

A dynamical system implementation for linear and convex optimization problems

D. Aeyels and F. De Smet

*SYSTeMS Research Group,
Dept. of Electrical Energy, Systems and Automation,
Ghent University,
Technologiepark-Zwijnaarde 914, 9052 Zwijnaarde, Belgium
E-mail: Dirk.Aeyels@UGent.be, Filip.DeSmet@UGent.be*

Abstract

It is shown that, for a strictly convex cost function and linear constraints, the minimization of the cost can be implemented as a dynamical system. First we consider a separable cost function; later on the technique is generalized to minimization problems with a non-separable cost function. The dynamical system has a unique equilibrium point which is globally asymptotically stable. The equilibrium point corresponds to the minimum cost under the given constraints. We show that the dynamical system approach remains applicable when the cost function is linear, and therefore includes linear programming problems. When the coefficient matrix of the linear constraints is sparse, a distributed implementation is possible. We also discuss the relation with the Lagrange dual problem.

As an illustration we consider the minimum cost flow problem. The state variables of the dynamical system can be chosen to correspond with

A dynamical system implementation for optimization

the accumulation of the commodity in the nodes. In a distributed implementation the flow through an arc is determined by the state variables associated with its start and end node. In practical situations (e.g. compartmental systems) the state variables may correspond to physical quantities (e.g. water level in a water distribution system or the queue length of the buffer of packages waiting to be processed by a router in a computer network). The functions determining the flow are independent of the inflow or outflow at the nodes and therefore the dynamical system implementation is robust with respect to the external inflow or outflow. The distributed nature of the optimizing strategy makes for a convenient practical implementation. We also discuss the cases of non-convex cost functions to illustrate the potential of the method.

1 Introduction

In this paper we will show how the solution of a strictly convex optimization problem [3] with linear constraints can be implemented as the equilibrium point of a dynamical system, and we formulate conditions guaranteeing global convergence to this equilibrium point. We first investigate the case where the cost function is separable, i.e. can be written as a sum of functions of one variable. As a special case we will consider linear cost functions. In section 6 we show that when the coefficient matrix of the constraints is sparse, a distributed implementation is possible. Section 7 deals with the relation of our approach with the Lagrange dual problem. In section 8, as an illustration, the minimum cost flow problem is discussed from the dynamical system point of view. (When restricting to the minimum cost flow problem, some aspects of our approach may be compared to [6, 7]). The generalization of the results to non-separable cost functions is discussed in section 9.

2 Problem formulation

We consider a separable cost function $U : \mathbb{R}^M \rightarrow \mathbb{R} : y \mapsto U(y)$:

$$U(y) \triangleq \sum_{i=1}^M U_i(y_i), \quad \forall y \in \mathbb{R}^M,$$

where the functions $U_i : \mathbb{R} \rightarrow \mathbb{R}$ are continuously differentiable, strictly convex (i.e. U_i' is increasing), and have a (unique, because of the strict convexity) minimum. The variables y_i are subject to linear constraints:

$$Ay = b, \quad \forall y \in \mathbb{R}^M,$$

where $b \in \mathbb{R}^N$ and $A \in \mathbb{R}^{N \times M}$ has full row rank. The problem is to minimize $U(y)$ under the constraints $Ay = b$, which we will refer to as the problem (\mathcal{P}) . Since U has a minimum and is strictly convex, it is radially unbounded, and the restriction of U to the set $\{y \in \mathbb{R}^M : Ay = b\}$ therefore also has a minimum, and there will be a *unique* $y \in \mathbb{R}^M$, solving (\mathcal{P}) .

Remark 1. Hard bounds on the components y_i can be taken into account by relaxing the conditions on U_i . For instance, setting $U_i(y_i) = y_i^2, \forall y_i > 0$, and $U_i(y_i) = ay_i^2, \forall y_i \leq 0$, then considering the limit $a \rightarrow \infty$ will result in the bound $y_i \geq 0$. However, if these bounds are inconsistent with the constraints $Ay = b$, the problem is not feasible. For the analysis we will maintain the restriction on U_i to be well-defined on \mathbb{R} .

Introducing Lagrange multipliers $\lambda_i, i \in \{1, \dots, N\}$, and setting the partial derivative of the Lagrangian

$$\mathcal{U}(y, \lambda) \triangleq \sum_{i=1}^M U_i(y_i) + \sum_{j=1}^N \lambda_j \left(b_j - \sum_{i=1}^M A_{ji} y_i \right)$$

to y_i equal to zero results in

$$U'_i(y_i) - \sum_{j=1}^N \lambda_j A_{ji} = 0, \quad (1)$$

leading to

$$y_i = U_i'^{-1} \left(\sum_{j=1}^N \lambda_j A_{ji} \right) \triangleq \tilde{y}_i(\lambda), \quad \forall i \in \{1, \dots, M\}, \quad (2)$$

thereby defining the function $\tilde{y} : \mathbb{R}^N \rightarrow \mathbb{R}^M$. The unique solution $y \in \mathbb{R}^M$, solving (\mathcal{P}) , will be equal to the unique solution y to the set of conditions $Ay = b$ and $y = \tilde{y}(\lambda)$ for some $\lambda \in \mathbb{R}^N$. (Because of the strict convexity of U there is only one extremum, and this implies uniqueness of the solution to this set of conditions.)

3 Dynamical system description

To determine the values of the λ_i and y_i corresponding to the solution of (\mathcal{P}) , we will associate λ -values to the state $x(t) \in \mathbb{R}^N$ of a dynamical system, through Lipschitz continuous functions $\tilde{\lambda}_i : \mathbb{R}^N \rightarrow \mathbb{R}^N$, $i \in \{1, \dots, N\}$, such that, as $t \rightarrow \infty$, $x(t)$ converges to an equilibrium state x^e , where $\tilde{y}(\tilde{\lambda}(x^e))$ solves (\mathcal{P}) . For $\tilde{y} \circ \tilde{\lambda}$ to be properly defined on \mathbb{R}^N , we require (because of (2)) that the image of the function $\sum_{j=1}^N A_{ji} \tilde{\lambda}_j$ is contained in the image of U'_i , for all $i \in \{1, \dots, M\}$. We propose the dynamical system:

$$\dot{x}(t) = b - A\tilde{y}(\tilde{\lambda}(x(t))), \quad \forall t \in \mathbb{R}. \quad (3)$$

Notice that, if x^e is an equilibrium point of (3), then, with $y \triangleq \tilde{y}(\tilde{\lambda}(x^e))$, the constraints $Ay = b$ are fulfilled. By definition of the function \tilde{y} , the condition

(2) is also fulfilled with $\lambda = \tilde{\lambda}(x^e)$, and therefore y solves the problem (\mathcal{P}) .

4 Stability

In this section we introduce additional conditions to guarantee that any solution of the dynamical system (3) converges to an equilibrium point, and therefore asymptotically implements the solution of (\mathcal{P}) .

For the remainder of this paper we will assume that the functions $\tilde{\lambda}_i$ are continuously differentiable and that the functions $U'_i : \mathbb{R} \rightarrow \mathbb{R}$ are surjective, removing the restriction (stated in the previous section) on the image of $\sum_{j=1}^N A_{ji} \tilde{\lambda}_j$. Denote by $\frac{\partial \tilde{\lambda}}{\partial x}$ the Jacobian matrix of the vector function $\tilde{\lambda}$, i.e. $(\frac{\partial \tilde{\lambda}}{\partial x})_{ij} \triangleq \frac{\partial \tilde{\lambda}_i}{\partial x_j}$, and set $\mathcal{S}_{\tilde{\lambda}} \triangleq \frac{1}{2} \left(\frac{\partial \tilde{\lambda}}{\partial x} + \frac{\partial \tilde{\lambda}^T}{\partial x} \right)$. We require that

(C1) $\mathcal{S}_{\tilde{\lambda}}(x) \geq 0$ (i.e. $\mathcal{S}_{\tilde{\lambda}}(x)$ is positive semi-definite), $\forall x \in \mathbb{R}^N$,

(C2) the set $\{x \in \mathbb{R}^N : \det(\mathcal{S}_{\tilde{\lambda}}(x)) = 0\}$ consists of isolated points,

(C3) the function $\|\tilde{\lambda}\|$ is radially unbounded in \mathbb{R}^N , i.e. $\|x\| \rightarrow \infty$ implies that $\|\tilde{\lambda}(x)\| \rightarrow \infty$,

(where $\|x\| = \sqrt{x^T x}$, $\forall x \in \mathbb{R}^N$).

Consider the function $V : \mathbb{R}^N \rightarrow \mathbb{R}$ defined by

$$V(\lambda) \triangleq - \sum_{i=1}^N b_i \lambda_i + \sum_{j=1}^M \int_0^{\sum_{k=1}^N \lambda_k A_{kj}} U_j'^{-1}(\sigma) d\sigma. \quad (4)$$

Since

$$\begin{aligned} \frac{\partial V}{\partial \lambda_i}(\lambda) &= -b_i + \sum_{j=1}^M A_{ij} U_j'^{-1} \left(\sum_{k=1}^N \lambda_k A_{kj} \right) \\ &= -b_i + \sum_{j=1}^M A_{ij} \tilde{y}_j(\lambda), \end{aligned}$$

and thus, with x a solution of (3),

$$\frac{\partial V}{\partial \lambda_i}(\tilde{\lambda}(x(t))) = -\dot{x}_i(t),$$

it follows that the derivative of $V \circ \tilde{\lambda} \circ x$ satisfies

$$\begin{aligned} \frac{d(V \circ \tilde{\lambda} \circ x)}{dt}(t) &= -\dot{x}^T(t) \frac{\partial \tilde{\lambda}}{\partial x}(x(t)) \dot{x}(t) \\ &= - \left(b - A\tilde{y}(\tilde{\lambda}(x(t))) \right)^T \mathcal{S}_{\tilde{\lambda}}(x(t)) \left(b - A\tilde{y}(\tilde{\lambda}(x(t))) \right) \\ &\leq 0, \quad \text{by (C1),} \end{aligned}$$

$\forall t \in \mathbb{R}$, and $V \circ \tilde{\lambda} : \mathbb{R}^N \rightarrow \mathbb{R}$ is therefore a Lyapunov function for the system (3).

Remark 2. If $\mathcal{S}_{\tilde{\lambda}}(x)$ is positive definite everywhere in \mathbb{R}^N , one can introduce a new metric, represented by the matrix function $\mathcal{S}_{\tilde{\lambda}}$ (i.e. $\langle v(x), w(x) \rangle = v^T(x) \mathcal{S}_{\tilde{\lambda}}(x) w(x)$, for any two vector functions v and w in \mathbb{R}^N), and then $V \circ \tilde{\lambda}$ is a gradient function for the system (3) with respect to this metric.

Proposition 1. *Under the assumption (C3) the function $V \circ \tilde{\lambda}$ is radially unbounded.*

The proof is given in the appendix, section A.1.

Since the Lyapunov function $V \circ \tilde{\lambda}$ is radially unbounded, every solution x of (3) has a non-empty, bounded positive limit set, which is contained in the largest invariant subset S of

$$\left\{ x \in \mathbb{R}^N : - \left(b - A\tilde{y}(\tilde{\lambda}(x)) \right)^T \mathcal{S}_{\tilde{\lambda}}(x) \left(b - A\tilde{y}(\tilde{\lambda}(x)) \right) = 0 \right\}.$$

If there are $x \in S$ which are not equilibrium points of (3), then they satisfy $\det(\mathcal{S}_{\tilde{\lambda}}(x)) = 0$, and thus by (C2) these points are isolated from each other. They are also isolated from other points in S . If they would not be isolated in S , then

there would exist $x^i \in S$, $i \in \mathbb{N}$, converging to x , and satisfying $b - A\tilde{y}(\tilde{\lambda}(x^i)) = 0$, $\forall i \in \mathbb{N}$. By continuity of $\tilde{y} \circ \tilde{\lambda}$, x would also satisfy $b - A\tilde{y}(\tilde{\lambda}(x)) = 0$ and be an equilibrium point of (3), leading to a contradiction. Therefore x is also isolated in S , and by the invariance of S it follows again that x is an equilibrium point of (3), again leading to a contradiction. Consequently S consists of the equilibrium points of (3). We now show that (3) cannot have more than one equilibrium point.

Assume there are two different equilibrium points $x^{e,1}$ and $x^{e,2}$ of (3). Since both $x^{e,1}$ and $x^{e,2}$ implement the *unique* solution of the problem (\mathcal{P}) , it follows from (2) that $\tilde{\lambda}^T(x^{e,1})A = \tilde{\lambda}^T(x^{e,2})A$, and thus $\tilde{\lambda}(x^{e,1}) = \tilde{\lambda}(x^{e,2})$ since A has full row rank. On the other hand

$$\begin{aligned}
& (x^{e,1} - x^{e,2})^T \left(\tilde{\lambda}(x^{e,1}) - \tilde{\lambda}(x^{e,2}) \right) \\
&= \int_0^1 (x^{e,1} - x^{e,2})^T \frac{\partial \tilde{\lambda}}{\partial x}((1-t)x^{e,2} + tx^{e,1}) (x^{e,1} - x^{e,2}) dt \\
&= \int_0^1 (x^{e,1} - x^{e,2})^T \mathcal{S}_{\tilde{\lambda}}((1-t)x^{e,2} + tx^{e,1}) (x^{e,1} - x^{e,2}) dt \\
&> 0, \tag{5}
\end{aligned}$$

since the integrand is zero in isolated points only (within the interval $[0,1]$). This contradicts the fact that $\tilde{\lambda}(x^{e,1}) = \tilde{\lambda}(x^{e,2})$.

Consequently, the system (3) has a *unique* equilibrium point, which is globally asymptotically stable. Summing up, we obtain the following theorem:

Theorem 1. *Under the assumptions (C), the system (3) has a globally asymptotically stable equilibrium point, for which the image under the mapping $\tilde{y} \circ \tilde{\lambda}$ solves (\mathcal{P}) .*

5 Linear cost functions

A considerable part of the research on optimization deals with the case where the cost $U(y)$ is linear in y and the components y_i are subject to hard bounds [2, 5]. A common formulation is the following: minimize

$$u^T y$$

with respect to y , under the constraints

$$\begin{cases} Ay = b \\ y \geq 0, \end{cases}$$

where $b \in \mathbb{R}^N$ and $u, y \in \mathbb{R}^M$ are column vectors, $A \in \mathbb{R}^{N \times M}$ and $y \geq 0$ is a compact notation for $y_i \geq 0, \forall i \in \{1, \dots, M\}$. Because of the constraints a solution does not necessarily exist, and since the cost is not strictly convex, a solution may not be unique. Throughout this section we will only consider the case for which a solution exists, and we assume that it is unique (which corresponds to the generic case). We will adopt the same cost function $U(y) = u^T y$ and constraints $Ay = b$, but instead of imposing the condition $y \geq 0$, we will assume that the components y_i are restricted to an interval $[c_i, d_i]$, with c_i and d_i finite. For a specific, feasible problem with $d_i = +\infty$, for some $i \in \{1, \dots, M\}$, one can in practice always replace d_i by a finite but sufficiently large value, and therefore our approach can also be applied to the general case with bounds of the form $y_i \geq 0$ only. Taking into account remark 1, the problem can be considered as a limit case of (\mathcal{P}) , but it is not included in (\mathcal{P}) since the cost is not strictly convex.

The bound on the components y_i can be taken into account by writing the

functions U_i as

$$U_i(y_i) = \begin{cases} u_i y_i, & y_i \in [c_i, d_i]; \\ +\infty, & y_i \notin [c_i, d_i]. \end{cases}$$

The functions U_i' are not invertible, but by approximating U_i by strictly convex functions, we are led to the following definition of \tilde{y} :

$$\tilde{y}_i(\lambda) \triangleq \begin{cases} c_i, & \text{if } \sum_{j=1}^N \lambda_j A_{ji} < u_i; \\ d_i, & \text{if } \sum_{j=1}^N \lambda_j A_{ji} > u_i. \end{cases}$$

The corresponding differential equations (3) have discontinuous right hand sides, which raises issues with regard to the existence of solutions of the resulting system. Although approaches exist to extend the solution concept to discontinuous differential equations (see [4]), we favor a heuristic approach. Notice that there is a gradient function $V_{\mathcal{L}} \circ \tilde{\lambda}$ with

$$V_{\mathcal{L}}(\lambda) \triangleq -\lambda^T b + \frac{1}{2} (\lambda^T A - u^T) (c + d) + \frac{1}{2} |\lambda^T A - u^T| (d - c), \quad \forall \lambda \in \mathbb{R}^N,$$

(where the absolute value of a vector is taken componentwise) which corresponds to the definition of V in (4), except for a constant term (the lower bounds of the integrals in (4) were replaced by u_i). The function $V_{\mathcal{L}}$ is convex and piecewise linear and by the assumptions on the existence of a unique solution, $V_{\mathcal{L}}$ has an absolute minimum in some $\tilde{\lambda}(x^e) \in \mathbb{R}^N$. The point x^e is no longer an equilibrium point in the classical sense, since the right hand sides in (3), which in the present case can only take a finite number of values, will not be equal to zero in x^e (in the generic case we are considering). In a simulation with a constant time step the trajectory will approach x^e at first, but since the right hand sides in (3) will not approach zero, chattering occurs around x^e and as a consequence there is an error proportional to the size of the time step (see figure 1 for an illustration

concerning the minimum cost flow problem, which is discussed in section 8). If the time step is sufficiently small, this results in a good approximation of x^e . The error can be reduced by using a variable time step. Since the right hand sides of (3) are piecewise constant along a solution x , the time instances at which $\dot{x}(t)$ changes can be exactly calculated, and by adjusting the time step appropriately the exact trajectory of x can be obtained, as well as the exact value of x^e , resulting in an algorithmic approach.

Notice that x^e does not completely determine all M components of $\tilde{y}(\tilde{\lambda}(x^e))$ since $\tilde{y}_i(\lambda)$ is not properly defined if $(\lambda^T A)_i = u_i$ for some $i \in \{1, \dots, M\}$. In the generic case at most N equations $(\lambda^T A)_i = u_i$ ($i \in \{1, \dots, M\}$) can be satisfied, since λ has only N components. In case less than N equations are satisfied, there are infinitely many solutions for λ satisfying these equalities, contradicting the fact that (in the generic case) the minimum of $V_{\mathcal{L}}$ is unique. Since $\tilde{y}_i(\lambda)$ is determined by λ if and only if $(\lambda^T A)_i \neq u_i$, it follows that there are exactly N unknown components of the solution $y = \tilde{y}(\tilde{\lambda}(x^e))$, which can be derived from the constraints $Ay = b$, or they can be calculated as long term averages of the components $\tilde{y}(\tilde{\lambda}(x(t)))$ resulting from simulations. Considering that a simulation trajectory x will remain bounded (even though there is no convergence to x^e), it follows that $\int_0^t (b - A\tilde{y}(\tilde{\lambda}(x(t'))))dt'$ is bounded, and $y \triangleq \lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \tilde{y}(\tilde{\lambda}(x(t'))))dt'$ will satisfy $Ay = b$.

6 Distributed implementation

When the matrix A is sparse, a limited number of components of $\tilde{y}(\tilde{\lambda}(x(t)))$ may be sufficient to update the state variable $x_i(t)$ ($i \in \{1, \dots, N\}$) by (3), and a limited number of components of $\tilde{\lambda}(x(t))$ may be sufficient to calculate $\tilde{y}_j(\tilde{\lambda}(x(t)))$ ($j \in \{1, \dots, M\}$) from (2). Choosing $\tilde{\lambda}$ such that each component $\tilde{\lambda}_k(x(t))$ depends on a limited number of components of $x(t)$, facilitates a distributed

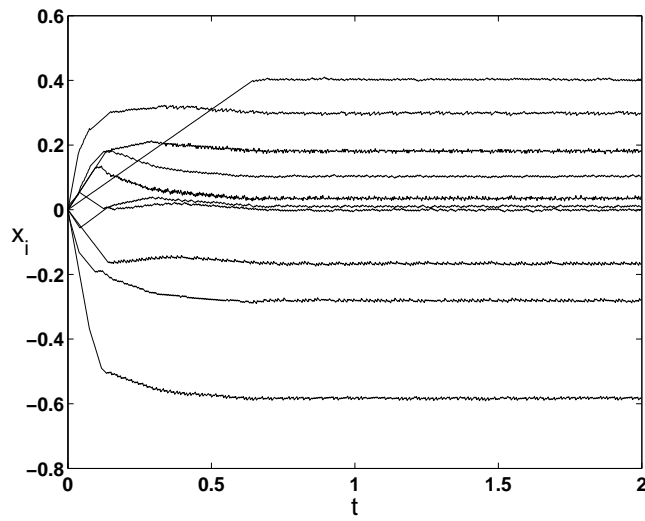


Figure 1: Simulation of the time evolution of x_i for a linear cost function with time step 0.003, $N = 10$, A equal to the incidence matrix of an all-to-all coupled directed network (see section 8), u_i and b_i chosen randomly from a uniform distribution on $[-1, 1]$ and $[-\frac{1}{2}, \frac{1}{2}]$ respectively (the values b_i were shifted to guarantee that they sum up to zero), $c_i = -1$ and $d_i = 1$.

implementation of the system (3). We therefore investigate the conditions (C) on $\tilde{\lambda}$ and $\mathcal{S}_{\tilde{\lambda}}$ when each of the functions $\tilde{\lambda}_i$ depends only on the state component $x_i(t)$:

$$\tilde{\lambda}_i(x) = \bar{\lambda}_i(x_i), \quad \forall x \in \mathbb{R}^N, \quad \forall i \in \{1, \dots, N\},$$

for some continuously differentiable functions $\bar{\lambda}_i : \mathbb{R} \rightarrow \mathbb{R}$.

The conditions (C) can be reformulated as follows.

1. $\bar{\lambda}'_i(x_i) \geq 0, \forall x_i \in \mathbb{R}, \forall i \in \{1, \dots, N\}$;
2. if $N > 1$, then $\bar{\lambda}'_i(x_i) > 0, \forall x_i \in \mathbb{R}, \forall i \in \{1, \dots, N\}$; if $N = 1$, then $\bar{\lambda}'_1$ is allowed to become zero in isolated points;
3. $\lim_{x_i \rightarrow \pm\infty} \bar{\lambda}_i(x_i) = \pm\infty, \forall i \in \{1, \dots, N\}$.

(Item 1 is equivalent to condition (C1). If item 1 holds, then item 2, resp. item 3, is equivalent to (C2), resp. (C3).)

7 Relation with the Lagrange dual problem

Notice that our approach also implies that the solution of (P) corresponds to the minimum of the function V . We will show that this result relates to the Lagrange dual problem [3, p. 223] to (P).

The Lagrange dual problem can be formulated as follows: maximize V_{dual} , with

$$V_{\text{dual}}(\lambda) \triangleq \min_{y \in \mathbb{R}^N} \mathcal{U}(y, \lambda), \quad \forall \lambda \in \mathbb{R}^N.$$

Minimizing \mathcal{U} over y again leads to (2), and thus

$$\begin{aligned} V_{\text{dual}}(\lambda) &= \mathcal{U}(\tilde{y}(\lambda), \lambda) \\ &= \lambda^T b + \sum_{i=1}^M \left(U_i(U_i'^{-1}((\lambda^T A)_i)) - (\lambda^T A)_i U_i'^{-1}((\lambda^T A)_i) \right). \end{aligned}$$

Since the derivative of the function mapping $\sigma \in \mathbb{R}$ to

$$U_i(U_i'^{-1}(\sigma)) - \sigma U_i'^{-1}(\sigma)$$

equals

$$-U_i'^{-1}(\sigma),$$

it follows that V_{dual} is, up to a constant term, equal to $-V$, and the minimization of the Lyapunov function V is equivalent to the maximization of the (Lagrange) dual function V_{dual} . In section 9 we discuss non-separable cost functions; we will explicitly use the dual function to construct a Lyapunov function, since it is not intuitively clear which Lyapunov function corresponds to the dynamical system associated with the optimization problem.

8 The minimum cost flow problem

As an illustration of the previous results we will consider the minimum cost flow problem [1, p. 9] (also referred to as the optimal distribution problem) with a separable and strictly convex cost function.

Consider a connected, directed network, with $N > 1$ nodes and M arcs, with at most one arc between each pair of nodes (and thus $N - 1 \leq M \leq \frac{1}{2}N(N - 1)$).

We define the incidence matrix $A \in \mathbb{R}^{N \times M}$ by

$$A_{ij} \triangleq \begin{cases} 1, & \text{if } i \text{ is the start node of the } j\text{th arc,} \\ -1, & \text{if } i \text{ is the end node of the } j\text{th arc,} \\ 0, & \text{if the } j\text{th arc does not start or end in node } i. \end{cases}$$

The rows of the matrix A sum up to zero, i.e. $\mathbf{1}^T A = 0$, with $\mathbf{1} \triangleq [1 \dots 1]^T \in \mathbb{R}^N$.

Because of the connectedness of the network, one can show that A has rank $N - 1$

(see e.g. [2, p. 282]), and therefore all solutions λ of the equation $\lambda^T A = 0$ are scalar multiples of $\mathbf{1}$.

Node i is assigned a value b_i , corresponding to inflow ($b_i \geq 0$) or outflow ($b_i \leq 0$) of the commodity under consideration. The k th arc carries a flow y_k from node i to node j ($A_{ik} = 1$, $A_{jk} = -1$) such that no accumulation takes place in the nodes:

$$Ay = b, \quad (6)$$

where A does not have full row rank since its rows sum up to zero. To guarantee the existence of a flow vector y satisfying $Ay = b$ we impose that $\mathbf{1}^T b = \sum_{i=1}^N b_i = 0$. Although this allows to drop one of the equalities in (6), we prefer to keep (6) in its present form as this allows a physical interpretation of the dynamical system (3) and a more symmetrical notation. The consequences of this choice will be discussed below.

8.1 The dynamical system

From equation (2) it follows that

$$\tilde{y}_k(\lambda) = U_k'^{-1}(\lambda_i - \lambda_j) \triangleq -f_{ij}(\lambda_j - \lambda_i) \triangleq f_{ji}(\lambda_i - \lambda_j), \quad (7)$$

where i is the start node and j is the end node of arc k . Denoting by $\mathcal{N}_i \subset \{1, \dots, N\}$ ($i \in \{1, \dots, N\}$) the set of the nodes connected to node i (irrespective of the direction of the connecting arc), and taking into account the definition of f_{ij} in (7), we can write (3) as

$$\dot{x}_i(t) = b_i + \sum_{j \in \mathcal{N}_i} f_{ij}(\tilde{\lambda}_j(x(t)) - \tilde{\lambda}_i(x(t))), \quad \forall t \in \mathbb{R}, \quad \forall i \in \{1, \dots, N\}. \quad (8)$$

In the equations (8), the time-derivative of the state-variable x_i equals the

excess flow in node i , leading to interpret x_i as the level of the commodity at node i . Since the right hand sides in (8) sum up to zero it follows that $\sum_{i=1}^N x_i$ is constant, i.e. since there is no net in- or outflow, there is conservation of the total amount of commodity in the network. This also implies that, independent of the choice of the functions $\tilde{\lambda}_i$, global convergence to a single equilibrium point is impossible unless we restrict the dynamical system (8) to the state space

$$L_C \triangleq \left\{ x \in \mathbb{R}^N : \sum_{i=1}^N x_i = C \right\},$$

with $C \in \mathbb{R}$.

For any choice of $b \in L_0$, the solution of the minimum cost flow problem satisfies $y^* = \tilde{y}(\lambda^*)$ for some $\lambda^* \in \mathbb{R}^N$ and unique optimal flows y_i^* . It follows that any $x^e \in \mathbb{R}^N$ satisfying $\tilde{\lambda}^T(x^e)A = \lambda^{*T}A$ (or, equivalently: $\tilde{\lambda}(x^e) - \lambda^*$ is a scalar multiple of $\mathbf{1}$, and therefore, $\tilde{\lambda}_j(x^e) - \tilde{\lambda}_i(x^e) = \lambda_j^* - \lambda_i^*, \forall i, j \in \{1, \dots, N\}$ with $i \neq j$) is an equilibrium point of (8), which implements the solution of the minimum cost flow problem. Since λ is not uniquely determined by the value of $\lambda^T A$, adaptation is required of the conditions (C) in section 4.

8.2 Stability

We retain conditions (C1) and (C2), but change condition (C3), resulting in the conditions (C') (with (C'1) equal to (C1) and (C'2) equal to (C2)):

(C'1) $\mathcal{S}_{\tilde{\lambda}}(x) \geq 0$ (i.e. $\mathcal{S}_{\tilde{\lambda}}(x)$ is positive semi-definite), $\forall x \in \mathbb{R}^N$,

(C'2) the set $\{x \in \mathbb{R}^N : \det(\mathcal{S}_{\tilde{\lambda}}(x)) = 0\}$ consists of isolated points,

(C'3) for each $C \in \mathbb{R}$ the function $\Delta\tilde{\lambda}$, with $\Delta\tilde{\lambda}(x) \triangleq \max_{i \neq j} |\tilde{\lambda}_j(x) - \tilde{\lambda}_i(x)|$, is radially unbounded in L_C .

Because of the connectedness of the network, the unboundedness of $\Delta\tilde{\lambda}$ is equivalent to the unboundedness of $\|\tilde{\lambda}^T A\|$, and the proof in section A.1 of the

appendix can be adjusted (see section A.2) to obtain the following adaptation of proposition 1 concerning $V \circ \tilde{\lambda}$, with V defined by (4):

Proposition 2. *Under the assumption (C'3) the restriction of the function $V \circ \tilde{\lambda}$ to the set L_C is radially unbounded for each $C \in \mathbb{R}$.*

Since the Lyapunov function $V \circ \tilde{\lambda}|_{L_C}$ is radially unbounded, we can conclude that every solution x of (8) with $x(0) \in L_C$ will converge to the largest invariant subset of

$$\left\{ x \in L_C : - \left(b - A\tilde{y}(\tilde{\lambda}(x)) \right)^T \mathcal{S}_{\tilde{\lambda}}(x) \left(b - A\tilde{y}(\tilde{\lambda}(x)) \right) = 0 \right\},$$

which consists of the equilibrium points of (8) belonging to L_C . We now show that there cannot be more than one equilibrium point in L_C .

Assume there are two different equilibrium points $x^{e,1}$ and $x^{e,2}$ of (8), both belonging to L_C . Since both $x^{e,1}$ and $x^{e,2}$ implement the unique solution of the minimum cost flow problem, it follows that $\tilde{\lambda}(x^{e,1}) - \tilde{\lambda}(x^{e,2}) = \gamma \mathbf{1}$, for some $\gamma \in \mathbb{R}$, and thus $(x^{e,1} - x^{e,2})^T (\tilde{\lambda}(x^{e,1}) - \tilde{\lambda}(x^{e,2})) = \gamma \sum_{i=1}^N (x_i^{e,1} - x_i^{e,2}) = \gamma C - \gamma C = 0$, contradicting (5) (which is based only on conditions (C1) = (C'1) and (C2) = (C'2)).

Consequently, the restriction of the system (8) to the state space L_C has a unique equilibrium point, which is globally asymptotically stable. Summing up, we obtain the following adaptation of theorem 1:

Theorem 2. *Under the assumptions (C'), every solution of the system (8) will converge to an equilibrium point, for which the image under the mapping $\tilde{y} \circ \tilde{\lambda}$ solves the minimum cost flow problem.*

An important observation is that, in practical problems where the x_i correspond to physical quantities which can be measured, and the flows can be controlled, one can implement (7) as a control law based on observation of x .

By choosing $\tilde{\lambda}_i$ independent of the vector b , the functions $\tilde{y} \circ \tilde{\lambda}$ are also independent of b , and in order to attain the minimizing flow distribution, the in- and outflow in the different nodes do not have to be known, and may also slowly (i.e. w.r.t. the transient time of the system (8)) vary in time. In other words, the corresponding practical implementation of the solution of the minimum cost flow problem is robust w.r.t. the inflow vector b .

8.3 Distributed implementation

We restrict each of the functions $\tilde{\lambda}_i$ to depend only on the state component $x_i(t)$:

$$\tilde{\lambda}_i(x) = \bar{\lambda}_i(x_i), \quad \forall x \in \mathbb{R}^N, \quad \forall i \in \{1, \dots, N\}.$$

The conditions (C') from section 8.2 again imply that $\bar{\lambda}'_i(x_i) > 0, \forall x_i \in \mathbb{R}, \forall i \in \{1, \dots, N\}$ (since $N > 1$). Under this condition, (C'3) is equivalent to the realization of (at least) one of the following conditions. (The proof is given in the appendix, section A.3.)

(D1) The functions $\bar{\lambda}_i$ all satisfy $\lim_{x_i \rightarrow +\infty} \bar{\lambda}_i(x_i) = +\infty$.

(D2) The functions $\bar{\lambda}_i$ all satisfy $\lim_{x_i \rightarrow -\infty} \bar{\lambda}_i(x_i) = -\infty$.

(D3) All but one of the functions $\bar{\lambda}_i$ satisfy $\lim_{x_i \rightarrow \pm\infty} \bar{\lambda}_i(x_i) = \pm\infty$.

Notice that the function $\tilde{y}_k \circ \tilde{\lambda}$ determines the flow y_k in terms of the ‘local variables’ x_i and x_j , which are associated to the nodes connected by the k th arc, and that \tilde{y}_k is determined by the ‘local cost’ U_k .

8.4 Non-convex cost functions

When U_i is not convex, U'_i is not monotone and therefore possesses no continuous inverse. Each monotone branch can be associated to an inverse, only defined in

some part of \mathbb{R} . Each combination of convex branches from different U_i may lead to (at most) one local minimum of $U(y)$ under (6), which can be implemented by the system (8) (but is not guaranteed to be an absolute minimum). However, the realization of the minimum cost may also involve a strictly concave branch, and then the corresponding equilibrium solution of (8) may be unstable. In this case a *stabilizing* distributed implementation is still possible but will depend on the values of the b_i , and will therefore not be robust w.r.t. changes of the b_i .

Example 1. Consider a network of three nodes with incidence matrix

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix},$$

i.e. arc 1 starts in node 2 and ends in node 3, arc 2 starts in node 3 and ends in node 1, arc 3 starts in node 1 and ends in node 2. For simplicity we will consider a distributed implementation with $\bar{\lambda}_i(x_i) \triangleq x_i, \forall x_i \in \mathbb{R}, \forall i \in \{1, 2, 3\}$, resulting in the system

$$\begin{aligned} \dot{x}_1(t) &= b_1 + U_2'^{-1}(x_3(t) - x_1(t)) - U_3'^{-1}(x_1(t) - x_2(t)) \\ \dot{x}_2(t) &= b_2 + U_3'^{-1}(x_1(t) - x_2(t)) - U_1'^{-1}(x_2(t) - x_3(t)) \\ \dot{x}_3(t) &= b_3 + U_1'^{-1}(x_2(t) - x_3(t)) - U_2'^{-1}(x_3(t) - x_1(t)), \end{aligned} \quad (9)$$

where we choose as cost functions

$$\begin{aligned} U_1(y_1) &\triangleq \frac{1}{2}y_1^2, \\ U_2(y_2) &\triangleq \frac{1}{2}y_2(-2 + |y_2|), \\ U_3(y_3) &\triangleq \frac{1}{2}y_3^2. \end{aligned}$$

For $y_2 < 0$ the cost is not convex, but for all choices of $b \in \mathbb{R}^3$, with $b_1 + b_2 +$

$b_3 = 0$, the cost $U(y)$ still has a minimal value under the constraints (6). The minimizing flows can be written as

$$\begin{aligned} y_1 &= \frac{1}{3} (b_1 - b_2 + 2 - |b_1 - b_3 + 1|) \\ y_2 &= \frac{1}{3} (2(b_1 - b_3 + 1) - |b_1 - b_3 + 1|) \\ y_3 &= \frac{1}{3} (b_2 - b_3 + 2 - |b_1 - b_3 + 1|). \end{aligned}$$

For $b_1 - b_3 + 1 \geq 0$ the minimum only involves convex branches and can be implemented by the system (9) with $U_1'^{-1}(x) = x$, $U_2'^{-1}(x) = 1 + x$, and $U_3'^{-1}(x) = x$. The eigenvalues associated with its Jacobian matrix are 0, -3 and -3 .

For $b_1 - b_3 + 1 < 0$ the minimum involves the concave branch of U_2 and the corresponding system (9) with $U_1'^{-1}(x) = x$, $U_2'^{-1}(x) = -1 - x$, and $U_3'^{-1}(x) = x$, has eigenvalues 0, -3 and 1, and is unstable.

An implementation by a single stable system, implementing the solution of the minimum cost flow problem, can be achieved e.g. by the choice $U_1'^{-1}(x) = x$, $U_2'^{-1}(x) = 1 + x + 2 \min(0, b_1 - b_3 + 1)$, and $U_3'^{-1}(x) = x$, but due to the dependence on the value of $b_1 - b_3$, there is no robustness w.r.t. b .

9 Optimization with a non-separable cost function

The relation of our approach of (\mathcal{P}) with the Lagrange dual problem discussed in section 7 allows us to consider the general case where U cannot be written as a sum of functions of one variable, i.e. U is non-separable. To a large extent, the discussion of this general case runs parallel with the treatment of (\mathcal{P}) .

We consider the more general problem (\mathcal{G}) to minimize the cost $U(y)$ ($y \in \mathbb{R}^M$, U continuously differentiable, strictly convex, and having a minimum value)

under the constraints $Ay = b$ ($A \in \mathbb{R}^{N \times M}$, $b \in \mathbb{R}^N$), where A has full row rank. Since U has a minimum and is strictly convex, it is radially unbounded, and the restriction of U to the set $\{y \in \mathbb{R}^M : Ay = b\}$ therefore also has a minimum, and there will be a unique $y \in \mathbb{R}^M$, solving (\mathcal{G}) .

Introducing Lagrange multipliers λ_i , $i \in \{1, \dots, N\}$, and setting the partial derivative of the Lagrangian

$$\mathcal{U}(y, \lambda) \triangleq U(y) + \sum_{j=1}^N \lambda_j \left(b_j - \sum_{i=1}^M A_{ji} y_i \right)$$

to y_i equal to zero results in

$$\frac{\partial \mathcal{U}}{\partial y_i}(y) - \sum_{j=1}^N \lambda_j A_{ji} = 0. \quad (10)$$

The vector function

$$\frac{\partial \mathcal{U}}{\partial y} \triangleq \left(\frac{\partial \mathcal{U}}{\partial y_1}, \dots, \frac{\partial \mathcal{U}}{\partial y_M} \right) : \mathbb{R}^M \rightarrow \mathbb{R}^M,$$

is injective. To see this, consider $y^a, y^b \in \mathbb{R}^M$, with $y^a \neq y^b$. Then the strict convexity of the function

$$v : \mathbb{R} \rightarrow \mathbb{R} : t \mapsto U(y^a + t(y^b - y^a)),$$

implies that $v'(0) \neq v'(1)$, and thus $\frac{\partial \mathcal{U}}{\partial y}(y^a) \neq \frac{\partial \mathcal{U}}{\partial y}(y^b)$.

We also assume that $\frac{\partial \mathcal{U}}{\partial y}$ is surjective. Then for each $\lambda \in \mathbb{R}^N$, equation (10) has a unique solution for y , which defines the function $\tilde{y} : \mathbb{R}^N \rightarrow \mathbb{R}^M : \lambda \mapsto y = \tilde{y}(\lambda)$. Imposing furthermore that U is twice differentiable, and that its Hessian matrix (i.e. the matrix consisting of the elements $\frac{\partial^2 U}{\partial y_i \partial y_j}$) is positive definite everywhere in \mathbb{R}^M , guarantees the existence of the derivatives $\frac{\partial \tilde{y}_i}{\partial \lambda_j}$.

The unique $y \in \mathbb{R}^M$, solving (\mathcal{G}) , will be equal to the unique solution y to the set of conditions $Ay = b$ and $y = \tilde{y}(\lambda)$ for some $\lambda \in \mathbb{R}^N$. With $y = \tilde{y}(\tilde{\lambda}(x^e))$, where x^e is an equilibrium point of (3), the constraints $Ay = b$ are fulfilled, and by definition of the function \tilde{y} , the condition (10) is also fulfilled with $y = \tilde{y}(\tilde{\lambda}(x^e))$ therefore solving the problem (\mathcal{G}) .

For the stability of (3) we impose the conditions (\mathcal{C}) from section 4. We will show that the function $V_{\mathcal{G}} \circ \tilde{\lambda}$, with

$$V_{\mathcal{G}}(\lambda) = -V_{\text{dual}}(\lambda) = -\min_{y \in \mathbb{R}^N} \mathcal{U}(y, \lambda) = -\mathcal{U}(\tilde{y}(\lambda), \lambda), \quad \forall \lambda \in \mathbb{R}^N, \quad (11)$$

satisfies the same properties as $V \circ \tilde{\lambda}$ in section 4. Since

$$\begin{aligned} \frac{\partial V_{\mathcal{G}}}{\partial \lambda_i}(\lambda) &= -\sum_{j=1}^M \frac{\partial \mathcal{U}}{\partial y_j}(\tilde{y}(\lambda), \lambda) \frac{\partial \tilde{y}_j}{\partial \lambda_i}(\lambda) - \frac{\partial \mathcal{U}}{\partial \lambda_i}(\tilde{y}(\lambda), \lambda) \\ &= -\sum_{j=1}^M \left(\frac{\partial \mathcal{U}}{\partial y_j}(\tilde{y}_j(\lambda)) - \sum_{k=1}^N \lambda_k A_{kj} \right) \frac{\partial \tilde{y}_j}{\partial \lambda_i}(\lambda) - \left(b_i - \sum_{j=1}^M A_{ij} \tilde{y}_j(\lambda) \right) \\ &= -b_i + \sum_{j=1}^M A_{ij} \tilde{y}_j(\lambda), \end{aligned}$$

it follows again that, with x a solution of (3),

$$\frac{\partial V_{\mathcal{G}}}{\partial \lambda_i}(\tilde{\lambda}(x(t))) = -\dot{x}(t),$$

and that $V_{\mathcal{G}} \circ \tilde{\lambda} : \mathbb{R}^N \rightarrow \mathbb{R}$ is a Lyapunov function for the system (3).

Proposition 3. *Under the assumption (C3) the function $V_{\mathcal{G}} \circ \tilde{\lambda}$ is radially unbounded.*

The proof is given in the appendix, section A.4. Repeating the arguments in section 4 we can derive the following theorem:

Theorem 3. *Under the assumptions (C), the system (3) has a globally asymp-*

totically stable equilibrium point, of which the image under the mapping $\tilde{y} \circ \tilde{\lambda}$ solves (\mathcal{G}) .

Remark 3. The Lyapunov function V defined in (4) can also be written in terms of the path integral of the vector field of (3) along a straight line from 0 to λ :

$$V(\lambda) = - \int_0^1 dt \lambda^T (-b + A\tilde{y}(\lambda t)), \quad \forall \lambda \in \mathbb{R}^N.$$

With the definition of \tilde{y} from this section, this expression is also equal to $V_{\mathcal{G}}(\lambda)$ (except for a constant term), but it is advantageous to consider (11) rather than the integral expression in order to prove that $V_{\mathcal{G}}$ is radially unbounded.

10 Conclusion

We have presented a general, alternative approach to deal with convex as well as linear optimization problems with linear constraints. We formulate explicit conditions guaranteeing that the optimum is implemented as the globally asymptotically stable equilibrium point of a dynamical system. When the cost is separable, there is a natural Lyapunov function governing the global convergence. Our approach is also applicable to optimization problems with a non-separable cost function, and using the relation with the Lagrange dual function, the equilibrium point implementing the minimal cost can again be shown to be globally asymptotically stable. When implementing linear programming problems, we end up with a dynamical system with discontinuous differential equations.

As an illustration we have considered the minimum cost flow problem, which allows a physical interpretation of the dynamical system. Each flow can be considered as a feedback control, which is a function of the commodity level in the nodes of the network only, and not of the external in- and outflow in the nodes.

When the coefficient matrix of the constraints (in the general optimization problem) is sparse, a distributed implementation is possible. In the distributed implementation of the minimum cost flow problem, the flow through an arc only depends on the commodity levels of its start and end node.

Acknowledgments

This paper presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors.

Filip De Smet is a Research Assistant of the Research Foundation - Flanders (FWO - Vlaanderen).

References

- [1] D. P. Bertsekas. *Network Optimization: Continuous and Discrete Models*. Athena Scientific, 1998.
- [2] D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization*. Number 6 in Athena Scientific Optimization and Computation Series. Athena Scientific, 1997.
- [3] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [4] A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*. Kluwer Academic Publishers, 1988.
- [5] S. I. Gass. *Linear Programming: Methods and Applications*. Dover Publications, 5 edition, 2003.
- [6] S. H. Low and D. E. Lapsley. Optimization flow control, i: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 7(6):861–874, 1999.
- [7] S. H. Low and R. Srikant. A mathematical framework for designing a low-loss, low-delay internet. *Networks and Spatial Economics*, 4:75–101, 2004.

A Appendix

A.1 Proof of proposition 1

Although proposition 1 is a special case of proposition 3, we insist on giving a separate proof which is not based on duality properties.

If $\|x\| \rightarrow \infty$, then $\|\tilde{\lambda}(x)\| \rightarrow \infty$. We will show that $\|\lambda\| \rightarrow \infty$ implies $V(\lambda) \rightarrow \infty$.

Since the function $U_i'^{-1}$ is increasing with $\lim_{\sigma \rightarrow \pm\infty} U_i'^{-1}(\sigma) = \pm\infty$, the function mapping $\mu_i \in \mathbb{R}$ to $\int_0^{\mu_i} U_i'^{-1}(\sigma) d\sigma - B_0 |\mu_i|$, with $B_0 > 0$ arbitrary, grows unbounded for $\mu_i \rightarrow \pm\infty$, for each $i \in \{1, \dots, M\}$. It follows that for any $B > 0$, there exists a $D > 0$ for which

$$\sum_{i=1}^M \int_0^{\mu_i} U_i'^{-1}(\sigma) d\sigma \geq B \|\mu\|, \quad \forall \mu \in \mathbb{R}^M, \text{ with } \|\mu\| \geq D.$$

Since A has full row rank, it follows that

$$\|\lambda^T\| = \left\| \lambda^T A A^T (A A^T)^{-1} \right\| \leq \|\lambda^T A\| \left\| A^T (A A^T)^{-1} \right\|,$$

and thus, for

$$B = (\|b\| + 1) \left\| A^T (A A^T)^{-1} \right\|,$$

there exists a $D > 0$, such that, $\forall \lambda \in \mathbb{R}^N$, with $\|\lambda^T A\| \geq D$,

$$\begin{aligned} V(\lambda) &= - \sum_{i=1}^N b_i \lambda_i + \sum_{j=1}^M \int_0^{\sum_{i=1}^N \lambda_i A_{ij}} U_j'^{-1}(\sigma) d\sigma \\ &\geq - \|b\| \|\lambda\| + (\|b\| + 1) \left\| A^T (A A^T)^{-1} \right\| \|\lambda^T A\| \\ &\geq \|\lambda\|. \end{aligned}$$

Since the condition $\|\lambda^T A\| \geq D$ is satisfied if $\|\lambda\| \geq D \left\| A^T (A A^T)^{-1} \right\|$, it follows

that $V \circ \tilde{\lambda}$ is radially unbounded in \mathbb{R}^N .

A.2 Proof of proposition 2

As shown in section A.1, for any $B > 0$, there exists a $D > 0$ for which

$$\sum_{i=1}^M \int_0^{\mu_i} U_i'^{-1}(\sigma) d\sigma \geq B \|\mu\|, \quad \forall \mu \in \mathbb{R}^M, \text{ with } \|\mu\| \geq D.$$

Since each component of $\lambda^T A$ is equal to the difference between the components of λ associated to the start and end node of an arc, and since each pair of nodes can be connected by a path of at most $N - 1$ arcs, it follows that

$$\|\tilde{\lambda}^T(x)A\| \geq \frac{\Delta\tilde{\lambda}(x)}{N-1}, \quad \forall x \in \mathbb{R}^N.$$

For

$$B = (N-1) \left(\sum_{i=1}^N |b_i| + 1 \right),$$

there exists a $D > 0$, such that, $\forall x \in \mathbb{R}^N$, with $\|\tilde{\lambda}^T(x)A\| \geq D$ (taking into account that $\sum_{i=1}^N b_i = 0$),

$$\begin{aligned} V(\tilde{\lambda}(x)) &= - \sum_{i=1}^N b_i \left(\tilde{\lambda}_i(x) - \tilde{\lambda}_1(x) \right) + \sum_{j=1}^M \int_0^{\sum_{i=1}^N \tilde{\lambda}_i(x) A_{ij}} U_j'^{-1}(\sigma) d\sigma \\ &\geq - \sum_{i=1}^N |b_i| \Delta\tilde{\lambda}(x) + (N-1) \left(\sum_{i=1}^N |b_i| + 1 \right) \|\tilde{\lambda}^T(x)A\| \\ &\geq \Delta\tilde{\lambda}(x). \end{aligned}$$

The condition $\|\lambda^T(x)A\| \geq D$ is satisfied if $\Delta\tilde{\lambda}(x) \geq (N-1)D$.

Since $\|x\| \rightarrow \infty$, with $x \in L_C$, implies $\Delta\tilde{\lambda}(x) \rightarrow \infty$, and thus also $V(\tilde{\lambda}(x)) \rightarrow \infty$, it follows that $V \circ \tilde{\lambda}$ is radially unbounded in L_C for each $C \in \mathbb{R}$.

A.3 Proof of the equivalence of the conditions (C'3) and

(D)

Under the assumption that $\bar{\lambda}'_i(x_i) > 0, \forall x_i \in \mathbb{R}, \forall i \in \{1, \dots, N\}$, we will show that the function $\Delta\tilde{\lambda}$ is radially unbounded in L_C , for all $C \in \mathbb{R}$, if and only if one of the three conditions (D) is satisfied.

- First assume that, for each $C \in \mathbb{R}$, $\Delta\tilde{\lambda}$, is radially unbounded in L_C . Pick $i_1, i_2 \in \{1, \dots, N\}$ arbitrary, with $i_1 \neq i_2$, set $x_i = 0, \forall i \in \{1, \dots, N\} \setminus \{i_1, i_2\}$, and set $x_{i_1} = -x_{i_2}$, such that $x \in L_0$. Consider the limit $x_{i_1} = -x_{i_2} \rightarrow +\infty$. Then $\Delta\tilde{\lambda}(x) \rightarrow \bar{\lambda}_{i_1}(x_{i_1}) - \bar{\lambda}_{i_2}(x_{i_2}) \rightarrow +\infty$, implying that

$$\lim_{x_{i_1} \rightarrow +\infty} \bar{\lambda}_{i_1}(x_{i_1}) = +\infty \quad \text{and/or} \quad \lim_{x_{i_2} \rightarrow -\infty} \bar{\lambda}_{i_2}(x_{i_2}) = -\infty. \quad (12)$$

If the condition (D1) does not hold, then there exists an $i_B \in \{1, \dots, N\}$ for which $\bar{\lambda}_{i_B}(x_{i_B})$ remains bounded as $x_{i_B} \rightarrow +\infty$, and thus, setting $i_1 = i_B$ and $i_2 \in \{1, \dots, N\} \setminus \{i_B\}$ in (12), it follows that

$$\lim_{x_i \rightarrow -\infty} \bar{\lambda}_i(x_i) = -\infty, \quad \forall i \in \{1, \dots, N\} \setminus \{i_B\}. \quad (13)$$

If neither conditions (D1) and (D2) hold, then it follows from (13) that $\bar{\lambda}_{i_B}(x_{i_B})$ also remains bounded as $x_{i_B} \rightarrow -\infty$, implying that

$$\lim_{x_i \rightarrow +\infty} \bar{\lambda}_i(x_i) = +\infty, \quad \forall i \in \{1, \dots, N\} \setminus \{i_B\}, \quad (14)$$

and (13) and (14) then imply (D3).

- Assume that one of the three conditions (D) holds. For each $i, j \in \{1, \dots, N\}$, with $i \neq j$, it follows that if $x_i \rightarrow +\infty$ and $x_j \rightarrow -\infty$, or

$x_i \rightarrow -\infty$ and $x_j \rightarrow +\infty$, then

$$\Delta\tilde{\lambda}(x) \geq |\bar{\lambda}_i(x_i) - \bar{\lambda}_j(x_j)| \rightarrow +\infty.$$

Therefore, for each $Q > 0$, there exists a $D_1 > 0$ such that, if $x_k \geq D_1$ and $x_l \leq -D_1$ for some $k, l \in \{1, \dots, N\}$, with $k \neq l$, then $\Delta\tilde{\lambda}(x) \geq Q$. On the other hand, for any $C \in \mathbb{R}$ there exists a $D_0 > 0$ such that, if $\|x\| \geq D_0$ and $x \in L_C$, there are $k, l \in \{1, \dots, N\}$, with $k \neq l$, satisfying $x_k \geq D_1$ and $x_l \leq -D_1$, and therefore $\Delta\tilde{\lambda}(x) \geq Q$. Consequently, for each $C \in \mathbb{R}$, $\Delta\tilde{\lambda}$ is radially unbounded in L_C .

A.4 Proof of proposition 3

Pick $\lambda \in \mathbb{R}^N \setminus \{0\}$ arbitrary, and set

$$y^* \triangleq A^T (AA^T)^{-1} \left(b + \frac{\lambda}{\|\lambda\|} \right).$$

Then

$$\begin{aligned} V_{\mathcal{G}}(\lambda) &= - \min_{y \in \mathbb{R}^N} (U(y) + \lambda^T (b - Ay)) \\ &\geq -U(y^*) - \lambda^T (b - Ay^*) \\ &= -U(y^*) + \|\lambda\|. \end{aligned}$$

Taking into account that $\|y^*\| \leq \left\| A^T (AA^T)^{-1} \right\| (\|b\| + 1)$ it follows that $U(y^*)$ is upper bounded by a constant $M_U \in \mathbb{R}$ (independent of y^*), and therefore

$$V_{\mathcal{G}}(\lambda) \geq -M_U + \|\lambda\|, \quad \forall \lambda \in \mathbb{R}^N,$$

showing that $V_{\mathcal{G}}$ is radially unbounded. Since $\|x\| \rightarrow \infty$ implies $\|\tilde{\lambda}(x)\| \rightarrow \infty$, it follows that $V_{\mathcal{G}} \circ \tilde{\lambda}$ is also radially unbounded.