Complexity of interior-point algorithms: tropical compuations

Mohammed Amine Bennouna, Yassine El Maazouz

December 2018

Our deep gratitude goes to prof. ***Xavier Allamigeon*** *& prof.* ***Stéphane Gaubert*** *for introducing us to the wonderful world of tropical geomerty and for all their time and dedication to help us complete this work.*

Abstract

Interior-point methods are a very import class of algorithms that solve linear and semi-definite optimization problems. Very successful in practice, they are widely believed to be strongly polynomial. In this report, we follow on a recently published paper proving that log-barrier interior point methods are not strongly polynomial by presenting a counter-example. We analyse the behaviour of a predictor-corrector algorithm on the counter-example instances to confirm the theoretical results. In particular, we present an explicit method of constructing an initial point for the algorithm over the field of generalized Puiseux series based on tropical characteristics of the central path.

1 Introduction and context

Let n , $m \in \mathbb{N}$ be two positive integers, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ two vectors and $A \in \mathbb{R}^{m \times n}$ a real $m \times n$ matrix. We are interested in solving linear programs which are optimization problems of the following form:

$$\text{minimize} \quad \langle c, x \rangle \quad \text{subject to } Ax \leq b \quad \text{and} \quad x \geq 0$$

where c is the cost vector and $x \geq 0$ mean that $x_i \geq 0$ for $1 \leq i \leq n$ (x is in the positive orthant).

By adding slack variables $w \in \mathbb{R}^m_{\geq 0}$ to x the linear program (P) can be transformed to the following form:

$$(P) \qquad \inf_{x \in \mathbb{R}^n, w \in \mathbb{R}^m} \langle c, x \rangle \quad \text{subject to } Ax + w = b \quad \text{and} \quad (x, w) \geq 0$$

We denote by $\mathcal{F}_P$ feasible set of the problem (P) is the set of vectors x, w that satisfy the constraints, formally:

$$\mathcal{F}_p = \{(x, w) \in \mathbb{R}^n \times \mathbb{R}^m : Ax + w = b, x \geq 0, w \geq 0\}$$

We call the strictly feasible set $\mathcal{F}_P^s$ the subset of $\mathcal{F}_P$ that corresponds to strict positivity constraints over x and w:

$$\mathcal{F}_p^s = \{(x, w) \in \mathbb{R}^n \times \mathbb{R}^m : Ax + w = b, x > 0, w > 0\}$$

.

The dual program of (P) is denoted (D) and is written as follows:

$$(D) \qquad \sup_{y\in\mathbb{R}^m, s\in\mathbb{R}^n} \langle b,y\rangle \quad \text{subject to } A^T y + s = c \quad \text{and} \quad (y,s)\geq 0$$

This dual program also gives rise to the corresponding feasible and strictly feasible sets :

$$\mathcal{F}_D = \{(y,s)\in\mathbb{R}^m\times\mathbb{R}^n : A^T y + s = c,\ y\geq 0,\ s\geq 0\}$$

$$\mathcal{F}_D^s = \{(y,s)\in\mathbb{R}^m\times\mathbb{R}^n : A^T y + s = c,\ y> 0\ s> 0\}$$

We define the primal-dual feasible set $\mathcal{F}$ and strictly feasible set $\mathcal{F}^s$ as follows:

$$\mathcal{F} = \mathcal{F}_P\times\mathcal{F}_D \text{ and } \mathcal{F}^s = \mathcal{F}_P^s\times\mathcal{F}_D^s$$

A familiar result in linear programming (see [2] for more details) states that $(x,w)\in\mathbb{R}^n\times\mathbb{R}^m$ is solution to (P) if and only if there exist a couple $(y,s)\in\mathbb{R}^m\times\mathbb{R}^n$ such that :

$$\begin{cases} A^T y + s = c,\ s\geq 0,\ y\geq 0 \\ Ax + w = b,\ x\geq 0\ w\geq 0 \\ \langle x,s\rangle = 0 \end{cases} \tag{1}$$

The same result holds for the dual: (y,s) is an optimal solution of (D) if and only if there exist $(x,w)\in\mathbb{R}^n\times R^m$ such that (1) holds. In this case we denote by $z=(x,w,y,s)$ the primal-dual solution of (1).

1.1 Logarithmic barrier and interior point algorithms

Positivity constraints are the main difficulty of the Primal-Dual linear program summarized in (1). To get rid of these conditions we penalize the points on the borders of the positive orthant by adding a barrier to the cost function that pushes the optimal solutions away from the boundaries of the feasible set $\mathcal{F}$ into the set $\mathcal{F}_s$. Formally we define a family of programs parametrized by $\mu>0$ as follows:

$$(P_\mu)\begin{cases} \inf\ \langle c,x\rangle + \mu.\text{lb}(x) \\ Ax+w=b, x\geq 0, w\geq 0 \end{cases} \qquad \text{where: } \text{lb}(x) = \begin{cases} -\sum_{i=1}^n \log(x_i),\ \text{if } x>0 \\ +\infty,\ \text{else} \end{cases}$$

The *log-barrier* function lb, *pushes* the solution into the positive orthan by *penalizing* the objective function when we get close to the borders, as $\text{lb}(x)\to+\infty$ when $x_i\to 0$.

The dual of this problem is written as:

$$(D_\mu)\begin{cases} \sup\ \langle b,y\rangle - \mu\text{lb}(s) \\ A^T y + s = c, y\geq 0, s\geq 0 \end{cases}$$

The primal-dual problem can be summarized (as we did above) in the following problem:

$$\begin{cases} A^T y + s = c,\ s> 0 \\ Ax + w = b,\ x>0, w>0 \\ x_i s_i = \mu, \forall 1\leq i\leq n \end{cases} \tag{2}$$

Solving $(P_\mu) - (D_\mu)$ for all $\mu > 0$ gives an application:

$$C : \mu \in]0, +\infty[\to z(\mu) := (x(\mu), w(\mu), s(\mu), y(\mu)) \in \mathcal{F}^s$$

We recall that $\mathcal{F}^s := \mathcal{F}^s_P \times \mathcal{F}^s_D$ is the primal-dual strictly feasible set.

Notice that (P_μ) and (D_μ) becomes (P) and (D) when $\mu \to 0$ and that $(C(\mu))_\mu$ is an algebraic curve. We show, under reasonable hypothesis, that we have as we may expect $z(\mu) \underset{\mu \to 0}{\longrightarrow} z^* = (x^*, w^*, y^*, s^*)$ where z^* is optimal primal-dual solution of the original problem. Therefore, the algebraic curve $(C(\mu))_{\mu \geq 0}$ (defined by the polynomial equations in (4)) is is path converging to the optimal solution of the problem as μ tends to 0 which is called the ***central path***. Log-barrier interior point algorithms consist of following this path to the optimal solution.

We now present as special type of interior point algorithm that we will will be using throughout in all that follows.

1.2 Predictor-Corrector algorithm

In this section we present the predictor corrector algorithm. We should stress that we consider the following linear program with equality constraints:

$$\inf_{x \in \mathbb{R}^n} \quad \langle c, x \rangle \quad \text{subject to } Ax = b \quad \text{and} \quad x \geq 0$$

whose dual problem is of the form:

$$\sup_{y \in \mathbb{R}^m, s \in \mathbb{R}^n} \quad \langle b, y \rangle \quad \text{subject to } A^T y + s = c \quad \text{and} \quad s \geq 0$$

We can transform a linear problem with inequality constraints to a problem of the above form by adding slack variables. We now describe the predictor corrector algorithm over a problem with equality constraints

Predictor-Corrector algorithm is an example of interior point methods. It is based on a opportunisticway of following the central path towards the optimal solution. On every step of the algorithm, we start by following a Newton direction to the optimal solution by making sure that we don't deviate too much from the central path (prediction) and then correct the path by getting closer to the central path (correction) which yields a piecewise approximation of the central path.

Formally, let's start by defining the *duality measure* that quantify how close we are from the optimal solution. For a given $z \in \mathcal{F}^s$, we define the duality measure as:

$$\bar{\mu}(z) = \frac{\langle x, s \rangle}{n}$$

Notice that $z \in \mathcal{F}^s$ is in the central path, if and only if $x * s = \bar{\mu}(z)e$, where $*$ is the Hadamard product of vectors ($(x * s)_i = x_i s_i$), and e is the vector in $\mathbb{R}^n$ whose entries are all 1.

In order to prevent the prediction step so as not to go so far from the central path, we need to define a neighbourhood around the central path. We will be using two neighbourhoods: one to limit the prediction step, and one to define a tolerance for the correction step. The neighbourhoods are defined by a parameter θ and can be of two types:

The norm-2 neighbourhood:

$$V_2(\theta) := \{z \in \mathcal{F}^s : \|x * s - \bar{\mu}(z)e\|_2 \leq \theta\bar{\mu}(z)\}$$

The norm-∞ neighbourhood:

$$V_\infty^-(\theta) := \{z \in \mathcal{F}^s : \|x * s - \bar{\mu}(z)e\|_\infty \leq \theta\bar{\mu}(z)\}$$

where $\|v\|_\infty = \max(0, \max_i(-v_i))$.

The norm-∞ neighbourhood can also be rewritten as:

$$V_\infty^-(\theta) := \{z \in \mathcal{F}^s : x * s \geq (1-\theta)\bar{\mu}(z)e\}$$

The following algorithm works for both types of neighbourhood and we refer to the neighbourhood as $V(\theta)$.

Algorithm 1 Predictor-corrector algorithm

We consider two sets $V(\theta)$ and $V(\theta')$ with $0 < \theta' < \theta$.
Suppose that the actual point is $z := (x, w, s, y) \in V(\theta') \subset \mathcal{F}^s$. The following point $z_+ \in V(\theta')$ is obtained by two steps.

1. **Prediction**
   (a) Compute Newton direction d from z toward the optimal solution such that we stay in the feasible set $\mathcal{F}^s$.
   (b) Compute the largest jump $\alpha \in [0, 1]$ such that $z' = z + \alpha d \in V(\theta)$.
   (c) Prediction point: z'.
2. **Correction**
   (a) Compute Newton direction d' from z' toward the central path such that we stay in the feasible set $\mathcal{F}^s$ and get closer to the narrow neighbourhood $V(\theta)$.
   (b) Correction point: $z_+ = z' + d'$.

In this part we explain how the Newton directions are computed.

1.a We compute $d = \begin{pmatrix} dx \\ dy \\ ds \end{pmatrix}$ as a solution of [2]:

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} dx \\ dy \\ ds \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -Xs \end{pmatrix}$$

Step 1.b

When $V = V_2$, we compute α by solving: $\|x(\alpha) * s(\alpha) - \bar{\mu}(z(\alpha))e\|_2^2 = \theta'^2\bar{\mu}(z(\alpha))^2$, where $z(\alpha) = z + \alpha d$.

When $V = V_\infty^-$, we need to compute:

$$\begin{aligned}\alpha^* &= \max\{\alpha \in [0,1] \; | X(\alpha)s(\alpha) \geq (1-\theta)\bar{\mu}(z(\alpha))e\} \\ &= \max\{\alpha \in [0,1] \; | \min_i x_i(\alpha)s_i(\alpha) = (1-\theta')\bar{\mu}(z(\alpha))\}\end{aligned} \tag{3}$$

In order to do that, we compute $\alpha^{**} = \min_i \min\{\alpha_i \in [0,1] \mid x_i(\alpha_i)s_i(\alpha_i) = (1-\theta')\bar{\mu}(z(\alpha_i))\}$ which is easier to compute (solving n second degree polynomial equations). Given lemma 2 in [1], we have $z + \alpha d \in V$ for all $\alpha \in [0, \alpha^*]$ and therefore $\alpha^* = \alpha^{**}$.

Step 2.b We compute $d' = \begin{pmatrix} d'x \\ d'y \\ d's \end{pmatrix}$ as a solution of [2]:

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} \begin{pmatrix} d'x \\ d'y \\ d's \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \bar{\mu}(z)e - Xs \end{pmatrix}$$

We use the values $\theta' = \frac{1}{2}$ and $\theta = \frac{1}{4}$ which are values for which the algorithm in known to converge (these values are not unique). For more details on the predictor corrector algorithm, we refer the reader to of the book [2, chapter 18] by Jean-Charles Gilbert.

2 Log-barrier interior point methods are not strongly polynomial

In their article, Stéphane Gaubert, Xavier Allamigeon and their co-authors prove that log-barrier Interior-Point methods are not strongly polynomial by presenting a counter example instance inspired by a Game-Theory problem.

The counter example is the follwing:

Given an integer number $r > 1$ and an Archimedean parameter t we consider the following Linear Program:

$$\begin{array}{rll} \text{Minimize} & x_1 & \\ \text{Subject to} & x_1 \leq t^2 & \\ & x_2 \leq t & \\ & x_{2j+1} \leq t\, x_{2j-1}\,, \; x_{2j+1} \leq t\, x_{2j}, & 1 \leq j < r \\ & x_{2j+2} \leq t^{1-1/2^j}(x_{2j-1} + x_{2j}), & 1 \leq j < r \\ & x_{2r-1} \geq 0\,, \; x_{2r} \geq 0 & \end{array} \tag{$\mathrm{LW}_r(t)$}$$

$\mathbf{LW}_r(t)$ has $2r$ variables and $3r+1$ constraints and can be transformed to fit the form of (P) by adding slack variables $(w_j)_{1 \leq j \leq r}$. When transformed, $\mathbf{LW}_r(t)$ has $5r+1$ variables and $3r+1$ constraints.

$\mathbf{LW}_r(t)$ can be rewritten in a matricial form as follows:

$$\begin{array}{rl} \text{Minimize} & c.x \\ \text{Subject to} & A(t)x \leq b(t) \\ & x \geq 0 \end{array}$$

2.1 Field of Puiseux series and Tropical Geometry tools

We denote by $\mathbb{K}$ the field of real and absolutely convergent Puiseux series, ie: the field of elements $\boldsymbol{f} = \sum_{\alpha\in\mathbb{R}} a_\alpha t^\alpha$ such that:

$$\begin{cases} (i)\ a_\alpha \in \mathbb{R} \text{ for all } \alpha \in \mathbb{R}, \\ (ii)\ supp(\boldsymbol{f}) = \{\alpha \in \mathbb{R} : a_\alpha \neq 0\} \text{ is either finite or has } -\infty \text{as the only accumulation point} \end{cases} \tag{4}$$

The definition of the field $\mathbb{K}$ guarantees that every series $\boldsymbol{f} = \sum_{\alpha\in\mathbb{R}} a_\alpha t^\alpha$ has a leading term $a_{\alpha_0} t^{\alpha_0}$ where alpha is the defined as $\alpha_0 := max(supp(\boldsymbol{f}))$. This allows us to define a total ordering on $\mathbb{K}$ where an element is positive if its leading term has a positive coefficient α.

The map $val : \mathbb{K} \to \mathbb{R} \cup \{-\infty\}$ that allows us to establish a link between the problem $\mathbf{LW}_r(t)$ and Tropical Geometry is defined by: $val(\boldsymbol{f})$ for $\boldsymbol{f} \in \mathbb{K}$ is the greatest element α_0 of $supp(\boldsymbol{f})$ or equivalently:

$$val(\boldsymbol{f}) = \lim_{t\to+\infty} \log_t |\boldsymbol{f}(t)|$$

This is a non trivial *valuation* over the field of generalized Puiseux series that has the following basic properties:

$\forall \boldsymbol{f}, \boldsymbol{g} \in \mathbb{K}$:

$$val(\boldsymbol{f} + \boldsymbol{g}) \leq \max(val(\boldsymbol{f}), val(\boldsymbol{g})) \quad \text{and} \quad val(\boldsymbol{f}\boldsymbol{g}) = val(\boldsymbol{f}) + val(\boldsymbol{g})\,. \tag{5}$$

where the first inequality becomes an equality if the leading terms of $\boldsymbol{f}$ and $\boldsymbol{g}$ don't cancel which is the case when $\boldsymbol{f}$, $\boldsymbol{g}$ are positive Puiseux series (i.e in $\mathbb{K}_+$). $val : (\mathbb{K}_+, +, \times) \to (\mathbb{T}, \oplus, \odot)$ where $\mathbb{T} := \mathbb{R} \cup \{-\infty\}$ is then some kind of morphisme.

We can then define Puiseux polyhedron as we do in a real numbers setting ([1] for a precise justification)

$$\boldsymbol{\mathcal{P}} = \{\boldsymbol{x} \in \mathbb{K}^d \colon \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}\,, \tag{6}$$

where $\boldsymbol{A} \in \mathbb{K}^{p\times d}$, $\boldsymbol{b} \in \mathbb{K}^p$ (with $p \geq 0$), and $\leq$ stands for the partial order over $\mathbb{K}^p$.

As explained in [1], this allows us to tropicalize the problem $\mathbf{LW}_r(t)$ and to compute it's exact tropical central path $\mathcal{C}^{trop}(\lambda)$ which in some sense describes the behaviour of the real central path when t is sufficiently large. The combinatorial analysis of the tropcial central path gives us an yields a lower bound on the curvature of the central path of the parametric family of linear problems over the reals $(\mathbf{LW}_r(t))_{t>0}$

The tropical central path is given by the equations:

for x :

$$\begin{aligned} x_1^\lambda &= \min(\lambda, 2) \\ x_2^\lambda &= 1 \\ x_{2j+1}^\lambda &= 1 + \min(x_{2j-1}^\lambda, x_{2j}^\lambda), \quad 1 \leq j < r \\ x_{2j+2}^\lambda &= (1 - 1/2^j) + \max(x_{2j-1}^\lambda, x_{2j}^\lambda), \quad 1 \leq j < r \end{aligned}$$

for the slack variables w :

$$\begin{array}{ll} & w_1^\lambda = 2,\ w_2^\lambda = 1,\ w_3^\lambda = 1 + \min(\lambda, 2) \\ \forall 1 < j < r & w_{3j}^\lambda = 1 + \min(w_{3j-3}^\lambda, w_{3j-2}^\lambda) \\ \forall 1 \leq j < r & w_{3j+1}^\lambda = 1 + w_{3j-1}^\lambda \\ & w_{3j+2}^\lambda = (1 - \frac{1}{2^j}) + \max(w_{3j} - 1, w_{3j-1}) \\ & w_{3r}^\lambda = \min(w_{3r-3}^\lambda, w_{3r-2}^\lambda) \\ & w_{3r+1}^\lambda = \max(w_{3r-3}^\lambda, w_{3r-2}^\lambda) - \frac{1}{2^{r-1}} \end{array}$$

For $z = (x, w, s, y)$ in the central path, we have $\begin{pmatrix} xs \\ wy \end{pmatrix} = \mu e$, and since the valuation of μ is λ, we deduce y^λ by tropicalizing this equality which yields:

$$\begin{array}{ll} & y_1^\lambda = \lambda - 2,\ y_2^\lambda = \lambda - 1,\ y_3^\lambda = \lambda - 1 - \min(\lambda, 2) \quad (7a) \\ \forall 1 < j < r & y_{3j}^\lambda = \max(y_{3j-3}^\lambda, y_{3j-2}^\lambda) - 1 \quad (7b) \\ \forall 1 \leq j < r & y_{3j+1}^\lambda = y_{3j-1}^\lambda - 1 \quad (7c) \\ & y_{3j+2}^\lambda = \min(y_{3j}^\lambda + 1, y_{3j-1}^\lambda) - (1 - \frac{1}{2^j}) \quad (7d) \\ & y_{3r}^\lambda = \max(y_{3r-3}^\lambda, y_{3r-2}^\lambda) \quad (7e) \\ & y_{3r+1}^\lambda = \min(y_{3r-3}^\lambda, y_{3r-2}^\lambda) + \frac{1}{2^{r-1}} \quad (7f) \end{array}$$

When plotting in a plane the points $(x_{2r-1}^\lambda, x_{2r}^\lambda)$ for different values of the tropical parameter λ we expect to get a staircase form with an exponential number of steps.

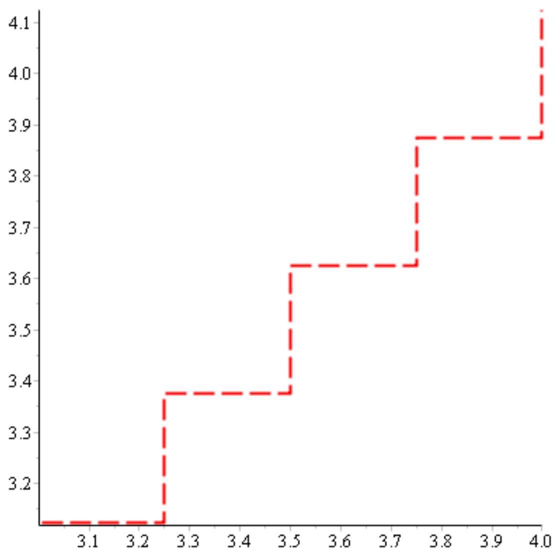

Figure 1: The form of the two last tropical x-coordinates

Our objective is to reproduce experimentally the theoretical results of [1] by running a Predictor-Corrector algorithm on instances of $\mathbf{LW}_r(t)$ and plotting an approximation of the tropical path given by the *valuation* ($log_t(.)$ when t is large) of the prediction/correction points following the primal-dual central path.

3 Constructing an initial point

To run the predictor corrector algorithm on our counter-example, we need to find an initial point z_0 to start the algorithm that verifies the following conditions:

1. $z_0 \in V(\theta')$. In particular, we need to have $z_0 \in \mathcal{F}^s$.
2. $\bar{\mu}(z_0)$ large enough. We need to be sufficiently far in the central path to be able to see all the steps.

Condition 1 is the most challenging to satisfy. A natural way to proceed is to start from a point from the tropical path – which we know exactly its form for all λ – and try to lift it into a real point z such that $\bar{\mu}(z) \approx t^\lambda$ and z close to the central path. In order to do that, we need to go through Puiseux series:

$$\begin{array}{lll} & \text{Tropical} & : z^\lambda = \mathcal{C}^{trop}(\lambda) \\ \longrightarrow & \text{Puiseux series} & : \boldsymbol{z} \mid \boldsymbol{z} \in \boldsymbol{\mathcal{F}^s},\ val(\boldsymbol{z}) = z^\lambda \\ \longrightarrow & \text{Real point} & : z = \boldsymbol{z}(t) \mid z \in V(\theta),\ \bar{\mu}(z) \approx t^\lambda \end{array}$$

for t sufficiently large.

In all what follows, we consider $\lambda \in]0,2]$ and $z^\lambda = \mathcal{C}^{trop}(\lambda)$. We will start therefore by analyzing the set $\{\boldsymbol{z} \mid \boldsymbol{z} \in \boldsymbol{\mathcal{F}^s},\ val(\boldsymbol{z}) = z^\lambda\}$.

In order to satisfy $\boldsymbol{z} \in \boldsymbol{\mathcal{F}^s}$, we will exploit the form of the conditions. We recall that $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{s}, \boldsymbol{y})$ and

$$\boldsymbol{z} \in \boldsymbol{\mathcal{F}^s} \iff \begin{cases} \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w} = \boldsymbol{b} \\ \boldsymbol{s} - \boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{c} \\ \boldsymbol{x}, \boldsymbol{w}, \boldsymbol{s}, \boldsymbol{y} > 0 \end{cases}$$

Given the form of the conditions, we construct $\boldsymbol{x}$ and $\boldsymbol{y}$ from the tropical point and define $\boldsymbol{s} := \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{c}$ and $\boldsymbol{w} := \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$ to guarantee the two first feasibility conditions.

In particular we will seek to construct such points by *lifting* the tropical points.

Definition 1. *Let z^λ be a tropical point in the tropical path. We say that a Puisieux serie point $\boldsymbol{z}$ is lifted from z^λ if there exist α and β such that*

$$\forall i,\ \boldsymbol{x}_i = \alpha_i t^{x_i^\lambda}$$
$$\forall i,\ \boldsymbol{y}_i = \beta_i t^{y_i^\lambda}$$

and $\boldsymbol{s} := \boldsymbol{A}^T\boldsymbol{y} + \boldsymbol{c}$, $\boldsymbol{w} := \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$.

Note that a lifted point $\boldsymbol{z}$ verify in particular $val(\boldsymbol{z}) = z^\lambda$, but not necessarily that $\boldsymbol{z} \in \boldsymbol{\mathcal{F}^s}$. The following result ensures that focusing only on lifted points does not reduce our chances of getting into the feasible set.

For a given vector of Puiseux series $\boldsymbol{x}$, we denote by $\bar{\boldsymbol{x}}$ the Puiseux series vector obtained by taking the leading monomial on each coordinate.

Proposition 2. *Let $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{y}, \boldsymbol{s})$ be a feasible Puiseux series.*
$\tilde{\boldsymbol{z}} := (\bar{\boldsymbol{x}},\ \tilde{\boldsymbol{w}} := \boldsymbol{b} - \boldsymbol{A}\bar{\boldsymbol{x}},\ \bar{\boldsymbol{y}},\ \tilde{\boldsymbol{s}} := \boldsymbol{c} + \boldsymbol{A}\bar{\boldsymbol{y}})$ is also a feasible point in $\mathbb{K}$ (i.e $\tilde{\boldsymbol{z}} \in \boldsymbol{\mathcal{F}^s}$)

Proof. We have by definition of $\boldsymbol{w}$ and $\bar{\boldsymbol{w}}$:

$$
\begin{aligned}
\boldsymbol{w}_1 &= t^2 - \boldsymbol{x}_1 \\
\boldsymbol{w}_2 &= t - \boldsymbol{x}_2 \\
\boldsymbol{w}_{3j} &= t\,\boldsymbol{x}_{2j-1} - \boldsymbol{x}_{2j+1} \\
\boldsymbol{w}_{3j+1} &= t\,\boldsymbol{x}_{2j} - \boldsymbol{x}_{2j+1} \\
\boldsymbol{w}_{3j+2} &= t^{1-1/2^j}(\boldsymbol{x}_{2j-1} + \boldsymbol{x}_{2j}) - \boldsymbol{x}_{2j+2}
\end{aligned}
\tag{8}
$$

and :

$$
\begin{aligned}
\tilde{\boldsymbol{w}}_1 &= t^2 - \bar{\boldsymbol{x}}_1 \\
\tilde{\boldsymbol{w}}_2 &= t - \bar{\boldsymbol{x}}_2 \\
\tilde{\boldsymbol{w}}_{3j} &= t\,\bar{\boldsymbol{x}}_{2j-1} - \bar{\boldsymbol{x}}_{2j+1} \\
\tilde{\boldsymbol{w}}_{3j+1} &= t\,\bar{\boldsymbol{x}}_{2j} - \bar{\boldsymbol{x}}_{2j+1} \\
\tilde{\boldsymbol{w}}_{3j+2} &= t^{1-1/2^j}(\bar{\boldsymbol{x}}_{2j-1} + \bar{\boldsymbol{x}}_{2j}) - \bar{\boldsymbol{x}}_{2j+2}
\end{aligned}
\tag{9}
$$

since $\bar{\boldsymbol{x}}$ is the Puiseux series of the leading terms of $\boldsymbol{x}$, the positivity of $\boldsymbol{w}$ yields that of $\tilde{\boldsymbol{w}}$. The same argument proves, using $\tilde{\boldsymbol{s}} := \boldsymbol{c} + \boldsymbol{A}\bar{\boldsymbol{y}}$, that $\tilde{\boldsymbol{s}}$ is a positive Puiseux series.

□

In order for a lifted Puiseux series $\boldsymbol{z}$ to be feasible, we only need to ensure that $\boldsymbol{x} > 0$, $\boldsymbol{w} > 0$, $\boldsymbol{y} > 0$, $\boldsymbol{s} > 0$. Notice that the two pair of conditions $\boldsymbol{x} > 0$, $\boldsymbol{w} > 0$ and $\boldsymbol{y} > 0$, $\boldsymbol{s} > 0$ are independent, therefore, we can state a feasibility condition on each pair independently.

It was showed in [1] that fixing a sequence of positive numbers $\alpha_0 = \frac{1}{2} > \alpha_1 > \cdots > \alpha_{r-1} > 0$ and defining $\boldsymbol{x}$ as

$$
\boldsymbol{x}_{2j+1} := \alpha_j t^{x^{\lambda}_{2j+1}}\,, \qquad \boldsymbol{x}_{2j+2} := \alpha_j t^{x^{\lambda}_{2j+2}}\,, \qquad (0 \le j < r) \tag{10}
$$

ensures the feasibility of the lifted pair $\boldsymbol{x}$, $\boldsymbol{w}$.

We will now state a necessary and sufficient condition for the feasibility of the pair $\boldsymbol{s} := \boldsymbol{c} + \boldsymbol{A}\boldsymbol{y}$, $\boldsymbol{y}$.

Proposition 3. *Let $\boldsymbol{z}$ be a lifted Puiseux series lifted from the tropical point. Let β be such that $\boldsymbol{y}_i := \beta_i t^{y^{\lambda}_i}$. The lifted pair ($\boldsymbol{s} := \boldsymbol{c} + \boldsymbol{A}\boldsymbol{y}$, $\boldsymbol{y}$) is feasible for every $\lambda \le 2$ if and only if:*

$$
\beta > 0
$$

$$
\begin{cases} 1 > \beta_3 + \beta_5 \\ \beta_2 > \beta_4 + \beta_5 \end{cases}
$$

$$
\forall j \begin{cases} \min(\beta_{3j}, \beta_{3j+1}) > \beta_{3j+3} + \beta_{3j+5} \\ \beta_{3j+2} > \beta_{3j+4} + \beta_{3j+5} \end{cases}
$$

$$
\begin{cases} \min(\beta_{3r-3}, \beta_{3r-2}) > \beta_{3r} \\ \beta_{3r-1} > \beta_{3r+1} \end{cases}
$$

Proof. Suppose β verifies the condition. We only have to prove that $\boldsymbol{s}, \boldsymbol{y} > 0$ to ensure the feasibility of the lifted pair. $\boldsymbol{y} > 0$ results directly from the positivity of β. Let us now prove the positivity of $\boldsymbol{s}$. By definition of the lifted pair, we have

$$\begin{aligned}
& \boldsymbol{s}_1 = \boldsymbol{y}_1 - t\boldsymbol{y}_3 - t^{\frac{1}{2}}\boldsymbol{y}_5 + 1 \\
& \boldsymbol{s}_2 = \boldsymbol{y}_2 - t\boldsymbol{y}_4 - t^{\frac{1}{2}}\boldsymbol{y}_5 && (11)\\
\forall 1 \leq j < r-1 \quad & \boldsymbol{s}_{2j+1} = \boldsymbol{y}_{3j} + \boldsymbol{y}_{3j+1} - t\boldsymbol{y}_{3j+3} - t^{1-\frac{1}{2^j}}\boldsymbol{y}_{3j+5} && (12)\\
& \boldsymbol{s}_{2j+2} = \boldsymbol{y}_{3j+2} - t\boldsymbol{y}_{3j+4} - t^{1-\frac{1}{2^j}}\boldsymbol{y}_{3j+5} && (13)\\
& \boldsymbol{s}_{2r-1} = \boldsymbol{y}_{3r-3} + \boldsymbol{y}_{3r-2} - \boldsymbol{y}_{3r} && (14)\\
& \boldsymbol{s}_{2r} = \boldsymbol{y}_{3r-1} - \boldsymbol{y}_{3r+1} && (15)
\end{aligned}$$

We will start by proving the following inequalities on y^λ that derive directly from equations 7:

$$\begin{aligned}
& 0 \geq \max(y_3^\lambda + 1, y_5^\lambda + \frac{1}{2}) && (16a)\\
& y_2^\lambda \geq \max(1 + y_4^\lambda, \frac{1}{2} + y_5^\lambda) && (16b)\\
\forall 1 \leq j < r-1 \quad & \max(y_{3j}^\lambda, y_{3j+1}^\lambda) \geq \max(1 + y_{3j+3}^\lambda, 1 - \frac{1}{2^j} + y_{3j+5}^\lambda) && (16c)\\
& y_{3j+2}^\lambda \geq \max(1 + y_{3j+4}^\lambda, 1 - \frac{1}{2^j} + y_{3j+5}^\lambda) && (16d)\\
& \max(y_{3r-3}^\lambda, y_{3r-2}^\lambda) \geq y_{3r} && (16e)\\
& y_{3r-1}^\lambda \geq y_{3r+1}^\lambda && (16f)
\end{aligned}$$

Because $\lambda \leq 2$, from 7a and 7d we have $y_3^\lambda + 1 = 0$ and $y_5^\lambda + \frac{1}{2} = \min(0, \lambda - 1) \leq 0$, hence $0 \geq \max(y_3^\lambda + 1, y_5^\lambda + \frac{1}{2})$ wich proves 16a.

7c and 7d yields to $y_4^\lambda + 1 = y_2^\lambda$ and $y_5^\lambda + \frac{1}{2} = \min(0, \lambda - 1) \leq \lambda - 1 = y_2^\lambda$ which proves 16b.

For $1 \leq j < r-1$, we have thanks to 7c and 7e, $1 + y_{3j+3}^\lambda = \max(y_{3j}^\lambda, y_{3j}^\lambda)$ and $y_{3j+5}^\lambda + 1 - \frac{1}{2^j} = \min(y_{3j+3}^\lambda + 1, y_{3j+2}^\lambda) \leq 1 + y_{3j+3}^\lambda = \max(y_{3j}^\lambda, y_{3j+1}^\lambda)$ proving 16c.

7c and 7e yield to $y_{3j+4}^\lambda + 1 = y_{3j+2}^\lambda$ and $y_{3j+5}^\lambda + 1 - \frac{1}{2^j} = \min(y_{3j+3}^\lambda + 1, y_{3j+2}^\lambda) \leq y_{3j+2}^\lambda$ which leads to 16d.

The two inequalities for 16e have already been proved.

From 7d we have $y_{3r-1}^\lambda = \min(y_{3r-3}^\lambda + 1, y_{3r-4}^\lambda) - (1 - \frac{1}{2^{r-1}})$. To prove 16f, it suffice to show that $y_{3r+1}^\lambda \leq y_{3r-3}^\lambda + \frac{1}{2^{r-1}}$ and $y_{3r+1}^\lambda \leq y_{3r-4}^\lambda + -(1 - \frac{1}{2^{r-1}})$. The former is a consequence of 7f and the latter is a consequence of 7c applied to $r - 1$.

16f is a direct consequence of 7f.

The inequalities 16 ensure that there exist a possible β such that $\boldsymbol{s} > 0$. In fact, it is a bare reformulation of the fact that the leading monomial of every coordinate of $\boldsymbol{s}$ must have a coefficient of the form $\sum_{j \in J+} \beta_j - \sum_{j \in J-} \beta_j$ where $J+, J- \subset [3r+1]$ and $J+ \neq \emptyset$ (the elements in inequations 16 being the powers of the terms in the expression of $\boldsymbol{s}$). Since $\beta > 0$, this is obviously a necessary condition for the positivity of $\boldsymbol{s}$. The rest of the conditions on β simply state that for all $\boldsymbol{s}$ coordinates leading monomial's coefficient of the form $\sum_{j \in J+} \beta_j - \sum_{j \in J-} \beta_j$, we have $\sum_{j \in J+} \beta_j \geq \sum_{j \in J-} \beta_j$ for the worst case scenario of $J+$, $J-$ (i.e $|J+| = 1$ and $|J-|$ maximal).

Reciprocally, let $(\boldsymbol{s}, \boldsymbol{y})$ be a lifted feasible point. $\boldsymbol{y} > 0$ implies that $\beta > 0$. We show by choosing particular cases of $\lambda < 2$, that each of the inequalities in 16 is saturated in a given case. Therefore, we reach the worse case scenarios on $J+$, $J-$ for all the coordinates of $\boldsymbol{s}$ proving therefore the necessity of the conditions over β.

$\square$

Remark 4. We denote ${}^sA := (\text{sign}(A_{i,j}))_{i,j}$. For a matrix $M \in \mathbb{R}^{m\times n}$ and a vector $v \in \mathbb{R}^n$, we define the product $\bullet$ as $(M \bullet v)_j = \min_k \frac{(M_{jk}+1)}{2}\beta_k + \sum_k \frac{(M_{jk}-1)}{2}\beta_k$ for all j. The previous conditions on β can be rewritten as:

$$\beta > 0,\ {}^sA \bullet \beta + c > 0$$

.

Remark 5. The previous formula suggest a generalization of the results to every standard linear problem. While the condition being sufficient is easy to generalize, the reciprocal is not obviously true. We suggest therefore the following proposition.

Proposition 6. *Consider a standard linear problem*

$$(P)\begin{cases} \inf_x\ \ c.x \\ Ax = b \\ x \geq 0 \end{cases}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, *and* $A \in \mathbb{R}^{n\times m}$ *and let* $\lambda > 0$. *Suppose the corresponding Puiseux series strictly feasible set* $\boldsymbol{\mathcal{F}^s} \cap \{\boldsymbol{z} \in \mathbb{K} | val(\boldsymbol{z}) = \lambda\}$ *is non empty. Let* $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{y}, \boldsymbol{s})$ *be a lifted Puiseux series lifted from a tropical point in the central path. Let* β *be such that* $\boldsymbol{y}_i := \beta_i t^{y_i^\lambda}$ *for all* i. *If* $\beta > 0,\ {}^sA \bullet \beta + c > 0$ *then* $(\boldsymbol{y}, \boldsymbol{s})$ *is strictly feasible.*

Choosing α and β under the constraints fixed above yields a lifted feasible point $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{y}, \boldsymbol{s})$ in $\boldsymbol{\mathcal{F}}$. For a large enough t, we know therefore that $\boldsymbol{z}(t)$ is a feasible point in $\boldsymbol{\mathcal{P}}(t)$. The problem of feasibility being solved, we still need to get a point in the neighbourhood V. The conditions over β being necessary, we know that we didn't lose many possibilities for the choice of the initial point (The eventual loses being in the choice of α).
α and β must be chosen to ensure that the lifted point is in V. It is very difficult to find a choice of α and β to satisfy this constraint because the equations defining V are not simple.

One way to overcome this problem is to start from a feasible point in $\boldsymbol{\mathcal{P}}(t)$ and get closer to V following Newton directions (correction steps). Once we find a feasible point in V we can launch the Predictor-Corrector on an instance of $\mathbf{LW}_r(t)$.

4 Experimentation

The following plots are outputs of running the predictor corrector algorithm presented in section 1 starting with the initial point constructed in section 3. We run the algorithm till $\bar{\mu}(z) < 1$ where z is the current point.

When t become large enough the image of a the central path of the problem $\mathbf{LW}_r(t)$ under the function $\log_t$ takes the form of the tropical central path which is a staircase (and the convergence is proven to be uniform in [1]). The image below illustrates this convergence and we see that the image of the central path under $\log_t$ starts to stick to the staircase form of the tropical central path (in green) showing (as expected in [1]) that, for a sufficiently large t, the total curvature of the central path is bounded from below exponentially in r.

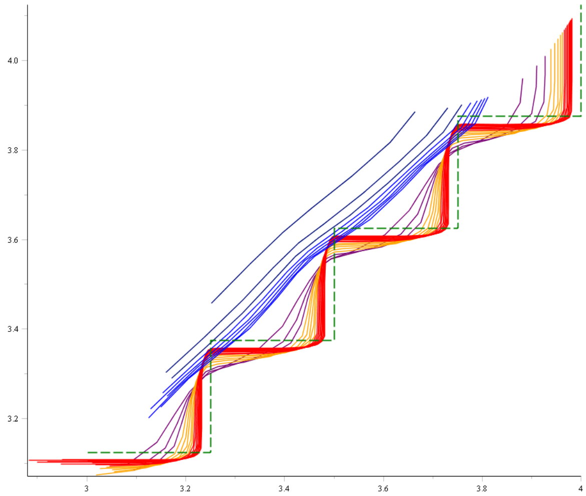

Figure 2: Deformation of the central path when t becomes large ($log_t - log_t$ scale of real central path) for r=4

The following plots show that the number of steps in the tropical central path is exponential in r and that the threshold $t_0(r)$ above which the staircase form appears is also exponential in r. On each r, we select t such that we can clearly see the staircase form of the path.

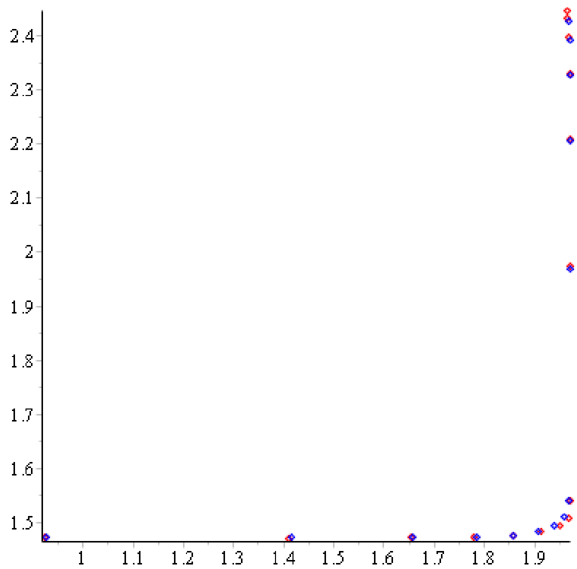

Figure 3: Projection on (x_{2r-1}, x_{2r}) of the Path followed by the predictor corrector on $\mathbf{LW}_r(t)$ for $r = 2$, $t = 10^{10}$. $log_t - log_t$ scale. *Red*: prediction steps, *blue*: correction steps.

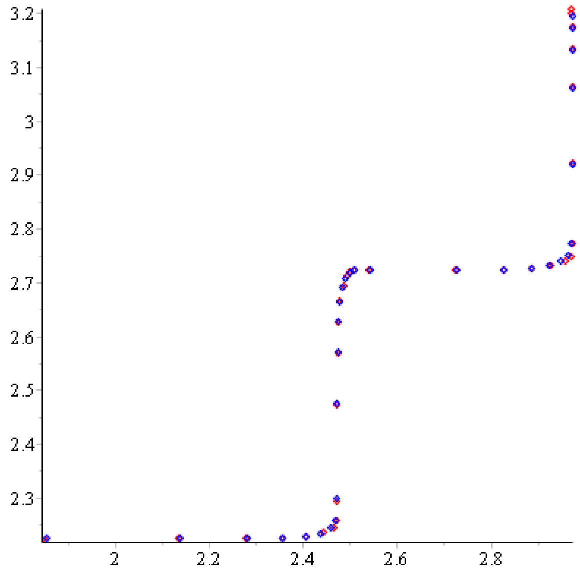

Figure 4: Projection on (x_{2r-1}, x_{2r}) of the Path followed by the predictor corrector on $\mathbf{LW}_r(t)$ for $r = 3$, $t = 10^{12}$. $log_t - log_t$ scale. *Red*: prediction steps, *blue*: correction steps.

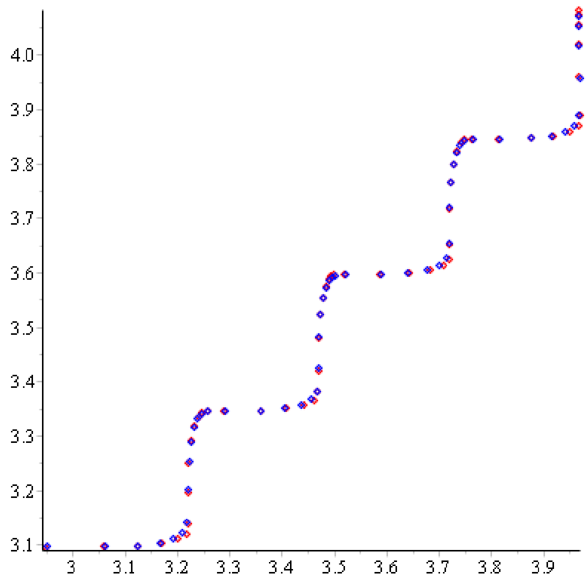

Figure 5: Projection on (x_{2r-1}, x_{2r}) of the Path followed by the predictor corrector on $\mathbf{LW}_r(t)$ for $r = 4$, $t = 10^{12}$. $log_t - log_t$ scale. *Red*: prediction steps, *blue*: correction steps.

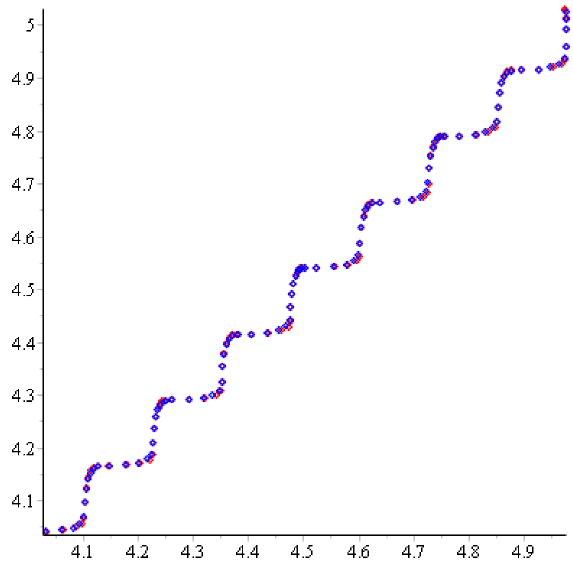

Figure 6: Projection on (x_{2r-1}, x_{2r}) of the Path followed by the predictor corrector on $\mathbf{LW}_r(t)$ for $r = 5$, $t = 10^{17}$. $log_t - log_t$ scale. *Red*: prediction steps, *blue*: correction steps.

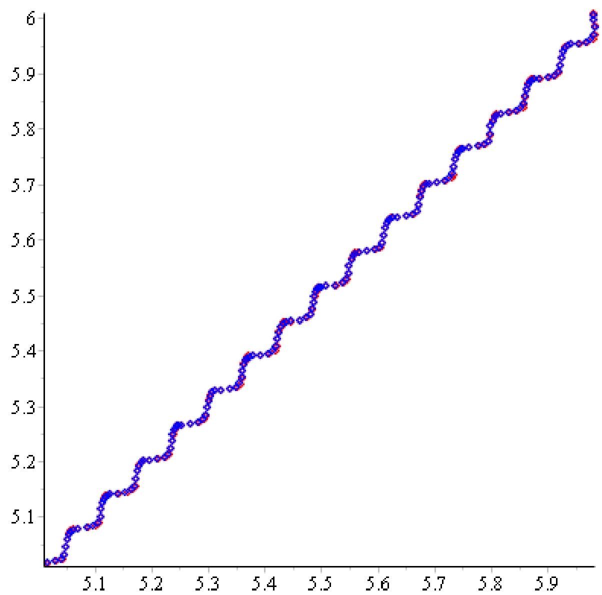

Figure 7: Projection on (x_{2r-1}, x_{2r}) of the Path followed by the predictor corrector on $\mathbf{LW}_r(t)$ for $r = 6$, $t = 10^{25}$. $log_t - log_t$ scale. *Red*: prediction steps, *blue*: correction steps.

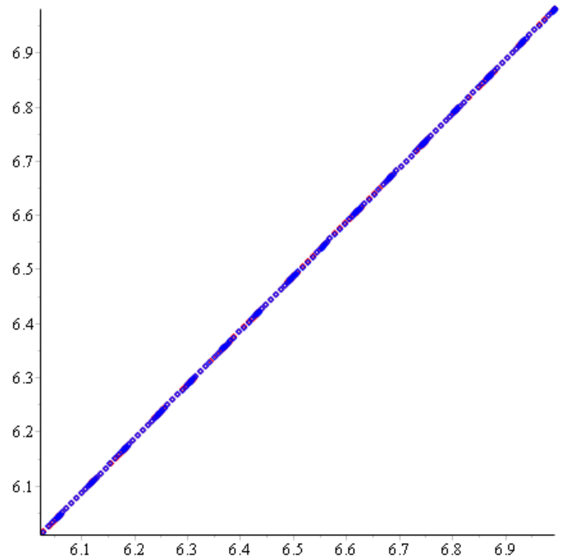

Figure 8: Projection on (x_{2r-1}, x_{2r}) of the Path followed by the predictor corrector on $\mathbf{LW}_r(t)$ for $r = 7$, $t = 10^{33}$. $log_t - log_t$ scale. *Red*: prediction steps, *blue*: correction steps.

Notice that in the last figure we cannot see the staircase form even with $t = 10^{33}$. This is due the fact that the threshold $t_0(r)$ increases exponentially with respect to r.

References

[1] Log-Barrier Interior Point Methods Are Not Strongly Polynomial X.Allamigeon, P.Benchimol, S.Gaubert, and M.JOSWIG

[2] https://who.rocq.inria.fr/Jean-Charles.Gilbert/ensta/optim.html