Analysis of three distributed continuous-time dynamical systems: application to power networks

Ashish Cherukuri
Advisor: Prof. Jorge Cortés

Seminar Talk
Indian Institute of Technology, Bombay
Jan 8th, 2016

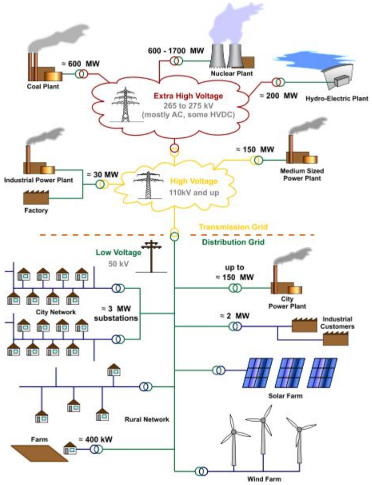
Cymer Center for Control Systems and Dynamics

UC San Diego
Mechanical and Aerospace Engineering
JACOBS SCHOOL OF ENGINEERING

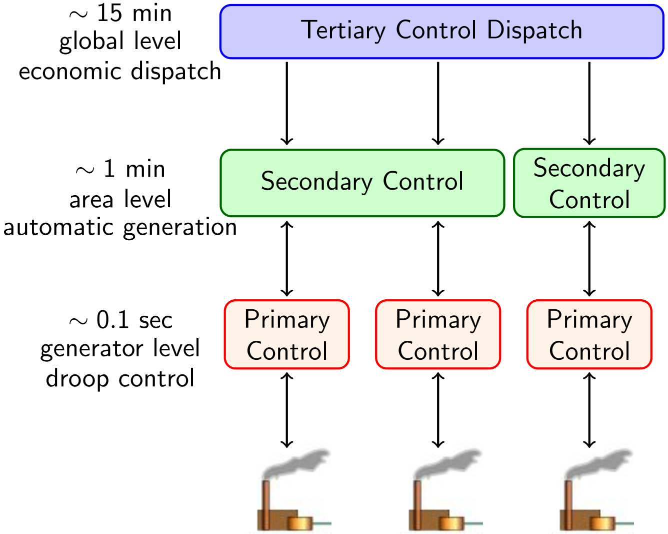
Electrical Power Network

Objectives

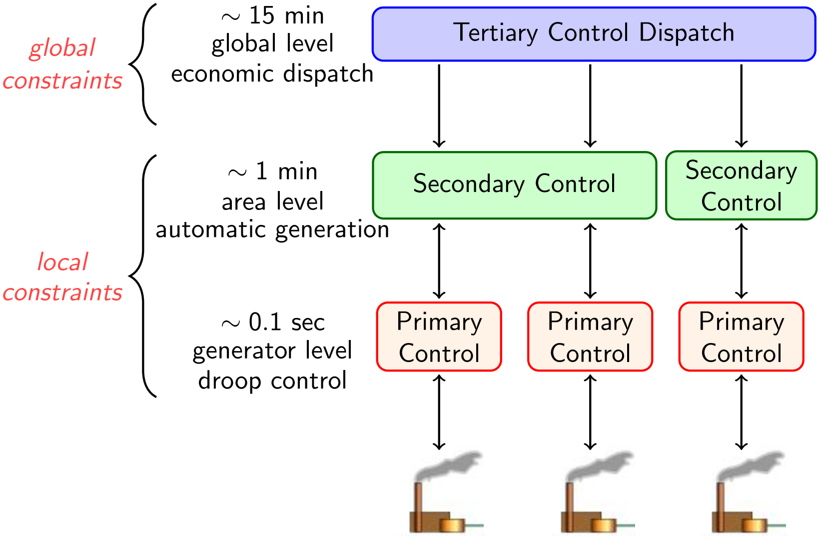
- Balance load and generation
- Restore nominal frequency
  - guarantee cost efficiency
  - satisfy physical constraints
  - ensure security & reliability



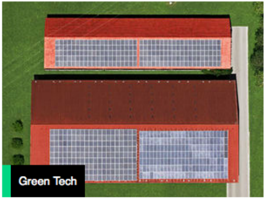
Electrical Power Network



Electrical Power Network



Future Power Grid: vertical to flat



How Rooftop Solar Can Stabilize the Grid

Following Germany's lead, California gives advanced inverters a bigger role in the grid

21 Jan

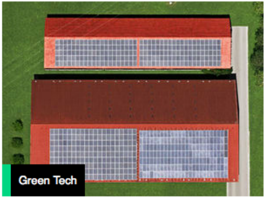
The Rise of the Personal Power Plant

Smart and agile power systems will let every home and business generate, store, and share electricity

28 May 2014

- Increase in Distributed Energy Resources (DERs)
  - wind turbines, solar PV, storages, microgrids etc
- Power generation – decentralized
- Large scale optimization problems

Future Power Grid: vertical to flat



How Rooftop Solar Can Stabilize the Grid

Following Germany's lead, California gives advanced inverters a bigger role in the grid

21 Jan

The Rise of the Personal Power Plant

Smart and agile power systems will let every home and business generate, store, and share electricity

28 May 2014

Distributed solutions

- Robust against failures
- Cater to dynamic demands
- Preserve "privacy"
- Provide plug-and-play

Outline

1. *Economic dispatch problem*
   - Problem statement
   - Relaxed problem and centralized algorithm
   - Robust distributed algorithm

2. *Analysis of Saddle-point dynamics*
   - Convex-Concave Functions
   - General Functions

3. *Analysis of Primal-dual dynamics*

Tertiary Control

Primary/Secondary Control

Problem Statement

Economic Dispatch (ED) Problem

$$\begin{aligned} \min \quad & f(P) = \sum_{i=1}^{n} f_i(P_i) \\ \text{s.t} \quad & \sum_{i=1}^{n} P_i = \mathbf{1}_n^\top P = P_l \\ & P_i^m \le P_i \le P_i^M, \ \forall i \end{aligned}$$

Problem Statement

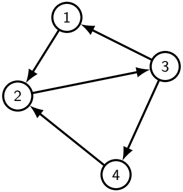
Economic Dispatch (ED) Problem

$$
\begin{aligned}
\min \quad & f(P) & \\
\text{s.t} \quad & \mathbf{1}_n^\top P = P_l & \textit{load condition} \\
& P_i^m \leq P_i \leq P_i^M, \ \forall i & \textit{box constraints}
\end{aligned}
$$

Problem Statement

Economic Dispatch (ED) Problem

$$
\begin{aligned}
\min \quad & f(P) \\
\text{s.t} \quad & \mathbf{1}_n^\top P = P_l & \textit{load condition} \\
& P_i^m \le P_i \le P_i^M, \forall i & \textit{box constraints}
\end{aligned}
$$



Communication network setup

- strongly connected weight-balanced digraphs
- generator i knows f_i and controls P_i
- generator i can send information to its in-neighbors

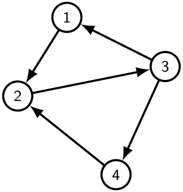
Assumptions: we do not consider

- line losses, transmission constraints
- ramp rates, valve-point effects, prohibited operating zones

Problem Statement

Economic Dispatch (ED) Problem

$$
\begin{aligned}
\min \quad & f(P) & \\
\text{s.t} \quad & \mathbf{1}_n^\top P = P_l & \color{red}{\textit{load condition}} \\
& P_i^m \le P_i \le P_i^M, \ \forall i & \color{red}{\textit{box constraints}}
\end{aligned}
$$



Objective: design distributed algorithm that

- solves the ED problem from any initial condition
- able to handle time-varying loads
- is robust to intermittent power generation

Overview of Literature

- quadratic cost function – consensus based [Zhang et al., 11; Kar&Hug, 12; Dominguez-Garcia et al., 12; Loia&Vacarro, 13; Binetti et al., 14b]
- general cost but no capacity bound [Xiao&Boyd, 06; Johansson, 09; Mudumbai et al., 12]
- regularized problem – suboptimal solution [Simonetto et al., 12]
- initialization or frequency feedback dependent [Pantoja et al., 14; Zhang et al., 14]
- general (nonconvex) problem - no theoretical guarantees
- distributed optimization [Nedich&Ozdaglar, 09; Johansson et al., 09; Wang&Elia, 10; Zhu&Martínez, 12; Gharesifard&Cortés 14]

Relaxed ED Problem: a motivation

Relaxed ED problem

$$\begin{aligned} \min \quad & f(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \end{aligned}$$

- Lagrangian:

  $$L(P,\nu) = f(P) + \nu(\mathbf{1}_n^\top P - P_l)$$

- KKT conditions:

  $$\nabla f(P_*) = -\nu_* \mathbf{1}_n \text{ and } \mathbf{1}_n^\top P_* = P_l$$

Agreement on *gradients* a solution!

Laplacian-gradient dynamics

Relaxed ED problem

$$\begin{aligned} \min \quad & f(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \end{aligned}$$

Laplacian-gradient dynamics

$$\dot{P} = -\mathsf{L}\nabla f(P)$$

Laplacian-gradient dynamics

Relaxed ED problem

$$\begin{aligned} \min \quad & f(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \end{aligned}$$

Laplacian-gradient dynamics

$$\dot{P} = -\mathsf{L}\nabla f(P)$$

- distributed implementation:

  $\dot{P}_i = -\sum_{j \in \mathcal{N}_i} a_{ij}(\nabla f_i(P_i) - \nabla f_j(P_j))$

- load condition conserved:

  $\frac{d}{dt}(\mathbf{1}_n^\top P) = -\mathbf{1}_n^\top \mathsf{L}\nabla f(P) = 0$

- f nonincreasing:

  $\langle \nabla f, \dot{P} \rangle = -\nabla f(P)^\top \mathsf{L}\nabla f(P) \leq 0$

Laplacian-gradient dynamics

Relaxed ED problem

$$\begin{aligned} \min \quad & f(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \end{aligned}$$

Laplacian-gradient dynamics

$$\dot{P} = -\mathsf{L}\nabla f(P)$$

- distributed implementation:

  $\dot{P}_i = -\sum_{j \in \mathcal{N}_i} a_{ij}(\nabla f_i(P_i) - \nabla f_j(P_j))$

- load condition conserved:

  $\frac{d}{dt}(\mathbf{1}_n^\top P) = -\mathbf{1}_n^\top \mathsf{L}\nabla f(P) = 0$

- f nonincreasing:

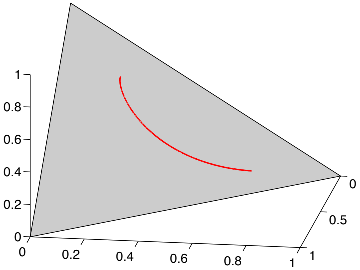
  $\langle \nabla f, \dot{P} \rangle = -\nabla f(P)^\top \mathsf{L} \nabla f(P) \le 0$

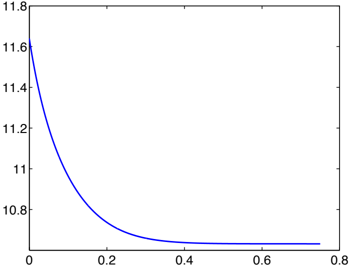
Theorem (Convergence of *Laplacian-gradient* dynamics)

The feasibility set is positively invariant and trajectories starting from a feasible point converge to the set of solutions of the relaxed ED problem

Laplacian-gradient dynamics: example

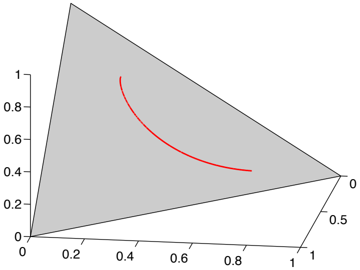
Anytime nature of dynamics

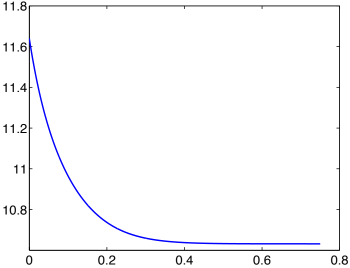




Laplacian-gradient dynamics: example

Anytime nature of dynamics





- How to incorporate box constraints? – *Exact penalty functions*
- How to make it initialization-free? – *Dynamic average consensus*

Reformulation using Exact Penalty Functions

ED Problem

$$\begin{aligned} \min \quad & f(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \\ & P_i^m \le P_i \le P_i^M, \forall i \end{aligned}$$

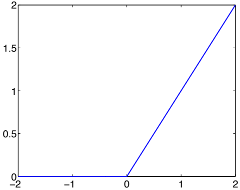
Modified ED Problem

$$\begin{aligned} \min \quad & f^\epsilon(P) = \sum_{i=1}^{n} f_i^\epsilon(P_i) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \end{aligned}$$

$$f_i^\epsilon(P_i) = f_i(P_i) + \tfrac{1}{\epsilon}([P_i - P_i^M]^+ + [P_i^m - P_i]^+)$$

where

$$[u]^+ = \begin{cases} 0 & \text{if } u \le 0 \\ u & \text{if } u > 0 \end{cases}$$



$[u]^+$

Reformulation using Exact Penalty Functions

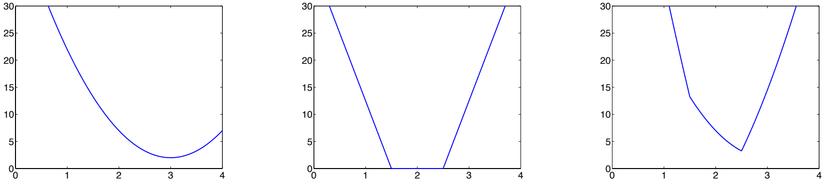
ED Problem

$$\begin{aligned} \min \quad & f(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \\ & P_i^m \leq P_i \leq P_i^M, \forall i \end{aligned}$$

Modified ED Problem

$$\begin{aligned} \min \quad & f^\epsilon(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \end{aligned}$$

$$f_i^\epsilon(P_i) = f_i(P_i) + \tfrac{1}{\epsilon}([P_i - P_i^M]^+ + [P_i^m - P_i]^+)$$



f_i $\quad + \quad$ Penalty function $\quad = \quad$ f_i^ϵ

Reformulation using Exact Penalty Functions

ED Problem

$$\begin{aligned} \min \quad & f(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \\ & P_i^m \leq P_i \leq P_i^M, \forall i \end{aligned}$$

Modified ED Problem

$$\begin{aligned} \min \quad & f^\epsilon(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \end{aligned}$$

$$f_i^\epsilon(P_i) = f_i(P_i) + \tfrac{1}{\epsilon}([P_i - P_i^M]^+ + [P_i^m - P_i]^+)$$

Proposition (Equivalence between optimizations)

The solutions of above problems coincide for $\epsilon \in \mathbb{R}_{>0}$ such that

$$\epsilon < \frac{1}{2 \max_{P \in \mathcal{F}_{\mathrm{ED}}} \|\nabla f(P)\|_\infty}$$

Reformulation using Exact Penalty Functions

ED Problem

$$
\begin{aligned}
\min \quad & f(P) \\
\text{s.t} \quad & \mathbf{1}_n^\top P = P_l \\
& P_i^m \leq P_i \leq P_i^M, \ \forall i
\end{aligned}
$$

Modified ED Problem

$$
\begin{aligned}
\min \quad & f^\epsilon(P) \\
\text{s.t} \quad & \mathbf{1}_n^\top P = P_l
\end{aligned}
$$

$$
f_i^\epsilon(P_i) = f_i(P_i) + \tfrac{1}{\epsilon}([P_i - P_i^M]^+ + [P_i^m - P_i]^+)
$$

$$
\partial f_i^\epsilon(P_i) = \begin{cases}
\{\nabla f_i(P_i) - \frac{1}{\epsilon}\} & P_i < P_i^m, \\
[\nabla f_i(P_i) - \frac{1}{\epsilon}, \nabla f_i(P_i)] & P_i = P_i^m, \\
\{\nabla f_i(P_i)\} & P_i^m < P_i < P_i^M, \\
[\nabla f_i(P_i), \nabla f_i(P_i) + \frac{1}{\epsilon}] & P_i = P_i^M, \\
\{\nabla f_i(P_i) + \frac{1}{\epsilon}\} & P_i > P_i^M.
\end{cases}
$$

Reformulation using Exact Penalty Functions

ED Problem

$$\begin{aligned} \min \quad & f(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \\ & P_i^m \leq P_i \leq P_i^M, \forall i \end{aligned}$$

Modified ED Problem

$$\begin{aligned} \min \quad & f^\epsilon(P) \\ \text{s.t} \quad & \mathbf{1}_n^\top P = P_l \end{aligned}$$

$$f_i^\epsilon(P_i) = f_i(P_i) + \tfrac{1}{\epsilon}([P_i - P_i^M]^+ + [P_i^m - P_i]^+)$$

$$-\nu_* \mathbf{1}_n \in \partial f^\epsilon(P_*) \quad \text{and} \quad \mathbf{1}_n^\top P_* = P_l$$

Laplacian-nonsmooth-gradient dynamics

Relaxed ED problem

$$\min \quad f(P)$$
$$\text{s.t} \quad \mathbf{1}_n^\top P = P_l$$

Laplacian-gradient dynamics

$$\dot{P} = -\mathsf{L}\nabla f(P)$$

Modified ED Problem

$$\min \quad f^\epsilon(P)$$
$$\text{s.t} \quad \mathbf{1}_n^\top P = P_l$$

Laplacian-nonsmooth-gradient dynamics

$$\dot{P} \in -\mathsf{L}\partial f^\epsilon(P)$$

where $\partial f^\epsilon(P) = \partial f_1^\epsilon(P_1) \times \cdots \times \partial f_n^\epsilon(P_n)$

Laplacian-nonsmooth-gradient dynamics: analysis

$$\dot{P} \in -\mathsf{L}\partial f^{\epsilon}(P)$$

Theorem (Convergence of $\mathsf{L}\partial$ dynamics)

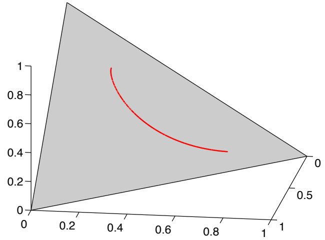
The feasibility set $\{P \in \mathbb{R}^n \mid \mathbf{1}_n^\top P = P_l \text{ and } P_i^m \leq P_i \leq P_i^M, \forall i\}$ is strongly positively invariant under the $\mathsf{L}\partial$ dynamics. Starting from a feasible point the trajectories converge to the solutions of the ED problem.

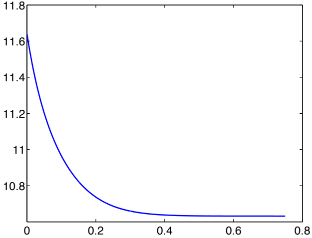
- f^{ϵ} is monotonically nonincreasing – *Anytime nature!*

[A. Cherukuri & S. Martínez & J. Cortés, ACC 2014]

[A. Cherukuri & J. Cortés, TCNS 2015]

How to handle initialization?





- ~~How to incorporate box constraints? – *Exact penalty functions*~~
- How to make it initialization-free? – *Dynamic average consensus*
  - *Laplacian-nonsmooth-gradient + dac* dynamics

Centralized Global (Asymptotic) Solution

`Laplacian-nonsmooth-gradient + lm` dynamics

$$\dot{P} \in -\mathsf{L}\partial f^{\epsilon}(P) + \frac{1}{n}(P_l - \mathbf{1}_n^\top P)\mathbf{1}_n$$

- Mismatch between load and total generation decreases exponentially

$$\frac{d}{dt}(P_l - \mathbf{1}_n^\top P) = -(P_l - \mathbf{1}_n^\top P)$$

- On load satisfaction, it reduces to `Laplacian-nonsmooth-gradient`

Theorem (Convergence of $\mathsf{L}\partial$+`lm` dynamics)

Trajectory of $\mathsf{L}\partial$+`lm` dynamics starting from any point in $\mathbb{R}^n$ converge to the solutions of the ED problem

Technical Analysis of $\mathsf{L}\partial+\mathtt{lm}$ dynamics

Using *refined LaSalle* invariance principle for *differential inclusions*

Theorem (refined LaSalle , Arsie & Ebenbauer (2010))

For $f : \mathbb{R}^n \to \mathbb{R}^n$ locally Lipschitz, $\mathcal{S} \subset \mathbb{R}^n$ closed embedded submanifold of $\mathbb{R}^n$, let $t \mapsto \varphi(t)$ be bounded solution of $\dot{x} = f(x)$ with omega-limit set $\Omega(\varphi)$. If

- $\Omega(\varphi) \subset \mathcal{S}$
- *$W : \mathcal{O} \to \mathbb{R}$ continuously differentiable on open neighborhood $\mathcal{O}$ of $\mathcal{S}$ such that $\mathcal{L}_f W \leq 0$ on $\mathcal{S}$*
- *$\mathcal{E} = \{x \in \mathcal{S} \mid 0 = \mathcal{L}_f W(x)\}$ belongs to a level set of W*

then $\Omega(\varphi) \subset \mathcal{E}$

Two LaSalle functions for $\mathsf{L}\partial+\mathtt{lm}$ dynamics

- $V_1(P) = (P_l - \mathbf{1}_n^\top P)^2$
- $V_2(P) = f^\epsilon(P)$

How to make L∂+lm distributed?

Laplacian-nonsmooth-gradient dynamics

$$\dot{P}_i \in \sum_{j \in \mathcal{N}_i} a_{ij}(\partial f_j^{\epsilon}(P_j) - \partial f_i^{\epsilon}(P_i)) + \nu_1 z_i$$

Generation levels $P_1, \dots, P_n$ ⇓

⇑ Load mismatch estimate $z_1, z_2, \dots, z_n$

dynamic average consensus (dac)

$$\dot{z}_i = -\alpha z_i + \beta \sum_{j \in \mathcal{N}_i} (z_j - z_i) - v_i + \nu_2 (P_l e_{ri} - P_i)$$

$$\dot{v}_i = \alpha\beta \sum_{j \in \mathcal{N}_i} (z_i - z_j)$$

- Each unit i has estimator $z_i \in \mathbb{R}$ tracking average signal $t \mapsto \frac{1}{n}(P_l - \mathbf{1}_n^\top P(t))$

Interconnected systems

- *bottom* component estimates evolving load mismatch given generation
- *top* component adjusts generation levels based on optimization of objective & estimate of load mismatch

Load Mismatch along $\mathsf{L}\partial+\boldsymbol{\mathtt{dac}}$ *dynamics*

Let $x_1 = \mathbf{1}_n^\top P - P_l$ be the mismatch, $x_2 = \dot{x}_1$

Because of dynamic average consensus we get

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\nu_1\nu_2 & -\alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Second-order *exponentially* stable linear system – hence *ISS*

Correctness Guarantees

Theorem (Convergence of $\mathsf{L}\partial+\mathtt{dac}$ dynamics)

For $\alpha, \beta, \nu_1, \nu_2 > 0$ with

$$\frac{\nu_1}{\beta \nu_2 \lambda_2(\mathsf{L}+\mathsf{L}^\top)} + \frac{\nu_2^2 \lambda_{\max}(\mathsf{L}^\top \mathsf{L})}{2\alpha} < \lambda_2(\mathsf{L}+\mathsf{L}^\top)$$

trajectories of $\mathsf{L}\partial+\mathtt{dac}$ dynamics starting with $\mathbf{1}_n^\top v = 0$ converge to $\{(P, z, v) \mid P \text{ solution of ED problem}, z = 0, v = \nu_2(P_l e_r - P)\}$

[A. Cherukuri & J. Cortés, Allerton 2014]

[A. Cherukuri & J. Cortés, Automatica, submitted 2014]

Technical and Robustness Analysis

Proof via *refined LaSalle* Invariance Principle for *differential inclusions*

$$V_1(P, z, v) = \nu_1\nu_2(P_l - \mathbf{1}_n^\top P)^2 + \nu_1^2(\mathbf{1}_n^\top z)^2$$
$$V_2(P, z, v) = f^\epsilon(P) + \frac{1}{2}\Big(\nu_1\nu_2\|z\|^2 + \|v + \alpha z - \nu_2(P_l e_r - P)\|^2\Big)$$

Technical and Robustness Analysis

Proof via *refined LaSalle* Invariance Principle for *differential inclusions*

$$V_1(P, z, v) = \nu_1\nu_2(P_l - \mathbf{1}_n^\top P)^2 + \nu_1^2(\mathbf{1}_n^\top z)^2$$
$$V_2(P, z, v) = f^\epsilon(P) + \frac{1}{2}\Big(\nu_1\nu_2\|z\|^2 + \|v + \alpha z - \nu_2(P_l e_r - P)\|^2\Big)$$

Performance guarantees (L∂+dac dynamics)

- global convergence
- load mismatch dynamics is ISS
- dynamic loads tracked with ultimate bound
- robust to intermittent generation

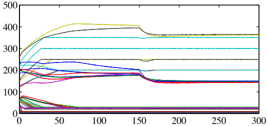
Illustration of Algorithm Performance

IEEE 118 bus example with *54 generators*

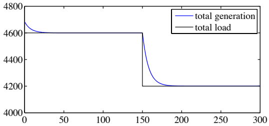
Quadratic cost: $f_i(P_i) = a_i + b_iP_i + c_iP_i^2$

$a_i \in [6.88, 74.33]$, $b_i \in [8.3391, 37.6968]$, and $c_i \in [0.0024, 0.0697]$

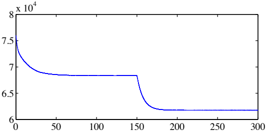
Communication topology is ring digraph with few additional edges



Power allocation



Load mismatch



Total cost

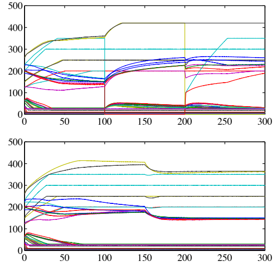
Illustration of Algorithm Performance

IEEE 118 bus example with *54 generators*

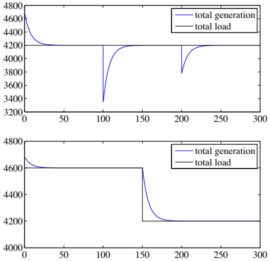
Quadratic cost: $f_i(P_i) = a_i + b_i P_i + c_i P_i^2$

$a_i \in [6.88, 74.33]$, $b_i \in [8.3391, 37.6968]$, and $c_i \in [0.0024, 0.0697]$

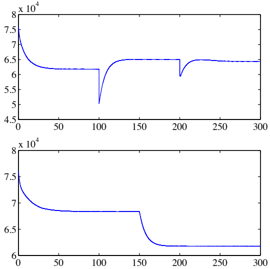
Communication topology is ring digraph with few additional edges



Power allocation



Load mismatch



Total cost

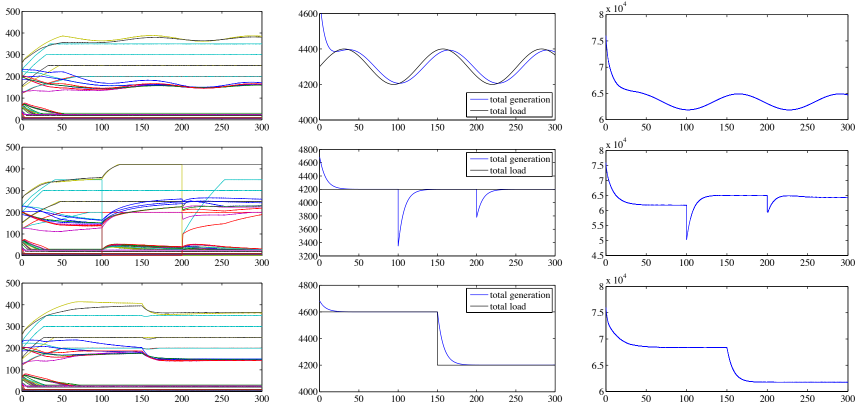
Illustration of Algorithm Performance

IEEE 118 bus example with *54 generators*

Quadratic cost: $f_i(P_i) = a_i + b_i P_i + c_i P_i^2$

$a_i \in [6.88, 74.33]$, $b_i \in [8.3391, 37.6968]$, and $c_i \in [0.0024, 0.0697]$

Communication topology is ring digraph with few additional edges



Power allocation | Load mismatch | Total cost

Summary

Conclusions

- distributed algorithm for *global* constraint problem
- exact penalty functions, `dac`, refined LaSalle
- *switching* communication topologies possible
- robustness to *intermittent* generation

Future work

- Stochastic dispatch
  - load, costs, min-(max-)capacities are *random variables*
  - robust or stochastic optimization
- Learning in electricity markets
  - generators are *strategic*
  - *selfish learning* by repeated play

Outline

Basic question

Gradient descent

Let $f : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^1$ & convex

$$\dot{x} = -\nabla f(x)$$

bdd trajectories converge to minimizers

Gradient ascent

Let $f : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^1$ & concave

$$\dot{x} = \nabla f(x)$$

bdd trajectories converge to maximizers

Basic question

Gradient descent

Let $f : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^1$ & convex

$$\dot{x} = -\nabla f(x)$$

bdd trajectories converge to minimizers

Gradient ascent

Let $f : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^1$ & concave

$$\dot{x} = \nabla f(x)$$

bdd trajectories converge to maximizers

Gradient descent + Gradient ascent

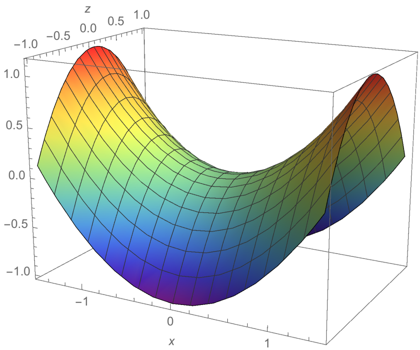
Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be $\mathcal{C}^1$ & convex-concave
(for any $(\bar{x}, \bar{z})$, $x \mapsto F(x, \bar{z})$ is convex & $z \mapsto F(\bar{x}, z)$ concave)

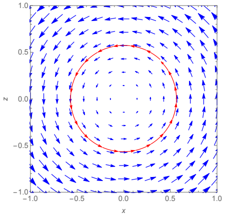
$$\dot{x} = -\nabla_x F(x, z)$$
$$\dot{z} = \nabla_z F(x, z)$$

Do bdd trajectories converge to (min-max) saddle points?

Saddle point: $F(x_*, z) \leq F(x_*, z_*) \leq F(x, z_*)$ for all $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^m$

A picture is worth a thousand words





$F(x,z) = xz$ & $(0,0)$ is a saddle pt.

$$\dot{x} = -z$$
$$\dot{z} = x$$

Motivation

Distributed convex optimization

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) = 0 \end{aligned}$$

- aggregate cost: $f(x) = \sum_{i=1}^{n} f_i(x_i)$
- local constraints: g_i only depends on x_i and $\{x_j\}_{j \in N(i)}$

- Lagrangian: $L(x, \lambda) = f(x) + \lambda^\top g(x)$, convex-concave in (x, λ)
- Primal-dual optimizers $\Leftrightarrow$ saddle points of L
- "gradient descent + gradient ascent" on L is distributed!

Convergence to saddle points of L?

Problem statement

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be $\mathcal{C}^1$, write saddle-point dynamics X_{sp},

$$\begin{aligned}\dot{x} &= -\nabla_x F(x, z)\\ \dot{z} &= \nabla_z F(x, z)\end{aligned}$$

When do trajectories of X_{sp} *converge to* Saddle$(F) \subset \mathbb{R}^n \times \mathbb{R}^m$?

What is already there

- Arrow & Hurwitz & Uzawa (1959): F convex-concave & strict in either
- Wang & Elia (2011): Lagrangian strictly convex in primal
- Fiejer & Paganini (2010): Projection in z-dynamics
- Ratliff & Burden & Sastry (2013): (Pos., Neg.) definite Hessian at NE

Problem statement

Let $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be $\mathcal{C}^1$, write saddle-point dynamics X_{sp},

$$\begin{aligned}\dot{x} &= -\nabla_x F(x, z)\\ \dot{z} &= \nabla_z F(x, z)\end{aligned}$$

When do trajectories of X_{sp} *converge to* Saddle$(F) \subset \mathbb{R}^n \times \mathbb{R}^m$?

Our focus

1. beyond strict convexity-concavity
2. beyond convexity-concavity
3. local vs global convergence
4. continuum of saddle points + convergence to a point
5. complementary conditions

Convexity-concavity-based convergence

Proposition (Local asymptotic stability via strict convexity-concavity)

If F is locally strictly convex-concave on $\mathrm{Saddle}(F)$ *then,* $\mathrm{Saddle}(F)$ *is locally asymptotically stable under X_{sp} and convergence is to a point.*

Proof sketch:

- LaSalle function: $V(x,z) = \frac{1}{2}(\|x - x_*\|^2 + \|z - z_*\|^2)$
- Lie derivative:

$$\begin{aligned} \mathcal{L}_{X_{\mathrm{sp}}} V(x,z) &= -(x - x_*)^\top \nabla_x F(x,z) + (z - z_*)^\top \nabla_z F(x,z) \\ &\leq 0 \end{aligned}$$

- *Stable equilibrium $\implies$ convergence to a point*

Convexity-concavity-based convergence

Proposition (Local asymptotic stability via strict convexity-concavity)

If F is locally strictly convex-concave on $\text{Saddle}(F)$ *then,* $\text{Saddle}(F)$ *is locally asymptotically stable under X_{sp} and convergence is to a point.*

Proposition (Local asymptotic stability via convexity-linearity)

If F is locally convex-concave on $\text{Saddle}(F)$*, linear in z, and*

- *for each* $(x_*, z_*) \in \text{Saddle}(F)$*, there exists a neighborhood* $\mathcal{U}_{x_*} \subset \mathbb{R}^n$ *of* x_* *where, if* $F(x, z_*) = F(x_*, z_*)$ *with* $x \in \mathcal{U}_{x_*}$*, then* $(x, z_*) \in \text{Saddle}(F)$*,*

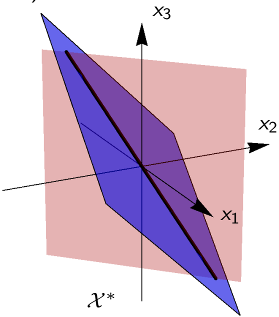
then $\text{Saddle}(F)$ *is locally asymptotically stable under X_{sp} and convergence is to a point.*

Convexity-linearity: example

Constrained optimization on $\mathbb{R}^3$

$$\begin{aligned} \text{minimize} \quad & (x_1 + x_2 + x_3)^2 \\ \text{subject to} \quad & x_1 = x_2 \end{aligned}$$

- Optimizers: $\mathcal{X}^* = \{x \in \mathbb{R}^3 \mid 2x_1 + x_3 = 0, x_2 = x_1\}$
- Lagrangian: $L(x, z) = (x_1 + x_2 + x_3)^2 + z(x_1 - x_2)$
- $\text{Saddle}(L) = \mathcal{X}^* \times \{0\}$



Convexity-linearity: example

Constrained optimization on $\mathbb{R}^3$

$$\begin{aligned} \text{minimize} \quad & (x_1 + x_2 + x_3)^2 \\ \text{subject to} \quad & x_1 = x_2 \end{aligned}$$

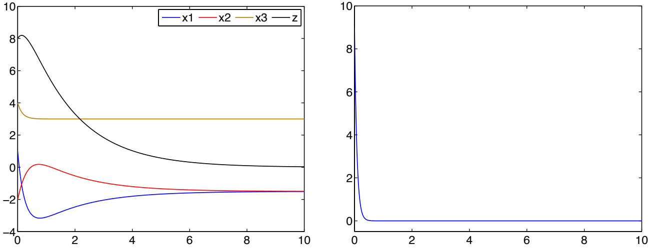
- Optimizers: $\mathcal{X}^* = \{x \in \mathbb{R}^3 \mid 2x_1 + x_3 = 0, x_2 = x_1\}$
- Lagrangian: $L(x, z) = (x_1 + x_2 + x_3)^2 + z(x_1 - x_2)$
- $\text{Saddle}(L) = \mathcal{X}^* \times \{0\}$
- Augmented Lagrangian: $\tilde{L}(x, z) = L(x, z) + (x_1 - x_2)^2$
- $\tilde{L}$ globally convex-concave, linear in z, and meets the third criteria
- $\tilde{L}$ is *NOT* strictly convex-concave

Convexity-linearity: example

Constrained optimization on $\mathbb{R}^3$

$$\begin{aligned} \text{minimize} \quad & (x_1 + x_2 + x_3)^2 \\ \text{subject to} \quad & x_1 = x_2 \end{aligned}$$

- X_{sp} for Augmented Lagrangian $\tilde{L}(x, z) = L(x, z) + (x_1 - x_2)^2$



(x, z) $\qquad$ $\tilde{L}$

Linearization-based convergence

Proposition (Local asymptotic stability via linearization)

For F being $\mathcal{C}^3$, let $\mathrm{Saddle}(F)$ be a p-dimensional manifold. Assume that DX_{sp} at each point in $\mathrm{Saddle}(F)$ has no eigenvalues in the imaginary axis other than 0, which is semisimple with multiplicity p. Then, $\mathrm{Saddle}(F)$ is locally asymptotically stable under X_{sp} and convergence is to a point.

Linearization-based convergence

Proposition (Local asymptotic stability via linearization)

For F being $\mathcal{C}^3$, let $\mathrm{Saddle}(F)$ be a p-dimensional manifold. Assume that DX_{sp} at each point in $\mathrm{Saddle}(F)$ has no eigenvalues in the imaginary axis other than 0, which is semisimple with multiplicity p. Then, $\mathrm{Saddle}(F)$ is locally asymptotically stable under X_{sp} and convergence is to a point.

Proof sketch:

$$DX_{\mathrm{sp}} = \begin{bmatrix} -\nabla_{xx} F & -\nabla_{xz} F \\ \nabla_{zx} F & \nabla_{zz} F \end{bmatrix}_{(x_*, z_*)}$$

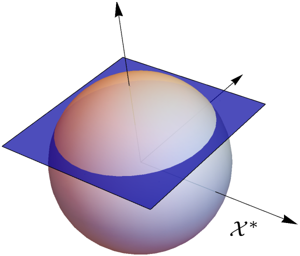
- Saddle point property $\Rightarrow DX_{\mathrm{sp}} + DX_{\mathrm{sp}}^\top \preceq 0$
- $\mathrm{Re}(\lambda_i(DX_{\mathrm{sp}})) \leq \lambda_{\max}(\frac{1}{2}(DX_{\mathrm{sp}} + DX_{\mathrm{sp}}^\top)) \leq 0$
- Now apply center manifold theory

Linearization: example

Constrained optimization on $\mathbb{R}^3$

$$\begin{aligned} \text{minimize} \quad & (\|x\| - 1)^2 \\ \text{subject to} \quad & x_3 = 0.5 \end{aligned}$$

- Optimizers: $\mathcal{X}^* = \{x \in \mathbb{R}^3 \mid x_3 = 0.5, x_1^2 + x_2^2 = 0.75\}$
- Lagrangian: $L(x, z) = (\|x\| - 1)^2 + z(x_3 - 0.5)$
- $\text{Saddle}(L) = \mathcal{X}^* \times \{0\}$
- The Jacobian of X_{sp} satisfies the hypotheses

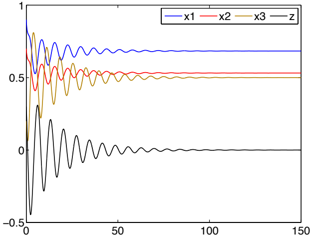


Linearization: example

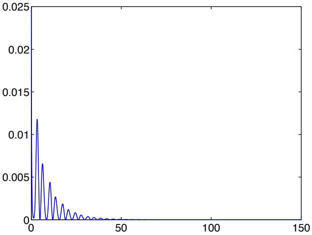
Constrained optimization on $\mathbb{R}^3$

$$\begin{array}{rl} \text{minimize} & (\|x\| - 1)^2 \\ \text{subject to} & x_3 = 0.5 \end{array}$$

- X_{sp} for Lagrangian $L(x, z) = (\|x\| - 1)^2 + z(x_3 - 0.5)$



(x, z)



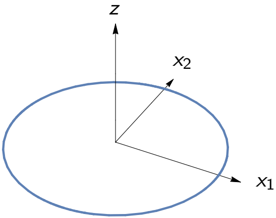
L

Yet more to explore ...

Consider $F : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$,

$$F(x,z) = (\|x\| - 1)^4 - z^2\|x\|^2$$

- $\text{Saddle}(F) = \{(x,z) \mid \|x\| = 1,\ z = 0\}$ *1-d manifold*
- Jacobian of X_{sp} has 0 eigenvalue with *multiplicty 2*



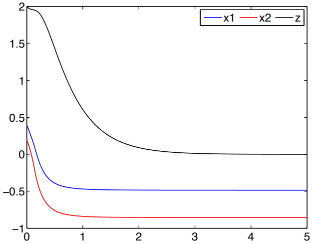
Saddle(F)

Yet more to explore . . .

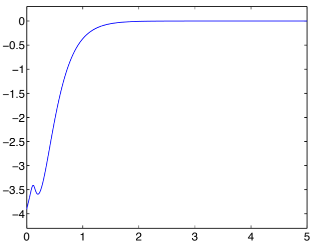
Consider $F : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$,

$$F(x, z) = (\|x\| - 1)^4 - z^2\|x\|^2$$

- $\text{Saddle}(F) = \{(x, z) \mid \|x\| = 1,\ z = 0\}$ *1-d manifold*
- Jacobian of X_{sp} has 0 eigenvalue with *multiplicty 2*



(x, z)



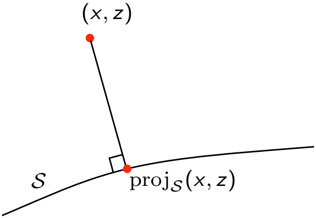
F

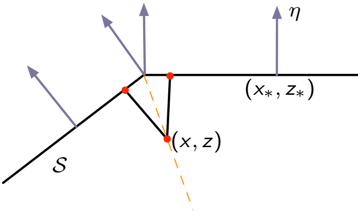
Proximal calculus

V might not be decreasing but $d_{\mathcal{S}}$ is!

$$d_{\mathcal{S}}(x,z) = \min_{(x_*,z_*)\in\mathcal{S}} \|(x,z) - (x_*,z_*)\|$$

$$\mathrm{proj}_{\mathcal{S}}(x,z) = \{(x_*,z_*) \in \mathcal{S} \mid \|(x,z) - (x_*,z_*)\| = d_{\mathcal{S}}(x,z)\}$$



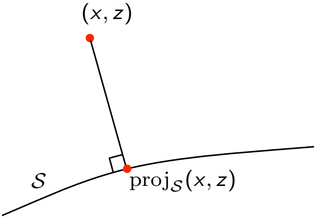


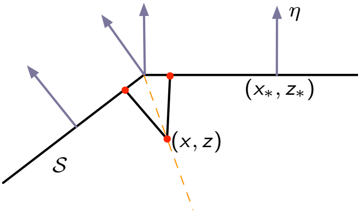
Proximal calculus

V might not be decreasing but $d_{\mathcal{S}}$ is!

$$d_{\mathcal{S}}(x,z) = \min_{(x_*,z_*)\in\mathcal{S}} \|(x,z)-(x_*,z_*)\|$$

$$\mathrm{proj}_{\mathcal{S}}(x,z) = \{(x_*,z_*)\in\mathcal{S} \mid \|(x,z)-(x_*,z_*)\| = d_{\mathcal{S}}(x,z)\}$$





$d_{\mathcal{S}}$ is locally Lipschitz and regular

$$\partial d^2_{\mathcal{S}}(x,z) = \mathrm{co}\{2(x-x_*; z-z_*) \mid (x_*,z_*)\in \mathrm{proj}_{\mathcal{S}}(x,z)\}$$

Does convexity-concavity along proximal normal to Saddle(F) *help?*

Proximal normal-based convergence

Proposition (Asymptotic stability via proximal normals)

For F being $\mathcal{C}^2$, assume that for every (x_*, z_*) and every proximal normal $\eta = (\eta_x, \eta_z)$ at (x_*, z_*) with $\|\eta\| = 1$, it holds that $\lambda \mapsto F(x_* + \lambda\eta_x, z_*)$ is convex and $\lambda \mapsto F(x_*, z_* + \lambda\eta_z)$ is concave with

$$F(x_* + \lambda\eta_x, z_*) - F(x_*, z_*) \geq k_1\|\lambda\eta_x\|^{\alpha_1}$$
$$F(x_*, z_* + \lambda\eta_z) - F(x_*, z_*) \leq -k_2\|\lambda\eta_z\|^{\beta_1}$$

Proximal normal-based convergence

Proposition (Asymptotic stability via proximal normals)

For F being $\mathcal{C}^2$, assume that for every (x_*, z_*) and every proximal normal $\eta = (\eta_x, \eta_z)$ at (x_*, z_*) with $\|\eta\| = 1$, it holds that $\lambda \mapsto F(x_* + \lambda\eta_x, z_*)$ is convex and $\lambda \mapsto F(x_*, z_* + \lambda\eta_z)$ is concave with

$$
\begin{aligned}
F(x_* + \lambda\eta_x, z_*) - F(x_*, z_*) &\geq k_1 \|\lambda\eta_x\|^{\alpha_1} \\
F(x_*, z_* + \lambda\eta_z) - F(x_*, z_*) &\leq -k_2 \|\lambda\eta_z\|^{\beta_1}
\end{aligned}
$$

and, for all $t \in [0, 1]$,

$$
\begin{aligned}
&\|\nabla_{xz} F(x_* + t\lambda\eta_x, z_* + \lambda\eta_z) - \nabla_{xz} F(x_* + \lambda\eta_x, z_* + t\lambda\eta_z)\| \\
&\qquad\qquad\qquad\qquad \leq L_x \|\lambda\eta_x\|^{\alpha_2} + L_z \|\lambda\eta_z\|^{\beta_2}
\end{aligned}
$$

Proximal normal-based convergence

Proposition (Asymptotic stability via proximal normals)

For F being $\mathcal{C}^2$, assume that for every (x_*, z_*) and every proximal normal $\eta = (\eta_x, \eta_z)$ at (x_*, z_*) with $\|\eta\| = 1$, it holds that $\lambda \mapsto F(x_* + \lambda\eta_x, z_*)$ is convex and $\lambda \mapsto F(x_*, z_* + \lambda\eta_z)$ is concave with

$$F(x_* + \lambda\eta_x, z_*) - F(x_*, z_*) \geq k_1 \|\lambda\eta_x\|^{\alpha_1}$$
$$F(x_*, z_* + \lambda\eta_z) - F(x_*, z_*) \leq -k_2 \|\lambda\eta_z\|^{\beta_1}$$

and, for all $t \in [0, 1]$,

$$\|\nabla_{xz} F(x_* + t\lambda\eta_x, z_* + \lambda\eta_z) - \nabla_{xz} F(x_* + \lambda\eta_x, z_* + t\lambda\eta_z)\| \\ \leq L_x \|\lambda\eta_x\|^{\alpha_2} + L_z \|\lambda\eta_z\|^{\beta_2}$$

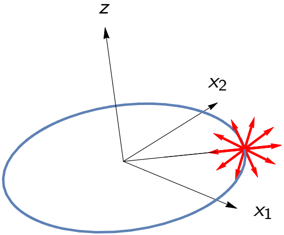
Then, Saddle(F) *is locally asymptotically stable under X_{sp} if*

(either $L_x = 0$ or $\alpha_1 \leq \alpha_2 + 1$) AND (either $L_z = 0$ or $\beta_1 \leq \beta_2 + 1$).

Proximal normal: example

$$F(x,z) = (\|x\| - 1)^4 - z^2\|x\|^2$$

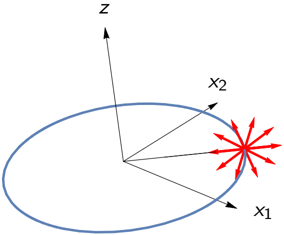
- $\mathrm{Saddle}(F) = \{(x,z) \mid \|x\| = 1,\ z = 0\}$
- $(x_*, z_*) = (\cos\theta, \sin\theta, 0)$, where $\theta \in [0, 2\pi)$
- $\eta = (\eta_x, \eta_z) = ((a_1\cos\theta, a_1\sin\theta), a_2), \quad a_1^2 + a_2^2 = 1$



Proximal normal: example

$$F(x,z) = (\|x\| - 1)^4 - z^2\|x\|^2$$

- $\text{Saddle}(F) = \{(x,z) \mid \|x\| = 1,\ z = 0\}$
- $(x_*, z_*) = (\cos\theta, \sin\theta, 0)$, where $\theta \in [0, 2\pi)$
- $\eta = (\eta_x, \eta_z) = ((a_1\cos\theta, a_1\sin\theta), a_2), \quad a_1^2 + a_2^2 = 1$



- $\lambda \mapsto F(x_* + \lambda\eta_x, z_*) = (\lambda a_1)^4$ is convex with $\alpha_1 = 4$
- $\lambda \mapsto F(x_*, z_* + \lambda\eta_z) = -(\lambda a_2)^2$ is concave with $\beta_1 = 2$
- $L_x = 0$, $L_z \neq 0$ and $\beta_2 = 1$

Summary

The story doesn't end here but the time does!

[Cherukuri & Gharesifard & Cortés, SICON, submitted 2015]

Conclusions

- convexity-concavity
- convexity-linearity
- linearization
- proximal normal

$$\boxed{V(x,z) = \frac{1}{2}(\|x - x_*\|^2 + \|z - z_*\|^2)}$$

$$\boxed{d_{\mathcal{S}}^2(x,z) = \min_{(x_*,z_*)\in\mathcal{S}}(\|x - x_*\|^2 + \|z - z_*\|^2)}$$

Summary

The story doesn't end here but the time does!

[Cherukuri & Gharesifard & Cortés, SICON, submitted 2015]

Conclusions

- convexity-concavity
- convexity-linearity
- linearization
- proximal normal

$$\boxed{V(x,z) = \frac{1}{2}(\|x - x_*\|^2 + \|z - z_*\|^2)}$$

$$\boxed{d_{\mathcal{S}}^2(x,z) = \min_{(x_*,z_*)\in\mathcal{S}}(\|x - x_*\|^2 + \|z - z_*\|^2)}$$

Future work

- other asymptotic behaviors [Holding & Lestas, CDC 2014]
- matrix flows [Helmke & Moore, "Opt. & Dyn. Systems"]
- robustness analysis
- finite-length trajectories
- gradient conjecture of René Thom for saddle-point dynamics

Primal-dual dynamics

For inequalities, dual optima are nonnegative:

$$\dot{x} = -\nabla_x F(x,z)$$
$$\dot{z} = [\nabla_z F(x,z)]_z^+$$

$$[a]_b^+ = \begin{cases} a & \text{if } a \geq 0 \text{ or } b > 0 \\ 0 & \text{otherwise} \end{cases}$$

Existing results on convergence:

- Arrow & Hurwitz & Uzawa (1959): Direct method with Taylor approximation – *limits further analysis*
- Fiejer & Paganini (2010): Indirect method using hybrid automata theory – *continuity not satisfied*

Primal-dual dynamics

For inequalities, dual optima are nonnegative:

$$\dot{x} = -\nabla_x F(x, z)$$
$$\dot{z} = [\nabla_z F(x, z)]^+_z$$

$$[a]^+_b = \begin{cases} a & \text{if } a \geq 0 \text{ or } b > 0 \\ 0 & \text{otherwise} \end{cases}$$

Existing results on convergence:

- Arrow & Hurwitz & Uzawa (1959): Direct method with Taylor approximation – *limits further analysis*
- Fiejer & Paganini (2010): Indirect method using hybrid automata theory – *continuity not satisfied*

Our *contribution* is a *novel proof* methodology:

- consider solutions in Caratheodory sense
- model as a projected dynamical system
- use LaSalle Invariance Principle for Caratheodory systems

[A. Cherukuri & E. Mallada & J. Cortés, SIAM CT 2015]

[A. Cherukuri & E. Mallada & J. Cortés, SCL, 2015]

Thank you. Comments or questions?

Jorge Cortés

Bahman Gharesifard

Enrique Mallada