\lambda^2}{2} E\left[\xi_t^2 e^{\lambda|\xi_t|} \mid X_{t-1}^{(i)}\right]. \\
&\quad \leq 1 + \frac{\lambda^2}{2} E\left[\xi_t^2 e^{\lambda|\xi_t|} \mid X_{t-1}^{(i)}\right] \\
&\quad = 1 + \frac{\lambda^2}{2}\int_0^{+\infty} \left(2x + \lambda x^2\right) e^{\lambda x} \Pr\left[|\xi_t| > x \mid X_{t-1}^{(i)}\right] dx \\
&\quad \leq 1 + 2\lambda^2 \int_0^{+\infty} \left(2x + \lambda x^2\right) \exp\left(\lambda x - \frac{2}{\rho^2 n}(x - \varepsilon_i)^2\right) dx \\
&\quad = 1 + 2\lambda^2 e^{\lambda\varepsilon_i + \frac{\lambda^2\rho^2 n}{8}} \int_0^{+\infty} \left(2x + \lambda x^2\right) \exp\left(-\frac{2}{\rho^2 n}\left(x - \varepsilon_i - \frac{\lambda\rho^2 n}{4}\right)^2\right) dx.
\end{aligned}$$

For the second inequality we use $E\left[\xi_t \mid X_{t-1}^{(i)}\right] \leq 0$, this is because $(Y_t^{(i)})_t$ is a supermartingale (Lemma 10). After that we use the lemmas 6 then 12. And now, we will upper bound the integral that we found using (iii) from the lemma 7 with $a = \frac{2}{\rho^2 n}$ and $b = \varepsilon_i + \frac{\lambda\rho^2 n}{4}$, and using $\varepsilon_i = \frac{\rho\sqrt{d_{i+1}}}{33} \leq \frac{\rho\sqrt{n}}{33}$,

$$
\begin{aligned}
&\int_0^{+\infty} \left(2x+\lambda x^2\right)\exp\left(-\frac{2}{\rho^2 n}\left(x-\varepsilon_i-\frac{\lambda\rho^2 n}{4}\right)^2\right)dx \\
&\quad\le \frac{\rho^2 n}{2}+\sqrt{\frac{\pi}{2}}\rho\sqrt{n}\left[2\varepsilon_i+\frac{\lambda\rho^2 n}{2}+\lambda\left(\varepsilon_i+\frac{\lambda\rho^2 n}{4}\right)^2+\frac{\lambda\rho^2 n}{4}\right] \\
&\quad= \frac{\rho^2 n}{2}+\sqrt{\frac{\pi}{2}}\rho\sqrt{n}\left[2\varepsilon_i+\frac{3}{4}\lambda\rho^2 n+\lambda\left(\varepsilon+\frac{\lambda\rho^2 n}{4}\right)^2\right] \\
&\quad\le \frac{\rho^2 n}{2}+\sqrt{\frac{\pi}{2}}\rho\sqrt{n}\left[\frac{2}{33}\rho\sqrt{n}+\frac{3}{4}\lambda\rho^2 n+\lambda\left(\frac{\rho\sqrt{n}}{33}+\frac{\lambda\rho^2 n}{4}\right)^2\right] \\
&\quad= \rho^2 n\left[\frac{1}{2}+\sqrt{\frac{\pi}{2}}\left(\frac{2}{33}+\lambda\rho\sqrt{n}\left(\frac{3}{4}+\left(\frac{1}{33}+\frac{\lambda\rho\sqrt{n}}{4}\right)^2\right)\right)\right]
\end{aligned}
$$

We deduce that

$$
\begin{aligned}
&E\left[\exp(\lambda\xi_t)\mid X_{t-1}^{(i)}\right] \\
&\le 1+2(\lambda\rho\sqrt{n})^2\left[\frac{1}{2}+\sqrt{\frac{\pi}{2}}\left(\frac{2}{33}+\lambda\rho\sqrt{n}\left(\frac{3}{2}+\left(\frac{1}{33}+\frac{\lambda\rho\sqrt{n}}{4}\right)^2\right)\right)\right]e^{\frac{\lambda\rho\sqrt{n}}{33}+\frac{\lambda^2\rho^2 n}{8}} \\
&= 1+f(\lambda).
\end{aligned}
$$

□

Lemma 14. *Let $\lambda > 0$. If $\lambda\rho\sqrt{n} \le 1$, then*

$$f(\lambda) \le 6(\lambda\rho\sqrt{n})^2.$$

Proof. Let $\lambda > 0$ such that $\lambda\rho\sqrt{n} \le 1$, Therefore, we have

$$f(\lambda) \le 2(\lambda\rho\sqrt{n})^2\left[\frac{1}{2}+\sqrt{\frac{\pi}{2}}\left(\frac{2}{33}+\frac{3}{2}+\left(\frac{1}{33}+\frac{1}{4}\right)^2\right)\right]e^{\frac{1}{33}+\frac{1}{8}} \le 6(\lambda\rho\sqrt{n})^2$$

□

Lemma 15. *Let $t \in \mathbb{N}$. If n is large enough, then for any $x, \lambda > 0$, we have*

$$\Pr\left[X_t^{(i)} \le \varepsilon_i t - x\right] \le \exp\left(-\lambda x + tf(\lambda)\right).$$

Proof. Let $x, \lambda > 0$. Given the result of lemma 13, the corollary 3 gives

$$\Pr\left[Y_t^{(i)} - Y_0^{(i)} \geq x\right] \leq \exp\left(-\lambda x + tf(\lambda)\right).$$

But since $X_0^{(i)}$ is positive, we have

$$(X_t^{(i)} \leq \varepsilon_i t - x) \implies (X_t^{(i)} \leq \varepsilon_i t - x + X_0^{(i)}) \iff (Y_t^{(i)} - Y_0^{(i)} \geq x).$$

Hence

$$\Pr[X_t^{(i)} \leq \varepsilon_i t - x] \;\leq\; \Pr\left[Y_t^{(i)} - Y_0^{(i)} \geq x\right] \;\leq\; \exp\left(-\lambda x + tf(\lambda)\right).$$

□

In the following, we will denote for each $i \in [1..\ell]$:

$$t_i = \begin{cases} \left\lceil \dfrac{66\sqrt{n}}{\rho \log n} \right\rceil & \text{if } d_{i+1} \leq \dfrac{n}{(\log n)^2}, \\[2ex] \left\lceil \dfrac{66\sqrt{d_{i+1}}}{\rho} \right\rceil & \text{if } d_{i+1} > \dfrac{n}{(\log n)^2}. \end{cases}$$

Theorem 16. *If $\rho = O((\sqrt{n}\log n)^{-1})$, then for all $i \in [1..\ell]$, we have*

$$\Pr\left[X_{t_i}^{(i)} \leq d_{i+1}\right] \;\leq\; \exp\left(-\Omega(\log n)\right).$$

Proof. Let $c > 0$ be a positive constant such that $\rho \leq \frac{c}{\sqrt{n}\log n}$ for n sufficiently large. Let $i \in [1..\ell]$.

• If $d_{i+1} \leq \dfrac{n}{(\log n)^2}$, let $x = \dfrac{\varepsilon_i}{2} \times \dfrac{66\sqrt{n}}{\rho \log n} = \dfrac{\sqrt{n d_{i+1}}}{\log n}$ and $\lambda = \mu \dfrac{(\log n)^{3/2}}{\sqrt{n}}$, with $\mu = (792c)^{-1}\sqrt{\beta}$. We remind that $\forall j \in [1..\ell], d_j \geq K = \beta \log n$.

Since $d_{i+1} \leq \dfrac{n}{(\log n)^2}$, we have the following inequality

$$\varepsilon_i t_i - x \geq 2x - x = x = \frac{\sqrt{n d_{i+1}}}{\log n} \geq d_{i+1}.$$

For n sufficiently large, we have $\lambda \rho \sqrt{n} \leq c\sqrt{\dfrac{\log n}{n}} \leq 1$. Using lemma 14, we have

$$\begin{aligned} t_i f(\lambda) &\leq \left(66\frac{\sqrt{n}}{\rho\log n}+1\right)\left(6\lambda^2\rho^2 n\right) = 396\lambda^2(\rho\sqrt{n})\frac{n}{\log n}+o(1) \\ &\leq 396\lambda^2 c\frac{n}{(\log n)^2}+o(1) = 396c\mu^2\log n + o(1); \end{aligned}$$

$$\lambda x = \frac{\mu(\log n)^{3/2}}{\sqrt{n}} \times \frac{\sqrt{nd_{i+1}}}{\log n} = \mu\sqrt{d_{i+1}\log n} \geq \mu\sqrt{\beta}\log n;$$

$$\begin{aligned} \lambda x - t_i f(\lambda) &\geq \mu\sqrt{\beta}\log n - 396c\mu^2\log n + o(1) = \mu\log n\left[\sqrt{\beta}-396c\mu\right]+o(1) \\ &= \frac{\mu\sqrt{\beta}}{2}\log n + o(1) = \Omega(\log n). \end{aligned}$$

- And now, if $d_{i+1} > \frac{n}{(\log n)^2}$, let $x = \frac{\varepsilon_i}{2} \times \frac{66\sqrt{d_{i+1}}}{\rho} = d_{i+1}$ and $\lambda = \frac{1}{\sqrt{d_{i+1}}} < \frac{\log n}{\sqrt{n}}$. As in the previous case, we have

$$\varepsilon_i t_i - x \geq 2x - x = x = d_{i+1}.$$

For n sufficiently large, we have $\lambda\rho\sqrt{n} \leq \frac{c}{\sqrt{n}} \leq 1$.
Using lemma 14, we get

$$\begin{aligned} t_i f(\lambda) &= O\left(\frac{\sqrt{d_{i+1}}}{\rho}\lambda^2\rho^2 n\right) = O\left(\sqrt{nd_{i+1}}(\rho\sqrt{n})\lambda^2\right) \\ &= O\left(\frac{\sqrt{nd_{i+1}}}{\log n}\frac{1}{d_{i+1}}\right) = O\left(\frac{\sqrt{n}}{\sqrt{d_{i+1}}\log n}\right) = O(1). \end{aligned}$$

because $\rho\sqrt{n} = O(\frac{1}{\log n})$ and $d_{i+1} > \frac{n}{(\log n)^2}$. Hence

$$\lambda x - t_i f(\lambda) = \sqrt{d_{i+1}} + O(1) \geq \frac{n}{(\log n)^2} + O(1) = \Omega(\log n).$$

- We proved that in both cases, we can choose x and λ such that $\varepsilon_i t - x \geq d_{i+1}$ and $\lambda x - tf(\lambda) \geq \Omega(\log n)$. Using the lemma 15, we deduce that

$$\begin{aligned} \Pr\left[X_{t_i}^{(i)} \leq d_{i+1}\right] &\leq \Pr\left[X_{t_i}^{(i)} \leq \varepsilon_i t_i - x\right] \\ &\leq \exp\left(-\lambda x + t_i f(\lambda)\right) = \exp(-\Omega(\log n)) \end{aligned}$$

□

Theorem 17. *We consider a run of the 2-MMAS$_{ib}$ on the* OneMax *function. Let T be the first time such that $D_t \leq K = \beta log(n)$ or a frequency drops below 1/3. If $\rho = O((\sqrt{n}\log n)^{-1})$ then*

$$\Pr\left[T \leq 33(5+\sqrt{2})\frac{\sqrt{n}}{\rho}\right] \geq 1 - \exp(-\Omega(\log n)).$$

Proof. Let $i \in [1..\ell]$, we remind that

$$T_i = \min\{t \in \mathbb{N} \ : \ X_t^{(i)} \geq d_{i+1} \mid X_0^i \geq 0\}.$$

It immediately comes that $\forall t \in \mathbb{N},\ T_i > t \Rightarrow X_t^{(i)} \leq d_{i+1}$, and hence

$$\Pr[T_i > t_i] \leq \Pr\left[X_{t_i}^{(i)} \leq d_{i+1}\right] \leq \exp(-\Omega(\log n));$$
$$\Pr[T_i \leq t_i] = 1 - \Pr[T_i > t_i] \geq 1 - \exp(-\Omega(\log n)).$$

This means that for each $i \in [1..\ell]$, with high probability, we have $T_i \leq t_i$. Naturally, we will try to prove that with high probability $T \leq \sum_{i=1}^{\ell} t_i$. Let's first upper bound this sum :

$$\begin{aligned}\sum_{i=1}^{\ell} t_i &\leq \sum_{i=1}^{\ell}\left(\frac{66\sqrt{n}}{\rho\log n} + \frac{66\sqrt{d_{i+1}}}{\rho}\right) = 66\frac{\sqrt{n}}{\rho}\left[\frac{\ell}{\log n} + \sum_{i=1}^{\ell}\left(\frac{1}{\sqrt{2}}\right)^{i+1}\right] \\ &< 66\frac{\sqrt{n}}{p}\left[\frac{\log_2(n)}{\log n} + \frac{1}{2(1-1/\sqrt{2})}\right] \leq 66\frac{\sqrt{n}}{p}\left[\frac{3}{2} + \frac{2+\sqrt{2}}{2}\right] \\ &= 33(5+\sqrt{2})\frac{\sqrt{n}}{\rho}.\end{aligned}$$

Now, we have the implications

$$(\forall i \in [1,\ell], T_i \leq t_i) \Rightarrow T \leq \sum_{i=1}^{\ell} t_i \Rightarrow T \leq 33(5+\sqrt{2})\frac{\sqrt{n}}{\rho}.$$

And finally

$$\begin{aligned}\Pr\left[T \leq 33(5+\sqrt{2})\frac{\sqrt{n}}{\rho}\right] &\geq \Pr[\forall i \in [1,\ell], T_i \leq t_i] = \prod_{i=1}^{\ell}\Pr[T_i \leq t_i] \\ &\geq \prod_{i=1}^{\ell}(1 - \exp(-\Omega(\log n))) \\ &= 1 - \log n \exp(-\Omega(\log n)) \\ &= 1 - \exp(-\Omega(\log n)).\end{aligned}$$

□

Theorem 18. *We consider a run of the 2-MMAS$_{ib}$ on the a jump function $\mathcal{F}$ with jump size $k \leq 6\log n$. Let T be the first time such that $D_t \leq K = 12\log(n)$ or a frequency drops below 1/3. If $\rho = O((\sqrt{n}\log n)^{-1})$ then*

$$\Pr\left[T \leq 33(5+\sqrt{2})\frac{\sqrt{n}}{\rho}\right] \geq 1 - O(\frac{1}{n}).$$

Proof. The proof uses the same idea in the proof of Lemma 16 in [Doe20]. We consider a modified process that optimizes $\mathcal{F}$ until $D_t \leq K$ then switches to OneMax. If this modified process verifies the claim, then the original process as well.

We couple this process with a run of the algorithm on OneMax, by sampling for every iteration t, and $j \in \{1,2\}$ a vector $r^{t,j} \hookrightarrow U([0,1]^n)$, then for every sample x taking $x_i^{t,j} = 1$ if and only if $r^{t,j} \leq f_{i,t}$

The two processes are identical until we sample in the gap region. Let $\widetilde{T} = 33(5+\sqrt{2})\frac{\sqrt{n}}{\rho}$ and p the probability of not sampling in the gap region in the first $\widetilde{T}$ iterations.

Then, from Theorem 17, with probability atleast $1 - \exp(-\Omega(\log n)) - p$ we reach K or drop below 1/3 in the first $\widetilde{T}$ iterations.

And before switching the OneMax in the modified process, the probability of sampling in the gap region is :

$$\Pr\left[d(x^{t,j}) \leq k\right] \leq \Pr\left[d(x^{t,j}) \leq \frac{D_t}{2}\right] \leq \exp\left(-\frac{1}{8}D_t\right) \leq \exp\left(-\frac{1}{8}K\right) \leq n^{-1.5}.$$

Which means that $p \leq 2\widetilde{T}n^{-1.5} = O(\frac{1}{n})$ by union bound. □

References

[Doe20] Benjamin Doerr. The runtime of the compact genetic algorithm on jump functions. *Algorithmica*, 2020.

[DS03] Marco Dorigo and Thomas Stützle. *The Ant Colony Optimization Metaheuristic: Algorithms, Applications, and Advances*, pages 250–285. Springer US, Boston, MA, 2003.

[DZ20] B. Doerr and W. Zheng. Sharp bounds for genetic drift in estimation of distribution algorithms. *IEEE Transactions on Evolutionary Computation*, 24(6):1140–1149, 2020.

[FGL15] Xiequan Fan, Ion Grama, and Quansheng Liu. Exponential inequalities for martingales with applications. *Electron. J. Probab.*, 20:22 pp., 2015.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963.

[JW99] Thomas Jansen and Ingo Wegener. On the analysis of evolutionary algorithms – a proof that crossover really can help. In *European Symposium on Algorithms, ESA 1999*, pages 184–193. Springer, 1999.

[NSW10] Frank Neumann, Dirk Sudholt, and Carsten Witt. A few ants are enough: Aco with iteration-best update. In *Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation*, GECCO '10, page 63–70. Association for Computing Machinery, 2010.