



LUND UNIVERSITY

LCCC focus period and workshop on Dynamics and Control in Networks

Como, Giacomo; Rantzer, Anders; Westin, Eva

2015

Document Version:

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Como, G., Rantzer, A., & Westin, E. (2015). *LCCC focus period and workshop on Dynamics and Control in Networks*. (Technical Reports TFRT-7640). Department of Automatic Control, Lund Institute of Technology, Lund University.

Total number of authors:

3

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00



Workshop: Dynamics and Control in Networks

LUND CENTER FOR CONTROL OF COMPLEX ENGINEERING SYSTEMS
LUND UNIVERSITY



LCCC focus period and workshop on Dynamics and Control in Networks

15-17 Oct 2014

Old Bishop's Palace at Biskopsgatan 1 in Lund

Scientific Committee

Giacomo Como, Lund University, Sweden

Tryphon Georgiou, University of Minnesota, USA

Steven Low, CalTech, USA

Asuman Ozdaglar, Massachusetts Institute of Technology, USA

Anders Rantzer, Lund University, Sweden, LCCCcoordinator

Rodolphe Sepulchre, University of Cambridge, UK

Sandro Zampieri, Università di Padova, Italy

Organizing Committee

Giacomo Como

Anders Rantzer

Eva Westin

MAILING ADDRESS

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND, SWEDEN

VISITING ADDRESS

Institutionen för Reglerteknik
Ole Römers väg 1
232 63 LUND

TELEPHONE

+46 46 222 87 87

FAX

+46 46 13 81 18

GENERIC E-MAIL ADDRESS

control@control.lth.se

WWW

www.lccc.lth.se

Printed: Media-Tryck, Lund, Sweden, Mars 2015

ISSN 0280-5316

ISRN LUTFD2/TFRT--7640--SE

Content

1. Introduction	5
1.1 Scientific Theme	5
2. Workshop	6
2.1 Workshop Seminars	6
2.2 Focus Period Seminars	7
2.3 Selected Open Problems	7
3. Panel discussion	8
Appendix A PROGRAM	9
Appendix B PARTICIPANTS	11
Appendix C PRESENTATIONS	13
Controllability for Complex Networks: Metrics and Limitations Sandro Zampieri , Università di Padova, Italy	13
Differentially positive systems Rodolphe Sepulchre , University of Cambridge, UK	24
A survey of classical and recent results in RLC circuit synthesis Malcolm Smith , University of Cambridge, UK	33
Synchronization and evolvability in nonlinear networks Jean-Jacques Slotine , Massachusetts Institute of Technology, USA	43
Community Detection: Fundamental Limits and Optimal Sampling Algorithms Alexandre Proutiere , Kungliga Tekniska Högskolan, Sweden	44
Community detection in stochastic block models via spectral methods Laurent Massoulié , Microsoft Research-INRIA Joint Centre	59
Complex Energy Systems Michael Chertkov , Los Alamos National Laboratory, USA	70
Semidefinite relaxation of optimal power flow Steven H. Low , CalTech, US	90
Distributed Alternating Direction Method of Multipliers for Multi-agent Optimization Asuman Ozdaglar , Massachusetts Institute of Technology, USA	109
Learning graphical models: hardness and tractability Devavrat Shah , Massachusetts Institute of Technology, USA	120
Load Balancing Using Limited State Information R. Srikant , University of Illinois at Urbana Champain, USA	130
On the eigenvalues of large structured matrices and the scalable stability of networks Glenn Vinnicombe , University of Cambridge, UK	138
Toward standards for dynamics in future electric energy systems: The basis for plug-and-play industry paradigm Marija Ilic , Carnegie Mellon University, USA	142

Plug-and-Play Control and Optimization in Microgrids Florian Dörfler , ETH Zurich, Switzerland	156
Lyapunov Approach to Consensus Problems Angelia Nedic , University of Illinois at Urbana Champaign, USA	166
Majority Consensus by local polling Moez Draief , Imperial College, UK	172
Robust Networked Stabilization: When Gap Meets Two-Port Li Qiu , Hong Kong University of Science and Technology	186
Schroedinger bridges: steering of stochastic systems classical and quantum Tryphon Georgiou , University of Minnesota, USA	191
Randomized averaging algorithms, when can errors and unreliabilities just be ignored? Julien Hendrickx , Université Catholique de Louvain, Belgium	206
Systemic Risk in Finance Munther A. Dahleh , Massachusetts Institute of Technology, USA	216
On Distributed Control of Dynamical Flow Networks Giacomo Como , Lund University, Sweden	223
Control of convex-monotone systems Anders Rantzer , Lund University	237

1. Introduction and organization

The LCCC focus period on Dynamics and Control in Networks took place in Lund in October 2014. It spanned over five weeks with a 3-day workshop in the middle (October 15-17). The highly cross-disciplinary research theme was chosen by the LCCC board to support the strategic vision of the Centre.

The event was organized by the LCCC faculty members Giacomo Como and Anders Rantzer, and local practical arrangements were handled by Eva Westin. A scientific committee was formed, composed by Tryphon Georgiou (University of Minnesota), Steven Low (CalTech), Asuman Ozdaglar (Massachusetts Institute of Technology), Rodolphe Sepulchre (University of Cambridge), Sandro Zampieri (Università di Padova), and the organizers. The scientific committee selected other fifteen world-leading researchers in the field and invited them to give a seminar at the workshop, extend their stay during the focus period compatibly with their commitments, and nominate outstanding young researchers for an extended stay. Seventeen such researchers were selected by the scientific committee and invited to join the Center for periods of 3 to 5 weeks and to give a seminar during the focus period.

In addition to their seminar, all the participants were invited to submit open problems that could motivate and inspire the research community. The initiative was meant as a first step of a wider project to collect open research problems in Control Systems in conjunction with the 2015-16 thematic year in Control Theory and its Applications organized at the Institute for Mathematics and its Applications (IMA). The submitted problems underwent a peer review process and some selected ones were presented during the workshop. The workshop was concluded by a panel discussion moderated by Anders Rantzer.

1.1 SCIENTIFIC THEME

Networks, ranging from infrastructure (such as transportation, communication, and distribution networks) to social and economical networks, play an increasingly central role in many aspects of our lives. Many of the key interactions in these networks are inherently dynamic and the network connectivity and feedback pathways affect robustness and functionality. Such concepts are at the core of a new and rapidly evolving frontier of Dynamical Systems and Control.

In this new technological and scientific realm, the modeling and representation of systems and the role of feedback need to be reevaluated. Traditional thinking, which is limited to a small number of feedback loops, is no longer applicable, e.g., because of its lack of scalability. Feedback control and stability of network dynamics require new approaches. Decentralized control and distributed optimization, as well as layered architectures and plug-and-play paradigms have become of central interest.

The October 2014 LCCC Focus Period and Workshop brought together leading experts and outstanding young researchers with different scientific backgrounds in disciplines such as Control, Power Systems, Operation Research, Communication Networks, and Statistics, to share their results, compare their approaches and attempt to define the goals and boundaries of this exciting new research field.

2. Workshop

2.1 WORKSHOP SEMINARS

Controllability for Complex Networks: Metrics and Limitations

Sandro Zampieri (Università di Padova)

Differentially positive systems

Rodolphe Sepulchre, University of Cambridge, UK

A survey of classical and recent results in RLC circuit synthesis

Malcolm Smith, University of Cambridge, UK

Synchronization and evolvability in nonlinear networks

Jean-Jacques Slotine, Massachusetts Institute of Technology, USA

Community Detection: Fundamental Limits and Optimal Sampling Algorithms

Alexandre Proutiere, Kungliga Tekniska Högskolan, Sweden

Community detection in stochastic block models via spectral methods

Laurent Massoulié, Microsoft Research-INRIA Joint Centre

Complex Energy Systems

Michael Chertkov, Los Alamos National Laboratory, USA

Semidefinite relaxation of optimal power flow

Steven H. Low, CalTech, USA

Distributed Alternating Direction Method of Multipliers for Multi-agent Optimization

Asuman Ozdaglar, Massachusetts Institute of Technology, USA

Learning graphical models: hardness and tractability

Devavrat Shah, Massachusetts Institute of Technology, USA

Load Balancing Using Limited State Information

R. Srikant, University of Illinois at Urbana Champaign, USA

On the eigenvalues of large structured matrices and the scalable stability of networks

Glenn Vinnicombe, University of Cambridge, UK

Toward standards for dynamics in future electric energy systems

Marija Ilic, Carnegie Mellon University, USA

Plug-and-Play Control and Optimization in Microgrids

Florian Dörfler, ETH Zurich, Switzerland

Lyapunov Approach to Consensus Problems

Angelia Nedic, University of Illinois at Urbana Champaign, USA

Majority Consensus by local polling

Moez Draief, Imperial College, UK

Robust Networked Stabilization: When Gap Meets Two-Port

Li Qiu, Hong Kong University of Science and Technology

Schroedinger bridges: steering of stochastic systems classical and quantum

Tryphon Georgiou, University of Minnesota, USA

Randomized averaging algorithms, when can errors and unreliabilities just be ignored?

Julien Hendrickx, Université Catholique de Louvain, Belgium

Systemic Risk in Finance

Munther A. Dahleh, Massachusetts Institute of Technology, USA

On Distributed Control of Dynamical Flow Networks

Giacomo Como, Lund University, Sweden

Control of convex-monotone systems

Anders Rantzer, Lund University, Sweden

2.2 FOCUS PERIOD SEMINARS

Robust stability of positive systems: a convex characterization

Marcello Colombino (ETH Zurich)

A new approach to sequential rate-distortion problems and applications

Takashi Tanaka (Massachusetts Institute of Technology)

Stability and power sharing in microgrids

Johannes Schiffer (TU Berlin)

Differentially positive systems

Fulvio Forni (University of Liege)

Submodularity and controllability in complex dynamical networks

Tyler Summers (ETH Zurich)

A scalable approach to the design of large networks

Richard Pates (University of Cambridge)

Bounded disturbance amplification for mass chains with passive interconnections

Kaoru Yamamoto (University of Cambridge)

Gossip coverage control for robotic networks: dynamical systems on the space of partitions

Ruggero Carli (Università di Padova)

A communication/control co-design paradigm for networked control systems

Wei Chen (Hong Kong University of Science and Technology)

Resilience of natural gas networks during conflicts, crises and disruptions

Rui Carvalho (University of Cambridge)

New approaches to identification and control of nonlinear systems

Ian Manchester, University of Sydney

On the push-sum algorithm with unreliable communication

Balázs Gerencsér (Université Catholique de Louvain)

Resilience and cascading failures in large-scale networks

Wilbert Rossi (Politecnico di Torino)

Universal convexification via risk aversion

Dwijotham Krishnamurthy (CalTech)

Distributed and event-based state estimation

Sebastian Trimpe, (Max Planck Institute for Intelligent Systems)

Stability of linear consensus processes

Ji Liu (University of Illinois at Urbana Champaign)

Control of evolutionary games on networks

James R. Riehl (University of Groningen)

2.3 SELECTED OPEN PROBLEMS

Two Problems on Minimality in RLC Circuit Synthesis

Timothy H. Hughes¹, Jason Z. Jiang, and Malcolm C. Smith

The Classical n-Port Resistive Synthesis Problem

Michael Z. Q. Chen

Synchronization of oscillators: Feasibility and Non-local analysis

Julien Hendrickx Florian Dorfler

Modeling for plug and play control in strongly coupled nonlinear networks

Marija D. Ilic and Xia Miao

Stability of Passivity Based Control for Power Systems and Power Electronics

Kevin Bacovchin and Marija D. Ilic

On the asymptotic behavior of differentially positive systems

Fulvio Forni and Rodolphe Sepulchre

3. Panel Discussion

The workshop was concluded by a panel discussion moderated by Anders Rantzer. The panel was composed by Munther Dahleh (Massachusetts Institute of Technology) Marija Ilic (Carnegie Mellon University), and Glenn Vinnicombe (Cambridge University). The panelists were asked to address the following points:

1. We have had lectures on control, energy, community detection and financial risks. What are the important common issues and the important differences? Do we have enough of a common language?
2. How do we stay relevant and where do we set the boundaries of our discipline?
3. What teaching curriculum do we need to address the common issues?

Concerning point 1, the consensus was that the topics could be classified in two relatively homogeneous categories. On the one hand, in large-scale physical systems such as power, transportation, and financial networks, there are measurable flows among the different units and there is an infrastructure posing constraints on such flows. Issues of efficiency and robustness/resilience are ubiquitous in this area. On the other hand, there are problems of community detection, consensus, and opinion dynamics in social networks where the central issue is to estimate from data the actual network structure and decision mechanisms in place.

The centrality of ideas such as the design of local decision makers achieving a global objective and layering, e.g., exploiting time-scale separations that are inherent in many complex systems was also brought up.

The discussion on Point 2 highlighted the tension between the need for the community not

to loose its rigorous methodology while being close to the problem. The importance and challenges of having good contact with practitioners were also brought up.

Point 3 stimulated the debate on how to go back to fundamentals and get rid of what is no longer so central. The need for engineers with systems knowledge was highlighted along with the fact that computer science and communication should be fundamental parts of control engineering curricula.

PROGRAM LCCC October 2014

Workshop on Dynamics and Control in Networks

Wednesday, October 15

- 08:30 Registration at the Old Bishop's Palace at Biskopsgatan 1 in Lund
- 08:45 Opening remarks
Giacomo Como and Anders Rantzer, Lund University, Sweden
- 09:00 Controllability for Complex Networks: Metrics and Limitations
Sandro Zampieri, Università di Padova, Italy
Differentially positive systems
Rodolphe Sepulchre, University of Cambridge, UK
- 10:20 Coffee
- 10:40 A survey of classical and recent results in RLC circuit synthesis
Malcolm Smith, University of Cambridge, UK
Synchronization and evolvability in nonlinear networks
Jean-Jacques Slotine, Massachusetts Institute of Technology, USA
- 12:00 Lunch
- 14:00 Community Detection: Fundamental Limits and Optimal Sampling Algorithms
Alexandre Proutiere, Kungliga Tekniska Högskolan, Sweden
Community detection in stochastic block models via spectral methods
Laurent Massoulié, Microsoft Research-INRIA Joint Centre
- 15:20 Coffee
- 15:40 Complex Energy Systems
Michael Chertkov, Los Alamos National Laboratory, USA
Semidefinite relaxation of optimal power flow
Steven H. Low, CalTech, USA

Thursday, October 16

- 09:00 Distributed Alternating Direction Method of Multipliers for Multi-agent Optimization
Asuman Ozdaglar, Massachusetts Institute of Technology, USA
Learning graphical models: hardness and tractability
Devavrat Shah, Massachusetts Institute of Technology, USA
- 10:20 Coffee
- 10:40 Load Balancing Using Limited State Information
R. Srikant, University of Illinois at Urbana Champaign, USA
On the eigenvalues of large structured matrices and the scalable stability of networks
Glenn Vinnicombe, University of Cambridge, UK
- 12:00 Lunch



- 14:00 Toward standards for dynamics in future electric energy systems:
 The basis for plug-and-play industry paradigm
Marija Ilic, Carnegie Mellon University, USA
 Plug-and-Play Control and Optimization in Microgrids
Florian Dörfler, ETH Zurich, Switzerland
- 15:20 Coffee
- 15:40 Lyapunov Approach to Consensus Problems
Angelia Nedic, University of Illinois at Urbana Champaign, USA
 Majority Consensus by local polling
Moez Draief, Imperial College, UK
- 18:15 Buses leave from Bangatan (green arrow on map)
- 19:00 Symposium Dinner at Långa bryggan, Bjärreds saltsjöbad

Friday, October 17

- 09:00 Robust Networked Stabilization: When Gap Meets Two-Port
Li Qiu, Hong Kong University of Science and Technology
 Schroedinger bridges: steering of stochastic systems classical and quantum
Tryphon Georgiou, University of Minnesota, USA
- 10:20 Coffee
- 10:40 Randomized averaging algorithms, when can errors and unreliabilities just be ignored?
Julien Hendrickx, Université Catholique de Louvain, Belgium
 Systemic Risk in Finance
Munther A. Dahleh, Massachusetts Institute of Technology, USA
- 12:00 Lunch
- 14:00 On Distributed Control of Dynamical Flow Networks
 Giacomo Como, Lund University, Sweden
 Control of convex-monotone systems
 Anders Rantzer, Lund University, Sweden
- 15:20 Coffee
- 15:40 Panel Discussion
Anders Rantzer (moderator)
- 17:00 Final remarks – end of the workshop



Appendix B

PARTICIPANTS LCCC Focus Period 2014

Workshop on Dynamics and Control in Networks

Amir Aminifar	Linköping University	amir.aminifar@liu.se
Karl Johan Åström	Lund University	karl_johan.astrom@control.lth.se
Bo Bernhardsson	Lund University	bo.bernhardsson@control.lth.se
Ruggero Carli	University of Padova	carlirug@dei.unipd.it
Rui Carvalho	University of Cambridge	r.carvalho@statslab.cam.eng.ac.uk
Wei Chen	Hong Kong Univ. of Science and Tech.	wchenust@ust.hk
Michael Chertkov	Los Almos National Laboratory	chertkov@lanl.gov
Marcello Colombino	ETH Zürich	mcolombi@control.ee.ethz.ch
Giacomo Como	Lund University	giacomo.como@control.lth.se
Munther Dahleh	MIT	dahleh@mit.edu
Florian Dörfler	University of California at Los Angeles	dorfler@seas.ucla.edu
Moez Draief	Imperial College	m.draief@imperial.ac.uk
Fulvio Forni	ETH Zürich	fforni@ulg.ac.be
Basilio Gentile	University of Liège	gentileb@control.ee.ethz.ch
Tryphon Georgiou	University of Minnesota	tryphon@umn.edu
Balázs Gerencsér	University of Louvain	balazs.gerencser@uclouvain.be
Camille Hamon	Royal Institute of Technology	camilleh@kth.se
Julien Hendrickx	University of Louvain	julien.hendrickx@uclouvain.be
Marija Ilic	Carnegie Mellon University	milic@ece.cmu.edu
Sei Zhen Khong	Lund University	seizhen@control.lth.se
Dwijotham Krishnamurthy	CalTech	dvij@cs.washington.edu
Carolina Lidström	Lund University	carolina.lidstrom@control.lth.se
Ji Liu	Univ. of Illinois at Urbana Champaign	jiliu@illinois.edu
Steven Low	CalTech	slow@caltech.edu
Daria Madjidian	Lund University	daria.madjidian@control.lth.se
Martina Maggio	Lund University	martina.maggio@control.lth.se
Ian Manchester	University of Sydney	ian.manchester@sydney.edu.au
Laurent Massoulié	Microsoft Research INRIA	laurent.massoulie@inria.fr
Victor Millnert	Lund University	victor.millnert@control.lth.se
Sidhant Misra	MIT	sidhant@mit.edu
Angela Nedic	Univ. of Illinois at Urbana Champaign	angelia@illinois.edu
Gustav Nilsson	Lund University	gustav.nilsson@control.lth.se
Asuman Ozdaglar	MIT	asuman@mit.edu
Alessandro Papadopoulos	Lund University	alessandro.papadopoulos@control.lth.se
Francesca Parise	ETH Zürich	parisef@control.ee.ethz.ch
Richard Pates	University of Cambridge	rtp22@cam.ac.uk

Magnus Perninge	Lund University	magnus.perninge@control.lth.se
Bala Kameshwar Poolla	ETH Zürich	bpoolla@control.ee.ethz.ch
Alexandre Proutiere	Royal Institute of Technology	alepro@kth.se
Li Qiu	Hong Kong Univ. of Science and Tech.	eeqiu@ece.ust.hk
Maben Rabi	Chalmers	maben.rabi@chalmers.se
Anders Rantzer	Lund University	anders.rantzer@control.lth.se
James R. Riehl	University of Groningen	j.r.riehl@rug.nl
Wilbert Samuel Rossi	Tech. University of Torino	wilbert.rossi@polito.it
Henrik Sandberg	Royal Institute of Technology	hsan@kth.se
Johannes Schiffer	TU Berlin	schiffer@control.tu-berlin.de
Rodolphe Sepulchre	University of Cambridge	r.sepulchre@eng.cam.ac.uk
Devavrat Shah	MIT	devavrat@mit.edu
Jean-Jacques Slotine	MIT	jjs@mit.edu
Malcolm Smith	University of Cambridge	mcs@eng.cam.ac.uk
R. Srikant	Univ. of Illinois at Urbana Champaign	rsrikant@illinois.edu
Yvonne Stuerz	ETH Zürich	stuerzy@control.ee.ethz.ch
Tyler Summers	ETH Zürich	tsummers@control.ee.ethz.ch
Sebastian Trimpe	Max Planck Inst. for Intelligent Systems	strimpe@tuebingen.mpg.de
Olof Troeng	Lund University	olof.troeng@control.lth.se
Glenn Vinnicombe	University of Cambridge	ky255@cam.ac.uk
Eva Westin	Lund University	eva.westin@control.lth.se
Björn Wittenmark	Lund University	bjorn.wittenmark@control.lth.se
Kuang Xu	Microsoft Research INRIA	kuangxu@gmail.com
Kaoru Yamamoto	University of Cambridge	gv@eng.cam.ac.uk
Sandro Zampieri	University of Padova	zampi@dei.unipd.it



Appendix C

PRESENTATIONS

CONTROLLABILITY FOR COMPLEX NETWORKS: METRICS AND LIMITATIONS

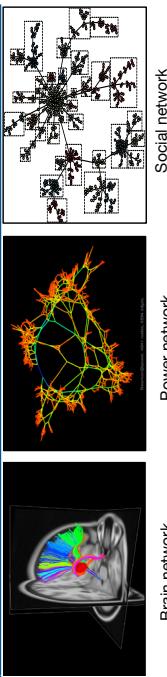
Sandro Zampieri, Università di Padova, Italy

This presentation will consider the problem of controlling complex networks, i.e., the joint problem of selecting a set of control nodes and of designing a control input to steer a network to a target state. For this problem, we propose a metric to quantify the difficulty of the control problem as a function of the required control energy and we derive bounds based on the system dynamics (network topology and weights) to characterize the tradeoff between the control energy and the number of control nodes. Our findings show several control limitations and properties. For instance, for symmetric networks, if the number of control nodes is constant, then the control energy increases exponentially with the number of network nodes. On the other hand, if the number of control nodes is a fixed fraction of the network nodes, then certain networks can be controlled with constant energy independently of the network dimension.

Asymmetric networks highlights richer scenarios. Indeed, it can be shown that, while asymmetric but isotropic networks are still difficult to control, sufficiently anisotropic networks are instead easy to control.



Growing Number of Applications



Sandro Zampieri

Università' di Padova

In collaboration with
Fabio Pasqualetti - University of California at Riverside
Francesco Bullo - University of California at Santa Barbara

Controllability metrics, limitations and algorithms for complex networks

- ✓ Y. Y. Liu, J. J. Slotine, and A. L. Barabasi, "Controllability of complex networks," *Nature* 2011.
- ✓ N. J. Cowan, E. J. Chastain, D. A. Vilhena, J. S. Freudenberg, and C. T. Bergstrom, "Nodal dynamics, not degree distributions, determine the structural controllability of complex networks," *PLoS ONE*, 2012.
- ✓ G. Yan, J. Ren, Y.-C. Lai, C.-H. Lai, and B. Li, "Controlling complex networks: How much energy is needed?" *Physical Review Letters*, 2012.
- ✓ J. Sun and A. E. Motter, "Controllability transition and nonlocality in network control," *Physical Review Letters*, 2013.
- ✓ F. L. Contesini, T. H. Summers, and J. Lygeros, "Submodularity of Energy Related Controllability Metrics," *Arxiv Preprint*, 2014.
- ✓ A. Olshevsky, "Minimal Controllability Problems," *Arxiv preprint*, 2014.

Problem formulation

Consider a linear system

$$x(t+1) = Ax(t) + Bu(t)$$

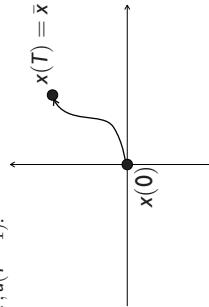
where A is sparse $n \times n$ matrix (the interactions between the states is described by a graph) and

$$B = [e_{i_1} \dots e_{i_n}]$$

where

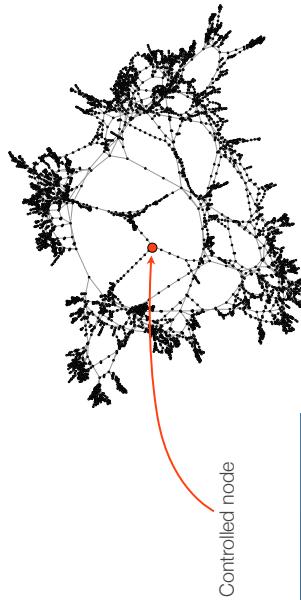
$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \Rightarrow \quad i$$

CONTROLLABILITY is the possibility of steering the state from the initial state $x(0) = 0$ to an arbitrary final state $x(T) = \bar{x}$ by applying a suitable input sequence $u(0), u(1), \dots, u(T-1)$.



Need of a controllability metric

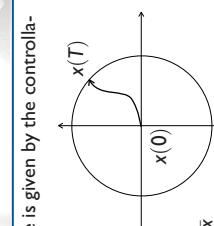
Assume that the graph is strongly connected and that it has all the self-loops.
Then to have that the resulting system is controllable (generically in the non-zero entries of A) it is enough that only one state is controlled.



Need of a controllability metric

QUESTION: How controllable is the resulting system?

A controllability metric



One metric for describing the controllability degree is given by the controllability Gramian

$$W_T := \sum_{t=0}^{T-1} A^t B B^T (A^T)^t$$

Interpretation: For driving

$$x(0) = 0 \longrightarrow x(T) = \bar{x}$$

where $\|\bar{x}\|_2 = 1$, we need an input $u(0), u(1), \dots, u(T-1)$ with L^2 energy $\bar{x}^T W_T^{-1} \bar{x}$.

The energy to drive the state from zero to a norm one state (in the worst case) is given by

$$\text{Energy} = \frac{1}{\lambda_{\min}(W_T)}$$

A controllability metric

A controllability metric



$$W_T := \sum_{t=0}^{T-1} A^t B B^T (A^T)^t \quad \text{Controllability Gramian}$$

$$\text{Energy} = \frac{1}{\lambda_{\min}(W_T)}$$

Small $\lambda_{\min}(W_T)$ \longrightarrow Small controllability degree

Large $\lambda_{\min}(W_T)$ \longrightarrow Large controllability degree

A controllability metric

Alternative controllability metrics

$$\mathbf{W}_T := \sum_{t=0}^{T-1} \mathbf{A}^t \mathbf{B} \mathbf{B}^T (\mathbf{A}^T)^t$$

$$\lambda_{\min}(\mathbf{W}_T)$$

$$\frac{1}{n} \text{tr}(\mathbf{W}_T^{-1})$$

$$\frac{1}{n} \ln \det(\mathbf{W}_T)$$

$$\frac{1}{n} \text{tr}(\mathbf{W}_T)$$

THEOREM

Assume the matrix \mathbf{A} symmetric. Fix any constant $0 < \mu < 1$ and let

$$n(\mu) := |\{\lambda \in \lambda(\mathbf{A}) : |\lambda|^2 \leq \mu\}|$$

Then

$$\lambda_{\min}(\mathbf{W}_T) \leq \frac{1}{\mu(1-\mu)} \mu^{n(\mu)/m}$$

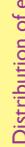
9

Few Nodes Cannot Control Symmetric Complex Networks

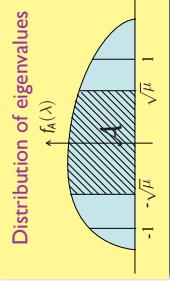
10

Few Nodes Cannot Control Symmetric Complex Networks

Distribution of eigenvalues
 $f_A(\lambda)$



$$\lambda_{\min}(\mathbf{W}_T) \leq \frac{1}{\mu(1-\mu)} \mu^{A \frac{n}{m}}$$



Distribution of eigenvalues
 $f_A(\lambda)$

$$\text{Let } F_A(\bar{\lambda}) := \frac{|\{\lambda \in \lambda(\mathbf{A}) : \lambda \leq \bar{\lambda}\}|}{n}$$

$$\text{and } f_A(\bar{\lambda}) := \frac{dF_A(\bar{\lambda})}{d\lambda}$$

Then

$$n(\mu) := |\{\lambda \in \lambda(\mathbf{A}) : |\lambda|^2 \leq \mu\}| = n \int_{-\sqrt{\mu}}^{\sqrt{\mu}} f_A(\lambda) d\lambda = A n$$

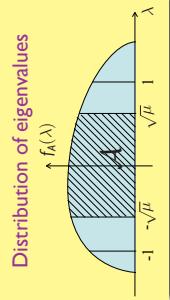
Hence the number $n(\mu)$ typically grows linearly in the network cardinality n .

$$\lambda_{\min}(\mathbf{W}_T) \leq \frac{1}{\mu(1-\mu)} \mu^{A \frac{n}{m}} \quad \mu < 1$$

11

12

Few Nodes Cannot Control Symmetric Complex Networks

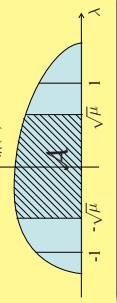


Few Nodes Cannot Control Symmetric Complex Networks

$\mu < 1$

$$\lambda_{min}(W_T) \leq \frac{1}{\mu(1 - \mu)} \mu^{A_{\frac{n}{m}}}$$

Distribution of eigenvalues



✓ For fixed number m of control nodes, the controllability degree decreases exponentially in the network cardinality n .

✓ To have a fixed controllability degree, number m of control nodes must grow linearly in the network cardinality n .

Example: symmetric line graph

$\mu < 1$

$$A = \begin{bmatrix} a & b & 0 & \cdots & 0 \\ b & a & b & \cdots & 0 \\ 0 & b & a & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a \\ 0 & 0 & 0 & \cdots & b \end{bmatrix}$$

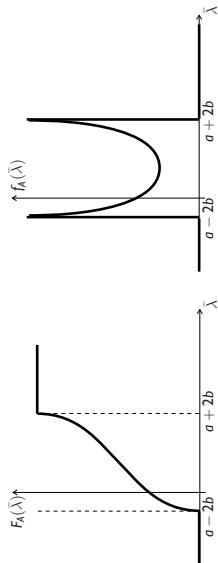
$$\lambda(A) = \left\{ a + 2b \cos\left(\frac{\pi k}{n+1}\right) : k = 1, 2, \dots, n \right\}$$

13

Example: symmetric line graph

$f_A(\lambda) = 1 - \frac{1}{\pi} \arccos\left(\frac{\bar{\lambda} - a}{2b}\right)$

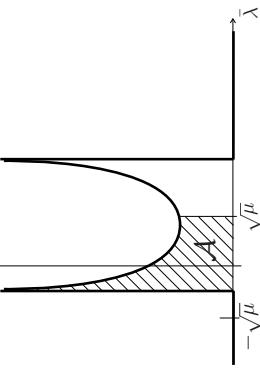
$$f_A(\bar{\lambda}) = \frac{1/\pi}{\sqrt{1 - \left(\frac{\bar{\lambda} - a}{2b}\right)^2}}$$



14

Example: symmetric line graph

$f_A(\lambda)$



15

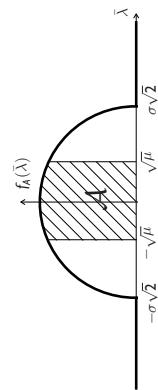
16

Example: Symmetric random matrix

Assume A is a symmetric matrix with a_{ij} iid random variables $\mathbb{E}[a_{ij}] = 0$ and $\mathbb{E}[a_{ij}^2] = \sigma^2/\sqrt{n}$.

Then (Wigner's semi-circle law)

$$f_A(\bar{\lambda}) \rightarrow W(\bar{\lambda})$$



$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \cdots & a_{nn} \end{bmatrix}$$

For asymmetric networks the situation is more complex

- For "isotropic" networks (networks with no preferential directions) it seems that the situation is the same as for symmetric networks, namely they are difficult to control.

- For "anisotropic" networks (networks with a preferential direction) it seems that few nodes can indeed control large scale networks.

17

Asymmetric Complex Networks

THEOREM

Assume the matrix A diagonalizable and let V an eigenvector matrix. Fix any constant $0 < \mu < 1$ and let

$$n(\mu) := |\{\lambda \in \lambda(A) : |\lambda|^2 \leq \mu\}|$$

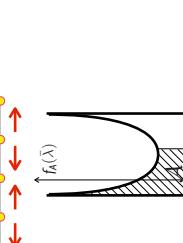
Then

$$\lambda_{\min}(W_T) \leq \text{cond}(V)^2 \frac{1}{\mu(1-\mu)} \mu^{n(\mu)/m}$$

where $\text{cond}(V) = \sigma_{\max}(V)/\sigma_{\min}(V)$ is the condition number of V .

18

Example: Asymmetric isotropic line graph



It can be found an eigenvector matrix V such that $\text{cond}(V) = 2$.

$$\lambda(A) = \left\{ \frac{1}{3} + \frac{2}{3} \cos \frac{k\pi}{n+1}, \quad k = 1, \dots, n \right\}$$

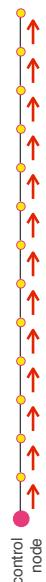
Consequence: If the condition number stays bounded in the network dimension, than the network remains difficult to control.

This network is difficult to control as a symmetric network.

19

20

Example: Asymmetric anisotropic line graph



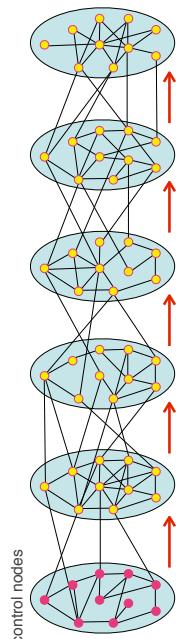
$$A = \begin{bmatrix} a & b & 0 & \cdots & 0 & 0 \\ c & a & b & \cdots & 0 & 0 \\ 0 & c & a & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & b \\ 0 & 0 & 0 & \cdots & c & a \end{bmatrix}$$

$$\lambda(A) = \left\{ a + 2\sqrt{bc} \cos\left(\frac{\pi}{n+1}k\right) : k = 1, 2, \dots, n \right\}$$

Given a , if c is sufficiently larger than b , then this network can be controlled with finite energy by the node on the extreme regardless the network dimension.

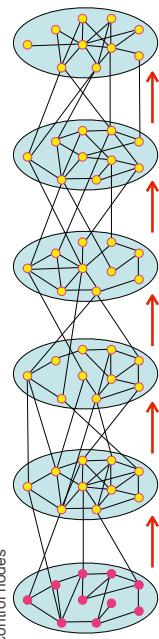
Exploiting spatial instability

control
node



$$A = \begin{bmatrix} A_1 & B_1 & 0 & \cdots & 0 & 0 \\ C_1 & A_2 & B_2 & \cdots & 0 & 0 \\ 0 & C_2 & A_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{\ell-1} & B_{\ell-1} \\ 0 & 0 & 0 & \cdots & C_{\ell-1} & A_\ell \end{bmatrix}$$

Extension to more general graphs: controllable graphs



control nodes

$$A = \begin{bmatrix} A_1 & B_1 & 0 & \cdots & 0 & 0 \\ C_1 & A_2 & B_2 & \cdots & 0 & 0 \\ 0 & C_2 & A_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{\ell-1} & B_{\ell-1} \\ 0 & 0 & 0 & \cdots & C_{\ell-1} & A_\ell \end{bmatrix}$$

Extension to more general graphs: uncontrollable graphs



control nodes

Assume A is a stochastic matrix (diffusion dynamics, consensus, ...). This means that

$$A\mathbf{1} = \mathbf{1}$$

where $\mathbf{1}$ is the vector with entries equal to one. Let w be a vector such that

$$w^T A = w^T \quad w^T \mathbf{1} = 1$$

which means that w is the invariant measure of the Markov chain associated with A . It is known that the entries of w represent the nodes "centrality" in the network (the bigger the more important).

Given the matrices A_i , if the matrices C_i are sufficiently larger than the matrices B_i , then this network can be controlled with finite energy by the nodes on the extreme subgraph regardless the network dimension.

In the symmetric case the entries of w are $1/n$.



Controllers positioning

Decoupled control strategy: "Divide et Impera"

Network partitioning Partition $\mathcal{V} = \{1, \dots, n\}$ into N disjoint sets $\mathcal{V}_1, \dots, \mathcal{V}_N$.

After relabeling of states and inputs, the matrices read as

$$A = \begin{bmatrix} A_1 & & A_{1N} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_N \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & B_N \end{bmatrix},$$

The networks dynamics can be written as the interconnection of N subsystems of the form

$$\dot{x}(t+1) = \underbrace{A_i \dot{x}(t)}_{\text{local dynamics}} + \underbrace{\sum_{j \in \mathcal{N}_i} A_j x_j(t)}_{\text{interconnections}} + \underbrace{\sum_{j \in \mathcal{N}_i} B_{ij} u_j(t)}_{\text{local controls}},$$

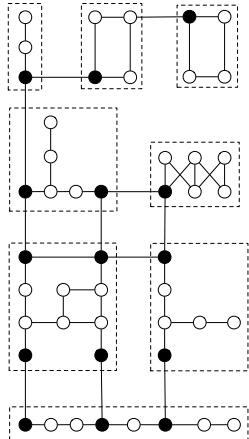
where $i \in \{1, \dots, N\}$ and $\mathcal{N}_i := \{j : A_{ij} \neq 0\}$.

Controllers positioning

Decoupled control strategy

Partition the network into disjoint connected parts.

1. Partition the network into disjoint connected parts.
2. Select boundary nodes as control nodes.



25

Controllers positioning

Decoupled control strategy

3. Apply the inputs

$$u(t) := v_i(t) - \sum_{j \in \mathcal{N}_i} B_j^T A_j x_j(t)$$

This control law yields N decoupled subsystems

$$x_i(t+1) = A_i x_i(t) + B_i v_i(t)$$

4. Choose v_i which minimizes the energy to steer the subsystem to the desired substrate.

Local controllability

$$\Lambda := \begin{bmatrix} \lambda_{\min}^{-1}(W_{1,T}) & 0_2 & \cdots & 0 \\ 0 & \lambda_{\min}^{-1}(W_{2,T}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{\min}^{-1}(W_{N,T}) \end{bmatrix}.$$

Coupling strength

$$\Delta := \begin{bmatrix} 1 & \|A_{12}\|_2 & \cdots & \|A_{1N}\|_2 \\ \|A_{21}\|_2 & 1 & \cdots & \|A_{2N}\|_2 \\ \vdots & \vdots & \ddots & \vdots \\ \|A_{N1}\|_2 & \|A_{N2}\|_2 & \cdots & 1 \end{bmatrix}.$$

Controllers positioning



26

Controllers positioning

Local controllability

$$\Lambda := \begin{bmatrix} \lambda_{\min}^{-1}(W_{1,T}) & 0_2 & \cdots & 0 \\ 0 & \lambda_{\min}^{-1}(W_{2,T}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{\min}^{-1}(W_{N,T}) \end{bmatrix}.$$

Coupling strength

$$\Delta := \begin{bmatrix} 1 & \|A_{12}\|_2 & \cdots & \|A_{1N}\|_2 \\ \|A_{21}\|_2 & 1 & \cdots & \|A_{2N}\|_2 \\ \vdots & \vdots & \ddots & \vdots \\ \|A_{N1}\|_2 & \|A_{N2}\|_2 & \cdots & 1 \end{bmatrix}.$$

27

28

Controllers positioning

Theorem If we choose a decoupled control law

then we obtain

$$\lambda_{\min}(W_T) \geq \frac{(1 - \bar{\lambda}_{\max})^2}{\|\Lambda\|_\infty \|\Delta\|_\infty^2}$$

where

$$\bar{\lambda}_{\max} = \max\{\lambda_{\max}(A_i) : i \in \{1, \dots, N\}\} < 1$$

Controllers positioning

Theorem If we choose a decoupled control law

then we obtain

$$\lambda_{\min}(W_T) \geq \frac{(1 - \bar{\lambda}_{\max})^2}{\|\Lambda\|_\infty \|\Delta\|_\infty^2} = \frac{(1 - \bar{\lambda}_{\max})^2}{\|\Lambda\|_\infty \|\Delta\|_\infty^2} = \frac{\text{local convergence speed}}{\text{(local controllability)/(coupling strength)}}$$

where

$$\bar{\lambda}_{\max} = \max\{\lambda_{\max}(A_i) : i \in \{1, \dots, N\}\} < 1$$

Controllers positioning

Theorem If we choose a decoupled control law

then we obtain

$$\lambda_{\min}(W_T) \geq \frac{(1 - \bar{\lambda}_{\max})^2}{\|\Lambda\|_\infty \|\Delta\|_\infty^2} = \frac{\text{local convergence speed}}{\text{(local controllability)/(coupling strength)}}$$

where

$$\bar{\lambda}_{\max} = \max\{\lambda_{\max}(A_i) : i \in \{1, \dots, N\}\} < 1$$

For high controllability degree:

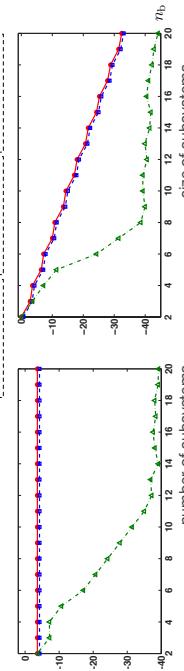
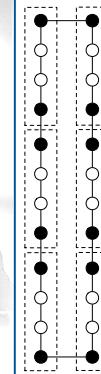
1. **Particular so that $\|\Delta\|_\infty$ are small (weakly coupled subsystems)**
2. **Select local control nodes so that $\|\Lambda\|_\infty$ is small (large local controllability)**

With decoupled control, global controllability becomes a local property

29

30

Circulant graph



λ_{\min} with the decoupled control strategy
 λ_{\min} theoretical lower bound
 λ_{\min} with random positioning

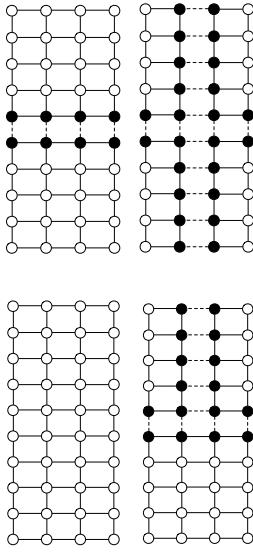
31

32

Network partition and selection of the control nodes

Until the desired number of control nodes have been selected:

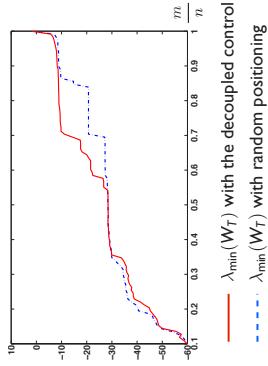
1. Bisect the least controllable subsystem via Fiedler partitioning.
2. Include boundary nodes in the control set.



Selection of the control nodes



Power grid with 118 nodes

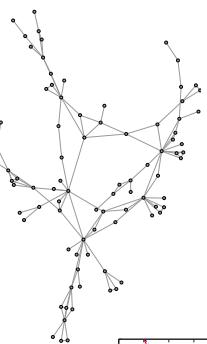


Examples



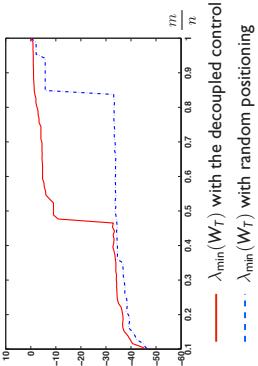
34

Examples

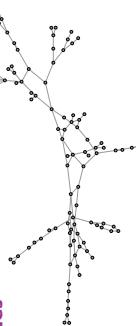


33

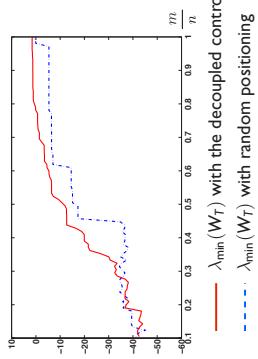
Epidemics network with 86 nodes



Social network with 105 nodes



Epidemics network with 86 nodes



Examples

35

36

Conclusions

- Similar results for **observability**
- For symmetric (isotropic) networks we need to control **a fixed fraction** of nodes
- For anisotropic networks it is enough to control **a fixed number** of nodes
- **Random** positioning works pretty well
- **Phase transition** can be noticed (critical fraction of controlled nodes)
- There are a lot of open problems:
 - Understanding isotropic and anisotropic networks
 - Controllability of random and of structured graphs
 - Performance of random positioning
 - Phase transition
 - Different metrics



37



38



Thank you

39

DIFFERENTIALLY POSITIVE SYSTEMS**Rodolphe Sepulchre, University of Cambridge, UK**

The talk will introduce differentially positive systems, as systems whose linearization along an arbitrary trajectory is positive. A generalization of Perron Frobenius theory is developed in this differential framework to show that the property induces a (co-) order that strongly constrains the asymptotic behavior of solutions. The results illustrate that behaviors constrained by local order properties extend beyond the well-studied class of linear positive systems and monotone systems, which both require a constant cone field and a linear state space. Joint work with Fulvio Forni.

Preprint on arxiv: <http://arxiv.org/abs/1405.6298>.



The main reference

Differential positivity

Differentially positive systems

F. Forni, R. Sepulchre

24 May 2014

Abstract

The paper introduces and studies differentially positive systems, that is, systems whose linearization along an arbitrary trajectory is positive. A generalization of Perron–Frobenius theory is developed in this differential framework to show that the property induces a (conal) order that strongly constrains the asymptotic behavior of solutions. The results illustrate that behaviors constrained by local order properties extend much beyond the well-studied class of linear positive systems and monotone systems, which both require a constant cone field and a linear state space.

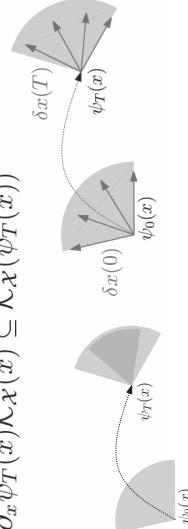
2

This talk: why did we study that property?

Differential positivity

Compared to contraction, a cone replaces the ball...

A linear map is (Lyapunov) stable if it leaves a ball invariant.
 A dynamical system is differentially stable (non expanding) if its linearization along an arbitrary trajectory is Lyapunov stable.



3

A linear map is (Lyapunov) stable if it leaves a ball invariant.

A dynamical system is differentially stable (non expanding) if its linearization along an arbitrary trajectory is Lyapunov stable.



In systems and control: Lohmiller & Slotine (1998), and many others since then...
 Terminology: contraction, convergence, incremental stability, ...

Textbook Perron-Frobenius theory

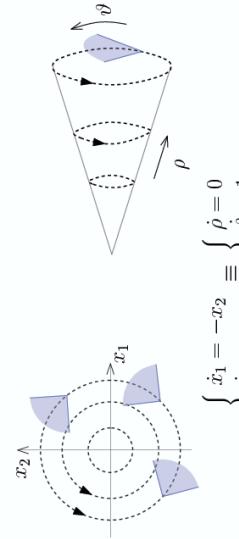
Perron–Frobenius theorem

From Wikipedia, the free encyclopedia.

In linear algebra, the Perron–Frobenius theorem, proved by Oskar Perron (1907) and Georg Frobenius (1912), asserts that a real square matrix with positive entries has a unique largest real eigenvalue and that the corresponding eigenvector has strictly positive components, and also asserts a similar statement for certain classes of nonnegative matrices. That theorem has important applications to probability theory (ergodicity of Markov chains), to the theory of dynamical systems (entropies of finite type), to economics (Okonek's theorem, Leontief's input-output model),^[1] to demography (Leslie population age distribution model),^[2] to Internet search engines,^[3] and even ranking of football teams.^[4]

↑ From Wikipedia, the free encyclopedia.

As a geometric concept, positivity is not antagonist to oscillations



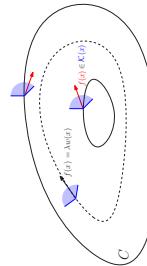
6

5

An obstacle to think of positivity as a geometric property?

Strict differential positivity and nonlinear oscillations

Corollary 2: Under the assumptions of Theorem 3, consider an open, forward invariant region $C \subseteq \mathcal{X}$ that does not contain any fixed point. If the vector field $f(x) \in \text{int} \mathcal{K}_{\mathcal{X}}(x)$ for any $x \in C$, then there exists a unique attractive periodic orbit contained in C .



Plan for this talk

- Motivation (1): positivity and networks

- Motivation (2): positivity and monotonicity

- Motivation (3): positivity on nonlinear spaces

- Theorem 3 is a *differential* version of Perron–Frobenius theory.
- The corollary is akin to Poincaré–Bendixson theorem for planar systems.

- Strict differential positivity, similarly to the topology of the plane, enforces a one-dimensional asymptotic behavior.

8

Recycling a few old slides...

Classical linear consensus theory

Linear consensus algorithms are linear time-varying systems

Consensus theory and Hilbert metric

R. Sepulchre
University of Liege, Belgium

LCCC workshop
January 2010

$$\dot{x}(t+1) = A(t)x(t), \quad x(t) \in \mathbb{R}^n$$

where for each t $A(t)$ is row stochastic, i.e.

A is nonnegative: $a_{ij} \geq 0$

each row sums to one: $A(t)\mathbf{1} = \mathbf{1}$

Uniform convergence to α_1 ("consensus": $x_i = \cdots = x_j$)
is proven under uniform connectivity / irreducibility
(Tsitsiklis, Jadbabaie et al., Moreau, ...)

Convergence analysis and Lyapunov functions

Tsitsiklis (1986) observed that

$$V(x) = \max_{1 \leq i \leq n} x_i - \min_{1 \leq i \leq n} x_i$$

is non increasing along the flow.
Uniform convergence is established by showing the strict decay of over a finite horizon.

It is known that no common quadratic Lyapunov exists in general.
(See Olshevsky & Tsitsiklis 08 for a discussion)

Birkhoff Theorem

Let K a closed solid cone in X a Banach space, with partial ordering \preceq .

A is positive if A maps $\overset{\circ}{K}$ to $\overset{\circ}{K}$
 A is monotone if $x \preceq y \Rightarrow Ax \preceq Ay$

Theorem (G. Birkhoff, 1957):

Positive linear monotone mappings contract the Hilbert metric in $\overset{\circ}{K}$.

The contraction coefficient is $\tanh \frac{1}{4} \Delta(A)$

Note: Perron-Frobenius follows from contraction mapping theorem

Conclusions

Conic geometries are adapted to consensus theory ...
 Quadratic Lyapunov functions aren't ...

Contraction analysis of linear consensus

Consider the displacements dynamics from (49) given by
 $\dot{x} = A(t)\delta x$, and the horizontal Finsler-Lyapunov function

$$V(x, \delta x) := \max_i \delta x_i - \min_i \delta x_i, \quad (50)$$

that coincides with the classical consensus function adopted in [30], [53] lifted to the tangent space. See [43] for its

A differential Lyapunov framework for contraction analysis.
 F. Forni and R.S., TAC 2014.

Tsitsiklis Lyapunov function is a measure of contraction of the Hilbert metric.

Birkhoff theorem (positive monotone operators contract the Hilbert metric)
 applies to more general cones, e.g. the SDP cone.

Opens the way to a consensus theory in noncommutative spaces,
 with a number of possible applications.

*How to bridge the gap between contraction measures
 and the i/o approach to consensus?*

Consensus theory connects to contraction analysis by interpreting

- Hilbert metric as a *Finsler-Lyapunov function* to study contraction
- the consensus (=Perron-Frobenius) direction as a symmetry to be factored out in the contraction analysis
- Projective contraction + row-stochasticity as horizontal contraction

Plan for this talk

- Motivation (1): positivity and networks
- Motivation (2): positivity and monotonicity
- Motivation (3): positivity on nonlinear spaces
- Motivation (4): positivity and interconnections

IEEE TRANSACTIONS ON AUTOMATIC CONTROL, VOL. 48, NO. 10, OCTOBER 2003

Monotone Control Systems

David Angeli and Eduardo D. Sontag, Fellow, IEEE

1. INTRODUCTION

ONE OF THE most important classes of dynamical systems in theoretical biology is that of *monotone systems*. Among the classical references in this area are the textbook by Smith [27] and the papers [14] and [15] by Hirsh and [26] by Smale. Monotone systems are those for which trajectories preserve a partial ordering on states. They include the subclass of *cooper-*

Definition II.1: A controlled dynamical system $\phi : \mathbb{R}_{\geq 0} \times X \times \mathcal{U}_{\infty} \rightarrow X$ is *monotone* if the following implication holds for all $t \geq 0$:

$$u_1 \succeq u_2, x_1 \succeq x_2 \Rightarrow \phi(t, x_1, u_1) \succeq \phi(t, x_2, u_2).$$

Reading the paper to the end...

Remark VIII.3. Looking at cooperativity as a notion of “incremental positivity” one can provide an alternative proof of the infinitesimal condition for cooperativity, based on the positivity of the variational equation. Indeed, assume that each system (35) is a positive linear time-varying system, along trajectories of (1). Pick arbitrary initial conditions $\xi_1 \succcurlyeq \xi_2 \in X$ and inputs $u_1 \geq u_2$. Let $\Phi(h) := \phi(t, \xi_2 + h(\xi_1 - \xi_2), u_2 + h(u_1 - u_2))$. We have (see, e.g., [28, Th. 1]) that $\phi(t, \xi_1, u_1) - \phi(t, \xi_2, u_2) = \Phi(1) - \Phi(0) = \int_0^1 \Psi'(h) dh = \int_0^1 z_h(t, \xi_1 - \xi_2, u_1 - u_2) dh$, where z_h denotes the solution of (35) when $(\partial f / \partial u)(x, u)$ and $(\partial f / \partial u)(x, u)$ are evaluated along $\phi(t, \xi_2 + h(\xi_1 - \xi_2), u_2 + h(u_1 - u_2))$. Therefore, by positivity, and monotonicity of the integral, we have $\phi(t, \xi_1, u_1) - \phi(t, \xi_2, u_2) \geq 0$, as claimed. □

We remark that monotonicity with respect to other orthants corresponds to generalized positivity properties for linearizations, as should be clear by Corollary III.3.

APPENDIX A

The importance of

Monotone Dynamical Systems

M.W. Hirsch*

Hal Smith †

We will see that the long-term behavior of monotone systems is severely limited. Typical conclusions, valid under mild restrictions, include the following:

- If all forward trajectories are bounded, the forward trajectory of almost every initial state converges to an equilibrium.
- There are no attracting periodic orbits other than equilibria, because every attractor contains a stable equilibrium.
- In \mathbb{R}^3 , every compact limit set that contains no equilibrium is a periodic orbit that bounds an invariant disk containing an equilibrium.
- In \mathbb{R}^2 , each component of any solution is eventually increasing or decreasing.

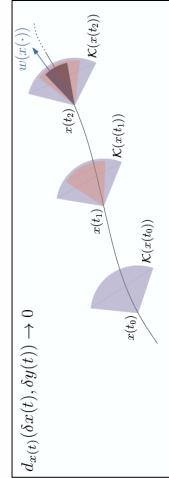
An analyst viewpoint on Perron-Frobenius theory

A differential geometric viewpoint on PF theory

VI. DIFFERENTIAL PERRON-FROBENIUS THEORY

A. Contraction of the Hilbert metric

Bushell [10] (after Birkhoff [7]) used the Hilbert metric on cones to show that the strict positivity of a mapping guarantees contraction among the rays of the cone, opening the way to many contraction-based results in the literature of positive operators [10], [30], [39], [8], [26].



B. The Perron-Frobenius vector field

The Perron-Frobenius vector of a strictly positive linear map is a fixed point of the projective space. Its existence is a consequence of the contraction of the Hilbert metric. [10]. To exploit the

The main result: the PF vector field determines the asymptotic behavior

Plan for this talk

(Theorem 3)

Suppose that the trajectories of Σ are bounded. Then, for every $\xi \in \mathcal{X}$, the ω -limit set $\omega(\xi)$ satisfies one of the following two properties:

- (i) The vector field $f(x)$ is aligned with the Perron-Frobenius vector field $w(x)$ for each $x \in \omega(\xi)$, and $\omega(\xi)$ is either a fixed point or a limit cycle or a set of fixed points and connecting arcs;
- (ii) The vector field $f(x)$ is nowhere aligned with the Perron-Frobenius vector field $w(x)$ for each $x \in \omega(\xi)$, and either $\liminf_{t \rightarrow \infty} |\partial_x \psi(t, 0, x)| w(x) | \psi(t, 0, x) = \infty$ or $\lim_{t \rightarrow \infty} f(\psi_t(x)) = 0$.

• Motivation (1): positivity and networks

• Motivation (2): positivity and monotonicity

• Motivation (3): positivity on nonlinear spaces

• Motivation (4): positivity and interconnections

22

Consensus on nonlinear spaces $\star\ddagger$

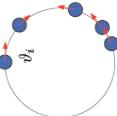
R. Sepulchre

Dept. of Electrical and Computer Engineering, Institut Montefiore, E28, Université de Liège, 4000 Liège Sart-Tilman, Belgium

A B S T R A C T

Consensus problems have attracted significant attention in the control community over the last decade. They act as a rich source of new mathematical problems pertaining to the growing field of cooperative and distributed control. This paper is an introduction to consensus problems whose underlying state-space is not a linear space, but instead a highly symmetric nonlinear space such as the circle and other relevant generalizations. A geometric approach is shown to highlight the connection between several fundamental models of consensus, synchronization, and coordination, to raise significant global convergence issues not present in linear models, and to be relevant for a number of engineering applications, including the design of planar or spatial coordinated motions.

\ddagger © 2011 Elsevier Ltd. All rights reserved.



Differential positivity and consensus on nonlinear spaces

Current work:

Positivity is the local contraction property of the consensus rule

“move towards the average of your neighbors”

Inferring the cone field from the space geometry.

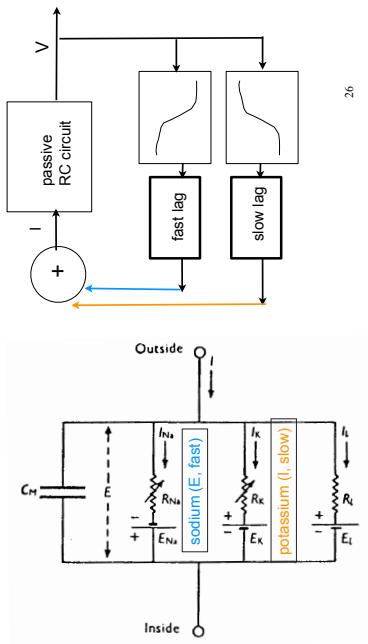
(Kuramoto model, phase synchronization, ...)

24

Plan for this talk

- Motivation (1): positivity and networks
- Motivation (2): positivity and monotonicity
- Motivation (3): positivity on nonlinear spaces
- Motivation (4): positivity and interconnections

Hodgkin-Huxley electrical circuit is a mixed feedback interconnection of monotone systems



26

25

Differential positivity and interconnection of monotone systems

Conclusion: differential positivity

=

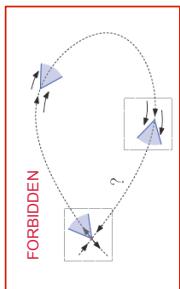
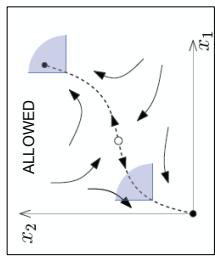
smooth patching of local orders

Current work:

Inferring the cone field from monotonicity of the blocks +
the interconnection structure
(Negative feedback is not cone-preserving)
(Biological oscillators, bursters, ...)

27

How weak is differential positivity ?



Local ordering is a weak property.
Smooth global patching is a demanding property.

**A SURVEY OF CLASSICAL AND RECENT RESULTS IN RLC
CIRCUIT SYNTHESIS****Malcolm Smith, University of Cambridge, UK**

The talk will recall some of the main results of classical circuit synthesis: Foster's Reactance theorem, the constructions of Brune and Darlington, reactance extraction, and the Bott-Duffin procedure. An introduction will also be provided to the enumerative method for RLC synthesis and the recent results of Hughes and Smith, Jiang and Smith..



A survey of classical and recent results in RLC circuit synthesis

Malcolm C. Smith

Cambridge University Engineering Department

Workshop on "Dynamics and Control in Networks"
Lund University
15–17 October 2014

Force-current analogy

Mechanical	Electrical
	 spring
	 mass
	 damper
	inductor
	capacitor
	resistor

Mass element only has one terminal—a fundamental restriction for synthesis.

What is the Most General “Passive” Suspension?

Conceptual step: replace the spring and damper with a “black-box”.



Can we characterise the properties of the “most general” such mechanism?

Question

Is it possible to construct a physical device such that the relative acceleration between its endpoints is proportional to the applied force?

$$F = b(\ddot{x}_2 - \ddot{x}_1)$$

Yes! A new word “inertier” was invented to describe such a device.

M.C. Smith, 2002, Synthesis of Mechanical Networks: The Inertier, *IEEE Trans. on Automat. Contr.*, 47, 1648–1662.

A survey of classical and recent results in RLC circuit synthesis

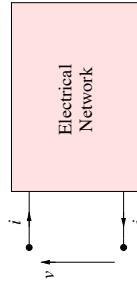
Mechanical Network Synthesis

Theorem
It is possible to build a passive mechanism **of small mass** whose impedance (velocity/force) is any rational positive-real function.

Proof
Bott-Duffin, force-current analogy + ideal inerter: $F = b(\dot{x}_1 - \dot{x}_2)$, where physical embodiments must satisfy:

- ▲ Inertance b (kg) is independent of mass;
- ▲ Inertance is independent of travel.

Driving-point impedance and admittance



Synthesis methods

LC only:

- ▲ Foster (1924)

RC and LC:

- ▲ Cauer et al.
- ▲ Brune (1931)
- ▲ Darlington (1939)
- ▲ Youla and Tissi (1966)

- RLC only:
- ▲ Bott and Duffin (1949)

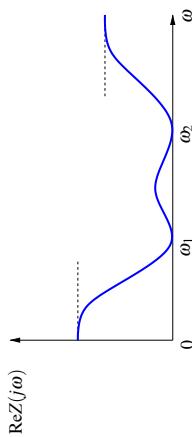
Ballscrew inerter made at Cambridge University Engineering Department (2003) - flywheel removed



Mass $\approx 1 \text{ kg}$, Inertance (adjustable) = 60–180 kg

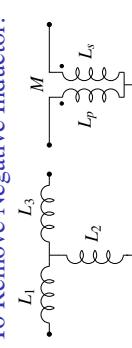
Minimum functions

A **minimum function** $Z(s)$ is a positive-real function with no poles or zeros on $j\mathbb{R} \cup \{\infty\}$ and with the real part of $Z(j\omega)$ equal to 0 at one or more frequencies.

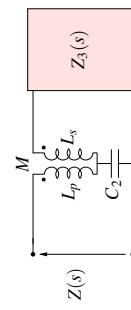


Problem: $\text{sign}(L_1 L_3) < 0$.

To Remove Negative Inductor:

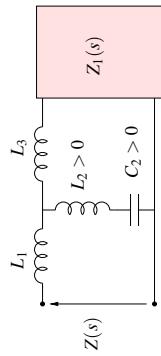


It turns out that: $L_p, L_s > 0$ and $\frac{M^2}{L_p L_s} = 1$ (unity coupling coefficient).
Realisation for completed Brune cycle:



The Brune Cycle

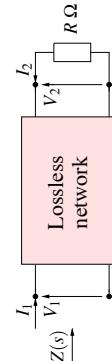
Let $Z(s)$ be a minimum function with no poles or zeros on $j\mathbb{R} \cup \{\infty\}$. It can be shown that the following decomposition is possible with Z_1 positive-real of lower degree than Z .



Problem: $\text{sign}(L_1 L_3) < 0$.

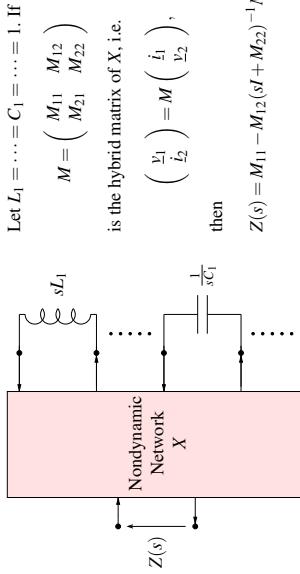
Darlington synthesis

Darlington showed that *any* positive-real $Z(s)$ could be realised by a lossless two-port (containing inductors, capacitors and transformers) terminated in a single resistor.



Darlington, S., "Synthesis of reducte 4-poles which produce prescribed insertion loss characteristics," J. Math. Phys., Vol. 18, 257-253, Sep. 1939.

Synthesis via reactance extraction

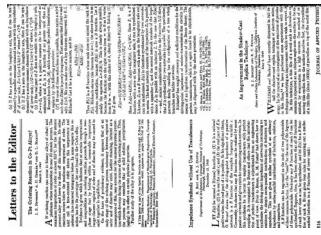


D.C. Youla and P. Tissi, "N-Port Synthesis via Reactance Extraction, Part I," *IEEE International Convention Record*, 183-205, 1966.

Lund, 15 October 2014

17

Bott-Duffin Synthesis



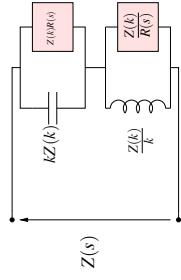
18

R. Bott and R.J. Duffin showed that transformers were unnecessary in the synthesis of positive-real functions. (1949)

Bott-Duffin Construction (cont.)

We now write:

$$\begin{aligned} Z(s) &= \frac{kZ(k)R(s) + Z(k)s}{k+sR(s)} = \frac{kZ(k)R(s)}{k+sR(s)} + \frac{Z(k)s}{k+sR(s)} \\ &= \frac{1}{\frac{Z(k)R(s)}{k+sR(s)} + \frac{s}{k+sR(s)}} + \frac{\frac{Z(k)s}{k+sR(s)}}{\frac{Z(k)R(s)}{k+sR(s)} + \frac{s}{k+sR(s)}}. \end{aligned}$$



is positive-real for any $k > 0$ (Richard's transformation).

$$R(s) = \frac{kZ(s) - sZ(k)}{kZ(k) - sZ(s)}$$

If $Z(s)$ is positive-real then

If $Z(s)$ is a minimum function with $Z(j\omega_1) = jX_1$ ($\omega_1 > 0$). (Assume $X_1 > 0$.) Then we can find a k s.t. $R(s)$ has a zero at $s=j\omega_1$.

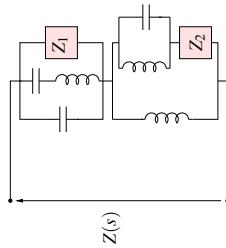
Lund, 15 October 2014

19

20

Bott-Duffin Construction (cont.)

We can write: $\frac{1}{Z(k/R(s))} = \text{const} \times \frac{s}{s^2 + \omega_0^2} + \frac{1}{R(s)}$ etc.



$$\delta(Z_1(s)) = \delta(Z_2(s)) = \delta(Z(s)) - 2 \text{ where } \delta = (\text{McMillan}) \text{ degree.}$$

New approach - the concept of a regularity

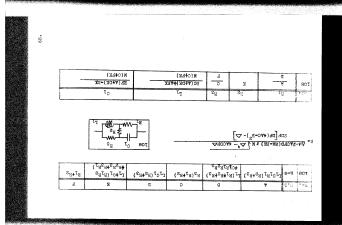
A positive-real function $Z(s)$ is defined to be *regular* if the smallest value of $\text{Re}(Z(j\omega))$ or $\text{Re}(Z^{-1}(j\omega))$ occurs at $\omega = 0$ or $\omega = \infty$.

Theorem

106 out 108 Ladehheim networks are regular.
6 series-parallel networks are a “generating set” for these 106 regular networks.
2 remaining bridge networks do not realise all the remaining biquadratic positive-real functions.

J.Z. Jiang and M.C. Smith, 2011, Regular Positive-Real Functions and Five-Element Network Synthesis for Electrical and Mechanical Networks, *IEEE Trans on Autom. Contr.*, **56**, pp. 1271–1280.

Enumerative approach—Ladenheim’s master’s thesis (1948)

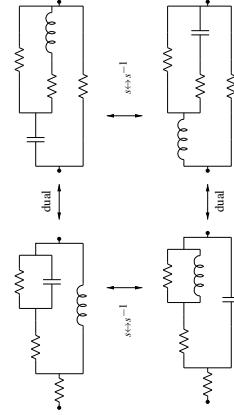


Ladenheim considered all networks with at most five elements and at most two reactive elements, and reduced the whole set to 108 networks (1948).

Questions not answered:

- What is the totality of biquadratics which may be realised?
- How many different networks are needed?

Network quartets



Rest of talk based on the recent paper:

T. H. Hughes and M. C. Smith, Algebraic criteria for circuit realisations, Mathematical System Theory—Festschrift in Honor of Uwe Helmke on the Occasion of his Sixtieth Birthday, Knut Hüper and Jochen Trumpf (eds.), CreateSpace, 2013, pp. 211–228.

Lund, 15 October 2014

A survey of classical and recent results in RLC circuit synthesis

29

M.C. Smith

Hankel matrix

Assume $Z(s)$ is proper and is realised with p inductors and q capacitors.

Suppose $n = \deg(Z(s)) = p + q$ (minimally reactive). Let

$$Z(s) = h_{-1} + \frac{h_0}{s} + \frac{h_1}{s^2} + \frac{h_2}{s^3} + \dots$$

and define the finite Hankel matrices

$$\mathcal{H}_k = \begin{bmatrix} h_0 & h_1 & \dots & h_{k-1} \\ h_1 & h_2 & \dots & h_k \\ \vdots & \ddots & \ddots & \vdots \\ h_{k-1} & h_k & \dots & h_{2k-2} \end{bmatrix}.$$

Then

$$\mathcal{H}_n = V_o \left(-\Lambda^{-1} \Sigma \right) V_o^T$$

where

$$\begin{aligned} \Lambda &= \text{diag}\{L_1, \dots, L_p, C_1, \dots, C_q\} \\ \Sigma &= (I_p + -I_q) \\ V_0 &\text{ non-singular} \end{aligned}$$

Lund, 15 October 2014

31

Synthesis via reactance extraction

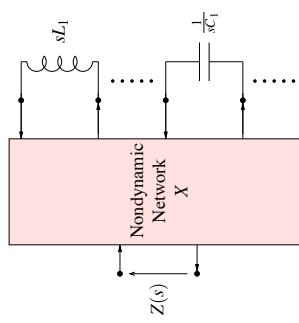
Let $L_1 = \dots = C_1 = \dots = 1$. If

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

is the hybrid matrix of X , i.e.

$$\begin{pmatrix} v \\ \frac{v}{l_1} \\ l_2 \end{pmatrix} = M \begin{pmatrix} i \\ \frac{i}{l_1} \\ \frac{i}{l_2} \end{pmatrix}$$

$$Z(s) = M_{11} - M_{12}(sI + M_{22})^{-1}M_{21}$$



D.C. Youla and P. Tessi, "N-Port Synthesis via Reactance Extraction, Part I", *IEEE International Convention Record*, 183–205, 1966.
Lund, 15 October 2014

30

M.C. Smith

Signature of the Hankel matrix

For the (proper) impedance

$$Z(s) = h_{-1} + \frac{h_0}{s} + \frac{h_1}{s^2} + \frac{h_2}{s^3} + \dots$$

where $n = \deg(Z(s))$. Then

$$\begin{aligned} p &= \# \text{ inductors} &= \# \text{ neg. eigs. of } \mathcal{H}_n \\ q &= \# \text{ capacitors} &= \# \text{ pos. eigs. of } \mathcal{H}_n \end{aligned}$$

Algebraic condition:

$$\begin{aligned} q &= \mathbf{P}(1, |\mathcal{H}_1|, \dots, |\mathcal{H}_n|), \\ p &= \mathbf{V}(1, |\mathcal{H}_1|, \dots, |\mathcal{H}_n|) \end{aligned}$$

where $\mathbf{P}(\cdot, \dots, \cdot)$ is the number of permanences of sign and $\mathbf{V}(\cdot, \dots, \cdot)$ is the number of variations of sign in the sequence.

32

Lund, 15 October 2014

Cauchy index

For a proper impedance $Z(s) \dots$

Corollary 1. $q - p = \sigma(\mathcal{H}_n)$ where σ denotes the signature.

Corollary 2. $q - p = I_{-\infty}^{+\infty} Z(s)$ where $I_{-\infty}^{+\infty} Z(s)$ is the difference between the number of jumps of $Z(s)$ from $-\infty$ to $+\infty$ and the number of jumps from $+\infty$ to $-\infty$ as s is increased in \mathbb{R} from $-\infty$ to $+\infty$ (Cauchy index).

Biquadratic functions

$$Z(s) = \frac{as_2s^2 + as + a_0}{bs_2s^2 + bs + b_0}.$$

$$|\mathcal{L}_2| = b_2a_1 - b_1a_2, \quad |\mathcal{L}_4| = (b_2a_1 - b_1a_2)(b_1a_0 - b_0a_1) - (b_2a_0 - b_0a_2)^2.$$

$$\begin{aligned} q &= \mathbf{P}(1, |\mathcal{L}_2|, |\mathcal{L}_4|), \\ p &= \mathbf{V}(1, |\mathcal{L}_2|, |\mathcal{L}_4|). \end{aligned}$$

Corollary. Whether the reactive elements are of the same or different kind is determined by the sign of the resultant $|\mathcal{L}_4|$.
Stated (without proof) by Foster (1962), as noted by Kalman (2010).

Sylvester matrix

Write

$$Z(s) = \frac{a(s)}{b(s)} = \frac{a_n s^n + a_{n-1} s^{n-1} + \dots + a_0}{b_n s^n + b_{n-1} s^{n-1} + \dots + b_0}.$$

Define the Sylvester matrices

$$\mathcal{S}_{2k} = \begin{bmatrix} b_n & b_{n-1} & \dots & b_{n-k+1} & b_{n-k} & \dots & b_{n-2k+1} \\ a_n & a_{n-1} & \dots & a_{n-k+1} & a_{n-k} & \dots & a_{n-2k+1} \\ 0 & b_n & \dots & b_{n-k+2} & b_{n-k+1} & \dots & b_{n-2k+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_n & b_{n-1} & \dots & b_{n-k} \\ 0 & 0 & \dots & a_n & a_{n-1} & \dots & a_{n-k} \end{bmatrix}.$$

Then $|\mathcal{S}_{2k}| = P_n^{2k} |\mathcal{H}_k|$ (Gantmacher).
Also, $|\mathcal{S}_{2n}| = \text{resultant of } a(s) \text{ and } b(s)$.

Thank you!

SYNCHRONIZATION AND EVOLVABILITY IN NONLINEAR NETWORKS

Jean-Jacques Slotine, Massachusetts Institute of Technology, USA

Computation, measurement, and synchronization are key issues in complex networks. Vast nonlinear networks are encountered in biology, for instance, and in neuroscience, where for most tasks the human brain grossly outperforms engineered algorithms using computational elements 7 orders of magnitude slower than their artificial counterparts. We show that nonlinear dynamic systems tools, and in particular contraction analysis, yield simple but highly non-intuitive insights about such issues, and that they also suggest systematic mechanisms to build progressively more refined networks through stable accumulation of functional building blocks and motifs.



**COMMUNITY DETECTION: FUNDAMENTAL LIMITS AND
OPTIMAL SAMPLING ALGORITHMS****Alexandre Proutiere, Kungliga Tekniska Högskolan,
Sweden**

We consider networks consisting of a finite number of non-overlapping communities. To extract these communities, the interaction between pairs of nodes may be sampled from a large available data set, which allows a given node pair to be sampled several times. When a node pair is sampled, the observed outcome is a binary random variable, equal to 1 if nodes interact and to 0 otherwise. The outcome is more likely to be positive if nodes belong to the same communities. For a given budget of node pair samples or observations, we wish to jointly design a sampling strategy (the sequence of sampled node pairs) and a clustering algorithm that recover the hidden communities with the highest possible accuracy. We consider both non-adaptive and adaptive sampling strategies, and for both classes of strategies, we derive fundamental performance limits satisfied by any sampling and clustering algorithm. In particular, we provide necessary conditions for the existence of algorithms recovering the communities accurately as the network size grows large. We also devise simple algorithms that accurately reconstruct the communities when this is at all possible, hence proving that the proposed necessary conditions for accurate community detection are also sufficient. The classical problem of community detection in the stochastic block model can be seen as a particular instance of the problems consider here. But our framework covers more general scenarios where the sequence of sampled node pairs can be designed in an adaptive manner. We provide new results for the stochastic block model, and extends the analysis to the case of adaptive sampling. This is a joint work with Se-Young Yun.



Community detection in networks

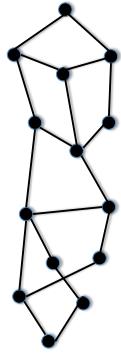
Objective: Extract K communities in a network of n nodes from random observations. $K \ll n$. Here finite K, n very large

Community Detection via Random and Adaptive Sampling

Se-Young Yun (MSRI-INRIA)
 Marc Lelarge (INRIA-ENS)
 Alexandre Proutiere (KTH)

Observations

1. A graph of interactions / similarities



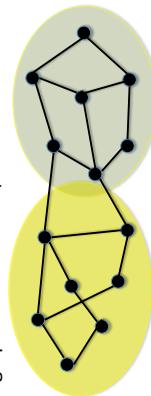
2. General sampling framework

Community detection in networks

Objective: Extract K communities in a network of n nodes from random observations. $K \ll n$. Here finite K, n very large

Observations

1. A graph of interactions / similarities



2. General sampling framework

Applications

- Social networks: recommendation systems, targeted advertisement
- Biology: the role of proteins
- Distributed computing: balanced partitions
- Communication networks: caching, pro-active resource allocation (user mobility)
- ...

Related work

Outline

Afora, Rao, Newman, Coja-Oghlan, Jerrum, Chen, Frieze, McSherry, Dyer, Sorkin, Kannan, Vempala, Vetta, Fortunato, Decelle, Krzakala, Karp, Condon, Reichart, Sanghavi, Nadakuditi, Girvan, Mosei Sly, Rohe, Chatterjee, Yu, Massoulié, Lelarge, Vazirani, Karger, Feld, Fischer, Kleinberg, Gibson, Raghavan, Hopcroft, Khan, Kulic, Santo, Wellman, Hogan, Berg, White, Boorman, Kelley, Xie, Kumar, Mathieu, Schudy, Alon, Krivelevich, Sudakov, Xu, Achlioptas, Kahale, Feige, Zdeborová, Carson, Giesen, Mitsche, Shamsi, Tsur, Hassibi, Oymak, Ames, Parrilo, Holland, Laskley, Pothen, Simon, Liou, Girvan, Chauhan, Leone, Ball, Karrer, ...

1. Community detection in the Stochastic Block Model
2. General sampling framework
 - a. Fundamental limits
 - b. Optimal Algorithms
3. Streaming, Memory Limited Algorithms

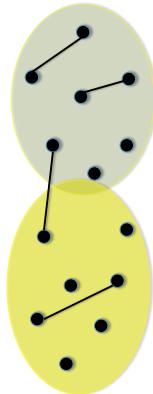
Stochastic Block Model (SBM)



1. Community Detection in the Stochastic Block Model

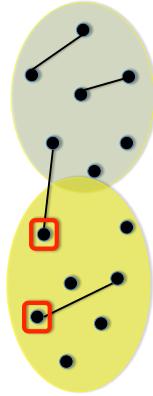
- The graph is built by considering each pair of nodes once
 - If in the same community: put an edge with probability p
 - Else: put an edge with probability $q < p$

Stochastic Block Model (SBM)



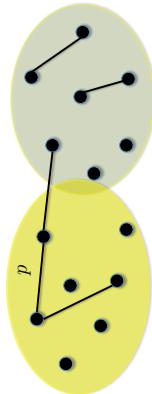
- The graph is built by considering each pair of nodes once
 - If in the same community: put an edge with probability p
 - Else: put an edge with probability $q < p$

Stochastic Block Model (SBM)



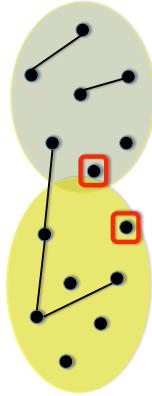
- The graph is built by considering each pair of nodes once
 - If in the same community: put an edge with probability p
 - Else: put an edge with probability $q < p$

Stochastic Block Model (SBM)



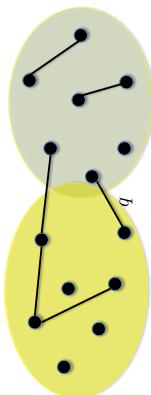
- The graph is built by considering each pair of nodes once
 - If in the same community: put an edge with probability p
 - Else: put an edge with probability $q < p$

Stochastic Block Model (SBM)



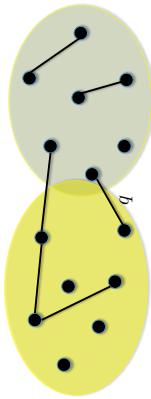
- The graph is built by considering each pair of nodes once
 - If in the same community: put an edge with probability p
 - Else: put an edge with probability $q < p$

Stochastic Block Model (SBM)



- The graph is built by considering each pair of nodes once
 - If in the same community: put an edge with probability p
 - Else: put an edge with probability $q \leq p$

Stochastic Block Model (SBM)



- Network size: n nodes, n tends to ∞
 - Sparse interaction: $p, q = o(1)$
 - Very sparse $p, q \sim 1/n$ (**Massoulie's talk**)
 - Sparse $p, q \sim f(n)/n$, $f(n) = \omega(1)$
 - Dense interaction: $p, q = O(1)$

Performance metrics

Proportion of misclassified nodes under π : $\varepsilon^\pi(n)$

- Asymptotic detection:** an algorithm *detects* the clusters if it does better than the algorithm that randomly assigns nodes to clusters

- Accurate asymptotic detection:** an algorithm π is asymptotically accurate if $\lim_{n \rightarrow \infty} \mathbb{E}[\varepsilon^\pi(n)] = 0$

Theorem (Mossel-Neeman-Sly 2012)
If $a - b < \sqrt{2(a + b)}$, then asymptotic detection is impossible.

Conjectured by Decelle-Krzakala-Moore-Zdeborova 2012

Theorem (Massoulie 2013)
If $a - b > \sqrt{2(a + b)}$, then there exists an algorithm leading to clusters that are positively correlated with the true clusters.

Asymptotic Detection in the SBM

- Two communities of equal sizes, sparse case $p = \frac{a}{n}$, $q = \frac{b}{n}$

Non-rigorous Spectral Analysis

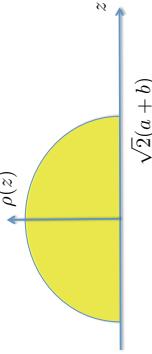
- Average adjacency matrix

$$\mathbb{E}[A] = \frac{1}{2}(a+b)11^T + \frac{1}{2}(a-b)uu^T$$

$$1 = \frac{1}{\sqrt{n}}(1, \dots, 1)^T, \quad u = \frac{1}{\sqrt{n}}(1, \dots, 1, -1, \dots, 1, -1, \dots, -1)^T$$

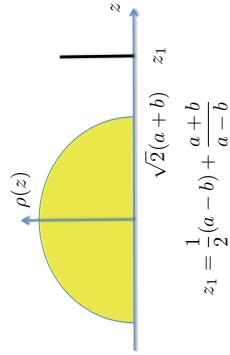
- Noisy observation: $A = \mathbb{E}[A] + X$

- Spectral density of noise matrix X (Wigner semicircle law)



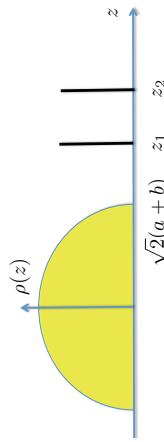
Non-rigorous Spectral Analysis

- Spectral density of the modularity matrix: $\frac{1}{2}(a-b)uu^T + X$



Non-rigorous Spectral Analysis

- Spectral density of the observed matrix:



Examples of algorithms

- Maximum Likelihood (NP hard problem)

- Exact solution: Belief Propagation
 - Compressed sensing: relaxation
- Spectral method
 - Provide a rank-K approximation of the adjacency matrix (+ Trimming + Post-processing)

- Communities are detectable if $z_1 > \sqrt{2(a+b)}$
- Method: find z_1 and the corresponding eigenvector u

Open problems

- Very sparse graphs: condition for asymptotic detection with more than two communities
- General graphs: condition on p, q, n for asymptotically accurate detection? ([this talk](#))
- General graphs: what is the optimal scaling of $\varepsilon^\pi(n)$?

2. General Sampling Framework

Sampling Framework



Sampling Strategies

- **Non-adaptive Random Strategies**
 - The pair of nodes sampled in round t does not depend on past observations, and is chosen uniformly at random
 - S1: sampling with replacement
 - S2: sampling without replacement
- **Adaptive Strategies**
 - The pair of nodes sampled in round t depends on past observations
 - Large data set available: many *samples* for the interaction of each pair of nodes
 - Sample for a given node pair: Bernoulli with mean p if nodes are in the same cluster, with mean q otherwise
 - Sample budget: $T = n(n - 1)/2$
- Classical SBM: random sampling without replacement, and

Objectives

Fundamental limits

- Performance metric: proportion of misclassified nodes $\varepsilon(n, T)$
Asymptotically accurate detection: $\lim_{n \rightarrow \infty} \mathbb{E}[\varepsilon(n, T)] = 0$
- Non-adaptive sampling:
 - Necessary conditions on n, T, p, q for the existence of asymptotically accurate algorithms
 - Asymptotically accurate clustering algorithms
- Adaptive sampling:
 - Necessary conditions on n, T, p, q for the existence of asymptotically accurate joint sampling and clustering algorithms
 - Asymptotically accurate sampling and clustering algorithms

Non-adaptive sampling:

$$\begin{aligned}\kappa_1(n, T) &= T \frac{2(n-2)}{n(n-1)} \min\{KL(q, p), KL(p, q)\} \\ &\quad + 2 \sqrt{\frac{4T(n-2)}{n(n-1)} \left[\min\{q, 1-p\} \left(\log \frac{p(1-q)}{q(1-p)} \right)^2 + \left(\log(\min\{\frac{p}{q}, \frac{1-q}{1-p}\}) \right)^2 \right]}\end{aligned}$$

Theorem Under random sampling strategy S1 or S2, for any clustering algorithm π , we have:

$$\mathbb{E}[\varepsilon^\pi(n, T)] \geq \frac{1}{8} \exp(-\kappa_1(n, T)),$$

Fundamental limits

- Non-adaptive sampling - necessary conditions for asymptotically accurate detection:

$$\frac{T}{n} = \omega(1), \quad \frac{T}{n} \min(KL(q, p), KL(p, q)) = \omega(1),$$

- Dense interaction: $p, q = \Theta(1)$

$$T(p - q)^2 / (pn) = \omega(1)$$

- Sparse interaction: $p, q = o(1)$

$$T(p - q)^2 / (pn) = \omega(1)$$

Fundamental limits

Non-adaptive sampling:

$$\kappa_1(n, T) = T \frac{2(n-2)}{n(n-1)} \min\{KL(q, p), KL(p, q)\}$$

$$+ 2 \sqrt{\frac{4T(n-2)}{n(n-1)} \left[\min\{q, 1-p\} \left(\log \frac{p(1-q)}{q(1-p)} \right)^2 + \left(\log(\min\{\frac{p}{q}, \frac{1-q}{1-p}\}) \right)^2 \right]}$$

Theorem Under random sampling strategy S1 or S2, for any clustering algorithm π , we have:

$$\mathbb{E}[\varepsilon^\pi(n, T)] \geq \frac{1}{8} \exp(-\kappa_1(n, T)),$$

Fundamental limits

Adaptive sampling:

Theorem For asymptotically accurate detection, we need:
 $\min\{p, 1-q\} \frac{T}{n} = \Omega(1)$ and $\frac{T}{n} \max(KL(q, p), KL(p, q)) = \omega(1)$.

- Example: $p = \frac{\log n}{n}$, $q = \frac{\sqrt{\log n}}{n}$
- Non-adaptive sampling: $\frac{T}{n} = \omega\left(\frac{n}{\log(n)}\right)$
- Adaptive sampling: $\frac{T}{n} = \Omega\left(\frac{n}{\log(n)}\right)$

Algorithms for non-adaptive sampling

- Spectral algorithms (extension of Coja-Oghlan's algorithm)
 1. From random samples, build an observation matrix
 2. Trimming (remove nodes with too many interactions)
 3. Spectral decomposition (find the largest eigenvalues and corresponding eigenvectors)
 4. Greedy improvement (for each node compare the number of interactions with the various clusters)

Performance

Theorem Assume that:

$$\frac{(p-q)^2}{p} \frac{\alpha T}{n} = \omega(1), \quad \frac{(p-q)^2}{p} \frac{\alpha T}{n} \geq \log(p \frac{T}{n}).$$

Then with high probability:

$$\varepsilon^{SP}(n, T) \leq 8 \exp\left(-\frac{(p-q)^2}{20p} \frac{\alpha T}{n}\right).$$

- The algorithm is asymptotically accurate under the necessary conditions for accurate detection in the case of random sampling
- The necessary conditions for accurate detection are tight!

Algorithms for adaptive sampling

- Spatial coupling idea: find reference kernels and build the clusters from these kernels
- 1. Kernels: select $n/\log(n)$ nodes and use $T/5$ samples to classify these nodes (using the previous spectral algorithm)
- 2. Select one of remaining nodes. Sample $T/3n$ pairs between the selected nodes to each kernel. Classify the node.
- 3. Repeat 2. until no remaining node or budget

Performance

Theorem Assume that:

$$\frac{(p-q)^2}{p+q} \frac{T}{n} = \Omega(1), \quad \frac{T}{n} \max(KL(q, p), KL(p, q)) = \omega(1).$$

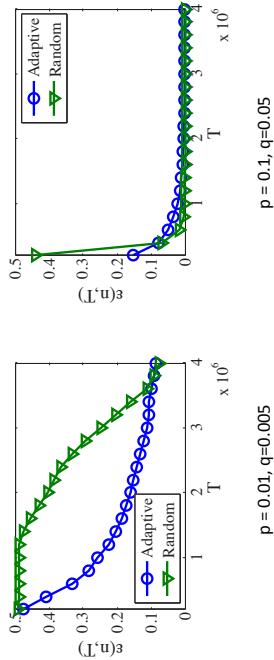
Then with high probability:

$$\varepsilon^{ASP}(n, T) \leq \exp\left(-\frac{T}{6n} (KL(q, p) + KL(p, q))\right).$$

- The algorithm is asymptotically accurate under the necessary conditions for accurate detection in the case of adaptive sampling
- The necessary conditions for accurate detection are tight!

Non-adaptive vs. adaptive sampling

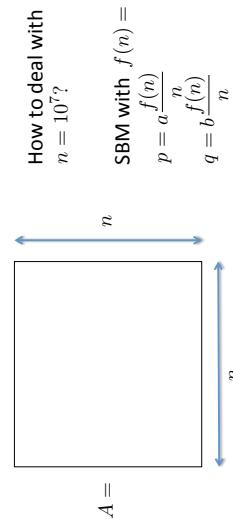
- $n = 4000$



3. Streaming, Memory-limited Algorithms

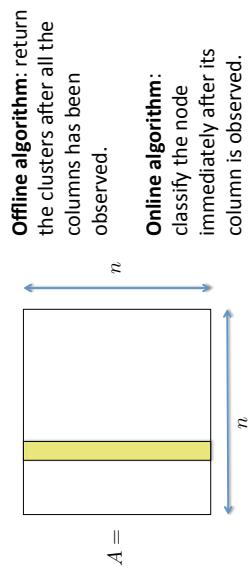
Memory and Streaming issues

- Storing and manipulating the adjacency matrix in RAM could be impossible (matlab cannot handle 5000 node networks!)
- Data about a node may arrive sequentially (one column at a time – e.g. recommender systems)



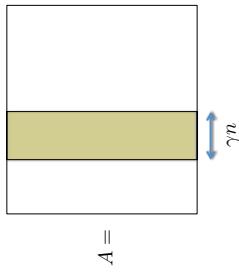
Memory and Streaming issues

- Storing and manipulating the adjacency matrix in RAM could be impossible (matlab cannot handle 5000 node networks!)
- Data about a node may arrive sequentially (one column at a time – e.g. recommender systems)



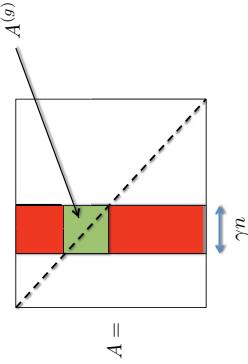
Classification with Partial Information

- Observe a proportion γ of columns
- Conditions on $\gamma, f(n)$ to classify the corresponding nodes or all nodes asymptotically accurately?



Classification with Partial Information

- Observe a proportion γ of columns
- Conditions on $\gamma, f(n)$ to classify the corresponding nodes or all nodes asymptotically accurately?



Fundamental limits

Theorem Assume that $\sqrt{\gamma}f(n) = o(1)$. Then asymptotic detection is impossible.

Algorithm for green nodes

- Indirect edges through red nodes



Theorem (i) If there is an algorithm classifying green nodes asymptotically accurately, then $\sqrt{\gamma}f(n) = \omega(1)$.
(ii) If there is an algorithm classifying all nodes asymptotically accurately, then $\gamma f(n) = \omega(1)$.

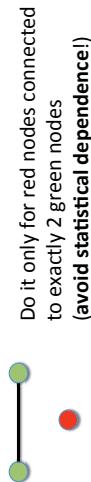
Remark: if one uses information about green nodes only, then it is possible to classify these nodes only if $\gamma f(n) = \omega(1)$. We have to use side information provided by red nodes.

Algorithm for green nodes

- Indirect edges through red nodes
- 

Algorithm for green nodes

- Indirect edges through red nodes



Algorithm for green nodes

- Indirect edges through red nodes
- 

Algorithm for green nodes

- Result: a new adjacency matrix A'
- Algorithm: spectral method for $A^{(g)}$ and A' + keep the most informative matrix (with the highest normalized K -th eigenvalue)
- Result: a new adjacency matrix A'
- Algorithm: spectral method for $A^{(g)}$ and A' + keep the most informative matrix (with the highest normalized K -th eigenvalue)

Theorem Assume that $\sqrt{f}(n) = \omega(1)$. The above algorithm classifies the green nodes asymptotically accurately.

Algorithm for red nodes

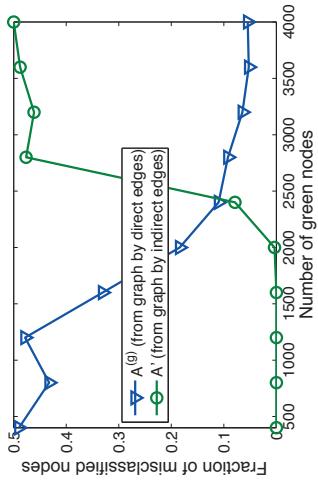
- Use the clusters of green nodes as kernels for the red nodes as in the adaptive sampling algorithm

Theorem Assume that $\gamma f(n) = \omega(1)$. The above algorithm classifies the red nodes asymptotically accurately.

The algorithms are optimal (see fundamental limits) – efficient use of side information provided by red nodes.

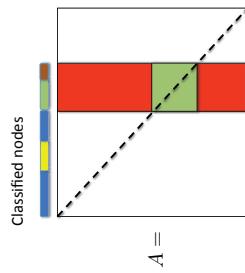
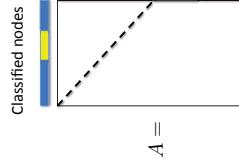
Example

$$n = 1000,000 ; p = 0.005, q = 0.001$$



Offline Memory-efficient Algorithm

- Sequentially treat blocks of columns
 - Place the block of B columns in the memory
 - Classify the corresponding nodes using previous algorithm



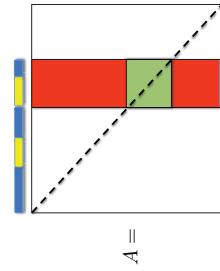
Offline Memory-efficient Algorithm

- Sequentially treat blocks of columns
 - Place the block of B columns in the memory
 - Classify the corresponding nodes using previous algorithm

Offline Memory-efficient Algorithm

- Sequentially treat blocks of columns
 - Place the block of B columns in the memory
 - Classify the corresponding nodes using previous algorithm
 - Merge the obtained clusters with the previously observed clusters, and erase the block

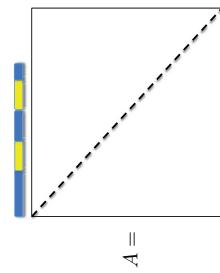
Classified nodes



Offline Memory-efficient Algorithm

- Sequentially treat blocks of columns
 - Place the block of B columns in the memory
 - Classify the corresponding nodes using previous algorithm
 - Merge the obtained clusters with the previously observed clusters, and erase the block

Classified nodes



Offline Memory-efficient Algorithm

- Memory-Performance trade-off

Theorem With block size $B = \frac{h(n)n}{\min(f(n), n^{1/3})\log(n)}$, and memory $\Theta(nh(n) + n)$.
Assume that $h(n) = \omega\left(\frac{\log(n)}{\min(f(n), n^{1/3})}\right)$, then after observing

$$T = \omega\left(\frac{n}{\min(f(n), n^{1/3})}\right) \text{ columns, we have: w.h.p.}$$

$$\varepsilon^\pi(n, T) = O\left(\exp(-T \frac{\min(f(n), n^{1/3})}{n})\right)$$

Offline Memory-efficient Algorithm

- Memory-Performance trade-off
- Example: $f(n) = \log(n)^2$
 - linear memory $h(n) = 1$
 - After observing $T = n/\log(n)$, the proportion of misclassified nodes decays faster than $1/n$

Online Memory-efficient Algorithm

Conclusions

- Sequentially treat blocks of columns
 - Place the block of B columns in the memory
 - Classify the corresponding nodes using previous algorithm
 - Merge the obtained clusters with the clusters of the first block, and erase the block
 - Return the classification for the block, and erase it
- A generic sampling framework extending the SBM
 - Necessary conditions for asymptotically accurate detection
 - Asymptotically optimal joint sampling and clustering algorithms
 - Our results hold in any regime! (Sparse or dense)
 - Memory limited, streaming algorithm
 - Everything can be done with linear memory or even sublinear memory
 - Memory-performance trade-off
- BTW, similar results for more general models than SBM
- Cloud algorithms with communication constraints?

Theorem With block size $B = \frac{h(n)n}{\min(f(n), n^{1/3})\log(n)}$, and memory $\Theta(nh(n))$.
 Assume that $h(n) = \omega\left(\frac{\log(n)}{\min(f(n), n^{1/3})}\right)$, then the above algorithm is asymptotically accurate.

Sublinear memory!

Thanks!

Papers

Yun-Proutiere, COLT 2014

Yun-Lelarge-Proutiere, NIPS 2014

COMMUNITY DETECTION IN STOCHASTIC BLOCK MODELS VIA SPECTRAL METHODS

Laurent Massoulié, Microsoft Research-INRIA Joint Centre

Community detection consists in identification of groups of similar items within a population. In the context of online social networks, it is a useful primitive for recommending either contacts or news items to users. We will consider a particular generative probabilistic model for the observations, namely the so-called stochastic block model, and generalizations thereof. We will describe spectral transformations and associated clustering schemes for partitioning objects into distinct groups. Exploiting results on the spectrum of random graphs, we will establish consistency of these approaches under suitable assumptions, namely presence of a sufficiently strong signal in the observed data. We will also discuss open questions on phase transitions for cluster detectability in such models when the signal becomes weak. In particular we will introduce a novel spectral method which provably allows detection of communities down to a critical threshold, thereby settling an open conjecture of Decelle, Krzakala, Moore and Zdeborová



COMMUNITY DETECTION IN STOCHASTIC BLOCK MODELS VIA SPECTRAL METHODS

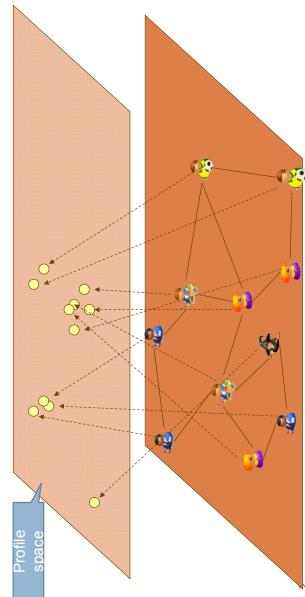
Laurent Massoulié (MSR-Inria Joint Centre, Inria)

based on joint work with:

Dan Tomozei (EPFL), Marc Lelarge (Inria), Jiaming Xu (UIUC)

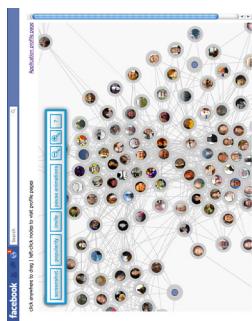
Community Detection

- identification of groups of similar objects within overall population
- Closely related objectives: clustering and embedding



Application 1: contact recommendation in
online social networks

Supporting data: e.g. OSN's friendship graph



→ recommend members of user's implicit community

Application 2: content recommendation to
users of Netflix-like system

Supporting data: user-content ratings matrix

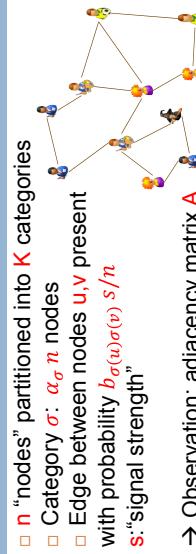
User / Movie	f_1	f_2	...	f_m
u_1	?	**		***
u_2	****	?	?	?
...				
u_n	*****	**	**	**

Use content communities to support recommendation
“users who liked this also liked...”

Outline

- The Stochastic Block Model
 - With labels
 - With general types
 - Performance of Spectral Methods
 - "rich signal" case
 - The weak signal case: sparse observations
 - Phase transition on detectability
 - A modified spectral method

The Stochastic Block Model [Holland-Laskey-Leinhardt'83]

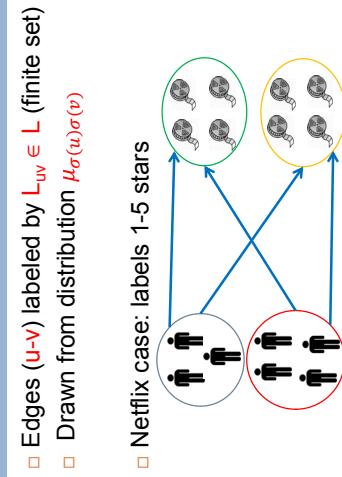


$$\rightarrow \text{Observation: adjacency matrix } \mathbf{A} = \begin{matrix} & \text{+ Noise matrix} \\ \begin{matrix} \text{A} = & \end{matrix} & \end{matrix}$$

Outline

- The Stochastic Block Model
 - The Stochastic Block Model
 - With labels
 - With general types
 - Performance of Spectral Methods
 - "rich signal" case
 - The weak signal case: sparse observations
 - Phase transition on detectability
 - A modified spectral method

The Labeled Stochastic Block Model



The SBM with general types [Aldous'81; Lovász'12]

- User type $\sigma(u)$ i.i.d. $\sim P$ in general set (e.g. uniform on $[0, 1]$)
- Edge $(u-v)$ present w.p. $b_{\sigma(u)\sigma(v)} s/m$ for "kernel" b

e.g. $b_{x,y} = F(x-y)$

Edges $(u-v)$ labeled by $L_{uv} \in \mathbb{L}$ (finite set)

Drawn from distribution $t_{\sigma(u)\sigma(v)}$

- Technical assumptions: compact type set and continuity of symmetric functions b and μ

Outline

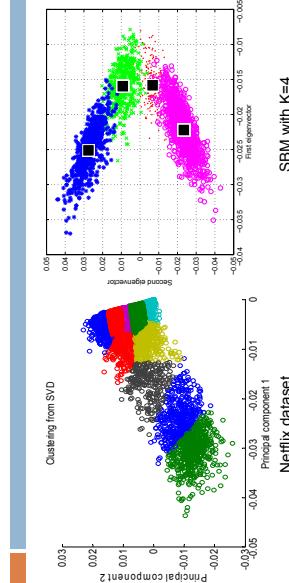
- The Stochastic Block Model
 - With labels
 - With general types
- Performance of Spectral Methods
 - "rich signal" case
 - The weak signal case: sparse observations
 - Phase transition on detectability
 - A modified spectral method

Spectral Clustering

- From Matrix A extract R normalized eigenvectors x_i corresponding to R largest eigenvalues $|\lambda_1| \geq \dots \geq |\lambda_R|$
- Form R -dimensional node representatives

$$y_u = \sqrt{i}(\alpha_i(u))_{i=1\dots R}$$
- Group nodes u according to proximity of spectral representatives y_u

Illustration for $R=2$



Result for “logarithmic” signal Strength s

Assume $s = \Omega(\log(n))$ and clusters are distinguishable, i.e.

$$\forall \sigma \neq \sigma' \exists \tau \text{ such that } b_{\sigma\tau} \neq b_{\sigma'\tau}$$

\rightarrow Then spectrum of A consists of

- R eigenvalues λ_i of order $\Omega(s)$ ($R \leq K$) and
- $n-R$ eigenvalues λ_i of order $O(\sqrt{s})$

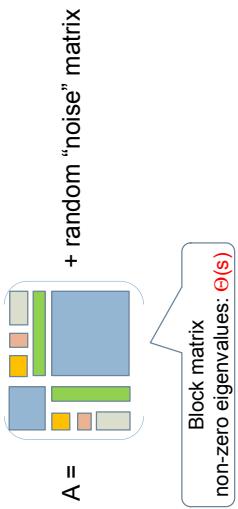
Node representatives y_u based on top R eigenvectors

x_i :

Cluster according to underlying “blocks” except for negligible fraction of nodes

Proof arguments

Control spectral radius of noise matrix
+ perturbation of matrix eigen-elements



Spectral separation properties “à la Ramanujan”

s -regular graph Ramanujan if

$$\lambda := \max(|\lambda_2|, |\lambda_n|) \leq 2\sqrt{s-1}$$

[Lubotzky-Phillips-Sarnak 88]

[Friedman'08]: random s -regular graph verifies whp

$$\lambda = 2\sqrt{s-1} + o(1)$$

[Feige-Ofek'05]: for Erdős-Rényi graph $G(n, s/n)$ and $s = \Omega(\log n)$, then whp $\lambda = O(\sqrt{s})$
Also: $\rho(A - \bar{A}) = O(\sqrt{s})$

Spectral separation properties “à la Ramanujan”

Corollary: in SBM with $s = \Omega(\log n)$, whp
 $\rho(A - \bar{A}) = O(\sqrt{s}) \rightarrow A$'s leading eigen-elements close to those of \bar{A}

For $s = \Theta(1)$, $\rho(A - \bar{A}) \sim C \sqrt{\frac{\log n}{\log \log n}}$
→ spectral separation is lost

Result for “logarithmic” signal strength \mathbf{s} – Labeled SBM

Random projection method: transform categorical labels into numerical data

For each label \mathbf{l} generate $\mathbf{W}(\mathbf{l})$ i.i.d. uniform on $[0, 1]$
Perform Spectral clustering on matrix $\{A_{ij}\mathbf{W}(L_{ij})\}$

→ Under modified distinguishability condition
 $\forall \sigma \neq \sigma', \exists \tau, \ell$ such that $b_{\sigma\tau}v_{\sigma\tau}(\ell) \neq b_{\sigma'\tau}v_{\sigma'\tau}(\ell)$

Same result holds as in unlabeled scenario

Discrepancy between SBM with small K and Netflix

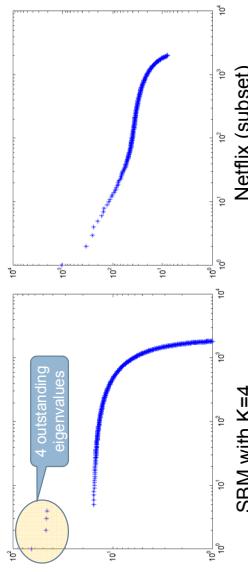
Random projection method: transform categorical labels into numerical data

For each label \mathbf{l} generate $\mathbf{W}(\mathbf{l})$ i.i.d. uniform on $[0, 1]$
Perform Spectral clustering on matrix $\{A_{ij}\mathbf{W}(L_{ij})\}$

→ Under modified distinguishability condition
 $\forall \sigma \neq \sigma', \exists \tau, \ell$ such that $b_{\sigma\tau}v_{\sigma\tau}(\ell) \neq b_{\sigma'\tau}v_{\sigma'\tau}(\ell)$

Same result holds as in unlabeled scenario

Eigenvalue distributions



→ motivates consideration of SBM with general types

SBM with general types

SBM with general types:
Spectral properties for logarithmic \mathbf{s}

- User types $\sigma(u)$ i.i.d. $\sim P$ from general set (e.g. uniform on $[0, 1]$)
- Edge $(u-v)$ present w.p. $b_{\sigma(u)\sigma(v)} \mathbf{s}/n$ for “kernel” \mathbf{b}
e.g. $b_{x,y} = F(x-y)$



- Edges $(u-v)$ labeled by $L_{uv} \in \mathbf{L}$ (finite set)
 - Drawn from distribution $\mu_{\sigma(u)\sigma(v)}$
- Form matrix $\{A_{ij}\mathbf{W}(L_{ij})\}$ from random projections $\mathbf{W}(\mathbf{l})$ of labels

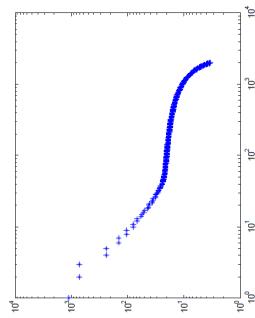
□ Eigenvalue convergence: $\mathbf{s}^{-1}\lambda_i^{(n)} \rightarrow \lambda_i$

□ Eigenvector convergence: $x_i(u) \rightarrow \varphi_i(\sigma_u)$



SBM with general types: Spectral properties for logarithmic s

- Flexible model
- power-law spectra (convolution operator + Fourier analysis)
- better matches to Netflix data



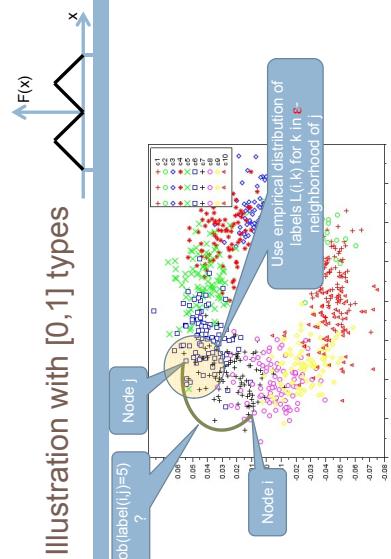
SBM with general types: estimation for logarithmic s

- For fixed R form R -dimensional node representatives

$$y_u = \sqrt{n} \left\{ \frac{\lambda_k}{\lambda_1} x_k(u) \right\}_{k=1 \dots R}$$

→ Embeds nodes according to pseudo-distance d_R that “captures geometry” of hidden node types $\sigma(u)$ with embedding accuracy controlled by “residual energy” $\varepsilon_R := \sum_{k>R} \lambda_k^{-2}$ of operator’s spectrum

Illustration with $[0, 1]$ types



Embedding allows consistent estimation of label distributions

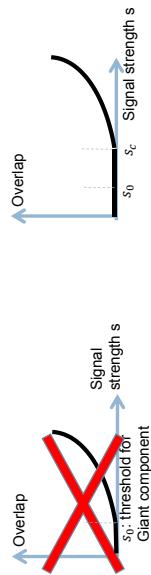
Outline

- The Stochastic Block Model
 - With labels
 - With general types
- Performance of Spectral Methods
 - “rich signal” case
 - Phase transition on detectability
 - A modified spectral method
- The weak signal case: sparse observations

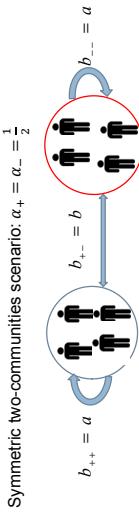
Weak signal strength: $s = \Theta(1)$

- Correct classification of all but negligible fraction of nodes impossible (isolated nodes...)
- Assess performance of clustering $\hat{\sigma}$ by overlap metric:

$$\text{ov}(\hat{\sigma}) = \frac{1}{n} \sum_{u=1}^n \{ \sigma_u = \hat{\sigma}_u \} - \max_k (\alpha_k)$$



Weak signal strength : $s=1$



Symmetric two-communities scenario: $\alpha_+ = \alpha_- = \frac{1}{2}$

Conjecture (Decelle-Krzakala-Moore-Zdeborova 2011):

For $\tau := \frac{(\alpha+b)^2}{2(\alpha+b)} < 1$, overlap tends to zero for any $\hat{\sigma}$

→ Proven by Mossel-Neeman-Sly 2012!

For $\tau > 1$, positive overlap can be achieved

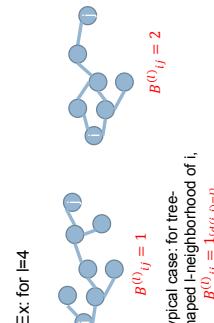
(by Belief Propagation [DKMZ 2011]; by "spectral redemption" [Zhang 2013])

No method proven to achieve positive overlap until Nov'13

Detection by modified spectral method

Form matrix $B^{(l)}; B^{(l)}_{ij} = \# \text{ of self-avoiding paths of length } l$

Ex: for $l=4$



Typical case: for tree-shaped l -neighborhood of i ,
 $B^{(l)}_{ij} = 1_{\{d(i,j)=l\}}$

Main result: spectral structure of $B^{(l)}$ for $\tau > 1$ & path length $l \sim c \log(n)$,

Let $\alpha = \frac{\alpha+b}{2}, \beta = \frac{\alpha-b}{2}$ (hence $\tau = \frac{\beta^2}{\alpha}$) eigenvalues of

$a/2$	$b/2$	$a/2$
$b/2$	$a/2$	$b/2$

- Top eigenvalue $\sim \tilde{\theta}(\alpha')$, top eigenvector \mathbf{y}^* ($y^{(B^{(l)})e} \sim |y| |B^{(l)}e|$)
- 2nd eigenvalue $\geq \tilde{\Omega}(\beta')$, 2nd eigenvector \mathbf{z}^* ($[z^{(B^{(l)})e}] \sim |z| |B^{(l)}e|$)
- 3rd eigenvalue $= O(n^\varepsilon \sqrt{\alpha'})$ for all $\varepsilon > 0$

Special separation
"à la Ramamurthy"

□ 2nd eigenvector \mathbf{z} of $B^{(l)}$ positively correlated with spin vector σ

→ Hence positive overlap obtained by estimate $\hat{\sigma}(u) = \begin{cases} +1 & \text{if } z_u \sqrt{n} > T \\ -1 & \text{if } z_u \sqrt{n} \leq T \end{cases}$

For suitable threshold T

Proof elements 1) matrix expansion

- Expected adjacency matrix
$$\bar{A} = \frac{\alpha}{n} \left[\frac{1}{2}(ce' + \sigma\sigma') - I \right] + \frac{b}{2n}(ce' - \sigma\sigma')$$

- Centered simple path adjacency matrix

$$\Delta_{ij}^{(\ell)} := \sum_{i_0 \in P_{ij}} \prod_{t=1}^{\ell} (A - \bar{A})_{i_{t-1} i_t}$$

→ Expansion: $B^{(\ell)} = \Delta^{(\ell)} + \sum_{m=1}^{\ell} (\Delta^{(\ell-m)} \bar{A} B^{(m-1)}) - \sum_{m=1}^{\ell} \Gamma_{\ell,m}$

"small" terms

"Smallness" of matrix coefficients

- Trace method: $\rho(M)^{2k} \leq \text{Trace}(M^{2k})$
 - + combinatorics (à la [Füredi-Komlós '81])
- Here: count contributions of concatenations of simple paths

→ Bounds on spectral radii: whp, for all $\varepsilon > 0$

$$\begin{aligned} \rho(\Delta^{(\ell)}) &\leq n^{\varepsilon-1} \alpha^{(\ell+m)/2}, \quad m = 1, \dots, \ell. \\ \rho(\Gamma_{\ell,m}) &\leq n^{\varepsilon-1} \alpha^{(\ell+m)/2}. \end{aligned}$$

Proof elements 2) Quasi-deterministic growth of node neighborhoods



→ then whp:

$$\begin{aligned} S_t(i) &= \alpha^{t-1} S_t(i) + \tilde{o}(\alpha^{t/2}) \\ D_t(i) &= \beta^{t-1} D_t(i) + \tilde{o}(\alpha^{t/2}) \end{aligned}$$

Proof: Chernoff bounds on binomial random variables

Corollary: For $m \leq l$, whp

$$\begin{aligned} \sup_{|x|=1, x^t B^{(l)}, e=x^r B^{(l)}, \sigma=0} |e^t B^{(m-1)} x| &= \tilde{o}(\sqrt{n} \alpha^{(m-1)/2}) \\ \sup_{|x|=1, x^t B^{(l)}, e=x^r B^{(l)}, \sigma=0} |\sigma^t B^{(m-1)} x| &= \tilde{o}(\sqrt{n} \alpha^{(m-1)/2}) \end{aligned}$$

Weak Ramanujan property

□ Previous results combined give

$$\sup_{|x|=1, x^t B^{(\ell)}, e=x^r B^{(\ell)}, \sigma=0} |B^{(\ell)} x| \leq n^{\varepsilon} \alpha^{\ell/2}.$$

$$B^{(\ell)} = \Delta^{(\ell)} + \sum_{m=1}^{\ell} (\Delta^{(\ell-m)} \bar{A} B^{(m-1)}) - \sum_{m=1}^{\ell} \Gamma_{\ell,m}$$

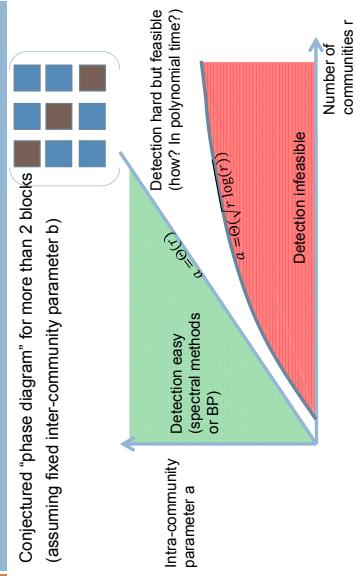
Use spectral radius bounds

$$\text{Express in terms of } e, \sigma, \bar{A} = \frac{a}{n} \left[\frac{1}{2}(ce' + \sigma\sigma') - I \right] + \frac{b}{2n}(ce' - \sigma\sigma')$$

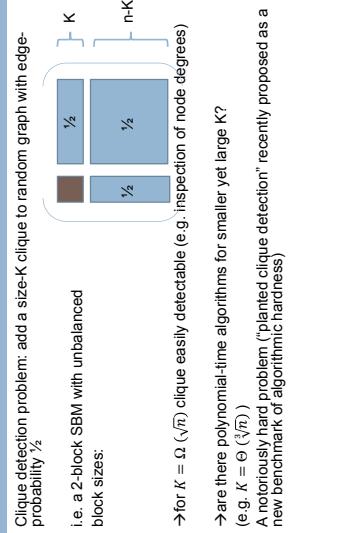
$$\begin{aligned} \text{Use bounds from quasi-deterministic growth} \\ \sup_{|x|=1, x^t B^{(\ell)}, e=x^r B^{(\ell)}, \sigma=0} |e^t B^{(m-1)} x| \\ \sup_{|x|=1, x^t B^{(\ell)}, e=x^r B^{(\ell)}, \sigma=0} |\sigma^t B^{(m-1)} x| \end{aligned}$$

Remaining mysteries about SBM's

(1)



(2)



Conclusions and Outlook

- "Vanilla" spectral methods efficient for strong (logarithmic) signal strength
- Alternatives needed at low signal strength
 - Belief propagation conjectured optimal
 - Spectral approach on path-expanded matrix proven optimal down to "easy/hard" transition
- Computationally efficient methods for "hard" cases?
 - Detection in SBM = rich playground for analysis of computational complexity with methods of statistical physics
 - Does SBM model correctly real-life data?

References

- D. Tomozei, L.M., distributed user profiling via spectral methods, ACM Sigmetrics'10
- M. Lelarge, L.M., J. Xu, Reconstruction in the labelled stochastic block model, ITW'13
- J. Xu, L.M., M. Lelarge, Edge label inference in generalized SBM: from spectral theory to impossibility results, COLT'14
- L.M., Community detection thresholds and the weak Ramanujan property, ACM STOC'14

Remaining mysteries about SBM's

SBM with general types:
estimation for logarithmic s

$$\text{Define Distance } d^2(x, y) = \int [K(x, z) - K(y, z)]^2 P(dz)$$

captures model structure

Verifies $0 \leq d_R \leq d$

$$\text{And } \iint [d^2(x, y) - d_R^2(x, y)] P(dx)P(dy) = \varepsilon_R$$

Consistency result for logarithmic s

Inference of label distribution based on

R -dimensional embedding

Empirical measures on ε -neighborhoods

For fraction of $1 - \sqrt{\varepsilon_R}$ node pairs, estimation
error verifies

$$\lim_{\varepsilon \rightarrow 0} (\lim_{\varepsilon_R \rightarrow 0} \text{Error}) = 0$$

COMPLEX ENERGY SYSTEMS**Michael Chertkov, Los Alamos National Laboratory, USA**

Today's energy systems, such as electric power grids and gas grids, already demonstrate complex nonlinear dynamics where, e.g., collective effects in one exert uncertainty and irregularities on other. These collective dynamics are not well understood and are expected to become more complex tomorrow as the grids are pushed to reliability limits, interdependencies grow, and appliances become more intelligent and autonomous. Tomorrow's will have to integrate the intermittent power from wind and solar farms whose fluctuating outputs create far more complex stress on power grid operations, often dependent, e.g. in providing fast regulation control, on the gas supply. Conversely, one anticipates significant effect of the wind-following gas fired turbines on reliability of the gas grid. Guarding against the worst of those perturbations will require taking protective measures based on ideas from optimization, control, statistics and physics.

In this talk we introduce a few of the physical, optimization and control principles and phenomena in today's energy grids and those that are expected to play a major role in tomorrow's grids.

We illustrate the new science of the energy grids on three examples: (a) discussing an efficient and highly scalable Chance Constrained Optimal Power Flow algorithm providing risk-aware control of the power transmission system under uncertainty associated with fluctuating renewables (wind farms); (b) describing effect of the intermittent power generation on reliability and compression control of the gas grid operations; and (c) briefly discussing examples of interdependencies, reliability troubles and solutions in the low level (distribution) grids.





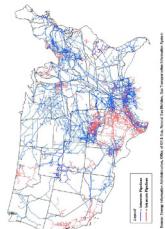
Complex Energy Systems

Misha Chertkov
LANL/DOE:OE + LANL/DTRA & NMC/NSF/ECCS

Los Alamos National Laboratory & New Mexico Consortium

Lund, Oct 15, 2014

Gas Grids



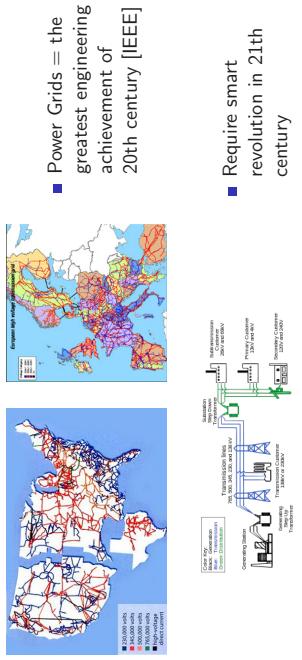
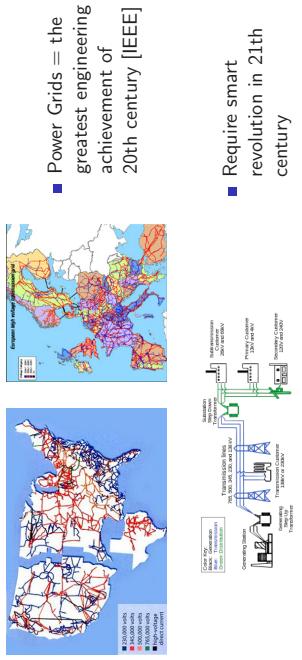
energy grids = electric+gas+heat

- Energy Hubs (local)
- Energy Interconnections (long distance)
- Vision of Future Energy Networks (green field)

- Power and Gas Grids = Basic Energy Grids
- **Fluctuations** & perturbances test the grids
- Calling for new measures and understanding of **reliability**
- Need to deal with emerging **interdependencies**

- Power Grids = the greatest engineering achievement of 20th century [IEEE]
- Require smart revolution in 21th century

Theme(s) of the talk



Complex Energy Systems

Complex Energy Systems

Power System Reliability: from Instanton to Chance Constrained OPF

Gas System Reliability: from OGF to controlling pressure fluctuations

Outline

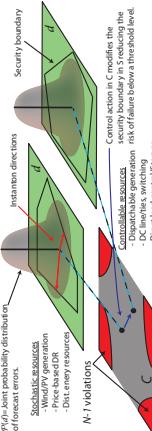
Reliability Measure of Power System Under Uncertainty



- Stochastic/uncontrollable participants
(e.g., renewables) fluctuate
- Just the standard " $N-1$ "-security gives no guarantees under uncertainty

Fault Induced Delayed Voltage Recovery: a power distribution trouble

Instantons in Power Systems: M.C. F. Pan, M. Stepanov (2010); MC, FP, MS, R. Baldick (2011); S. S. Baghsorkhi, I. Hiskens (2012)



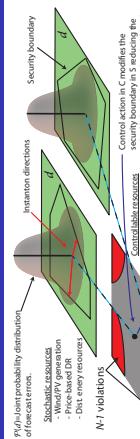
Complex Energy Systems

Power System Reliability: from Instanton to Chance Constrained OPF

Gas System Reliability: from OGF to controlling pressure fluctuations

Reliability Measure of Power System Under Uncertainty

Chance Constrained Re-dispatch



- Step one (distance to failure): compute the instantons to find the total probability of stochastic failure = P_{Fail}
- Step two: if $P_{Fail} > \text{threshold}$ re-dispatch controllable resources so that $P_{Fail} < \text{threshold}$ at minimum cost

Step two — can be built in real-time operations

- e.g. E. Karangelas, P. Panditaci, L. Wehenkel, 'Whither probabilistic security management for real-time operation of power systems?', IREP 2013

Chance Constrained Re-dispatch

Complex Energy Systems
 └ Power System Reliability: from Instanton to Chance Constrained OPF
 └ Chance Constrained Optimum Power Flows

Standard re-dispatch = Optimum Power Flow

- Or ... instead of Steps one and two one can follow another path =>
- Incorporate probabilistic security directly into optimization

D. Bienstock, MC S. Hannett (Columbia/LANL)
 R. Bent, DB. MC

CC-OPF = make sure that generation is re-dispatched at minimum cost such that $\text{failures} : P_{\text{failure}} < \text{threshold}$

Related, independent work

- E. Sjödin, D. F. Gayme and U. Topcu, *Risk-Mitigated Optimal Power Flow for Wind Powered Grids*, ACC 2012.
- L. Roald, F. Oldewurtel, T. Krause and G. Andersson, *Analytical Reformulation of Security Constrained Optimal Power Flow with Probabilistic Constraints*, Proceedings of the Grenoble PowerTech, Grenoble, France, June 2013.
- M. Vrakopoulou, K. Margellos, J. Lygeros and G. Andersson, *A Probabilistic Framework for Reserve Scheduling and N-1 Security Assessment of Systems with High Wind Power Penetration*, to appear IEEE Transactions on Power Systems.

Constrained (thermal + generation limits) OPF

$$\min_{\rho, \theta} \underbrace{c(\rho)}_{\text{cost of generation}} \quad \left| \begin{array}{l} B\theta = p - d \\ \beta_{ij}(\theta_i - \theta_j) \leq u_{ij}, \forall (i,j) \\ P_g^{\min} \leq P_g \leq P_g^{\max}, \forall g \end{array} \right. \quad \begin{array}{l} [\text{Power flows}] \\ [\text{Thermal limits}] \\ [\text{Generation constraints}] \end{array}$$

$\bullet p = \text{vector of generations } \in \mathcal{R}^n, \quad d = \text{vector of loads } \in \mathcal{R}^n$
 $\bullet B \in \mathcal{R}^{n \times n}, \quad - \quad \text{bus susceptance matrix}$



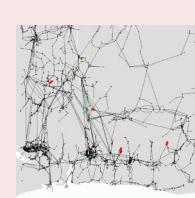
Complex Energy Systems
 └ Power System Reliability: from Instanton to Chance Constrained OPF
 └ Chance Constrained Optimum Power Flows

How does OPF handle (renewable) fluctuations?

Want to improve the standard OPF

Complex Energy Systems
 └ Power System Reliability: from Instanton to Chance Constrained OPF
 └ Chance Constrained Optimum Power Flows

- standard automatic control [affine, possibly changing rates]
 - aware of security (limits)
 - not too conservative
 - computationally practicable



Complex Energy Systems
 └ Power System Reliability: from Instanton to Chance Constrained OPF
 └ Chance Constrained Optimum Power Flows

Experiment: Bonneville Power Administration data, Northwest US

- Automatic frequency control: primary [seconds] + secondary [AGC, 1-2 minute]
- Generator output varies up or down proportionally to aggregate change

Complex Energy Systems	Power System Reliability: from Instanton to Chance Constrained OPF		
	Power Flow Eqs.	[for mean forecast]	
	Chance Constraints	Generation satisfies Chance Constraints	

OPF vs Chance Constrained-OPF

Standard OPF (Dispatch for the mean forecast, not aware of fluctuations)

$$\min_p \quad c(p) \quad \text{cost of generation}$$

Power Flow Eqs.	Chance Constraints for Generation
Generation limits	Chance Constraints for line Flows

Chance Constrained OPF (fluctuations aware dispatch)

$$\min_{\bar{p}, \alpha} \mathbb{E}[c(\bar{p}, \alpha)] \quad \text{Power Flow Eqs. [for mean forecast]}$$

Power Flow Eqs.	Chance Constraints for Generation
Generation limits	Chance Constraints for line Flows

Chance Constraints for Line Flows:

$$\forall (i, j) \in \mathcal{E} : \quad \text{Prob}(|f_{ij}| > f_{ij}^{\max}) < \varepsilon_{ij}.$$

Interpretation: overload is allowed for ϵ -fraction of "time".

- \bar{p} - generation re-dispatch for beginning of the period: α - proportional rates for the period

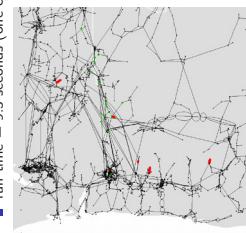
Back to motivating example

Complex Energy Systems	Power System Reliability: from Instanton to Chance Constrained OPF		
	Power Flow Eqs.	[for mean forecast]	
	Chance Constraints	Generation satisfies Chance Constraints	

Back to motivating example

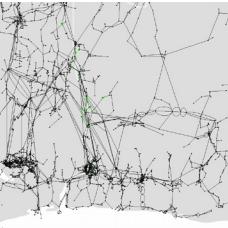
BPA case

- standard OPF: cost 235603, 7 lines unsafe $\geq 8\%$ of the time
- CC-OPF: cost 237297, every line safe $\geq 98\%$ of the time
- run time = 9.5 seconds (one cutting plane!)



BPA case

- standard OPF: cost 235603, 7 lines unsafe $\geq 8\%$ of the time
- CC-OPF: cost 237297, every line safe $\geq 98\%$ of the time
- run time = 9.5 seconds (one cutting plane!)



Experiments with CC-OPF

- CC-OPF succeeds where standard OPF fails
- Cost of Reliability [CC-OPF saving over standard OPF]
- CC-OPF is not a naive fix. [Changes are nonlocal]
- ...

Performance of the Method



Enhancements of CC-OPF

Performance of cutting-plane method on a number of large cases.

Case	Buses	Generators	Lines	Time (s)	Iterations	Barrier iterations
BPA	2209	176	2866	5.51	2	256
Polish1	327	327	2896	13.64	13	535
Polish2	2746	388	3514	30.16	25	1431
Polish3	3120	349	3693	25.45	23	508

Rapid convergence on realistic networks

Typical convergence behavior of cutting-plane algorithm on a large instance.

Iteration	Max rel. error	Objective
1	1.2e-1	7.0933e6
4	1.3e-3	7.0934e6
7	1.9e-3	7.0934e6
10	1.0e-4	7.0964e6
12	8.9e-7	7.0965e6

► [details](#)



Summary (CC-OPF) + Extensions:

- DC PF + affine control + independent fluctuations \Rightarrow conic (tractable) optimization
- Specialized cutting-plane algorithm proves effective
- Commercial solvers do not
- Algorithm efficient even in cases with thousands of buses/lines
- Algorithm can be made robust with respect to data errors
- Allows to account for synchronization constraints

Path Forward (work in progress)

- AC generalizations, convexifications (e.g. FDPF, Energy Function based appr.)
- Ramp Constraints
- Multi-stage CC-OPF
- combined CC-OPF & Unit Commitment



Out of Sample Tests – can handle either of the two cases

- True distribution is non-Gaussian, but our Gaussian distribution is close
- Parameters of the Gaussian distributions, μ_i, σ_i^2 are mis-estimated
- Robust (ambiguous) CC-OPF
- CC-OPF which is robust with respect to parameters of the Gaussian distribution from a range
- Allows convex tractable reformulation

"Non-linear" OPF & sync-CC OPF [R. Bent, D. Bienstock, MC 2013]

- Convex AC-OPF (lossless, constant voltage - based on an exercise from Boyd, Vandenberghe [book], ► [details](#))
- Synchronization constrained CC-OPF (based on MC work with F. Dorfler & F. Bullo [PNAS, 2013], ► [details](#))
- Voltage constraints — work in progress



- Consider CC-OPF, or other type of dispatch, responding to uncertainty/wind
- Gas turbines [fast to ramp up/down, relatively clean] are
- producers of electricity, follow wind
- consumers of gas: inducing/transferring fluctuations/stress to the gas network

- Study interdependences ...
- Start from analysis of gas system under uncertainty (e.g. caused by the wind-induced correlations)

2 min crash course on the hydro (gas) dynamics

- single pipe; not tilted (gravity is ignored); constant temperature
- ideal gas, $\rho \sim \rho$ – pressure and density are in a linear relation
- all fast transients are ignored – gas flow velocity is significantly slower than the speed of sound, $u \ll c_s$
- turbulence is modeled through turbulent friction; mass flow, $\phi = u\rho$, are averaged across the pipe's crosssection

$$\left. \begin{array}{l} \partial_t \rho + \partial_x(u\rho) = 0 \\ \text{conservation of mass} \\ \partial_t(\rho u) + \partial_x(\rho u^2) + \partial_x \rho = -\frac{\rho u |u|}{2D} f \\ \text{conservation of energy} \end{array} \right\} \Rightarrow$$

Approximations ... allowing to resolve flows analytically (lump description)

Stationary, balanced regime [standard]

$\phi = \text{const.}$, $p_m^2 = x\beta\phi/\phi|/D$

$\phi = \phi_{\text{std}}(x) + \delta\phi(t,x)$, $p = p_{\text{std}}(x) + \delta p(t,x)$

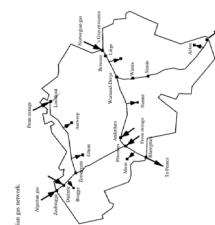
Unbalanced, linearized line-pack [non-standard]

$\phi = \phi_{\text{std}}(x) + \delta\phi(t,x)$

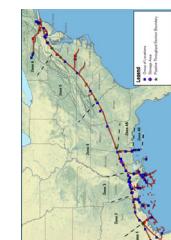
Complex Energy Systems
└ Gas System Reliability: from OGF to controlling pressure fluctuations

Gas Flows. Steady (balanced) Case.

Optimal Gas Flow



Relief gas network.



$$\begin{aligned} &\text{with compressors, } \alpha_{ij} \leq 1 \\ &p_i^2 \xrightarrow{r} \alpha_{ij} \xrightarrow{(1-r)} p_j^2 \end{aligned}$$

$$\begin{aligned} &\text{Gas Flow Equations: } (\sum_i q_i) = 0, \quad \bar{q}_j = L_j \beta_{ij} / D_{ij} \\ &\forall (i,j) : \quad \alpha_{ij}^2 = \frac{\beta_{ij}^2 + (1-r)\beta_{ij}\phi_{ij}^2}{\rho_i^2 - r\alpha_{ij}\phi_{ij}^2} \\ &\forall i : \quad q_i = \sum_{j:(i,j) \in \mathcal{E}} \bar{q}_j - \sum_{j:(j,i) \in \mathcal{E}} \phi_{ij} \end{aligned}$$

$$\begin{aligned} &0 < m = (\gamma - 1)/\gamma < 1, \gamma \text{ gas heat capacity ratio (thermodynamics)} \\ &m u_{ij} \phi_{ij} |\phi_{ij}| (1 - r) u_{ij} \phi_{ij} |\phi_{ij}| \end{aligned}$$

- The problem is convex on trees (many existing gas transmission systems are trees) ⇔ through Geometric Programming (log-function transformation)

Minimizing the cost of compression (~ work applied externally to compress)

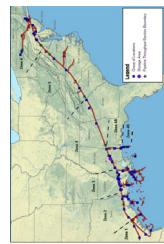
$$\begin{aligned} &\min_{\alpha_{ij}} \sum_{(i,j)} \alpha_{ij} \left(\frac{\sum_i \phi_{ij}}{n_{ij}} \left(\alpha_{ij}^m - 1 \right) + \right. \\ &\quad \left. \forall (i,j) : \quad \alpha_{ij}^2 = \frac{\beta_{ij}^2 + (1-r)\alpha_{ij}\phi_{ij}^2}{\rho_i^2 - r\alpha_{ij}\phi_{ij}^2} \right. \\ &\forall i : \quad 0 \leq \underline{p}_i \leq p_i \leq \bar{p}_i \\ &\forall (i,j) : \quad \alpha_{ij} \leq \bar{\alpha}_{ij} \end{aligned}$$

- $q_i = \sum_{j:(i,j) \in \mathcal{E}} \bar{q}_j - \sum_{j:(j,i) \in \mathcal{E}} \phi_{ij}$

Gas Flows. Steady (balanced) Case.



Relief gas network.



Gas System Reliability: from OGF to controlling pressure fluctuations

Gas Flow Equations:

$$\sum_i q_i = 0, \quad \bar{q}_j = L_j \beta_{ij} / D_{ij}$$

$$\forall (i,j) : \quad \alpha_{ij}^2 = \frac{\beta_{ij}^2 + (1-r)\beta_{ij}\phi_{ij}^2}{\rho_i^2 - r\alpha_{ij}\phi_{ij}^2}$$

$$\forall i : \quad q_i = \sum_{j:(i,j) \in \mathcal{E}} \bar{q}_j - \sum_{j:(j,i) \in \mathcal{E}} \phi_{ij}$$

Gas System Reliability: from OGF to controlling pressure fluctuations

Gas Flow Equations:

$$\forall (i,j) : \quad \alpha_{ij}^2 = \frac{\beta_{ij}^2 + (1-r)\beta_{ij}\phi_{ij}^2}{\rho_i^2 - r\alpha_{ij}\phi_{ij}^2}$$

$$\forall i : \quad q_i = \sum_{j:(i,j) \in \mathcal{E}} \bar{q}_j - \sum_{j:(j,i) \in \mathcal{E}} \phi_{ij}$$

- Gas Flow Equations: $(\sum_i q_i) = 0, \quad \bar{q}_j = L_j \beta_{ij} / D_{ij}$

$$\forall (i,j) : \quad \alpha_{ij}^2 = \frac{\beta_{ij}^2 + (1-r)\beta_{ij}\phi_{ij}^2}{\rho_i^2 - r\alpha_{ij}\phi_{ij}^2}$$

$$\forall i : \quad q_i = \sum_{j:(i,j) \in \mathcal{E}} \bar{q}_j - \sum_{j:(j,i) \in \mathcal{E}} \phi_{ij}$$

- $q_i = \sum_{j:(i,j) \in \mathcal{E}} \bar{q}_j - \sum_{j:(j,i) \in \mathcal{E}} \phi_{ij}$

- S. Misra, M. W. Fisher, S. Backhaus, R. Bent, M.C. F. Pan, Optimal compression in natural gas networks: a geometric programming approach, IEEE TCNS 2014

Complex Energy Systems
 ↳ Gas System Reliability: from OGF to controlling pressure fluctuations

Re-cap of the OGF & pressure reliability studies

- Geometric programming offers an efficient way of solving the steady/balanced OGF over tree structures
- Dynamic and stochastic GF (with line pack) is solved perturbatively.
 - Shows diffusive, spatially inhomogeneous line pack jitter – extremal at the points of the flow reversal.
- Path Forward
 - Steady OGF over graphs with loops
 - Other OGF formulations, e.g. max-throughput
 - Generalize stochastic-line-pack-GF to discrete models with loops
 - Extend it to reliability-aware stochastic and non-stationary optimizations

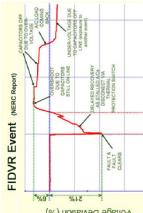
Complex Energy Systems
 ↳ Fault Induced Delayed Voltage Recovery: a power distribution trouble

Recorded Distribution/Transmission Voltage Events

- TVA, Blistering Sat of Aug 22, 1987: Cascading Voltage Collapse in West Tennessee. Fault at 115kV switch. Cleared in 1s. Continued into 161kV and 600kV lines for 10-15s. Resulted in the loss of 700MW in Memphis. Motor Jitter started and drawn large amount of reactive power even after the fault was cleared.
- 1988 event in Florida reported in "Air Conditioner Response to Transmission Fault" by J. W. Shaffer in 1987 ... In the last ten years there have been at least eight events in which normally loaded (in 2-3 cycles), multi-phase events in Southern Florida have caused a significant drop in customer load (20-40MW).
- 1990 Egypt ... 1999 metro area Atlanta, Arizona, Southern California ... NERC Planning Committee White Paper on "Fault Induced Delayed Voltage Recovery" by Transmission Issues Subcommittee

- delays (between cause and the result)
- nonlinearity of loads plays a significant role
- many inductive motors **simultaneously** affected
- initiated (fault) at the **transmission-to-distribution** interface, matures within distribution, cascades into transmission

Fault in Southern California
 Time



Fault in Southern California
 Time

Complex Energy Systems
 ↳ Fault Induced Delayed Voltage Recovery: a power distribution trouble

Modeling extended FIDVR

- Observed in feeders with many induction motors (air-conditioning)
- Uncontrolled depressed voltage can spread causing a larger outage
- Hypotheses (Hiskens, Leshture, Chassin, ...): the events are caused by many air conditioners stalled
- Modeling the event is a challenge

courtesy of D. Kostyrev and B. Leshture
 Our Contribution - Modeling of FIDVR over extended feeder

- Observation (simulations – consistent with measurements): soliton-like propagation of "stalled" phase/front
- Coarse-grained (reduced) PDE modeling of the "extended" FIDVR

Fault in Southern California
 Time



courtesy of D. Kostyrev and B. Leshture
 Our Contribution - Modeling of FIDVR over extended feeder

- However, ...
 - the faults (in spite of being cleared fast) may be of **high risks for power distribution**

Complex Energy Systems
 ↳ Gas System Reliability: from OGF to controlling pressure fluctuations

- Fluctuations in power sources (renewable and interdependences) lead to more frequent (then in the past) interruptions
- In particular **voltage faults**
- Cleared faults do not cause major direct effect on balanced transmission

Complex Energy Systems
 ↳ Gas System Reliability: from OGF to controlling pressure fluctuations

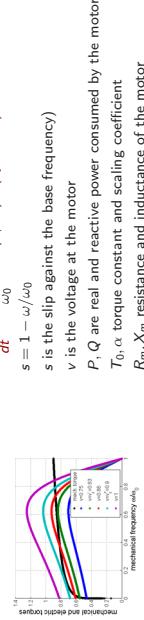
Modeling Individual Motor Popovic, Hiskens, Hill '98

Complex Energy Systems
 └ Fault Induced Delayed Voltage Recovery: a power distribution trouble
 └ Individual Motor: Two bus. Hysteresis.

Modeling Individual Motor Popovic, Hiskens, Hill '98

minimal model of the motor
 Complex Energy Systems
 └ Fault Induced Delayed Voltage Recovery: a power distribution trouble
 └ Individual Motor: Two bus. Hysteresis.

$$\begin{aligned} P &= \frac{sR_m v^2}{R_m^2 + s^2 X_m^2} \\ Q &= \frac{s^2 X_m v^2}{R_m^2 + s^2 X_m^2} \\ \mu \frac{d}{dt} \omega &= \frac{P}{\omega_0} - T_0 (\omega/\omega_0)^\alpha \quad (\text{dynamics}) \\ s &= 1 - \omega/\omega_0 \end{aligned}$$



Explanation for "jumped" FIDVR

- Hysteresis: The motor is trapped in the stalled (low-voltage) state!
- First order phase transition. Bifurcation (stability). Spinodal points.

Explanation for "jumped" FIDVR

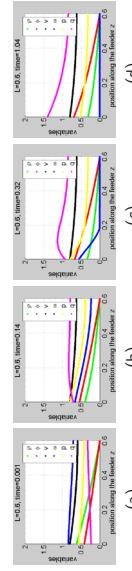
- Hysteresis: The motor is trapped in the stalled (low-voltage) state!
- First order phase transition. Bifurcation (stability). Spinodal points.

Feeder with Many (distributed) Inductive Motors

Complex Energy Systems
 └ Fault Induced Delayed Voltage Recovery: a power distribution trouble
 └ Feeder with Many Distributed Motors

Dynamics/Transitions in an Extended Feeder (I)

Example of a Large Fault → feeder is stalled (Movie Large Fault)



- Pre fault
- (b) Immediately past fault
- (c) Later in the process
- (d) The feeder is fully stalled

Spatially-continuous version of
 Dist Flow [Baran, Wu (1989)]
 Reduced model of the "extended" feeder
 Easy to analyze dynamics: PDE.

$$\begin{aligned} \partial_x \rho &= -\rho - r \frac{\rho^2 + \phi^2}{v^2} \\ \partial_x \phi &= -q - x \frac{\rho^2 + \phi^2}{v^2} \\ v \partial_z v &= -(r\rho + q)\phi \\ \rho &= \frac{sR_m v^2}{R_m^2 + s^2 X_m^2} \\ q &= \frac{s^2 X_m v^2}{R_m^2 + s^2 X_m^2} \\ \mu \frac{d}{dt} \omega &= \frac{\rho}{\omega_0} - \gamma_0 \left(\frac{\omega}{\omega_0} \right)^\alpha \\ \nu(0) = 1, \rho(L) &= \phi(L) = 0 \end{aligned}$$

Dynamics/Transitions in an Extended Feeder (II)

Example of a Small Fault → feeder is partially stalled (Movie Small Fault)

(a) Immediately past fault
 (b) Later in the process
 (c) Front advances
 (d) Stabilized, part. stalled
 • dynamics of restoration

Complex Energy Systems

- ▀ Fault Induced Delayed Voltage Recovery: a power distribution trouble
- ▀ Feeder with Many Distributed Motors

Complex Energy Systems

- ▀ Fault Induced Delayed Voltage Recovery: a power distribution trouble
- ▀ Recap of the "extended" FIDVR story

The 1+1 (space+time) continuous model of distribution

- ▀ Integrating multiple bi-stable individual motors into power flow
- ▀ Emergence of multiple spatially-extended states/transitions

Hysteresis

Sof-Similar Transients

... to do ...

- ▀ Inhomogeneity (disorder), stochasticity (noise); what is the probability that the feeder with a given level of disorder will recover?
- ▀ Effects of other devices, e.g. distributed generation and control (PV) ...
- ▀ Possible cascade – from feeder to feeder (within substation) ... to transmission
- ▀ What is the least control effort needed to avoid a FIDVR event/cascade following a given type of fault? ⇔ Voltage control through
 - ▀ tap changers, e.g. incorporating in the framework of "primary voltage control of active distributed networks" by Christakou, Tomozei, Le Boudec, Paolone (2014)
 - ▀ inverters [distributed power electronics]

Conclusions Drawn from Experiments/Numerics concern

Take Home high-level message: illustrated today on a few examples

New Science of Complex Power/Energy System Engineering

- ▀ Old approach focused on individual devices, deterministic – remained valid ... but
- ▀ New complexities (renewables, fluctuations, interdependencies) need to be controlled through ...
- ▀ Better understanding of the multi-scale, probabilistic science (applied physics/math, operation research, other IT disciplines) ...
- ▀ To enable better practical control and optimization of tomorrow grids

Chance Constrained OPF

- ▀ New reliability measure for uncertainty – the instants
- ▀ Efficient and Scalable
- ▀ Chance-Constrained OPF

Gas Reliability with Line Pack

- ▀ Geometric/Signomial Programming for Optimum Gas Flow
- ▀ Line-Pack Jitter of Pressure Fluctuations

Distributed Fault induced Delayed Voltage Recovery

- ▀ Soliton-like front describing FIDVR
- ▀ PDF approach ⇒ coarse-grained estimation

DC-approximation

Computing line flows

Frequency Control (quasi-static proxy)

Take Home High-level message illustrated today on a few examples

New Science of Complex Power/Energy System Engineering

DC approximation

Complex Energy Systems
└ Appendix: CC-OPF
└ DC approximation

Complex Energy Systems
└ Appendix: CC-OPF
└ CC-OPF – details

Complex Energy Systems
└ Appendix: CC-OPF
└ Chance Constrained OPF

Complex Energy Systems
└ Appendix: CC-OPF
└ CC-OPF – details

Complex Energy Systems
└ Appendix: CC-OPF
└ Chance Constrained OPF

Complex Energy Systems
└ Appendix: CC-OPF
└ Chance Constrained OPF

- Old approach focused on individual devices, deterministic – remained valid ... but
- New **complexities** (renewables, fluctuations, interdependencies) need to be controlled through ...
- Better understanding of the **multi-scale, probabilistic science** (applied physics/math, operation research, other IT disciplines) ...
- To enable better practical control and optimization of tomorrow grids

Thank You!

For each generator i , two parameters:

- \bar{p}_i = mean output
- α_i = response parameter

Real-time output of generator i :

$$\rho_i = \bar{p}_i - \alpha_i \sum_j \Delta \omega_j$$

where $\Delta \omega_j$ = change in output of renewable j (from mean).

$$\sum_i \alpha_i = 1$$

~ primary + secondary control

◀ Chance Constrained OPF

- (0) The amplitude of the complex potentials are all fixed to the same number (unity, after trivial re-scaling) $\forall a : u_a = 1$.
- (1) $\forall \{a, b\} : |\varphi_a - \varphi_b| \ll 1$ - phase variation between any two neighbors on the graph is small
- (2) $\forall \{a, b\} : \rho_{ab} \ll x_{ab}$ - resistive (real) part of the impedance is much smaller than its reactive (imaginary) part. Typical values for the r/x is in the $1/27 \sim 1/2$ range.
- (3) $\forall a : \rho_a \gg q_a$ - the consumed and generated powers are mainly real, i.e. reactive components of the power are much smaller than their real counterparts

It leads to

- Linear relation between powers and phases (at the nodes): $\hat{B}\varphi = \rho$
- Losses are ignored: $\sum_a \rho_a = 0$
- B - graph Laplacian constructed of line susceptances

◀ DC-OPF

wind power at bus i : $\mu_i + \mathbf{w}_i$

DC approximation

- $B\theta = \bar{\rho} - d$
- $+ (\mu + \mathbf{w} - \alpha \sum_{j \in G} \mathbf{w}_j)$
- $\theta = B^+ (\bar{\rho} - d + \mu) + B^+ (I - \alpha e^T) \mathbf{w}$
- flow is a linear combination of bus power injections:

$$\mathbf{f}_{ij} = \beta_{ij}(\theta_i - \theta_j)$$

◀ Chance Constrained OPF

Computing line flows

Complex Energy Systems
[└ Appendix: CC-OPF](#)
[└ CCC-OPF -details](#)

$$\begin{aligned}\mathbf{f}_{ij} &= \beta_{ij} \left((B_i^+ - B_j^+)^\top (\bar{p} - d + \mu) + (A_i - A_j)^\top \mathbf{w} \right), \\ A &= B^+ (I - \alpha \mathbf{e}^\top)\end{aligned}$$

Given distribution of wind can calculate moments of line flows:

$$E\mathbf{f}_{ij} = \beta_{ij} (B_i^+ - B_j^+)^\top (\bar{p} - d + \mu)$$

$$\text{var}(\mathbf{f}_{ij}) := s_{ij}^2 \geq \beta_{ij}^2 \sum_k (A_{ik} - A_{jk})^2 \sigma_k^2$$

(assuming independence)

■ and higher moments if necessary

[◀ Chance Constrained OPF](#)

Complex Energy Systems
[└ Appendix: CC-OPF](#)
[└ CCC-OPF -details](#)

Formulation [convex]:

Choose mean generator outputs and control to minimize expected cost, with the probability of line overloads kept small.

$$\begin{aligned}\min_{\bar{\rho}, \alpha} \quad & \mathbb{E}[c(\bar{\rho})] \\ \text{s.t.} \quad & \sum_{i \in G} \alpha_i = 1, \quad \alpha \geq 0 \\ & B\delta = \alpha, \delta_n = 0 \\ & \sum_{i \in G} \bar{\rho}_i + \sum_{i \in W} \textcolor{red}{\mathbf{f}_{ij}} = \sum_{i \in D} d_i \\ & \bar{f}_{ij} = \beta_{ij}(\bar{\theta}_i - \bar{\theta}_j), \\ & B\bar{\theta} = \bar{p} + \textcolor{blue}{\mathbf{u}} - d, \quad \bar{\theta}_n = 0 \\ & s_{ij}^2 \geq \beta_{ij}^2 \sum_{k \in W} \textcolor{red}{\sigma_k^2} (B_{ik}^+ - B_{jk}^+ - \delta_i + \delta_j)^2 \\ & |\bar{f}_{ij}| + s_{ij} \phi^{-1}(1 - \epsilon_{ij}) \leq \bar{f}_{ij}^{\max}\end{aligned}$$

[◀ Chance Constrained OPF](#)

Chance constraints to deterministic constraints

- chance constraint: $P(\bar{f}_{ij} > f_{ij}^{\max}) < \epsilon_{ij}$ **and** $P(\bar{f}_{ij} < -f_{ij}^{\max}) < \epsilon_{ij}$
- from moments of \mathbf{f}_{ij} , can get conservative approximations using e.g. Chebyshev's inequality

- for Gaussian wind, can do better, since \mathbf{f}_{ij} is Gaussian :

$$|E\mathbf{f}_{ij}| + \text{var}(\mathbf{f}_{ij}) \phi^{-1}(1 - \epsilon_{ij}) \leq \bar{f}_{ij}^{\max}$$

[◀ Chance Constrained OPF](#)

Cutting-Plane Method

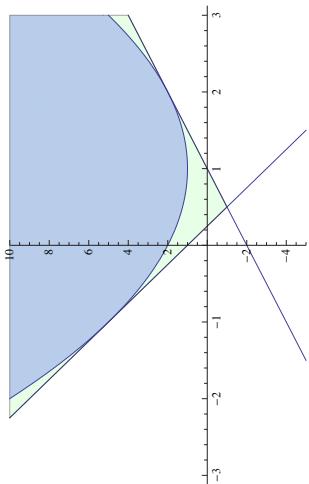
New Solutions still violates conic constraint

Complex Energy Systems
[└ Appendix: CC-OPF](#)
[└ Cutting-Plane Method](#)

[◀ Chance Constrained OPF](#)

Cutting-Plane Method

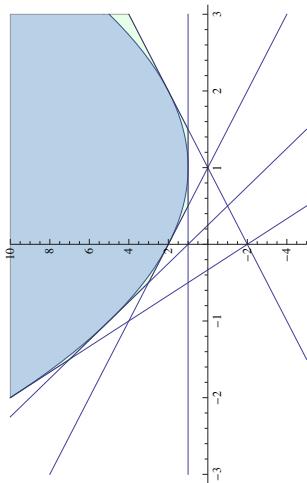
Separate again



Complex Energy Systems
└ Appendix: CC-OPF
└ Cutting-Plane Method

Cutting-Plane Method

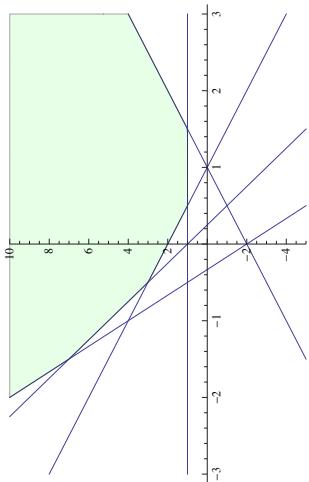
... which approximate the conic constraint



◀ Chance Constrained OPF

Cutting-Plane Method

We might end up with many linear constraints



◀ Chance Constrained OPF

Complex Energy Systems
└ Appendix: CC-OPF
└ Cutting-Plane Method

Polish 2003-2004 winter peak case

- 2746 buses, 3514 branches, 8 wind sources
- 5% penetration and $\sigma = .3\mu$ each source

CPLEX: the optimization problem has

- 36625 variables
- 38507 constraints, 6242 conic constraints
- 128533 nonzeros, 87 dense columns

◀ Chance Constrained OPF

◀ Chance Constrained OPF

Complex Energy Systems
 └ Appendix: CC-OPF
 └ Cutting Plane Method

- CPLEX:
 - total time on 16 threads = 3393 seconds
 - "optimization status 6"
 - solution is wildly infeasible

Gurobi:

- time: 31.1 seconds
- "Numerical trouble encountered"



Polish 2003-2004 case

CPLEX: "opt status 6"

Gurobi: "numerical trouble"

Example run of cutting-plane algorithm:

Iteration	Max rel. error	Objective
1	1.2e-1	7.0933e6
4	1.3e-3	7.0934e6
7	1.9e-3	7.0934e6
10	1.0e-4	7.0964e6
12	8.9e-7	7.0965e6

Total running time: 32.9 seconds

◀ CC-OPF Performance

Complex Energy Systems
 └ Appendix: CC-OPF
 └ Cutting Plane Method

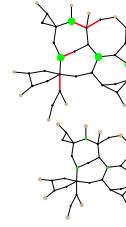
Experiments with CC-OPF (I)



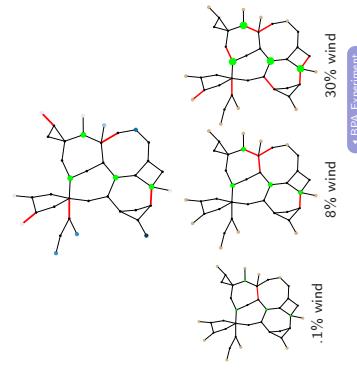
CC-OPF is not a naive fix.

(Changes are nonlocal.)

- 39-bus case. Darker shades of blue indicating generators with greater change from CC-OPF to standard OPF.



◀ BPA Experiment



◀ BPA Experiment

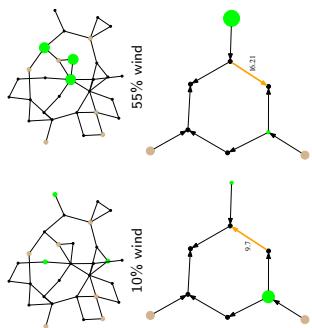
- What is the penetration that can be tolerated (without upgrading)?
- 39-bus case. Three levels of penetration. Standard OPF is infeasible for three level of penetrations. CC-OPF is infeasible only with the penetration level > 30 + %.
- Cost of Reliability [CC-OPF saving over standard OPF]
 ■ 39-bus case under standard solution. Even with a 10% buffer on the line flow limits (for the average wind), five lines exceed their limit over 5% of the time with 30% penetration (right). The penetration must be decreased to 5% before the lines are relieved, but at great cost (left). The CC-OPF model is feasible for 30% penetration at a cost of 264,000. The standard solution at 5% penetration costs 1,275,020 almost 5 times as much.

Experiments with CC-OPF (II)

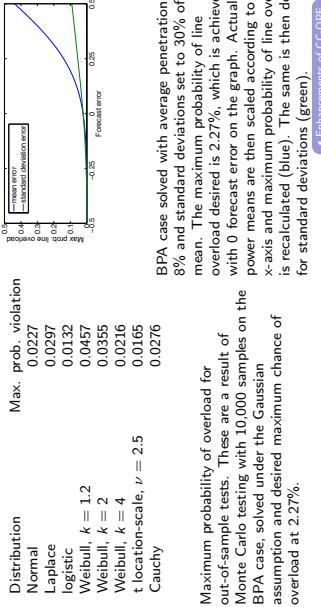
Experiments with CC-OPF (III)

Out of Sample Tests

- Which sites to place wind-farms?
- 30 bus case with three wind farms.
- Placement on the right is preferable.
- CC-OPF finds the nodes where the entire network is less susceptible to fluctuations.



- CC-OPF valid configurations may show significant (allowed!) variability, e.g. flow reversal.
- 9-bus case, 25% average penetration - two significantly different flows.



• Enhancements of CC-OPF

Constant voltage, lossless, security constrained OPF:

Complex Energy Systems
└ Appendix: CC-OPF
└ Convex AC-OPF

PF through Optimization [lossless, constant voltage]

Boyd & Vandenberghe (add. ex. for convex opt. – 2012):
Suppose you solve the convex optimization problem:

$$\begin{aligned} \min_{p, \theta} \quad & \underbrace{\sum_{ij \in \mathcal{L}} \beta_{ij} \Psi(p_{ij})}_{\rho - \text{line flows}} \quad , \quad \Psi(p) = \int_{-1}^p \arcsin(y) dy \\ \text{s.t.} \quad & \underbrace{\sum_{ij \in \mathcal{L}} \beta_{ij} p_{ij} - \sum_{ij \in \mathcal{L}} \beta_{ij} \bar{p}_{ij}}_{\text{reactive losses in lines}} = p_i - d_i \quad \forall i \in \mathcal{B} \quad (*) \\ & |\rho_{ij}| < 1 \quad \text{for each line } ij \\ & \text{network flow conservation} \end{aligned}$$

Then: If θ_j is the optimal dual for (*), $\rho_{ij} = \sin(\theta_i - \theta_j)$.

The opt. is dual to the Energy Function opt.

How can we incorporate this methodology into OPF-type problems?

• CC-OPF enhancements
• CC-OPF enhancements

Theorem: "Exact" AC-OPF [BBC 2013]

Suppose you solve the convex optimization problem:

$$\min_{\rho, \theta, \delta \geq 0} c(\rho) + D \sum_{j \in \mathcal{L}} \beta_j \Psi(\rho_{ij}) - K \sum_{j \in \mathcal{L}} \beta_{ij} \log(\delta_{ij})$$

s.t.

$$\sum_{j: j \notin \mathcal{L}} \beta_j \rho_{ij} - \sum_{j: j \in \mathcal{L}} \beta_j \rho_{ji} = p_i - d_i \quad \forall i \in \mathcal{B}$$

$$|\rho_{ij}| + \min\{1, u_{ij}/\beta_{ij}\} \delta_{ij} < \min\{1, u_{ij}/\beta_{ij}\} \quad \text{for each line } ij$$

$$P_g^{\min} \leq P_g \leq P_g^{\max} \quad \text{for each generator } g$$

For appropriate positive constants D (small) and K (large). Then if a feasible solution is found

- The optimal ρ_{ij} are approximate optimal flows [with line flow limits obeyed]
- $\rho_{ij} \approx \sin(\vartheta_i - \vartheta_j)$
- ρ_{ij} = optimal duals to (**)

• CC-OPF enforcement

Based on Dörfler, Chertkov, Bullo 2013: an approximation

$$\min_{\rho, \vartheta} c(\rho)$$

s.t. $\sum_{j: j \in \mathcal{L}} \beta_{ij} (\vartheta_i - \vartheta_j) = p_i - d_i \quad \forall i \in \mathcal{B}$

$$|\vartheta_i - \vartheta_j| < \min\{1, u_{ij}/\beta_{ij}\} \quad \text{for each line } ij$$

[criterion for existence of solution, assumes strong damping]

• $\beta_{ij} \ll 1/\rho_{ij}$

■ The ϑ are auxiliary variables only.

■ Exact on trees, very accurate for almost all realistic cases tested

■ In experiments, $\beta_{ij}(\vartheta_i - \vartheta_j)$ provides a close approximation to the lossless (active) AC power flow on each line ij

■ (But does not provide phase angles)

AC-OPF [loseless, constant voltage] formulation

AC-OPF [loseless, constant voltage] formulation

$$\min_{\rho, \vartheta} c(\rho)$$

s.t.
$$\begin{aligned} \sum_{j: j \in \mathcal{L}} \beta_{ij} \rho_{ij} - \sum_{j: j \in \mathcal{L}} \beta_{ji} \rho_{ji} &= p_i - d_i \\ |\rho_{ij}| + \min\{1, u_{ij}/\beta_{ij}\} \delta_{ij} &< \min\{1, u_{ij}/\beta_{ij}\} \quad \text{for each line } ij \end{aligned}$$

• CC-OPF enforcement

Incorporation into chance-constrained problem:

AC-OPF [loseless, constant voltage] formulation

On any line ij , we replace $\sin(\vartheta_i - \vartheta_j)$ with the quantity $\vartheta_i - \vartheta_j$

So 'sync' constraint $|\sin(\vartheta_i - \vartheta_j)| \leq \gamma_{ij}$ becomes $|\vartheta_i - \vartheta_j| \leq \gamma_{ij}$

But in either case the constraint is stochastic

Results in a (conic) convex optimization

■ Chance-constrained version: $P(|\vartheta_i - \vartheta_j| > \gamma_{ij}) < \epsilon_{ij}$

All (thermal, gen, sync) Chance Constraints accounted

■ Results in the convex (conic) optimization

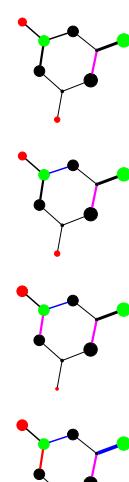
■ Similar to DC CC-OPF – extra sync Chance-constraints added

• CC-OPF enforcement

Complex Energy Systems
 ↳ Appendix: CC-OPF
 ↳ Synchronization-Constraint CC-OPF

Thermal and Sync Aware CC-OPF: Experiments (I)

Competition of sync and thermal risks guides iterations



- The case of $\beta_{ij}/\beta_{ij} \leq 1$ but $\varepsilon = 10^{-4} \ll \varepsilon_f = 10^{-2}$.
- 1st, 8th, 11th and 13th (final) iteration steps shown.
- Nodes: Loads = black, wind farms = green, regular generators = red
- Lines: sync+therm = red; only sync = magenta; only therm = blue, no viol. = black
- Scaling – with actual values or means

► CC-OPF enhancements

Chance-constrained, thermal and sync-aware (approximate) OPF:
 Choose mean generator outputs and control to minimize expected cost, with the probability of line overloads and phase angle excursions kept small. (abridged)

$$\begin{aligned} & \min_{\bar{P}, \alpha} \mathbb{E}[c(\bar{P})] \\ \text{s.t. } & \sum_{i \in G} \alpha_i = 1, \quad \alpha \geq 0 \\ & \bar{B}\delta = \alpha \\ & \sum_{ij \in \mathcal{L}} \beta_{ij} (\bar{\vartheta}_i - \bar{\vartheta}_j) = \bar{p}_i + \mu_i - d_i \end{aligned}$$

- $P(\beta_{ij} |\vartheta_i - \vartheta_j| > u_{ij}) \leq \epsilon_1$ for each line ij
- $P(|\vartheta_i - \vartheta_j| > \gamma_{ij}) \leq \epsilon_2$ for each line ij
- $P(\mathbf{p}_g < \mathbf{p}_g^{\min} \text{ or } \mathbf{p}_g^{\max} < \mathbf{p}_g) \leq \epsilon_3 \leq \epsilon_1$ for each generator g

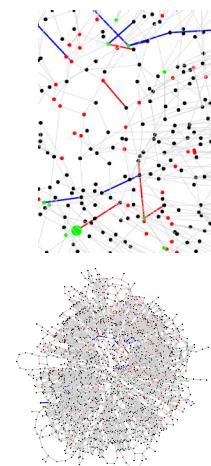
Again: a conic optimization problem

► CC-OPF enhancements

Thermal and Sync Aware CC-OPF: Experiments (II)

Pattern(s) of Sync Warnings

- Qualitative value in studying the warning patterns

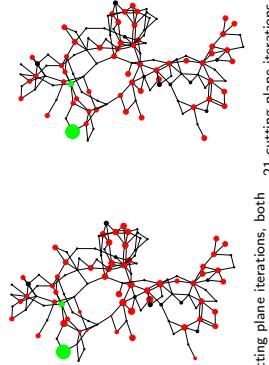


Polish case completed in 11 iterations. Sync overload dominated. Red lines = sync overloaded with probability $\in [10^{-4}; 10^{-2}]$. Blue lines = weaker overload. Scaling according to cons/prod and mean flows within the optimal solution.

► CC-OPF enhancements

Thermal and Sync Aware CC-OPF: Experiments (III)

Sensitivity of the optimal solution to risk awareness



- two slightly different config. of loads
- results distinctly different [cost and distr. of gen.]
- red – regulated generation
- green – renewables

► CC-OPF enhancements

9 cutting plane iterations, both sync and thermal conditions violated [less uniform]

21 cutting plane iterations, only sync conditions violated

Sync in Pics

Complex Energy Systems
 ↳ Appendix: CC-OPF
 ↳ Synchronization-Constrained CC-OPF

Dist(ributed) Flow Representation [Baran, Wu '89]

graph-linear Element $k = 1, \dots, N$ of the distribution feeder

$$\begin{aligned} P_{k+1} - P_k &= p_k - r_k \frac{v_k^2 + Q_k^2}{v_k^2} \\ Q_{k+1} - Q_k &= q_k - x_k \frac{v_k^2 + Q_k^2}{v_k^2} \\ v_{k+1}^2 - v_k^2 &= -2(v_k p_k + x_k q_k) - (r_k^2 + x_k^2) \frac{v_k^2 + Q_k^2}{v_k^2} \end{aligned}$$

$k = 0, \dots, N, v_0 = 1$

$P_{n+1} = Q_{n+1} = 0$

- nonlinear AC over a line
- generalizable to a tree
- P_k, Q_k real and reactive powers flowing through the segment k
- p_k, q_k, v_k powers injected/consumed and voltage at the bus k

Complex Energy Systems
 ↳ Appendix: Distribution Power Flows
 ↳ Dynamics/Transitions in Distributed Feeder (Aux)

Dynamics/Transitions in Distributed Feeder (III)

Example of a Short Fault (\downarrow to full recovery) (Movie Recovery)

Complex Energy Systems
 ↳ Appendix: Distribution Power Flows
 ↳ Dynamics/Transitions in Distributed Feeder (Aux)

ODE with mixed boundary conditions

$$v(0) = 1, \rho(L) = \phi(L) = 0$$

$\nu(t) = 1, \rho(z) = \phi(z) = 0$

From Algebraic Eqs. on a (linear) Graph to Power Flow ODEs

$$0 = p + \underbrace{\beta \partial_r (v^2 \partial_r v) + g v (\beta^2 v^2 - v(\beta v)^2)}_{\text{balance of real power}}, \quad 0 = q + \underbrace{\beta \partial_r (v^2 \partial_r v) - g \partial_r (v^2 \partial_r \theta)}_{\text{balance of reactive power}}$$

$\rho = \underbrace{-\beta v^2 \partial_r \theta - g v \partial_r v}_{\text{real power density flowing through the segment}} - \underbrace{\beta \partial_r v + g^2 \partial_r \theta}_{\text{reactive power density flowing through the segment}}$

$0 = \underbrace{\frac{\partial}{\partial z} (r^2 + \rho^2)}_{\text{real transport}} - \underbrace{r^2 + \rho^2}_{\text{real dissipation}} - \underbrace{\frac{\partial}{\partial z} \phi}_{\text{reactive consumption}} - \underbrace{\frac{g^2}{v^2}}_{\text{reactive dissipation}}$

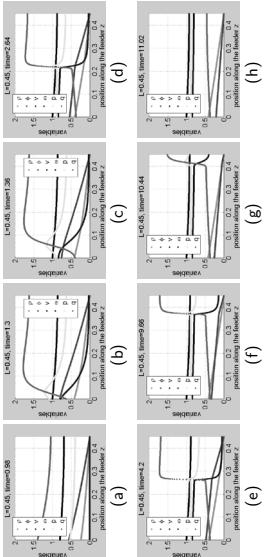
(a) Pre fault
 (b) Post voltage drop at the header. Leads to a fully stalled phase.
 (c) Fault is cleared. Front of recovery is advancing towards the tail.

► Fault Dynamics

Complex Energy Systems
 ↳ Appendix: Distribution Power Flows
 ↳ Dynamics/Transitions in Distributed Feeder (Aux)

Dynamics/Transitions in Distributed Feeder (IV)

- From Stalled to Normal
 - (Movie Recovery)



Of interest: 'Soliton'-like shape; voltage profile is (almost) frozen



Fault Dynamics



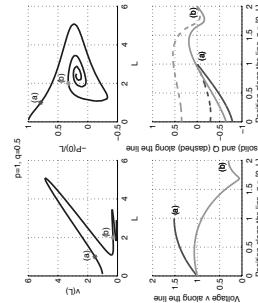
D. Wang, K. Turitsyn, MC (2012)

Feeder with Distributed Generation

Complex Energy Systems
 ↳ Appendix: Distribution Power Flows
 ↳ Unusual Effect(s) of Distributed Generation

Effect of Distributed Generation

- PV systems inject both p and q
- New regulations will require ride-through-low-voltage capability
- If the distributed generation is too large, multiple low-voltage states will appear
- Prediction of a potential trouble: after a fault the system may be trapped in the low-voltage state (similar to FIDVR)
- The only (normal) way to get out of the trouble is to disconnect the PVs



The effect is DISTRIBUTED!
 not seen in the two-node model

Complex Energy Systems
 ↳ Appendix: Distribution Power Flows
 ↳ FIDVR – questions & challenges

What can one do at the distribution level to mitigate FIDVR?

- Monitor/learn/model distributed motor parameters
- Control voltage at the head of the line (rise it when needed)
- Distributed reactive control

Why should System Operator worry about FIDVR?

- Simple restoration of the transmission network may not drive the circuits back to a running state.
- A transmission fault \Rightarrow correlated dynamical response in multiple distribution feeders \Rightarrow individual circuits stalled. Specific to each circuit, there is an energy barrier to the transition back to a running state.
- Once a spatially-correlated stalled state exists, the state of the transmission grid has now fundamentally changed.

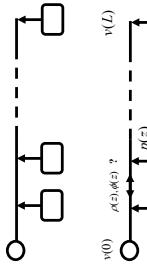
What can the system operator do about FIDVR and related?

- Consider FIDVR as yet another (and much less analyzed!) transient stability issue/contingency
- Attempt to predict (monitoring short voltage faults within the transmission) ... and pull it back to normal without relying (or with minimal reliance) on the distribution level protection and response

Complex Energy Systems
 ↳ Appendix: Distribution Power Flows
 ↳ Unusual Effect(s) of Distributed Generation

Effect of Distributed Generation

- Smart Grid Scenario: Significant Penetration of Photo-Voltaic Systems
- Many Consumers feed back to the system
- (Normally) voltage raises down the feeder and feeder exports, $\rho(0), \eta(0) < 0$
- And what if a fault occurs?

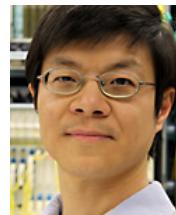


Complex Energy Systems
 ↳ Appendix: Distribution Power Flows
 ↳ Unusual Effect(s) of Distributed Generation

SEMIDEFINITE RELAXATION OF OPTIMAL POWER FLOW**Steven H. Low, CalTech, USA**

The optimal power flow (OPF) problem seeks to optimize a certain objective, such as power loss, generation cost or user utility, subject to Kirchhoff's laws, power balance as well as capacity, stability and contingency constraints on the voltages and power flows. It is a fundamental problem that underlies many power system operations and planning. It is nonconvex and many algorithms have been proposed to solve it approximately. A new approach via semidefinite relaxation of OPF has been developed in the last few years.

In this tutorial, we present a bus injection model and a branch flow model, formulate OPF within each model, and prove their equivalence. The complexity of OPF formulated here lies in the quadratic nature of power flows, i.e., the nonconvex quadratic constraints on the feasible set of OPF. We characterize these feasible sets and design convex supersets that lead to three different convex relaxations based on semidefinite programming (SDP), chordal extension, and second-order cone programming (SOCP). When a convex relaxation is exact, an optimal solution of the original nonconvex OPF can be recovered from every optimal solution of the relaxation. We summarize three types of sufficient conditions that guarantee the exactness of these relaxations. Finally, we extend the convex relaxations to multiphase unbalanced radial networks that are common in distribution systems.



Convex Relaxation of OPF

Acknowledgment

Caltech

- M. Chandy, J. Doyle, M. Farivar, L. Gan, B. Hassibi,
Q. Peng, T. Teeraratkul, C. Zhao

Former

- S. Bose (Cornell), L. Chen (Colorado), D. Gayme
(JHU), J. Lavaei (Columbia), L. Li (Harvard), U.
Topcu (Upenn)

SCE

- A. Auld, J. Castaneda, C. Clark, J. Gooding, M.
Montoya, S. Shah, R. Sherick



Steven Low

Computing + Math Sciences
Electrical Engineering



October 2014
Lund, Sweden

Optimal power flow (OPF)

OPF is solved routinely to determine

- How much power to generate where
- Parameter setting, e.g. taps, VARs
- Market operation & pricing

Non-convex and hard to solve

- Huge literature since 1962
- Common practice: DC power flow (LP)
- Also: Newton-Raphson, interior point, ...

Optimal power flow (OPF)

■ bus injection model, branch flow model

3 convex relaxations

- SDP, chordal, second-order cone (SOCP)
- Relation among them

Sufficient conditions for exact relaxation

- Radial: 3 main conditions
- Mesh: phase shifters



Summary: OPF (bus injection model)

$$\begin{aligned} \min & \quad \text{tr } CVV^* \\ \text{subject to} & \quad \underline{s}_j \leq \text{tr}(Y_j VV^*) \leq \bar{s}_j \\ & \quad |V_j|^2 \leq \bar{V}_j \end{aligned}$$

nonconvex QCQP



Summary: OPF (branch flow model)

$$\begin{aligned} \min & \quad f(x) \\ \text{over} & \quad x := (S, I, V, s) \\ \text{s. t.} & \quad \underline{s}_j \leq s_j \leq \bar{s}_j \\ & \quad \underline{v}_j \leq v_j \leq \bar{v}_j \end{aligned}$$

$$\left[\begin{array}{l} \text{branch flow} \\ \text{model} \end{array} \right] \left\{ \begin{array}{l} \sum_{i \rightarrow j} \left(S_{ij} - \bar{s}_{ij} |I_{ij}|^2 \right) - \sum_{j \rightarrow k} S_{jk} = s_j \\ V_j = V_i - z_{ij} I_{ij} \quad S_{ij} = V_i I_{ij}^* \end{array} \right.$$

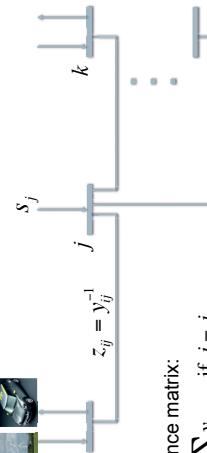
nonconvex



Bus injection model



details



admittance matrix:

$$Y_{ij} := \begin{cases} \sum_{k \sim i} y_{ik} & \text{if } i = j \\ -y_{ij} & \text{if } i \sim j \\ 0 & \text{else} \end{cases}$$

Y specifies topology of G and
impedance z on lines



Recap

Bus injection model

$$s_j = \text{tr}(Y_j V V^*)$$

Branch flow model

$$V_i - V_j = z_{ij} I$$

$$S_{ij} = V I_{ij}^*$$

$$\sum_{j \rightarrow k} S_{jk} = \sum_{i \rightarrow j} \left(S_{ij} - z_{ij} |I_{ij}|^2 \right) + s_j$$

$$(V, s) \in \mathbb{C}^{2(n+1)}$$

solution set
V



Equivalence

Theorem: $V \equiv \tilde{X}$

- BIM and BFM are equivalent in this sense
- Any result in one model is in principle provable in the other,
- ... but some results are easier to formulate or prove in one than the other
- BFM seems to be much more numerically stable (radial networks)

$$(S, I, V, s) \in \mathbb{C}^{2(m+n+1)}$$

solution set
X

$$(S, I, V, s) \in \mathbb{C}^{2(m+n+1)}$$

solution set
X



OPF: bus injection model

$$\begin{aligned} & \min_{\substack{\text{gen cost,} \\ \text{power loss}}} && V^* C V \\ & \text{over} && (V, s) \\ & \text{subject to} && \underline{s}_j \leq s_j \leq \bar{s}_j \quad \underline{V}_j \leq |V_j| \leq \bar{V}_j \\ & && S_j = \text{tr} \left(Y_j^H V V^H \right) \end{aligned}$$

$$\begin{aligned} & \min_{\substack{\text{gen cost,} \\ \text{power loss}}} && V^* C V \\ & \text{over} && (V, s) \\ & \text{subject to} && \underline{s}_j \leq s_j \leq \bar{s}_j \quad \underline{V}_j \leq |V_j| \leq \bar{V}_j \\ & && S_j = \text{tr} \left(Y_j^H V V^H \right) \end{aligned}$$

power flow equation



OPF: bus injection model



Recap

Bus injection model

$$s_j = \text{tr} (Y_j V V^*)$$

Branch flow model

$$V_i - V_j = z_{ij} I$$

$$S_{ij} = V I_{ij}^*$$

$$\sum_{j \rightarrow k} S_{jk} = \sum_{i \rightarrow j} \left(S_{ij} - z_{ij} |I_{ij}|^2 \right) + s_j$$

$$(V, s) \in \mathbb{C}^{2(n+1)}$$

solution set
V



Equivalence

Theorem: $V \equiv \tilde{X}$

- BIM and BFM are equivalent in this sense
- Any result in one model is in principle provable in the other,
- ... but some results are easier to formulate or prove in one than the other
- BFM seems to be much more numerically stable (radial networks)

$$(S, I, V, s) \in \mathbb{C}^{2(m+n+1)}$$

solution set
X

$$(S, I, V, s) \in \mathbb{C}^{2(m+n+1)}$$

solution set
X

OPF: bus injection model

$$\begin{aligned} & \min_{\substack{\text{gen cost,} \\ \text{power loss}}} && V^* C V \\ & \text{over} && (V, s) \\ & \text{subject to} && \underline{s}_j \leq s_j \leq \bar{s}_j \quad \underline{V}_j \leq |V_j| \leq \bar{V}_j \\ & && S_j = \text{tr} (Y_j^H V V^H) \end{aligned}$$

$$\begin{aligned} & \min_{\substack{\text{gen cost,} \\ \text{power loss}}} && V^* C V \\ & \text{over} && (V, s) \\ & \text{subject to} && \underline{s}_j \leq s_j \leq \bar{s}_j \quad \underline{V}_j \leq |V_j| \leq \bar{V}_j \\ & && S_j = \text{tr} (Y_j^H V V^H) \end{aligned}$$

$$\begin{aligned} & \min_{\substack{\text{gen cost,} \\ \text{power loss}}} && V^* C V \\ & \text{over} && (V, s) \\ & \text{subject to} && \underline{s}_j \leq s_j \leq \bar{s}_j \quad \underline{V}_j \leq |V_j| \leq \bar{V}_j \\ & && S_j = \text{tr} (Y_j^H V V^H) \end{aligned}$$

OPF: bus injection model

OPF: bus injection model



$$\begin{aligned}
 & \min \quad \text{tr } C V V^* \\
 & \text{subject to} \quad \underline{s}_j \leq \text{tr} \left(Y_j V V^* \right) \leq \bar{s}_j \quad \underline{v}_j \leq |V_j|^2 \leq \bar{v}_j \\
 & \quad \quad \quad \text{s. t.}
 \end{aligned}$$

quadratically constrained QP (QCQP)
nonconvex, NP-hard

OPF: branch flow model



$$\begin{aligned}
 & \min \quad f(x) \\
 & \text{over} \quad x := (S, I, V, s) \\
 & \text{s. t.} \quad \underline{s}_j \leq s_j \leq \bar{s}_j \quad \underline{v}_j \leq v_j \leq \bar{v}_j
 \end{aligned}$$

quadratically constrained QP (QCQP)
nonconvex, NP-hard

OPF: branch flow model



$$\begin{aligned}
 & \min \quad f(x) \\
 & \text{over} \quad x := (S, I, V, s) \\
 & \text{s. t.} \quad \underline{s}_j \leq s_j \leq \bar{s}_j \quad \underline{v}_j \leq v_j \leq \bar{v}_j
 \end{aligned}$$

$$\begin{aligned}
 & \min \quad f(x) \\
 & \text{over} \quad x := (S, I, V, s) \\
 & \text{s. t.} \quad \underline{s}_j \leq s_j \leq \bar{s}_j \quad \underline{v}_j \leq v_j \leq \bar{v}_j
 \end{aligned}$$

$$\begin{aligned}
 & \sum_{i \rightarrow j} \left(S_{ij} - z_{ij} |I_{ij}|^2 \right) - \sum_{j \rightarrow k} S_{jk} = s_j \\
 & V_j = V_i - z_{ij} I_{ij} \quad S_{ij} = V_i I_{ij}^* \\
 & \text{branch flow model}
 \end{aligned}$$

nonconvexity

Outline



Other features



Security constraint OPF

- Solve for operating points after each single contingency (N-1 security)
- N sets of variables and constraints, one for each contingency

Unit commitment

- Discrete variables

Stochastic OPF

- Chance constraints $\Pr(\text{bad event}) < \varepsilon$

Other constraints

- Line flow, line loss, stability limit, ...
... OPF in practice is a lot harder

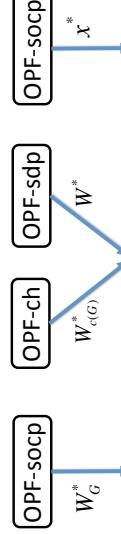
Optimal power flow (OPF)

- bus injection model, branch flow model

3 convex relaxations

- SDP, chordal, second-order cone (SOCP)
- Relation among them
- Sufficient conditions for exact relaxation
- Radial: 3 main conditions
- Mesh: phase shifters

details



What are semidefinite relaxations of OPF?

How to check & recover global optimal ?

details





Literature

Convex relaxation of OPF

relaxation	model	first proposed	first analyzed
SOCP	BIM	Jabr 2006 TPS	
SDP	BIM	Bai et al 2008 EPES	Lavaei, Low 2012 TPS
Chordal	BIM	Bai, Wei 2011 EPES Jabr 2012 TPS	Mohzahn et al 2013 TPS Bose et al 2014 TAC

Low: Convex relaxation of OPF (I, II), IEEE Trans Control of Network Systems, 2014



Basic idea

$$\min_{\mathbf{V}} \text{tr} C \mathbf{V} \mathbf{V}^*$$

$$\text{subject to } \underline{s}_j \leq \text{tr} (\mathbf{Y} \mathbf{V} \mathbf{V}^*) \leq \bar{s}_j \quad \underline{v}_j \leq |\mathbf{V}_j|^2 \leq \bar{v}_j$$

\mathbf{V}

All complexity due to nonconvexity of \mathbf{V}

Relaxations:

- design convex supersets of \mathbf{V}
- minimize cost over convex supersets

All complexity due to nonconvexity of \mathbf{V}

Relaxations:

- design convex supersets of \mathbf{V}
- minimize cost over convex supersets

Convex relaxation of OPF

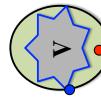
relaxation	model	first proposed	first analyzed
SOCP	BIM	Jabr 2006 TPS	Jabr 2006 TPS
SDP	BIM	Bai et al 2008 EPES	Bai et al 2008 EPES
Chordal	BIM	Bai, Wei 2011 EPES Jabr 2012 TPS	Bai, Wei 2011 EPES Jabr 2012 TPS

Low: Convex relaxation of OPF (I, II), IEEE Trans Control of Network Systems, 2014

$$\min_{\mathbf{V}} \text{tr} C \mathbf{V} \mathbf{V}^*$$

$$\text{subject to } \underline{s}_j \leq \text{tr} (\mathbf{Y} \mathbf{V} \mathbf{V}^*) \leq \bar{s}_j \quad \underline{v}_j \leq |\mathbf{V}_j|^2 \leq \bar{v}_j$$

\mathbf{V}



Basic idea



Basic idea



$$\begin{aligned} \min \quad & \text{tr } C V V^* \\ \text{subject to} \quad & \underline{s}_j \leq \text{tr} (Y_j V V^*) \leq \bar{s}_j \quad \underline{v}_j \leq |V_j|^2 \leq \bar{v}_j \\ & \underbrace{\qquad\qquad\qquad}_{\mathbf{V}} \end{aligned}$$

Approach

1. Three equivalent characterizations of \mathbf{V}
2. Each suggests a lift and relaxation

• What is the relation among different relaxations?
 • When will a relaxation be exact?

Feasible sets



$$\begin{aligned} \min \quad & \text{tr } C V V^* \\ \text{subject to} \quad & \underline{s}_j \leq \text{tr} (Y_j V V^*) \leq \bar{s}_j \quad \underline{v}_j \leq |V_j|^2 \leq \bar{v}_j \\ & \underbrace{\qquad\qquad\qquad}_{\mathbf{V}} \end{aligned}$$

quadratic in V
linear in W

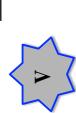
Equivalent problem:

$$\begin{aligned} \min \quad & \text{tr } C W \\ \text{subject to} \quad & \underline{s}_j \leq \text{tr} (Y_j W) \leq \bar{s}_j \quad \underline{v}_i \leq W_{ii} \leq \bar{v}_i \\ & W \geq 0, \text{ rank } W = 1 \quad \text{convex in } W \\ & \qquad\qquad\qquad \text{except this constraint} \end{aligned}$$

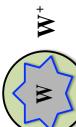
Equivalent feasible sets



$$\mathbf{V} := \left\{ V : \text{satisfies quadratic constraints} \right\}$$



instead of n variables
solve for n^2 vars!



$$\mathbf{W} := \left\{ W : \text{satisfies linear constraints} \right\} \cap \left\{ W \geq 0 \right\}$$

idea: $W = VV^*$

Feasible set



only $n+2m$ vars!

$$\begin{aligned} \text{linear in } (W_{ij}, W_{jk}) \quad & \underbrace{\qquad\qquad\qquad}_{\sum_{k \in \mathcal{N}_j} y_{jk}^* (|V_j|^2 - V_j V_k^*)} \quad \underbrace{\qquad\qquad\qquad}_{W_{ij}} \quad \underbrace{\qquad\qquad\qquad}_{W_{jk}} \\ & \sum_{k \in \mathcal{N}_j} y_{jk}^* (|V_j|^2 - V_j V_k^*) : \text{only } |V_j|^2 \text{ and } V_j V_k^* \end{aligned}$$

corresponding to edges (j,k) in G !

$$\begin{aligned} \min \quad & \text{tr } C V V^* \\ \text{subject to} \quad & \underline{s}_j \leq \text{tr} (Y_j V V^*) \leq \bar{s}_j \quad \underline{v}_j \leq |V_j|^2 \leq \bar{v}_j \\ & \underbrace{\qquad\qquad\qquad}_{\mathbf{V}} \end{aligned}$$

Feasible set

only $n+2m$ vars !

$$\text{linear in } (W_{ji}, W_{jk}) \quad \downarrow \quad \sum_{k:j\neq k} y^*_k (|V_j|^2 - V_j V_k^*)$$

partial matrix W_G defined on G

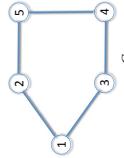
$$W_G := \{[W_G]_{ij}, [W_G]_{jk}, [W_G]_{ki} \mid i, j, k \in G\}$$

Kirchoff's laws depend directly only on W_G

Example

$$\sum_{k:j\neq k} y^*_k (|V_j|^2 - V_j V_k^*)$$

$$W_G = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \\ W_{41} & W_{42} & W_{43} \\ W_{51} & W_{52} & W_{53} \end{bmatrix}$$



$$W = \begin{bmatrix} W_{11} & W_{12} & W_{13} & W_{14} & W_{15} \\ W_{21} & W_{22} & W_{23} & W_{24} & W_{25} \\ W_{31} & W_{32} & W_{33} & W_{34} & W_{35} \\ W_{41} & W_{42} & W_{43} & W_{44} & W_{45} \\ W_{51} & W_{52} & W_{53} & W_{54} & W_{55} \end{bmatrix}$$

SDP solves for $W \in \mathbb{C}^{n^2}$
 n^2 variables

Feasible sets

$$\text{OPF} \quad \mathbf{V} := \left\{ V \Big| \underline{s}_j \leq \text{tr}(Y_j V V^*) \leq \bar{s}_j, \underline{v}_j \leq V_j \leq \bar{v}_j \right\} \cap \{W \geq 0, \text{rank-1}\}$$

$$\text{SDP} \quad \mathbf{W} := \left\{ W \Big| \underline{s}_j \leq \text{tr}(Y_j W) \leq \bar{s}_j, \underline{v}_j \leq W_{jj} \leq \bar{v}_j \right\} \cap \{W \geq 0, \text{rank-1}\}$$

first idea:

$$\mathbf{W}_G := \left\{ W_G \Big| \underline{s}_j \leq \text{tr}(Y_j W_G) \leq \bar{s}_j, \underline{v}_j \leq [W_G]_{jj} \leq \bar{v}_j \right\} \cap \{W_G \geq 0, \text{rank-1}\}$$

\mathbf{W}_G is equivalent to \mathbf{V} when G is chordal
 Not equivalent otherwise

Feasible sets

$$\text{OPF} \quad \mathbf{V} := \left\{ V \Big| \underline{s}_j \leq \text{tr}(Y_j V V^*) \leq \bar{s}_j, \underline{v}_j \leq V_j \leq \bar{v}_j \right\}$$

$$\text{SDP} \quad \mathbf{W} := \left\{ W \Big| \underline{s}_j \leq \text{tr}(Y_j W) \leq \bar{s}_j, \underline{v}_j \leq W_{jj} \leq \bar{v}_j \right\} \cap \{W \geq 0, \text{rank-1}\}$$

depend on all entries of W
 depend only on W_G

Equivalent feasible sets

Cycle condition

$\mathbf{W}_G := \left\{ \begin{array}{l} W_{ij}, W_{jk} : (j,k) \text{ in } G \\ \text{satisfy linear constraints} \end{array} \right\} \cap \left\{ \begin{array}{l} W(j,k) \geq 0 \text{ rank-1,} \\ \text{cycle cond on } \angle W_{jk} \end{array} \right\}$

idea: $W_G = (VV^* \text{ only on } G)$

$\mathbf{W}_{c(G)} := \left\{ \begin{array}{l} W_{ij}, W_{jk} : (j,k) \text{ in } c(G) \\ \text{satisfy linear constraints} \end{array} \right\} \cap \left\{ \begin{array}{l} W_{c(G)} \geq 0 \text{ rank-1} \end{array} \right\}$

idea: $W_{c(G)} = (VV^* \text{ on } c(G))$

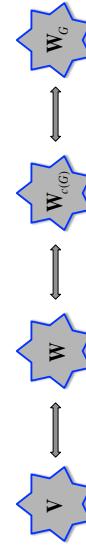
matrix completion [Grone et al 1984]

$\mathbf{W} := \{W : \text{satisfies linear constraints}\} \cap \{W \geq 0 \text{ rank-1}\}$

idea: $W = VV^*$

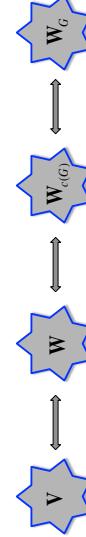
local $W_G(j,k) \geq 0, \text{ rank } W_G(j,k) = 1, \quad (j,k) \in E,$
 $\sum_{(j,k) \in c} \angle [W_G]_{jk} = 0 \quad \text{mod } 2\pi \quad \xrightarrow{\text{cycle cond}}$

Equivalent feasible sets



Theorem: $V = W = W_{c(G)} = W_G$

Equivalent feasible sets



Theorem: $V = W = W_{c(G)} = W_G$

Given $W_G \in \mathbf{W}_G$ or $W_{c(G)} \in \mathbf{W}_{c(G)}$ there is
unique completion $W \in \mathbf{W}$ and unique $V \in \mathbf{V}$

Bose, Low, Chandy Allerton 2012
Bose, Low, Teeraratkul, Hassibi TAC2014

Can minimize cost over any of these sets, but ...



Relaxations

$$\mathbf{W}_G := \left\{ \begin{array}{l} W_{ij}, W_{jk} : (j,k) \text{ in } G \\ \text{satisfy linear constraints} \end{array} \right\} \cap \left\{ \begin{array}{l} W(j,k) \geq 0 \\ \text{cycle condition} \end{array} \right\}$$

idea: $W_G = (V^* \text{ only on } G)$

$$\mathbf{W}_{c(G)} := \left\{ \begin{array}{l} W_{ij}, W_{jk} : (j,k) \text{ in } c(G) \\ \text{satisfy linear constraints} \end{array} \right\} \cap \left\{ \begin{array}{l} W_{c(G)} \geq 0 \\ \text{cycle condition} \end{array} \right\}$$

idea: $W_{c(G)} = (VV^* \text{ on } c(G))$

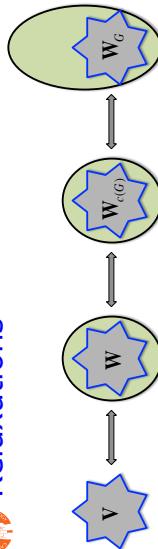
matrix completion [Grone et al 1984]

$$\mathbf{W} := \{W : \text{satisfies linear constraints}\} \cap \{W \geq 0\}$$

idea: $W = VV^*$



Relaxations



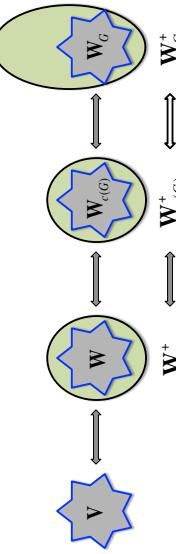
Theorem

- Radial $G : V \subseteq W^+ \equiv W_{c(G)}^+ \equiv W_G^+$
- Mesh $G : V \subseteq W^+ \equiv W_{c(G)}^+ \subseteq W_G^+$

Bose, Low, Chandy, Allerton 2012
Bose, Low, Teeraraku, Hassibi TAC 2014



Convex relaxations



OPF

$$\min_V C(V) \text{ subject to } V \in \mathbf{V}$$

OPF-sdp:

$$\min_W C(W_G) \text{ subject to } W \in \mathbb{W}^+$$

OPF-ch:

$$\min_{W_{c(G)}} C(W_{c(G)}) \text{ subject to } W_{c(G)} \in \mathbb{W}_{c(G)}^+$$

OPF-socp:

$$\min_{W_G} C(W_G) \text{ subject to } W_G \in \mathbb{W}_G^+$$

Theorem

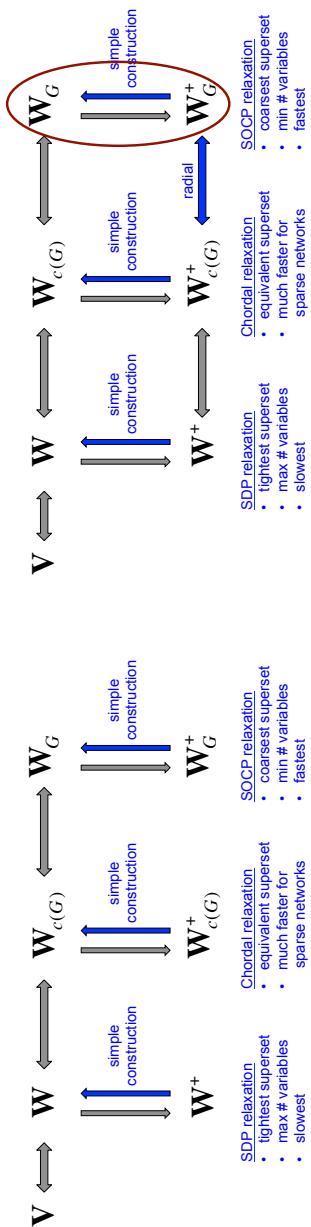
- Radial $G : V \subseteq W^+ \equiv W_{c(G)}^+ \equiv W_G^+$
- Mesh $G : V \subseteq W^+ \equiv W_{c(G)}^+ \subseteq W_G^+$

For radial networks: always solve SOCP !

Recap: convex relaxations

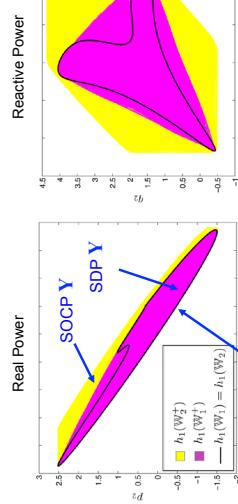


Recap: convex relaxations



For radial network: always solve SOCP !

Examples



Without PS: SDP vs SOCP

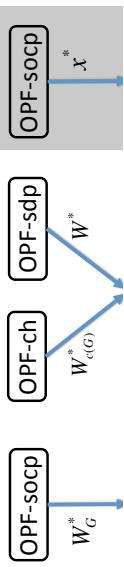


Test case	Objective values (\$/ht)		Running times (sec)	
	SDP	SOCP	SDP	SOCP
9 bus	5297.4	5297.4	0.2	0.2
14 bus	8081.7	8075.3	0.2	0.2
30 bus	574.5	573.6	0.4	0.3
39 bus	41889.1	41881.5	0.7	0.3
57 bus	41738.3	41712.0	1.3	0.3
118 bus	129668.6	129372.4	6.9	0.6
300 bus	720031.0	71906.5	109.4	1.8
2383 bus	1840270	1789500.0	-	155.3

Bose, Low, Teeraratkul, Hassibi TAC 2014

SDP not scalable
SOCP inexact

Examples

Test case	Objective values (\$/hr)		Running times (sec)		
	SDP/ch	SOCP	SDP	chordal	SOCP
9 bus	5297.4	5297.4	0.2	0.2	0.2
14 bus	8081.7	8075.3	0.2	0.2	0.2
30 bus	574.5	573.6	0.4	0.3	0.3
39 bus	41889.1	41881.5	0.7	0.3	0.3
57 bus	41738.3	41712.0	1.3	0.5	0.3
118 bus	129668.6	129372.4	6.9	0.7	0.6
300 bus	720031.0	719006.5	109.4	2.9	1.8
2383 bus	1840270	1789500	-	1005.6	155.3

SDP not scalable
SDP inexact

Branch flow model



Branch flow model

SOCP relaxation

$$\sum_{j \rightarrow k} S_{jk} = \sum_{i \rightarrow j} \left(S_{ij} - z_{ij} |I_{ij}|^2 \right) + s_j$$

$$V_i - V_j = z_{ij} I_{ij}$$

$$V_i I_{ij}^* = S_{ij}$$

$$(S, I, V, s) \in \mathbf{C}^{2(m+n+1)}$$

$$\sum_{j \rightarrow k} P_{jk} = \sum_{i \rightarrow j} \left(P_{ij} - r_{ij} |I_{ij}|^2 \right) + p_j$$

$$\sum_{j \rightarrow k} Q_{jk} = \sum_{i \rightarrow j} \left(Q_{ij} - x_{ij} |I_{ij}|^2 \right) + q_j$$



Branch flow model

SOCP relaxation

$$\sum_{j \rightarrow k} S_{jk} = \sum_{i \rightarrow j} \left(S_{ij} - z_{ij} \ell_{ij} \right) + s_j$$

$$V_i - V_j = z_{ij} I_{ij}$$

$$V_i \ell_{ij} = |S_{ij}|^2$$

$$(S, \ell, v, s) \in \mathbf{R}^{3(m+n+1)}$$



$$\ell_{ij} := |Y_{ij}|^2$$

$$v_i := |Y_i|^2$$

$$\sum_{j \rightarrow k} S_{jk} = \sum_{i \rightarrow j} \left(S_{ij} - z_{ij} \ell_{ij} \right) + s_j$$

$$v_i - v_j = 2 \operatorname{Re}(z_{ij}^* S_{ij}) - |z_{ij}|^2 \ell_{ij}$$

Branch flow model

Branch flow model SOCP relaxation

$$\begin{aligned} \sum_{\substack{j=k \\ j \rightarrow k}} S_{j \bar{k}} &= \sum_{i \rightarrow j} \left(S_{ij} - z_{ij} |I_{ij}|^2 \right) + S_j & \sum_{\substack{j=k \\ j \rightarrow k}} S_{\bar{j}k} &= \sum_{i \rightarrow j} \left(S_{ij} - z_{ij} \ell_{ij} \right) + S_j \\ V_i - V_j &= z_{ij} I_{ij} & v_i - v_j &= 2 \operatorname{Re} \left(z_{ij}^* S_{ij} \right) - \left| z_{ij} \right|^2 \ell_{ij} \\ V_i I_{ij}^* &= S_{ij} & v_i \ell_{ij} &\geq \left| S_{ij} \right|^2 \end{aligned}$$

$$(S, I, V, s) \in C^{2(m+n+1)}$$



Branch flow model

$$\mathbf{X}^+ := \left\{ x : \begin{array}{l} \text{satisfies linear} \\ \text{constraints} \end{array} \right\} \cap \left\{ \ell_{jk} v_j \geq |S|^2 \right\} \quad \text{soc}$$

$$C := \left\{ \begin{array}{l} \ell_{jk} v_j = |S|^2 \\ \text{cycle cond on } x \end{array} \right\}$$

Theorem $X = X^+ \cap C$

Branch flow model

power flow solutions: $x := (S, \ell, v, s)$ satisfy

$$\sum_{j=k}^i S_{jk} = S_{ij} - z_{ij}\ell_{ij} + S_j$$

$$\begin{aligned} v_i - v_j &= 2 \operatorname{Re} \left(\frac{{}^*S_{ij}}{z_{ij}} \right) - |z_{ij}|^2 \ell_{ij} \\ \ell_{ij} v_i &= |S_{ij}|^2 \end{aligned}$$

$$(S, I, V, s) \in C^{2(m+n+1)}$$



Branch flow model

power flow solutions: $x := (S, \ell, v, s)$ satisfy

$$\sum_{j=k}^i S_{jk} = S_{ij} - z_{ij}\ell_{ij} + S_j$$

$$v_i - v_j = 2 \operatorname{Re} \left(\frac{{}^*S_{ij}}{z_{ij}} \right) - |z_{ij}|^2 \ell_{ij}$$

$$(S, \ell, v, s) \in \mathbf{R}^{3(m+n+1)}$$



Cycle condition

A relaxed solution x satisfies the cycle condition if

$$\exists \theta \text{ s.t. } B\theta = \beta(x) \mod 2\pi$$

incidence matrix;
depends on topology

$$\beta_{jk}(x) := \angle(v_j - z_{jk}^H S_{jk})$$

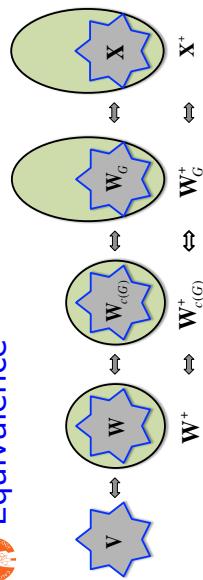
Cycle condition



BFM: SOCP relaxation of OPF

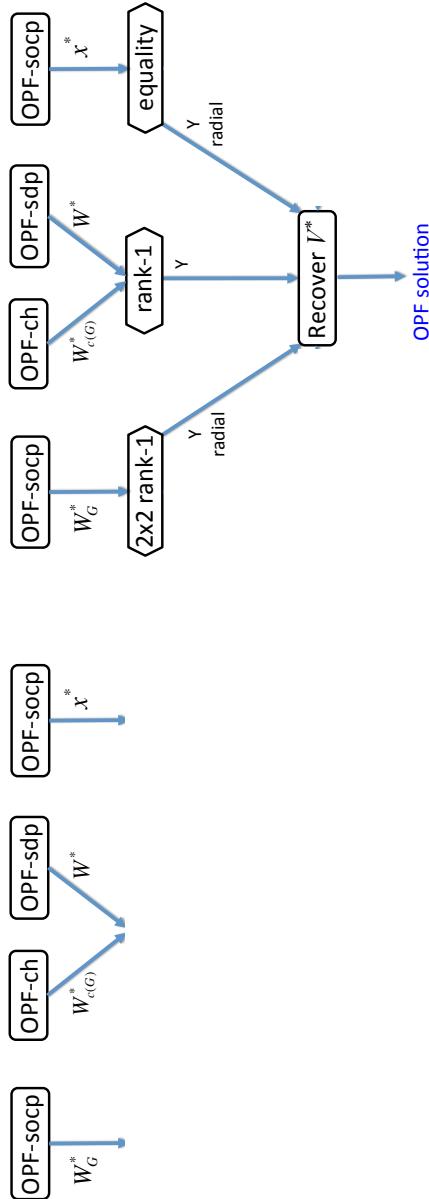


$$\text{OPF: } \min_{x \in X} f(x)$$

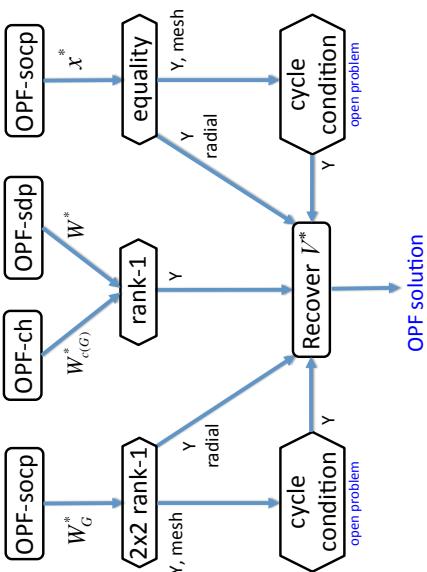


$$\text{SOCP: } \min_{x \in X^+} f(x)$$

Theorem
 $W_G = X$ and $W_G^+ = X^+$



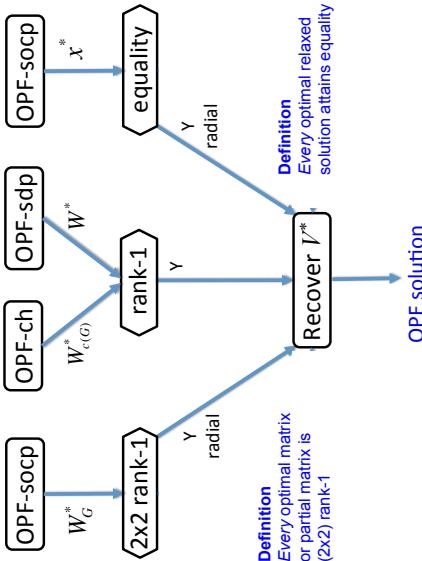
Outline

Exact relaxation



A relaxation is **exact** if an optimal solution of the original OPF can be recovered from every optimal solution of the relaxation



Definition
Every optimal matrix or partial matrix is (2×2) rank-1

Definition
Every optimal relaxed solution attains equality

Summary of sufficient condns



1. QCQP over tree

QCQP (C, C_k)			
type	condition	model	reference
A	power injections	BIM, BFM	[25], [26], [27], [28], [29]
B	voltage magnitudes	BFM	[30], [31], [32], [33], [34]
C	voltage angles	BIM	[35], [36]

TABLE I: Sufficient conditions for radial (tree) networks.

network	condition	reference	remark
with phase shifters	type A, B, C	[17], Part III, [37]	equivalent to radial networks
direct current	type A	[17], Part I, [19], [38]	assumes nonnegative voltages
	type B	[39], [40]	assumes nonnegative voltages

TABLE II: Sufficient conditions for mesh networks

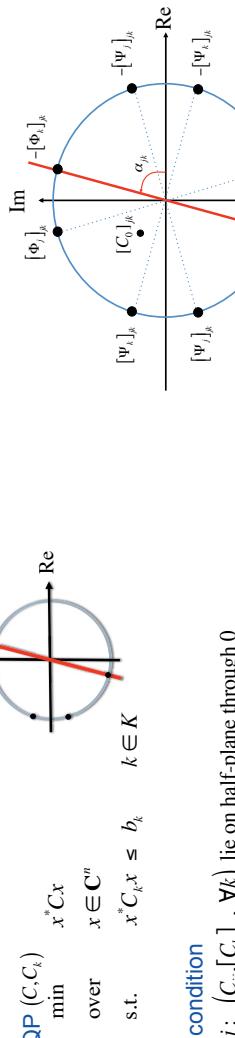
graph of QCQP

$$\begin{aligned} G(C, C_k) \text{ has edge } (i, j) &\Leftrightarrow \\ C_{ij} \neq 0 \text{ or } [C_k]_{ij} \neq 0 &\text{ for some } k \end{aligned}$$

QCQP over tree

$$G(C, C_k) \text{ is a tree}$$

1. Linear separability



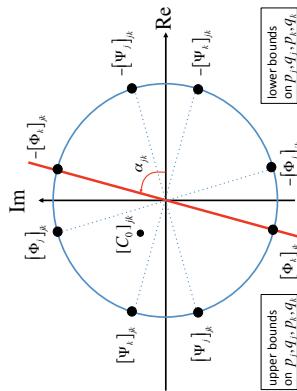
Key condition

$$i \sim j: (C_{ij}, [C_k]_{ij}, \forall k) \text{ lie on half-plane through 0}$$

Theorem
SOCP relaxation is exact for
QCQP over tree

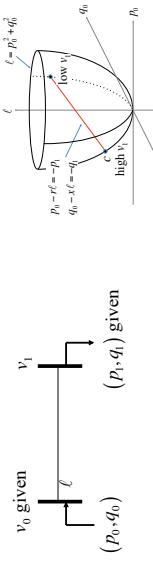
Bose et al 2012
Sjöqvist, Lavafel 2013

Implication on OPF



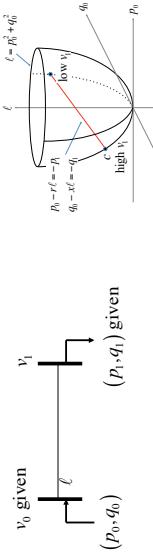
Not both lower & upper bounds on real & reactive powers at both ends
of a line can be finite

2. Voltage upper bounds

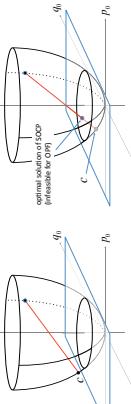


when there is no voltage constraint
 • feasible set: 2. intersection pts
 • relaxation: line segment
 • exact relaxation: c is optimal

2. Voltage upper bounds



voltage lower bound (upper bound on l) does not affect relaxation



(a) Voltage constraint not binding

(b) Voltage constraint binding

2. Voltage upper bounds



$$\begin{aligned} \text{OPF: } & \min_{x \in X} f(x) \quad \text{s.t. } \underline{v} \leq v \leq \bar{v}, \quad s \in \Sigma \\ \text{SOCP: } & \min_{x \in X} f(x) \quad \text{s.t. } \underline{v} \leq v \leq \bar{v}, \quad s \in \Sigma \end{aligned}$$

Key condition:

- $L(s) \leq \bar{v}$
- **Jacobian condition**
 $\bar{\Delta}_1, \dots, \bar{\Delta}_l, z_{l+1} > 0$ for all $1 \leq l \leq l' < k$

voltages if network were lossless
 if upward current were reduced
 then all subsequent powers dec

2. Voltage upper bounds



$$\begin{aligned} \text{OPF: } & \min_{x \in X} f(x) \quad \text{s.t. } \underline{v} \leq v \leq \bar{v}, \quad s \in \Sigma \\ \text{SOCP: } & \min_{x \in X} f(x) \quad \text{s.t. } \underline{v} \leq v \leq \bar{v}, \quad s \in \Sigma \end{aligned}$$

Key condition:

- $L(s) \leq \bar{v}$
- **Jacobian condition**
 $\bar{\Delta}_1, \dots, \bar{\Delta}_l, z_{l+1} > 0$ for all $1 \leq l \leq l' < k$

satisfied with large margin in

IEEE circuits and SCE circuits

Theorem
 SOCP relaxation is exact for
 radial networks

Gan, Li, Topcu, Low TAC2014

Gan, Li, Topcu, Low TAC2014

MULTI-AGENT OPTIMIZATION

Asuman Ozdaglar, Massachusetts Institute of Technology, USA

We consider a network of agents solving a global optimization problem, where the objective function is the sum of privately known local convex objective functions. Recent literature presented subgradient based methods for distributed solution of this problem with $O(1/\sqrt{k})$ rate of convergence (where k is the iteration number). In this talk, we present distributed Alternating Direction Method of Multipliers (ADMM) based methods for solving this problem over undirected and directed networks. We present convergence rate estimates that show that these methods converge at rate $O(1/k)$ and highlight the dependence of performance on network structure.



Distributed Alternating Direction Method of Multipliers for Multi-agent Optimization

Asu Ozdaglar

Laboratory for Information and Decision Systems
Operations Research Center
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology

Lund Workshop on Dynamics and Control in Networks
October, 2014

Distributed Multi-agent Optimization

- Many of these problems can be represented within the general formulation:
- A set of agents (nodes) $\{1, \dots, N\}$ connected through a network (graph).

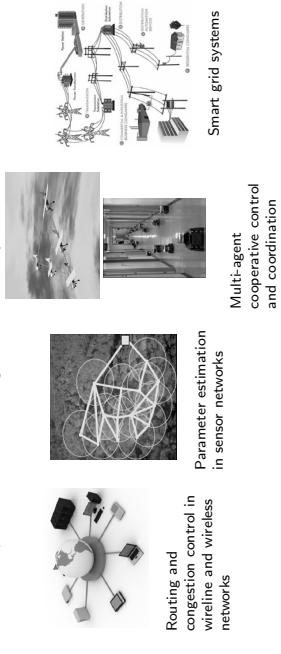
- The goal is to cooperatively solve

$$\begin{aligned} \min_{\mathbf{x}} \quad & \sum_{i=1}^N f_i(x_1, \dots, x_n) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbb{R}^n, \\ & f_i(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R} \text{ is a convex} \\ & \text{(possibly nonsmooth) function,} \\ & \text{known only to agent } i. \end{aligned}$$

- Since such systems often lack a centralized processing unit, algorithms for this problem should involve each agent performing computations locally and communicating only with neighbors.

Motivation

- Many networks are large-scale and comprise of agents with local information and heterogeneous preferences.
- This motivated much interest in developing distributed schemes for control and optimization of multi-agent networked systems.



Machine Learning Example

- A network of agents $i = 1, \dots, N$.
- Each agent i has access to local feature vectors A_i and output b_i .
- System objective: train weight vector \mathbf{x} to

$$\min_{\mathbf{x}} \sum_{i=1}^{N-1} L(A'_i \mathbf{x} - b_i) + p(\mathbf{x}),$$

for some loss function L (on the prediction error) and penalty function p (on the complexity of the model).

- Example: Least-Absolute Shrinkage and Selection Operator (LASSO):

$$\min_{\mathbf{x}} \sum_{i=1}^{N-1} \|A'_i \mathbf{x} - b_i\|_2^2 + \lambda \|\mathbf{x}\|_1.$$

- Other examples from ML estimation, low rank matrix completion, image recovery [Schiolas, Ribeiro, Giannakis 08], [Recht, Fazel, Parrilo 10], [Steidl, Teuber 10].

3

4

Literature: Parallel and Distributed Optimization

Introduction

Literature: Alternating Direction Method of Multipliers

Introduction

- Lagrangian relaxation and dual optimization methods:
 - Dual gradient ascent, coordinate ascent methods.
 - Parallel computation and optimization:
 - [Titsiklis, Bertsekas, Athans 86], [Bertsekas and Tsitsiklis 89].
 - [Jadbabaie, Lin, Morse 03], [Olfati-Saber, Murray 04], [Boyd et al. 05], [Olshevsky, Tsitsiklis 07], [Fagnani, Zampieri 09].
 - Consensus and cooperative control:
 - [Globus, Marrocco 75], [Fortin, Glowinski 83], [Gabay 83], [Eckstein-Bertsekas 92].
 - Recent tutorials: [Boyd et al. 10], [Eckstein 12].
 - Relation to proximal point algorithm: [Rockafellar 76, 76], [Luque 84].
 - Decentralized estimation and compressive sensing applications: [Schizas, Ribeiro, Giannakis 08], [Mota, Xavier, Aguiar, Puschel 11].
 - Multi-agent optimization
 - Distributed primal subgradient methods [Nedic, Ozdaglar 07].
 - Distributed dual averaging methods [Duchi, Aganwal Wainwright 12].
 - These methods converge at the rate $O(1/\sqrt{k})$, where k is the number of iterations.

This Talk

Standard ADMM

- We present distributed ADMM-type algorithms for multi-agent optimization.
 - Distributed ADMM over undirected networks [Shtern, Wei, and Ozdaglar [\[14\]](#)].
 - Distributed ADMM over directed networks [Makhoul and Ozdaglar [\[14\]](#)].
 - In both cases, we show that these algorithms converge at the $O(1/k)$ rate.

- Standard ADMM solves a separable problem, where decision variable decomposes into two (linearly coupled) variables:
$$\begin{aligned} & \min_{x,y} f(x) + g(y) \\ & \text{s.t. } Ax + By = c. \end{aligned}$$

• Consider an Augmented Lagrangian function:

$$\begin{aligned} & \text{s.t.} \\ & Ax + By = c \end{aligned}$$

卷之八

- ADMM: approximate version of classical Augmented Lagrangian method.
 - Primal variables: approximately minimize augmented Lagrangian through a single-pass coordinate descent (in a Gauss-Seidel manner)
 - Dual variable: updated through gradient ascent.

7

Standard ADMM

More specifically, updates are as follows:

$$\begin{aligned} x^{k+1} &= \operatorname{argmin}_x L_\beta(x, y^k, p^k), \\ y^{k+1} &= \operatorname{argmin}_y L_\beta(x^{k+1}, y, p^k), \\ p^{k+1} &= p^k - \beta(Ax^{k+1} - By^{k+1} - c). \end{aligned}$$

- Each minimization involves (quadratic perturbations of) functions f and g separately.

- In some applications, these minimizations are easy (quadratic minimization, l_1 minimization, which arises in Huber fitting, basis pursuit, LASSO, total variation denoising).

- Best known convergence rate: $O(1/k)$ [He, Yuan 11].¹

¹Under stronger assumptions (strong convexity, Lipschitz gradient), ADMM converges linearly [Goldfarb et. al 10], [Deng, Yin 12], [Hong, Luo 12].

Special Case Study: 2-agent Optimization Problem

- Multi-agent optimization problem with two agents:

$$\begin{aligned} \min_{x_1, x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & x_1 = x_2. \end{aligned}$$

- ADMM applied to this problem yields:



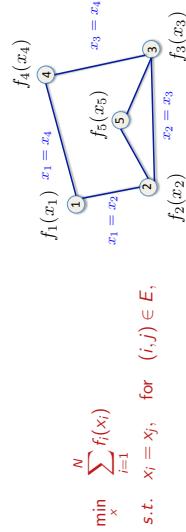
$$x_1^{k+1} = \operatorname{argmin}_{x_1} f_1(x_1) + f_2(x_2^k) - (p_{12}^k)'(x_1 - x_2^k) + \frac{\beta}{2} \|x_1 - x_2^k\|_2^2$$

$$x_2^{k+1} = \operatorname{argmin}_{x_2} f_1(x_1) - (p_{12}^k)'x_1 + \frac{\beta}{2} \|x_1 - x_2^k\|_2^2$$

ADMM for Multi-agent Optimization Problem

- Multi-agent optimization can be reformulated in the ADMM framework:

- Consider a set of agents $V = \{1, \dots, N\}$ in an undirected connected graph $G = (V, E)$.
- $\mathcal{N}(i)$: agent i 's neighborhood.
- We introduce a local copy x_i in \mathbb{R}^n for each of the agents and impose $x_i = x_j$ for all $(i, j) \in E$.

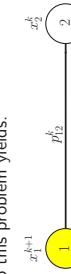


Special Case Study: 2-agent Optimization Problem

- Multi-agent optimization problem with two agents:

$$\begin{aligned} \min_{x_1, x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & x_1 = x_2. \end{aligned}$$

- ADMM applied to this problem yields:



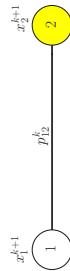
Special Case Study: 2-agent Optimization Problem

Introduction

- Multi-agent optimization problem with two agents:

$$\begin{aligned} \min_{x_1, x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & x_1 = x_2. \end{aligned}$$

- ADMM applied to this problem yields:



$$x_2^{k+1} = \arg\min_{x_2} \quad f_1(x_1^{k+1}) + f_2(x_2) - (\rho_{12}^k)(x_1^{k+1} - x_2) + \frac{\beta}{2} \|x_1^{k+1} - x_2\|_2^2$$

- Multi-agent optimization problem with two agents:

$$\begin{aligned} \min_{x_1, x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & x_1 = x_2. \end{aligned}$$

- ADMM applied to this problem yields:



$$\begin{aligned} x_2^{k+1} &= \arg\min_{x_2} \quad f_1(x_1^{k+1}) + f_2(x_2) - (\rho_{12}^k)(x_1^{k+1} - x_2) + \frac{\beta}{2} \|x_1^{k+1} - x_2\|_2^2 \\ x_2 &= \arg\min_{x_2} \quad f_2(x_2) + (\rho_{12}^k)x_2 + \frac{\beta}{2} \|x_1^{k+1} - x_2\|_2^2 \end{aligned}$$

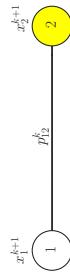
Special Case Study: 2-agent Optimization Problem

Introduction

- Multi-agent optimization problem with two agents:

$$\begin{aligned} \min_{x_1, x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & x_1 = x_2. \end{aligned}$$

- ADMM applied to this problem yields:



Special Case Study: 2-agent Optimization Problem

Introduction

- Multi-agent optimization problem with two agents:

$$\begin{aligned} \min_{x_1, x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & x_1 = x_2. \end{aligned}$$

- ADMM applied to this problem yields:



$$\begin{aligned} x_2^{k+1} &= \rho_{12}^k - \beta(x_1^{k+1} - x_2^{k+1}) \\ x_2 &= \rho_{12}^k - \beta(x_1 - x_2) \end{aligned}$$

Multi-agent Optimization Problem: Reformulation

11

Synchronous ADMM

- Multi-agent optimization problem with two agents:

$$\begin{aligned} \min_{x_1, x_2} \quad & f_1(x_1) + f_2(x_2) \\ \text{s.t.} \quad & x_1 = x_2. \end{aligned}$$

- ADMM applied to this problem yields:



$$\begin{aligned} x_2^{k+1} &= \rho_{12}^k - \beta(x_1^{k+1} - x_2^{k+1}) \\ x_2 &= \rho_{12}^k - \beta(x_1 - x_2) \end{aligned}$$

- The reformulated problem can be written compactly as

$$\begin{aligned} \min_{x \in \mathbb{R}^{2Mn}, z \in Z} \quad & F(x) = \sum_{i=1}^N f_i(x_i) \\ \text{s.t.} \quad & Dx + z = 0. \end{aligned} \tag{1}$$

$$Z = \{z \in \mathbb{R}^{2Mn} \mid z_{ij} = z_j, \text{ for } (i, j) \text{ in } E\}.$$

- Assumption: The optimal solution set of this problem is nonempty.

Introduction

Introduction

11

12

Distributed Implementation

Synchronous ADMM

Rate of Convergence

- Each agent i maintains x_i^k and p_j^k, z_j^k , for $j \in \mathcal{N}(i)$.

- At iteration k ,

- Agent i updates the primal variable x_i^{k+1} as

$$x_i^{k+1} \in \underset{x_i}{\operatorname{argmin}} f_i(x_i) - \sum_{j \in \mathcal{N}(i)} (\rho_j^k)^T x_i + \frac{\beta}{2} \sum_{j \in \mathcal{N}(i)} \|x_i - z_j^k\|^2.$$

The value x_j^{k+1} is sent to all neighbors.

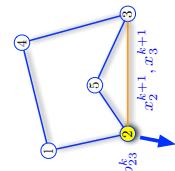
- Agent i computes

$$z_{ij}^{k+1} = z_{ji}^{k+1} = \frac{1}{2}(x_i^{k+1} + x_j^{k+1}).$$

- Agent i computes

$$p_j^{k+1} = p_j^k - \frac{\beta}{2}(x_i^{k+1} - x_j^{k+1}),$$

for all neighbors in j in $\mathcal{N}(i)$.



Synchronous ADMM

- Let $\{\bar{x}^k, z^k, p^k\}$ be the sequence generated by the synchronous ADMM algorithm. We define the ergodic averages:

$$\bar{x}^k = \frac{1}{k} \sum_{l=0}^{k-1} x^l, \quad \bar{z}^k = \frac{1}{k} \sum_{l=0}^{k-1} z^l.$$

We also define the ergodic time average of the residual: $\bar{r}^k = D\bar{x}^k + \bar{z}^k$.

Theorem

Let (x^*, z^*, p^*) be a primal-dual optimal solution for problem (1). The following hold at each iteration k :

$$|F(\bar{x}^k) - F(x^*)| \leq \frac{1}{2\beta k} \left(\beta^2 \|z^0 - z^*\|^2 + \max \left\{ \|\rho^0\|^2, \|\rho^0 - 2p^*\|^2 \right\} \right),$$

$$|\bar{r}(k)| \leq \frac{1}{2\beta k} \left(\beta^2 \|z^0 - z^*\|^2 + (\|\rho^0 - p^*\| + 1)^2 \right).$$

15

Synchronous ADMM

Proof of idea

- Using optimality conditions and dual update, we obtain for all $\rho \in \mathbb{R}^{2Mn}$,

$$\begin{aligned} F(x^{k+1}) - F(x^*) - \rho^k r^{k+1} &\leq \frac{1}{2\beta} \left(\|\rho^k - \rho\|^2 - \|\rho^{k+1} - \rho\|^2 \right) \\ &\quad + \frac{\beta}{2} \left(\|z^k - z^*\|^2 - \|z^{k+1} - z^*\|^2 \right). \end{aligned}$$

- First bound follows from adding this over a window and using convexity to generate function values at ergodic averages.

- Second bound follows from picking $\rho = p^* - \frac{r^k}{\|r^k\|}$.

- One can use $\|\rho^k - p^*\|^2 + \beta^2 \|z^k - z^*\|^2$ as a Lyapunov function to show that the sequence $\{x^k, z^k, p^k\}$ converges to a primal-dual optimal solution of problem (1).

Synchronous ADMM

Network Effect

- The performance depends on the network topology through $\|\rho^*\|$.

Theorem

Let (x^*, z^*, p^*) be an optimal primal-dual solution for problem (1). Then,

$$\|\rho^*\|^2 \leq \frac{2C^2}{\rho_2(L(G))},$$

where C is a bound on the norms of all vectors in $\partial F(x^*)$ and $\rho_2(L(G))$ is the second smallest positive eigenvalue of the Laplacian matrix $L(G)$ of the underlying graph.²

- The performance depends on the algebraic connectivity of the graph: the more connected it is, the larger $\rho_2(L(G))$ is, the better the performance.

² $L(G)$ is a matrix with elements $[L(G)]_{ij} = \begin{cases} \text{degree}(i) & i = j, \\ -1 & (i, j) \in E. \end{cases}$

Dependence of Convergence Rate on Network

Theorem

Let (x^*, z^*) be a primal optimal solution for problem (1). Then with initialization $x^0 = 0, z^0 = 0$ and $\rho^0 = 0$, we have

$$\begin{aligned} |F(\bar{x}^k) - F(x^*)| &\leq \frac{M\beta ||x^*||^2}{kN} + \frac{4C^2}{k\beta \rho_2(L(G))}, \\ ||\bar{x}^k|| &\leq \frac{1}{2k\beta} \left(\frac{C\sqrt{2}}{\sqrt{\rho_2(L(G))}} + 1 \right)^2 + \frac{M\beta ||x^*||^2}{kN}, \end{aligned}$$

where M is the number of edges.

Distributed ADMM over Directed Networks

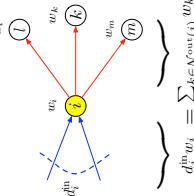
- So far, we have assumed communication among nodes bidirectional.
- This does not hold when nodes have different local interference patterns and different power transmission levels.
- Question:** Can we design a distributed ADMM-like algorithm over directed networks with fast convergence? [Makhdoumi, Ozdaglar 14]
 - Every agent updates primal and dual variables based on information received from incoming links.

- Key idea:** Normalize the information entering a node with the *right weight* so that in the limit the network becomes **balanced** (i.e., amount of information entering and leaving a node is equalized).

- Two approaches:
 - Push-sum [Kempe, Dobra, Gehrke 03], [Nedich, Olshevsky 13,14].
 - Leads to nonlinear updates and a non convergent algorithm for ADMM.
 - Graph balancing: [Gharesifard, Cortes 09], [Priolo et al. 13], [Rikos et al. 14].

Weight Balanced Graph

- Let $G = (V, E)$ be a strongly connected directed network.
- $N^m(i)$: in-neighbors of node i ; $N^{out}(i)$: out-neighbors of node i .
- Each node has a weight w_i representing how much he scales his incoming information.
- The node weights (w_1, \dots, w_N) make the network **balanced** if for each node i , total weight of i 's incoming information is equal to the total weight of its outgoing information.



Distributed Balancing

- Laplacian of the directed graph L : $L_{ii} = d_i^in$ and $L_{ij} = -1, j \in N^m(i)$.
- The balancing weight vector $\mathbf{w} = (w_1, \dots, w_N)'$ is the left eigenvector of L corresponding to eigenvalue 0, i.e., $\mathbf{w}L = 0$; or for $i = 1, \dots, N$,

$$(w_1, \dots, w_i, \dots, w_N) \begin{bmatrix} -1 & & & \\ & \ddots & & \\ & & d_i^in & \\ & & \vdots & \ddots \end{bmatrix} \Rightarrow d_i^in w_i = \sum_{k \in N^{out}(i)} w_k.$$

- We will use a slight variation of [Priolo et al. 13] to design a distributed balancing algorithm.
- For each node i , let $0 < \delta_i < \frac{1}{d_i^in}$.
- Let $C = I - \Delta L$, where $\Delta = \text{diag}(\delta_1, \dots, \delta_N)$.

Distributed Balancing Algorithm

Convergence of Balancing Algorithm

Each agent i maintains weight vector $\tilde{\mathbf{w}}_i^k \in \mathbb{R}^N$. At iteration k :

- The weight matrix $\tilde{W}^k = [\tilde{\mathbf{w}}_1^k; \dots; \tilde{\mathbf{w}}_N^k]$ is updated as
- $$\tilde{W}^{k+1} = C\tilde{W}^k.$$

- This corresponds to each agent i updating his weight vector $\tilde{\mathbf{w}}_i^k$ using [local information](#) from his in-neighbors:

$$\tilde{\mathbf{w}}_i^{k+1} = (1 - \delta_i d_i^{in})\tilde{\mathbf{w}}_i^k + \sum_{j \in \mathcal{N}^i(i)} \delta_i \tilde{\mathbf{w}}_j^k.$$

- Agent i sends $\tilde{\mathbf{w}}_i^{k+1}$ to all his out-neighbors.

Theorem
Starting from $\tilde{W}^0 = I$, we have

$$\lim_{k \rightarrow \infty} \tilde{W}^k = \lim_{k \rightarrow \infty} C^k = \mathbf{1}\mathbf{v}' = \begin{bmatrix} v_1 & v_2 & \dots & v_N \\ v_1 & v_2 & \dots & v_N \\ \vdots & \vdots & \ddots & \vdots \\ v_1 & v_2 & \dots & v_N \end{bmatrix},$$

where \mathbf{v}' is the left eigenvector of C corresponding to eigenvalue 1. Convergence is exponentially fast with exponent given by the second largest eigenvalue of L .

Proof. Follows from Perron-Frobenius Theorem (C is a primitive matrix since G is strongly connected).

- Hence the weight vector of agent i converges to $\tilde{\mathbf{w}}_i^\infty = [v_1, \dots, v_N]'$.
- The balancing weight for agent i is $w_i^\infty = \delta_i \tilde{\mathbf{w}}_i^\infty(\mathbf{1}) = \delta_i v_i$.

$$v_i = (1 - \delta_i d_i^{in})v_i + \sum_{j \in \mathcal{N}^i(i)} \delta_j v_j \Rightarrow d_i^{in} \delta_i v_i = \sum_{j \in \mathcal{N}^i(i)} \delta_j v_j$$

23

Implementation of ADMM over Directed Graphs

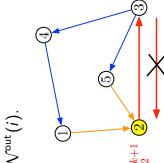
- Consider the distributed ADMM iteration over undirected networks.

- At iteration k , agent i updates the primal variable x_i^k as

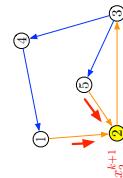
$$\begin{aligned} x_i^{k+1} \in \arg\min_{x_i} f(x_i) - \sum_{j \in \mathcal{N}^i(i)} (\mathbf{p}_j^k)' x_i + \frac{\beta}{2} \sum_{j \in \mathcal{N}^i(i)} \|x_i - z_j^k\|^2 \\ - \sum_{j \in \mathcal{N}^{\text{out}}(i)} (\mathbf{p}_j^k)' x_i + \frac{\beta}{2} \sum_{j \in \mathcal{N}^{\text{out}}(i)} \|x_i - z_j^k\|^2. \end{aligned}$$

Fix:

- This requires node i to receive information from $j \in \mathcal{N}^{\text{out}}(i)$.



- Use only incoming information in the update.



24

Distributed ADMM over Directed Networks

Distributed ADMM over Directed Networks

- Each agent i maintains $x_i^k, \tilde{w}_i^k (w_i^k)$ and p_{ij}^k, z_{ij}^k , for $j \in \mathcal{N}^{in}(i)$.
- At iteration k ,

- a Agent i updates the primal variable x_i^k as

$$x_i^{k+1} \in \underset{x_i}{\operatorname{argmin}} f(x_i) - w_i^k \sum_{j \in \mathcal{N}^{in}(i)} (\rho_{ij}^k)^\top x_i + w_i^k \frac{\beta}{2} \sum_{j \in \mathcal{N}^{in}(i)} \|x_i - z_j^k\|^2,$$

The values x_i^{k+1} and \tilde{w}_i^k are sent to all out-neighbors.

- b For all $j \in \mathcal{N}^{in}(i)$, agent i computes

$$z_{ij}^{k+1} = \frac{1}{2}(x_i^{k+1} + x_j^{k+1}),$$

$$\rho_{ij}^{k+1} = \rho_{ij}^k - \frac{\beta}{2}(x_i^{k+1} - x_j^{k+1}),$$

$$p_{ij}^{k+1} = p_{ij}^k - \frac{\beta}{2}(x_i^{k+1} - x_j^{k+1}),$$

$$x_2^{k+1} = \delta(\tilde{w}_i^{k+1}(i),$$

$$\text{d Agent } i \text{ lets } w_i^{k+1} = \delta(\tilde{w}_i^{k+1}(i),$$

$$\text{where } \tilde{w}_i^{k+1} = (1 - d_i^{in} \delta_i) \tilde{w}_i^k + \sum_{j \in \mathcal{N}^{in}(i)} \delta_i \tilde{w}_j^k$$

$$p_{i2}^{k+1} = p_{i2}^k - \frac{\beta}{2}(x_2^{k+1} - x_2^{k+1})$$

$$p_{i2}^{k+1} = (1 - d_i^{in} \delta_i) p_{i2}^k + \sum_{j \in \mathcal{N}^{in}(i)} \delta_i p_{j2}^k$$

Rate of Convergence

- Let $\{x^k, z^k, p^k, w^k\}$ be the sequence generated by the Directed ADMM algorithm. We define the ergodic average:

$$\bar{x}^k = \frac{1}{k} \sum_{i=0}^{k-1} x^i.$$

- Assume the sequence $\{x^k\}$ is bounded.

Theorem
Let (x^*, z^*, p^*) be a primal-dual optimal solution for problem (1). The following hold at each iteration k :

$$|F(\bar{x}^k) - F(x^*)| \leq \frac{1}{k} D, \quad \sum_{(i,j) \in E} |\bar{x}_i^k - \bar{x}_j^k| \leq \frac{1}{k} C,$$

where the constants C and D depends on number of nodes N , number of edges M , $\|p^*\|$, and the second largest eigenvalue of Laplacian.

Distributed ADMM over Directed Networks

Distributed ADMM over Directed Networks

- Each agent i maintains $x_i^k, \tilde{w}_i^k (w_i^k)$ and p_{ij}^k, z_{ij}^k , for $j \in \mathcal{N}^{in}(i)$.
- At iteration k ,

- a Agent i updates the primal variable x_i^k as

$$x_i^{k+1} \in \underset{x_i}{\operatorname{argmin}} f(x_i) - w_i^k \sum_{j \in \mathcal{N}^{in}(i)} (p_{ij}^k)^\top x_i + w_i^k \frac{\beta}{2} \sum_{j \in \mathcal{N}^{in}(i)} \|x_i - z_j^k\|^2,$$

The values x_i^{k+1} and \tilde{w}_i^k are sent to all out-neighbors.

- b For all $j \in \mathcal{N}^{in}(i)$, agent i computes

$$z_{ij}^{k+1} = \frac{1}{2}(x_i^{k+1} + x_j^{k+1}),$$

$$p_{ij}^{k+1} = p_{ij}^k - \frac{\beta}{2}(x_i^{k+1} - x_j^{k+1}),$$

$$p_{ij}^{k+1} = p_{ij}^k - \frac{\beta}{2}(x_i^{k+1} - x_j^{k+1}),$$

$$x_2^{k+1} = \frac{1}{2}(x_1^{k+1} + x_3^{k+1}),$$

$$p_{i2}^{k+1} = p_{i2}^k - \frac{\beta}{2}(x_2^{k+1} - x_2^{k+1})$$

$$\tilde{w}_i^{k+1} = (1 - d_i^{in} \delta_i) \tilde{w}_i^k + \sum_{j \in \mathcal{N}^{in}(i)} \delta_i \tilde{w}_j^k$$

$$w_3^{k+1} = (1 - d_3^{in} \delta_3) w_3^k + \sum_{j \in \mathcal{N}^{in}(i)} \delta_3 w_j^k$$

Distributed ADMM over Directed Networks

Conclusions and Future Work

- We presented distributed ADMM-based algorithms for solving multi-agent optimization problems over undirected and directed networks.
- For general convex cost functions, we showed that these methods converge at the faster rate $O(1/k)$ and provided rate estimates that highlighted dependence on network structure.
- Simulation results illustrate the superior performance of ADMM (even for network topologies with slow mixing).
- This comes at the expense of a more complex local optimization step.

- Extensions and Future Work:

- Asynchronous distributed ADMM algorithms [Wei, Ozdaglar 13].
- Broadcast-based ADMM algorithms [Mahdoumi, Ozdaglar 14].
- Graph based distributed optimization over directed networks.
- ADMM-type algorithms for time-varying networks.
- Distributed and incremental second order methods [Gurbuzbalaban, Ozdaglar, Parrilo 14].

**LEARNING GRAPHICAL MODELS: HARDNESS AND
TRACTABILITY**

**Devavrat Shah, Massachusetts Institute of Technology,
USA**

Our understanding of the question of learning graphical model, despite it's fundamental importance, has been far from satisfactory. In this talk, I will discuss recent progresses we have made towards improving this status quo. Specifically, we will discuss the reasons for learning graphical model being hard, both computationally and statistically. We shall also discuss, somewhat surprising, conditions under which it becomes tractable. Time permitting, we shall discuss a variant of the 'standard' graphical model learning problem from literature that might be more relevant to practice.

This talk describes joint work with Guy Bresler and David Gamarnik, both at MIT.

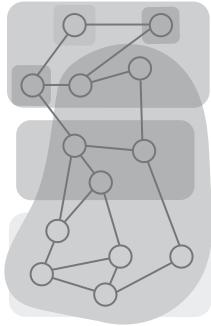


graphical models

**Learning graphical models:
hardness and tractability**

Guy Bresler David Gamarnik Devavrat Shah
MIT

$$G = (V, E) \quad |V| = p \quad |\partial i| \leq d$$



$$X_A \perp\!\!\!\perp X_B \mid X_S$$

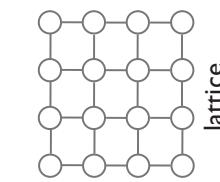
$$X_i \perp\!\!\!\perp X_{V \setminus \partial i \cup \{i\}} \mid X_{\partial i}$$

efficient inference

belief propagation can be used to do inference
historically people knew what models to use



hidden Markov model



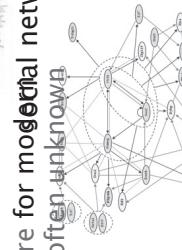
lattice
LDPC code

modern applications: unknown structure

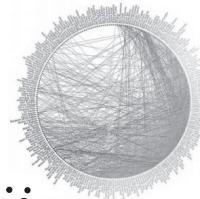


financial data

structure for modern neural network data
often unknown



gene regulatory network



learning graphical model

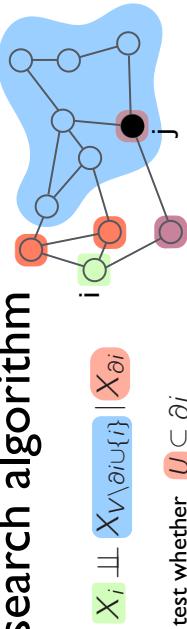
$G = (V, E)$
 $|V| = p$
 $|\partial i| \leq d$
 $X \in \{0, 1\}^p$

graphical model: $P(X) = \frac{1}{Z} \exp \left(\sum_{\{i,j\} \in E} \theta_{ij} X_i X_j + \sum_{i \in V} \alpha \right)$

data: $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ $X \sim P$ (i.i.d. samples)

task: reconstruct graph and parameters from the data
w.prob. $\rightarrow 1$ as $n, p \rightarrow \infty$

baseline: exhaustive search algorithm



test whether $U \subseteq \partial i$
 if $U \subseteq \partial i$ then for all $j \in U$ and $W \supseteq \partial i \setminus U$

$$|P(X_i = +1 | X_U = x_U, X_W = x_W) - P(X_i = +1 | X_U = \text{flip}_j(x_U), X_W = x_W)| \geq \frac{\tanh 2\alpha}{2e^{2(d-1)\beta}}$$

“ U passes test”

baseline: exhaustive search algorithm



test whether $U \subseteq \partial i$

if $U \subseteq \partial i$ then for all $j \in U$ and $W \supseteq \partial i \setminus U$

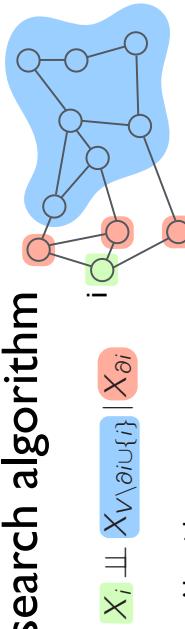
$$|P(X_i = +1 | X_U = x_U, X_W = x_W) - P(X_i = +1 | X_U = \text{flip}_j(x_U), X_W = x_W)| \geq \frac{\tanh 2\alpha}{2e^{2(d-1)\beta}}$$

Algorithm:
 for each node i
 test all possible neighborhoods U
 choose largest U passing test

“ U passes test”

Theorem:[Bresler-Mossel-Sly '08]
 algorithm recovers with prob. $1 - o(1)$ using
 $n = O(2^{2d} e^{(4\beta+n)d} \log p)$ samples, w runtime $\tilde{O}(p^{2d+1})$

“ U fails test”



Algorithm:
 for each node i
 test all possible neighborhoods U
 choose largest U passing test

baseline: exhaustive search algorithm



Algorithm:
 for each node i
 test all possible neighborhoods U
 choose largest U passing test

$$|P(X_i = +1 | X_U = x_U, X_W = x_W) - P(X_i = +1 | X_U = \text{flip}_j(x_U), X_W = x_W)| \geq \frac{\tanh 2\alpha}{2e^{2(d-1)\beta}}$$

Theorem:[Bresler-Mossel-Sly '08]
 algorithm recovers with prob. $1 - o(1)$ using
 $n = O(2^{2d} e^{(4\beta+n)d} \log p)$ samples, w runtime $\tilde{O}(p^{2d+1})$

“ U fails test”

our notion of computational efficiency

correlation decay

Theorem: [Bresler-Mossel-Sly '08]

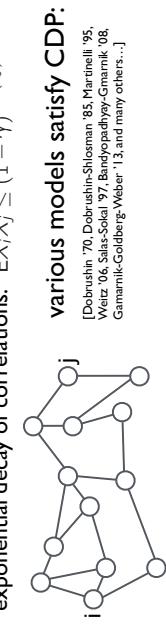
if have **correlation decay** and $E X_i X_j \geq \kappa$ for $\{i, j\} \in E$
 can learn using $n = O(2^{8d} e^{16\beta d} \log p)$ samples in time $O(np^2)$

exhaustive search: $p^{\Theta(d)}$

efficient:
 $f(d) p^c$:
 can be exponential
 indep. of d !

want to have
 no restrictions on
 graph structure

question: for what types
 of interactions can we
 learn efficiently?

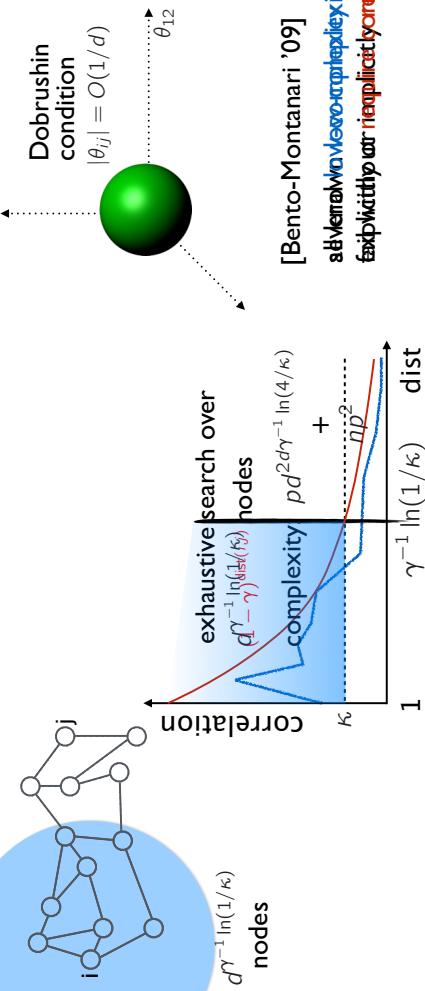


Theorem: if have **correlation decay** and $E X_i X_j \geq \kappa$ for $\{i, j\} \in E$:
 can learn using $n = O(2^{8d} e^{16\beta d} \log p)$ samples in time $O(np^2)$

correlation decay

other low-complexity
 approaches to learning:

- [Dobrushin '70, Dobrushin-Shlosman '85, Marinelli '95, Weitz '06, Salas-Sokal '97, Banerjee-Gamarnik '08, Gamarnik-Goldberg-Weber '13, and many others...]
- [Lee-Ganapathi-Koller '06]
- [Anandkumar-Tan-Luang-Willsky '12]
- [Ray-Singhvi-Shakkottai '12]
- [Wu-Srikant-Ni '13]
- [many many others]



repelling models

can we learn
efficiently without
correlation decay?

$$P(X) = \frac{1}{Z} \exp \left(\sum_{\{i,j\} \in E} \theta_{ij} X_i X_j + \sum_{i \in V} \theta_i X_i \right) \quad X \in \{0,1\}^p$$

$$\alpha \leq |\theta_{ij}| \leq \beta$$

repelling → $\theta_{ij} \leq -\alpha$
 $\theta_i \leq h$
 $\alpha \geq d(h + \ln 2)$

Theorem: [Bresler-Gamarnik-Shah '14a]

can learn these models with prob. $1 - o(1)$ using
 $n = O(2^{2d} e^{4\beta d} \log p)$ samples, with runtime $O(np^2)$

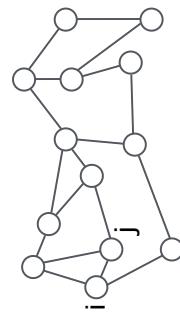
repelling models

Ex. Independent set model

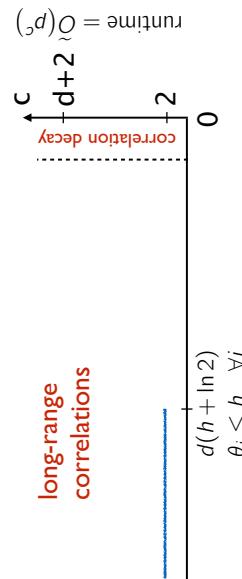
$$\alpha \rightarrow \infty, \text{ i.e. } \theta_{ij} \rightarrow -\infty$$

Key observation:

- i and j are neighbors: $(X_i, X_j) \neq (1, 1) \text{ w.p. 1}$
- i and j are not neighbors: $(X_i, X_j) = (1, 1) \text{ w.p. } \geq \gamma(d)$



repelling models



Theorem: [Bresler-Gamarnik-Shah '14a]

can learn these models with prob. $1 - o(1)$ using
 $n = O(2^{2d} e^{4\beta d} \log p)$ samples, with runtime $O(np^2)$

repelling models learning parameters

$$\alpha \geq d(h + \ln 2)$$

$$\alpha \geq (d - \tau)(h + \ln 2)$$

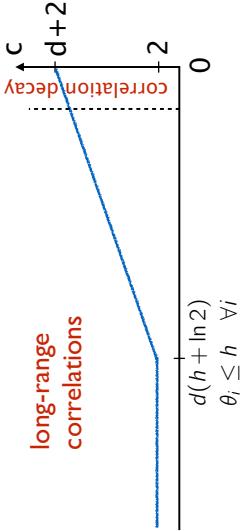
long-range correlations you know(s) the graph, learning parameters is

1. consider learning parameters

2. this allows to estimate **easy**

$$P(X_i = 1 | X_{\partial i} = 0) = \frac{e^{h_i}}{1 + e^{h_i}}$$

3. solve for h_i



Theorem: [Bresler-Gamarnik-Shah '14]

can learn these models with prob. $1 - o(1)$ using
 $n = O(2^{2d} e^{4\beta d} \log p)$ samples, with runtime
 ~~$O(\frac{(np)^2}{\epsilon})$~~
 $O(np^{2+\tau})$

can we do better in general

exhaustive search: $p^{\Theta(d)}$

a general approach to
simplifying:

reduce to sufficient statistics

Theorem: [Bresler-Gamarnik-Shah '14a]

No algorithm can do better than $p^{\Theta(d)}$ under the computation model of “statistical algorithms” in general.

reducing to sufficient statistics

$$P(X) = \frac{1}{Z} \exp \left(\sum_{\{(i,j)\} \in E} \theta_{ij} X_i X_j + \sum_{i \in V} \theta_i X_i \right) \quad X \in \{0,1\}^p$$

$$\begin{aligned} (\mu_i)_i &= (EX_i)_i = (P(X_i = i))_i \\ (\mu_{ij})_{ij} &= (EX_i X_j)_{ij} = (P(X_i = X_j = i))_{ij} \end{aligned}$$

try to estimate $\mu \mapsto \theta$ (feasible in principle!)

physicists try to estimate this map using various “expansions”

$$P(X) = \frac{1}{Z} \exp \left(\sum_{\{\{i,j\}\} \in E} \theta_{ij} X_i X_j + \sum_{i \in V} \theta_i X_i \right) \quad X \in \{0,1\}^p$$

(special case of repelling model)
(sufficient statistics)

Theorem: [Bresler-Gamarnik-Shah '14b] [Montanari '14]

learning parameters of graphical models from sufficient statistics is **NP-hard**

some remarks on proof

Reduction:

Suppose there exists efficient algorithm for $\mu \mapsto \theta$

Use it as a black-box to solve a known difficult problem

The difficult problem: given θ find corresponding $\mu \equiv \mu(\theta)$

For independent set with $\theta = 0$ this corresponds to

counting # of independent sets in G

a known hard (to approximate) problem

[Dyer-Frieze-Jerrum '02]

[Sly '10]

[Sly-Sun '12]

Reduction:

The difficult problem: given θ find corresponding $\mu \equiv \mu(\theta)$

Solve using black-box $\mu \mapsto \theta$

$$\mu(\theta) \in \arg \max_{\nu \in \mathcal{M}} (\nu, \theta) + H_{\text{ER}}(\nu)$$

Gradient ascent:

$$\mu^t \rightarrow \theta^t$$

$$\mu^{t+1} = \mu^t + \frac{1}{t} (\theta - \theta^t)$$

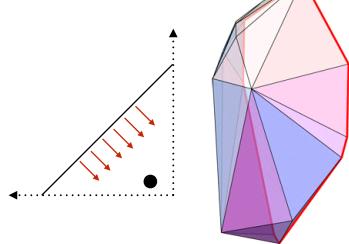

Key challenge:

μ^{t+1} needs to be projected on marginal polytope \mathcal{M}

some remarks on proof

what if algorithm naturally
avoids boundary

Lemma: [Bresler-Ganarnik-Shah '14b]
For the objective of interest, the
polytope boundary has an
inherent repulsion property



marginal polytope is
very complicated

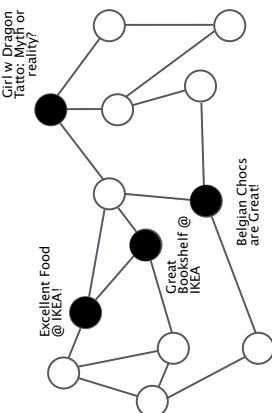
once you know the graph,
learning parameters is easy

graph tells you on which
higher order statistics to focus

learning from **sufficient statistics**
is probably not a good idea

what sort of data?

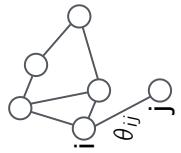
Social Behavior: Purchases, Likes, ...



so far: i.i.d. data
revisit original goal:
learning from data

dynamics over time

learning models from data



$$P(X) = \frac{1}{Z} \exp \left(\sum_{\{i,j\} \in E} \theta_{ij} X_i X_j + \sum_{i \in V} \theta_i X_i \right)$$

$$X \in \{0, 1\}^P \quad \alpha \leq |\theta_{ij}| \leq \beta$$

data: $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ i.i.d. samples
 n steps of some process

task: reconstruct graph and parameters from the data
w.prob. $\rightarrow 1$ as $n, p \rightarrow \infty$

slow mixing

i.i.d. sampling is **NP-hard** for some models
but Glauber dynamics defined for **any graphical model**

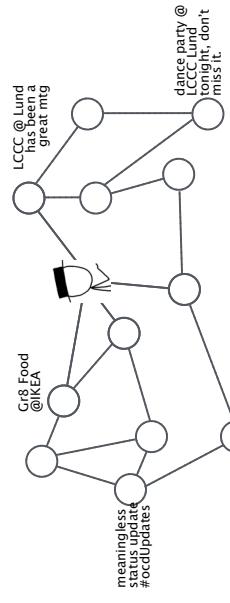
for models **without correlation decay**, the Glauber dynamics is known to **mix exponentially slowly** in p

Glauber dynamics

- 1. each node has a Poisson(1) clock

- 2. when clock rings, update variable according to

$$P(X_i = 1 | X_{\partial i}^t) = \frac{\exp \left(2 \sum_{j \in \partial i} \theta_{ij} X_j^t \right)}{1 + \exp \left(2 \sum_{j \in \partial i} \theta_{ij} X_j^t \right)}$$



efficient learning from the Glauber dynamics

Theorem: [Bresler-Gamarnik-Shah '14c] With $n = O(e^{4d\beta} \log p)$
samples per node, and runtime $O(np^2)$ can learn
any pairwise model even *without correlation decay*

learning theory:

- [Aldous-Yazirian '90]
- [Bartlett-Fischer-Hoffgen '94]
- [Bhattacharya-Mossel-'03]
- [Dahleh-Tsitsiklis-Zoumpoulis '13]
- [Netrapalli-Sanghavi '12]
- [O'Donnell-Servedio '03]

samples will be **far** from i.i.d.

estimating effect of a neighbor

imaginary scenario: node i updates,
then node j flips, then node i again
test for existence of an edge:

$$\exp(\theta_{ij}) = \frac{p^+(1-p^-)}{p^-(1-p^+)}$$

$$p^+ = \mathbb{P}(X_i = +1 | X_{\partial i \setminus j} = +1, X_j = +1)$$

$$p^- = \mathbb{P}(X_i = +1 | X_{\partial i \setminus j} = +1, X_j = -1)$$

this would require $\Omega(e^{d\beta} p^2)$ samples per node

a more delicate argument is required to get to $O_d(\log p)$

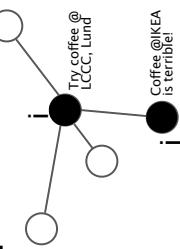
summary

correlation decay is not necessary to learn efficiently
however exhaustive algorithm seems the best in general

reducing to sufficient statistics is computationally suboptimal

observing dynamics over time can make things easy

insight: often makes sense to learn structure first
and only then estimate parameters



To/coffee @
ICCC, Lund
Coffee @IKEA
is terrible!

LOAD BALANCING USING LIMITED STATE INFORMATION**R. Srikant, University of Illinois at Urbana Champaign, USA**

Social network providers, search engine providers, and other big data companies operate massive data centers with thousands of processors handling data processing requests from incoming tasks. A central problem in such data centers is to design good load balancing algorithms using very little state information, where the state of each processor is the queue length of tasks waiting to be processed at that processor. We will present an algorithm which does remarkably well while sampling approximately only one queue per arriving task. This reduces the sampling requirement by half compared to previously known algorithms, while matching the delay performance of the best known algorithms. Joint work with Lei Ying and Xiaohan Kang



Load Balancing Using Limited State Information

R. Srikan†*
Joint work with Lei Ying* and Xiaohan Kang†

*University of Illinois at Urbana-Champaign
+Arizona State University

Data Centers are Large

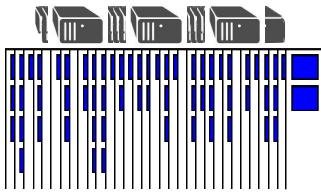


Yahoo! Hadoop cluster
42,000 nodes



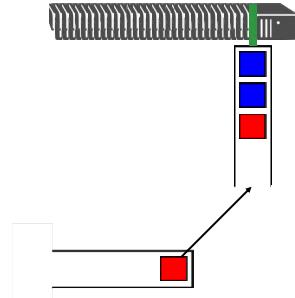
Load Balancing

n servers with unit service rate



- Arriving tasks have to be routed to a server
- Requirement: small delays
- Join-the-shortest-queue
- Expensive feedback overhead

Random Routing



- No overhead
- Delay $\sim \frac{1}{1-\rho}$
- ρ : Traffic Intensity, i.e., the ratio of the arrival rate of tasks to the maximum rate at which they can be processed by the servers

3

4

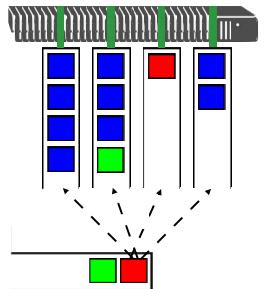
Power-of-Two-Choices

- Mitzenmacher 1996.
- Vvedenskaya, Dobrushin & Karpelevich 1996
- Delay, many servers limit

$$\sum_{l=1}^{\infty} \rho^{2^l - 2}$$

$$\rightarrow \log_2 \frac{1}{1-\rho} \quad (\rho \approx 1)$$

5



6

Key Question

- Question: sample **one** queue per arrival, **on average**, instead of two?
- Observation (Oosterhout et al., 2013): job arrivals occur in large **batches** (parallel tasks)



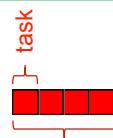
6

Model

- **n** servers
- Fixed batch size **m**

$$1 \leq m \leq n$$
- Poisson batch arrivals with rate

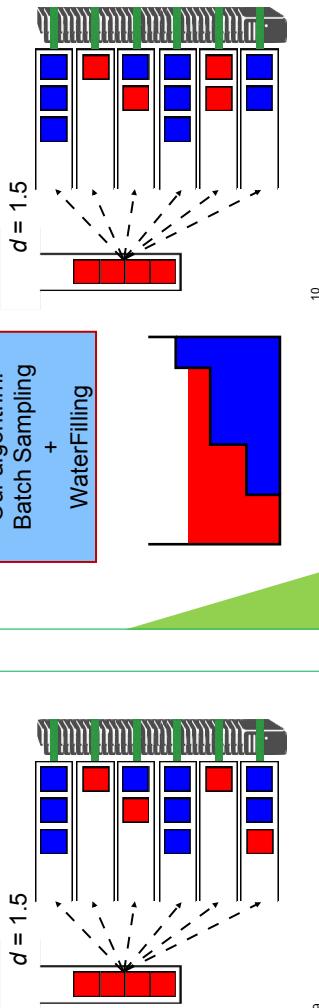
$$n\rho/m$$
- Note: Job is completed when all tasks in the job are completed
- Exponentially distributed service



Model and Main Results

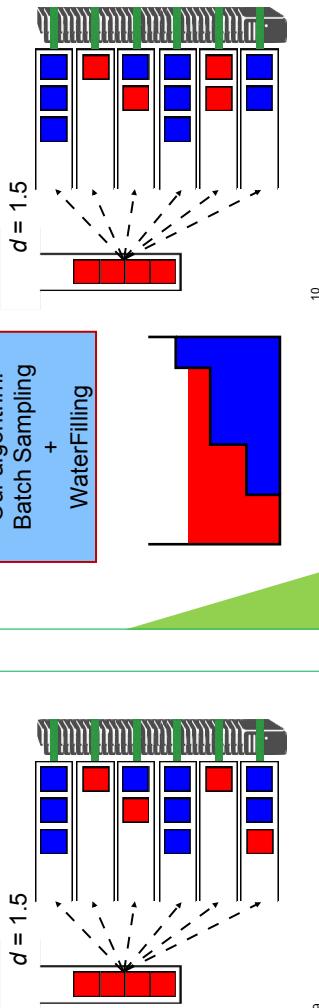
Batch-Sampling (BatchSamp)

- Ousterhout et al. 2013
- Probe ratio d
- One job in each of the smallest queues



Batch-Filling (BatchFill)

- Ousterhout et al. 2013
- Probe ratio d
- One job in each of the smallest queues



BatchSamp vs BatchFill

- Suppose we sample six queues to distribute five tasks; let's say that the queue lengths of the sampled queues are 0, 1, 2, 15, 20, 25
- Under BatchSamp, the resulting queue lengths are **1, 2, 3, 16, 21, 25**
- Under BatchFill, they are **2, 3, 3, 15, 20, 25**

BatchFill

- Many-server heavy-traffic delay

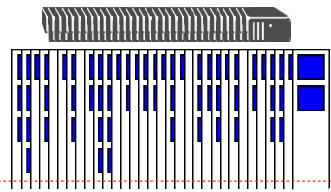
$$\frac{\log \frac{1}{1-p}}{\log(1+d)}$$

- Queue length upper-bounded by

$$\left\lceil \frac{\log \frac{1}{1-p}}{\log(1+\rho d)} \right\rceil$$

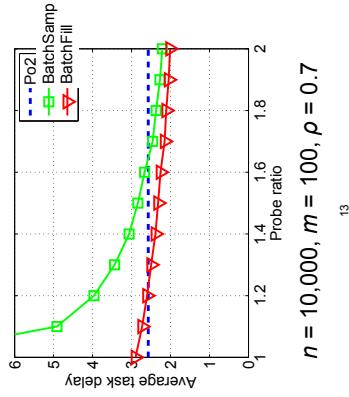
$$\rho = 0.9, d = 1.1, \text{ upper bound } 4$$

12

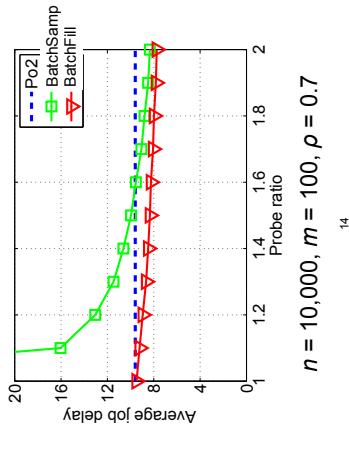


12

Simulation Results: Average Task Delay



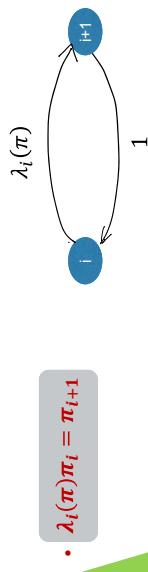
Simulation Results: Average Job Delay



Mean Field Approximation

General Idea

- Focus on a particular queue: Assume all other queues are in steady-state (stationary distribution π) and independent of each other
- The arrival rate when there are i tasks in the queue is a function of the state of the other queues, and hence a function of π



Calculating $\lambda_i(\pi)$: Po2

- Task arrival rate $n\rho$
- Prob(a particular queue is sampled) = $2/n$
- Probability being chosen $\pi_i/2 + \pi_{i+1} + \pi_{i+2} + \dots$
- $\lambda_i = \rho(\pi_i + 2\pi_{i+1} + 2\pi_{i+2} + \dots)$

Stationary Distribution: Po2

- Queue length distribution: $\pi_l = \rho^{2^l - 1} - \rho^{2^{l+1} - 1}$

$$\text{Delay: } \frac{\sum_{l=1}^{\infty} l \pi_l}{\rho} = \sum_{l=1}^{\infty} \rho^{2^l - 2}$$

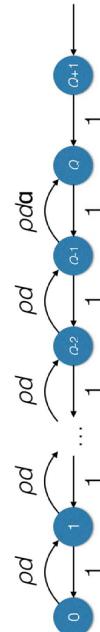
18

BatchSamp

- No waterfilling

- Reducible to finite states

$$0 < \alpha \leq 1$$



19

Why is the queue finite?

- $m d$ out of n queues are sampled

- We see the empirical distribution: $m d \pi_0$ empty queues, $m d \pi_1$ queues with one task, etc.
 $0, 0, 1, 1, 1, 1, \dots, j, j, j, j, \dots$

- The first m of these queues gets one task each

- Under the mean-field approximation, the number of tasks in the m^{th} queue is fixed, as a function of π

π for BatchSamp

- $\pi_0 = 1 - \rho$
- $\pi_i = (1 - \rho) \rho^d i!, \quad 1 \leq i \leq Q - 1$
- $\pi_Q = 1 - (1 - \rho)(\rho^Q d^{Q-1}) / (\rho d - 1)$

• Delay $\frac{\sum_{i=1}^{\infty} i \pi_i}{\rho} \approx \frac{\log \frac{1}{1-\rho}}{\log d}$ (heavy-traffic)

- When $d = 2$, asymptotically equivalent to Po2

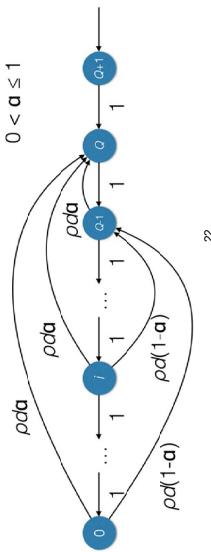
21

BatchFill

Cutoff queue length

$$Q = \left\lceil \frac{\log \frac{d-1}{\rho(1-\rho)}}{\log(\rho d)} \right\rceil$$

- Constant up-crossing rate to Q or $Q-1$, determined by the stationary distribution π



Justifying the Mean-Field Approximation

Cutoff queue length

$$Q = \left\lceil \frac{\log \frac{1}{1-\rho}}{\log(1 + \rho d)} \right\rceil$$

π for BatchFill

- $\pi_0 = 1 - \rho$
- $\pi_i = (1 - \rho) \rho d (1 + \rho d)^{i-1}, \quad 1 \leq i \leq Q - 1$
- $\pi_Q = 1 - (1 - \rho) (1 + \rho d)^{Q-1}$

• Delay $\frac{\sum_{i=1}^{\infty} i \pi_i}{\rho} \approx \frac{\log \frac{1}{1-\rho}}{\log(1 + d)}$ (heavy-traffic)

- Better asymptotic delay than Po2 when $d > 1$!

23

Ordinary Differential Equation (ODE) for BatchFill

- Deriving an ODE: Derivative is given by

$$\lim_{n \rightarrow \infty} \lim_{\delta \rightarrow 0^+} E \left(\frac{\text{change in state}}{\delta} \middle| \text{current state} \right)$$

- ODE

$$\frac{dx_i}{dt} = \begin{cases} -(1 + \rho d)x_i + x_{i+1} & i \leq \bar{X}_x - 2 \\ \rho d(1 - \alpha_x) \sum_{j=0}^{i-1} x_j - (1 + \rho d\alpha_x)x_i + x_{i+1} & i = \bar{X}_x - 1 \\ \rho d\alpha_x \sum_{j=0}^i x_j - x_i + x_{i+1} & i = \bar{X}_x \\ -x_i + x_{i+1} & \text{otherwise} \end{cases}$$

25

Global Asymptotic Stability of ODE

- $s(t)$ fraction of servers with queue size $\geq i$
- Lyapunov function

$$V(s) = \sum_{l=1}^{\infty} |s_l - \hat{s}_l|$$

- $s(t)$ converges to the equilibrium point for any $s(0)$

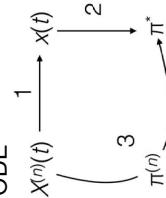
26

Justifying the Mean-Field Approximation

- The ODE approximation (in a finite time interval) works well. The deviation from the ODE goes to zero as n goes to infinity

- Global asymptotic stability of the ODE

- Interchange of limits



27

Conclusions

- One sample can be powerful in randomized load-balancing
- Batch arrivals can be exploited to reduce sampling overhead
- Extensions: (i) Variable batch sizes, (ii) Batch arrivals are not necessary, (iii) General service-time distributions and Processor Sharing

28

**ON THE EIGENVALUES OF LARGE STRUCTURED MATRICES
AND THE SCALABLE STABILITY OF NETWORKS**
Glenn Vinnicombe, University of Cambridge, UK

A class of scalable stability results for interconnections of linear dynamical systems depend ultimately on tools for locating the eigenvalues of large structured matrices using only local information. This talk will summarise some existing results in this area and introduce some open problems.



Scalable Design Rules for Physical Systems

Richard Pates¹, and Glenn Vinnicombe²,

$$e_k = Z_k(s) f_k$$

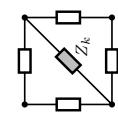
$$\text{or } e_k = \phi_k(f_k)$$

$$\xrightarrow[e_k]{}$$

$$\xleftarrow[f_k]{}$$

$$\xleftrightarrow{\text{Component}}$$

$$\xrightarrow{\quad}$$



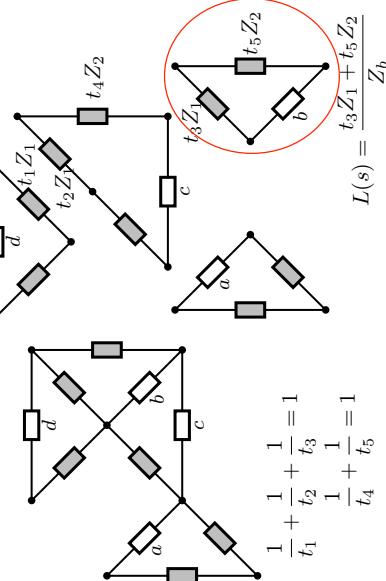
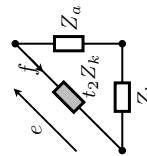
$$\frac{1}{t_1} + \frac{1}{t_2} = 1$$

$$e = t_2 Z_k(s) f$$

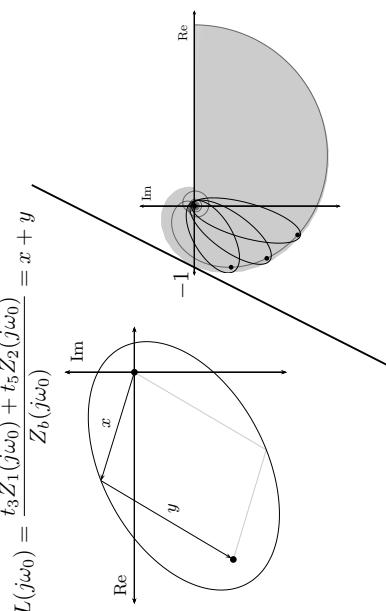
$$-f = \frac{1}{(Z_a(s) + Z_b(s))} e$$

$$\Rightarrow \text{use Nyquist with } L(s) = \frac{t_2 Z_k(s)}{Z_a(s) + Z_b(s)}$$

(we will be assuming that Z_k is open circuit stable and Z_a, Z_b short circuit stable).



$$L(j\omega_0) = \frac{t_3 Z_1(j\omega_0) + t_5 Z_2(j\omega_0)}{Z_b(j\omega_0)} = x + y$$



Network is stable if this *and all other* curves lies to the right of a globally specified line through -1

So, what's going on?

Return Ratio of form GA

$$G = \begin{bmatrix} g_1 & 0 & \cdots & 0 \\ 0 & g_2 & & \\ \vdots & & \ddots & \\ 0 & \cdots & & g_n \end{bmatrix}, \quad [A]_{ij} \in \mathbb{R}$$

If $A = A^T > 0$ then

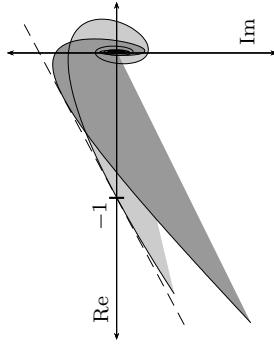
$$\sigma(GA) \in \text{Co}\{g_i\}\rho(A) \quad \text{e.g. V (2000)}$$

If $g_1 = g_2 = \cdots = g_n = g$ then

$$\sigma(GA) = g\sigma(A) \quad \text{e.g. Fax \& Murray (2001)}$$

e.g. $A = R \text{diag}\{l_i\}R^T$ or $A = \text{a Laplacian (consensus problems)}$.

Can do Popov too!



Can these be put together?

$$\begin{aligned} AGv &= \lambda v \\ Gv &= A^{-1}v \\ \implies v^*Gv &= \lambda v^*A^{-1}v \end{aligned}$$

So $W(G) \cap W(A^{-1}) = \emptyset \implies$ no eigenvalue at λ
(where $W(X) = \left\{ \frac{v^*Xv}{v^*v} : v \in \mathbb{C}^n, v \neq 0 \right\}$)

also

$$v^*G^*Gv = \lambda^2 v^*A^{-*}A^{-1}v$$

So $DW(G) \cap DW(A^{-1}) = \emptyset \implies$ no eigenvalue at λ

(Jönsson and Kao 2010, Lestas 2012)

(where $DW(X) = \left\{ \Re \frac{v^*Xv}{v^*v}, \Im \frac{v^*Xv}{v^*v}, \frac{v^*X^*Xv}{v^*v} : v \in \mathbb{C}^n, v \neq 0 \right\}$)

Also

$DW(G) \cap DW(A^{-1}) = \emptyset \iff DW(G^{-1}) \cap DW(A) = \emptyset$

What about neighbouring dynamics

Could consider $\sqrt{G}A\sqrt{G}$, but strongest results are in the bipartite case:

$$\begin{aligned} \text{e.g. } G &= \text{diag}(f_1, f_2, \dots, h_1, h_2, \dots) \quad A = [0 \ R \ R^T \ 0], \\ A_{ij} &\in \{-1, 0, 1\} \end{aligned}$$

$$\sigma(\text{diag}(g_i)R^T \text{diag}(h_i)R) \subset \text{Co}\{m_i h_i S(n_j g_j : R_{ij} \neq 0)\}$$

where $S(X) = \text{Co}(\sqrt{X})^2 \quad [\text{V (2002), Lestas \& V (2006)]}$

A better result is

$$\sigma(\text{diag}(g_i)R^T \text{diag}(h_i)R) \subset \text{Co}\{h_i E(n_j g_j : R_{ij} \neq 0)\}$$

where $E(x_i)$ is ellipse with foci 0, $\sum x_i$ and major axis
 $\sum |x_i|$. [Pates & V (2012)]

proof :

$$\begin{aligned} \sigma(\text{diag}(g_j)R^T \text{diag}(h_i)R) &= \sigma(\text{diag}(\sqrt{g_j})R^T \text{diag}(h_i)R \text{diag}(\sqrt{g_j})) \\ &= \sigma\left(\sum_i \text{diag}(\sqrt{g_j})R_i^T h_i R_i \text{diag}(\sqrt{g_j})\right) \\ &\quad \text{etc} \end{aligned}$$

Open question: Local conditions for eigenvalue locations

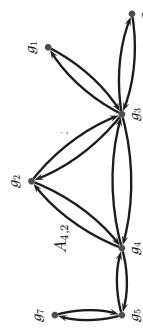
Richard Pates & Glenn Vinnicombe

$$G = \begin{bmatrix} g_1 & 0 & \dots & 0 \\ 0 & g_2 & & \\ \vdots & & \ddots & \\ 0 & \dots & & g_n \end{bmatrix}, \quad [A]_{ij} \in \{0, -1, 1\}$$

Identify A with the adjacency matrix of a *directed* graph, with $g_i \in \mathbb{C}$ labelling the nodes and R_{ij} labelling the edges. What can we say about the $\sigma(GA)$ in terms of *local information* about the cycles of A (including those of length 2)?

Each node knows all cycles it participates in, and for each of those cycles g_i, A_{ij}, n_i along that cycle.

Is it possible, for some region \mathcal{B} , to come up with a yes/no question such that if all nodes say "yes", based on their local information, then $\sigma(GA) \in \mathcal{B}$? What is the smallest \mathcal{B} (and associated question).



What we know (bipartite & symmetric)

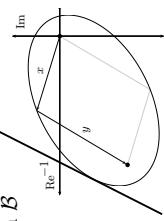
If

$$G = \text{diag}(f_1, f_2, \dots, f_k, h_1, h_2, \dots), A = [0 \ R; R^T \ 0], A_{ij} \in \{0, 1\}$$

$$\sigma(GA)^2 \subset \text{Co}\{f_i E(n_j h_j : R_{ij} \neq 0)\}$$

where $E(x_i)$ is ellipse with foci 0, $\sum x_i$ and major axis $\sum |x_i|$, n_j = in-degree of node h_j .

For \mathcal{B} being the region to the right of a given, line through -1 , and question is "does your ellipse lie in \mathcal{B} " If all answer "yes", then $\sigma(GA) \in \mathcal{B}$



**TOWARD STANDARDS FOR DYNAMICS IN FUTURE
ELECTRIC ENERGY SYSTEMS: THE BASIS FOR PLUG-AND-
PLAY INDUSTRY PARADIGM**

Marija Ilic, Carnegie Mellon University, USA



In this talk we propose that by taking a step back and understanding the fundamental physics of interconnected electric power systems one can begin to think systematically about the role of both competitive and cooperative control, and their pros and cons for these rapidly evolving systems. To support this, we organize our presentation in several parts: 1) physics and operation of power grids; 2) the role of control; 3) systematic specification of performance objectives (dynamic standards) for plug-and-play operations; 4) use of cyber for implementing standards (sensing, communications and control architectures for supporting implementation of these standards); and, 5) illustration of how several industry problems can be solved when following proposed plug-and-play standards for dynamics (use of power electronics for transient stabilization during faults and sudden equipment failures; use of PMUs for ensuring voltage and frequency quality by an intelligent balancing authority (iBA) – generalization of today's control areas; storage control in microgrids/systems with small inertia).

In the first part we give a somewhat new, dynamic systems view, of physical operation in several qualitatively different power grids (bulk power regulated; bulk power with markets; hybrid mix of emerging grids; micro-grids for developing countries; and micro-grids for developed countries). The physical operation dictates how to define internal states of given (groups of) physical components and the interaction variables between different (groups of) physical components. Based on understanding physical principles, we propose a new state space model of an interconnected grid comprising (groups of) different physical components. This model has a transparent physical interpretation, and, it is, therefore used as the basis for explicit performance specifications in terms of interactions of components. Performance specifications become standards for plug-and-play dynamic interactions of components (behavior) over several time horizons within an otherwise complex dynamical grid, which, when followed, ensure system-wide performance. We emphasize that the proposed approach is a framework for thinking about the necessary specifications in future electric energy systems, which enables both choice of technology at the (groups of) component levels and, minimal interaction specifications to align these with the performance of the overall system. It is not a specific method. This framework could possibly overcome the roadblock of integrating distributed local grid technologies with the bulk power grid without running into major coordinating complexity. Specific proposed approaches to distributed control for smart grids are interpreted.



Outline

- Overview of current NERC standards and evolving standards for wind and solar plants
- Issues with current standards
- Our proposal :
 - ❖ Plug-and-play (TCP/IP) like protocols/standards
 - ❖ Introduction of intelligent Balancing Authority (IBAs)
- Examples of iBAs
- Theoretical foundations for new standards (TCP/IP like)
- Proof-of-concept examples of controller designs which meet such protocols

Marija Ilic milic@ece.cmu.edu
 Electric Energy Systems Group (EEESG) <http://www.eesg.ece.cmu.edu/>, Director
 Invited talk, LCCC Workshop
 Lund, Sweden
 October 15-17, 2014



Toward standards for dynamics in future electric energy systems— The basis for plug-and-play industry paradigm

NERC standards Transmission Planning Standards

- System simulations and associated assessments are needed periodically to ensure that reliable systems are developed that meet specified performance (<http://www.nerc.com>)

Category	Contingencies	System Stable and Both Thermal and Voltage Limits Within Applicable Rating	Loss of Demand
A	No contingency	Yes	No
B	Event(s) resulting in the loss of a single element.	Yes	No
C	Event(s) resulting in the loss of two or more (multiple) elements	Yes	Planned/ Controlled
D	Extreme event resulting in (two or more (multiple) elements removed or Cascading out of service,	Evaluate for risks and consequences. May involve substantial loss of customer demand and generation in a widespread area or areas. - Portions or all of the interconnected systems may or may not achieve a new stable operating point. - Evaluation of these events may require joint studies with neighboring systems.	

Evolving standards for Wind and Solar Generation Technologies

- voltage/var control/regulation
- voltage ride-through
- power curtailment and ramping
- primary frequency regulation
- inertial response

http://www.nerc.com/files/2012_IGTF_Task_1-3.pdf
 NERC 2012 Special Assessment: Interconnection Requirements for Variable Generation September 2012

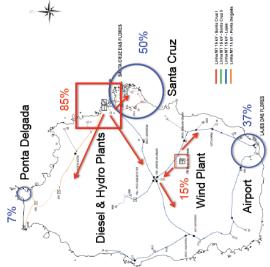


Need for a new paradigm

- ❖ Today's industry approach—the worst case approach, inefficient and does not rely on on-line automation and regulation other than energy feed-forward economic dispatch
- ❖ Emphasis on large-scale time-domain system simulations for transient stability, voltage, collapse, power flow feasibility, etc
- ❖ Primary control is constant gain tuned assuming no dynamic interactions with the rest of the system
- ❖ Existing and emerging system-level unacceptable interactions; no incentives for ``smarts'' of modules

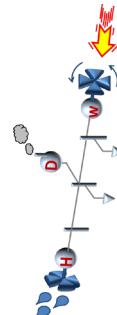
 Carnegie Mellon

From old to new paradigm—Flores Island Power System, Portugal [11]



 Carnegie Mellon

Controllable components—today's operations (very little dynamic control, sensing)



*Sketch by Milos Cvetkovic
 Carnegie Mellon

Two Bus Equivalent of the Flores Island Power System



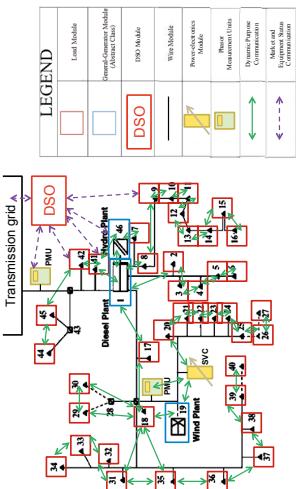
	Generator	Diesel	From Diesel to Load Bus
$x_0 [\mu\text{u}]$	8.15	$R [\mu\text{u}]$	$S_c = 100 \text{ MVA}$
$t_0 [\mu\text{u}]$	0.3071	$V_b [\text{V}]$	$V_b = 15 \text{ kV}$
$t_1 [\mu\text{u}]$	0.1695		

	AVR	Diesel	From Diesel to Load Bus
$x_0' [\mu\text{u}]$	0.5917	$K_d [\mu\text{u}]$	$S_c = 100 \text{ MVA}$
$T_{av} [\mu\text{u}]$	40	$T_d [\mu\text{u}]$	$V_b = 15 \text{ kV}$
$T_{av} [\text{s}]$	0.02	$T_d [\text{s}]$	
$T_{av} [\text{s}]$	2.35	$T_d [\text{s}]$	
$J [\text{kg}]$	2.35	$K_d [\mu\text{u}]$	
$J [\text{s}]$	0.226	$K_d [\mu\text{u}]$	
$f [\text{s}]$	0.005	$T_d [\text{s}]$	
$v_0 [\mu\text{u}]$	1	$T_d [\text{s}]$	
$v_1 [\mu\text{u}]$	0.8527	$T_d [\text{s}]$	
$v_2 [\mu\text{u}]$	0.7432	$T_d [\text{s}]$	
$v_3 [\mu\text{u}]$	0	$T_d [\text{s}]$	
$p_m [\mu\text{u}]$	0.03	$T_d [\text{s}]$	
$p_n [\mu\text{u}]$	0.01	$T_d [\text{s}]$	
$q_m [\mu\text{u}]$	0	$T_d [\text{s}]$	
$q_n [\mu\text{u}]$	1	$T_d [\text{s}]$	

Base values \$S_c = 100 \text{ MVA}, V_b = 15 \text{ kV}\$
 Carnegie Mellon

 Carnegie Mellon

Information exchange in the case of Flores---new (lots of dynamic control and sensing)



Possible dynamical problems seen by particular dynamic components

Types of Component	Dynamical problems						
	Small Signal Instab.	Transient Instab.	SSR	SSCI	Freq. Instab.	Volt. Instab.	Power flow imbalance
Synchronous generators	?	?	?	?	?	?	?
Wind generators	?	?	?	?	?	?	?
Solar plants	?	?	?	?	?	?	?
FACTS	?	?	?	?	?	?	?
Storage	?	?	?	?	?	?	?

Table 1.



Our proposal: TCP/IP like standards

- ❖ Given specified disturbances and range of operating conditions within a known system:

- specified with e.g. voltage, power
- similar to LVRT curves for wind turbines

- with specified duration

- ❖ All components (synchronous gens, wind gens) should guarantee that they would not create any of the problems in Table 1. (Clear objectives goals for components, assigned responsibility for system reliability)

- ❖ Two key questions: Q1 -- Why does it matter?
Q2 -- Can this be technically done?

Not one way to achieve these!



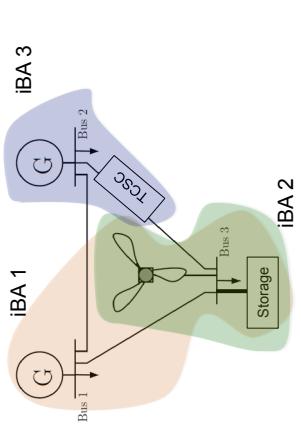
Not one way to meet the standards -iBAs

- ❖ iBAs (Intelligent Balancing Authorities)
 - = Single component or group of components which meet the desired objectives: Given specified disturbances their components do not cause any of the dynamical problems in Table 1.
 - = Dynamic notion of Control Areas—Intelligent Balancing Authorities (IBAs)
- ❖ iBAs would utilize advanced control design methods to meet the protocol; could be either decentralized or wide area control (cooperative control to save on number of controllers and energy used within the iBA)
- ❖ Huge potential for exploiting efficiently new technologies like storage and FACTS and at the same time have guaranteed system performance

S.Baroš, Milic Intelligent Balancing Authorities (IBAs) for Transient Stabilization of Large Power Systems IEEE PES General Meeting 2014

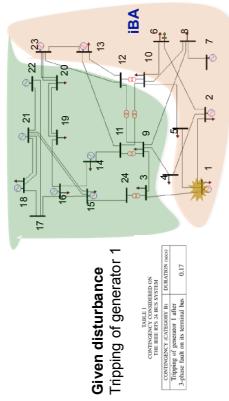


A1: Examples of iBAs—it matters for ensuring both reliable and efficient operations [13]



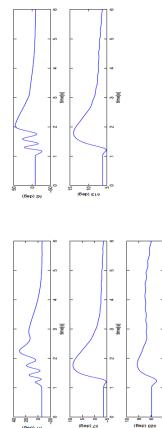
Carnegie Mellon¹³
IEEE

Possible to create iBAs for meeting transient stability distributed standard



S. Baros, Millc intelligent Balancing Authorities (iBAs) for Transient Stabilization of Large Power Systems IEEE PES General Meeting 2014
Carnegie Mellon¹⁴
IEEE

Rotor angle response of iBA generators



(b) Transient stabilization of critical generators $i=1,2,7,13,23$ with iBA-based control in low-load scenario

Carnegie Mellon¹⁵
IEEE

Q2: Can we have a unifying theoretically sound approach to TCP/IP like standards for smart grids? [12]

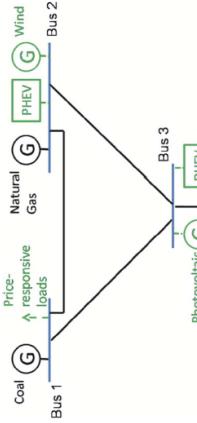


Fig. 5. Small example of the future electric energy system.

Carnegie Mellon¹⁶
IEEE

Basic functionalities

- ❖ Simple transparent TCP/IP like functionalities
- ❖ Transparency based on a unifying modular modeling of network system dynamics
- ❖ **Provable performance-difficult**
- ❖ **Proposal—use interaction variables to specify family of standards sufficient to avoid operating problems**
 - Measure of how well modules balance themselves in steady state
 - Measure of rate of exchange of stored energy between a module and the rest of the system over different time horizons

 Carnegie Mellon 17

Unifying modeling and control approach—use of multi-scale interaction variables

- ❖ Standards/protocols --- specifications of module interactions for plug-and-play operations; architectures define how are sets of protocols organized
- ❖ Cyber design for managing multi-layered interactions
- ❖ New physics-based modeling and control as the basis for **interaction variables-based protocols**
- ❖ Illustrations of possible standards-based enhancements (transient stabilization using power electronics switching; storage control in micro-grids)

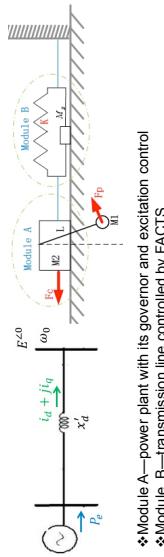
 Carnegie Mellon 18

New physics-based modeling and control using interaction variables [12]

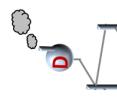
- ❖ Mechanical system representation of electric power grids
- ❖ Physics-based state transformation for multi-layered dynamics
- ❖ Defining interaction variables over different time horizons (to capture bounds on change in stored energy over T and on the rate of change of stored energy)
- ❖ Multi-layered specifications for interaction variable-based standards

 Carnegie Mellon 19

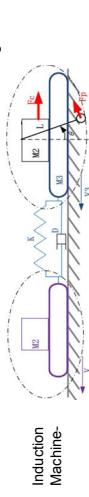
Two-module power system in Flores



- ❖ Module A—power plant with its governor and excitation control
- ❖ Module B—transmission line controlled by FACTS



Interconnected power system and its mechanical representation



Conventional generator

Analogy Table

Analogies		
	Mechanical Quantity	Power System Analogue
❖ Component level	Force, F_p	Mechanical Power Input, P_m
	Velocity, V	Current, I
	Force, F_c	Exciter input, F_{exc}
	Force between M_2 & M_3 , F_{32}	Back EMF
❖ System level		
	Mechanical Quantity	Power System Analogue
	Velocity, V_3	Terminal Voltage, V
	Interaction Force, F_s	Current, I

Carnegie Mellon²¹ ↗Carnegie Mellon²² ↗Carnegie Mellon²³ ↗Carnegie Mellon²⁴ ↗

Very intriguing questions

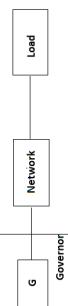
- ❖ Can this system synchronize around inverted pendulum position? Is this acceptable? What should be a standard/protocol for plug-and-play?
- ❖ Which controller works? Why?
- ❖ Answers depend on the actual operating conditions (close to the stable or unstable equilibrium); on control laws applied to F_p and/or F_c ; if the equilibrium is not known (stabilization) energy-based controller will work, while constant gain controller will not. It will also depend on the load model used. The control effort and quality of response vary.
- ❖ Not all “smarts” are the same in smart grids!

Three different architectures of interest

- ❖ Bulk power system can be represented as an infinite bus (ideal voltage source)
- ❖ Flores Island (micro-grid)—diesel power plant has dynamics—hydro and wind negative constant power load
- ❖ Flores Island (micro-grid)-diesel power plant dynamics and the load is constant impedance

Carnegie Mellon²³ ↗Carnegie Mellon²⁴ ↗

Governor of AC Generator



- ❖ Governor is the main synchronizing controller.
- ❖ In mechanical representation, F_p responds to the deviations of speed from synchronous speed.
- ❖ Does not synchronize around inverted pendulum position.
- ❖ It does synchronize when F_p also responds to $x_3 \rightarrow x_3$ (phase lock loop).

Carnegie Mellon 25

Constant power load model



- ❖ **Unstable Equilibrium Set**
- $\begin{cases} \dot{\theta}_1^{(2)} = 2.7419 \\ \dot{\theta}_2^{(2)} = -149.348 \end{cases}$
- ❖ **System Matrix**
- $A_{sysw} = \begin{bmatrix} 0 & 377 \\ 1.398 & -0.0221 \end{bmatrix} \quad \begin{cases} \lambda_1 = 22.9569 \\ \lambda_2 = -22.9569 \end{cases}$
- ❖ **Participation Factor**
- $P = \begin{bmatrix} 0.5 & 0.4998 \\ 0.4998 & 0.5 \end{bmatrix}$
- ❖ θ and ω contribute same to the unstable eigenvalue which means governor should take care of θ as well as ω (PI control)

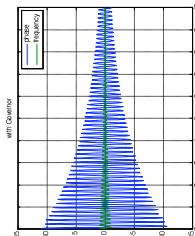
Carnegie Mellon 26

Synchronized system to constant power load



- ❖ **Governor Design**
- $Governor: K_s(\omega - \omega^*) + K_r(\theta - \theta^*)$
- ❖ **System Matrix**
- $A_{load} = \begin{bmatrix} 0 & 377 \\ -3.03 & -0.46 \end{bmatrix}$
- $\begin{cases} \lambda_1 = -0.23 + 33.77j \\ \lambda_2 = -0.23 - 33.77j \end{cases}$
- ❖ **System response**

Stable



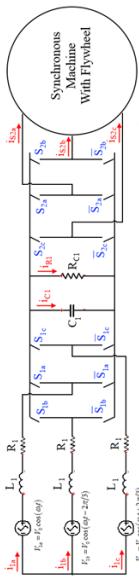
Carnegie Mellon 27

Second control –fast Fc

- ❖ Exciter in power plant
- ❖ Fc responds to acceleration of rotating pendulum
- ❖ Mechanical analogy—control acceleration of the pendulum pivot
- ❖ Internal interactions non-unique—can control with F_p , F_c or a combination (competitive at the module A level)
- ❖ Module B could also control stiffness of the spring ***FACTS is controllable inverter** to help stabilize module A (cooperative control)

Carnegie Mellon 28

Gets complicated fast— Variable Speed Drives for Flywheels

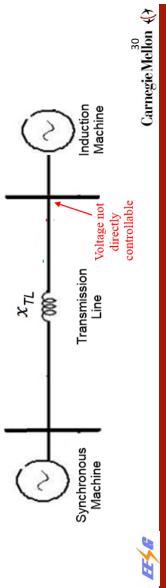


- ❖ Stator voltages of synchronous machine with flywheel are NOT assumed to be directly controllable
- ❖ The only controllable inputs are the switch positions of the power electronics
- ❖ Use AC/DC/AC converter to regulate the speed of the flywheel to a different frequency than the grid frequency

Source: K. D. Bachovchin, M. D. Ilic, "Passivity-Based Control Using Three Time-Scale Separations of Variable Speed Drives for Flywheel Energy Storage Systems," IEEE Working Paper No. R-WP-6-2014, October 2014. Carnegie Mellon 29

Standalone vs. Interconnected Variable Speed Drives

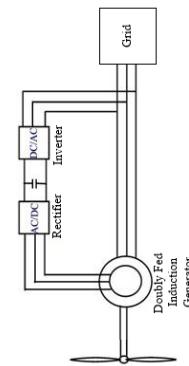
- ❖ For standalone variable speed drives, the voltage of the stator windings is directly controllable
- ❖ For interconnected systems, the voltage of the stator windings is not directly controllable and depends on the dynamics of the rest of the system
- ❖ Need power electronics (AC/DC/AC converter) in order to control the stator voltages



Carnegie Mellon 30

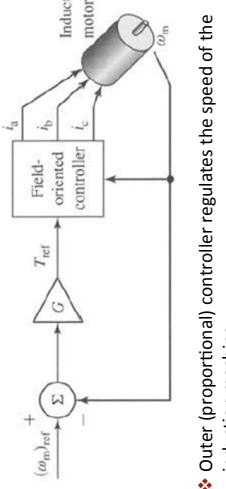
Doubly-fed induction machines (DFIG)—wind power plant

- ❖ AC/DC/AC converter interfacing between the stator windings and the rotor windings of the DFIG
- ❖ The rotor voltage is controlled through switches in the AC/DC/AC converters



Carnegie Mellon 31

Field Oriented Control for Variable Speed Drives



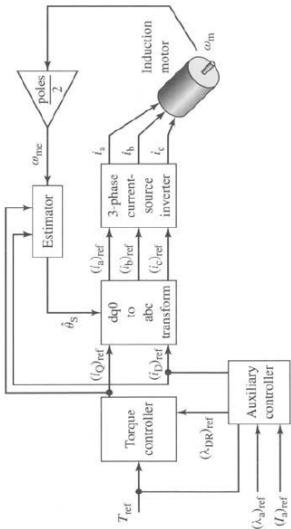
- ❖ Outer (proportional) controller regulates the speed of the induction machine
- ❖ Inner (field oriented) controller regulates the torque to the set point given by outer controller

Source: A. Fitzgerald, C. Kingsley Jr., and S. Umans, *Electric Machinery*, Piscataway, NJ: McGraw-Hill, 2003.

Carnegie Mellon 32

Carnegie Mellon 33

Field Oriented Controller Block Diagram

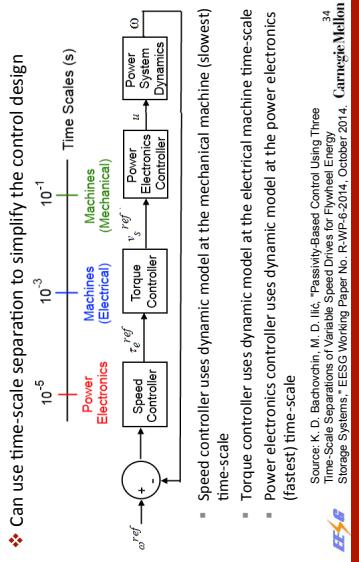


Source: A. Fitzgerald, C. Kingsley Jr., and S. Umans, *Electric Machinery*. Piscataway, NJ: McGraw-Hill, 2003.



Carnegie Mellon

Controller Implementation--detailed

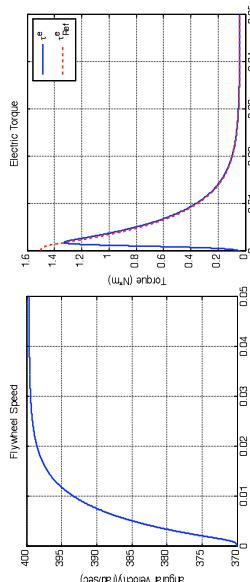


Source: K. D. Bachovchin, M. D. Ilic, "Passivity-Based Control Using Three Time-Scale Separations of Variable Speed Drives for Flywheel Energy Storage Systems," EESG Working Paper No. R-WP-6-2014, October 2014. Carnegie Mellon



Carnegie Mellon

Control design—can be done



Can control speed of the flywheel to set point $\omega_2 = 400 \text{ rad/sec}$



Carnegie Mellon

Must simplify!

- ❖ Utilities are having hard time adding all these new components and their smarts for simulating system-wide dynamics
- ❖ Is there a “smarter” way to model and define modular functionalities so that the interconnected system meets system-level performance (Table 1)?

❖ 80% of each solution is modeling (Petr Kokotovic, Challenges in Control Theory, Santa Clara, circa 1982)

EE&G
Carnegie Mellon

- ❖ Utilities are having hard time adding all these new components and their smarts for simulating system-wide dynamics
- ❖ Is there a “smarter” way to model and define modular functionalities so that the interconnected system meets system-level performance (Table 1)?

❖ 80% of each solution is modeling (Petr Kokotovic, Challenges in Control Theory, Santa Clara, circa 1982)

EE&G
Carnegie Mellon

From acceleration to stored energy

- ❖ Basis for Furuta stabilization of pendulum around the inverted equilibrium; actuator dynamics neglected, acceleration control Fc. No Fp. Requires lots of effort
- ❖ A better approach: Energy approach Astrom, Furuta
- ❖ For power systems this would mean: Avoiding blackout by changing the logic of Fc when close to voltage "collapse"

 Carnegie Mellon 37

Dynamics of interconnected system in new state space



- ❖ Standard state space of interconnected system

$$\dot{\bar{X}}_A = f_A(\bar{X}_A, Z_A, P_A, u_A)$$

$$\dot{Z}_A = f_{ZA}(\bar{X}_A, Z_A, P_B)$$

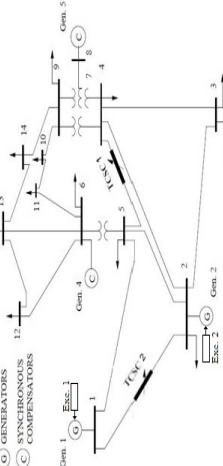
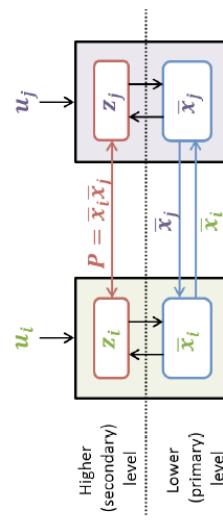
$$\dot{P}_A = f_{PA}(\bar{X}_A, P_A, \dot{P}_B)$$

$$\dot{Z}_B = f_{ZB}(Z_B, P_A, u_B)$$

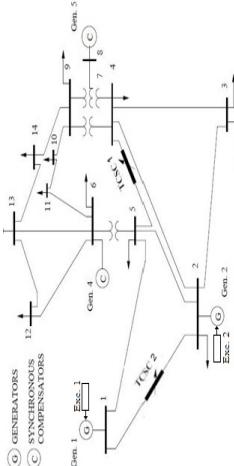
$$\dot{P}_B = f_{PB}(P_B, \dot{P}_A)$$

 Carnegie Mellon 38

Interaction variable-based two-layer model



Nonlinear cooperative control of FACTS in the new state space to synchronize generator 2

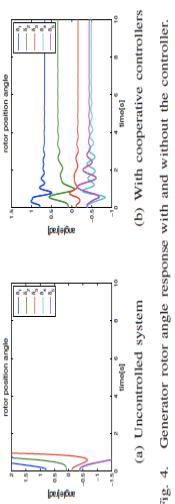


- 1. IEEE 14 bus system with two TCSC and five excitation controllers.

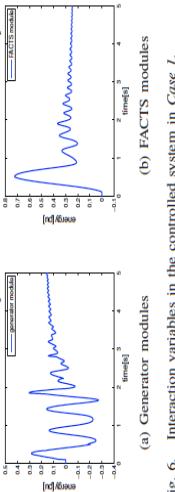
 Carnegie Mellon 39

 Carnegie Mellon 40

Fig. 1. An example of a system comprised of two modules i and j .



(a) Uncontrolled system
 (b) With cooperative controllers



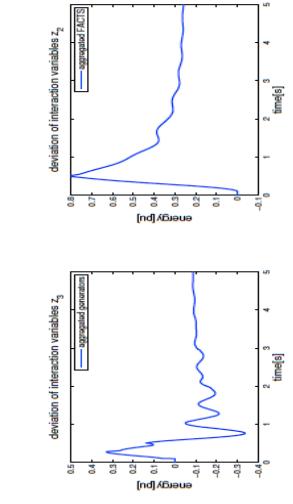
(a) Generator modules
 (b) FACTS modules

Carnegie Mellon

Fig. 5. Interaction variables without any controllers in Case 1.

Fig. 6. Interaction variables in the controlled system in Case 1.
 (a) Generator modules
 (b) FACTS modules

Fig. 7. Interaction variables of the controlled system in Case 2.
 (a) Generator modules
 (b) FACTS modules



(a) Generator modules
 (b) FACTS modules

Carnegie Mellon

Fig. 8. Aggregated interaction variables in Case 2.
 (a) Generator modules
 (b) FACTS modules

Interaction variables in new state space

- ❖ System dynamics **separable** into multi-layer system:
internal layer and interaction layer;
- ❖ **Rate of change of stored energy zero when module disconnected**
- ❖ Natural evolution of control area specifications (ACE) – linearized, steady-state notion; can cooperate for AGC
- ❖ Essential for iBA-level transient stabilization, power flow feasibility, SSCI, SSR standards (non-existent today)
- ❖ **Using this modelling framework**, different control strategy can be used and designed: competitive or cooperative control
- ❖ Note on synchronization—in the new state space entirely decentralized controller

Carnegie Mellon

Passivity-based AC/DC/AC Converter control



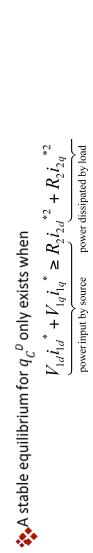
- ❖ Choose to directly regulate the direct and quadrature components of the load and source currents
- ❖ Desired capacitor charge has dynamics

$$\begin{aligned} \dot{q}_C^D &= \left(q_C^D - \left(q_C^D - C \cdot R_{L1} \cdot R_{L2} \cdot \left(i_{L1}^{*2} + i_{L2}^{*2} \right) \right) \right) \cdot C \cdot R_C \\ &= - \frac{\left(q_C^D - \left(q_C^D - C \cdot R_{L1} \cdot R_{L2} \cdot \left(i_{L1}^{*2} + i_{L2}^{*2} \right) \right) \right)}{CR \cdot q_C^D} \end{aligned}$$

Source K. D. Bachrachin, M. D. Ilic, "Automated Passivity-Based Control Law Derivation for Electrical Circuit-Lagrange Systems and Demonstration on Three-Phase AC/DC/AC Converter", EESG Working Paper No. R-WP-5-2014, August 2014.



Existence of stable equilibrium

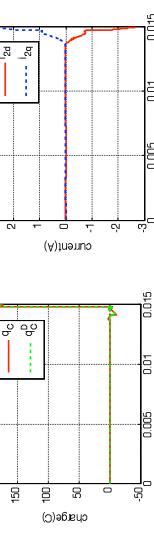


$V_{in} \cdot i_{Ld}^{*} + V_{in} \cdot i_{Lq}^{*} \geq R_{Ld} \cdot i_{Ld}^{*} + R_{Lq} \cdot i_{Lq}^{*}$

power supplied by load

power input by source

If this condition is violated, then the passivity-based controller will be unstable



Source K. D. Bachrachin, M. D. Ilic, "Automated Passivity-Based Control Law Derivation for Electrical Circuit-Lagrange Systems and Demonstration on Three-Phase AC/DC/AC Converter", EESG Working Paper No. R-WP-5-2014, August 2014.

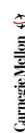


Interaction variables between complex modules

- ❖ Need to understand both module-level dynamics and inter-module dynamics
- ❖ Cooperative control between modules w/o detailed knowledge of internal dynamics is possible according to functional performance quantifiable in terms of interaction variables requirements for several time horizons.
- ❖ Competitive control of individual modules without much information exchange with the rest of the system also possible?
- ❖ Potential for sub-synchronous control instabilities (instabilities created by power electronics controllers on wind power plants and FACTS controllers in Texas power grid) managed by protocols avoiding problems in Table 1.

Conclusions

- ❖ Our proposal: Interaction variable-based
- ❖ Standards/protocols for **interactive** IBAs can set the basis for plug-and-play in smart grids—bounds on stored energy change and on rate of change of stored energy for T of interest
- ❖ Standards need to define transparent protocols for all dynamic components
 - Complexity of smart grids can be managed this way
 - At the same time system performance is guaranteed
- ❖ With current NERC standards system performance cannot be mapped into responsibilities of different components



References

- [1] Ilic, M., "Toward a Multi-Layered Architecture of Large-Scale Complex Systems for Reliable and Efficient Man-Made Infrastructures, Proc. of the MIT/ESD Symposium, Cambridge, MA, March 29-31, 2004.
- [2] Marija Ilic, IT-Enabled Electricity Services: The Missing Piece of the Environmental Puzzle, Issues in Technology and Policy, 2011, AP Seminar Series, MIT, Jan.,19,2011
- [3] Yang (Weng), Statistical and Inter-temporal Methods Using Embeddings for Nonlinear AC Power System State Estimation, PhD thesis, CMU ECE, August 2014
- [4] Elinor Ostrom, et al, A General Framework for Analyzing Sustainability of Socio-Ecological Systems, Science 325, 419 (2009).
- [5] Ilic, M., E. Allen, J. Chapman, C. King, J. Lang, and E. Litvinov, Preventing Future Blackouts by Means of Enhanced Electric Power Systems Control: From Complexity to Order, Proc. of the IEEE, vol. 93, no. 11, pp. 1920-1941, November 2005.
- [6] Ilic, Maria and Liu, Zhiqian, "A New Method for Selecting Best Locations of PMUs for Robust Automatic Voltage Control (AVC) and Automatic Flow Control (AFC), IEEE PES 2010, Minneapolis, MN, July 25-29, 2010.
- [7] Milos Oretkovic, "Power-Electronics-Enabled Transient Stabilization of Power Systems, PhD thesis, CMU, ECE Department, May 2014
- [8] Kevin D. Bachovchin, "Electromechanical Design and Applications in Power Grids of Flywheel Energy Storage Systems, PhD thesis, CMU, ECE Dept. May 2015
- [9] M. Prica and M. Ilic, "Maximizing Reliable Service by Coordinated Islanding and Load Management in Distribution Systems under Stress," Proc. of the 16th Power Systems Computation Conference, July 14-18, 2008, Glasgow, Scotland.
- [10] Siripha Junakarn, "Differentiated Reliability Options Using Distributed Generation and System Reconfiguration, PhD thesis, CMU, EPP Department, May 2015.
- [11] Ilic, M. D. Le, Xiu, Q. (co-Eds) Engineering IT-Enabled Sustainable Electricity Services: The Tale of Two Low-Cost Green Azores Islands, Springer August 2013
- [12] Ilic, et al, CDC 2014 tutorial.
- [13] S. Baros, M. Ilic, intelligent Balancing Authorities (IBAs) for Transient Stabilization of Large Power Systems, IEEE PES General Meeting 2014
- [14] Smart Grids and Future Electric Energy Systems, 18-618, CMU, ECE, Fall 2014.
- [15] Maria Ilic, Integration of Heterogeneous Small Test Beds for Emulating Large Scale smart Power Grids: The Emphasis on Cyber Architectures, MSCPEs2014 Workshop, invited keynote Berlin, CP week, April 14, 2014
- [16] Ilic, Maria; Xie, Le; Joo, Jhi-Young, "Efficient Coordination of Wind Power and Price-Responsive Demand Part I: Theoretical Foundations, IEEE Trans. on Power Systems.
- [17] Maria Ilic, Jhi-Young Jho, Le Xie, Maria Prica, Niklas Rotering, A Decision-Making Framework and Simulator for Sustainable Energy Services, IEEE Transactions on Sustainable Energy vol. 2, No. 1, January 2011, pp. 37-49.

Carnegie Mellon

ED

Carnegie Mellon

PLUG-AND-PLAY CONTROL AND OPTIMIZATION IN**MICROGRIDS****Florian Dörfler, ETH Zurich, Switzerland**

Microgrids are low-voltage electrical distribution networks, heterogeneously composed of distributed generation, storage, load, and managed autonomously from the larger transmission network. Modeled after the hierarchical control architecture of power transmission systems, a layering of primary, secondary, and tertiary control has become the standard operation paradigm for microgrids. Despite this superficial similarity, the control objectives in microgrids across these three layers are varied and ambitious, and they must be achieved while allowing for robust plug-and-play operation and maximal flexibility, without hierarchical decision making and time-scale separations. In this seminar, we explore control strategies for these three layers and illuminate some possibly-unexpected connections and dependencies among them. We build upon a first-principle model and different decentralized primary control strategies such as droop, quadratic droop, and virtual oscillator control. We motivate the need for additional secondary regulation and study centralized, decentralized, and distributed secondary control architectures. We find that averaging-based distributed controllers using communication among the generation units offer the best combination of flexibility and performance. We further leverage these results to study constrained AC economic dispatch in a tertiary control layer. Surprisingly, we show that the minimizers of the economic dispatch optimization problem are in one-to-one correspondence with the set of steady-states reachable by droop control. This equivalence results in simple guidelines to select the droop coefficients, which include the known criteria for power sharing. Finally, we illustrate the performance and robustness of our designs through hardware experiments.

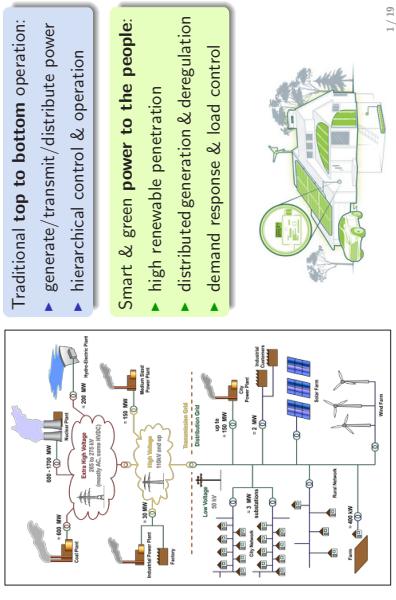


Plug-and-Play Control and Optimization in Microgrids

Florian Dörfler
ETH Zürich

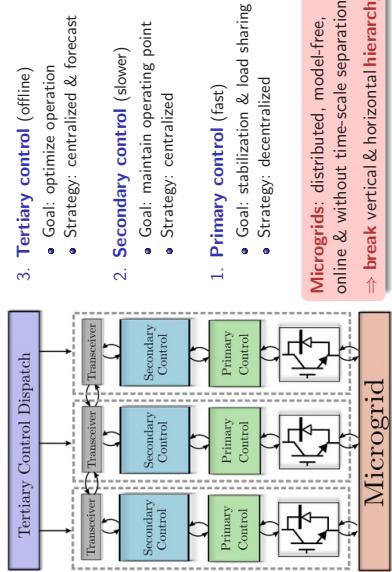
LCCC Dynamics and Control in Networks Workshop

Paradigm shifts in the operation of power networks

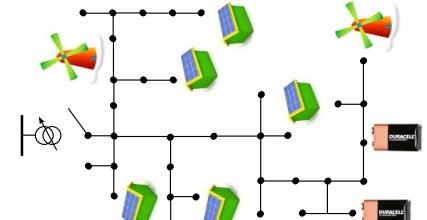


1 / 19

Conventional control architecture from bulk power nwks



3 / 19



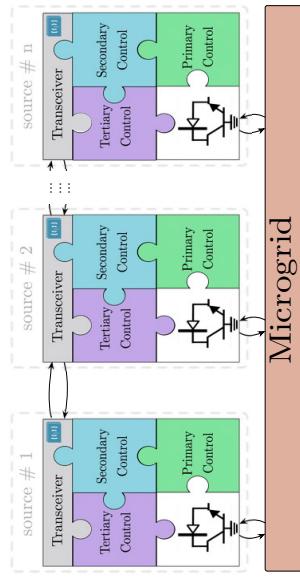
2 / 19

Microgrids

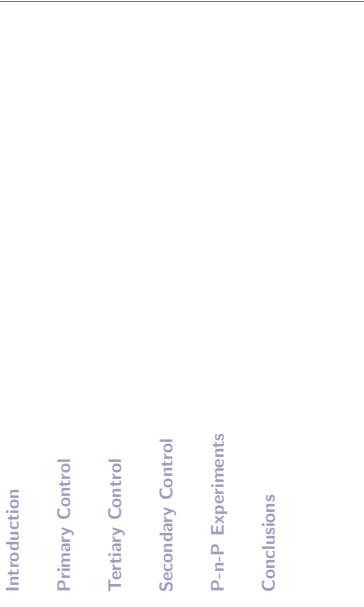
Structure	Applications	Benefits	Operational challenges
<ul style="list-style-type: none"> low-voltage distribution networks grid-connected or islanded autonomously managed 	<ul style="list-style-type: none"> hospitals, military, campuses, large vehicles, & isolated communities 	<ul style="list-style-type: none"> naturally distributed for renewables flexible, efficient, & reliable 	<ul style="list-style-type: none"> volatile dynamics & low inertia plug n' play & no central authority

Outline

A preview – plug-and-play control and optimization
flat hierarchy, distributed, no time-scale separations, & model-free ...



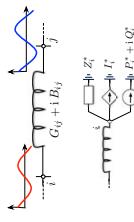
4 / 19



4 / 19

Modeling: a microgrid is a circuit

- ① synchronous (& acyclic) **AC circuit**
with harmonic waveforms $E_j e^{i(\theta_j + \omega_j t)}$



- ② **ZIP loads:** constant impedance, current, & power $P_i^* + iQ_i^*$ (today)



$$\text{injection} = \sum \text{power flows}$$

- ③ **coupling** via Kirchhoff & Ohm

- ④ decoupling: $P_i \approx P_i(\theta)$ & $Q_i \approx Q_i(E)$
(near operating point)

- ⑤ purely inductive lines $G/B \approx 0$
(can be relaxed to $G/B = \text{const.}$)

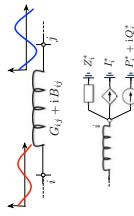
- ⑥ active power: $P_i = \sum_j B_{ij} E_i E_j \sin(\theta_i - \theta_j) + G_{ij} E_i E_j \cos(\theta_i - \theta_j)$
► trigonometric active power flow: $P_i(\theta) = \sum_j \beta_j \sin(\theta_i - \theta_j)$

- ⑦ reactive power: $Q_i = -\sum_j B_{ij} E_i E_j \cos(\theta_i - \theta_j) + G_{ij} E_i E_j \sin(\theta_i - \theta_j)$
► polynomial reactive power flow: $Q_i(E) = -\sum_j \beta_j E_i E_j E_i$ (not today)

5 / 19

Modeling: a microgrid is a circuit

- ① synchronous (& acyclic) **AC circuit**
with harmonic waveforms $E_j e^{i(\theta_j + \omega_j t)}$



- ② **ZIP loads:** constant impedance, current, & power $P_i^* + iQ_i^*$ (today)



$$\text{injection} = \sum \text{power flows}$$

- ③ **coupling** via Kirchhoff & Ohm

- ④ decoupling: $P_i \approx P_i(\theta)$ & $Q_i \approx Q_i(E)$
(near operating point)

- ⑤ purely inductive lines $G/B \approx 0$
(can be relaxed to $G/B = \text{const.}$)

5 / 19

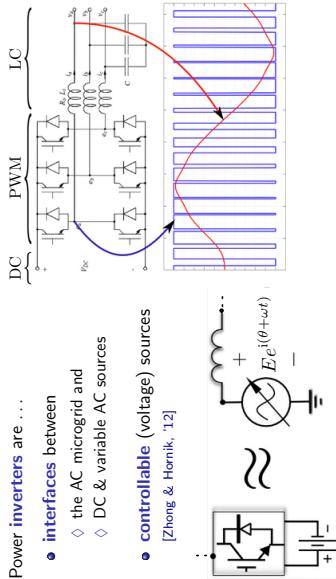
Modeling: sources interfaced with inverters

(all results also apply to synchronous machines & frequency-dependent loads)

Power inverters are ...

- **interfaces between**
 - ◊ the AC microgrid and
 - ◊ DC & variable AC sources
- **controllable (voltage) sources**

[Zhong & Hornik, 12]



primary control

Decentralized primary control of active power

Inverters are controlled to emulate the physics of synchronous generators. [Chandorkar et al. '93]

Intuition: Recall...

$$P_i(\theta) = \sum_{j=1}^n B_{ij} \sin(\theta_i - \theta_j)$$

P/θ droop control:
 $(\omega_j - \omega^*) \propto (P_i^* - P_i(\theta))$

$$D_i \dot{\theta}_i = P_i^* - P_i(\theta)$$

Putting the pieces together...
differential-algebraic closed loop

network physics

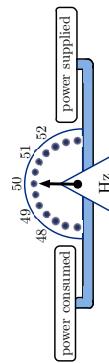
load power balance: $P_i^* = \sum_j B_{ij} \sin(\theta_i - \theta_j)$

source injections: $P_i(\theta) = \sum_j B_{ij} \sin(\theta_i - \theta_j)$

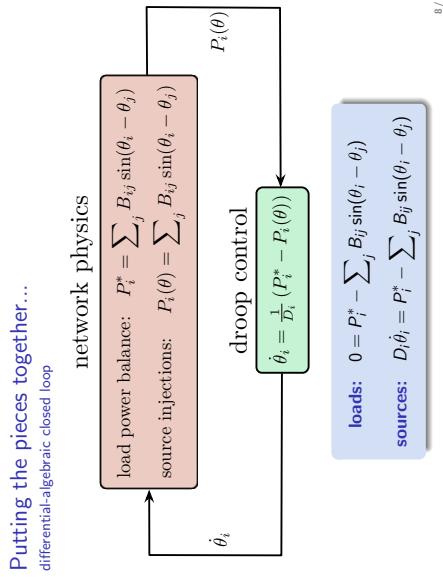
$\dot{\theta}_i = \frac{1}{D_i} (P_i^* - P_i(\theta))$

droop control

loads: $0 = P_i^* - \sum_j B_{ij} \sin(\theta_i - \theta_j)$
sources: $D_i \dot{\theta}_i = P_i^* - \sum_j B_{ij} \sin(\theta_i - \theta_j)$



6 / 19



7 / 19

Closed-loop stability under droop control

Theorem: stability of droop control [J. Simpson-Porco, FD, & F. Bullo, '12]
 \exists unique & exp. stable frequency sync \iff active power flow is feasible

Main proof ideas and some further results:

$$\omega_{\text{sync}} = \omega^* + \frac{\sum_{\text{inverters}} P_i^* + \sum_{\text{loads}} P_i^*}{\sum_{\text{inverters}} D_i}$$

(\propto power balance)

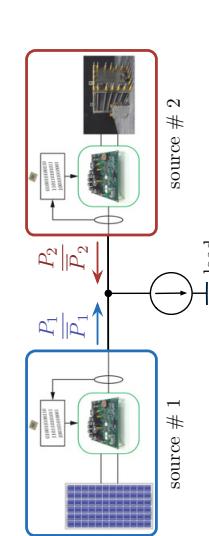
tertiary control

- synchronization frequency:
 $P_i = \begin{cases} P_i^* & \text{for loads} \\ P_i^* - D_i(\omega_{\text{sync}} - \omega^*) & \text{for inverters} \end{cases}$
 (depend on D_i & P_i^*)
- unique steady-state branch flows: $\xi_{ij} = B_{ij} \sin(\theta_i^* - \theta_j^*) \Rightarrow B_{ij} \geq \xi_{ij}$
 $(P_j \mapsto \xi_{ij})$

9 / 19

Objective I: decentralized proportional load sharing

- 1) Inverters have **injection constraints**: $P_i(\theta) \in [0, \bar{P}_i]$
- 2) Load must be **serviceable**: $0 \leq \left| \sum_{\text{loads}} P_j^* \right| \leq \sum_{\text{inverters}} \bar{P}_j$
- 3) **Fairness**: load should be shared proportionally: $P_i(\theta)/\bar{P}_i = P_j(\theta)/\bar{P}_j$



10 / 19

Objective I: decentralized proportional load sharing

- 1) Inverters have **injection constraints**: $P_i(\theta) \in [0, \bar{P}_i]$
- 2) Load must be **serviceable**: $0 \leq \left| \sum_{\text{loads}} P_j^* \right| \leq \sum_{\text{inverters}} \bar{P}_j$
- 3) **Fairness**: load should be shared proportionally: $P_i(\theta)/\bar{P}_i = P_j(\theta)/\bar{P}_j$

Theorem: fair proportional load sharing [J. Simpson-Porco, FD, & F. Bullo, '12]

Let the droop coefficients be selected **proportionally**:

$$D_i/\bar{P}_i = D_j/\bar{P}_j \quad \& \quad P_i^*/\bar{P}_i = P_j^*/\bar{P}_j$$

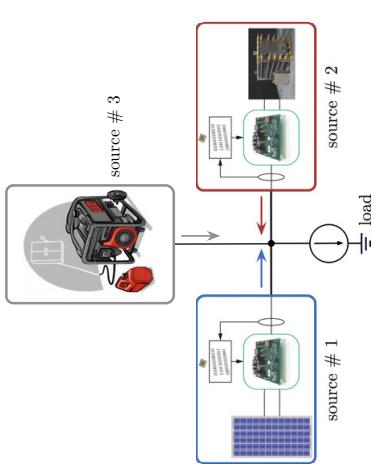
The the following statements hold:

- i) Proportional load sharing: $P_i(\theta)/\bar{P}_i = P_j(\theta)/\bar{P}_j$
- ii) Constraints met: $0 \leq \left| \sum_{\text{loads}} P_j^* \right| \leq \sum_{\text{inverters}} \bar{P}_j \iff P_i(\theta) \in [0, \bar{P}_i]$

10 / 19

Objective I: fair proportional load sharing

proportional load sharing is not always the right objective



Objective II: optimal economic dispatch
minimize the total accumulated generation

$$\begin{aligned} & \text{minimize } \theta \in \mathbb{T}^n, u \in \mathbb{R}^n \\ & \text{subject to} \\ & \quad \text{inverter power balance: } P_i^* + u_i = P_i(\theta) \\ & \quad \text{load power balance: } P_i^* = P_i(\theta) \\ & \quad \text{branch flow constraints: } |\theta_i - \theta_j| \leq \gamma_{ij} < \pi/2 \\ & \quad \text{inverter injection constraints: } P_i(\theta) \in [0, \bar{P}_i] \end{aligned}$$

Problem is generally non-convex and feasible only if the load is serviceable
In conventional power system operation, the economic dispatch is

- solved offline, in a centralized way, & with a model & load forecast
- solved online, in a decentralized way, & without knowing a model

In an autonomously managed microgrid, the economic dispatch should be

- solved offline, in a centralized way, & with a model & load forecast
- solved online, in a decentralized way, & without knowing a model

Objective II: optimal economic dispatch

minimize the total accumulated generation

$$\begin{aligned} & \text{minimize } u \in \mathbb{T}^n, u \in \mathbb{R}^n \\ & \text{subject to} \\ & \quad \text{inverter power balance: } P_i^* + u_i = P_i(\theta) \\ & \quad \text{load power balance: } P_i^* = P_i(\theta) \\ & \quad \text{branch flow constraints: } |\theta_i - \theta_j| \leq \gamma_{ij} < \pi/2 \\ & \quad \text{inverter injection constraints: } P_i(\theta) \in [0, \bar{P}_i] \end{aligned}$$

Problem is generally non-convex and feasible only if the load is serviceable

In conventional power system operation, the economic dispatch is

- solved offline, in a centralized way, & with a model & load forecast
- solved online, in a decentralized way, & with a model & load forecast

In an autonomously managed microgrid, the economic dispatch should be

- solved online, in a decentralized way, & without knowing a model

12 / 19

Objective II: decentralized dispatch optimization

Insight: droop-controlled microgrid = decentralized primal algorithm

Theorem: optimal droop

The following statements are equivalent:

- (i) the economic dispatch with cost coefficients α_i is **strictly feasible** with global minimizer (θ^*, u^*) .
- (ii) \exists droop coefficients D_i such that the microgrid possesses a unique & locally exp. stable sync'd solution θ satisfying $P_i(\theta) \in [0, \bar{P}_i]$.

If (i) & (ii) are true, then $\theta_i \sim \theta^*$, $u_i^* = -D_i(\omega_{\text{sync}} - \omega^*)$, & $D_i \alpha_i = D_j \alpha_j$.

- similar results hold for the general **constrained** case
- similar results in transmission nwks with DC flow [E. Mallada & S. Low, '13] & [N. Li, L. Chen, C. Zhao, & S. Low '13] & [X. Zhang & A. Papachristodoulou, '13] & [M. Andreasson, D. V. Dimarogonas, K. H. Johansson, & H. Sandberg, '13] & ...

13 / 19

Objective II: decentralized dispatch optimization

Insight: droop-controlled microgrid = decentralized primal algorithm

Theorem: optimal droop

The following statements are equivalent:

- (i) the economic dispatch with cost coefficients α_i is **strictly feasible** with global minimizer (θ^*, u^*) .
- (ii) \exists droop coefficients D_i such that the microgrid possesses a unique & locally exp. stable sync'd solution θ satisfying $P_i(\theta) \in [0, \bar{P}_i]$.

If (i) & (ii) are true, then $\theta_i \sim \theta^*$, $u_i^* = -D_i(\omega_{\text{sync}} - \omega^*)$, & $D_i \alpha_i = D_j \alpha_j$.

- similar results hold for the general **constrained** case
- similar results in transmission nwks with DC flow [E. Mallada & S. Low, '13] & [N. Li, L. Chen, C. Zhao, & S. Low '13] & [X. Zhang & A. Papachristodoulou, '13] & [M. Andreasson, D. V. Dimarogonas, K. H. Johansson, & H. Sandberg, '13] & ...

12 / 19

secondary control

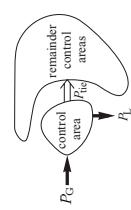
Secondary frequency control in power networks

Problem: steady-state frequency deviation ($\omega_{sync} \neq \omega^*$)

Solution: integral control [Chandorkar et. al. '93, Lopes et.al. '05, Bevrani '09, ...]

Interconnected Systems

- **Centralized** automatic generation control (AGC)



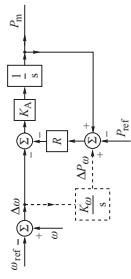
compatible with econ. dispatch
[N. Li, L. Chen, C. Zhao, & S. Low '13]

is *globally* stabilizing
[C. Zhao, E. Mallada, & F.D. '14]

14 / 19

Isolated Systems

- **Decentralized** PI control



is *globally* stabilizing
[C. Zhao, E. Mallada, & F.D. '14]

14 / 19

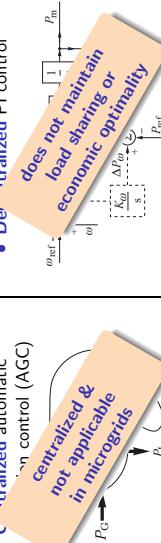
Secondary frequency control in power networks

Problem: steady-state frequency deviation ($\omega_{sync} \neq \omega^*$)

Solution: integral control [Chandorkar et. al. '93, Lopes et.al. '05, Bevrani '09, ...]

Interconnected Systems

- **Centralized** automatic generation control (AGC)



compatible with econ. dispatch

is *globally* stabilizing

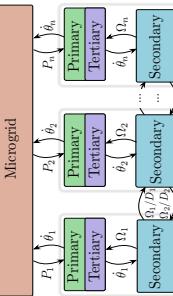
Microgrids require **distributed** (1) secondary control strategies.

14 / 19

Distributed Averaging PI (DAPI) control

$$D_i \dot{\theta}_i = P_i^* - P_i(\theta) - \Omega_i$$

$$k_i \dot{\Omega}_i = D_i \dot{\theta}_i - \sum_{j \subseteq \text{inverters}} a_{ij} \cdot \left(\frac{\Omega_i}{D_i} - \frac{\Omega_j}{D_j} \right)$$



Theorem: stability of DAPI

- [J. Simpson-Porco, F.D. & F. Bullo, '12]
primary droop controller works
 \iff
secondary DAPI controller works
- no tuning & no time-scale separation: $k_i, D_i > 0$
 - distributed & modular: connected comm. \subseteq inverters
 - recovers primary op. cond. (load sharing & opt. dispatch)
 \Rightarrow plug'n'play implementation

15 / 19

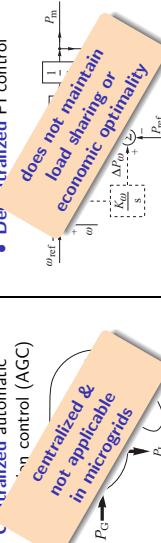
Secondary frequency control in power networks

Problem: steady-state frequency deviation ($\omega_{sync} \neq \omega^*$)

Solution: integral control [Chandorkar et. al. '93, Lopes et.al. '05, Bevrani '09, ...]

Interconnected Systems

- **Centralized** automatic generation control (AGC)



compatible with econ. dispatch

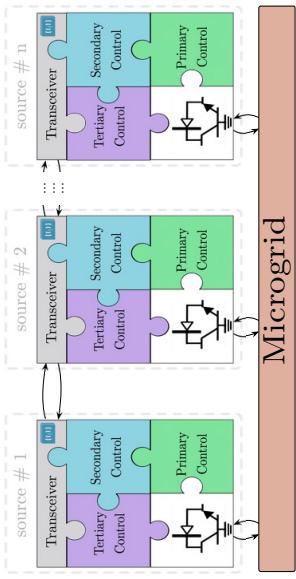
is *globally* stabilizing

Microgrids require **distributed** (1) secondary control strategies.

14 / 19

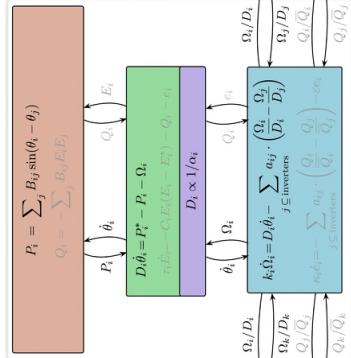
plug-and-play experiments

Plug'n'play architecture
flat hierarchy, distributed, no time-scale separations, & model-free



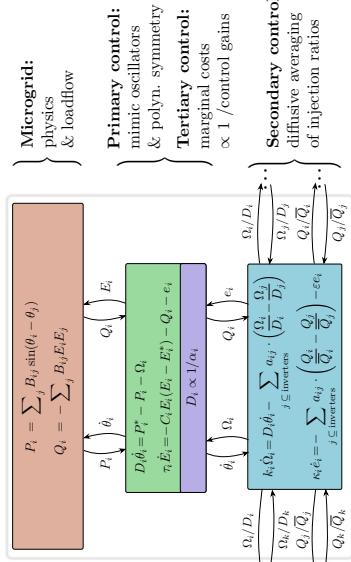
16 / 19

Plug'n'play architecture
recap of detailed signal flow (active power only)



16 / 19

Plug'n'play architecture
similar results in the reactive case



16 / 19

Microgrid:
physics & loadflow

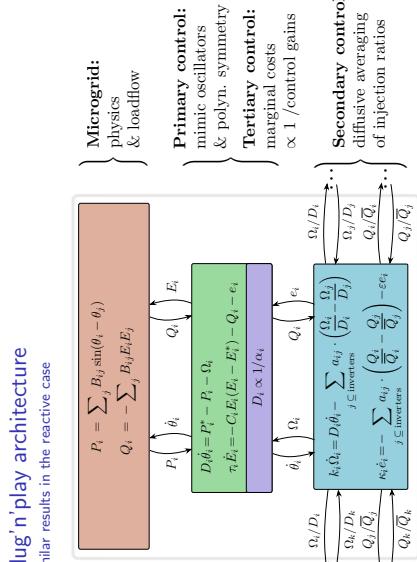
Primary control:
mimic oscillators & polyn. symmetry
Tertiary control:
marginal costs $\propto 1/\text{control gains}$



16 / 19

Microgrid:
physics & loadflow

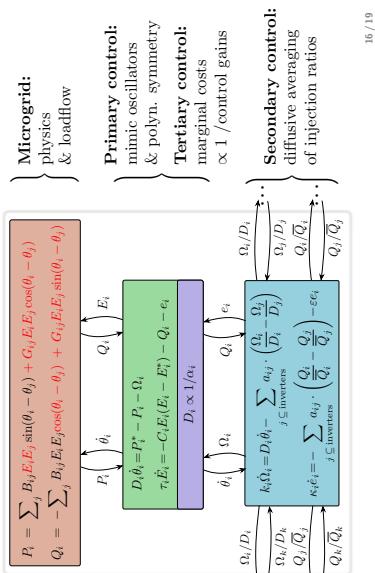
Primary control:
mimic oscillators & polyn. symmetry
Tertiary control:
marginal costs $\propto 1/\text{control gains}$



16 / 19

Plug'n'play architecture

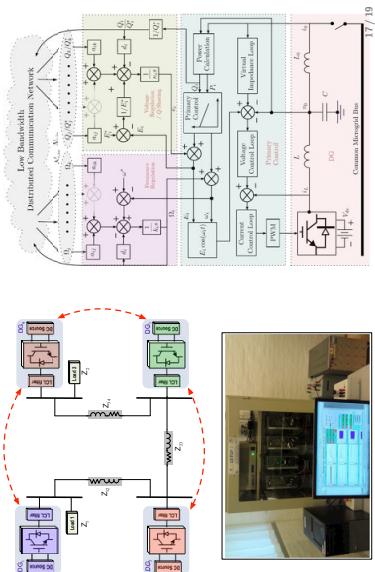
experiments also work well in the coupled & lossy case



16 / 19

Experimental validation of control & opt. algorithms

in collaboration with Q. Shafee & J.M. Guerrero @ Aalborg University

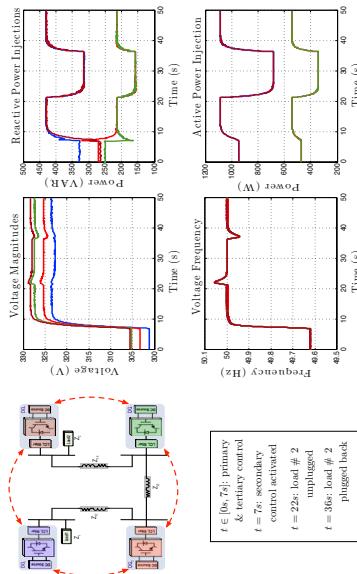


17 / 19

conclusions

Experimental validation of control & opt. algorithms

frequency/voltage regulation & active/reactive load sharing



18 / 19

Conclusions

Summary

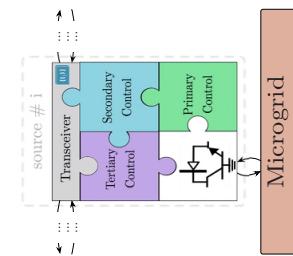
- primary P/θ droop control
- fair proportional load sharing & economic dispatch optimization
- distributed secondary control strategies based on averaging
- experimental validation

Further results

- reactive power control
- virtual oscillator control

Open conjecture

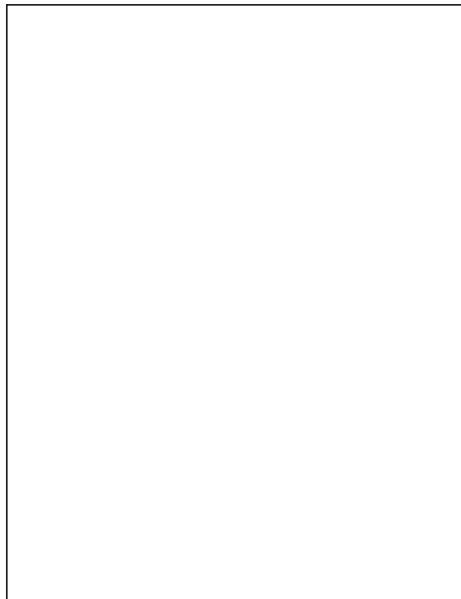
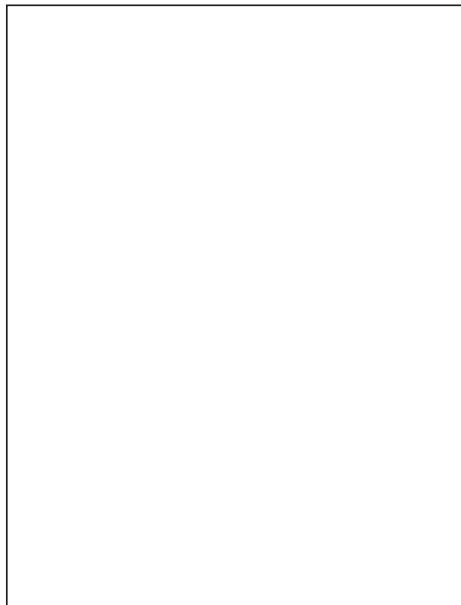
- solve these problems without comm



10 / 19

Acknowledgements

	A. Hamadéh
	B. Gentile
	H. Bouattour
	Q. Shafiee
	J. Simpson-Porco
	S. Dhopde
	B. Johnson
	J. Zhao
	S.Y. Çaliskan
	P. Tabuada
	M. Rungger
	F. Bullo
	M. Todiscoato



LYAPUNOV APPROACH TO CONSENSUS PROBLEMS

**Angelia Nedic, University of Illinois at Urbana Champaign,
USA**

This talk is focused on the weighted-averaging dynamic for unconstrained and constrained consensus problems. Through the use of a suitably defined adjoint dynamic, quadratic Lyapunov comparison functions are constructed to analyze the behavior of weighted-averaging dynamic. As a result, new convergence rate results are obtained that capture the graph structure in a novel way. In particular, the exponential convergence rate is established for unconstrained consensus with the exponent of the order of $O(1/(m\log 2m))$. Also, the exponential convergence rate is established for constrained consensus, which extends the existing results limited to the use of doubly stochastic weight matrices.



Motivation

LYAPUNOV APPROACH TO CONSENSUS PROBLEMS

Angelia Nedić

angelia@illinois.edu

Industrial and Enterprise Systems Engineering Department
and Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

joint work with Ji Liu (ECE, UIUC)

Say we are given an $m \times m$ row-stochastic matrix A , think of it as a transition matrix of a Markov chain over m states.

Suppose that the chain is ergodic, so that

$$\lim_{t \rightarrow \infty} A^t = \mathbf{1},$$

where $\mathbf{1}$ is the stationary distribution of the chain and every state has positive charge, i.e., $i > 0$ for all i .

Convergence rate for $\|\mathbf{1} - \mathbf{1}^\top A^t\|_1$ is known to be exponential

$$\|\mathbf{1}^\top A^t - \mathbf{1}\|_1 \geq c_1 \cdot e^{-c_2 t} \quad \text{for all } t > 0,$$

where c_1, c_2 are some universal constants that do not depend on m .

What is the best known convergence rate (in terms of the dependence on m)?

- For deterministic graphs? $f(m) = O(n^2)$
- For random graphs? $f(m) = O(m \log_2 m)$

Unconstrained Consensus Problem

- Consider a system consisting of m agents (nodes, sensors, robots, etc), represented by a set $[m] = \{1, \dots, m\}$
- We assume that a sequence $\{G_t\}$ of directed graphs is given externally, where each graph G_t represents communication structure among the agents, where $G_t = ([m], E_t)$

We consider the unconstrained (scalar) consensus problem, formalized as follows.
[Unconstrained Consensus] *Design a distributed algorithm achieving the communication structure given by graph G_t at each time t and ensuring that, for every set of initial values $x_i(0) \in \mathbb{R}$, $i \in [2, m]$, the following limiting behavior emerges:*

$$\lim_{t \rightarrow \infty} x_i(t) = c \quad \text{for all } i \in [2, m] \text{ and some } c \in \mathbb{R}.$$

Back to Weighted Averaging Models for Consensus

- The literature is vast: Tsitsiklis 1984, Jadbabaie Morse and Lin 2003 and the stream of publications that followed

- Unconstrained consensus problem - can be used as a mechanism for information diffusion in networks to deal with optimization, tracking, estimation, learning in networks

- Understanding of the convergence rate is critical for making use of the consensus as a building block for other algorithms over the network AN and Ozdaglar 2007, 2008, 2009, 2010

- Convergence rate results did not explicitly capture the network structure

- Alternative approach by using a different Lyapunov function to measure the progress of the consensus-dynamics allows us to capture the rate in terms of network structure

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Literature: VAST

Roughly speaking there are three approaches

- Push-sum or Ratio Consensus Algorithm (discussed yesterday)

- Laplacian-Based Algorithm:

$$x(t+1) = I - \frac{1}{L(t)} x(t)$$

where $x(t) = [x_1(t), \dots, x_m(t)]^T$ and $0 \in \mathbf{m}$ (see Jadbabaie et al. 2003)

- Weighted-Averaging Algorithm

$$x(t+1) = A(t)x(t)$$

where $A(t)$ is a row-stochastic matrix with sparsity pattern matching the graph G_t structure (Tsitsiklis 1984)

Most of the Laplacian-based algorithms require that each $L(t)$ is also symmetric which implicitly require bidirectional communication links.

Weighted-averaging algorithm gets around this limitation.

4

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Weighted-Averaging Algorithm

$$x(t+1) = A(t)x(t),$$

where the weight matrices $A(t)$ are row-stochastic and compliant with the graph G_t structure (to be discussed soon)

The existing analysis of the weighted-averaging is based on studying the behavior of the left-matrix products

$$x(t) = A(t)A(t-1) \cdots A(s+1)A(0)x(0) \quad \text{for } t \geq 0,$$

In particular, when the matrices $A(t)A(t-1) \cdots A(1)A(0)$ converge to a rank one matrix, the iterates $x(t)$ converge to a consensus

Some conditions on the graphs G_t and the matrices $A(t)$ that convergence are: (Tsitsiklis 1984; Nedić and Ozdaglar 2009; Nedić, Olshevsky, Ozdaglar, Tsitsiklis 2009)

Assumption Let $\{G_t\}$ be a graph sequence and $\{A(t)\}$ be a sequence of $m \times m$ matrices that satisfy the following conditions:

(a) **Graph Compliance** Each $A(t)$ is a stochastic matrix that is compliant with the graph G_t , i.e., $A_{ij}(t) > 0$ when $(j, i) \in E_t$ for all t .

(b) **Aperiodicity** The diagonal entries of each $A(t)$ are positive, $A_{ii}(t) > 0$ for all t and $t \neq 1m$.

(c) **Uniform Positivity** There is a scalar $\lambda > 0$ such that $A_{ij}(t) \geq \lambda$ whenever $A_{ij}(t) > 0$.

(d) **Irreducibility** Each G_t is strongly connected.

5

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Convergence Rate

Under the preceding Assumption we have

$$\lim_{d \downarrow 1} A(t) \cdots A(k+1)A(k) = \mathbf{1} \quad \text{for all } t \neq 0,$$

where each $\mathbf{1}$ is stochastic vector and $\mathbf{1} = [1, \dots, 1]^T$.

Furthermore, the convergence rate is geometric: for all $t \neq k \neq 0$,

$$\mathbf{1}A(t) \cdots A(k+1)A(k) - \mathbf{1} \quad [k]: \mathbf{1} \cdot Cq^{t-k},$$

where the constants $C > 0$ and $q \in (0, 1)$ depend only on m and n .

When the matrices $A(t)$ are doubly stochastic, we have for all $t \neq k \neq 0$,

$$A(t) \cdots A(k+1)A(k) - \frac{\mathbf{1}}{m} \mathbf{1}^T \quad 1 - \frac{1}{2m^2}.$$

Refs. Tsitsiklis 1984, AN and Ozdaglar 2009; AN, Olshevsky, Ozdaglar, and Tsitsiklis 2009

These and other existing rate results are not explicitly capturing the structure of the graph G_t , such as the longest shortest path or the maximum node degrees for example.

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lyapunov Approach

October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Convergence Rate: Lyapunov Approach

October 15–17, 2014

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

New rate results are possible by adopting dynamic system point of view and applying Lyapunov approach

This approach allows us to characterize the convergence of the weighted-averaging algorithm with a more explicit dependence on the graph structure than the existing results. In particular, we work with a quadratic Lyapunov comparison function proposed by Touri 2011, and Touri and AN 2014

In this approach, an absolute probability sequence of matrices $A(t)$ play a critical role in the construction of a Lyapunov comparison function and in establishing its rate of decrease along the iterates of the algorithm.

Definition For a chain of row-stochastic matrices $\{A(t)\}$ we say that the vector sequence $\{\mathbf{1}(t)\}$ is an absolute probability sequence of the chain if each $\mathbf{1}(t)$ is a probability vector and

$$\mathbf{1}(t) = \mathbf{1}(t+1)A(t) \quad \text{for all } t \geq 0.$$

NOTE The absolute probability sequence is an adjoint dynamic for the weighted-averaging dynamic

6

7

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Assumptions

An undirected tree \mathcal{T} is weakly spanning tree of a directed graph $G = ([m], E)$, if for every $\{i, j\} \in \mathcal{T}$ we have either $(i, j) \in E$ or $(j, i) \in E$.

Assumption 1

Let $\{G_t\}$ be a graph sequence and $\{A(t)\}$ be a matrix sequence such that:

(a) (*Partial Irreducibility*)

Each graph G_t contains a weakly spanning tree and each $A(t)$ is a stochastic matrix that is compliant with a weakly spanning tree T_t of G_t , i.e., $A_{ij}(t) > 0$ whenever $(j, i) \in T_t$ for all $t \geq 0$.

(b) (*Asymptoticity*)

The diagonal entries of each $A(t)$ are positive, $A_{ii}(t) > 0$ for all t , and $i \in [m]$. There is a scalar $\rho > 0$ such that $A_{ii}(t) \leq \rho$ and $A_{ij}(t) \leq \rho$ for all $(i, j) \in T_t$ and for all $t \geq 0$.

(c) (*Partial Uniform Positivity*)

The matrix sequence $\{A(t)\}_{t \geq 0}$ has an absolute probability sequence $\{\pi(t)\}$ that is uniformly bounded away from zero, i.e., there is $\bar{\pi} \in (0, 1)$ such that $\pi_i(t) \geq \bar{\pi}$ for all i and t .

(d) (*Adjoint Dynamics / Absolute Probability Sequence*)

The matrix sequence $\{A(t)\}_{t \geq 0}$ has an absolute probability sequence $\{\pi(t)\}$ that is uniformly bounded away from zero, i.e., there is $\bar{\pi} \in (0, 1)$ such that $\pi_i(t) \geq \bar{\pi}$ for all i and t .

8

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Proof Sketch

Proof: Based on the use of Lyapunov Function

$$V(x, \dot{x}) = \sum_{i=1}^m (x_i - \bar{x}_i)^2$$

and the relation

$$V(x(t+1), \dot{x}(t+1)) = V(x(t), \dot{x}(t)) - D(t),$$

where

$$D(t) = \frac{1}{2} \sum_{i=1}^m (t+1) \sum_{j=1}^m A_{ij}(t) A_i(t) (\dot{x}_j(t) - x_j(t))^2.$$

Abound for the decrement is obtained (using Assumption 1)

$$D(t) \leq \frac{p^0(t)}{2} \max_{j \in [m]} (x_j(t) - \bar{x}(t))^2 \quad \text{for } t \geq 0,$$

$p^0(t)$ is the maximum number of links in any of the shortest paths in the tree T_t of Assumption 1(a). We also use

$$\dot{q}_i(t)x(t) = q_0 x(t)$$

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Basic Result

Theorem Under Assumption 1, for the iterates $\{x(t)\}$ generated by the weighted-averaging algorithm with any initial vector $x(0) \in \mathbb{R}^n$, we have for any $t \geq k \geq 0$,

$$\sum_{i=1}^m q_i(t) x_i(t) - (0) \hat{x}(t) \leq 2 - \frac{2}{p^2} \sum_{j=k}^{t-k} q_j(k) - (0) \hat{x}(0),$$

where

- (a) $\hat{x}(t)$ is the uniform lower bound on the entries in $A(t)$ (Assumptions 1(c))
- (b) > 0 is the uniform lower bound on the absolute probability vectors for all $t \geq 0$.
- (c) > 0 is the uniform lower bound on the entries of the absolute probability vectors for all $t \geq 0$.
- (d) $\hat{x}(t)$ (Assumption 1(d)),
- while
- $\hat{p}^2(s) = \max_{i \in [m]} p^2(s)$
- $\hat{p}^2(s)$ is the longest shortest path in the tree T_s of Assumption 1(a).

9

Lund University LCCC Workshop on Dynamics and Control in Networks October 15–17, 2014

Implications of Theorem 1: Doubly Stochastic

Let Assumption 1 hold, and assume also that the weight matrices $A(t)$, $t \geq 0$, are doubly stochastic. Then, we have $\langle 1 \rangle = \frac{1}{m} \mathbf{1}$ and the relation of Theorem 1 reduces to:

$$\mathbf{k}\pi(t) - \bar{x}(t) \mathbf{k}^2 \leq \mathbf{k}\pi(t) - \bar{x}(0) \mathbf{k}^2, \quad (1)$$

with $\bar{x}(0) = \frac{\mathbf{k}\pi(0)}{m}$. Since the maximum path length in a tree is no larger than $m-1$, i.e., $p^k(s) \leq m-1$, it follows that

$$\mathbf{k}\pi(t) - \bar{x}(t) \mathbf{k}^2 \leq \mathbf{k}\pi(t) - \bar{x}(0) \mathbf{k}^2 \leq \frac{2}{m(m-1)} \mathbf{k}\pi(t) - \bar{x}(0) \mathbf{k}^2.$$

Thus, when π does not depend on m , the convergence rate has dependency of $O(m^2)$ in terms of the number m of agents, which is the same as the rate result in our earlier work (see slide 4).

Think of $\pi = \frac{1}{\max_i |\mathcal{N}_i(t)|}$. Given m , what graphs on m nodes will be the best?

10

Implications of Theorem 1: New Bound

Suppose now that we want to construct the graphs G_t such that Assumption 1 holds and we want to get the most favorable rate dependency on m . In this case, the following result is valid.

Theorem 1 *There is a sequence $\{G_t\}$ of regular undirected graphs such that for all $x(0) \in \mathbb{R}^m$ and all $t \geq 0$*

$$2x(t) - \tilde{x}(0)12^2 2^{t-k} 2x_t(k) - \tilde{x}(0)12^2,$$

with $\tilde{x}(0) = \frac{x_2(0)}{m}$ and

$$q = 1 - \frac{1}{4m^2 \log_2 m^2}$$

The rate of this order is provided in Diaconis and Stroock 1991 (Proposition 3 and Example 2.3) for a static ergodic Markov Chain.



satisfied, and the estimate in (1) reduces to

$$kx(t) - \tilde{x}(0)1k^2 k - 1 - \frac{1}{4\pi m^2} \frac{kx(k) - \tilde{x}(0)1k^2}{\log_2 m^2}.$$

The result follows by noting that $d = \log_2 m$.

Theorem 1 shows that the exponential convergence rate with the ratio of the order $1 - O(\frac{1}{m \log_2 m})$ is achievable for consensus on some tree-like regular undirected graphs. This improves the best known bound with the ratio of the order $1 - O(\frac{1}{m^2})$ for undirected graphs and doubly stochastic matrices (see slide 4).

Proof

We will construct an undirected graph sequence $\{G_t\}$ that satisfies Assumption 1.

Let $m = 2^d$ for some integer $d \geq 1$. Let t be arbitrary but fixed time. Select $2^d - 1$ agents and construct an undirected binary tree with these agents as nodes. Next, add one extra agent as a root with a single child (see left plot in the Figure). Thus, each agent i except for the root and the leaf agents has the degree equal to 3.

Consider, now connecting all leaf-nodes with undefined edges (see middle plot in Figure). Now, all leaf-agents have degree equal to 3 except for the far most left and far most right agents, each of which has the degree equal to 2. Connect these two agents to the root node (see right plot in Figure). In this way, the far most left and far most right leaf agents, as well as the root agent have degree 3. In the resulting regular undirected graph, we let $A_{ij}(t) = \frac{1}{4}$ for all $j \in N_i(t) \setminus \{i\}$ and for all i , so that $\|A(t)\| = \frac{1}{4}$. The shortest path from the root agent to any other agent in the graph is at most $2^d/2$ (going down from the root of the tree to the nodes at the depth $2^d/2$, and going through the leaf nodes to reach those that are the depth larger than d).

Using the same construction, for all times t , we have that $\{A(t)\}$ is a sequence of doubly stochastic matrices, and therefore $\|A(t)\| = \frac{1}{m}$ for all t . Thus, Assumption 1 is

12

Implication for Matrix Sequence

We next consider an implication of Theorem 1 for the convergence of matrix products

$$A(t:k) = A(t) \cdots A(k+1)A(k) \quad \text{for all } t < k < 0,$$

where $A(t:k) = A(k)$ whenever $t = k$.

Theorem 2 *If Assumption 1 holds, then for all $t < k < 0$,*

$$A(t:k) - I = (k)^{p/2} 0 \quad 1 - \frac{1}{4\pi m^2} \frac{I - 1}{\rho^2} (k)^{p/2},$$

where I is the identity matrix.

Corollary 3 *Under Assumption 1, the sequence $\{A(t)\}$ is ergodic:*

$$\lim_{t \rightarrow 0} A(t) \cdots A(k) = I \quad \text{for all } k < 0.$$

14

Lund University	LCCC Workshop on Dynamics and Control in Networks	October 15–17, 2014
	Constrained Consensus	

Assumption 2 We are given a collection of sets $X_i \setminus \mathbb{R}^n$ which closed and convex, and their intersection is nonempty, i.e., $X = \bigcap_{i=1}^m X_i \neq \emptyset$.
The constrained consensus problem is as follows.

[Constrained Consensus] Assuming that each agent knows only its set X_i , design a distributed algorithm obeying the communication structure given by graph G_t at each time t and ensuring that, for every set of initial values $\mathbf{x}_i(0) \setminus \mathbb{R}^n$, $i \setminus [m]$, the following limiting behavior emerges:

$$\lim_{t \rightarrow \infty} \mathbf{x}_i(t) = c \quad \text{for all } i \setminus [m] \text{ and some } c \setminus X.$$

We have considered this problem in AN, Oodaghifar, and Parrilo 2010, proposing the following algorithm

$$\begin{aligned} \mathbf{w}_i(t+1) &= \sum_{j=1}^m A_{ij}(t) \mathbf{x}_j(t), \\ \mathbf{x}_i(t+1) &= P_{X_i}[\mathbf{w}_i(t+1)], \end{aligned} \quad (2)$$

with a restriction to the use of doubly stochastic matrices

Lund University	LCCC Workshop on Dynamics and Control in Networks	October 15–17, 2014
	Convergence Rate	

We have results for the sets X_i that satisfy a certain regularity condition which relates the distances from a given point to the sets X_i with the distance from the point to the intersection set $X = \bigcap_{i=1}^m X_i$.

In particular, since $X \setminus X_i$ for all i , it follows that

$$\text{dist}(x, X) \leq \text{dist}(x, X_i) \quad \text{for all } x \setminus \mathbb{R}^n \text{ and } i \setminus [m]. \quad (3)$$

In our analysis, we need an upper bound on $\text{dist}(x, X)$ in terms of the distances $\text{dist}(x, X_i)$, $i \setminus [m]$.

We will use the following definition of set regularity.

Let $Z \setminus \mathbb{R}^n$ be a nonempty set. We say that a (finite) collection of closed convex sets $\{Y_i, i \setminus \mathcal{T}\}$ is regular (in Euclidean norm) with respect to the set Z , if there is a constant $r \setminus 1$ such that

$$\text{dist}(y, Y) \leq r \max_{i \in \mathcal{T}} \{\text{dist}(y, Y_i)\} \quad \text{for all } y \setminus Z.$$

We refer to the scalar r as a regularity constant. When the preceding relation holds with $Z = \mathbb{R}^n$, we say that the sets $\{Y_i, i \setminus \mathcal{T}\}$ are uniformly regular. In view of relation (3) it follows that the regularity constant r must satisfy $r \setminus 1$.

Lund University	LCCC Workshop on Dynamics and Control in Networks	October 15–17, 2014
	Convergence	

The following result proves that the iterates of the algorithm converge to a common point in the set X .

Theorem 4 Let Assumption 1 and Assumption 2 hold. Then, the sequences $\{\mathbf{x}_i(t)\}_{i \setminus [m]}$ of the algorithm (2) are bounded, i.e., there is a scalar $R > 0$ such that

$$\Delta x_i(t) \leq 2 \quad \text{for all } i \setminus [m] \text{ and all } t \geq 0,$$

and they converge to a common point $x^* \setminus X$:

$$\lim_{t \rightarrow \infty} \mathbf{x}_i(t) = x^* \quad \text{for some } x^* \setminus X \text{ and for all } i \setminus [m].$$

Some notes on the determining

Boundedness of iterates: for all $y \setminus X$ and all $t \geq 0$,

$$\begin{aligned} &\sum_{i=1}^m (t) 2 \mathbf{x}_i(t) - y^2 \geq \sum_{j=1}^m f_j(t) 2 \mathbf{x}_j(t) - y^2 \\ &\text{yields (under Assumption 1) for any } y \setminus X, \text{ all } t \text{ and } i, \\ &2 \mathbf{x}_i(t) 2 \geq 2 y^2 + \frac{1}{p} \max_j 2 \mathbf{x}_j(t) - y^2 \end{aligned}$$

17

Lund University	LCCC Workshop on Dynamics and Control in Networks	October 15–17, 2014
	Convergence	

Theorem 5 Let Assumption 1 and Assumption 2 hold. Assume further that the sets $\{X_i, i \setminus [m]\}$ are regular, with a regularity constant $r \setminus 1$, with respect to a ball $B(O, r)$ which contains all the iterates $\{\mathbf{x}_i(t)\}$ generated by the algorithm (2). Consider the following Lyapunov comparison function:

$$V(t, y) = \sum_{i=1}^m (t) \mathbf{x}_i(t) - y^2. \quad (4)$$

Then, the Lyapunov comparison function ' $\mathbf{P}_r(\mathbf{v}(t))$ ' decreases at a geometric rate:

$$\dot{V}(t+1, \mathbf{v}(t+1)) \leq \frac{1 - p^2(r+1)^2}{p^2(r+1)^2} V(t, \mathbf{v}(t)) \quad \text{for all } t \geq 0,$$

where

$$\mathbf{u}(t) = \sum_{i=1}^m (t) \mathbf{x}_i(t), \quad \mathbf{v}(t) = \mathbf{P}_r[\mathbf{u}(t)], \quad \text{for all } t \geq 0. \quad (5)$$

and the scalars ' $r \setminus 2$ ', ' $(0, 1)$ ' and the integer ' $p^2 \setminus 2$ ' are the same as in Theorem 1.

Theorem 6 Under the assumptions of Theorem 5, for all $t \geq 0$,

$$\text{dist}^2(\mathbf{x}_i(t), X) \leq \frac{1}{1 - \frac{4p^2}{r^2}(r+1)^2} \cdot (0, \mathbf{v}(0)),$$

where $\mathbf{v}(0) = \mathbf{P}_r[\mathbf{u}(0)]$ with $\mathbf{u}(0) = \sum_{j=1}^m (0) \mathbf{x}_j(0)$.

MAJORITY CONSENSUS BY LOCAL POLLING**Moez Draief, Imperial College, UK**

We consider a network where each of n anonymous nodes holds some initial binary opinion. Our goal is to design efficient and fully-distributed algorithms that enable the network to reach a state where all the nodes hold the initial majority opinion. This problem has received a lot of interest from a number of communities such as distributed computing, communication networks, biological and chemical networks and social networks.

In this talk, I will present two distributed procedures for solving the majority consensus. In the first one we allow nodes to hold undecided states, provide a simple dynamics for solving the exact majority consensus problem on general connected graphs and derive bounds on the time to convergence of the algorithm. The second algorithm lets nodes interact with more than one neighbour at a time and show that, for some family of graphs of interest, and provided there is sufficient bias in the population towards the major it opinion, this algorithm solves majority consensus with high probability. We compute its asymptotic convergence time.

This is a joint work with M. Vojnovic and M. Abdullah.



Aggregating through Majority

Majority Consensus by Local Polling

Moez Draief

Imperial College London

Lund Workshop on Dynamics and Control in Networks
October 2014

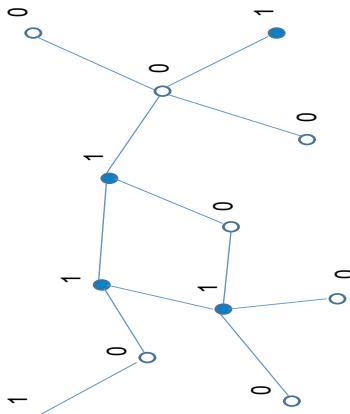
1785, Marquis de Condorcet's weak law of large numbers

- in a large population of voters, and each one independently votes correctly with probability $\alpha > 1/2$
- as population size grows, probability that the outcome of a majority vote is correct converges to one

Information is efficiently aggregated

Aggregating in a network

Moez Draief Majority Consensus by Local Polling



Binary majority consensus

Moez Draief Majority Consensus by Local Polling

Desired outcome and metrics

- Nodes end with opinion held by majority of nodes
- Node can probe neighbours and update opinion accordingly using little (constant) memory
- Probability of error (convergence to incorrect consensus)
- Time to convergence

Applications

- Occurrence of a given event in cooperative decision making
- Voting in distributed systems
- Routine to solve more elaborate distributed decision making instances

Moez Draief Majority Consensus by Local Polling

Moez Draief Majority Consensus by Local Polling

Local Majority Dynamics

Outline

$G = (V, E)$ simple connected graph on $|V| = n$ vertices

Each vertex either **red** (1) or **blue** (0).

Initial proportion of blues is $\alpha \in (1/2, 1)$

GOAL: Local algorithm for inferring the majority state.

- Does the graph settle into one colour?

- If so, how does the graph structure and the initial distribution affect which colour wins?

- How long does it take?

Continuous-time Interaction Model

Mozz Draief	Majority Consensus by Local Polling	Mezz Draief	Majority Consensus by Local Polling

- Connected undirected graph $G = (V, E)$, $|V| = n$

- αn nodes hold 0 and $(1 - \alpha)n$ nodes hold 1, $\alpha \in (1/2, 1)$

- Nodes i and j interact at rate $q_{ij} = q_{ji}$, $q_{ij} \neq 0$ iff $(i, j) \in E$

Markov chain

- $(X_t)_{t \geq 0}$ continuous-time Markov chain with rate matrix Q ,
- $$q_{ii} = -\sum_{j \neq i} q_{ij}$$
- $(\pi_t)_{t \in V}$ stationary distribution is uniform on V . Mixing time:

$$\mathbb{P}_j(X_t = i) - 1/n = O\left(e^{-\lambda_2(Q)t}\right)$$

where $\lambda_2(Q) = \inf\{\sum_{i,j} q_{ij}(x_i - x_j)^2/2, \|x\| = 1, x^T \mathbf{1} = 0\}$

Theorem [Hassini-Peleg '01]

- The number of nodes in state 1 is a martingale.
- Probability of reaching (wrong) consensus at 1 is $1 - \alpha$.
- Time to convergence of voter model $O(n(\lambda_2(Q)))$.

Mozz Draief	Majority Consensus by Local Polling	Mezz Draief	Majority Consensus by Local Polling

Mozz Draief	Majority Consensus by Local Polling	Mezz Draief	Majority Consensus by Local Polling

Distributed averaging: Proof

Small memory

Let $R(t) = \|x(t)\|^2$. When an i, j interaction takes place $R(t)$ reduces by $(x_i - x_j)^2/2$.

$$\begin{aligned}\mathbb{E}(dR(t) \mid x(t) = x) &= \sum_{i,j} q_{ij} \left(2 \left(\frac{x_i + x_j}{2} \right)^2 - (x_i^2 + x_j^2) \right) \\ &= - \sum_{i,j} q_{ij} \frac{(x_i - x_j)^2}{2}\end{aligned}$$

(Assume that $\sum_i x_i(0) = 0$)

In particular

$$\mathbb{E}\|x(t)\|^2 \leq \|x(0)\|^2 e^{-\lambda_2(Q)t}$$

Could we use less memory and still guarantee small error?

Theorem: Impossibility

- Connected undirected graph $G = (V, E)$, $|V| = n$,
- αn nodes in 0 and $(1 - \alpha)n$ nodes in 1, $\alpha \in (1/2, 1)$,
- $2\alpha - 1$ is the voting margin.

No 1-bit distributed algorithm can solve the majority consensus problem.

Land, Belew, "No perfect two-state cellular automata for density classification exists", PRL 74, 5148-5150, 1995

Mezö Dániel

Majority Consensus by Local Polling

Ternary Consensus

- αn nodes hold 0 and $(1 - \alpha)n$ nodes hold 1,
- Additional state e for undecided nodes, $q_{ij} = 1/n, \forall i, j$

Theorem [PVV '09]

Probability of reaching wrong consensus 1. For n large,

$$P_{\text{error}} = (1 + o(1))^{2 - D(\alpha || \frac{1}{2})n}$$

where $D(\alpha || \frac{1}{2})$ is the Kullback-Leibler divergence. T time to convergence, $\mathbb{E}(T) = (1 + o(1)) \log n$.

- Results (seem to) hold for expander but fail for the line.
- Generalises beyond binary consensus [Babaee, Drajef '14]

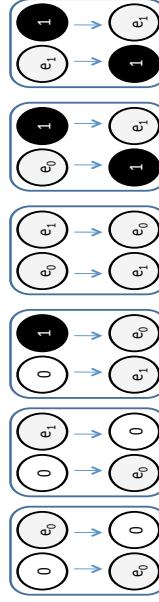
Kashyap, Basar, Srikanth, "Quantized consensus" Automatica, 1192-1203, 2007

Béreznitz, Thiran, Vetterli, "Interval consensus: From quantized gossip to voting", ICASSP 2009

Mezö Dániel

Majority Consensus by Local Polling

Averaging-like updates: States $0 < \theta_0 < e_1 < 1$.
Rules: Swaps + Annihilation



Binary Consensus with two undecided states

Kashyap, Basar, Srikanth, "Quantized consensus" Automatica, 1192-1203, 2007

Béreznitz, Thiran, Vetterli, "Interval consensus: From quantized gossip to voting", ICASSP 2009

Mezö Dániel

Majority Consensus by Local Polling

Proof: Spectrum of Q_S

Proof

Note that

$$Q_S = \begin{bmatrix} \text{diag}(q_{ii}, i \in S) & \mathbf{0} \\ (q_{ij})_{j \in S^c, j \in S} & (q_{jj})_{j \in S^c} \end{bmatrix}$$

- First $(q_{ii} = -\sum_{j \neq i} q_{ij})$, $i \in S$ are eigenvalues of Q_S
- The remaining eigenvalues correspond to eigenvectors $\underline{x} = (\underbrace{0, \dots, 0}_{S^c}, \underbrace{\underline{x}}_{S^c})^T$. Let $W \subset S^c$, for $i \in W, x_i \neq 0$

$$\begin{aligned} -\lambda &= \underline{x}^T Q_S \underline{x} \\ &= \sum_{i \in W} \sum_{j \in S} q_{ij} x_j^2 + \sum_{i \in W, j \in S^c \setminus W} q_{ij} x_i^2 + \frac{1}{2} \sum_{i, j \in W} q_{i,j} (x_i - x_j)^2 \end{aligned}$$

Complete graph

Moaz Draief Majority Consensus by Local Polling

$$\mathbb{E}(Y(t)) = \mathbb{E}[e^{\lambda(t)} Y(0)]$$

where $\lambda(t) = Q_{S_k}(t - t_k) + \sum_{l=0}^{k-1} Q_S(t_{l+1} - t_l)$
By Jensen's and matrix norm inequalities,

$$\|\mathbb{E}(Y(t))\|_2 \leq \mathbb{E} \left[\|\mathbb{E}^{Q_{S_k}(t-t_k)} \| \prod_{l=0}^{k-1} \|e^{Q_S(t_{l+1}-t_l)}\| \|Y(0)\|_2 \right] \leq \sqrt{n} e^{-\delta(G, \alpha)t}$$

Therefore, by Cauchy-Schwartz, we have

$$\mathbb{P}(Y(t) \neq \mathbf{0}) \leq \sum_{i \in V} \mathbb{E}(Y_i(t)) \leq n e^{-\delta(G, \alpha)t}$$

We conclude since $\mathbb{E}(T_0) = \int_0^\infty \mathbb{P}(Y(t) \neq \mathbf{0}) dt$.

Various initial conditions

Moaz Draief Majority Consensus by Local Polling

- $|S_0| - |S_1| = (2\alpha - 1)\eta$, α a constant larger than $1/2$

$$\mathbb{E}(T_1) = \frac{1}{2\alpha - 1} \log(\eta) + O(1).$$

- If $|S_0| = |S_1|$

$$\mathbb{E}(T_1) = \frac{\pi^2}{6} \eta(1 + o(1)).$$

- $\mu_n = (|S_0| - |S_1|)/n$ is strictly positive but small ($o(1)$),

$$\mathbb{E}(T_1) = \frac{1}{\mu_n} \log(n) + O(1).$$

Corollary

An application of the theorem to complete graph $q_{ij} = \frac{1}{n-1}$ for all $i \neq j$, yields

$$\mathbb{E}(T) \leq 2 \frac{1}{2\alpha - 1} \log(n).$$

Exact asymptotics

A direct analysis of the dynamics of the 1st phase

$$\mathbb{E}(T_1) = \frac{n-1}{|S_0| - |S_1|} (H_{|S_1|} + H_{|S_0|-|S_1|} - H_{|S_0|})$$

where $H_k = \sum_{l=1}^k \frac{1}{l}$

Moaz Draief Majority Consensus by Local Polling

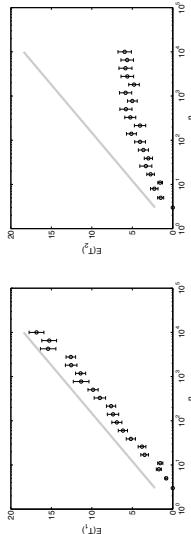
Moaz Draief Majority Consensus by Local Polling

Complete Graph: Theory v. Simulation

Star

- **Star Network:** $q_{i,i} = q_{i,1} = \frac{1}{n-1}$, $i \neq 1$ and $q_{i,j} = 0$, $i, j \neq 1$.

$$\mathbb{E}(T_i) \leq \frac{1}{2\alpha-1} n \log(n).$$



Mean Draft! Majority Consensus by Local Polling

ER-graph

- **ER-graph:** $q_{i,j} = \frac{1}{np_n} X_{i,j}$ i.i.d. Bernoulli r.v. with mean $C \frac{\log(n)}{n}$, $C > \frac{2}{2\alpha-1}$, for h^{-1} the inverse of $h(x) = x \log(x) + 1 - x$,

$$\mathbb{E}(T_i) \leq \frac{1}{(2\alpha-1)h^{-1}\left(\frac{2}{C(2\alpha-1)}\right)} \log(n) + O(1)$$

- Path: $\mathbb{E}(T_i) \leq \frac{16(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$

- Ring: $\mathbb{E}(T_i) \leq \frac{4(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$.

Mean Draft! Majority Consensus by Local Polling

Majority Consensus by Local Polling

- **Star Network:** $q_{i,i} = q_{i,1} = \frac{1}{n-1}$, $i \neq 1$ and $q_{i,j} = 0$, $i, j \neq 1$.

$$\mathbb{E}(T_i) \leq \frac{1}{2\alpha-1} n \log(n).$$

$$\mathbb{E}(T_i) = \frac{1}{(2\alpha-1)(3-2\alpha)} n \log(n) + O(n)$$

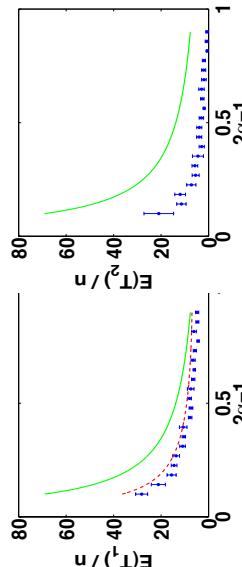
- **ER-graph:** $q_{i,j} = \frac{1}{np_n} X_{i,j}$ i.i.d. Bernoulli r.v. with mean $\rho_n = C \frac{\log(n)}{n}$, $C > \frac{2}{2\alpha-1}$, for h^{-1} the inverse of $h(x) = x \log(x) + 1 - x$,

$$\mathbb{E}(T_i) \leq \frac{1}{(2\alpha-1)h^{-1}\left(\frac{2}{C(2\alpha-1)}\right)} \log(n) + O(1)$$

- Path: $\mathbb{E}(T_i) \leq \frac{16(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$
- Ring: $\mathbb{E}(T_i) \leq \frac{4(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$.

Mean Draft! Majority Consensus by Local Polling

Majority Consensus by Local Polling



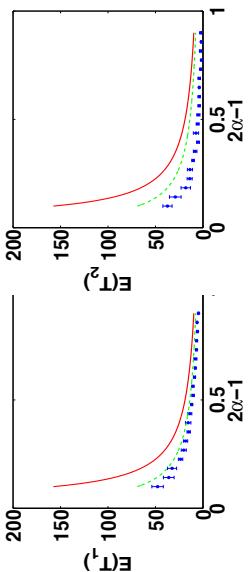
Mean Draft! Majority Consensus by Local Polling

Majority Consensus by Local Polling

Mean Draft! Majority Consensus by Local Polling

Majority Consensus by Local Polling

Path and Ring



Summary

- Star Network:** $q_{1,i} = q_{i,1} = \frac{1}{n-1}$, $i \neq 1$ and $q_{i,j} = 0$, $i, j \neq 1$.
 $\mathbb{E}(T_1) \leq \frac{1}{2(n-1)} n \log(n)$. Using, direct calculation

$$\mathbb{E}(T_1) = \frac{1}{(2\alpha - 1)(3 - 2\alpha)} n \log(n) + O(n)$$

- **ER-graph:** $q_{i,j} = \frac{1}{np_n} X_{i,j}$ i.i.d. Bernoulli r.v. with mean $\frac{\log(n)}{n}$, $C > \frac{2}{2\alpha - 1}$, for h^{-1} the inverse of $h(x) = x \log(x) + 1 - x$,

$$\mathbb{E}(T_i) \leq \frac{1}{(2\alpha - 1)h^{-1}\left(\frac{2}{c(2\alpha - 1)}\right)} \log(n) + O(1)$$

- **Path:** $\mathbb{E}(T_i) \leq \frac{16(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$
 - **Ring:** $\mathbb{E}(T_i) \leq \frac{4(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1).$

At $t = 0$, each vertex of G is blue independently with **constant probability** $\alpha \in (1/2, 1)$.

Local Majority

We then run $\lambda/7^{k^k}$ on G . Choose k odd ($k \geq 5$ in what follows).

- At each time t , each vertex v polls k neighbours uar , and assumes majority colour
- If v doesn't have k neighbours, poll all, or all minus one

What is the probability that there will be a red consensus?

How long does it take to reach consensus?

Upper bound on the expected convergence time for a number of distributed for solving Majority consensus

- Bounds based on the location of the spectral gap of rate matrix (generalised-cut: quick for expander graphs). For binary consensus, expected convergence time critically depends on the voting margin

Application to particular network topologies: complete graphs, stars, ER graph, paths, cycles.

Graphs of a given degree sequence

Nice degree sequences

Let $V = [n]$

$\mathcal{G}_n(\mathbf{d})$: the set of connected simple graphs with degree sequence $\mathbf{d} = (d_1, d_2, \dots, d_n)$, where d_i is the degree of vertex $i \in V$.

Need some restrictions on degree sequence to make it graphical, e.g., $\sum d_i$ is even

卷之三

e effective minimum degree

卷之三

(iii) Let $d \geq 5$ be such that $|V_d| = \kappa n + o(n)$. We call d

Need not be a constant; can have $d \rightarrow \infty$ as $n \rightarrow \infty$

Not necessarily the minimum degree (though it can be)
Can have "little" vertices with smaller degree, as long as not too many of them.

(iv) Number of little vertices $\sum_{j=\delta}^{d-1} |V_j| = O(n^{\frac{1}{11}})$; a vertex *i* is little if $d_i < d - 1$

Moez Draief Majority Consensus by Local Polling

卷之三

11

1

1

110

1

Results: informal statement

Suppose G is typical with effective min degree d . If we run $\lambda \vee P^k$ then

Upper bound

If $d/k = O(1)$ and α is 'not too close' to $1/2$, then **whp**, correct consensus is reached within $(A \log_k d) \log_k n$ steps ($A \leq 5$ and $A \rightarrow 1$ if $k \rightarrow \infty$)

Lower bound

Any algorithm where a vertex keeps its colour if same as all neighbours, will take at least $\log_d \log_d n$ steps to reach correct consensus, **whp**

Bias condition

" α is not too close to $1/2$ " means

$$\left[\left(1 + \frac{1}{\sqrt{k}} \right) 2 \right]^{\frac{2}{k-2}} \alpha(1-\alpha) < 1/4$$

Since $\alpha \neq 1/2 \Rightarrow \alpha(1-\alpha) < 1/4$, so inefficiency is in

$$\left[\left(1 + \frac{1}{\sqrt{k}} \right) 2 \right]^{\frac{2}{k-2}}$$

$K = 5$ needs $1 - \alpha < 0.143$

$K = 20$ needs $1 - \alpha < 0.350$

$K = 100$ needs $1 - \alpha < 0.437$

Compare with other works

Moaz Draiif Majority Consensus by Local Polling

Moaz Draiif Majority Consensus by Local Polling

Compare with other works

E. Mossel, J. Neeman, O. Tamuz ('14) Study local majority on d -regular λ -expanders. Show sufficient bias implies certain correct consensus.

better bias condition but only regular graphs, no timing information, full polling only

Y. Kanoria and A. Montanari ('10) Study local majority on d -regular infinite tree. Give bias conditions for convergence to majority
+better bias condition, -only infinite regular graph

Moaz Draiif Majority Consensus by Local Polling

Moaz Draiif Majority Consensus by Local Polling

Compare with other works

Typical graphs

E. Mossel, J. Neeman, O. Tamuz ('14) Study local majority on d -regular λ -expanders. Show sufficient bias implies certain correct consensus.

better bias condition but only regular graphs, no timing information, full polling only

Y. Kanoria and A. Montanari ('10) Study local majority on d -regular infinite tree. Give bias conditions for convergence to majority

+better bias condition, -only infinite regular graph

J. Cruise and A. Ganesh ('10) Study (m,d)-generalisation of local majority on complete graphs with unit rate exponential on each vertex. Give exponential decay error probability and $O(\log n)$ timing

+stronger error probability, -only complete graph

Typical graphs: For a nice degree sequence \mathbf{d} , the space $\mathcal{G}_n(\mathbf{d})$ is the set of nice graphs

We do not analyse for the whole space, only for those graphs called **typical**

Informally, G is typical if it is nice and:

- most vertices are locally tree-like
- little vertices and very high-degree vertices, should they exist, are far from each other and small cycles

Let $\mathcal{G}'_n(\mathbf{d}) \subset \mathcal{G}_n(\mathbf{d})$ be the typical graphs, then
 $|\mathcal{G}'_n(\mathbf{d})| / |\mathcal{G}_n(\mathbf{d})| \rightarrow 1$ as $n \rightarrow \infty$

Modified Majority

Mezz Draft!

Majority Consensus by Local Polling

Modified majority protocol

Mezz Draft!

Majority Consensus by Local Polling

- Let $\mathcal{T} = G[v, c \log_k \log_k n]$.
At $t+1$, each $x \in V$ randomly picks a $x(k)$ -subset of neighbours $N_x(t+1)$
- $x \notin \mathcal{T}$ then x becomes at $t+1$ the majority colour of the vertices in $N_x(t+1)$.

$$X_{t+1}^{\text{MMP}^k(v,s)}(x) = \mathbf{1}_{\left\{ \left(\sum_{y \in N_x(t+1) \setminus \{\text{Par}(x)\}} X_t^{\text{MMP}^k(y,s)}(y) \right) > x(k)/2 \right\}}.$$

- For a vertex v , let $X_v(t)$ be the indicator v is red at time t under MMP^k . Let $k = 2r + 1$.
- At time $t = 0$, for each level 2 (i.e., leaf) vertex v , $\mathbb{P}(X_v(0) = 0) = p_0 = 1 - \alpha$
 - At time $t = 1$, for each level 1 vertex v $\mathbb{P}(X_v(1) = 0) = p_1 = \mathbb{P}(\text{Bin}(2r, p_0) \geq r)$
 - At time $t = 2$, for each level 0 vertex v (i.e., the root) $\mathbb{P}(X_v(2) = 0) = p_2 = \mathbb{P}(\text{Bin}(2r, p_1) \geq r)$

$$X_{t+1}^{\text{MMP}^k(v,s)}(x) = \mathbf{1}_{\left\{ \left(\sum_{y \in N_x(t+1) \setminus \{\text{Par}(x)\}} X_t^{\text{MMP}^k(y,s)}(y) \right) > x(k)/2 \right\}}.$$

Mezz Draft!

Majority Consensus by Local Polling

Mezz Draft!

Majority Consensus by Local Polling

Modified majority protocol

If height of the tree is H , then given p_t , at $t + 1$, for v at distance $H - t - 1$ from root,

$$\mathbb{P}(X_\nu(t+1) = 0) = \rho_{t+1} = \mathbb{P}(\text{Bin}(2r, \rho_t) \geq r)$$

and we get a rapidly decaying sequence $p_0 > p_1 > \dots > p_t$

with $\rho_0 = \alpha \gg \rho_t$ when t large
When $t = \Omega(\log \log n)$, ρ_t is very small and we conclude by
union bound over all n vertices

The root will have the correct colour.
Now we are left to deal with vertices not locally tree-like...

卷之三

Summary

Related work

Biology, Volcanoes, and the Environment: Proceedings of a Conference

- Asymptotic correct and efficient consensus using local polling. What happens for other values of k ? [Cooper-Eisässer-Radzicki'14]

- Analysis for a sparse family of graphs and dense E-R graphs.

Gullane - f - 1 - 1 - 1 - 1 - 1 -

Becchetti, Clementi, Natale, Pasquale, Silvestri, trevisan,
Simple dynamics for plurality consensus. SPAA 2014.

Manitou Consonants | local Brillat
Draief 06/07

Theorem: Erdős-Renyi graphs

Let $p = \frac{c \log n}{n}$ where $c > 2 + \epsilon$ for some constant $\epsilon > 0$, $k \geq 5$ and $\nu = \lfloor \frac{k-1}{\nu} \rfloor$. Run MIP^k on $G \in \mathcal{G}(n, p)$. Let $A = \frac{1+\epsilon}{\log_k((k-1)-\log_2 \nu)}$ where $\epsilon > 0$ is a small constant. Subject to condition (1), we have

$$\left[\left(1 + \frac{1}{\sqrt{2\nu}} \right) 2 \right]^{\frac{1}{\nu-1}} 4\alpha(1-\alpha) < 1$$

by time $A \log_k n$, \mathcal{MP}^k will have reached consensus on the initial majority **whp**.

THE JOURNAL OF CLIMATE

**ROBUST NETWORKED STABILIZATION: WHEN GAP MEETS
TWO-PORT****Li Qiu, Hong Kong University of Science and Technology**

In robust networked stabilization, if the plant and controller uncertainties are described by the gap metric (or $\|\cdot\|_{\text{gap}}$ metric), and the communication network between the plant and the controller is described by a two-port with an uncertain transmission matrix, magic happens.



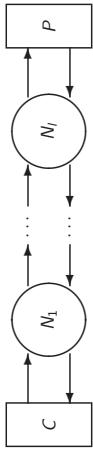
Networked control system (NCS)

Networked Robust Stabilization: when gap meets two-port

Guoxiang Gu and Li Qiu

Louisiana State University
Hong Kong University of Science and Technology

October 2014



- The plant and controller are uncertain and the uncertainty is described by the **gap metric**.
- The communication network is a cascade of **two-port networks**, modeling bidirectional transmission with relays.
- We only consider SISO plants and controllers in this talk, for the sake of simplicity.

Guoxiang Gu and Li Qiu [LSU and HKUST] Networked Robust Stabilization October 2014 1 / 15 October 2014 1 / 15

Gap metric: a review

(Zames; Vidyasagar; Georgiou and Smith; Qiu and Davison; Vinnicombe, from 1980 to ~1995.)

- P is LTI and possibly unstable.
- Graph of P

$$\mathcal{G}_P = \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in \mathcal{H}_2 \times \mathcal{H}_2 : y = Pu \right\}.$$

a subspace of $\mathcal{H}_2 \times \mathcal{H}_2$.

- Gap metric between P_1 and P_2 .

$$\delta(P_1, P_2) = \|\Pi_{\mathcal{G}_{P_1}} - \Pi_{\mathcal{G}_{P_2}}\|.$$

- Uncertain systems can be described by gap balls

$$\mathcal{B}(P, r) = \{\tilde{P} : \delta(P, \tilde{P}) \leq r\}.$$

- Gap ball viewed as rotation of the graph:

$$\{\tilde{P} : \mathcal{G}_{\tilde{P}} = (I + \Delta)\mathcal{G}_P, \|\Delta\|_\infty \leq r\} \subset \mathcal{B}(P, r).$$

The maximal rotating angle is $\arcsin r$.

Guoxiang Gu and Li Qiu [LSU and HKUST] Networked Robust Stabilization October 2014 3 / 15 October 2014 3 / 15

Gap metric: a review (continued)

Guoxiang Gu and Li Qiu [LSU and HKUST] Networked Robust Stabilization

October 2014

1 / 15

October 2014

2 / 15

October 2014

3 / 15

October 2014

4 / 15

October 2014

5 / 15

October 2014

6 / 15

October 2014

7 / 15

October 2014

8 / 15

October 2014

9 / 15

October 2014

10 / 15

October 2014

11 / 15

October 2014

12 / 15

October 2014

13 / 15

October 2014

14 / 15

October 2014

15 / 15

October 2014

16 / 15

October 2014

17 / 15

October 2014

18 / 15

October 2014

19 / 15

October 2014

20 / 15

October 2014

21 / 15

October 2014

22 / 15

October 2014

23 / 15

October 2014

24 / 15

October 2014

25 / 15

October 2014

26 / 15

October 2014

27 / 15

October 2014

28 / 15

October 2014

29 / 15

October 2014

30 / 15

October 2014

31 / 15

October 2014

32 / 15

October 2014

33 / 15

October 2014

34 / 15

October 2014

35 / 15

October 2014

36 / 15

October 2014

37 / 15

October 2014

38 / 15

October 2014

39 / 15

October 2014

40 / 15

October 2014

41 / 15

October 2014

42 / 15

October 2014

43 / 15

October 2014

44 / 15

October 2014

45 / 15

October 2014

46 / 15

October 2014

47 / 15

October 2014

48 / 15

October 2014

49 / 15

October 2014

50 / 15

October 2014

51 / 15

October 2014

52 / 15

October 2014

53 / 15

October 2014

54 / 15

October 2014

55 / 15

October 2014

56 / 15

October 2014

57 / 15

October 2014

58 / 15

October 2014

59 / 15

October 2014

60 / 15

October 2014

61 / 15

October 2014

62 / 15

October 2014

63 / 15

October 2014

64 / 15

October 2014

65 / 15

October 2014

66 / 15

October 2014

67 / 15

October 2014

68 / 15

October 2014

69 / 15

October 2014

70 / 15

October 2014

71 / 15

October 2014

72 / 15

October 2014

73 / 15

October 2014

74 / 15

October 2014

75 / 15

October 2014

76 / 15

October 2014

77 / 15

October 2014

78 / 15

October 2014

79 / 15

October 2014

80 / 15

October 2014

81 / 15

October 2014

82 / 15

October 2014

83 / 15

October 2014

84 / 15

October 2014

85 / 15

October 2014

86 / 15

October 2014

87 / 15

October 2014

88 / 15

October 2014

89 / 15

October 2014

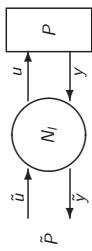
90 / 15

October 2014

91 / 15

Plant with two-port distortion

System with $+$, $-$, \times , \div uncertainty.



- Linear fractional transformation (LFT)

$$\tilde{P} = \frac{(1 + \Delta_x)P + \Delta_+}{1 + \Delta_+ + \Delta_- P}.$$

- Graph of the distorted system

$$\tilde{\mathcal{G}}_{\tilde{P}} = (I + \Delta)\mathcal{G}_P.$$

The same type of rotation as in the gap ball!

Two-port transmission model of NCS

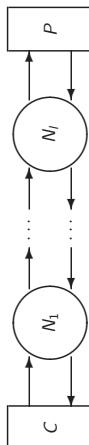


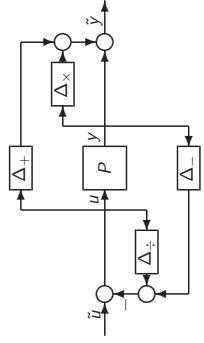
Figure: Networked control system.

- \tilde{P}, \tilde{C} are only known to belong to gap balls

$$\tilde{P} \in \mathcal{B}(P, r_P), \tilde{C} \in \mathcal{B}(C, r_C).$$

- N_i is only known to have transmission matrix

$$A_i = I + \Delta_i, \|\Delta_i\|_\infty \leq r_i.$$



- Graph of the distorted system

$$\tilde{P} = \frac{(1 + \Delta_x)P + \Delta_+}{1 + \Delta_+ + \Delta_- P}.$$

- Graph of the distorted system

$$\tilde{\mathcal{G}}_{\tilde{P}} = (I + \Delta)\mathcal{G}_P.$$

The same type of rotation as in the gap ball!

Growing Gu and Li Qiu (LSU and HKUST) Networked Robust Stabilization

October 2014 9 / 15 Growing Gu and Li Qiu (LSU and HKUST) Networked Robust Stabilization

October 2014 10 / 15 Growing Gu and Li Qiu (LSU and HKUST) Networked Robust Stabilization

October 2014 11 / 15 Growing Gu and Li Qiu (LSU and HKUST) Networked Robust Stabilization

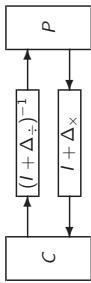
$$\arcsin r_P + \arcsin r_C + \sum_{i=1}^l \arcsin r_i < \arcsin \|GoF(P, C)\|_\infty^{-1}.$$

- The NCS is robustly stable iff

$$\min_{\tilde{C}} \|GoF(P, C)\|_\infty.$$

$$\begin{aligned} &\text{Growing Gu and Li Qiu (LSU and HKUST)} \\ &\text{Networked Robust Stabilization} \\ &\text{October 2014 12 / 15 Growing Gu and Li Qiu (LSU and HKUST) Networked Robust Stabilization} \end{aligned}$$

Hammerstein-Wiener model of NCS



- Δ_x and Δ_{\div} are nonlinear time-varying systems satisfying

$$\|\Delta_x\|_\infty \leq r, \quad \|\Delta_{\div}\|_\infty \leq r.$$

- Two-port network with diagonal transmission matrix.
- Δ_x and Δ_{\div} can be used to model logarithmic quantizations.
- The NCS is robustly stable iff

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \Delta_{\div} & 0 \\ 0 & \Delta_x \end{bmatrix} G_o F$$

is stably invertible for all Δ_x and Δ_{\div} .

Conclusions

- Trade-off between the capacities of the down-link and the up-link channels.
- Optimal robust networked control linking the history.
- H_2 vs H_∞ theory

μ -synthesis of Hammerstein-Wiener system

- Introducing scaling: the closed-loop system is robustly stable iff

$$\begin{aligned} & \inf_{\gamma \in (0, \infty)} \left\| \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix} G_o F(P, C) \begin{bmatrix} 1 & 0 \\ 0 & \gamma^{-1} \end{bmatrix} \right\|_\infty \\ &= \inf_{\gamma \in (0, \infty)} \left\| \begin{bmatrix} \frac{1-P\bar{C}}{\gamma} & \frac{-\gamma^{-1}C}{1-P\bar{C}} \\ \frac{1-P\bar{C}}{\gamma} & \frac{1-P\bar{C}}{1-P\bar{C}} \end{bmatrix} \right\|_\infty \\ &= \inf_{\gamma \in (0, \infty)} \|G_o F(\gamma P, \gamma^{-1} C)\|_\infty < \frac{1}{r}. \end{aligned}$$

- Optimal design
- The function $\gamma \mapsto \inf_C \|G_o F(\gamma P, \gamma^{-1} C)\|_\infty$ was mistakenly conjectured to be unimodal.

\square

SCHROEDINGER BRIDGES: STEERING OF STOCHASTIC SYSTEMS CLASSICAL AND QUANTUM
Tryphon Georgiou, University of Minnesota, USA

The classical Schroedinger bridge seeks the most likely probability law for a diffusion process, in path space, that matches marginals at two end points in time; the likelihood is quantified by the relative entropy between the sought law and a prior, and the law dictates a controlled path that abides by the specified marginals. Schreodinger proved that the optimal steering of the density between the two end points is effected by a multiplicative functional transformation of the prior; this transformation represents an automorphism on the space of probability measures and has since been studied by Fortet, Beurling and others. A similar question can be raised for processes evolving in a discrete time and space as well as for processes defined over non-commutative probability spaces. Ultimately, in all of the above Schrodinger's question relates to a corresponding stochastic control problem to steer a stochastic system from an initial to a final distribution. In the talk we will begin with a treatment of the Schroedinger bridge problem for Markov chains, Quantum channels, and then for classical stochastic systems where we will present in the linear-quadratic case (linear dynamics, Gaussian evolution) the solution to the minimum energy steering problem in closed form. We will discuss potential applications in active suppression of noise and in steering swarms of inertial diffusive particles. The presentation will be based on joint work with Michele Pavon and with Yongxin Chen.



Schrödinger Bridges classical and quantum evolution

Tryphon Georgiou
University of Minnesota

Joint work with Michele Pavon and Yongxin Chen

LCCC
Workshop on Dynamic and Control in Networks
Lund, October 2014

- History of Schrödinger bridges
- Bridges for Markov chains
- The Hilbert metric
- Bridges for quantum (TPTP) evolutions
- Bridges for Gauss-Markov process

Schrodinger 1931/32:
The time reversal of the laws of nature

Kolmogoroff:
The reversibility of the statistical laws of nature

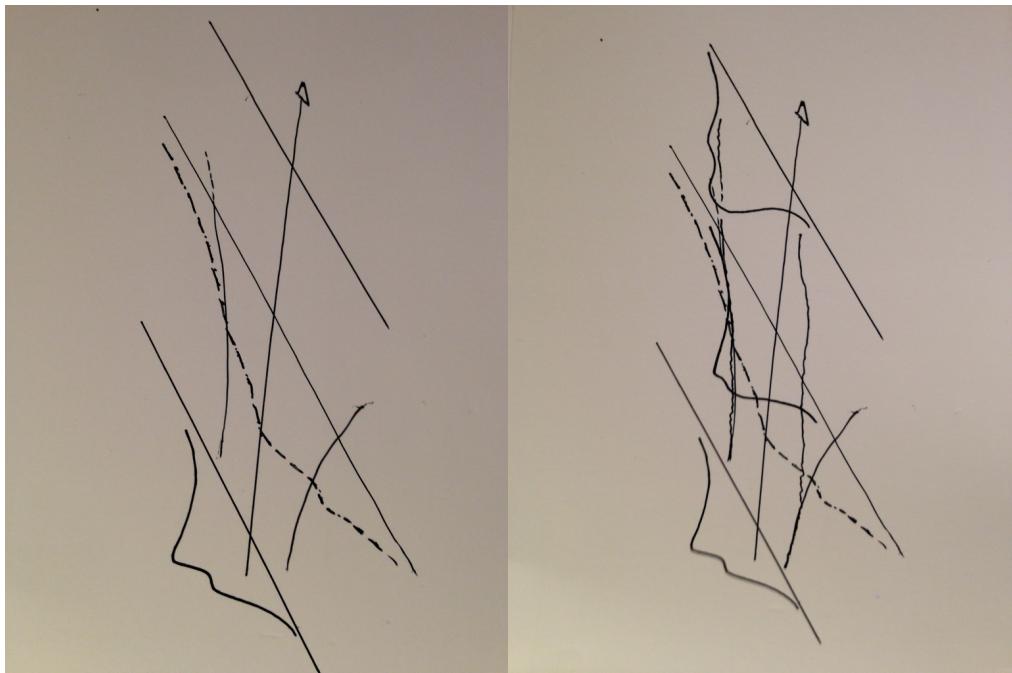
Bernstein 1932
Fortet 1940
Beurling 1960
Jamison 1974/75
Follmer 1988

Hilbert metric:

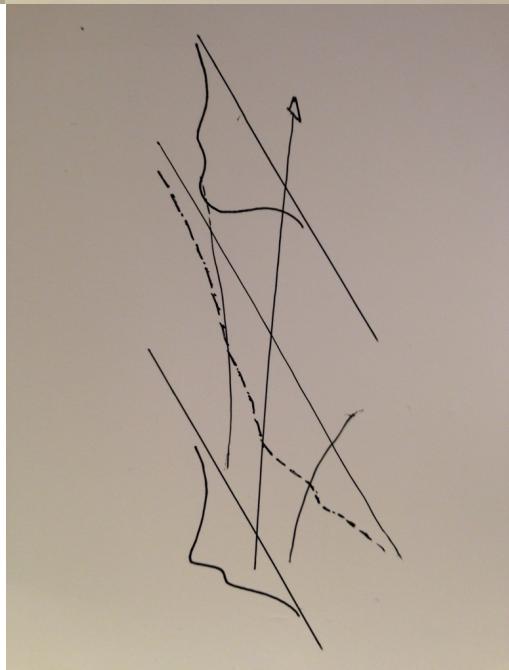
Hilbert 1895
Birkhoff 1957
Bushell 1973

Sepulchre, Sarlette, Rouchon 2010
Reeb, Kastoryano, Wolf 2011

connections to Nelson's stochastic mechanics
Zambelli, Wakolbinger, Dai Pra, Pavon, Ticcozzi
and others



- Schrodinger 1931/1932: suppose a large number of Brownian particles observed at two different times to evolve between two empirical distributions. What is the most likely intermediate distribution at any point in time?



Given initial and final distribution $p_0(x)$, $p_T(x)$ and transition $p(x, y)$

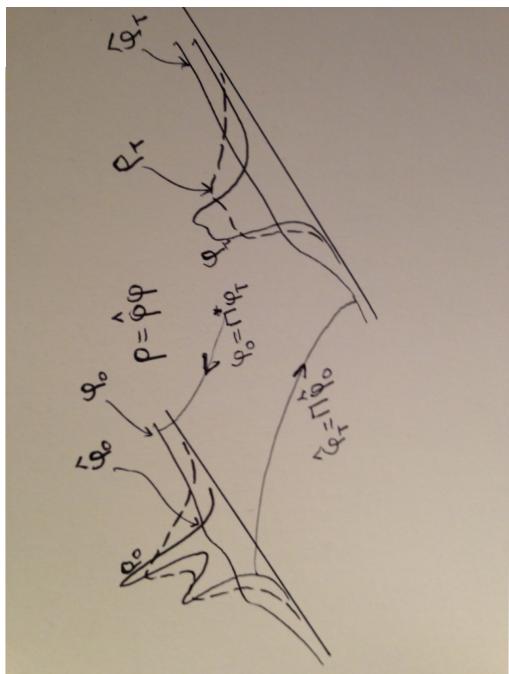
Schrödinger hypothesised that

$$p_T(\cdot) \neq \int p(x, \cdot) p_0(x) dx \\ =: \Pi_{0T}(p_0(x))$$

$$\Pi_{0t} : q_0(x) \rightarrow q_t(x)$$

$$\frac{\partial q_t(x)}{\partial t} = \frac{1}{2} \frac{\partial^2 q_t(x)}{\partial x^2}, \quad q(0, x) = q_0(x).$$

Large deviations
Sample paths
Relative entropy
Stochastic Control



Schrödinger system

discretized time, space, N particles
Schrödinger's approximation
optimized, lagrange multipliers

the most likely joint density and transition probability

$$D^*(x_0, x_T) = \hat{\phi}(x_0)p(x_0, x_T)\phi(x_T) \quad \text{and} \quad p^*(x_0, x_T) = p(x_0, x_T) \frac{q(x_T)}{q(x_0)}$$

$$p_0(x_0) = \hat{\phi}(x_0) \underbrace{\int p(x_0, x_T)\phi(x_T)dx_T}_{\phi(x_0)} = \hat{\phi}(x_0) \underbrace{\prod_{t=0}^{T-1} \underbrace{\int \phi(x_T)}_{\hat{\phi}(x_T)} \underbrace{\int \phi(x_T)dx_T}_{\phi(x_T)}}_{\hat{\phi}(x_0)}$$

$$p_T(x_T) = \phi(x_T) \underbrace{\int p(x_0, x_T)\phi(x_0)dx_0}_{\phi(x_T)} = \phi(x_T) \underbrace{\prod_{t=0}^{T-1} \underbrace{\int \phi(x_0)}_{\hat{\phi}(x_0)} \underbrace{\int \phi(x_0)dx_0}_{\phi(x_0)}}$$

and $\boxed{p_t(x_t) = \hat{\phi}_t(x_t)\phi_t(x_t)}$ where
 $\hat{\phi}_t(x_t) := \prod_{i=t}^{T-1} \hat{\phi}(x_i)$
 $\phi_t(x_t) := \prod_{i=t}^{T-1} \phi(x_i)$

Schrödinger system

Schrödinger: there exists a solution “except possibly for very nasty p_0, p_T because the question leading to the pair of equations is so reasonable.”

Existence/uniqueness

Fortet 1940

Résolution d'un système d'équations de M. Schrödinger

Berning 1960

An automorphism of product measures

Markov chains

$\{1, \dots, N\}$ states, $x = (x_0, x_1, \dots, x_T)$ sample path
 Π_t stochastic matrices, $t \in \{1, \dots, T\}$
 P in probability induced by Π_t 's on $\{1, \dots, N\}^{T+1}$

$$P(x_0, \dots, x_T) = P(x_0, x_T)P(x_1, \dots, x_{T-1} | x_0, x_T)$$

Schrödinger question

given p_0, p_T
 $p_T \neq \Pi_T \dots \Pi_1 p_0$

$$\text{find } Q(x_0, \dots, x_T) = Q(x_0, x_T)Q(x_1, \dots, x_{T-1} | x_0, x_T)$$

such that
 $\sum_{x_T} Q(x_0, x_T) = p_0(x_0)$
 $\sum_{x_0} Q(x_0, x_T) = p_T(x_T)$
and minimizes the relative entropy

$$\sum_{x_0, x_T} Q \log \frac{Q}{P} = \sum_{x_0, x_T} Q(x_0, x_T) \log \frac{Q(x_0, x_T)}{P(x_0, x_T)} + \sum_{x_0, x_T} Q(\cdot | x_0, x_T) \log \frac{Q(\cdot | x_0, x_T)}{P(\cdot | x_0, x_T)} Q(x_0, x_T)$$

Lagrangian

$$\begin{aligned} L(Q) &= \sum_{x_0, x_T} Q(x_0, x_T) \log \frac{Q(x_0, x_T)}{P(x_0, x_T)} \\ &\quad + \sum_{x_0} \lambda(x_0) \left(\sum_{x_T} Q(x_0, x_T) - p_T(x_T) \right) \\ &\quad + \sum_{x_T} \mu(x_T) \left(\sum_{x_0} Q(x_0, x_T) - p_0(x_0) \right) \end{aligned}$$

$$\lambda(x_0) \sim \hat{\phi}_0$$

$$\mu(x_T) \sim \hat{\phi}_T$$

such that with $\Pi = \Pi_T \dots \Pi_1$

Schrödinger system
 $\hat{\phi}_T = \Gamma \hat{\phi}_0$
 $\hat{\phi}_0 = \Gamma^\dagger \hat{\phi}_T$
 $p_0 = \phi_0 \circ \hat{\phi}_0$
 $p_T = \phi_T \circ \hat{\phi}_T$

if there is a solution

$$\Pi^* = D_{\phi_T} \Pi D_{\phi_0}^{-1}$$

$$[Q(x_0, x_T)]_{x_0, x_T} = D_{\phi_T} \Pi D_{\phi_0}$$

Hilbert metric

S real Banach space
 K closed solid cone in S

$$x \preceq y \Leftrightarrow y - x \in K,$$

$$\begin{aligned} M(x, y) &:= \inf \{\lambda \mid x \preceq \lambda y\} \\ m(x, y) &:= \sup \{\lambda \mid \lambda y \preceq x\}. \end{aligned}$$

define the Hilbert metric:

$$d_H(x, y) := \log \left(\frac{M(x, y)}{m(x, y)} \right).$$

Examples:

- i) positive cone in \mathbb{R}^n
- ii) positive definite Hermitian matrices

d_H -gain bound of positive maps

Π is a positive map:

$$\Pi : K \setminus \{0\} \rightarrow K \setminus \{0\}.$$

Projective diameter

$$\Delta(\Pi) := \sup\{d_H(\Pi(x), \Pi(y)) \mid x, y \in K \setminus \{0\}\}$$

Contraction ratio, or gain/ H -norm

$$\|\Pi\|_H := \inf\{\lambda \mid d_H(\Pi(x), \Pi(y)) \leq \lambda d_H(x, y), \text{ for all } x, y \in K \setminus \{0\}\}.$$

Birkhoff-Bushell theorem

Let Π positive, monotone, homogeneous of degree m , i.e.,

$$x \preceq y \Rightarrow \Pi(x) \preceq \Pi(y),$$

and

$$\Pi(\lambda x) = \lambda^m \Pi(x),$$

then

$$\|\Pi\|_H \leq m.$$

For the special case where Π is also linear, the (possibly stronger) bound

$$\|\Pi\|_H = \tanh\left(\frac{1}{4}\Delta(\Pi)\right)$$

also holds.

Solution of the Schrödinger system

Consider

$$\varphi_0 \xrightarrow{\Pi} \varphi_T$$

$$\varphi_0(x_0) = \frac{\mathbf{p}_0(x_0)}{\varphi_0(x_0)} \uparrow \quad \downarrow \quad \varphi_T(x_T) = \frac{\mathbf{p}_T(x_T)}{\varphi_T(x_T)}$$

$$\varphi_0 \xleftarrow{\Pi^\dagger} \varphi_T$$

where

$$\mathcal{D}_T : \varphi_0 \mapsto \hat{\varphi}_0(x_0) = \frac{\mathbf{p}_0(x_0)}{\varphi_0(x_0)}$$

$$\mathcal{D}_T : \hat{\varphi}_T \mapsto \mathcal{P}_T(x_T) = \frac{\mathbf{p}_T(x_T)}{\hat{\varphi}_T(x_T)}$$

are componentwise division of vectors $\Rightarrow d_H$ -isometries!

The composition

$$\hat{\varphi}_0 \xrightarrow{\Pi} \hat{\varphi}_T \xrightarrow{\mathcal{P}_T} \varphi_T \xrightarrow{\Pi^\dagger} \varphi_0 \xrightarrow{\mathcal{D}_0} (\hat{\varphi}_0)_{\text{next}}$$

is strictly contractive in the Hilbert metric.

Lemma
Let $\Pi \geq_e 0$ (element-wise positive) stochastic matrix
 p_0, p_T probability vectors

then $\|\Pi\|_H < 1$.

proof

i) $\Delta(\Pi) = \sup\{d_H(\Pi(x), \Pi(y)) \mid x, y \in K \setminus \{0\}\}$

remains the same if we restrict x, y to be probability vectors

ii) $d_H(\Pi(x), \Pi(y)) < \infty \quad \forall x, y$.

iii) the probability simplex is compact.

Sinkhorn's theorem

If $\Pi >_e 0$,
then $\exists a_i, b_j$
such that $[\pi_{ij}a_i b_j]_{i,j}$ doubly stochastic.

TPCP maps $\{\mathcal{E}_t, 0 \leq t \leq T-1\}$
with Kraus representation

$\mathcal{E}_t : \sigma_t \mapsto \sigma_{t+1} = \sum_i E_{t,i} \sigma_t E_{t,i}^\dagger, \quad t = 0, 1, \dots, T-1.$

Consider the composition
 $\mathcal{E}_{0:T} := \mathcal{E}_{T-1} \circ \dots \circ \mathcal{E}_0.$
 initial and a final ρ_0 and ρ_T

Problem
 Find $\mathcal{F}_{0:T} = \mathcal{F}_{T-1} \circ \dots \circ \mathcal{F}_0$ such that
 $\mathcal{F}_{0:T}(\rho_0) = \rho_T.$
 and \mathcal{F} "close to" \mathcal{E}

Quantum analogues

Density matrices: $\mathfrak{D} = \{\rho \geq 0 \mid \text{trace}(\rho) = 1\}$

TPTP: $\mathcal{E} : \mathfrak{D} \rightarrow \mathfrak{D} : \rho \longrightarrow \sigma = \sum_{i=1}^{n_E} E_i \rho E_i^\dagger$
 with

$$\sum_{i=1}^{n_E} E_i^\dagger E_i = I$$

i.e., $\mathcal{E}^\dagger(I) = I$

\mathcal{E} is positivity improving: if $\rho \geq 0 \Rightarrow \mathcal{E}(\rho) > 0$

Reference quantum evolution

"rank-1" corrections

$$\mathcal{F}_t(\cdot) = \chi_{t+1} \left(\mathcal{E}_t(\chi_t^{-1}(\cdot) \chi_t^{-\dagger}) \right) \chi_{t+1}^\dagger$$

i.e., $\mathcal{F}_t = \Phi_{t+1} \circ \mathcal{E}_t \circ \Phi_t^{-1}$ where
 Φ are rank-1 Kraus maps, $n_\Phi = 1$

Corresponds to the commutative case via: $\chi^\dagger \chi = \phi$

Quantum version of Sinkhorn's thm

Suppose $\mathcal{E}_{0,T}$ is positivity improving
 Then, \exists observables ϕ_0, ϕ_T such that,
 for any factorization

$$\begin{aligned}\phi_0 &= \chi_0^\dagger \chi_0, \text{ and} \\ \phi_T &= \chi_T^\dagger \chi_T,\end{aligned}$$

the map

$$\mathcal{F}(\cdot) := \chi_T \left(\mathcal{E}_{0,T}(\chi_0^{-1}(\cdot) \chi_0^{-\dagger}) \right) \chi_T^\dagger$$

is a *doubly stochastic* Kraus map,
 in that $\mathcal{F}(I) = I$ as well as $\mathcal{F}^\dagger(I) = I$.

Proof

$$\begin{array}{ccc}\hat{\phi}_0 & \xrightarrow{\mathcal{E}_{0,T}} & \hat{\phi}_T \\ \hat{\phi}_0 = \phi_0^{-1} & \uparrow & \downarrow \quad \phi_T = \hat{\phi}_T^{-1} \\ \phi_0 & \xleftarrow{\mathcal{E}_{0,T}^\dagger} & \phi_T\end{array}$$

The composition map

$$\mathcal{C} : (\hat{\phi}_0)_{\text{starting}} \xrightarrow{\mathcal{E}_{0,T}} \hat{\phi}_T \xrightarrow{(\cdot)^{-1}} \phi_T \xrightarrow{\mathcal{E}_{0,T}^\dagger} \phi_0 \xrightarrow{(\cdot)^{-1}} (\hat{\phi}_0)_{\text{next}}$$

is strictly contractive
 the steps are identical

General case

Given $\mathcal{E}_{0,T}^\dagger$ and ρ_0 and ρ_T
 if $\exists \phi_0, \phi_T, \hat{\phi}_T, \hat{\phi}_r$ solving

$$\begin{aligned}\mathcal{E}_{0,T}^\dagger(\phi_T) &= \phi_0, \\ \mathcal{E}_{0,T}^\dagger(\hat{\phi}_0) &= \hat{\phi}_r, \\ \rho_0 &= \lambda \hat{\phi}_0 \phi_0^\dagger, \\ \rho_T &= \lambda \hat{\phi}_T \phi_T^\dagger.\end{aligned}$$

Then, for any factorization

$$\begin{aligned}\phi_0 &= \chi_0^\dagger \chi_0, \text{ and} \\ \phi_T &= \chi_T^\dagger \chi_T,\end{aligned}$$

the map

$$\mathcal{F}(\cdot) := \chi_T \left(\mathcal{E}_{0,T}(\chi_0^{-1}(\cdot) \chi_0^{-\dagger}) \right) \chi_T^\dagger$$

is a quantum bridge for $(\mathcal{E}_{0,T}, \rho_0, \rho_T)$, namely $\mathcal{F}(I) = I$
 and $\mathcal{F}^\dagger(\rho_0) = \rho_T$.

The quantum Schrödinger system has a solution
 for arbitrary ρ_0, ρ_T

Sing in the proof:
 $\phi \mapsto \hat{\phi}$ and $\phi \mapsto \hat{\phi}$ are not isometries, e.g.,

$$\begin{aligned}D_T : \hat{\phi}_T \mapsto \phi_T &= \left(\rho_T^{1/2} \left(\rho_T^{-1/2} \phi^{-1} \rho_T^{-1/2} \right)^{1/2} \rho_T^{1/2} \right)^2 \\ D_0 : \phi_0 \mapsto \hat{\phi}_0 &= (\phi_0)^{1/2} \rho(\phi_0)^{1/2}\end{aligned}$$

Extensive numerical evidence that
 the composition has a fixed point
 Software for numerical experimentation
<http://www.ece.umn.edu/~georgiou/papers/schrodinger-bridge/>

Pinned bridge

$\mathcal{E}_{0:T}$ positivity improving and two pure states

$$\rho_0 = v_0 v_0^\dagger \text{ and } \rho_T = v_T v_T^\dagger$$

(i.e., v_0, v_T are unit norm vectors), define

$$\phi_0 := \mathcal{E}(v_0 v_0^\dagger)$$

$$\phi_T := v_T v_T^\dagger$$

and

$$\mathcal{F}^\dagger(\cdot) := \phi_T^{1/2} \mathcal{E}(\phi_0^{-1/2}(\cdot) \phi_0^{-1/2}) \phi_T^{1/2}$$

(where, clearly, $\phi_T^{1/2} = \phi_T = v_T v_T^\dagger$). Then, \mathcal{F}^\dagger is TPTP and satisfies the marginal conditions

$$\rho_T = \mathcal{F}^\dagger(\rho_0).$$

Example

$$\begin{aligned}\mathcal{E}(\cdot) &= E_1(\cdot)E_1^\dagger + E_2(\cdot)E_2^\dagger + E_3(\cdot)E_3^\dagger \\ E_1 &= \begin{bmatrix} \sqrt{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\frac{1}{2}} \end{bmatrix}, \quad E_3 = \begin{bmatrix} 0 & \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} & 0 \end{bmatrix} \\ \rho_0 &= \begin{bmatrix} 1/4 & 0 \\ 0 & 3/4 \end{bmatrix} \text{ and } \rho_T = \begin{bmatrix} 2/3 & 0 \\ 0 & 1/3 \end{bmatrix}\end{aligned}$$

$$\begin{aligned}\phi_0 &= \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \\ \phi_1 &= \begin{bmatrix} 2/3 & 0 \\ 0 & 1/3 \end{bmatrix} \\ \phi_2 &= \begin{bmatrix} 1/2 & 0 \\ 0 & 3/2 \end{bmatrix} \\ \phi_3 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\end{aligned}$$

$$F_1 = \begin{bmatrix} \sqrt{2/3} & 0 \\ 0 & 0 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{1/3} \end{bmatrix}, \quad F_3 = \begin{bmatrix} 0 & \sqrt{2/3} \\ \sqrt{1/3} & 0 \end{bmatrix}$$

Recap

Hilbert metric \Rightarrow constructive existence proofs for

- i) classical Schrödinger systems
- ii) quantum Sinkhorn version (uniform marginals)
- iii) general case open

Final topic:

Schrödinger bridges for "degenerate" classical linear stochastic systems

\equiv a new type of optimal control problem

Optimal steering of state-densities

min relative entropy \leftrightarrow Clebsch \leftrightarrow minimum energy stochastic control

$$dx = bdt + dw \text{ diffusion}$$

$$dx = (b+u)dt + dw \text{ controlled diffusion}$$

$$\begin{aligned}\min\{E\{\|u\|^2\} \mid p_0, p_T\} &\sim \text{relative entropy from prior} \\ (\text{dai PrA})\end{aligned}$$

our interest:
inertial particles, cooling of oscillators

$$\begin{aligned}dx &= vdt \\ dw &= (b+u)dt + dw \text{ controlled degenerate diffusion}\end{aligned}$$

Optimal steering of state-densities

$$dx(t) = A(t)x(t)dt + B(t)u(t)dt + B(t)dw(t)$$

Given initial and terminal (target) Gaussian densities with covariances Σ_0, Σ_T .

Find $u(t)$ with $t \in [0, T]$ that steers the system from the initial to the target state density and minimizes

$$E\left\{\int_0^T u(t)'u(t)dt\right\}$$

Optimal steering of state-densities

Theorem (Gauss-Markov Schrödinger bridge):

There exists a unique solution to the following (analogue of the Schrödinger system)

$Q(T), P(0)$ values for matrices satisfying

$$\Sigma_0^{-1} = Q(0)^{-1} + P(0)^{-1}$$

$$\Sigma_T^{-1} = Q(T)^{-1} + P(T)^{-1}$$

and $Q(0), P(T)$ obtained via

$$\dot{Q}(t) = A(t)Q(t) + Q(t)A(t)' + B(t)B(t)'$$

$$\dot{P}(t) = A(t)P(t) + P(t)A(t)' - B(t)B(t)'$$

with $Q(t)$ invertible over $[0, T]$.

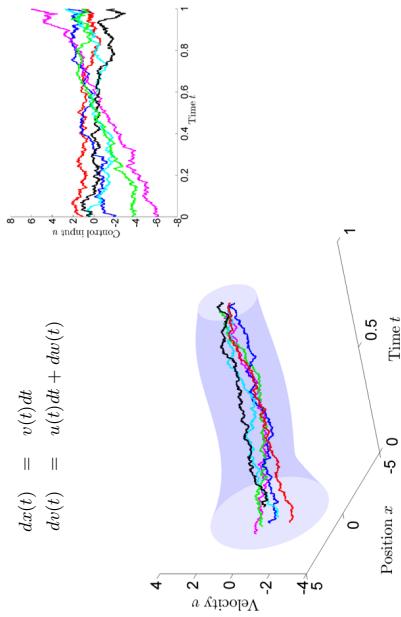
The optimal control is $u(t) = -B(t)'Q(t)^{-1}x(t)$

The controlled degenerate diffusion is the closest to the uncontrolled diffusion in the relative entropy sense.

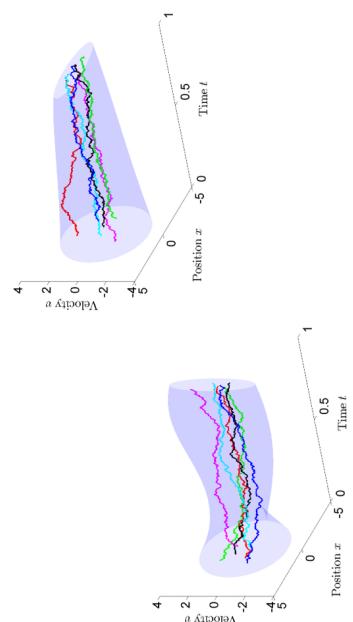
$$\begin{aligned} Q(0) &= N(T, 0)^{1/2}S_0^{1/2}\left(S_0 + \frac{1}{2}I - \left(S_0^{1/2}S_T S_0^{1/2} + \frac{1}{4}I\right)^{1/2}\right)^{-1}S_0^{1/2}N(T, 0)^{1/2} \\ &+ \frac{1}{4}I \end{aligned}$$

$N(T, 0)$ is the controllability Grammian.

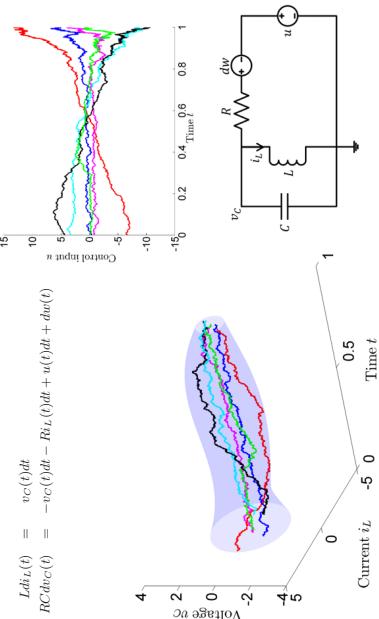
Gauss Markov model for inertial particles



Gauss Markov model for inertial particles



Gauss Markov model for Nyquist-Johnson noise driven oscillator



Gauss Markov model for inertial particles:
state-cost \sim particles with losses

$$dX(t) = f(X(t), t)dt + \sigma(X(t), t)d\omega(t)$$

$$\inf_{(\bar{\rho}, \bar{u})} \int_{\mathbb{R}^N} \int_0^T \left[\frac{1}{2} \|u\|^2 + V(x, t) \right] \bar{\rho}(x, t) dt dx,$$

$$\frac{\partial \bar{\rho}}{\partial t} + \nabla \cdot ((f + \sigma u) \bar{\rho}) = \frac{1}{2} \sum_{i,j=1}^N \frac{\partial^2 (a_{ij} \bar{\rho})}{\partial x_i \partial x_j},$$

$$a_{ij}(x, t) = \sum_k \sigma_{ik}(x, t) \sigma_{kj}(x, t)$$

$$\bar{\rho}(0, x) = \rho_0(x), \quad \bar{\rho}(T, y) = \rho_T(y).$$

$$\varphi(x, 0)\bar{\varphi}(x; 0) = \bar{\rho}_0(x), \quad \varphi(x, T)\bar{\varphi}(x, T) = \bar{\rho}_T(x).$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (f(x, t)\rho) + V(x, t)\rho = \frac{1}{2} \sum_{i,j=1}^N \frac{\partial^2 (a_{ij} \rho)}{\partial x_i \partial x_j},$$

$$w^*(x, t) = \sigma \nabla \log \varphi(x, t),$$

$$\frac{\partial \bar{\varphi}}{\partial t} + \nabla \cdot ((f + \sigma \nabla \log \varphi) \bar{\rho}) = \frac{1}{2} \sum_{i,j=1}^N \frac{\partial^2 (a_{ij} \bar{\rho})}{\partial x_i \partial x_j},$$

Controllability of Fokker-Planck - Linear-Gaussian

$$dx(t) = Ax(t)dt + Bu(t)dt + B_1dw(t)$$

with $x(0) = x_0$ a.s.

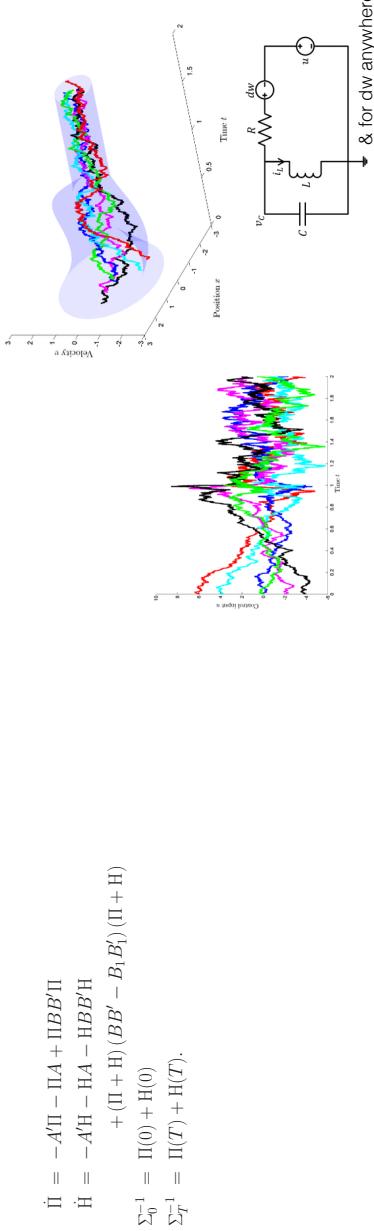
Thm: (A,B) controllable is sufficient to steer the system from any initial Gaussian distribution to a final one at $t=T$.

Thm: A Gaussian state-pdf can be "sustained" with constant state-feedback iff the state covariance satisfies

$$(A - BK)\Sigma + \Sigma(A' - K'B') + B_1B_1' = 0$$

$$\text{equivalently, } \text{rank} \begin{bmatrix} A\Sigma + \Sigma A' + B_1B_1' & B \\ B & 0 \end{bmatrix} = \text{rank} \begin{bmatrix} 0 & B \\ B & 0 \end{bmatrix}$$

Schrödinger system



Compare with conditions for:

- i) steering the system to a given state - controllability
- ii) steering within the positive cone?
- iii) maintaining the state at a given value

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$0 = (A - BK)\xi + Bu$$

Open problem

Density matrices: e.g.

$$\mathfrak{D} = \{\rho \geq 0 \mid \text{symmetric } \rho \in \mathbb{R}^{n \times n} \text{ with } \text{trace}(\rho) = 1\}$$

E_i with $i = 1, \dots, n_{\mathcal{E}}$ and $\sum_{i=1}^{n_{\mathcal{E}}} E_i^\dagger E_i = I$
(typically $n_{\mathcal{E}} \sim n^2$)

for "positivity-improving": $\rho \geq 0 \Rightarrow \mathcal{E}(\rho) > 0$

$$\text{TPTP: } \mathcal{E} : \mathfrak{D} \rightarrow \mathfrak{D} : \rho \longrightarrow \sigma = \sum_{i=1}^{n_{\mathcal{E}}} E_i \rho E_i^\dagger$$

Thank you for your attention

Problem: Prove that the iteration:

$$\begin{aligned} \mathcal{E} : \hat{\phi}_0 &\mapsto \hat{\phi}_{\mathcal{E}} = \mathcal{E}(\hat{\phi}_0) \\ D_T : \hat{\phi}_T &\mapsto \hat{\phi}_{T'} = \left(\frac{1}{\rho_T^{1/2}} \left(\rho_T^{-1/2} \hat{\phi}_T \rho_T^{-1/2} \right)^{1/2} \rho_T^{1/2} \right)^2 \end{aligned}$$

$$\mathcal{E}^1 : \hat{\phi}_T \mapsto \hat{\phi}_0 = \mathcal{E}^1(\hat{\phi}_T)$$

$$D_0 : \hat{\phi}_0 \mapsto \hat{\phi}_0 = (\hat{\phi}_0)^{1/2} \rho(\hat{\phi}_0) \hat{\phi}_0^{1/2}$$

has an attractive fixed point.

Software for numerical experimentation

http://www.eeh.van.esch.org/papers/schroedinger_beta/

$$\begin{aligned} \text{http://arxiv.org/abs/1405.6650} \\ \text{Positive contraction mappings for classical and quantum Schrödinger systems} \end{aligned}$$

$$\begin{aligned} \text{http://arxiv.org/abs/1407.3421} \\ \text{Stochastic bridges of linear systems} \end{aligned}$$

$$\begin{aligned} \text{http://arxiv.org/abs/1410.1605} \\ \text{Optimal steering of inertial particles diffusing anisotropically with losses} \end{aligned}$$

$$\begin{aligned} \text{arxiv.org/abs/1408.2222} \\ \text{Optimal steering of a linear stochastic system to a final probability distribution} \\ \text{arxiv.org/abs/1410.3447} \\ \text{Optimal steering of a linear stochastic system to a final probability distribution, Part II} \end{aligned}$$

**RANDOMIZED AVERAGING ALGORITHMS, WHEN CAN
ERRORS AND UNRELIABILITIES JUST BE IGNORED?****Julien Hendrickx, Université Catholique de Louvain,
Belgium**

We consider randomized discrete-time consensus systems that preserve the average "on average", and provide a new upper bound on the mean square deviation of the final consensus value from the initial average.

We show that a certain asymptotic accuracy can be guaranteed when there are few simultaneous interactions or when the simultaneous interactions are sufficiently uncorrelated. Our results are easily applicable to many classes of systems, and we particularize them to various algorithms having been proposed in the literature, obtaining bounds that match or outperform all previously available ones..



(Linear) Averaging

Agents have values $x_i(t)$

Randomized averaging algorithms:
when can errors & unreliability just be ignored?

Paolo Frasca & Julien Hendrickx

Lund – LCCC workshop Oct 2014

Averaging iteration:

$$x_i(t+1) = \sum_j a_{ij}(t) x_j(t)$$

$$\begin{array}{l} a_{ij}(t) \geq 0 \\ \sum_j a_{ij}(t) = 1 \end{array}$$

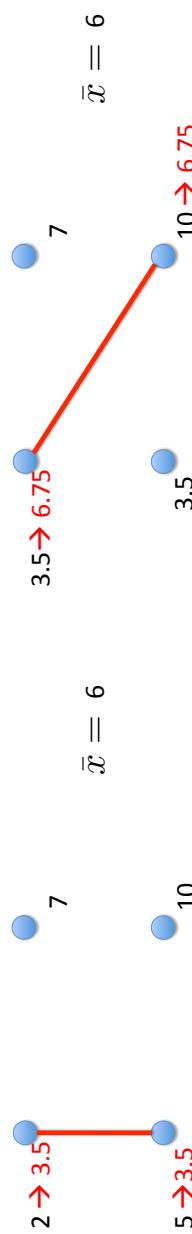
If average preserving (ex: if $a_{ij} = a_{ji}$) and enough interactions

$$x_i \rightarrow \bar{x} = \frac{1}{n} \sum x_i(0)$$

1 Basis for many decentralized algorithms

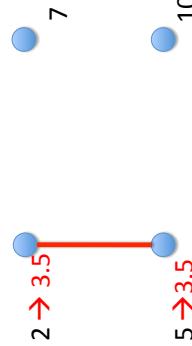
2

Ex: Gossip



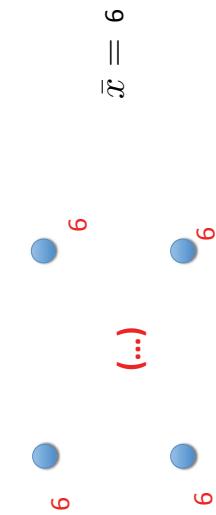
3

Ex: Gossip



4

Ex: Gossip



Convergence to the average

What if transmissions are asymmetric/uncertain?

5

Communication requirements for averaging

- Simple consensus:

$$x(t+1) = A(t)x(t)$$

$A(t)$ stochastic
Average preserved if $A(t)$ doubly stochastic. $1^T A(t) = 1^T$. Needs
- **symmetric / balanced** interactions
- therefore **synchronous** interactions
- Push sum [Kempe et al 03], based on mass preservation
Directed asynchronous communications OK
But needs **reliability**: agents must know
 - if message received,
 - how many agents receive message

6

Unpredictability and Unreliability

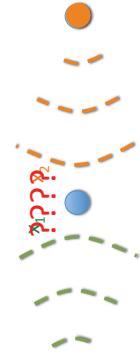
- Directed & uncertain communication
- Packet loss
- Broadcasts at random times
- Collisions due to interferences



7

Unpredictability and Unreliability

- Directed & uncertain communication
- Packet loss
- Broadcasts at random times
- Collisions due to interferences



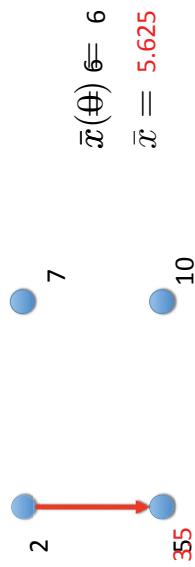
8

Unpredictability and Unreliability: 2 approaches

- 1) Verification and correction mechanisms
idea: errors may create big trouble

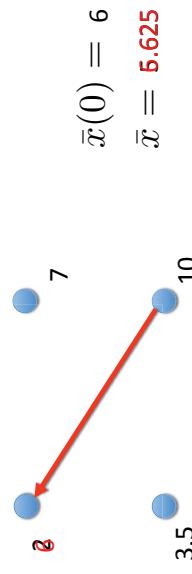
2) Ignore issues, assume average behavior and hope for the best
Idea: errors will cancel out

Naïve extension: Asymmetric Gossip



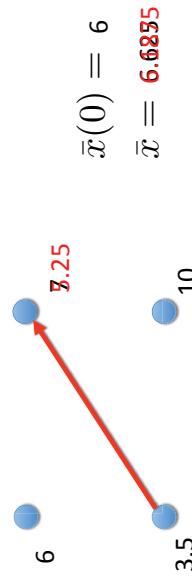
9

Naïve extension: Asymmetric Gossip



10

Naïve extension: Asymmetric Gossip



10

Does not converge to average, but seems close...

Guarantee?

11

12

Outline

- Introduction
 - **Formulation and previous results**
 - Formulation
 - Exact approach
 - Convergence-based approach
 - Our approach
 - Particularization and Applications
 - Conclusions

13

Problem formulation

$$x(t+1) = A(t)x(t)$$

where $A(t) \in \mathbb{R}_+^{n \times n}$

- Averaging matrices ($A(t)\mathbf{1} = \mathbf{1}$)
- **i.i.d. random variables**

- Our approach
- Particularization and Applications
- Conclusions

14

Problem formulation

$$x(t+1) = A(t)x(t)$$

where $A(t) \in \mathbb{R}_+^{n \times n}$

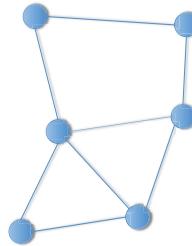
- Averaging matrices ($A(t)\mathbf{1} = \mathbf{1}$)
- **i.i.d. random variables**
- preserve average on expectation ($\mathbf{1}^T E A(t) = \mathbf{1}^T$)

$$\Rightarrow E\bar{x}(t) = \bar{x}(0)$$

Convergence to "consensus" w.p. 1 under conditions
 $x(t) \rightarrow \mathbf{1}x_\infty$ with $E x_\infty = \bar{x}(0)$

Goal: bound **standard asymptotic error** $E(x_\infty - \bar{x}(0))^2$

Examples: nominal network



1. Synchronous symmetric updates,

$$x_i(t+1) = x_i(t) + \sum_j a_{ij}(x_j(t) - x_i(t))$$

but some messages are lost

2. Nodes wake up and update at random times

3. Nodes wake up, send messages to neighbor(s) at random time

15

Exact approach

Standard asymptotic error computed exactly using $x(0)$ and V : Standard asymptotic error computed exactly using $x(0)$ and V :

$$E(A^T V A) = V \quad 1^T V 1 = 1 \quad E(A^T V A) = V \quad 1^T V 1 = 1$$

But, very hard to obtain expression for V

→ Theoretical computation only for specific cases (involved!) → Theoretical computation only for specific cases (involved!)

Ex: [Fagnani & Zampieri, SICON 2009]

Fixed averaging algorithm with packet losses
deployed on Cayley graphs of Abelian groups

But, very hard to obtain expression for V

→ Theoretical computation only for specific cases (involved!) → Numerical verification for given system (hard)

- Numerical issues
- May need enumeration of all A
often many (one for each possible combination of updates)

Exact approach

Standard asymptotic error computed exactly using $x(0)$ and V : Standard asymptotic error computed exactly using $x(0)$ and V :

$$E(A^T V A) = V \quad 1^T V 1 = 1 \quad E(A^T V A) = V \quad 1^T V 1 = 1$$

But, very hard to obtain expression for V

→ Theoretical computation only for specific cases (involved!) → Theoretical computation only for specific cases (involved!)

Ex: [Fagnani & Zampieri, SICON 2009]

Fixed averaging algorithm with packet losses
deployed on Cayley graphs of Abelian groups

But, very hard to obtain expression for V

→ Theoretical computation only for specific cases (involved!) → Numerical verification for given system (hard)

- Numerical issues
- May need enumeration of all A
often many (one for each possible combination of updates)

17

Convergence based approach

$$Var(t) \text{ Individual variance } E \frac{1}{n} \sum (x_i(t) - \bar{x}(t))^2$$

$$err(t) \text{ Expected group error } E(\bar{x}(t) - \bar{x}(0))^2$$

$$\textbf{Idea} \quad \Delta err(t) \leq \frac{K}{n} Var(t)$$



Convergence based approach

$$err(\infty) \leq O\left(\frac{K}{n\lambda}\right) Var(0)$$

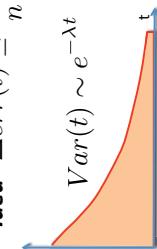
λ linked to algebraic connectivity of "network"

Conclusion: keep λ large when n grows → expanders, etc.

But, expanders not easy to build physically (range constraints, etc.)

and λ often not critical in experimental results

$$err(\infty) \leq O\left(\frac{K}{n\lambda}\right) Var(0)$$



λ linked to algebraic connectivity of "network"
Ex: [Fagnani & Frasca 2011]

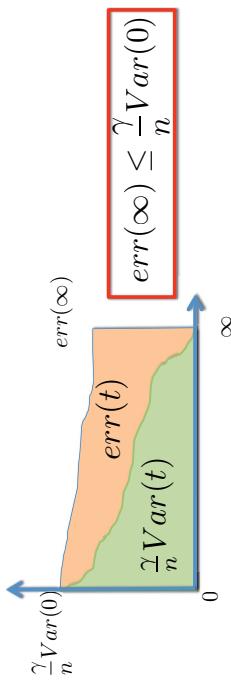
20

Outline

- Introduction
- Formulation and previous results
- **Our approach**
- Particularization and Applications
- Conclusion and further works

Our approach: conservation

- Find γ s.t. $err(t) + \frac{\gamma}{n} Var(t)$ non-increasing



- If γ independent of n (or $O(n)$), “**accuracy**” for large n
- Not always the case, exist systems with large errors

2.1

How to pick γ ?

$$err(t) + \frac{\gamma}{n} Var(t) \text{ non-increasing}$$

$$\gamma \text{ valid if } E(L^T \mathbf{1} \mathbf{1}^T L) \leq \gamma E(L + L^T - L^T L)$$

$$\text{with } L(t) = I - A(t)$$

Appears involved, but

- Generic γ can be found for large classes of systems
- Results very easy to use
- Often leads to strong bounds

2.3

Outline

- Introduction
- Formulation and previous results
- Our approach
- Particularization and Applications
- Conclusions

Particularization and Applications

1. **Limited updates**
2. **Uncorrelated updates**
3. **Uncorrelated broadcasts**
4. **Limited correlations**

2.4

1. Limited updates

$$\begin{aligned} A_{\max} &\geq \sum_i \sum_{j \neq i} a_{ij}(t) && \text{Max total simultaneous updates} \\ a_{ii}^{\min} &\leq a_{ii}(t) && \text{Min. self confidence} \end{aligned}$$

Then $\gamma = \frac{A_{\max}}{a_{ii}^{\min}}$

$$\text{err}(\infty) \leq \frac{1}{n} \left(\frac{A_{\max}}{a_{ii}^{\min}} \right) V(0)$$

"Accuracy" if self confidence + amount of updates $< \leq_5 n$

Asymmetric gossip

$$A_{\max} \geq \sum_i \sum_{j \neq i} a_{ij}(t) \rightarrow \text{err}(\infty) \leq \frac{1}{n} \left(\frac{A_{\max}}{a_{ii}^{\min}} \right) V(0)$$

Iteration: At each t, **1 random node** i chooses * j and updates to

$$x_i(t) + q(x_j(t) - x_i(t))$$

$$\begin{aligned} A_{\max} &= q && \rightarrow \\ a_{ii}^{\min} &= 1 - q && \text{With symmetric probabilities} \end{aligned}$$

Accuracy for large n. Bound similar to [Fagnani & Zampieri 08]

²⁵*

Broadcast gossip

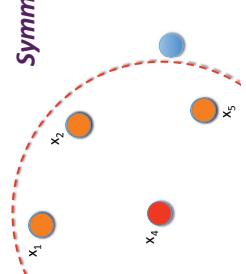
- At each t, a node j wakes up with uniform probability
- j transmits x_i to all within listening range (assumed symmetric)
- update $x_i(t+1) = x_i(t) + q(x_j(t) - x_i(t))$

$$q \in (0, 1)$$

Previous results

- Involved, rely on spectral quantities related to network defined by listening ranges
- No general guarantee of accuracy for large n

Symmetry $\rightarrow E(\bar{x}(t)) = \bar{x}(0)$



$$\text{Ex: } \text{err}(\infty) \leq O\left(\frac{d_{\max}^2}{\lambda_1 n}\right) V(0)$$

d_{\max} : Largest degree
 λ_1 : Smallest nonzero eigenvalue laplacian

[Aysoy et al. 2009], [Fagnani & Delvenne 2010], [Fagnani and P. Frasca. 2012a,b]

Broadcast gossip: our approach

- At each t, a node j wakes up with uniform probability
- j transmits x_j to all within listening range (assumed symmetric)
- update $x_i(t+1) = x_i(t) + q(x_j(t) - x_i(t))$

Limited updates rule:

$$\begin{aligned} A_{\max} &= qd_{\max} \\ a_{ii}^{\min} &= 1 - q \end{aligned} \quad \rightarrow \quad err(\infty) \leq \frac{d_{\max}}{n} \left(\frac{q}{1-q} \right) V(0)$$

- Accuracy for large n if $d_{\max} < o(n)$
- Independent of graph spectrum

20

2. Independent updates

- At each t, a node j wakes up with uniform probability
- j transmits x_j to all within listening range (assumed symmetric)
- update $x_i(t+1) = x_i(t) + q(x_j(t) - x_i(t))$

If nodes make independent decisions about updates and weights
 a_{ij}, a_{kl} Uncorrelated if $i \neq k$ Then

$$\gamma = \frac{1 - a_{ii}^{\min}}{a_{ii}^{\min}}$$

- Always accurate when $n \rightarrow \infty$ if a_{ii}^{\min} bounded

30

Synchronous Asymmetric Gossip

Iteration: Every node i chooses neighbor j_i and updates to

$$x_i(t+1) = (1 - q)x_i(t) + qx_{j_i}(t)$$

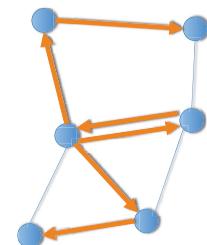
Probabilities s.t. $E(\bar{x}(t)) = \bar{x}(0)$

Synchronous Asymmetric Gossip

Iteration: Every node i chooses neighbor j_i and updates to

$$x_i(t+1) = (1 - q)x_i(t) + qx_{j_i}(t)$$

Probabilities s.t. $E(\bar{x}(t)) = \bar{x}(0)$



Previous results

- Related to eigenvalues of graph Laplacian
- Accuracy for particular cases

[Fagnani & Zampieri, 2008]

31

32

Synchronous Asymmetric Gossip

Iteration: Every node i chooses neighbor j_i and updates to

$$x_i(t+1) = (1-q)x_i(t) + qx_{j_i}(t)$$

Probabilities s.t. $E(\bar{x}(t)) = \bar{x}(0)$

$$\boxed{\text{Indep updates: } err(\infty) \leq \frac{\gamma}{n} Var(0) \quad \gamma = \frac{1 - a_{ii}^{\min}}{a_{ii}^{\min}}}$$

If a_{ij} correlated to more than $4C$ other coefficients

$$\boxed{err(\infty) \leq \frac{1}{n} \left(\frac{q}{1-q} \right) V(0)}$$

For all graphs, independently of spectral quantities

33

3. Independent broadcasts

$$err(\infty) \leq \frac{\gamma}{n} Var(0)$$

• 2. Independent decisions about updates and weights

$$\gamma = \frac{1 - a_{ii}^{\min}}{a_{ii}^{\min}}$$

• 3. Independent broadcast decisions (i.e. columns in A)

$$\boxed{\gamma = \frac{\substack{\text{Max importance} \\ \text{of a node}}}{\substack{a_{col}^{\max} \\ a_{ij}^{\min}}}} \quad a_{col}^{\max} \geq \sum_{i:j \neq j} a_{ij}(t)$$

Minimal self confidence

34

4. General correlation rule

$$err(\infty) \leq \frac{\gamma}{n} Var(0)$$

If no a_{ij} correlated to more than $4C$ other coefficients

$$\boxed{\gamma = \frac{C}{a_{ii}^{\min}}}$$

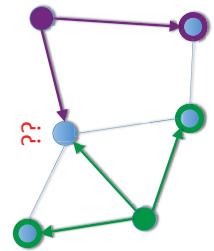
Broadcast with collisions

At every time:

- Some nodes awake and broadcast their values
- Nodes receiving one value update
- Nodes receiving two or more values do not

Weights and probabilities
s. t. $E(\bar{x}(t)) = \bar{x}(0)$

- Challenges:**
- Multiple updates
 - Multiple correlations



35

36

Broadcast with collisions

At every time:

- Some nodes awake and broadcast their values
- Nodes receiving one value update
- Nodes receiving two or more values do not

Previous results

- Results on Cayley graphs of Abelian Groups

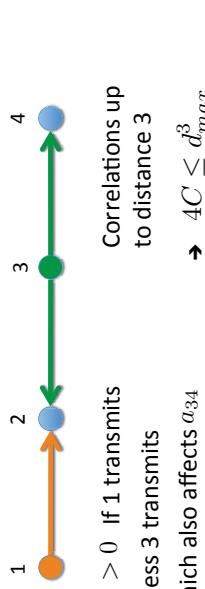
- Accuracy for certain sparse networks when n grows
[Fagnani & Frasca, 2011]

37

Broadcast with collisions

$$\gamma = \frac{C}{a_{ii}^{\min}} \quad \begin{array}{l} \text{if } \text{no } a_{ij} \text{ correlated to more} \\ \text{than } 4C \text{ other coefficients} \end{array}$$

How many correlations?



38

Broadcast with collisions

$$\gamma = \frac{C}{a_{ii}^{\min}} \quad \begin{array}{l} \text{if } \text{no } a_{ij} \text{ correlated to more} \\ \text{than } 4C \text{ other coefficients} \end{array}$$

Correlations up to distance 3 $4C \leq d_{max}^3$

$$err(\infty) \leq \frac{d_{\max}^3}{n} \frac{1}{a_{ii}^{\min}} V(0)$$

Accuracy for large n if $d_{max} \leq o(n^{1/3})$

Outline

- Introduction
- Formulation and previous results
- Our approach
- Particularization and Applications
- **Conclusions**

39

40

Summary

- $\frac{\gamma}{n}$ group error + individual var. non-increasing*
 - γ often easy to compute
 - Bound independent of convergence speed and spectral properties; only local properties
 - Prove asymptotical accuracy of many schemes
 - Equals or outperforms (almost) all previous ad hoc results
- * actually $\frac{\gamma}{n+\gamma}$

Summary

$$\text{err}(\infty) \leq \frac{\text{Largest "correlated event"} - \text{Min "self confidence"}}{n} \frac{\text{Var}(0)}{n}$$

- Ignoring problems and hoping for the best is OK if
 - Limited correlations (*often OK in multi-agent systems*)
 - Sufficient self-confidence for agents
- Robustness of large class algorithms with respect to important fluctuations

42

Possible developments

- Expected average not entirely preserved
- Correlation between different times
- More detail → Less conservative bounds
- Other algorithms

Thank you for your attention!

References

Paolo Frasca and J.-H. On the mean square error of randomized averaging algorithms, *Automatica* 2013 arXiv:1111.4572v1

Paolo Frasca and J.-H. Large network consensus is robust to packet losses and interferences, Proceedings of ECC 2013, Zurich, July 2013

43

44

SYSTEMIC RISK IN FINANCE

Munther A. Dahleh, Massachusetts Institute of Technology, USA

We study welfare of competitive equilibria in an economy with banks runs and costly fire sales, formalizing the importance of "animal spirits" for financial stability. These forces are not the product of irrationality, but a consequence of uncertainty of a self-fulfilling nature. Short-term debt issued by banks is safe, and thus functions as private money. Feedback between financial constraints and market prices results in a systemic externality not internalized by banks in the process of money creation: during credit booms aggregate short-term debt is excessive; during recessions it is insufficient. With a multi-stage game formulation we show that there are exactly two equilibria corresponding to a Boom-Bust scenarios.

Work is co-authored with Diego Fejer.



Systemic Risk in Financial Systems

Diego Feijer and Munther Dahleh
LIDS, MIT

Cascades and Systemic Risk

Systemic Risk is a term used to describe fragility in interconnected systems that result in cascades of failures due to either relatively small shocks at the subsystem level or larger and more malicious types of disruptions affecting the whole system.



2

Market Expectations and Economic Activity

- Aggregate expectations play a key role in credit markets and macroeconomic activity
- Optimism results in **credit booms**; pessimism leads to **credit crunches**, and potential recession
 - Most recent financial crisis
- **Fluctuations** in market **sentiments** often seem to happen without apparent reason; they seem to be driven by **animal spirits**

Animal Spirits and Regulation

"History—including recent history—shows that without regulation, animal spirits will drive the economic activity to extremes."

G. Akerlof and R. Shiller

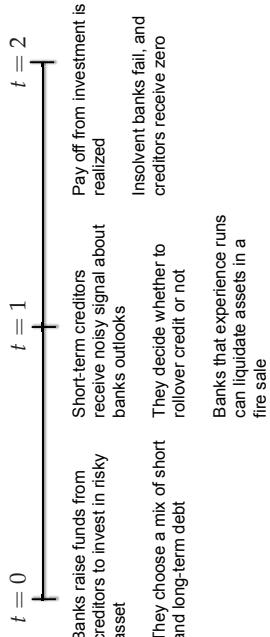
"Good Government and Animal Spirits." WSJ, Apr. 24, 2009

Related Literature

- Dominant theory to justify regulation is based on **fire sales** and **pecuniary externalities**.
 - Occasionally binding **collateral constraint**
 - **Over-borrowing** during credit booms
Caballero & Krishnamurthy'03; Lorenzonit'08; Bianchi'11; Stein'12
 - **Silent about credit crunches**

- In this model:
 - Collateral constraint + (binding) **credit risk constraint**
 - **Endogenous** state of the economy: **booms** and **crunches**

Preview of the Model



Households

- Initial endowment of consumption good
- Consume endowment at date 0, or invest in financial assets (lend to banks) and consume proceeds at date 2
- **Two financial assets:** $(m, 1 - m)$
 - Risky “bonds” with gross real return R_B (lend long-term)
 - **Riskless “money”** with gross real return R_M (lend short-term)
- Linear preferences over consumption:

$$U(\{C_t\}) = C_0 + \beta E[C_2] + \gamma m R_M$$

with $\beta + \gamma < 1$
- Derive utility from “monetary” services provided by **safe assets**

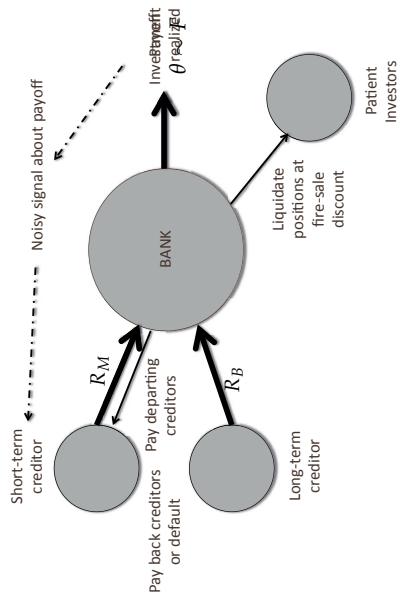
Banks

- Continuum of identical banks, with total mass one
- **Risky investment opportunity** at date 0, with real return $\theta \sim F$ with mean $\bar{\theta} > \frac{1}{\beta}$ and support over $[\theta_{\min}, \theta_{\max}]$
- Banks need 1 unit of consumption good to undertake investment. They raise money (short-term debt)
 - and finance issuing bonds (long-term debt): $(m, 1 - m)$
- Short-term debt claims need to be **rolled over** at date 1
 - To meet redemption demands, banks can sell any fraction of their assets in a **secondary market**.

Fire Sales and Patient Investors

- Sales yield $k\theta$ per unit liquidated. **Fire-sale discount** $k \in [0, 1]$
- Collateral constraint:** Short-term debt has to be **riskless**. If a bank issues m units of short-term finance, then
$$mR_M \leq k\theta_{\min}$$
- Patient investors receive **fixed endowment** $I > \frac{1}{\beta + \gamma}$ at **date 1**
- Investment technology: $g(\cdot)$ increasing and strictly concave
- No Crisis:** they invest all resources in new, late arriving projects that yield total output of $g(I)$
- Crisis:** they provide liquidity to banks purchasing assets at fire-sale discount, and invest the rest of resources in projects

Recap



Coordination Problem

- A bank **fails** at date 2 when unable to honor its liabilities
- Assumption (Zero recovery value).** In case of failure, **all creditors** holding debt claims at date 2 receive a payoff of 0
- Suppose the capital structure of a bank is $(m, 1 - m)$
- At date 1, θ is realized and its value revealed to banks and patient investors. Each creditor j receives a **noisy private signal**: $\theta_j = \theta + \epsilon_j$ with $\epsilon_j \sim \mathcal{U}[-\epsilon, \epsilon]$ iid
- If a fraction λ of creditors decide to withdraw, the bank needs to liquidate a fraction q : $q\theta = \lambda mR_M$. The bank fails if:

$$(1 - q)\theta < (1 - \lambda)mR_M + (1 - m)R_B$$

Defines: $\theta_{\text{run}}(\lambda)$

Panic Runs

- The bank fails if

$$\theta < \theta_{\text{run}}(\lambda) \equiv mR_M + (1 - m)R_B + \lambda mR_M \left(\frac{1}{k} - 1 \right)$$

Lower dominance region $[\theta_{\min}, \theta_{\text{run}}(0))$

Upper dominance region $(\theta_{\text{run}}(1), \theta_{\max}]$

- Switching strategy.** In the limit as $\epsilon \rightarrow 0$, the coordination game has a **unique (symmetric) equilibrium** in which agents run whenever their signal is below the threshold:

$$\theta^* = mR_M + (1 - m)R_B + mR_M \left(\frac{1}{k} - 1 \right)$$

Insolvency threshold

Self-fulfilling illiquidity fears may drive a solvent bank to failure

Prices

- **Credit Market:** households indifference condition

$$R_M = \frac{1}{\beta + \gamma} < R_B = \frac{1}{\beta(1 - F(\theta^*))}$$

Cheaper because of convenience yield

- **Fire Sales:** patient investors indifference condition; investment activities should yield same marginal returns

$$\frac{1}{k} = g'(I - mR_M)$$

Captures how funding decisions at the individual level affect the whole system

Sentiments

- **Assumption (Looting).** The owner of a bankrupt bank steals any remaining assets for personal consumption

- Taking k as given, banks choose m to maximize profits:

$$\Pi(m; k) = \left(\bar{\theta} - \frac{1}{\beta} \right) + m \left(\frac{1}{\beta} - R_M \right) - F(\theta^*)mR_M \left(\frac{1}{k} - 1 \right)$$

กิตติมศักดิ์ ภูมิพล อดุลยเดช

- **First term:** expected profits if only long-term debt
 - **Second term:** gains from moving to cheaper short-term debt
 - **Third term:** cost (risk) that comes along with short-term debt

Private Money Creation

- **Assumption (Looting).** The owner of a bankrupt bank steals any remaining assets for personal consumption

- Taking k as given, banks choose m to maximize profits:

$$\Pi(m; k) = \left(\bar{\theta} - \frac{1}{\beta} \right) + m \left(\frac{1}{\beta} - R_M \right) - F(\theta^*)mR_M \left(\frac{1}{k} - 1 \right)$$

กิตติมศักดิ์ ภูมิพล อดุลยเดช

- **Fire Sales:** patient investors indifference condition; investment activities should yield same marginal returns

$$\theta^* = \frac{\frac{m_n M}{k}}{1 - n} + \frac{1 - n}{\beta(1 - F(\theta^*))}$$

Captures how running decisions at the individual level affect the whole system

Competitive Equilibrium

- Definition: ACE is a pair (m^*, k^*) such that:

$\theta^* = \frac{mR_M}{k} + \frac{1-m}{\beta(1-F(\theta^*))}$ has two solutions

m=0 / / 0≤m≤1

- m^* maximizes:

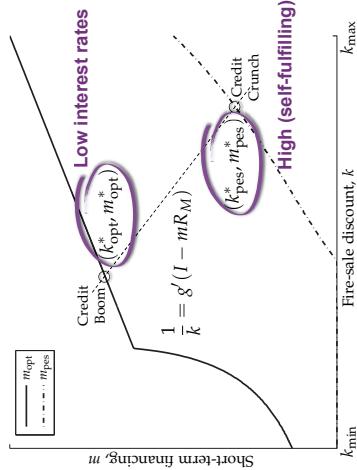
$$\left(\bar{\theta} - \frac{1}{\bar{\beta}}\right) + m\left(\frac{1}{\bar{\beta}} - R_M\right) = F(\theta_s^*)mR_M\left(\frac{1}{\bar{\beta}} - 1\right)$$

Self-fulfilling

subject [s: $s \in \{\text{opt, pes}\}$] (sentiment)

- and: $\frac{1}{l^*} = g'(I - m^* R_M)$

Credit Booms and Crunches



Fluctuations in market sentiment are driven by **animal spirits**

Welfare

- Proceeds from investments by banks and patient investors are rebated back to households

- Welfare at date 2 is $W(m)$

$$\begin{aligned} \frac{1}{\bar{\theta}} + m \left(\frac{1}{\bar{\beta}} - R_M \right) \\ + \left\{ (1 - F(\theta^*))g(I) + F(\theta^*)(g(I - mR_M) + mR_M) - \frac{I}{\bar{\beta}} \right\} \end{aligned}$$

- First term: net expected return to investment by banks
- Second term: monetary services
- Third term: net expected return to investment by patient investors

Welfare

- Proceeds from investments by banks and patient investors are rebated back to households

Planning Problem

- The planner solves:

$$\max_{m \in [0,1]} W(m) \text{ subject to:}$$

$mR_M \leq k \theta_{\min}$
 $\theta_s^* = \theta_s^*(m, k)$
 $\frac{1}{k} = g'(I - mR_M)$

↑ (collateral)
s ∈ {opt, pes} (sentiment)
fire sale discount

$$\begin{aligned} 1 - F(\theta^*) \\ F(\theta^*) \end{aligned}$$

- Consumption at date 2:

$$\begin{aligned} E[\theta | \theta > \theta^*] + g(I) \\ E[(1 - q)\theta | \theta < \theta^*] \\ + g(I - mR_M) + mR_M + E[q|\theta | \theta < \theta^*] \end{aligned}$$

- Welfare at date 2 is $W(m)$

$$\begin{aligned} \frac{1}{\bar{\theta}} + m \left(\frac{1}{\bar{\beta}} - R_M \right) \\ + \left\{ (1 - F(\theta^*))g(I) + F(\theta^*)(g(I - mR_M) + mR_M) - \frac{I}{\bar{\beta}} \right\} \end{aligned}$$

Inefficiency

- Theorem. CE, in general, not constrained efficient.

- Wedge between private and social solutions:

$$\tau_s^* = -\mu_s^P \frac{\theta_{\min}}{R_M} \frac{dk}{dm} + C(m) \frac{dF(\theta_s^*)}{dm} - mR_M \left(\frac{1}{k} - 1 \right) \frac{\partial F(\theta_s^*)}{\partial m}$$

$$\text{with } C(m) \equiv \left(g(I) - g(I - mR_M) - mR_M \right) \geq 0$$

- $\tau_{\text{opt}}^* \geq 0 \iff$ excessive money creation
- $\tau_{\text{pes}}^* \leq 0 \iff$ insufficient money creation

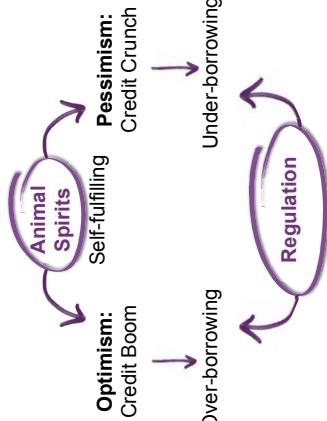
- Efficiency-restoring tax:

$$\max_m W(m) = \max_m \{ \Pi_s(m; k^*) - \tau_s^* m \}$$

Key Results

Sign of restoring tax depends on the collective sentiment

in the credit market:



Concluding Remarks

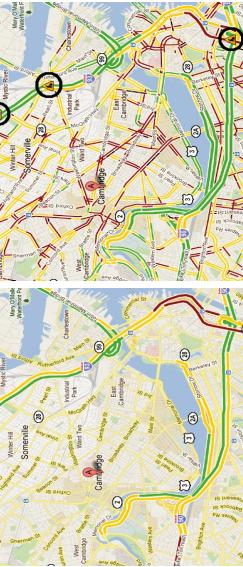
- The economy benefits from private money creation, but banks incentives are often distorted because they don't internalize cost of runs and fire sales
- State of the economy is endogenous: interaction between insolvency risk and rollover risk. **Multiple equilibria**
- Unified setting (boom/crunch) to justify interventions; direction depends on the sentiment in the credit market
- Highlights the need for dynamic regulatory frameworks
- Restoring tax/subsidy doesn't prevent ups/downs, but prevents the economy from swinging to extremes

questions

ON DISTRIBUTED CONTROL OF DYNAMICAL FLOW**NETWORKS****Giacomo Como, Lund University, Sweden**

This talk focuses on distributed control of dynamical flow networks. First, we show that optimal throughput and resilience can be achieved by feedback policies that depend only on local information and require no global knowledge of the network. Then, we prove how the optimal selection of a stable equilibrium and the optimal control of the transient can be cast as convex problems which are amenable to distributed solutions. Applications to arterial traffic control are discussed.





Large-scale infrastructure networks

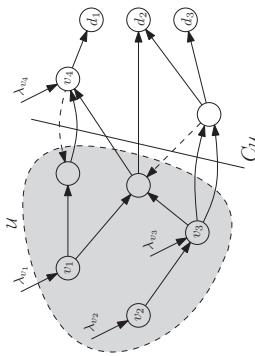
Resilient distributed control of dynamical flow networks

Giacomo Como
Department of Automatic Control
Lund University

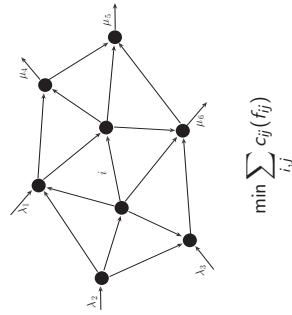
LCCC Workshop on Dynamics and Control in Networks
October 17, 2014

- typical Monday at 18:30
- Monday July 11, at 18:30
- good in business as usual, prone to disruptions
- cascade effects
- \Rightarrow network vulnerability $\gg \sum$ component vulnerabilities

Max-flow min-cut theorem ('56)



Optimal network flow



$$\exists \text{ feasible equilibrium flow} \iff \min_U \{C_U - \lambda_U\} > 0$$

$$\lambda_i + \sum_j f_{ij} - \sum_j f_{ji} - \mu_i = 0$$

- static, centralized, global information
- static, convex, rich duality theory

Wardrop equilibrium ('56)

Lighthill-Whitham-Richards traffic model ('55)

$$\frac{\partial}{\partial t} \rho + \frac{\partial}{\partial x} \Phi(\rho) = 0$$



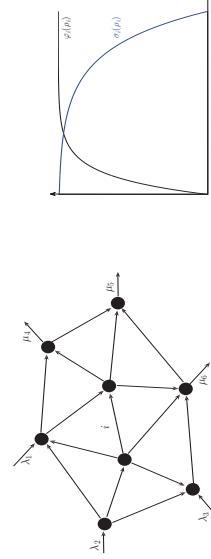
► 'the journey times in all routes actually used are equal and less than those which would be experienced on any unused route'

$f^W \Leftrightarrow$ Nash equilibrium of congestion game

► user optimum vs social optimum: price of anarchy

► using tolls on links allows one to align f^W with any desired f^*

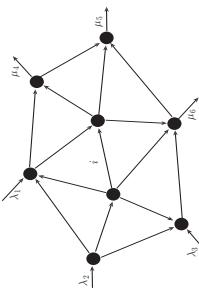
Dynamical flow networks



$$\dot{\rho}_i = \lambda_i + \sum_j f_{ji} - \sum_j f_{ij} - \mu_i$$

$$\mu_i + \sum_j f_{ij} \leq \varphi_i(\rho_i)$$

$$\sum_j f_{ji} \leq \sigma_i(\rho_i)$$



$$\dot{\rho}_i = \lambda_i + \sum_j f_{ji} - \sum_j f_{ij} - \mu_i$$

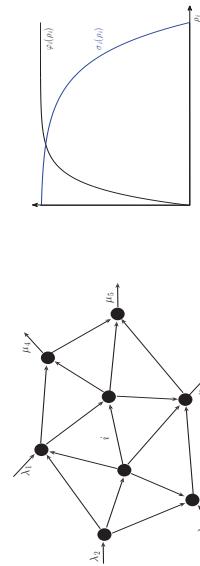
$$\mu_i + \sum_j f_{ij} \leq \varphi_i(\rho_i)$$

$$\sum_j f_{ji} \leq \sigma_i(\rho_i)$$

► Daganzo '92, Gomes&Horowitz '06, Pisařský&CanudasDeWit '12, Varaiya '08, Coogan&Arcak '14, Karafyllis&Papageorgiou '14, ...

Dynamical flow networks

Resilient decentralized routing in dynamical flow networks

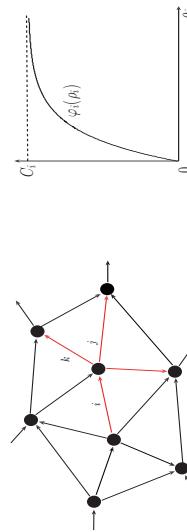


$$\dot{\rho}_i = \lambda_i + \sum_j f_{ij} - \sum_j f_{ji} - \mu_i$$

$$\mu_i + \sum_j f_{ij} \leq \varphi_i(\rho_i)$$

► goal: scalable design of f_{ij} with good performance and resilience

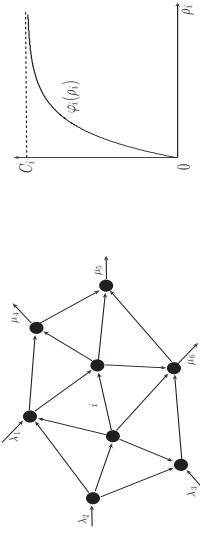
Decentralized routing



$$\dot{\rho}_i = \lambda_i + \sum_j f_{ij} - \sum_j f_{ji} - \mu_i$$

flow f_{ij} depends only on **local information** ρ^j

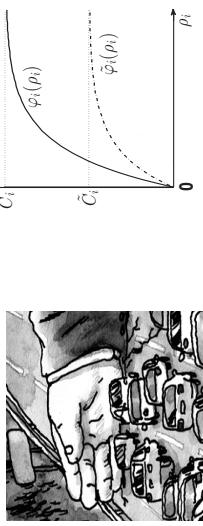
$$\begin{aligned} \nearrow f_{ij} &= \varphi_i(\rho_i) & R_{ij}(\rho) \\ \text{flow from } i \text{ to } j &\quad \text{demand on } i & \text{fraction routed to } j \end{aligned}$$



$$\dot{\rho}_i = \lambda_i + \sum_j f_{ij} - \sum_j f_{ji} - \mu_i$$

$$\mu_i + \sum_j f_{ij} \leq \varphi_i(\rho_i)$$

Resilience



$$\delta := \sum_i |\tilde{\lambda}_i - \lambda_i| + \sum_i |\tilde{C}_i - C_i|$$

► magnitude of perturbation

$$\gamma := \liminf_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \sum_i (\tilde{\lambda}_i(s) - \lambda_i(s)) ds$$

$$\begin{aligned} \dot{\rho}_i &= \lambda_i + \sum_j \tilde{f}_{ij} - \sum_j \tilde{f}_{ji} - \tilde{\mu}_i \\ \tilde{f}_{ij} &= \varphi_i(\rho_i) R_{ij}(\rho) \end{aligned}$$

Resilience



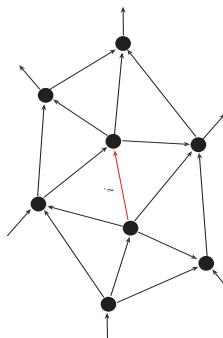
margin of resilience := $\inf \{\delta : \gamma > 0\}$

- ▶ δ := magnitude of perturbation
- ▶ γ := throughput loss

margin of resilience $\nu := \inf \{\delta : \gamma > 0\}$

- ▶ necessarily: $\nu \leq \min_{U} \{C_U - \lambda_U\}$
- ▶ Problem: max resilience ν over decentralized routing policies R

Resilience with fixed routing

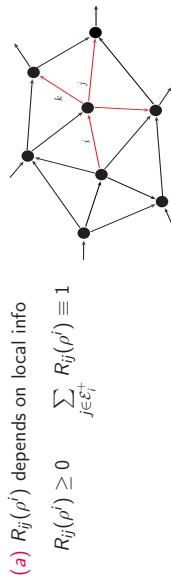


- ▶ $f_{ij} = R_{ij}\varphi_i(\rho_i)$ with constant R_{ij}
- ▶ start from equilibrium $f_i^* = \varphi_i(\rho_i^*)$

$$\nu = \underbrace{\min_i \{C_i - f_i^*\}}_{\text{link residual capacity}}$$

link residual capacity

Resilience with locally responsive routing



$$(a) R_{ij}(\rho^j) \text{ depends on local info}$$

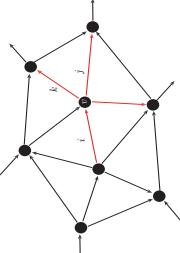
$$R_{ij}(\rho^j) \geq 0 \quad \sum_{j \in \epsilon_i^+} R_{ij}(\rho^j) \equiv 1$$

Resilience with locally responsive routing

Resilience with locally responsive routing

(a) $R_{ij}(\rho^i)$ depends on local info

$$R_{ij}(\rho^i) \geq 0 \quad \sum_{j \in \mathcal{E}_i^+} R_{ij}(\rho^i) \equiv 1$$



► Theorem [G.C., K.Savla,D.Acemoglu,M.Dahleh,E.Frazzoli,TAC'13]

(a)

$$\Rightarrow \nu \leq \min_v \underbrace{\sum_{j \in \mathcal{E}_v^+} (C_j - f_j^*)}_{\text{node residual capacity}}$$

initial equilibrium f^*

► Example

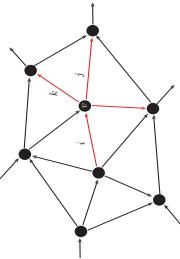
$$R_{ij}(\rho^i) = \frac{e^{-\beta(\rho_j + \alpha_j)}}{\sum_k e^{-\beta(\rho_k + \alpha_k)}}$$

$$R_{ij}(\rho^i) \geq 0 \quad \forall k \neq j$$

$$\rho_j \rightarrow \infty \Rightarrow R_{ij} \rightarrow 0$$

(b) $R_{ij}(\rho^i)$ depends on local info

$$R_{ij}(\rho^i) \geq 0 \quad \sum_{j \in \mathcal{E}_i^+} R_{ij}(\rho^i) \equiv 1$$



$$\frac{\partial}{\partial \rho_k} R_{ij} \geq 0 \quad \forall k \neq j$$

$$\rho_j \rightarrow \infty \Rightarrow R_{ij} \rightarrow 0$$

Resilience with locally responsive routing

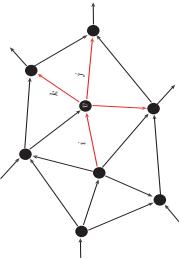
Resilience with locally responsive routing

(a) $R_{ij}(\rho^i)$ depends on local info

$$R_{ij}(\rho^i) \geq 0 \quad \sum_{j \in \mathcal{E}_i^+} R_{ij}(\rho^i) \equiv 1$$

$$\frac{\partial}{\partial \rho_k} R_{ij} \geq 0 \quad \forall k \neq j$$

$$\rho_j \rightarrow \infty \Rightarrow R_{ij} \rightarrow 0$$



► Theorem [G.C., K.Savla,D.Acemoglu,M.Dahleh,E.Frazzoli,TAC'13]

In acyclic networks

(a) and (b)

$$\Rightarrow \nu = \min_v \sum_{j \in \mathcal{E}_v^+} (C_j - f_j^*)$$

f^* initial equilibrium

f^* globally attractive

► Theorem [G.C., K.Savla,D.Acemoglu,M.Dahleh,E.Frazzoli,TAC'13]

$$(a) \text{ and } (b) \quad \Rightarrow \quad \nu = \min_v \sum_{j \in \mathcal{E}_v^+} (C_j - f_j^*)$$

► shocks (and information) propagate only downstream

sub-additive effects of perturbations

Resilience with locally responsive routing

(a) $R_{ij}(\rho^j)$ depends on local info

$$R_{ij}(\rho^i) \geq 0 \quad \sum_{j \in \mathcal{E}_i^+} R_{ij}(\rho^i) \equiv$$

(b) $\partial P_{\cdot \cdot} > 0 \quad \forall t \neq 0$

$$\partial_i \Rightarrow R^i \Rightarrow 0$$

卷之三

► Theorem [G.C., K.Savla, D.Acemoglu, M.Dahleh, E.Frazzoli, TAC'13]

$$(a) \text{ and } (b) \quad \Rightarrow \quad \nu = \min_{\nu} \sum_{j \in \mathcal{E}_v^+} (C_j - f_j^*)$$

► typically $\min_v \sum_{j \in \mathcal{E}_v^+} (C_j - f_j^*) \ll \min_U \{C_U - \lambda_U\}$

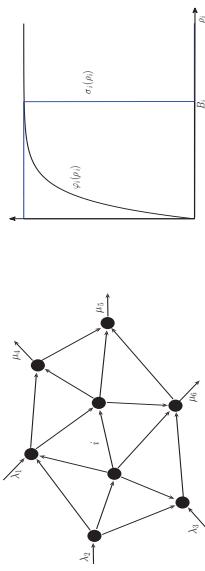
Dynamical flow networks with cascading failures

► Theorem [G.C., K.Savla, D.Acemoglu, M.Dahleh, E.Frazzoli, TAC'13]

$$(a) \text{ and } (b) \quad \Rightarrow \quad \nu = \min_v \sum_{j \in \mathcal{E}_v^+} (C_j - f_j^*)$$

► typically $\min \sum (C_j - f_i^*) \ll \min \{C_{\mathcal{Y}} - \lambda_{\mathcal{Y}}\}$

Dynamical flow networks with cascading failures



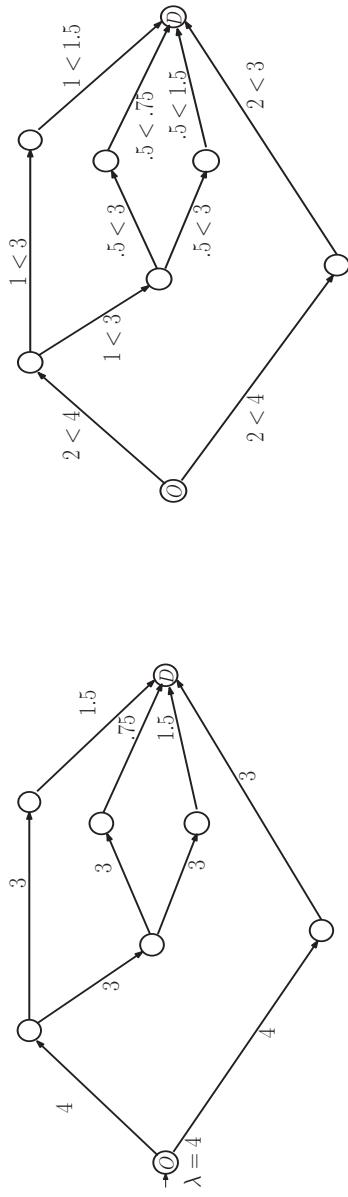
$$\dot{\rho}_i = \lambda_i + \sum_j f_{ij} - \sum_j f_{ij} - \mu_i$$

$$f_{ij} = \varphi_i(\rho_j) R_{ij}(\rho^j)$$

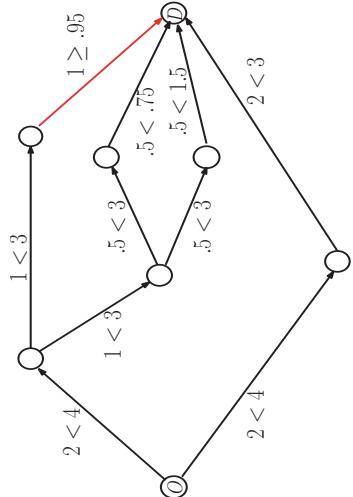
- link i goes down irreversibly when $\rho_i(t) = B_i$

Dynamical flow networks with cascading failures

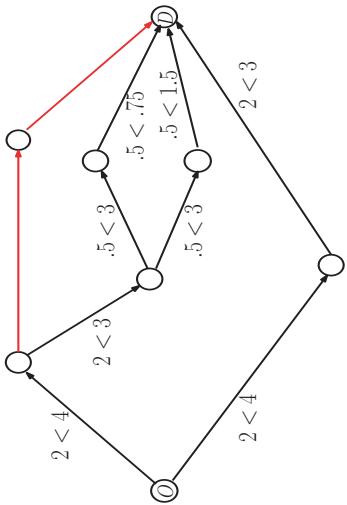
Dynamical flow networks with cascading failures



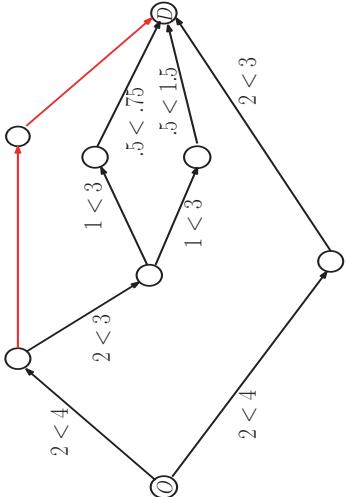
Dynamical flow networks with cascading failures



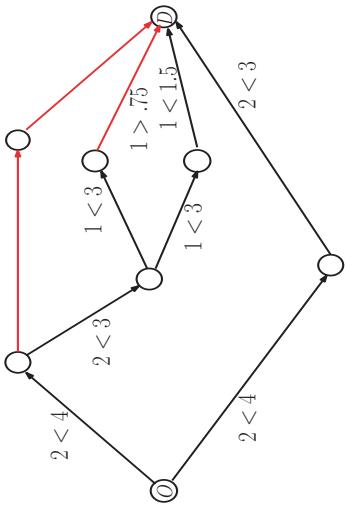
Dynamical flow networks with cascading failures



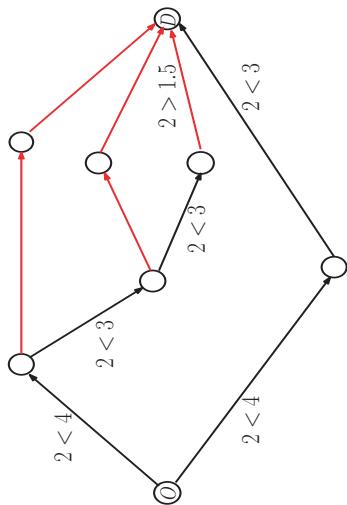
Dynamical flow networks with cascading failures



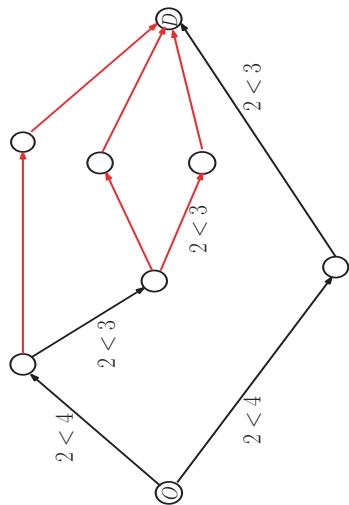
Dynamical flow networks with cascading failures



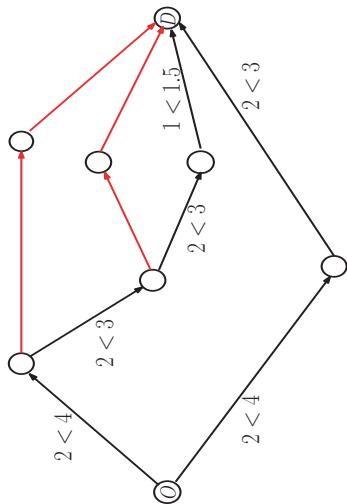
Dynamical flow networks with cascading failures



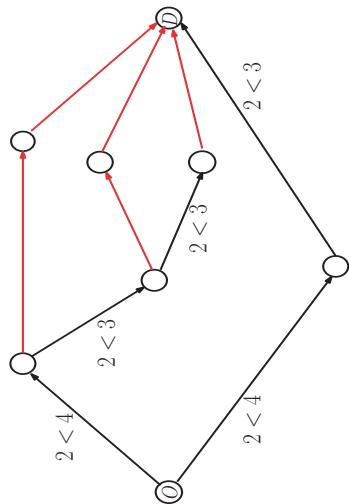
Dynamical flow networks with cascading failures



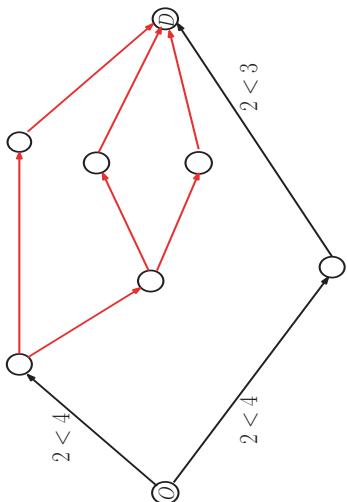
Dynamical flow networks with cascading failures



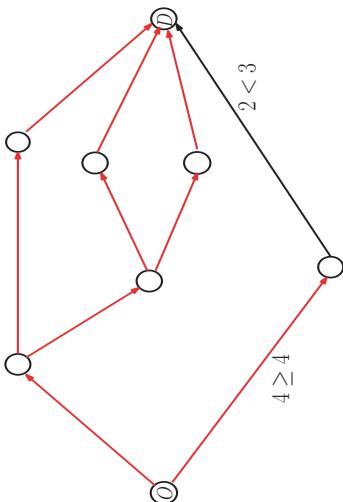
Dynamical flow networks with cascading failures



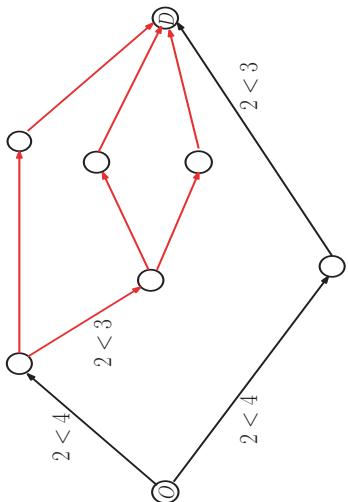
Dynamical flow networks with cascading failures



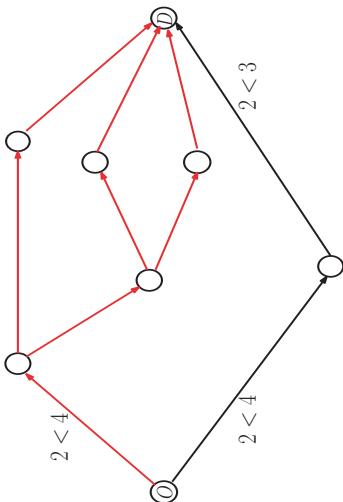
Dynamical flow networks with cascading failures



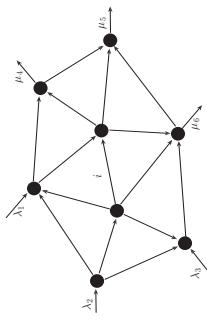
Dynamical flow networks with cascading failures



Dynamical flow networks with cascading failures



Dynamical flow networks with cascading failures



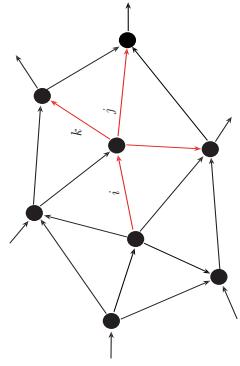
Theorem [K.Savla, G.C., M.Dahleh, 14] In acyclic networks

optimal resilience ν can be computed by dynamic programming

- shocks propagate both up- and downstream
- typically $\nu \ll \min_{\mathcal{U}} \{C_U - \lambda_U\}$

Decentralized routing with flow control

Decentralized routing with flow control



$$f_{ij} = \varphi_i(\rho_i) \max_{j \in \mathcal{E}_i^+} R_j(\rho_j)$$

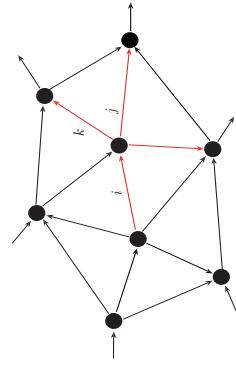
↑ flow from i to j

↑ $\varphi_i(\rho_i)$ $R_j(\rho_j)$

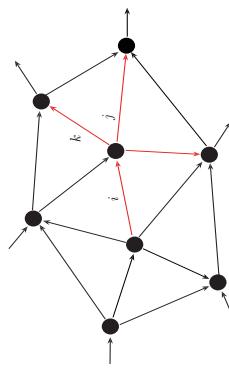
↑ max outflow from i

↑ fraction routed to j

Decentralized monotone routing with flow control



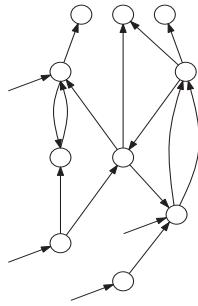
$$R_{ij}(\rho') = \frac{e^{-\beta(\rho_j + \alpha_{ij})}}{e^{-\beta(\rho_i + \alpha_i)} + \sum_{k \in \mathcal{E}_i^+} e^{-\beta(\rho_k + \alpha_k)}}$$



$$f_{ij} = \varphi_i(\rho_i) R_{ij}(\rho_j)$$

- relax $\sum_{j \in \mathcal{E}_i^+} R_{ij} = 1$ to $\sum_{j \in \mathcal{E}_i^+} R_{ij} \leq 1$, still decentralized

Decentralized monotone routing with flow control

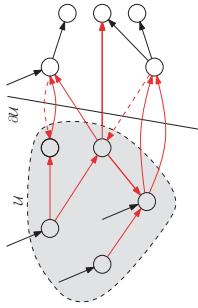


Theorem [G.C., E.Lovisari,K.Savla,TCONES'14] :

Decentralized monotone routing, both with finite and infinite buffer

- ▶ $\min_U \{C_U - \lambda_U\} > 0 \implies \exists \text{ equilibrium } \rho^* \text{ s.t. } \rho(t) \rightarrow \rho^* \quad \forall \rho(0)$

Decentralized monotone routing with flow control

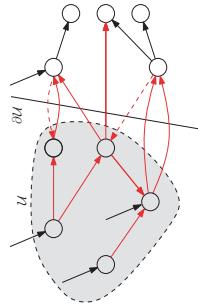


Theorem [G.C., E.Lovisari,K.Savla,TCONES '14] :

Decentralized monotone routing, infinite buffer

- ▶ $\min_U \{C_U - \lambda_U\} < 0 \implies \text{minimal throughput loss (graceful degradation)}$

Decentralized monotone routing with flow control

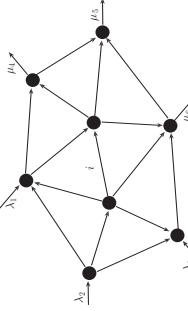


Theorem [G.C., E.Lovisari,K.Savla,TCONES '14] :

Decentralized monotone routing, finite buffer

- ▶ $\min_U \{C_U - \lambda_U\} < 0 \implies \text{links left of cut fail simultaneously}$

Beyond throughput and resilience



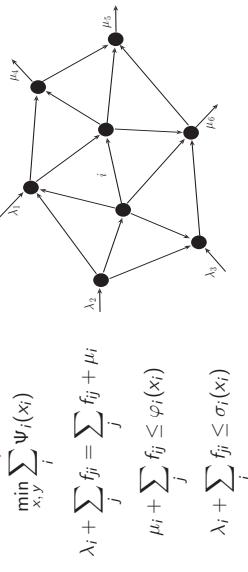
► decentralized routing + flow control achieves optimal resilience

implicitly propagating information through the network

- ▶ for other performance measures (e.g., delay at equilibrium) communication / distributed optimization layer necessary

What equilibrium?

Equilibrium optimization



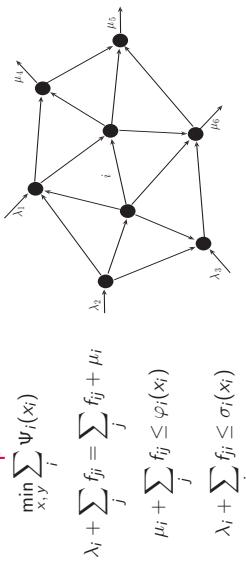
[E.Lovisari,G.C.,A.Rantzer,K.Savla,'14]

If φ_i , σ_i , and $-\psi_i$ are concave (linear):

- equilibrium optimization is convex (linear) problem
- suitable for distributed solutions (ADMM)

$$\alpha_{ij} = -\rho_j^* - \beta^{-1} \log(f_{ij}^*) \quad \alpha_{ii} = -\rho_i^* - \beta^{-1} \log(\varphi_i(\rho_i^*) - \sum_j f_{ij}^*)$$

Equilibrium optimization

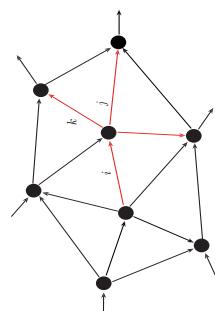


[E.Lovisari,G.C.,A.Rantzer,K.Savla,'14]

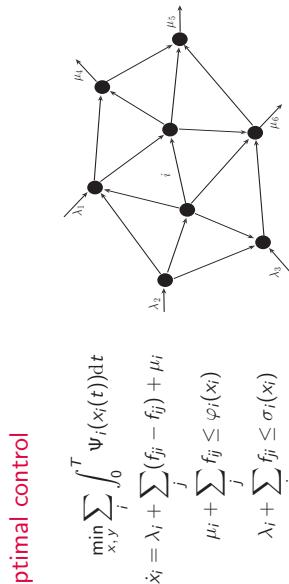
If φ_i , σ_i , and $-\psi_i$ are concave (linear):

- equilibrium optimization is convex (linear) problem
- analogous problem when routing is fixed (or constrained)

Optimal control



$$R_{ij}(\rho^i) = \frac{e^{-\beta(\rho_j+\alpha_{ij})}}{e^{-\beta(\rho_i+\alpha_{ii})} + \sum_{k \in \mathcal{E}_i^+} e^{-\beta(\rho_k+\alpha_{ik})}}$$



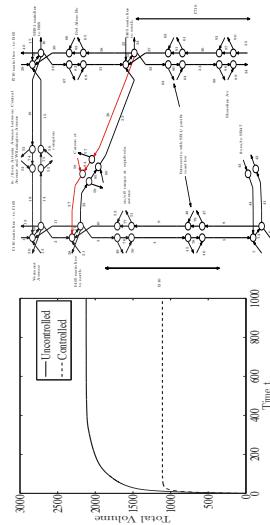
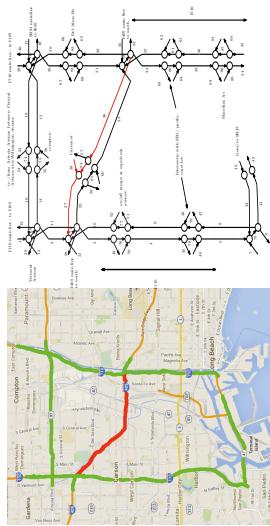
[E.Lovisari,G.C.,A.Rantzer,K.Savla,'14]

If φ_i , σ_i , and $-\psi_i$ are concave (linear):

- optimal control is convex (linear) problem

Highway net in Long Beach

Highway net in Long Beach



Conclusion

Dynamical flow networks:

- ▶ resilience and cascading failures with decentralized routing
- ▶ distributed equilibrium optimization and optimal control
- ▶ proofs: exploit monotonicity and l_1 -contraction

Dynamical flow networks beyond transportation:

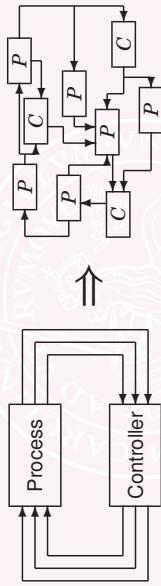
- ▶ production networks
- ▶ distribution networks

CONTROL OF CONVEX-MONOTONE SYSTEMS**Anders Rantzer, Lund University**

We define the notion of convex-monotone system and prove that for such systems the state trajectory is a convex function of the initial state and the input trajectory. This observation gives a useful class of nonlinear dynamical systems for which optimal trajectories can be performed by convex optimization. Applications to evolutionary dynamics of diseases and voltage stability in power networks are presented. In particular, first order convex optimization methods enable computation of optimal trajectories with a complexity that grows only linearly with the number of network edges.



Towards a Scalable Control Theory



Scalable Control of Convex-Monotone Systems
Anders Rantzer and Bo Bernhardsson
LCCC Linnaeus Center, Lund University, Sweden

Can we find distributed controllers by distributed computation?

Anders Rantzer, LCCC Linnaeus center | Scalable Control of Convex-Monotone Systems

Outline

- Positive and Convex-Monotone Systems

- Voltage Stability
- HIV and Cancer Treatment

Positive systems

A linear system is called *positive* if the state and output remain nonnegative as long as the initial state and the inputs are nonnegative:

$$\frac{dx}{dt} = Ax + Bu \quad y = Cx$$

Equivalently, A , B and C have nonnegative coefficients except for the diagonal of A .

Examples:

- Probabilistic models.
- Economic systems.
- Chemical reactions.
- Traffic Networks.

Anders Rantzer, LCCC Linnaeus center | Scalable Control of Convex-Monotone Systems

Anders Rantzer, LCCC Linnaeus center | Scalable Control of Convex-Monotone Systems

Anders Rantzer, LCCC Linnaeus center | Scalable Control of Convex-Monotone Systems

Positive Systems and Nonnegative Matrices

Classics:

Mathematics: Perron (1907) and Frobenius (1912)

Economics: Leontief (1936)

Books:

Nonnegative matrices: Berman and Plemmons (1979)

Dynamical Systems: Luenberger (1979)

Recent control related work:

Biology inspired theory: Angelil and Sontag (2003)

Synthesis by linear programming: Rami and Tadeo (2007)

Switched systems: Liu (2009), Fornasini and Valcher (2010)

Distributed control: Tanaka and Langbort (2010)

Robust control: Briat (2013)

Stability of Positive systems

Suppose the matrix A has nonnegative off-diagonal elements.
Then the following conditions are equivalent:

(i) The system $\frac{dx}{dt} = Ax$ is exponentially stable.

(ii) There exists a vector $\xi > 0$ such that $A\xi < 0$.
(The vector inequalities are elementwise.)

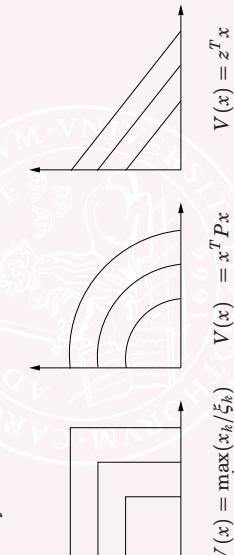
(iii) There exists a vector $z > 0$ such that $A^T z < 0$.

(iv) There is a diagonal matrix $P > 0$ such that
 $A^T P + PA < 0$

Lyapunov Functions of Positive systems

Solving the three alternative inequalities gives three different Lyapunov functions:

$$A\xi < 0 \quad A^T z < 0 \quad A^T P + PA < 0$$



$$V(x) = \max_k (x_k / \xi_k) \quad V(x) = x^T P x \quad V(x) = z^T x$$

A Scalable Stability Test for Positive Systems

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems



Stability of $\dot{x} = Ax$ follows from existence of $\xi_k > 0$ such that

$$\begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & a_{32} \\ a_{41} & 0 & a_{43} & a_{44} \end{bmatrix} \underbrace{\begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix}}_A < \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The first node verifies the inequality of the first row.

The second node verifies the inequality of the second row.

...
Verification is scalable!

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Scalable Control of Convex-Monotone Systems

A Distributed Search for Stabilizing Gains

Examples: Transportation Networks

$$\text{Suppose } \begin{bmatrix} a_{11} - \ell_1 & a_{12} & 0 & a_{14} \\ a_{21} + \ell_1 & a_{22} - \ell_2 & a_{23} & 0 \\ 0 & a_{32} + \ell_2 & a_{33} & a_{32} \\ a_{41} & 0 & a_{43} & a_{44} \end{bmatrix} \geq 0 \text{ for } \ell_1, \ell_2 \in [0, 1].$$

For stabilizing gains ℓ_1, ℓ_2 , find $0 < \mu_k < \xi_k$ such that

$$\begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} \\ a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{32} & a_{33} & a_{32} \\ a_{41} & 0 & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} < \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

and set $\ell_1 = \mu_1/\xi_1$ and $\ell_2 = \mu_2/\xi_2$. Every row gives a local test. Distributed synthesis by linear programming (gradient search).

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Externally Positive Systems

$\mathbf{G} \in \mathbb{R}\mathbb{H}_\infty^{m \times n}$ is called *externally positive* if the corresponding impulse response $g(t)$ is nonnegative for all t . The set of all such matrices is denoted $\mathbb{PH}_\infty^{m \times n}$.

Suppose $\mathbf{G}, \mathbf{H} \in \mathbb{H}_\infty^{m \times n}$. Then

- $\mathbf{GH} \in \mathbb{PH}_\infty^{n \times n}$
- $a\mathbf{G} + b\mathbf{H} \in \mathbb{PH}_\infty^{m \times n}$ when $a, b \in \mathbb{R}_+$.
- $\|\mathbf{G}\|_\infty = \|\mathbf{G}(0)\|$.
- $(I - \mathbf{G})^{-1} \in \mathbb{PH}_\infty^{n \times n}$ if and only if $\mathbf{G}(0)$ is Schur.

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Positively Dominated Systems

$\mathbf{G} \in \mathbb{R}\mathbb{H}_\infty^{m \times n}$ is called *positively dominated* if $|\mathbf{G}_{ijk}(i\omega)| \leq \mathbf{G}_{ijk}(0)$ for $\omega \in \mathbb{R}$. The set of all such matrices is denoted $\mathbb{DEH}_\infty^{m \times n}$.

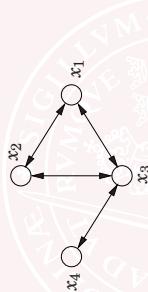
Suppose $\mathbf{G}, \mathbf{H} \in \mathbb{DEH}_\infty^{m \times n}$. Then

- $\mathbf{GH} \in \mathbb{DEH}_\infty^{n \times n}$
- $a\mathbf{G} + b\mathbf{H} \in \mathbb{DEH}_\infty^{m \times n}$ when $a, b \in \mathbb{R}_+$.
- $\|\mathbf{G}\|_\infty = \|\mathbf{G}(0)\|$.
- $(I - \mathbf{G})^{-1} \in \mathbb{DEH}_\infty^{n \times n}$ if and only if $\mathbf{G}(0)$ is Schur.

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Example 3: Mass-spring system



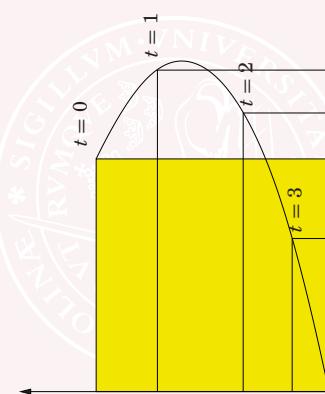
$$\begin{aligned} \ddot{x}_i + d_i \dot{x}_i + k_i x_i &= \sum_j \ell_{ij} (x_j - x_i) + w_i \\ \left(s^2 + d_i s + k_i + \sum_j \bar{\ell}_{ij} \right) X_i(s) &= \sum_j \left(\ell_{ij} X_j(s) + (\bar{\ell}_{ij} - \ell_{ij}) X_i(s) \right) + W_i(s) \\ X = (\mathbf{A} + \mathbf{E}L\mathbf{F})X + \mathbf{B}W \end{aligned}$$

The transfer matrices \mathbf{B} , \mathbf{E} and $\mathbf{A} + \mathbf{E}L\mathbf{F}$ are positively dominated for all $L \in \mathcal{D}$ provided that $d_i \geq k_i + \sum_j \bar{\ell}_{ij}$.

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Proof idea



The system

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = a$$

is a **monotone system** if its linearization is a positive system. It is a **convex monotone system** if every row of f is also convex.

Theorem. [Rantzer/Bernhardsson (2014)]

For a convex monotone system $\dot{x} = f(x, u)$, each component of the trajectory $\phi_t(a, u)$ is a convex function of (a, u) .

Max-separable Lyapunov Functions

$$\text{Max-separable: } V(x) = \max\{V_1(x_1), \dots, V_n(x_n)\}$$

Theorem. Let $\dot{x} = f(x)$ be a monotone system such that the origin globally asymptotically stable and the compact set $X \subset \mathbb{R}_+^n$ is invariant. Then there exist strictly increasing functions $V_k : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ for $k = 1, \dots, n$, such that $V(x) = \max\{V_1(x_1), \dots, V_n(x_n)\}$ satisfies

$$\frac{d}{dt} V(x(t)) = -V(x(t))$$

along all trajectories in X .

[Rantzer, Rüffer, Dirr, CDC-13]

Anders Rantzer, LCCC Linnaeus center

Convex-Monotone Systems

The system

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = a$$

is a **monotone system** if its linearization is a positive system. It is a **convex monotone system** if every row of f is also convex.

Theorem. [Rantzer/Bernhardsson (2014)]

For a convex monotone system $\dot{x} = f(x, u)$, each component of the trajectory $\phi_t(a, u)$ is a convex function of (a, u) .

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Outline

- Positive and Convex -Monotone Systems
- **Voltage Stability**
- HIV and Cancer Treatment

One Transmission Line



◦ Positive and Convex -Monotone Systems

Voltage Stability

The power $p = iu_2$ delivered to the load is upper bounded by

$$p = i(u_1 - Ri) \leq \frac{u_1^2}{4R}.$$

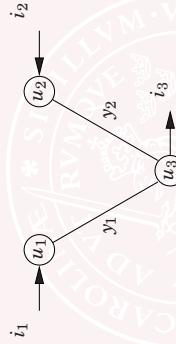
An active load:

$$\frac{di}{dt} = \frac{\hat{p}}{u_1 - Ri} - i.$$

where \hat{p} is the power demand.

Voltage collapse occurs if \hat{p} is too large!

Two Transmission Lines



Node 3 is an active load with

$$\frac{di_3}{dt} = \frac{\hat{p}(y_1 + y_2)}{y_1 u_1 + y_2 u_2 - i_3} - i_3$$

For constant generator voltages u_1 and u_2 , the load voltage $u_3 = y_1 u_1 + y_2 u_2 - i_3$ could shrink to zero in finite time, which means voltage collapse.

Arbitrary Networks

Voltages at generators u^G and loads u^L are mapped into external currents i^G and i^L according to

$$\begin{bmatrix} -i^G(t) \\ i^L(t) \end{bmatrix} = \begin{bmatrix} Y^{GG} & Y^{GL} \\ Y^{LG} & Y^{LL} \end{bmatrix} \begin{bmatrix} u^G(t) \\ u^L(t) \end{bmatrix}$$

The load model: $\frac{di_k^L}{dt}(t) = \frac{\hat{p}_k}{u_k^L(t)} - i_k^L(t)$ gives

$$\frac{di^L}{dt}(t) = \hat{p} \cdot [(Y^{LL})^{-1}(i^L - Y^{LG}u^G)] - i^L(t)$$

This system is convex-monotone with state i^L and input $-u^G$, so

$$i^G, -u^L, i^L, \frac{di^L}{dt} \text{ and } \frac{di^G}{dt}$$

are all convex functions of u^G

Outline

Combination Therapy is a Control Problem

- Positive and Convex-Monotone Systems
- Voltage Stability
- HIV and Cancer Treatment
 -

Evolutionary dynamics:

$$\dot{x} = \left(A - \sum_i u_i D^i \right) x$$

Each state x_k is the concentration of a mutant. (There can be hundreds!) Each input u_i is a drug dosage.

A describes the mutation dynamics without drugs, while D^1, \dots, D^m are diagonal matrices modeling drug effects.

Determine $u_1, \dots, u_m \geq 0$ with $u_1 + \dots + u_m \leq 1$ such that x decays as fast as possible!

[Jonsson, Ranzer,Murray, ACC 2014]

Optimizing Decay Rate

Stability of the matrix $A - \sum_i u_i D^i + \gamma I$ is equivalent to existence of $\xi > 0$ with

$$(A - \sum_i u_i D^i + \gamma I)\xi < 0$$

For row k , this means

$$A_k \xi - \sum_i u_i D_{k,k}^{i,\xi} + \gamma \xi_k < 0$$

or equivalently

$$\frac{A_k \xi}{\xi_k} - \sum_i u_i D_k^i + \gamma < 0$$

Maximizing γ is convex optimization in $(\log \xi_i, u_i, \gamma)$!

Evolutionary dynamics:

$$\dot{x}(t) = \left(A - \sum_i u_i(t) D^i \right) x(t)$$

Anders Ranzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Using Measurements of Virus Concentrations

Evolutionary dynamics:

$$\dot{x}(t) = \left(A - \sum_i u_i(t) D^i \right) x(t)$$

Can we get faster decay using time-varying $u(t)$ based on measurements of $x(t)$?

Anders Ranzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Using Measurements of Virus Concentrations

The evolutionary dynamics can be written as a convex monotone system:

$$\frac{d}{dt} \log x_k(t) = \frac{A_k x(t)}{x_k(t)} - \sum_i u_i(t) D_k^i$$

Hence the decay of $\log x_k$ is a convex function of the input and optimal trajectories can be found even for large systems.

$$A = \begin{bmatrix} -\delta & \mu & \mu & 0 \\ \mu & -\delta & 0 & \mu \\ 0 & 0 & -\delta & \mu \\ 0 & \mu & \mu & -\delta \end{bmatrix}$$

clearance rate $\delta = 0.24 \text{ day}^{-1}$, mutation rate $\mu = 10^{-4} \text{ day}^{-1}$ and replication rates for viral variants and therapies as follows

Virus variant	Therapy 1	Therapy 2	Therapy 3
Type 1 (x_1)	$D_1^1 = 0.05$	$D_1^2 = 0.10$	$D_1^3 = 0.30$
Type 2 (x_2)	$D_2^1 = 0.25$	$D_2^2 = 0.05$	$D_2^3 = 0.30$
Type 3 (x_3)	$D_3^1 = 0.10$	$D_3^2 = 0.30$	$D_3^3 = 0.30$
Type 4 (x_4)	$D_4^1 = 0.30$	$D_4^2 = 0.30$	$D_4^3 = 0.15$

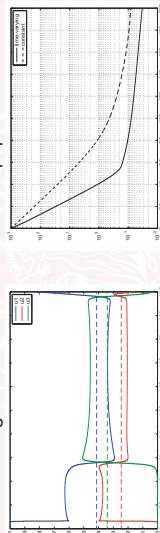
Example

Anders Rantzer, LCCC Linnaeus center

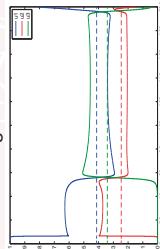
Scalable Control of Convex-Monotone Systems

Example

Total virus population:



Optimized drug doses:



Summary

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

- Scalability for Positive and Convex-Monotone Systems
- Voltage Stability
- HIV and Cancer Treatment

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems

Anders Rantzer, LCCC Linnaeus center

Scalable Control of Convex-Monotone Systems



LUND
UNIVERSITY

Department of Automatic Control
Box 118, 221 00 Lund, Sweden
www.lcc.lth.se

ISRN LUTFD2/TFRT--7640--SE
ISSN 0280-5316