# Analyzing Search Heuristics with Differential Equations

Tobias Friedrich
Hasso Plattner Institute
Potsdam, Germany

Timo Kötzing
Hasso Plattner Institute
Potsdam, Germany

Anna Melnichenko
Hasso Plattner Institute
Potsdam, Germany

## ABSTRACT

Drift Theory is currently the most common technique for the analysis of randomized search heuristics because of its broad applicability and the resulting tight first hitting time bounds. The biggest problem when applying a drift theorem is to find a suitable potential function which maps a complex space into a single number, capturing the essence of the state of the search in just one value.

We discuss another method for the analysis of randomized search heuristics based on the Theory of Differential Equations. This method considers the deterministic counterpart of the randomized process by replacing probabilistic outcomes by their expectation, and then bounding the error with good probability. We illustrate this by analyzing an Ant Colony Optimization algorithm (ACO) for the Minimum Spanning Tree problem (MST).

**ACM Reference format:**
Tobias Friedrich, Timo Kötzing, and Anna Melnichenko. 2017. Analyzing Search Heuristics with Differential Equations. In *Proceedings of GECCO '17 Companion, Berlin, Germany, July 15-19, 2017,* 2 pages.
DOI: http://dx.doi.org/10.1145/3067695.3075607

## 1 INTRODUCTION

While core theorems for drift theory date back to the 1980s [4], this technique found its first applications for the analysis of randomized search heuristics only with the introduction of the additive drift theorem in 2004 [5]. Since then, more and more variants were developed (see, for example, [2, 9]), aiming at convenient applicability. However, common to all these theorems, the complex state space of randomized search heuristics has to be captured in a single potential. Finding a good potential function for which the prerequisites of a chosen drift theorem hold is thus the key difficulty in analyzing randomized search heuristics with drift theorems.

Somewhat unnoticed in evolutionary computation, Wormald [12] proved in 1999 a theorem which took a different angle: if one replaces the probabilistic outcomes of a process by its expectation and projects its path, then, with a certain probability, the random process will closely follow this path; this path can be approximated by solving a corresponding differential equation. The interesting part of this observation (also implicit in many drift theorems) is that it generalizes to processes in many dimensions. Thus, a complex, high dimensional random process can be tracked with Wormald's theorem (see Section 2).

For use of *stochastic* differential equation in evolutionary computing, see [10, 13].



The first (and so far only) application of Wormald's theorem for the analysis of randomized search heuristics was made for a fixed-budget analysis of the (1+1) EA working on the OneMax function in 2015 [6]. In our work we want to illustrate how an application of Wormald's theorem can be used for an easy analysis of an Ant Colony Optimization (ACO) algorithm [11] on the Minimum Spanning Tree optimization problem (see Section 3 for details).

ACOs have been analyzed on MST before [8], as well as in many other areas [1, 3, 7, 8]. While most of these theoretical results ask for expected optimization times, we focus on the ACO's ability to maintain a probability distribution over good solutions. We show that for a specific set of instances the ACO will converge to a pheromone setting which corresponds to the uniform distribution over all minimum spanning trees (see Section 4).

## 2 WORMALD'S THEOREM

Consider a stochastic process $(Y^{(t)})_{t\geq 0}$, where each random variable $Y^{(t)}$ takes value in some set $S$. We use $H_t$ to denote a history of the process up to time $t$, i.e. $H_t = (Y^{(0)}, \ldots, Y^{(t)})$. Let $S^+$ denotes the set of all sequences $(Y^{(0)}, \ldots, Y^{(t)})$ such that $Y^{(t)} \in S$. Consider the following simplification of Wormald's theorem [12].

Theorem 2.1. *For some fixed $a \in \mathbb{N}$, let $(Y_i^{(t)})_{1\leq i\leq a, t\geq 0}$ be a stochastic process, such that $|Y_i^{(t)}| < m$ for all $H_t \in S^+$. Let $D$ be some bounded connected open set containing the closure of*

$$\{(0, z_1, \ldots, z_a) \mid Pr[Y_i^{(0)} = z_i m, 1 \leq i \leq a] \neq 0 \text{ for some } m\}.$$

*Assume the following three conditions hold, where for each $1 \leq i \leq a$ $f_i : \mathbb{R}_+ \times \mathbb{R}^a \to \mathbb{R}$ is continuous function.*

(i) *(Boundedness hypothesis)*
$\max_{1\leq i\leq a} |Y_i^{(t+1)} - Y_i^{(t)}| \leq 1$ *for any $t \geq 0$.*

(ii) *(Trend hypothesis)*
$E[Y_i^{(t+1)} - Y_i^{(t)} \mid H_t] = f_i(t/m, Y_1^{(t)}/m, \ldots, Y_a^{(t)}/m)$.

(iii) *(Lipschitz condition)*
$\exists L > 0\ \forall u = (u_1, \ldots, u_{a+1}), v = (v_1, \ldots, v_{a+1}) \in D$, *such that* $|f(u) - f(v)| \leq L \max_{1\leq i\leq k} |u_i - v_i|$.

*Then the following are true.*

(a) *For any $(0, \hat{z}_1, \ldots, \hat{z}_a) \in D$ the system of differential equations*

$$\frac{dz_i}{dx} = f_i(x, z_1, \ldots, z_a), \quad i = 1, \ldots, a$$

*has a unique solution in $D$ for $z_i : \mathbb{R} \to \mathbb{R}$ passing through $z_i(0) = \hat{z}_i, 1 \leq i \leq a$, and which extends to points arbitrary close to the boundary of $D$.*

(b) *Let $\lambda = \lambda(m) = o_m(1)$. For some constant $C > 0$, with probability $1 - O_m(\frac{1}{\lambda}\exp(-m\lambda^3))$,*

$$Y_i^{(t)} = mz_i(t/m) + O_m(\lambda m)$$

*uniformly for $0 \leq t \leq \sigma m$ and for each $i$, where $z_i(x)$ is the solution in (a) with $\hat{z}_i = \frac{1}{m} Y_i^{(0)}$, and $\sigma = \sigma(m)$ is the supremum of those $x$ to which the solution can be extended before reaching within $\ell^\infty$-distance $C\lambda$ of the boundary of $D$.*

## 3 ACO AND MST

We consider the case of the ACO algorithm running on a graph $G$ with $n$ nodes and $m$ edges to find a minimum spanning tree. Let $(\tau_i^{(t)})_{1 \leq i \leq m, t \geq 0}$ be the sequence of random variables describing the amount of pheromone on edge $i$ after $t$ iterations.

We use the simple ACO algorithm called 1-ANT described in [8]. The algorithm produces, in each iteration, a solution constructed by a Prim-based procedure on a so-called edge-weighted construction graph $G' = \langle V', E' \rangle$, $|V'| = |E| + 1$, where $G'$ is a directed graph with pheromone values $\tau : E \to [\rho, 1-\rho]$ on the edges, initial pheromone values $\tau_0 = 0.5$ and $\rho \in (0, 0.5)$ is an evaporation coefficient. The ant starts from the initial vertex of the graph $G'$ and moves between non-visited vertexes, choosing edges for the solution with probability proportional to its pheromone value. At the end of each iteration, the ACO updates pheromone values as follows, depending on whether an edge is in the ant's path $p_t$ of the solution:

$$\tau_i^{(t+1)} = \begin{cases} \tau_i^{(t)} \cdot (1-\rho) + \rho, & \text{if } e_i \in p_t; \\ \tau_i^{(t)} \cdot (1-\rho), & \text{if } e_i \in p_t. \end{cases}$$

## 4 ANALYSIS: DIFFERENTIAL EQUATIONS

We consider $C_n$, the cycle of size $n$ with all weights equal 1, as input graph of our problem. Standard methods give us an expected pheromone difference in one step:

$$E[\tau_i^{(t+1)} - \tau_i^{(t)} \mid H_t] = -\tau_i^{(t)} \rho + Pr[e_i \in p_t]\, \rho.$$

Let $\tau_0 \in [\rho, 1-\rho]$ be the initial pheromone on an edge. Assume that $n$ is fixed and model the pheromones as continuous function $z(t)$, thus getting the following system of differential equations.

$$\begin{aligned} z_i'(t) &= -z_i(t) + 1 - \sum_{i=1}^{n} \sum_{p_t \in P_i} \frac{\prod_{j \neq i} z_j(t)}{\prod_{k=1}^{n-1} \left( z_i(t) + \sum_{j=k}^{n-1} z_{i_j}(t) \right)}, \\ z_i(0) &= \tau_0, \qquad\qquad (1) \\ z_i(t) &\in D,\ i \in \{1, \ldots, n\}, \end{aligned}$$

where $P_i$ is the set of all permutations of the $n-1$ elements $e_{i_1}, \ldots, e_{i_{n-1}}$, $e_{i_j} \in E(G_n), i_j \neq i$, i.e. $P_i$ is a set of possible ant's path combinations, such that the path does not contain edge $e_i$; and $D = [0, 2] \times (0, 1)^n$.

Now, by Theorem 2.1, the track of pheromone values concentrates near the solution of system (1).

Theorem 4.1. *With probability $1 - O\left(\rho^{-1/4} \exp\left(-\rho^{-1/4}\right)\right)$,*

$$\tau_k^{(t)} = \left(\tau_0 - \frac{n-1}{n}\right) e^{-\rho t} + \frac{n-1}{n} + O(\rho^{1/4})$$

*uniformly for all $k \in \{1, \ldots, n\}$ and all $0 \leq t \leq \rho^{-1}$.*

The theorem was given for all pheromone values starting with the same value. Figure 1 shows that pheromone values come close to the solution tracks and goes to $\frac{n-1}{n}$ together even if initial values are very sparsely distributed. The figure also shows a very interesting non-monotone behavior of a pheromone values during it's approach of the limiting value of $2/3$.

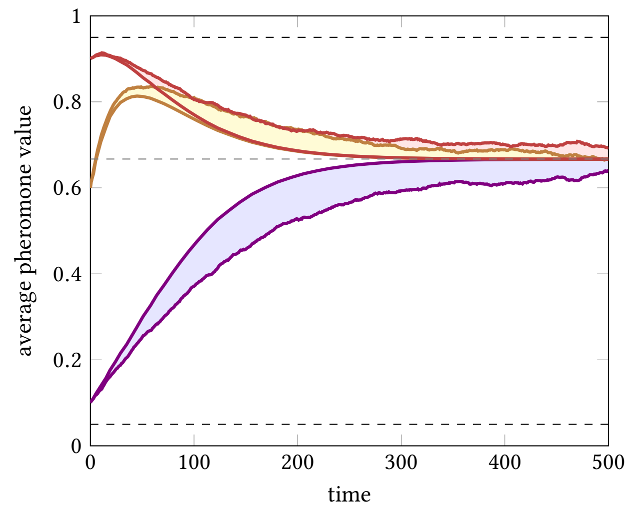


**Figure 1: Solution of the differential equations system (1) (smooth tracks) and ACO solution (ragged tracks) for the MST problem for the cycle of size 3. The filled areas show a difference between differential equation and algorithm solutions. The dashed lines are the ACO thresholds $\rho$ and $1-\rho$ and the limiting value $2/3$.**

## REFERENCES


[1] B. Doerr and D. Johannsen. Refined runtime analysis of a basic ant colony optimization algorithm. In *Proc. of IEEE CEC'07*, pages 501–507. IEEE Press, 2007.
[2] B. Doerr, D. Johannsen, and C. Winzen. Multiplicative drift analysis. *Algorithmica*, 64(4):673–697, 2012.
[3] B. Doerr, F. Neumann, D. Sudholt, and C. Witt. On the runtime analysis of the 1-ANT ACO algorithm. In *Proc. of GECCO'07*, pages 33–40. ACM, 2007.
[4] B. Hajek. Hitting-time and occupation-time bounds implied by drift analysis with applications. *Advances in Applied probability*, 14(03):502–525, 1982.
[5] J. He and X. Yao. A study of drift analysis for estimating computation time of evolutionary algorithms. *Natural Computing*, 3(1):21–35, 2004.
[6] J. Lengler and N. Spooner. Fixed budget performance of the (1+ 1) EA on linear functions. In *Proc. of FOGA'15*, pages 52–61. ACM, 2015.
[7] F. Neumann and C. Witt. Runtime analysis of a simple ant colony optimization algorithm. In *Proc. of ISAAC'06*, pages 618–627. Springer, 2006.
[8] F. Neumann and C. Witt. Ant colony optimization and the minimum spanning tree problem. *Theoretical Computer Science*, 411(25):2406–2413, 2010.
[9] P. S. Oliveto and C. Witt. Simplified drift analysis for proving lower bounds in evolutionary computation. *Algorithmica*, 59(3):369–386, 2011.
[10] T. Paixao and J. P. Heredia. An application of stochastic differential equations to evolutionary algorithms. In *Proc. of FOGA'17*, pages 3–12, 2017.
[11] T. Stützle and H. H. Hoos. MAX–MIN ant system. *Future generation computer systems*, 16(8):889–914, 2000.
[12] N. C. Wormald. The differential equation method for random graph processes and greedy algorithms. *Lectures on approximation and randomized algorithms*, pages 73–155, 1999.
[13] G. M. Yin, G. Rudolph, and H. Schwefel. Analyzing (1, $\lambda$) evolution strategy via stochastic approximation methods. *Evolutionary Computation*, 3(4):473–489, 1995.