
Daniel P. Palomar
Hong Kong University of Science and Technology (HKUST)

Gesualdo Scutari
State University of New York (SUNY) at Buffalo

International Workshop on Mathematical Issues in Information Sciences - MIIS 2012
Xi’an, China
July 7-13, 2012
Acknowledgement

The results on game theory and variational inequality theory are joint work with our collaborators:

• Sergio Barbarossa, INFOCOM Dept., University of Rome “La Sapienza”, Italy.

• Francisco Facchinei, Dept. of Computer and System Sciences, University of Rome “La Sapienza”, Italy.

• Jong-Shi Pang, Dept. of Industrial & Enterprise Systems Engineering, UIUC, IL, USA.
Outline of Tutorial

• Part I: Overview of Convex Optimization and Game Theory.

• Part II: Variational Inequalities: The Theory.

• Part III: Variational Inequalities: Applications.
Part I:

Overview of Convex Optimization
and Game Theory
Part I - Outline

- Convex optimization:
  - Convexity and convex optimization problems
  - Classes of convex problems
  - Lagrange duality and KKT conditions.

- Game theory:
  - Nash equilibrium and Pareto optimality
  - Existence/uniqueness theorems
  - Algorithms.
Convex Optimization
Optimization Problem

• General optimization problem in standard form:

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0 \quad i = 1, \ldots, m \\
& \quad h_i(x) = 0 \quad i = 1, \ldots, p
\end{align*}
\]

where

\[
x = (x_1, \ldots, x_n) \text{ is the optimization variable}
\]

\[
f_0 : \mathbb{R}^n \rightarrow \mathbb{R} \text{ is the objective function}
\]

\[
f_i : \mathbb{R}^n \rightarrow \mathbb{R}, \quad i = 1, \ldots, m \text{ are inequality constraint functions}
\]

\[
h_i : \mathbb{R}^n \rightarrow \mathbb{R}, \quad i = 1, \ldots, p \text{ are equality constraint functions.}
\]

• Goal: find an optimal solution \( x^* \) that minimizes \( f_0 \) while satisfying all the constraints.
Examples

Convex optimization is currently used in many different areas:

- circuit design (start-up named Barcelona in Silicon Valley)
- data fitting
- signal processing (e.g., filter design)
- communication systems (e.g., transceiver design, beamforming design, ML detection, power control in wireless)
- image processing (e.g., deblurring, compressive sensing, blind separation)
- biomedical applications (e.g., analysis of DNA)
- portfolio optimization (e.g., investment in assets)
**Example: Power Control in Wireless Networks**

- Consider a wireless network with $n$ logical transmitter/receiver pairs:

  - Goal: design the power allocation so that each receiver receives minimum interference from the other links.
• The signal-to-interference-plus-noise-ratio (SINR) at the $i$th receiver is

$$\text{sinr}_i = \frac{p_i G_{ii}}{\sum_{j \neq i} p_j G_{ij} + \sigma_i^2}$$

where

$p_i$ is the power used by the $i$th transmitter
$G_{ij}$ is the path gain from transmitter $j$ to receiver $i$
$\sigma_i^2$ is the noise power at the $i$th receiver.

• **Problem:** maximize the weakest SINR subject to power constraints $0 \leq p_i \leq p_i^{\text{max}}$:

$$\max_p \min_{i=1,\ldots,n} \frac{p_i G_{ii}}{\sum_{j \neq i} p_j G_{ij} + \sigma_i^2}$$

subject to $0 \leq p_i \leq p_i^{\text{max}} \quad i = 1, \ldots, n.$
Solving Optimization Problems

- General optimization problems are very difficult to solve (either a long computation time or not always finding the solution).

- Exceptions: least-squares problems, linear programming problems, and convex optimization problems.

- Least-squares (LS):

  \[
  \text{minimize} \quad \| Ax - b \|_2^2
  \]

- solving LS problems: closed-form solution \( x^* = (A^T A)^{-1} A^T b \) for which there are reliable and efficient algorithms; mature technology

- using LS: easy to recognize
• **Linear Programming (LP):**

\[
\begin{align*}
\min_{x} & \quad c^T x \\
\text{subject to} & \quad a_i^T x \leq b_i, \quad i = 1, \ldots, m
\end{align*}
\]

- solving LP problems: no closed-form solution, but reliable and efficient algorithms and software; mature technology
- using LP: not as easy to recognize as LS problems, a few standard tricks to converse problems into LPs

• **Convex optimization:**

\[
\begin{align*}
\min_{x} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq b_i, \quad i = 1, \ldots, m
\end{align*}
\]

- solving convex problems: no closed-form solution, but still reliable and efficient algorithms and software; almost a technology
- using convex optimization: often difficult to recognize, many tricks for transforming problems into convex form.
History Snapshot of Convex Optimization

- **Theory** (convex analysis): ca1900-1970 (e.g. Rockafellar)
- **Algorithms:**
  - 1947: simplex algorithm for linear programming (Dantzig)
  - 1960s: early interior-point methods (Fiacco & McCormick, Dikin)
  - 1970s: ellipsoid method and other subgradient methods
  - 1980s: polynomial-time interior-point methods for linear programming (Karmakar 1984)
- **Applications:**
  - before 1990s: mostly in operations research; few in engineering
  - since 1990: many new applications in engineering and new problem classes (SDP, SOCP, robust optim.)
Convex Sets and Functions
Convex Set

• A set $C \in \mathbb{R}^n$ is said to be convex if the line segment between any two points is in the set: for any $x, y \in C$ and $0 \leq \theta \leq 1$,

$$\theta x + (1 - \theta) y \in C.$$

• Examples: hyperplanes, halfspaces, polyhedra, balls, ellipsoids, convex cones, norm balls, norm cones, PSD cone, etc.
Convex Function

• A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be convex if the domain, $\text{dom} \ f$, is convex and for any $x, y \in \text{dom} \ f$ and $0 \leq \theta \leq 1$,

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y).$$

• $f$ is concave if $-f$ is convex.

• Second-order condition for convexity: $\nabla^2 f(x) \succeq 0 \quad \forall x \in \text{dom} \ f$
Examples on $\mathbb{R}$

Convex functions:
- affine: $ax + b$ on $\mathbb{R}$
- powers of absolute value: $|x|^p$ on $\mathbb{R}$, for $p \geq 1$ (e.g., $|x|$)
- powers: $x^p$ on $\mathbb{R}_{++}$, for $p \geq 1$ or $p \leq 0$ (e.g., $x^2$)
- exponential: $e^{ax}$ on $\mathbb{R}$
- negative entropy: $x \log x$ on $\mathbb{R}_{++}$

Concave functions:
- affine: $ax + b$ on $\mathbb{R}$
- powers: $x^p$ on $\mathbb{R}_{++}$, for $0 \leq p \leq 1$
- logarithm: $\log x$ on $\mathbb{R}_{++}$
Examples on $\mathbb{R}^n$

- **Affine functions** $f(x) = a^T x + b$ are convex and concave on $\mathbb{R}^n$.

- **Norms** $\|x\|$ are convex on $\mathbb{R}^n$ (e.g., $\|x\|_\infty$, $\|x\|_1$, $\|x\|_2$).

- **Quadratic functions** $f(x) = x^T P x + 2q^T x + r$ are convex $\mathbb{R}^n$ if and only if $P \succeq 0$.

- The **geometric mean** $f(x) = \prod_{i=1}^n x_i$ is concave on $\mathbb{R}^{n+}$.

- The **log-sum-exp** $f(x) = \log \sum_i e^{x_i}$ is convex on $\mathbb{R}^n$ (it can be used to approximate $\max_{i=1,...,n} x_i$).

- **Quadratic over linear**: $f(x, y) = x^2 / y$ is convex on $\mathbb{R}^n \times \mathbb{R}^{++}$. 
Examples on \( \mathbb{R}^{n \times n} \)

- **Affine functions:** (prove it!)
  \[
  f(X) = \text{Tr}(AX) + b
  \]
  are convex and concave on \( \mathbb{R}^{n \times n} \).

- **Logarithmic determinant function:** (prove it!)
  \[
  f(X) = \log \det(X)
  \]
  is concave on \( \mathbb{S}^n = \{X \in \mathbb{R}^{n \times n} \mid X \succeq 0\} \).

- **Maximum eigenvalue function:** (prove it!)
  \[
  f(X) = \lambda_{\text{max}}(X) = \sup_{y \neq 0} \frac{y^T X y}{y^T y}
  \]
  is convex on \( \mathbb{S}^n \).
Convex Optimization Problem

- Convex optimization problem in standard form:

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0 & i = 1, \ldots, m \\
& \quad Ax = b
\end{align*}
\]

where \( f_0, f_1, \ldots, f_m \) are convex and equality constraints are affine.

- **Local and global optima**: any locally optimal point of a convex problem is globally optimal.

- Most problems are not convex when formulated.

- Reformulating a problem in convex form is an art, there is no systematic way.
Classes of Convex Problems
Linear Program (LP)

\[
\begin{align*}
\text{minimize} \quad & c^T x + d \\
\text{subject to} \quad & Gx \leq h \\
& Ax = b
\end{align*}
\]

• Convex problem: affine objective and constraint functions.

• Feasible set is a polyhedron:
**Quadratic Program (QP)**

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} x^T P x + q^T x + r \\
\text{subject to} & \quad Gx \leq h \\
& \quad Ax = b
\end{align*}
\]

- Convex problem (assuming \( P \in S^n \succeq 0 \)): convex quadratic objective and affine constraint functions.

- Minimization of a convex quadratic function over a polyhedron:

![Diagram of a convex quadratic function over a polyhedron]
Second-Order Cone Programming (SOCP)

\[
\begin{align*}
\text{minimize} \quad & f^T x \\
\text{subject to} \quad & \|A_i x + b_i\| \leq c_i^T x + d_i \quad i = 1, \ldots, m \\
& Fx = g
\end{align*}
\]

- Convex problem: linear objective and second-order cone constraints
- For \(A_i\) row vectors, it reduces to an LP.
- For \(c_i = 0\), it reduces to a QCQP.
- More general than QCQP and LP.
Semidefinite Program (SDP)

minimize \( x^T c \)

subject to \( x_1 F_1 + x_2 F_2 + \cdots + x_n F_n \preceq G \)

\( Ax = b \)

- Convex problem: linear objective and linear matrix inequality (LMI) constraints.

- Observe that multiple LMI constraints can always be written as a single one.
Minimum Principle

• For an unconstrained convex optimization problem:

\[
\min_{x} f(x)
\]

an optimal solution \(x^*\) must satisfy \(\nabla f(x^*) = 0\).

• The extension of this optimality criterion to a constrained convex optimization problem

\[
\min_{x} f(x) \\
\text{subject to } x \in \mathcal{K}
\]

is the so-called minimum principle:

\[
(y - x^*)^T \nabla f(x^*) \geq 0 \quad \forall y \in \mathcal{K}.
\]
Geometrical Interpretation of the Minimum Principle

- A feasible $x^*$ that satisfies the minimum principle: $\nabla f(x^*)$ forms a nonobtuse angle with all feasible vectors $d = y - x^*$ emanating from $x^*$:

![Diagram showing the geometrical interpretation of the minimum principle]
Geometrical Interpretation of the Minimum Principle

- A feasible $x^*$ that does NOT satisfy the minimum principle: there are other feasible points $y \neq x^*$ such that $f(y) < f(x^*)$:
Lagrange Duality and KKT Conditions
Lagrangian

- Consider an optimization problem in standard form (not necessarily convex)

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 0 \quad i = 1, \ldots, m \\
& \quad h_i(x) = 0 \quad i = 1, \ldots, p
\end{align*}
\]

with variable \( x \in \mathbb{R}^n \), domain \( D \), and optimal value \( p^* \).

- The Lagrangian is a function \( L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R} \), with \( \text{dom } L = D \times \mathbb{R}^m \times \mathbb{R}^p \), defined as

\[
L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)
\]

where \( \lambda_i \) is the Lagrange multiplier associated with \( f_i(x) \leq 0 \) and \( \nu_i \) is the Lagrange multiplier associated with \( h_i(x) = 0 \).
Karush-Kuhn-Tucker (KKT) Conditions

Necessary (under constraint qualifications) and sufficient KKT conditions for convex problems:

1. primal feasibility: \( f_i(x^*) \leq 0, \ i = 1, \ldots, m, \ h_i(x^*) = 0, \ i = 1, \ldots, p \)

2. dual feasibility: \( \lambda^* \geq 0 \)

3. complementary slackness: \( \lambda_i^* f_i(x^*) = 0 \) for \( i = 1, \ldots, m \)

4. zero gradient of Lagrangian with respect to \( x \):

\[
\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0
\]
Elegant and Compact KKT Conditions

- We can make use of the compact notation $a \perp b \iff a^T b = 0$ to write the primal-dual feasibility and complementary slackness KKT conditions as

$$0 \leq \lambda_i^* \perp - f_i(x^*) \geq 0 \quad \text{for } i = 1, \ldots, m.$$  

- By stacking the $f_i$'s, $h_i$'s, and $\lambda_i^*$'s in vectors, we can finally write the KKT conditions in the following compact form:

$$0 \leq \lambda^* \perp - f (x^*) \geq 0$$

$$h(x^*) = 0$$

and

$$\nabla f_0(x^*) + \nabla f(x^*)^T \lambda^* + \nabla h(x^*)^T \nu^* = 0.$$
Game Theory
Game Theory and Nash Equilibrium

- Game theory deals with problems with interacting decision-makers (called players).

- Noncooperative game theory is a branch of game theory for the resolution of conflicts among selfish players, each of which tries to optimize his own objective function.

- A solution to such a competitive game is an equilibrium point where each user is unilaterally happy and does not want to deviate: it is the so-called **Nash Equilibrium** (NE).

- We will consider two important types of noncooperative games: Nash Equilibrium Problems (NEP) and Generalized NEP (GNEP).
Nash Equilibrium Problems (NEP)

• Mathematically, we can define a NEP as a set of coupled optimization problems:

\[
(G): \quad \begin{align*}
\text{minimize} & \quad f_i(x_i, x_{-i}) \\
\text{subject to} & \quad x_i \in \mathcal{K}_i \\
\end{align*}
\]

\[i = 1, \ldots, Q\]

where:

– \(x_i \in \mathbb{R}^{n_i}\) is the strategy of player \(i\)
– \(x_{-i} \triangleq (x_j)_{j \neq i}\) are the strategies of all the players except \(i\)
– \(\mathcal{K}_i \subseteq \mathbb{R}^{n_i}\) is the strategy set of player \(i\)
– \(f_i(x_i, x_{-i})\) is the cost function of player \(i\).

• How to define a solution of the game?
• **Solution of \( G \):** A (pure strategy) Nash Equilibrium (NE) is a feasible \( x^* = (x^*_i)_{i=1}^{Q} \) such that

\[
f_i(x^*_i, x^*_{-i}) \leq f_i(y_i, x^*_{-i}), \quad \forall y_i \in K_i, \quad \forall i = 1, \ldots, Q
\]

A NE is a strategy profile where every player is *unilaterally* happy.

• Life is not so easy:
  – A (pure strategy) NE may not exist or be unique
  – Even when the NE is unique, there is no guarantee of convergence of iterative (best-response) algorithms.

• **How to analyze a game?**

• **How to design convergent distributed algorithms?**
Minimum Principle for NEPs

- A NE is defined as a simultaneous solution of each of the single-player optimization problems: given the other players $x_{-i}^*$, $x_i^*$ must be the solution to

\[
\begin{align*}
\text{minimize} \quad & f_i(x_i, x_{-i}^*) \\
\text{subject to} \quad & x_i \in \mathcal{K}_i.
\end{align*}
\]

- The minimum principle for a convex game is then, for each $i = 1, \ldots, Q$:

\[
(y_i - x_i^*)^T \nabla f_i (x_i^*, x_{-i}^*) \geq 0 \quad \forall y_i \in \mathcal{K}_i.
\]
KKT Conditions for NEPs

• Suppose now that each set $\mathcal{K}_i$ is described by a set of equalities and inequalities:

$$\mathcal{K}_i = \{x_i : g_i(x_i) \leq 0, \ h_i(x_i) = 0\}.$$

• Then we can write the coupled KKT conditions for the convex NEP, for each $i = 1, \ldots, Q$, as

$$0 \leq \lambda_i^* \perp - g_i(x^*) \geq 0$$

$$h_i(x_i) = 0$$

and

$$\nabla f_i(x^*) + \nabla g_i^T(x_i^*) \lambda_i^* + \nabla h_i^T(x_i^*) \nu_i^* = 0.$$
Pareto Optimality

- Pareto optimality is the natural extension of the concept of optimality in classical optimization to multi-objective optimization.

- A point \( x \) is Pareto optimal if its set of objectives \((f_1(x), \ldots, f_Q(x))\) cannot be improved (element by element).

- Note that some objectives of a Pareto optimal point may be improved by another point, but not all the objectives at the same time.

- Therefore, a multi-objective optimization problem has associated a Pareto-optimal frontier rather than an optimal point.

- We can then compare the performance achieved by a NE versus that achieved by a Pareto optimal point in terms of social optimum. This is related to the concept of price of anarchy in game theory.
Solution Analysis: Existence of NE

• Theorem. [(existence theorem) Debreu-Fan-Glicksberg (1952)]

Given the game $\mathcal{G} = \langle \mathcal{K}, f \rangle$, suppose that:

– The action space $\mathcal{K}$ is compact and convex;

– The cost-functions $f_i(x_i, x_{-i})$ are continuous in $x \in \mathcal{K}$ and quasi-convex in $x_i \in \mathcal{K}_i$, for any given $x_{-i}$.

Then, there exists a pure strategy NE.

• Existence may also follow from the special structure of the game, for example:

– Potential games

– Supermodular games.
Solution Analysis: Uniqueness of NE for Convex Games

- **Theorem. [Rosen uniqueness theorem (1965)]** Given the game $G = \langle \mathcal{K}, f \rangle$, suppose that:
  - The action space $\mathcal{K}$ is compact and convex;
  - The cost-functions $f_i(x_i, x_{-i})$ are continuous in $x \in \mathcal{K}$ and quasi-convex in $x_i \in \mathcal{K}_i$, for any given $x_{-i}$;
  - The Diagonal Strict Convexity (DSC) property holds true: given
    \[ g(x, r) \triangleq (r_i \nabla_{x_i} f_i(x))_{i=1}^Q, \]
    
    DSC: $\exists r > 0 : (x - y)^T (g(x; r) - g(y; r)) > 0$, $\forall x, y \in \mathcal{K}, x \neq y$

Then, there exists a unique pure strategy NE.

- The DSC property is not easy to check and may be too restrictive (we want the conditions to be satisfied for all $x \in \mathcal{K}$).
Best Response

- Let define the best-response $B_i(x_{-i})$ of each player $i$ as the set of optimal solutions of player $i$’s optimization problem for any given $x_{-i}$:

$$
B_i(x_{-i}) \triangleq \{ x_i \in K_i : f_i(x_i, x_{-i}) \leq f_i(y_i, x_{-i}), \quad \forall y_i \in K_i \}.
$$
Solution Analysis: Uniqueness of NE for Standard Games

- **Standard function [Yates, 1995]**: A function $g : \mathcal{K} \rightarrow \mathbb{R}^m_+$ is said to be standard if it has the two following properties:
  
  - **Monotonicity**: $\forall x, y \in \mathcal{K}, \ x \leq y \Rightarrow g(x) \leq g(y)$ (component-wise)
  - **Scalability**: $\forall \alpha > 1, \forall x \in \mathcal{K}, \ g(\alpha x) \leq \alpha g(x)$.

- If the best-response map $B(x) = (B_i(x_{-i}))_{i=1}^Q$ is a standard function (requiring $B_i(x_{-i})$ be a single-valued function), then the NE of the game $\mathcal{G} = \langle \mathcal{K}, f \rangle$ is unique.

- The above requirement is quite strong and in general is not satisfied in practice by games of our interest.
Algorithms for Achieving NE

• Best-response (BR) based dynamics:
  – Convergence requires contraction or monotonicity of the BR map;
  – Contraction can be studied, e.g., using classical fixed-point theory if the BR is known in closed form;
  – Monotonicity is guaranteed, e.g., if the BR is a standard function (too restrictive).

• ODE approximation: the idea is to introduce a time-continuous ordinary differential equation (ODE), whose equilibrium coincides with the NE of the game:
  – Convergence to a NE is guaranteed under the globally asymptotic stability (GAS) of the equilibrium of the ODE;
  – Sufficient conditions for the DSC implies the GAS.
State of the Art in Comm and Netw: Pre-2006

- **Theory of “standard function”** \cite{Yat95}, ..., \cite{Sun05}
  - *Scalar* power control problems for CDMA/TDMA/FDMA ad-hoc/cellular systems \cite{Fos93}, \cite{Mit93}, \cite{Bam98}, ..., \cite{Sar02}, \cite{Xia05}, \cite{Alp05}
  - Vector power control problems are not “standard” \cite{Ros65}, \cite{Ich83}.

- **Games with special structure** \cite{Mon06}, \cite{Top98}
  - Potential games \cite{Kel98},...,\cite{Gun03}, \cite{Scu06},...
  - Supermodular games \cite{Alt03}, ..., \cite{Hua06}.

- **Unfortunately**... most of the games we are interested in do not have a special structure and/or the best response is not a standard function.
Summary

• We have seen the basic definitions of convex sets, functions, and problems with examples, as well as an overview of classes of convex problems such as LP, QP, SOCP, and SDP.

• We have introduced the minimum principle, Lagrange duality, and the KKT conditions.

• Within the context of game theory, we have defined the concepts of NE, Pareto optimality, and have formulated NEPs.

• We have given the state of the art for the solution analysis of NEPs as well as practical algorithms.
References

• Convex optimization:

• Game theory:
Part II:

Variational Inequalities: The Theory
Part II - Outline

- VI as a general framework
- Alternative formulations of VI: KKT conditions, primal-dual form
- Solution analysis of VI: existence, uniqueness, solution as a projection
- Algorithms for VI
- NEP as a VI: solution analysis and algorithms
- GNEP as a VI: variational equilibria and algorithms
VI as a General Framework
The VI Problem

• Given a set $\mathcal{K} \subseteq \mathbb{R}^n$ and a mapping $F : \mathcal{K} \rightarrow \mathbb{R}^n$, the VI problem $\text{VI}(\mathcal{K}, F)$ is to find a vector $x^* \in \mathcal{K}$ such that

$$(y - x^*)^T F(x^*) \geq 0 \quad \forall y \in \mathcal{K}.$$ 

• Let $\text{SOL}(\mathcal{K}, F)$ denote the solution set of $\text{VI}(\mathcal{K}, F)$. 

Geometrical Interpretation of the VI

- A feasible point \( x^* \) that is a solution of the VI\((\mathcal{K}, F)\): \( F(x^*) \) forms an acute angle with all the feasible vectors \( y - x^* \)

 feasable set \( \mathcal{K} \)

\[ \text{feasible set } \mathcal{K} \]

\[ \text{feasible set } \mathcal{K} \]

\[ F(x^*) \]

\[ x^* \]

\[ y \]

\[ y - x^* \]
Geometrical Interpretation of the VI

• A feasible point \( x^\star \) that is NOT a solution of the VI(\( \mathcal{K}, F \))

\[
\begin{align*}
\text{feasible set } & \mathcal{K} \\
y - x & \\
F(x) & \\
x & \\
y & 
\end{align*}
\]
Convex Optimization as a VI

- Convex optimization problem:

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad x \in \mathcal{K}
\end{align*}
\]

where \( \mathcal{K} \subseteq \mathbb{R}^n \) is a convex set and \( f : \mathbb{R}^n \to \mathbb{R} \) is a convex function.

- Minimum principle: The problem above is equivalent to finding a point \( x^* \in \mathcal{K} \) such that

\[
(y - x^*)^T \nabla f(x^*) \geq 0 \quad \forall y \in \mathcal{K} \iff \text{VI}(\mathcal{K}, \nabla f)
\]

which is a special case of VI with \( F = \nabla f \).
• It seems that a VI is more general than a convex optimization problem only when $\mathbf{F} \neq \nabla f$.

• But is it really that significative? The answer is affirmative.

• The $\text{VI}(\mathcal{K}, \mathbf{F})$ encompasses a wider range of problems than classical optimization whenever $\mathbf{F} \neq \nabla f$ ($\iff \mathbf{F}$ has not a symmetric Jacobian).

• Some examples of relevant problems that can be cast as a VI include NEPs, GNEPs, system of equations, nonlinear complementary problems, fixed-point problems, saddle-point problems, etc.
NEP as a VI

• The minimum principle for a convex game is, for each $i = 1, \ldots, Q$:

$$
(y_i - x_i^*)^T \nabla f_i (x_i^*, x_{-i}^*) \geq 0 \quad \forall y_i \in K_i.
$$

• We can write it in a more compact way as

$$(y - x^*)^T F (x^*) \geq 0 \quad \forall y \in K$$

by defining $K = K_1 \times \cdots \times K_Q$ and $F = (\nabla x_i f_i)_{i=1}^Q$ (note that $F \neq \nabla f$).

• Hence, the interpretation of the NEP as a VI:

$$
\min_{x_i \in K_i} f_i(x_i, x_{-i}), \quad \forall i = 1, \ldots, Q \quad \iff \quad \text{VI}(K, F).
$$
System of Equations as a VI

• In some engineering problems, we may not want to minimize a function but instead finding a solution to a system of equations:

\[ F(x) = 0. \]

• This can be cast as a VI by choosing \( K = \mathbb{R}^n \).

• Hence,

\[ F(x) = 0 \iff \text{VI}(\mathbb{R}^n, F). \]
Nonlinear Complementary Problem (NCP) as a VI

- The NCP is a unifying mathematical framework that includes linear programming, quadratic programming, and bi-matrix games.

- The NCP \( F \) is to find a vector \( x^* \) such that

\[
\text{NCP}(F) : \quad 0 \leq x^* \perp F(x^*) \geq 0.
\]

- An NCP can be cast as a VI by choosing \( \mathcal{K} = \mathbb{R}_+^n \):

\[
\text{NCP}(F) \quad \iff \quad \text{VI}(\mathbb{R}_+^n, F).
\]
Linear Complementary Problem (LCP) as a VI

• An LCP is just an NCP with an affine function: \( F(\boldsymbol{x}) = \mathbf{Mx} + \mathbf{q} \):

\[
\text{LCP}(\mathbf{M}, \mathbf{q}) : \quad 0 \leq \mathbf{x}^* \perp \mathbf{Mx}^* + \mathbf{q} \geq 0.
\]

• An LCP can be cast as a VI by choosing \( \mathcal{K} = \mathbb{R}_+^n \):

\[
\text{LCP}(\mathbf{M}, \mathbf{q}) \iff \text{VI}(\mathbb{R}_+^n, \mathbf{Mx} + \mathbf{q}).
\]

• Examples of LCP: Quadratic (linear) programming, Canonical Bimatrix games, Nash-Cournot games.
Quadratic (Linear) Programming as an LCP

- Consider the following QP (non necessarily convex):

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} x^T P x + c^T x + r \\
\text{subject to} & \quad Ax \leq b \\
& \quad x \geq 0
\end{align*}
\]

where \( P \in \mathbb{R}^{n \times n} \) is symmetric (the case \( P = 0 \) gives rise to an LP), \( A \in \mathbb{R}^{m \times n}, \ c \in \mathbb{R}^n, \ b \in \mathbb{R}^m, \ r \in \mathbb{R} \).

- The KKT are necessary conditions for the optimality (and sufficient if \( P \succeq 0 \))

\[
\begin{align*}
\lambda &= Px + c + A^T \mu \geq 0, \quad x \geq 0, \quad \lambda^T x = 0, \\
b - Ax &\geq 0 \quad \mu \geq 0 \quad \mu^T (b - Ax) = 0
\end{align*}
\]
The KKT conditions can be written more compactly as

\[
0 \leq x \perp Px + A^T \mu + c \geq 0 \\
0 \leq \mu \perp -Ax + b \geq 0
\]

which can be rewritten as \( \text{LCP}(M, q) \)

\[
0 \leq \begin{bmatrix} x \\ \mu \end{bmatrix} \perp \begin{bmatrix} P & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ \mu \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix} \geq 0
\]
Fixed-Point Problem as a VI

- In other engineering problems, we may need to find the (unconstrained) fixed-point of a mapping $G(x)$:

  $$\text{find } x \in \mathbb{R}^n \text{ such that } x = G(x).$$

- This can be easily cast as a VI by noticing that a fixed-point of $G(x)$ corresponds to a solution to a system of equations with function $F(x) = x - G(x)$:

  $$x = G(x) \iff \text{VI}({\mathbb R}^n, I - G).$$

- Similarly, for constrained fixed-point equations:

  $$\text{find } x \in \mathcal{K} \text{ such that } x = G(x) \iff \text{VI}(\mathcal{K}, I - G).$$
Saddle-Point Problem as a VI

- Given two sets $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{Y} \subseteq \mathbb{R}^m$, and a function $\mathcal{L}(x, y)$, the saddle point problem is to find a pair $(x^*, y^*) \in \mathcal{X} \times \mathcal{Y}$ such that

$$\mathcal{L}(x^*, y) \leq \mathcal{L}(x^*, y^*) \leq \mathcal{L}(x, y^*) \quad \forall (x, y) \in \mathcal{X} \times \mathcal{Y}$$

or, equivalently,

$$\mathcal{L}(x^*, y^*) = \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \mathcal{L}(x, y) = \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \mathcal{L}(x, y).$$

- Assuming $\mathcal{L}(x, y)$ continuously differentiable and “convex-concave”, if the saddle-point exists, then

$$(x - x^*)^T \nabla_x \mathcal{L}(x^*, y^*) \geq 0 \quad \forall x \in \mathcal{X}$$

$$(y - y^*)^T (- \nabla_y \mathcal{L}(x^*, y^*)) \geq 0 \quad \forall y \in \mathcal{Y}$$
which is equivalent to the $\text{VI}(\mathcal{K}, \mathbf{F})$ with

\[ \mathcal{K} = \mathcal{X} \times \mathcal{Y} \quad \text{and} \quad \mathbf{F} = \begin{bmatrix} \nabla_x \mathcal{L}(x, y) \\ -\nabla_y \mathcal{L}(x, y) \end{bmatrix}. \]
**Alternative Formulations of VI: KKT Conditions**

- Suppose that the (convex) feasible set $\mathcal{K}$ of $\text{VI}(\mathcal{K}, F)$ is described by a set of inequalities and equalities

$$\mathcal{K} = \{ x : g(x) \leq 0, \; h(x) = 0 \}$$

and some constraint qualification holds.

- Then $\text{VI}(\mathcal{K}, F)$ is equivalent to its KKT conditions:

$$0 = F(x) + \nabla g(x)^T \lambda + \nabla h(x)^T \nu$$

$$0 \leq \lambda \perp g(x) \leq 0$$

$$0 = h(x).$$
• To derive the KKT conditions it suffices to realize that if \( x \) is a solution to \( \text{VI}(\mathcal{K}, F) \) then it must solve the following convex optimization problem and vice-versa:

\[
\begin{align*}
\text{minimize} & \quad y^{T}F(x^*) \\
\text{subject to} & \quad y \in \mathcal{K}.
\end{align*}
\]

(Otherwise, there would be a point \( y \) with \( y^{T}F(x^*) < x^{*T}F(x^*) \) which would imply \( (y - x^*)^{T}F(x^*) < 0 \).)

• The KKT conditions of the VI follow from the KKT conditions of this problem noting that the gradient of the objective is \( F(x^*) \).
Alternative Formulations of VI: Primal-Dual Representation

- We can now capitalize on the KKT conditions of $\text{VI}(\mathcal{K}, F)$ to derive an alternative representation of the VI involving not only the primal variable $x$ but also the dual variables $\lambda$ and $\nu$.

- Consider $\text{VI}(\tilde{\mathcal{K}}, \tilde{F})$ with $\tilde{\mathcal{K}} = \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^p$ and

$$
\tilde{F}(x, \lambda, \nu) = \begin{bmatrix}
F(x) + \nabla g(x)^T \lambda + \nabla h(x)^T \nu \\
-g(x) \\
h(x)
\end{bmatrix}.
$$

- The KKT conditions of $\text{VI}(\tilde{\mathcal{K}}, \tilde{F})$ coincide with those of $\text{VI}(\mathcal{K}, F)$. Hence, both VIs are equivalent.
• VI($\mathcal{K}, F$) is the original (primal) representation whereas VI($\tilde{\mathcal{K}}, \tilde{F}$) is the so-called primal-dual form as it makes explicit both primal and dual variables.

• In fact, this primal-dual form is the VI representation of the KKT conditions of the original VI.

• The primal-dual form is useful to study GNEPs with shared constraints.
Generalization of VI

• **Generalized VI** $GVI(\mathcal{K}, F)$:
  
  - $\mathcal{K} \subseteq \mathbb{R}^n$ is a fixed set;
  - $F : \mathbb{R}^n \ni x \Rightarrow F(x) \subseteq \mathbb{R}^n$ is a set-valued map.

• **Quasi-VI** $QVI(\mathcal{K}, F)$:
  
  - $\mathcal{K} : \mathbb{R}^n \ni x \Rightarrow \mathcal{K}(x) \subseteq \mathbb{R}^n$ is a set-valued map;
  - $F : \mathbb{R}^n \ni x \rightarrow \mathbb{R}^n$ is a vector function.

• **Generalized Quasi-VI** $QVI(\mathcal{K}, F)$:
  
  - Both $\mathcal{K} : \mathbb{R}^n \ni x \Rightarrow \mathcal{K}(x) \subseteq \mathbb{R}^n$ and $F : \mathbb{R}^n \ni x \Rightarrow F(x) \subseteq \mathbb{R}^n$ are set-valued maps.
Solution Analysis of VI
Existence of a Solution

- **Theorem.** Let $\mathcal{K} \subseteq \mathbb{R}^n$ be compact and convex, and let $F : \mathcal{K} \rightarrow \mathbb{R}^n$ be continuous. Then, the VI$(\mathcal{K}, F)$ has a nonempty and compact solution set.

- Without boundedness of $\mathcal{K}$, the existence of a solution needs some additional properties on the vector function $F$ (as for optimization problems):
  - “Coercivity”
  - Monotonicity/P-properties.
Monotonicity Properties of $F$: Outline

- Monotonicity properties of vector functions.

- Convex programming - a special case: monotonicity properties are satisfied immediately by gradient maps of convex functions.

- In a sense, role of monotonicity in VIs is similar to that of convexity in optimization.

- Existence (uniqueness) of solutions of VIs and convexity of solution sets under monotonicity properties.
Monotonicity Properties of $F$ (I)

• A mapping $F : \mathcal{K} \to \mathbb{R}^n$ is said to be

(i) monotone on $\mathcal{K}$ if

$$(x - y)^T (F(x) - F(y)) \geq 0, \quad \forall x, y \in \mathcal{K}$$

(ii) strictly monotone on $\mathcal{K}$ if

$$(x - y)^T (F(x) - F(y)) > 0, \quad \forall x, y \in \mathcal{K} \text{ and } x \neq y$$

(iii) strongly monotone on $\mathcal{Q}$ if there exists constant $c_{sm} > 0$ such that

$$(x - y)^T (F(x) - F(y)) \geq c_{sm} \|x - y\|^2, \quad \forall x, y \in \mathcal{K}$$

The constant $c_{sm}$ is called strong monotonicity constant.
Monotonicity Properties of $F$ (II)

- Example of (a) monotone, (b) strictly monotone, and (c) strongly monotone functions:

(a) $F(x)$

(b) $F(x)$

(c) $F(x)$
Monotonicity Properties of $F$ (III)

- Relations among the monotonicity properties and connection with the Jacobian matrix $J_F(x)$ of $F$ defined on an open convex set $\mathcal{K}$:

  \[
  \text{strongly monotone} \iff \text{strictly monotone} \iff \text{monotone}
  \]

  \[
  J_F(x) - c I \succeq 0 \iff J_F(x) \succ 0 \iff J_F(x) \succeq 0
  \]

- For an affine map $F = Ax + b$, where $A \in \mathbb{R}^{n \times n}$ (not necessarily symmetric) and $b \in \mathbb{R}^n$, we have:

  \[
  \text{strongly monotone} \iff \text{strictly monotone} \iff A \succ 0
  \]

  \[
  \text{monotone} \iff A \succeq 0
  \]
Monotonicity Properties of $F$ (IV)

- If $F = \nabla f$, the monotonicity properties can be related to the convexity properties of $f$

\[\nabla f \text{ monotone} \iff \nabla^2 f \succeq 0\]
\[\nabla f \text{ strictly monotone} \iff \nabla^2 f \succ 0\]
\[\nabla f \text{ strongly monotone} \iff \nabla^2 f - cI \succeq 0\]
Monotonicity Properties of $F$ (IV)

- If $F = \nabla f$, the monotonicity properties can be related to the convexity properties of $f$

a) $f$ convex $\iff \nabla f$ monotone
b) $f$ strictly convex $\iff \nabla f$ strictly monotone
c) $f$ strongly convex $\iff \nabla f$ strongly monotone
Monotonicity Properties of $F$ (IV)

- If $F = \nabla f$, the monotonicity properties can be related to the convexity properties of $f$

  a) $f$ convex $\iff \nabla f$ monotone $\iff \nabla^2 f \succeq 0$

  b) $f$ strictly convex $\iff \nabla f$ strictly monotone $\iff \nabla^2 f \succ 0$

  c) $f$ strongly convex $\iff \nabla f$ strongly monotone $\iff \nabla^2 f - cI \succeq 0$
Why Are Monotone Mappings Important?

- Arise from important classes of optimization/game-theoretic problems.
- Can articulate existence/uniqueness statements for such problems and VIs.
- Convergence properties of algorithms may sometimes (but not always) be restricted to such monotone problems.
Theorem. Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a convex closed set, and let $F : \mathcal{K} \rightarrow \mathbb{R}^n$ be continuous.

(a) If $F$ is monotone on $\mathcal{K}$, the solution set of $\text{VI}(\mathcal{K}, F)$ is closed and convex (possibly empty).

(b) If $F$ is strictly monotone on $\mathcal{K}$, the $\text{VI}(\mathcal{K}, F)$ has at most one solution.

(c) If $F$ is strongly-monotone on $\mathcal{K}$, the $\text{VI}(\mathcal{K}, F)$ has a unique solution.

(d) If $F$ is Lipschitz continuous and strongly-monotone on $\mathcal{K}$, there exists a $c > 0$ such that

$$\|x - x^*\| \leq c \left\| F_{\mathcal{K}}^{\text{nat}}(x) \right\| \quad \forall x \in \mathcal{K}$$

where $x^*$ is the unique solution of the VI and $F_{\mathcal{K}}^{\text{nat}}(x) \triangleq x - \prod_{\mathcal{K}} (x - F(x))$ (note that $F_{\mathcal{K}}^{\text{nat}}(x^*) = 0$).
• Remarks:

− Strict monotonicity of $F$ does not guarantee the existence of a solution. For example, $F(x) = e^x$ is strictly monotone, but the $\text{VI}(\mathbb{R}, e^x)$ does not have a solution.

− Result in (d) provides an upper bound on the distance from the solution.

− For “partitioned VI” (i.e., $\mathcal{K} = \prod_{i=1}^{Q} \mathcal{K}_i$, with $\mathcal{K}_i \subseteq \mathbb{R}^{n_i}$ and $F = (F_i)^{Q}_{i=1}$, with $F_i : \mathcal{K}_i \rightarrow \mathbb{R}^{n_i}$), the existence and uniqueness results can be made weaker (there are also necessary and sufficient conditions).
Parallelism with Convex Problems

**Theorem.** Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a convex closed set, and let $f : \mathcal{K} \rightarrow \mathbb{R}$ be continuously differentiable ($\Rightarrow F \triangleq \nabla f$ is continuous). Consider the optimization problem

$$(P) : \quad \text{minimize} \quad f(x) \quad \forall x \in \mathcal{K}$$

(a) If $f$ is convex ($\iff F$ is monotone) on $\mathcal{K}$, the solution set of $(P)$ [the VI$(\mathcal{K}, F)$] is closed and convex.

(b) If $f$ is strictly convex ($\iff F$ is strictly monotone) on $\mathcal{K}$, problem $(P)$ [the VI$(\mathcal{K}, F)$] has at most one solution.

(c) If $f$ is strongly convex ($\iff F$ is strongly-monotone) on $\mathcal{K}$, problem $(P)$ [the VI$(\mathcal{K}, F)$] has a unique solution.
Characterization of Solution of VI as a Projection

- The solution of a VI($\mathcal{K}, F$) can be interpreted as the Euclidean projection of a proper map onto the convex set $\mathcal{K}$.

- Let $\mathcal{K} \subseteq \mathbb{R}^n$ be closed and convex and let $F : \mathcal{K} \rightarrow \mathbb{R}^n$ be arbitrary.

\[
  x^* \text{ is a solution of the VI($\mathcal{K}, F$)} \iff x^* = \prod_{\mathcal{K}} (x^* - F(x^*))
\]

- The fixed-point interpretation can be very useful for both analytical and computational purposes:
  - Basic existence results of a solution of the VI come from Brouwer fixed-point theorem applied to $\Phi(x) = \prod_{\mathcal{K}} (x - F(x))$
  - Development of a large family of iterative methods (e.g., projection methods).
• Graphical interpretation:
  
  - (a) $x^*$ is a solution of the $\text{VI}(\mathcal{K}, F)$ if and only if $x^* = \prod_{\mathcal{K}} (x^* - F(x^*))$;
  
  - (b) a feasible $x$ that is not a solution of the $\text{VI}$ and thus $x \neq \prod_{\mathcal{K}} (x^* - F(x^*))$. 

![Graphical representation of (a) and (b) with feasible set $\mathcal{K}$, $F(x^*)$, $x^* = \prod_{\mathcal{K}} (x^* - F(x^*))$, and $\prod_{\mathcal{K}} (x - F(x))$.]
Waterfilling: A New Look at an Old Result

• Capacity of parallel Gaussian channels \( \{\lambda_k\}_{k=1}^{N} \geq 0 \) under average power constraints:

\[
\begin{align*}
\text{maximize} \quad & \sum_{k=1}^{N} \log (1 + p_k \lambda_k) \\
\text{subject to} \quad & \sum_{k=1}^{N} p_k \leq P_T \\
& 0 \leq p_k \leq p_k^{\max}, \quad 1 \leq k \leq N
\end{align*}
\]

• Optimal power allocation: the waterfilling solution

\[
p_k^* = \left[ \mu - \lambda_k^{-1} \right] p_k^{\max}, \quad 1 \leq k \leq N
\]

where \( \mu \) is found to satisfy \( \sum_{k=1}^{N} p_k = P_T \).

• History: Shannon 1949, Holsinger 1964, Gallager 1968.

• For some problems, this waterfilling expression is not convenient.
New Interpretation of Waterfillings via VI

**Theorem** [ScuPhD’05, Luo-Pan’06, Scu-Pal-Bar’07]: The waterfilling \( p^* = \left[ \mu - \lambda^{-1} \right] p^\text{max} \) is the unique solution of the affine VI \((K, F)\), where

\[
K = \left\{ p \in \mathbb{R}^N : \sum_{i=1}^{N} p_k = P_T, \ 0 \leq p_k \leq p_k^\text{max} \text{ for all } k \right\},
\]

and

\[
F(p) = p + \lambda^{-1}.
\]

**Corollary:** The waterfilling solution can be rewritten as a projection

\[
p^* = \Pi_K \left[ -\lambda^{-1} \right]
\]

\[
p^* \in \text{SOL}(K, F) \iff p^* = \Pi_K \left[ p^* - F(p^*) \right]
\]
Algorithms for VI
Algorithms for VI

• Newton methods and globalization.

• Equation-based algorithms for Complementarity Problems.

• KKT-based approaches.

• Merit function-based approaches

• Monotone problems and beyond:
  – Projection methods;
  – Tikhonov and Proximal-point regularization;
  – Jacobi and Gauss-Seidel methods for distributed computation (partitioned VIs).
Projection Methods: Introduction

- Projection methods are conceptually simple methods for solving monotone VI($\mathcal{K}, \mathbf{F}$) for a convex closed set $\mathcal{K}$.

- Their advantages are
  - Easily implementable and computationally inexpensive (when $\mathcal{K}$ has structure that makes the projection onto $\mathcal{K}$ easy)
  - Suitable for large scale problems
  - They are often used as a sub-procedures in faster and more complex methods (enabling the moves into “promising” regions)

- Their main disadvantage is slow progress since they do not use higher order information.
Basic Idea of Projection Algorithms

- **Fact**: Recall that, given $\mathcal{K} \subseteq \mathbb{R}^n$ closed and convex, $\mathbf{x}^*$ is a solution of the VI($\mathcal{K}, F$) if and only if $\mathbf{x}^*$ is a fixed-point of the mapping $\Phi(\mathbf{x}) \triangleq \prod_{\mathcal{K}} (\mathbf{x} - F(\mathbf{x}))$, i.e., $\mathbf{x}^* = \Phi(\mathbf{x}^*)$ [note that $\Phi : \mathcal{K} \rightarrow \mathcal{K}$]

- The above fact motivates the following simple fixed-point based iterative scheme:

$$\mathbf{x}^{(n+1)} = \Phi(\mathbf{x}^{(n)}), \quad \mathbf{x}^{(0)} \in \mathcal{K}$$

which produces a sequence with accumulation points being fixed points of $\Phi$.

- If one could ensure that $\Phi$ is a contraction in some norm, then one could use fixed-point iterates to find a fixed point of $\Phi$ and hence, a solution to VI($\mathcal{K}, F$).
Basic Projection Method

Algorithm 1: Projection algorithm with constant step-size

(S.0) : Choose any \( x^{(0)} \in \mathcal{K} \), and the step size \( \tau > 0 \); set \( n = 0 \).

(S.1) : If \( x^{(n)} = \prod_{\mathcal{K}} (x^{(n)} - F(x^{(n)})) \), then: STOP.

(S.2) : Compute

\[
  x^{(n+1)} = \prod_{\mathcal{K}} \left(x^{(n)} - \tau F(x^{(n)})\right).
\]

(S.3) : Set \( n \leftarrow n + 1 \); go to (S.1).

• In order to ensure the convergence of the sequence \( \{x^{(n+1)}\}_{n=0}^{\infty} \) (or a subsequence) to a fixed point of \( \Phi \), one needs some conditions of the mapping \( F \) and the step size \( \tau > 0 \). (Note that instead of a scalar step size, one can also use a positive definite matrix.)
Basic Projection Method: Convergence Conditions

- **Theorem.** Let $\mathbf{F} : \mathcal{K} \rightarrow \mathbb{R}^n$, where $\mathcal{K} \subseteq \mathbb{R}^n$ is closed and convex. Suppose $\mathbf{F}$ is strongly monotone and Lipschitz continuous on $\mathcal{K}$:

\[
(\mathbf{x} - \mathbf{y})^T(\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})) \geq c_F \|\mathbf{x} - \mathbf{y}\|^2, \\
\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\| \leq L_F \|\mathbf{x} - \mathbf{y}\|
\]

and let

\[
0 < \tau < \frac{2c_F}{L_F^2}.
\]

Then, the mapping $\prod_{\mathcal{K}} (\mathbf{x}^{(n)} - \tau \mathbf{F}(\mathbf{x}^{(n)}))$ is a contraction in the Euclidean norm with contraction factor

\[
\eta = 1 - L_F^2 \tau \left(\frac{2c_F}{L_F^2} - \tau\right).
\]

Therefore, any sequence $\{\mathbf{x}^{(n)}\}_{n=0}^{\infty}$ generated by Algorithm 1 converges linearly to the unique solution of the VI($\mathcal{K}, \mathbf{F}$).
• Pros:
  – Convergence of the *whole* sequence to the unique solution of the VI.
  – Easy to implement (for “nice” $\mathcal{K}$): one step of projection.

• Cons:
  – Strong requirements on $\mathbf{F}$ (strongly monotonicity and Lipschitz property).
  – We do not always have access to $L_{\mathbf{F}}$ and $c_{\mathbf{F}}$. Hence, we do not know how small $\tau$ should be to ensure convergence.

• We can trade the cons with a higher computational complexity; this includes using a variable step-size, using the extra-gradient method, hyperplane projection methods, etc.
Tikhonov Regularization

- The basic idea is to approximate the given VI by a sequence of auxiliary VIs with “better properties” than the original VI.

- The mapping $F$ is approximated with a family of mappings $F_\epsilon$ such that $F_\epsilon \to F$ as $\epsilon \to 0$ (the convergence of the mappings is point-wise). More specifically, $F_\epsilon = F + \epsilon I$.

- When $F$ is monotone on $\mathcal{K}$, $F_\epsilon = F + \epsilon I$ is strongly monotone on $\mathcal{K}$ with constant $\epsilon > 0$.

- Thus, Tikhonov Regularization corresponds to solving a sequence of VIs of the form: given $\epsilon > 0$, find a $x_\epsilon \in \mathcal{K}$ such that

  $$(x - x_\epsilon)^T (F(x_\epsilon) + \epsilon x_\epsilon) \geq 0 \quad \forall x \in \mathcal{K}.$$
Algorithm 2: Exact Tikhonov Regularization

(S.0) : Choose any $x^{(0)} \in \mathcal{K}$, and a sequence $\{\epsilon_n\} \downarrow 0$; set $n = 0$.
(S.1) : If $x^{(n)} \in \text{SOL}(\mathcal{K}, F)$, then: STOP.
(S.2) : Compute $x^{(n+1)}$ as the unique solution of the VI $(\mathcal{K}, F_{\epsilon_n})$:

$$
(x - x^{(n+1)})^T \left( F \left( x^{(n+1)} \right) + \epsilon_n x^{(n+1)} \right) \geq 0 \quad \forall x \in \mathcal{K}
$$

(S.3) : Set $n \leftarrow n + 1$; go to (S.1).

- **Theorem.** Let $F : \mathcal{K} \rightarrow \mathbb{R}^n$ be continuous and monotone over $\mathcal{K}$, and let $\mathcal{K} \subseteq \mathbb{R}^n$ be closed and convex. Let $\text{SOL}(\mathcal{K}, F) \neq \emptyset$. Then, the sequence $\{x^{(n)}\}_{n=0}^{\infty}$ generated by the Exact Tikhonov Regularization converges to the least-norm solution of the VI $(\mathcal{K}, F)$. 
Proximal Point Methods (PPMs)

- The Tikhonov regularization suffers a potential computational drawback: when \( \epsilon \to 0 \) the perturbed VIs approach the original VI and thus it may become more and more difficult to solve them.

- PPMs provide a way of alleviating such difficulty.

- PPMs use a different perturbation than Tikhonov regularization that maintains \textit{uniformly} strong monotonicity of the approximate problems.

- The perturbation is of the form \( F_\tau = F + \tau (I - y) \), where \( \tau > 0 \) and \( y \) are given \([F \text{ monotone} \Rightarrow F_\tau \text{ strongly monotone}]\).

- In PPMs, at step \( n \), the perturbing function is given by \( \tau (x - x^{(n-1)}) \), for some \( \tau > 0 \), instead of \( \epsilon_n x^{(n)} \) (Tikhonov regularization).
Proximal Point Methods (PPMs)

Algorithm 3: Exact Proximal Point Method

(S.0): Choose any \( x^{(0)} \in \mathcal{K} \), let \( \{\rho_n\}_{n=0}^{\infty} \), and \( \tau > 0 \) be given; set \( n = 0 \).

(S.1): If \( x^{(n)} \) satisfies a suitable termination criterion: STOP.

(S.2): Compute \( z^{(n+1)} \) as the unique solution of the \( \text{VI}(\mathcal{K}, F + \tau(I - x^{(n)})) \).

(S.3): Set \( x^{(n+1)} \triangleq (1 - \rho_n)x^{(n)} + \rho_n z^{(n+1)} \).

(S.4): \( n \leftarrow n + 1 \); go to (S.1).

• Theorem. Let \( F : \mathcal{K} \rightarrow \mathbb{R}^n \) be continuous and monotone over \( \mathcal{K} \), let \( \mathcal{K} \subseteq \mathbb{R}^n \) be closed and convex, let \( \text{SOL}(\mathcal{K}, F) \neq \emptyset \) and let \( \rho_n \) be a sequence such that \( \{\rho_n\} \subset [R_m, R_M] \) with \( 0 < R_m \leq R_M < 2 \). Then, the sequence \( \{x^{(n)}\}_{n=0}^{\infty} \) generated by the Exact Proximal Point Method converges to a solution of the \( \text{VI}(\mathcal{K}, F) \).
Why Does Proximal Methods Work Better?

• Intuitively, the benefit of the proximal perturbation is that if the sequence \( \{x^{(n)}\} \) converges, the term \( \tau (x^{(n)} - x^{(n-1)}) \) approaches zero; thus \( \tau \) does not need to go to zero as in Tikhonov regularization \( \epsilon_n x^{(n)} \).

• As a result, we can maintain the “uniformly strong monotonicity” of the perturbed functions \( F + \tau (I - x^{(n-1)}) \) for all \( n \) even when \( F \) is just monotone.

• It turns out that the proximal idea has far reaching applications and leads to developments that are much more than just a modification of the Tikhonov regularization method.
NEP as a VI
NEP as a VI: Solution Analysis and Algorithms

- We have seen that the NEP

\[ \mathcal{G} = \langle \mathcal{K}, \mathbf{f} \rangle : \begin{array}{c}
\text{minimize} \\
\text{subject to}
\end{array} \begin{array}{c}
\sum_{i} f_i(x_i, x_{-i}) \\
x_i \in \mathcal{K}_i \end{array} \forall i = 1, \ldots, Q, \]

is equivalent (under mild conditions) to the VI(\mathcal{K}, \mathbf{F})

- Building on the VI framework, we can now derive conditions for the existence/uniqueness of a NE and devise distributed algorithms along with their convergence properties.

- The state-of-the-art-results are given in [Scu-Fac-Pan-Pal’11(sub)].
NEP as a VI: Existence and Uniqueness of a NE

Theorem. Given the NEP $\mathcal{G} = \langle \mathcal{K}, f \rangle$, suppose that each $\mathcal{K}_i$ is closed and convex, $f_i(x_i, x_{-i})$ is continuously differentiable and convex in $x_i$, for any $x_{-i}$, and let $F(x) \triangleq (\nabla x_i f_i(x))_{i=1}^Q$. Then the following statements hold.

(a) Suppose that for every $i$ the strategy set $\mathcal{K}_i$ is bounded. Then the NEP $[\text{the VI}(\mathcal{K}, F)]$ has a nonempty and compact solution set;

(b) Suppose that $F(x)$ is a strictly monotone function on $\mathcal{K}$. Then the NEP $[\text{the VI}(\mathcal{K}, F)]$ has at most one solution;

(c) Suppose that $F(x)$ is a strongly monotone on $\mathcal{K}$. Then the NEP $[\text{the VI}(\mathcal{K}, F)]$ has a unique solution.
Matrix Conditions for the Monotonicity of $F$

- We provide sufficient conditions for $F$ to be (strictly/strongly) monotone.

- Let introduce the matrices $J_{F_{\text{low}}}$ and $\Upsilon_F$, defined as

$$[J_{F_{\text{low}}}]_{ij} \triangleq \inf_{x \in \mathcal{K}} \begin{cases} \nabla^2_{x_ix_i} f_i(x), & \text{if } i = j, \\ -\frac{1}{2} \left( \left| \nabla^2_{x_ix_j} f_i(x) \right| + \left| \nabla^2_{x_jx_i} f_j(x) \right| \right), & \text{otherwise.} \end{cases}$$

and

$$[\Upsilon_F]_{ij} \triangleq \begin{cases} \alpha^\text{min}_i, & \text{if } i = j, \\ -\beta^\text{max}_{ij}, & \text{otherwise,} \end{cases}$$

where

$$\alpha^\text{min}_i \triangleq \inf_{x \in \mathcal{K}} \lambda_{\text{min}} \left( \nabla^2_{x_ix_i} f_i(x) \right) \quad \text{and} \quad \beta^\text{max}_{ij} \triangleq \sup_{x \in \mathcal{K}} \left\| \nabla^2_{x_jx_i} f_i(x) \right\|.$$
Proposition. [Scu-Fac-Pan-Pal’11(sub)] Let \( F(x) \triangleq (\nabla_{x_i} f_i(x))_{i=1}^Q \) be continuously differentiable with bounded derivatives on \( \mathcal{K} \). Then the following statements hold:

(a) If \( JF_{\text{low}} \) is a (strictly) copositive matrix or \( \Upsilon_F \) is a positive semidefinite (definite) matrix, then \( F \) is monotone (strictly monotone) on \( \mathcal{K} \);

(b) If \( JF_{\text{low}} \) or \( \Upsilon_F \) is a positive definite matrix, then \( F \) is strongly monotone on \( \mathcal{K} \), with strong monotonicity constant \( c_{\text{sm}}(F) \geq \lambda_{\text{least}}(JF_{\text{low}}) \) [or \( c_{\text{sm}}(F) \geq \lambda_{\text{least}}(\Upsilon_F) \)].

- Sufficient conditions for \( \Upsilon_F \succ 0 \) are

\[
\frac{1}{w_i} \sum_{j \neq i} w_j \frac{\beta_{ij}^{\max}}{\alpha_i^{\min}} < 1, \quad \forall i, \quad \frac{1}{w_j} \sum_{i \neq j} w_i \frac{\beta_{ij}^{\max}}{\alpha_i^{\min}} < 1, \quad \forall j.
\]
• Remarks:

– The uniqueness result stated in part (b)-(c) of the Theorem does not require that the set $\mathcal{K}$ be bounded;

– Note that if $\Upsilon_\mathbf{F} \succ 0$ (the NE is unique), it must be $\alpha_{i \min} = \inf_{z \in \mathcal{K}} \left[ \lambda_{\min}(\nabla^2_{x_i} f_i(z)) \right] > 0$ for all $i$. This implies the uniform strong convexity of $f_i(\cdot, x_{-i})$ for all $x_{-i} \in \mathcal{K}_{-i}$ (the optimal solution of each player’s optimization problem is unique);

– The uniqueness conditions are sufficient also for global convergence of best-response asynchronous distributed algorithms described later on.
Algorithms for NEPs

• Algorithms for *strongly monotone* NEPs: Totally asynchronous best-response algorithm [Scu-Fac-Pan-Pal’11(sub)].

• Algorithms for *monotone* NEPs: Proximal Decomposition Algorithms [Scu-Fac-Pan-Pal’11(sub)].
Best-response Algorithms: Basic Idea

- A NE $\mathbf{x}^*$ is defined as a “simultaneous” solution of each of the single-player optimization problems.

- Introducing the set of the optimal solutions to the $i$th optimization problem for any given $\mathbf{x}_{-i}$

$$
\mathcal{B}_i(\mathbf{x}_{-i}) \triangleq \{ \mathbf{x}_i \in \mathcal{K}_i : f_i(\mathbf{x}_i, \mathbf{x}_{-i}) \leq f_i(\mathbf{y}_i, \mathbf{x}_{-i}), \ \forall \mathbf{y}_i \in \mathcal{K}_i \}
$$

also termed as best-response map, the NE of the NEP can be reinterpreted as the fixed-points of the best-response map $\mathcal{B}(\mathbf{x}) \triangleq (\mathcal{B}_i(\mathbf{x}_{-i}))_{i=1}^Q$, i.e.,

$$
\mathbf{x}^* \in \mathcal{B}(\mathbf{x}^*).
$$
• This fact motivates the class of so-called “best-response algorithms”: fixed-point based iterative schemes where at every iteration each player updates his own strategy choosing a point in $B_i(x_{-i})$ (solving his own optimization problem given $x_{-i}$).

• If one could ensure that $B$ is a contraction in some norm, then one could use fixed-point iterates to find a fixed point of $B$ and hence, a solution to the NEP.
Asynchronous Best-Response Algorithm

- In the Asynchronous Best-Response Algorithm, users
  - update their strategy in a totally asynchronous way based on $B_i(x_{-i})$;
  - may update at arbitrary and different times and more or less frequently than others
  - may use an outdated measure of the strategies of the others.
Asynchronous Best-Response Algorithm

Algorithm 4: Asynchronous Best-Response Algorithm

(S.0) Choose any feasible starting point $x^{(0)} = (x_i^{(0)})_{i=1}^Q$; set $n = 0$.

(S.1) If $x^{(n)}$ satisfies a suitable termination criterion: STOP

(S.2) for $i = 1, \ldots, Q$, compute $x_i^{(n+1)}$ as

$$x_i^{(n+1)} = \begin{cases} 
\begin{aligned}
\text{argmin} & \quad f_i \left( x_i, x_{-i}^{(\tau_i(n))} \right), \\
& \quad \text{if } n \in T_i, \\
\end{aligned} \\
\begin{aligned}
x_i^{(n)} , & \quad \text{otherwise}
\end{aligned}
\end{cases}$$

(S.3) $n \leftarrow n + 1$; go to (S.1).
Theorem. [Scu-Fac-Pan-Pal’11(sub)] Given the NEP $\mathcal{G} = \langle \mathcal{K}, \mathbf{f} \rangle$, suppose that $\mathbf{Y}_F \succ 0$. Then, any sequence $\{x^{(n)}\}_{n=0}^{\infty}$ generated by the asynchronous best-response algorithm converges to the unique NE of $\mathcal{G}$, for any given updating feasible schedule of the players.

• Remarks:
  
  – Convergence conditions independent of players’ updating schedule;
  
  – Under $\mathbf{Y}_F \succ 0$, the best-response mapping $B(x)$ is a contraction (or equivalently, the function $\mathbf{F} = (\nabla x_i f_i)_{i=1}^{Q}$ is strongly monotone on $\mathcal{K}$);
  
  – If $f_i(\cdot, x_{-i})$ is not uniformly strongly convex for all $x_{-i} \in \mathcal{K}_{-i}$, (i.e., $\mathbf{Y}_F \not\succ 0$), the algorithm is not guaranteed to converge.

• How to deal with (non strongly) monotone NEPs?
Algorithms for Monotone NEPs: Main Idea

- We know how to solve a strongly monotone NEP. To solve a monotone NEP the proposed approach is then to make it strongly monotone by a proper regularization.

- The regularization has to be chosen so that:
  - one can recover somehow the NE of the original game from those of the regularized game;
  - one should be able to solve a regularized game via distributed algorithms.

- Ingredients:
  - Equivalence between the VI and the NEP;
  - Proximal decomposition algorithms for monotone VIs.
How to Make a Monotone NEP Strongly Monotone

• Given a monotone NEP $G = <\mathcal{K}, f>$, the NEP is equivalent to the monotone $\text{VI}(\mathcal{K}, F)$, with $F = (\nabla x_i f_i)_{i=1}^Q$.

• Consider the proximal regularization of the VI, given by $\text{VI}(\mathcal{K}, F + \tau (I - y))$, where $I : x \rightarrow x$ is the identity map, $y \in \mathbb{R}^n$ is a fixed vector and $\tau$ is a positive constant.

• $\text{VI}(\mathcal{K}, F + \tau (I - y))$ is strongly monotone and thus admits a unique solution, denoted by $S_{\tau}(y)$.

• The fixed-points $y^* = S_{\tau}(y^*)$ coincide with $\text{SOL}(\mathcal{K}, F)$ [and thus with the solution set of the monotone NEP $G = <\mathcal{K}, f>$].

• We can then compute the fixed-points of $S_{\tau}(\cdot)$ by the Proximal Decomposition Algorithm described before.
### Prox Decomposition Algorithms for Monotone NEP

**Algorithm 5: Exact Proximal Point Method for Monotone NEPs**

1. **(S.0)**: Choose any \( x^{(0)} \in \mathcal{K} \), let \( \{\rho_n\}_{n=0}^{\infty} \), and \( \tau > 0 \) be given; set \( n = 0 \).

2. **(S.1)**: If \( x^{(n)} \) satisfies a suitable termination criterion: STOP.

3. **(S.2)**: Solve \( \text{VI}(\mathcal{K}, \mathbf{F} + \tau(\mathbf{I} - x^{(n)})) \); let \( z^{(n+1)} = S_\tau(x^{(n)}) \).

4. **(S.3)**: Set \( x^{(n+1)} \triangleq (1 - \rho_n)x^{(n)} + \rho_nz^{(n+1)} \).

5. **(S.4)**: \( n \leftarrow n + 1 \); go to (S.1).

---

- **Theorem. [Scu-Fac-Pan-Pal’11(sub)]** Let \( \mathcal{G} = \langle \mathcal{K}, \mathbf{f} \rangle \) be a convex and monotone NEP. Let \( \{\rho_n\}_n \) be such that each \( \rho_n \subset [R_m, R_M] \) with \( 0 < R_m \leq R_M < 2 \). Then, the sequence \( \{x^{(n)}\}_{n=0}^{\infty} \) generated by the Exact Proximal Point Method converges to a NE of the game \( \mathcal{G} \).
• The key point becomes how to compute the unique solution \( S_{\tau}(x^{(n)}) \) of the regularized VI \( \mathcal{VI}(K, F + \tau(I - x^{(n)})) \) in a distributed way.

• \( \mathcal{VI}(K, F + \tau(I - x^{(n)})) \) is equivalent to the following regularized **strongly monotone** game

\[
\mathcal{G}_{x^{(n)}} : \begin{align*}
\min_{x_i} \quad & f_i(x_i, x_{-i}) + \frac{\tau}{2} \| x_i - x_i^{(n)} \|^2 \\
\text{subject to} \quad & x_i \in K_i, \quad \forall i = 1, \ldots, Q.
\end{align*}
\]

• It follows that:
  
  – \( S_{\tau}(x^{(n)}) \) is the unique NE of the strongly monotone \( \mathcal{G}_{x^{(n)}} \);
  
  – \( S_{\tau}(x^{(n)}) \) can be computed using the asynchronous best-response algorithms for strongly monotone NEPs; the global convergence is guaranteed if \( \tau \) is sufficiently large (bigger than a known quantity).
GNEP as a VI
GNEP: Basic Definitions

- The GNEP extends the classical NEP setting by assuming that each player’s strategy set can depend on the rival players’ strategies $x_{-i}$.

- Let $\mathcal{K}_i(x_{-i}) \subseteq \mathbb{R}^{n_i}$ be the feasible set of player $i$ when the other players choose $x_{-i}$. The aim of each player $i$, given $x_{-i}$, is to choose a strategy $x_i \in \mathcal{K}_i(x_{-i})$ that solves the problem

$$
\begin{align*}
\text{minimize} & \quad f_i(x_i, x_{-i}) \\
\text{subject to} & \quad x_i \in \mathcal{K}_i(x_{-i}).
\end{align*}
$$

- A Generalized Nash Equilibrium (GNE) is a feasible point $x^*$ such that the following holds for each player $i = 1, \ldots, Q$:

$$
f_i(x^*_i, x^*_{-i}) \leq f_i(x_i, x^*_{-i}), \quad \forall x_i \in \mathcal{K}_i(x^*_{-i}).$$
• Facts:

– The GNEP can be rewritten as a QVI;
– However, QVIs are much harder problems than VIs and only few results are available;
– Thus the GNEP, in its full generality, is almost intractable and also the VI approach does not help.

• We then restrict our attention to particular classes of (more tractable) equilibrium problems: the so-called GNEPs with jointly convex shared constraints.
A GNEP is termed as **GNEP with jointly convex shared constraints (JCSC)** if the feasible sets are defined as:

\[ \mathcal{K}_i(x_{-i}) \triangleq \left\{ x_i \in \overline{\mathcal{K}}_i : g(x_i, x_{-i}) \leq 0 \right\} \]

where:

- \( \overline{\mathcal{K}}_i \subseteq \mathbb{R}^{n_i} \) is the closed and convex set of individual constraints of player \( i \);
- \( g(x_i, x_{-i}) \leq 0 \) represents the set of shared coupling constraints (equal for all the players);
- \( g \triangleq (g_j)_{j=1}^m : \mathbb{R}^n \rightarrow \mathbb{R}^m \) is (jointly) convex in \( x \).

If there are no shared constraints the GNEP reduces to a NEP.
Geometrical Interpretation

- Let define the joint feasible set $\mathcal{K}$

$$
\mathcal{K} = \{ \mathbf{x} : g(x_i, x_{-i}) \leq 0, \ x_i \in \overline{\mathcal{K}_i}, \ \forall i = 1, \ldots, Q \}
$$

- It easy to see that

$$
\mathcal{K}_i(x_{-i}) = \{ x_i : (x_i, x_{-i}) \in \mathcal{K} \}.
$$
Connection Between GNEPs with JCSC and VIs

- GNEPs with JCSC are still very difficult problems, however at least some type of solutions can be studied and calculated using the VI approach.

- Given the GNEP with JCSC, let

\[ \mathcal{K} = \{ x : g(x_i, x_{-i}) \leq 0, \ x_i \in \overline{K}_i, \ \forall i \} \]

and consider the VI \((\mathcal{K}, F)\).

- **Lemma.** Every solution of the VI \((\mathcal{K}, F)\) is a solution of the GNEP with JCSC; the converse in general may not be true.

- The solutions of the GNEP that are also solution of the VI are termed as *Variational Equilibria (VE)*.
VEs: Game Theoretical Interpretation

• Recall that the VEs are solution of the VI, with

\[ \mathcal{K} = \left\{ \mathbf{x} : g(x_i, x_{-i}) \leq 0, \quad x_i \in \overline{\mathcal{K}}_i, \forall i \right\}, \quad \text{and} \quad F = (\nabla_x f_i)_{i=1}^Q. \]

• Connection between the VI(\(\mathcal{K}, F\)) and NEPs: \(\mathbf{x}^* \in \mathcal{K}\) is a VE if and only if \(\mathbf{x}^*\) along with a suitable \(\lambda^*\) is a solution of the NEP with pricing

\[ \mathcal{G}_{\lambda^*} : \quad \begin{array}{c}
\text{minimize} \\
\mathbf{x}_i
\end{array} \quad f_i(x_i, x_{-i}) + \lambda^{*T} g(x) \]

subject to \(x_i \in \overline{\mathcal{K}}_i \quad \forall i = 1, \ldots, Q, \)

and furthermore

\[ 0 \leq \lambda^* \perp g(x^*) \leq 0. \]
• Remarks:
  – We can interpret the $\lambda^*$ as prices paid by the players for the common “resource” represented by the shared constraints;
  – The complementarity condition says that we actually have to pay only when the resource becomes scarce;
  – Thus, the NEP with pricing can be seen as a “penalized” version of the GNEP with JCSC, where the shared constraints are enforced by making the players to pay the common price $\lambda^*$;
  – Mathematically, $\lambda^*$ is the KKT common multiplier of the shared constraints.

• We are now able to reduce the solution analysis & computation of a VE to that of the equilibrium of a NEP, to which we can in principle apply the theory developed so far.
VEs: Solution Analysis and Algorithms

• **Solution analysis**: Since the VEs are solution of a VI, one can derive existence and uniqueness conditions from the VI theory developed so far.

• **Algorithms**: Similarly, we can also devise algorithms for VEs based on the $\text{VI}(\mathcal{K}, F)$; however they will not be distributed since the set $\mathcal{K}$ does not have a Cartesian structure (there is a coupling among the strategies of the players).

• **How to attack the problem**: Building on the equivalence between the $\text{VI}(\mathcal{K}, F)$ and the NEP with pricing, we can overcome this issue. To do that, however, we still need some more work.
Toward Distributed Algs: A NCP Reformulation

- We rewrite the NEP with pricing as a VI whose feasible set has a Cartesian structure. For the sake of simplicity, we focus only on strongly monotone games $\mathcal{G}_F \succ 0$.

- **Step 1:** The NEP with pricing can be rewritten as

  $$
  G_\lambda : \quad \text{VI}(\overline{K}, F + \nabla g^T \lambda) \\
  0 \leq \lambda \perp g(x) \leq 0
  $$

  where $\overline{K} = \prod_{i=1}^Q \overline{K}_i$ and $F = (\nabla x_i f_i)_{i=1}^Q$.

- The $\text{VI}(\overline{K}, F + \nabla g^T \lambda)$ is strongly monotone and thus has a unique solution $x^*(\lambda)$ [the unique NE of $G_\lambda$].
• **Step 2**: We rewrite $\mathcal{G}_\lambda \cup \mathcal{CC}$ as a NCP. Let define the map

$$\Phi(\lambda) : \mathbb{R}_+^m \ni \lambda \to -g(x^*(\lambda))$$

which measures the violation of the shared constraints at $x^*(\lambda)$.

• **Theorem. [Scu-Fac-Pan-Pal’11(sub)]**  If $\Upsilon_F \succ 0$, the (strongly monotone) NEP with pricing in $\mathcal{G}_\lambda \cup \mathcal{CC}$ is equivalent to the NCP in the price tuple $\lambda$

$$\text{NCP}(\Phi) : \quad 0 \leq \lambda \perp \Phi(\lambda) \geq 0 \iff \mathcal{VI}(\mathbb{R}_+^m, \Phi).$$

• The NCP reformulation is instrumental to devise distributed algorithms. We can now use the algorithms developed so far for strongly monotone VIs.
Algorithm 6: Projection Algorithm with Variable Steps (PAVS)

(S.0): Choose any $\lambda^{(0)} \geq 0$; set $n = 0$.

(S.1): If $\lambda^{(n)}$ satisfies a suitable termination criterion: STOP.

(S.2): Given $\lambda^{(n)}$, compute $x^*(\lambda^{(n)})$ as the unique NE of $G_{\lambda^{(n)}}$:

$$x^*(\lambda^{(n)}) = \text{SOL}(\overline{K}; F + \nabla_x g_{\lambda^{(n)}}).$$

(S.3): Choose $\tau_n > 0$ and update the price vectors $\lambda$ according to

$$\lambda^{(n+1)} = \left[ \lambda^{(n)} - \tau_n \Phi \left( \lambda^{(n)} \right) \right]^+. $$

(S.4): Set $n \leftarrow n + 1$; go to (S.1).
• **Theorem. [Scu-Fac-Pan-Pal’11(sub)]** Suppose $\Upsilon_F \succ 0$. If the scalars $\tau_n$ are chosen so that $0 < \inf_n \tau_n \leq \sup_n \tau_n < 2 c_{\text{coc}}(\Phi)$, where $c_{\text{coc}}(\Phi) \triangleq \hat{c}_{\text{sm}}(F)/c_{\text{Lip}}^2(g)$, $c_{\text{Lip}}(g) \triangleq \max_{x \in \mathcal{K}} \|\nabla g(x)^T\|_2$, and $\hat{c}_{\text{sm}}(F)$ is the strongly-monotonicity constant of $F$, then the sequence $\{\lambda^{(n)}\}_{n=0}^{\infty}$ generated by the PAVS converges to a solution of the NCP($\Phi$).

• **Inner loop:** The NE $p^*(\lambda^{(n)})$ of $G_{\lambda^{(n)}}$ can be computed using the asynchronous best-response algorithms (convergence is guaranteed under $\Upsilon_F \succ 0$).

• **Algorithms for Monotone Games:** Following the same idea as for monotone NEPs we can make the monotone NEP with pricing strongly-monotone via Proximal regularization [Scu-Fac-Pan-Pal’11(sub)].
Part III: Variational Inequalities: Applications
Part III - Outline

• Application of NEP: Ad-Hoc Networks

• Application of GNEP: QoS Ad-Hoc Networks

• Application of NEP: Robust CR Systems with Individual Constraints

• Application of GNEP with Shared Constraints: Cognitive Radio Systems

• Application of GNEP with Shared Constraints: Routing in Communication Networks

• Other applications: smart grids, multi-portfolio optimization in finance, etc.
Application of NEP: Ad-Hoc Networks
Competitive Ad-Hoc Networks

- Consider a decentralized and competitive network of users fighting for the resources (i.e., spectrum):

\[
y_q = H_{qq} x_q + \sum_{r \neq q} H_{qr} x_r + w_q.
\]

- Baseband signal model: Vector Gaussian Interference Channel (IC)
The Frequency-Selective Case
• The channel matrices are diagonal: $H_{qr} = \text{diag}\left\{ (H_{qr}(k))_{k=1}^N \right\}$.

• The optimization variables correspond to the power allocation over the carriers: $p_q = \{p_q(k)\}_{k=1}^N$.

• There is a power budget for each user:

$$\sum_{k=1}^N p_q(k) \leq P_q.$$

• The payoff for user $q$ is the information rate:

$$r_q(p_q, p_{-q}) = \sum_{k=1}^N \log (1 + \sin r_q(k)) \quad \text{with} \quad \sin r_q(k) = \frac{|H_{qq}(k)|^2 p_q(k)}{1 + \sum_{r\neq q} |H_{qr}(k)|^2 p_r(k)}.$$

• The feasible set for the variables is: $\mathcal{P}_q = \left\{ p_q \in \mathbb{R}_+^N : \sum_{k=1}^N p_q(k) = P_q \right\}$.
Game Formulation for Ad-Hoc Networks

- Each of the $Q$ users selfishly maximizes its own rate subject to the constraints:

$$
\begin{align*}
G : & \quad \max_{p_q} \quad \sum_{k=1}^{N} \log (1 + \sin r_q(k)) \\
& \quad \text{subject to} \quad p_q \in \mathcal{P}_q \\
& \quad q = 1, \ldots, Q
\end{align*}
$$

- Given the strategies of the others $p_{-q}$, the best response for each user is the waterfilling solution:

$$
p_q^* = \text{wf}_q(p_{-q}) \triangleq (\mu_q - \text{interf}_q(p_{-q}))^+
$$

where

$$
\text{interf}_q(k; p_{-q}) \triangleq \frac{1 + \sum_{r \neq q} |H_{qr}(k)|^2 p_r(k)}{|H_{qq}(k)|^2} \quad k = 1, \ldots, N
$$
• Any NE is a simultaneous waterfilling for all users:

\[
p_q^* = \text{wf}_q (p_{-q}^*) \quad \forall q = 1, \ldots, Q \iff p^* = \text{wf} (p^*).
\]

• Main issues:

- Does a NE exist?
- Is the NE unique?
- How to reach a NE?

• Open problem for years: Different researchers have actively worked on this problem since 2001 [Yu-Gin-Ciof’01], [Yam-Luo’04], [ScuPhD’05], [Luo-Pan’06], [Scu-Pal-Bar’06], …

• VI provides the answers.
How to Attack This Problem?

- One option is to write the NEP using the minimum principle as a VI and try to show monotonicity properties for the function $F$.

- A more elaborate option is to write the KKT conditions of the NEP and, after some manipulations, rewrite the NEP as an Affine VI and try to show monotonicity properties.

- In this problem, since the best-response of the NEP can be written in closed-form, we could try to show a contraction mapping property.

- Interestingly, the fixed-point characterization of the solution of the AVI happens to be the alternative simple representation of the waterfilling solution as a projection.

- This results in showing some norm property on a matrix that is equivalent to showing strongly monotonicity of the AVI.
Why is so Difficult Studying This Game

- We need to prove that the waterfilling mapping is a contraction:

\[
\|\text{wf} \left( p^{(1)} \right) - \text{wf} \left( p^{(2)} \right) \| \leq c \| p^{(1)} - p^{(2)} \|
\]

with

\[
\text{wf} \left( p^{(1)} \right) = \left( \mu^{(1)} - M p^{(1)} \right)^+ \\
\text{wf} \left( p^{(2)} \right) = \left( \mu^{(2)} - M p^{(2)} \right)^+
\]

- Note that

\[
\mu^{(1)} = \mu \left( p^{(1)} \right) \quad \text{and} \quad \mu^{(2)} = \mu \left( p^{(2)} \right)
\]
Why is so Difficult Studying This Game

- We need to prove that the waterfilling mapping is a contraction:

\[ \| \text{wf} \left( p^{(1)} \right) - \text{wf} \left( p^{(2)} \right) \| \leq c \| p^{(1)} - p^{(2)} \| \]

- Using the definition of the wf:

\[ \| \text{wf} \left( p^{(1)} \right) - \text{wf} \left( p^{(2)} \right) \| = \left\| \left( \mu^{(1)} - Mp^{(1)} \right)^+ - \left( \mu^{(2)} - Mp^{(2)} \right)^+ \right\| \]

\[ \leq \left\| \left( \mu^{(1)} - \mu^{(2)} \right) - \left( Mp^{(1)} - Mp^{(2)} \right) \right\| \leq 0 \]

We are stuck !!!
Why is so difficult studying this game?

• We need to prove that the waterfilling mapping is a contraction:

\[ \| w_f \left( p^{(1)} \right) - w_f \left( p^{(2)} \right) \| \leq c \left| p^{(1)} - p^{(2)} \right| \]

• Using the interpretation of the \( w_f \) as a solution of the AVI:

\[ w_f \left( p^{(1)} \right) = \Pi_K \left( -M p^{(1)} \right) \]

\[ w_f \left( p^{(2)} \right) = \Pi_K \left( -M p^{(2)} \right) \]
Why is so difficult studying this game?

• We need to prove that the waterfilling mapping is a contraction:

\[ \| \text{wf} \left( p^{(1)} \right) - \text{wf} \left( p^{(2)} \right) \| \overset{?}{\leq} c \| p^{(1)} - p^{(2)} \| \]

• Using the interpretation of the \text{wf} as a solution of the AVI:

\[ \| \text{wf} \left( p^{(1)} \right) - \text{wf} \left( p^{(2)} \right) \| = \| \Pi_K \left( -Mp^{(1)} \right) - \Pi_K \left( -Mp^{(2)} \right) \| \]
\[ \leq \left\| M \left( p^{(2)} - p^{(1)} \right) \right\| \]
\[ \leq \| M \| \left\| p^{(2)} - p^{(1)} \right\| \]

We are done !!!
Existence and Uniqueness of the NE

• **Theorem [Scu-Pal-Bar’06]**: The solution set of the game is nonempty and compact. The NE is unique if

\[ \rho(H^{\text{max}}) < 1 \]

where

\[ [H^{\text{max}}]_{qr} \triangleq \begin{cases} \max_k \left\{ \frac{|H_{qr}(k)|^2}{|H_{qq}(k)|^2} \right\} & \text{if } q \neq r \\ 0 & \text{otherwise.} \end{cases} \]

• **Sufficient conditions**:  

**Low MUI** :  

\[ \sum_{r \neq q} \max_k \left\{ \frac{|H_{qr}(k)|^2}{|H_{qq}(k)|^2} \right\} < 1, \quad \forall q = 1, \ldots, Q. \]
State-of-the-Art Algorithms: Asynchronous IWFA

- Do the players reach an equilibrium if every one selfishly performs his waterfilling solution $wf_q (p_{-q})$ against the others?

- **Asynchronous IWFA** [Scu-Pal-Bar’06]: Users update the power allocation in a totally asynchronous way based on $wf_q (\cdot)$
  
  - users may update at arbitrary and different times and more or less frequently than others
  - users may use an outdated measure of interference
  - distributed implementation: local measures of the MUI & weak constraints on synchronization

- **Theorem** [Scu-Pal-Bar’06]: The asynchronous IWFA converges to the unique NE if $\rho (H^{\text{max}}) < 1$. 

Convergence of the IWFA

- Simultaneous IWFA vs. Sequential IWFA [under $\rho(\mathbf{H}^{\text{max}}) < 1$]
IWFA vs. Proximal Response Algorithm

- What happens when $\rho(H_{\text{max}}) > 1$?
IWFA vs. Proximal Response Algorithm

- What happens when $\rho(H_{\text{max}}) > 1$?
The MIMO Case
Ad-Hoc Networks - The MIMO Game

• Game theoretic formulation [Scu-Pal-Bar’08]:

\[
\max_{Q_q} \log \det \left( I + Q_q H_{qq}^\dagger R_{-q}^{-1}(Q_{-q}) H_{qq} \right)
\]

subject to

\[
Q_q \succeq 0, \quad \text{Tr} (Q_q) \leq P_q
\]

where \( R_{-q}(Q_{-q}) = R_{n_q} + \sum_{r \neq q} H_{rq} Q_r H_{rq}^\dagger. \)

• Can we similarly analyze this MIMO formulation?

• The answer is affirmative when the channel matrices are square and invertible.
Difficulties in the MIMO Case

- However, the answer is negative for the general case!!

- Some expected conjectures do not hold when the channel matrices are not square:
  - the WF as a projection does not follow from the square case simply replacing the inverse with some generalized inverse
  - the VI cannot be rewritten as a AVI and the reverse order law for generalized inverses does not hold
  - the multiuser WF is not a contraction under the conditions valid for the square case.

- Full characterization of the game for arbitrary MIMO channels (not square, not invertible) [Scu-Pal-Bar’09]:
  - Solution analysis: Existence and uniqueness of the NE
  - Distributed algorithms: Asynchronous MIMO IWFA.
Application of GNEP: QoS Ad-Hoc Networks
NEP Formulation of Ad-Hoc Networks

• The previous ad-hoc game formulation is based on each user maximizing his own utility function subject to local constraints:

\[
\begin{align*}
\text{maximize} & \quad \sum_{k=1}^{N} \log (1 + \sinr_{q}(k)) \\
\text{subject to} & \quad \sum_{k=1}^{N} p_{q}(k) \leq P_{q} \quad \forall q = 1, \ldots, Q
\end{align*}
\]

• The coupling among the users is only through the objective (not the constraints) since the SINR depends on the actions of all the other users:

\[
\sinr_{q}(k) = \frac{|H_{qq}(k)|^2 p_{q}(k)}{1 + \sum_{r \neq q} |H_{qr}(k)|^2 p_{r}(k)}.
\]
GNEP Formulation of QoS Ad-Hoc Networks

• Consider now a small and innocent variation of the previous game where we swap the role of the objective and power constraint:

\[
\begin{align*}
\text{minimize} & \quad \sum_{k=1}^{N} p_q(k) \\
\text{subject to} & \quad \sum_{k=1}^{N} \log (1 + \sin r_q(k)) \geq R_q \\
\end{align*}
\]

\[
\forall q = 1, \ldots, Q
\]

• In this formulation, each user is guaranteed a minimum Quality-of-Service (QoS) in terms of rate \( R_q \), while minimizing the transmitted power.

• From a mathematical perspective, this game is not a NEP anymore because now the constraint set is coupled among the users: it is now a Generalized NEP (GNEP).
• But it is not a GNEP with shared constraints...

• The analysis of this GNEP becomes very challenging.

• Even for the study of the existence & uniqueness of the Generalized NE (GNE) one needs to resort to very sophisticated mathematical tools like degree theory.

• However, it can be analyzed and distributed algorithms can be devised [Pan-Scu-Fac-Wan’08]:
  – the best response is still a waterfilling
  – the IWFA can still be used.
Convergence of the IWFA

The figure illustrates the convergence of the Iteratively Weighted Fenchel Averaging (IWFA) algorithm for different users under both simultaneous and sequential update strategies. The y-axis represents the rates of the algorithms, while the x-axis shows the number of iterations. The red dashed line represents the simultaneous IWFA, and the black solid line represents the sequential IWFA. The figure highlights the convergence behavior for users #1, #25, and #50.
Application of NEP: CR Systems with Individual Constraints
CR Systems

• Consider now an established network of primary users on top of which some secondary users play the previous game.

• Hierarchical CR networks
  – PU=Primary users (legacy spectrum holders)
  – SU=Secondary users (unlicensed users).

• Opportunistic communications: SUs can use the spectrum subject to inducing a limited interference or no interference at all on PUs.
Standard Waterfilling Does Not Work

- The standard iterative waterfilling algorithm does not work because it violates the interference constraint:

![Graph showing PSD of interference at the primary receiver with interference limit and constraint violated points marked.](image)
Signal Model for CR Systems

• Same signal model as for ad-hoc networks with the additional \textit{individual} per-carrier interference constraints:

\[
\left|h_{qp}^{(P,S)}(k)\right|^2 p_q(k) \leq I_{pq}(k), \quad k = 1, \ldots, N \quad p = 1, \ldots, P
\]

− \(|h_{qp}^{(P,S)}(k)|^2\) is the cross-channel gain between the \(q\)th secondary and the \(p\)th primary user

− \(I_{pq}(k)\) is the maximum level of interference tolerable by the primary user \(p\) from the secondary user \(q\) over the subchannel \(k\).

• Equivalently, we can write these constraints in terms of mask constraints

\[
p_q(k) \leq p_{q}^{\text{max}}(k) \triangleq \min_{p=1,\ldots,P} \frac{I_{pq}(k)}{|h_{qp}^{(P,S)}(k)|^2} \quad k = 1, \ldots, N.
\]
Game Formulation for CR Systems

• Each of the $Q$ users selfishly maximizes its own rate subject to the constraints:

$$\begin{align*}
\text{maximize} & \quad \sum_{k=1}^{N} \log (1 + \sinr_q(k)) \\
\text{subject to} & \quad \sum_{k=1}^{N} p_q(k) \leq P_q \\
& \quad 0 \leq p_q(k) \leq p_{q\max}(k), \quad \forall k \quad \forall q = 1, \ldots, Q
\end{align*}$$

• The best response in this case also has a nice and simple closed-form expression based on a modified waterfilling with clipping from above:

$$\tilde{\text{wf}}_q (p_{-q}) \triangleq [\mu_q - \text{interf}_q(p_{-q})]_0^{p_{q\max}}.$$

• Similar analysis and algorithms as for the previous game, based on our interpretation of $\text{wf}$ as a projection [Scu-Pal-Bar’06].
Application of NEP: Robust CR Systems with Individual Constraints
Robust Formulation of CR Systems

• In practice, it is difficult to have a very good estimate of the cross-channels between the secondary and primary users $h_{qp}^{(P,S)}(k)$.

• To make the design robust to uncertainties in these cross-channels, we will assume that the estimate contains some error $e_{qp}(k)$:

$$\hat{h}_{qp}^{(P,S)}(k) = h_{qp}^{(P,S)}(k) + e_{qp}(k)$$

where $e_{qp}$ is some unknown error bounded in norm: $\|e_{qp}\| \leq \varepsilon_{qp}$. 
Therefore, the original non-robust interference constraints become

\[
\left| h_{qp}(P,S)(k) \right|^2 p_q(k) \leq I_{pq}^{\text{peak}}(k) \quad \forall k = 1, \ldots, N
\]

\[
\sum_{k=1}^{N} \left| h_{qp}(P,S)(k) \right|^2 p_q(k) \leq I_{pq}^{\text{ave}}
\]

\[
\left| h_{qp}(P,S)(k) + e_{qp}(k) \right|^2 p_q(k) \leq I_{pq}^{\text{peak}}(k) \quad \forall k = 1, \ldots, N
\]

\[
\sum_{k=1}^{N} \left| h_{qp}(P,S)(k) + e_{qp}(k) \right|^2 p_q(k) \leq I_{pq}^{\text{ave}}
\]

\[
\forall e_{qp} \in \mathcal{D}_{qp} = \{ e_{qp} : \| e_{qp} \| \leq \varepsilon_{qp} \}.
\]
• The robust formulation leads to a NEP where each individual optimization problem contains an infinite number of constraints!

• This robust feasible set can be represented by a finite number of constraints after using the \textit{S-lemma} [Wan-Scu-Pal’10].

• The robust (convex) game formulation is, for each $q = 1, \ldots, Q$:

$$\begin{align*}
\text{maximize} \quad & \sum_{k=1}^{N} \log (1 + \sin r_q(k)) \\
\text{subject to} \quad & \sum_{k=1}^{N} p_q(k) \leq P_q \\
& 0 \leq p_q(k) \leq p_q^{\text{max}}(k), \quad \forall k \\
& \sum_{k=1}^{N} \frac{\mu_{qp} w_{qp}(k) |\hat{H}_{qp}^{(P,S)}(k)|^2}{\mu_{qp} w_{qp}(k) - p_q(k)} p_q(k) + \mu_{qp} \varepsilon_{qp}^2 \leq I_{pq}^{\text{ave}} \quad \forall p \\
& p_q(k) \leq \mu_{qp} w_{qp}(k) \quad \forall k, p \\
& \left( |\hat{h}_{qp}^{(P,S)}(k)| + \varepsilon_{qp}/\sqrt{w_{qp}(k)} \right)^2 p_q(k) \leq I_{pq}^{\text{peak}}(k) \quad \forall k, p
\end{align*}$$
Comparison of the robust and non-robust power allocation:
Application of GNEP with Shared Constraints: Cognitive Radio Systems
Signal Model for CR Systems

- Same signal model as for ad-hoc networks with the additional individual per-carrier interference constraints:

\[ |H_{qp}^{(P,S)}(k)|^2 p_q(k) \leq I_{pq}(k), \quad k = 1, \ldots, N \quad p = 1, \ldots, P \]

- \( |H_{qp}^{(P,S)}(k)|^2 \) is the cross-channel gain between the \( q \)th secondary and the \( p \)th primary user
- \( I_{pq}(k) \) is the maximum level of interference tolerable by the primary user \( p \) from the secondary user \( q \) over the subchannel \( k \).

- Equivalently, we can write these constraints in terms of mask constraints

\[ p_q(k) \leq p_q^{\text{max}}(k) \overset{\Delta}{=} \min_{p=1,\ldots,P} \frac{I_{pq}(k)}{|H_{qp}^{(P,S)}(k)|^2} \quad k = 1, \ldots, N. \]
• The interference constraints are now satisfied:

However... this method may be too conservative as the level of interference from each secondary user is limited *individually* in a conservative way.
Revised Signal Model for CR Systems

- The really important quantity is not the individual interference generated by each secondary user but the aggregate interference generated by all of them.

- We can then limit the aggregate interference:
  - per-carrier constraints:
    $$\sum_{q=1}^{Q} \left| H_{qp}^{(P,S)}(k) \right|^2 p_q(k) \leq I_p(k), \quad p = 1, \ldots, P, \quad k = 1, \ldots, N$$
  - interference-temperature limits
    $$\sum_{k=1}^{N} \sum_{q=1}^{Q} \left| H_{qp}^{(P,S)}(k) \right|^2 p_q(k) \leq I_p(k), \quad p = 1, \ldots, P$$
Proposed game theoretical formulation [Pan-Scu-Pal-Fac’10]:

\[ G : \max_{p_q \geq 0} \sum_{k=1}^{N} \log (1 + \sin r_q(k)) \]
subject to
\[ \sum_{k=1}^{N} p_q(k) \leq P_q \]
\[ \sum_{q=1}^{Q} \left| H_{qp}^{(P,S)}(k) \right|^2 p_q(k) \leq I_p(k), \quad p = 1, \ldots, P, \quad k = 1, \ldots, N. \]
• Indeed, this new reformulation achieves our goal without accidentally violating the limit:
or being too conservative:
How to Deal with the Coupling Constraint

• The price to pay for including the coupling constraints is twofold:
  
  – on a mathematical level, it complicates the analysis of the game and its design \( \Rightarrow \) no previous GT results can be used
  
  – on the practical side, this new game must include some mechanism to calculate the aggregate interference.

• For the mathematical analysis and design, we need more advance tools: VI theory for the analysis of GNEP with shared constraints.
Formulation of the Game with Coupling Constraint

• Recall the game formulation with the coupling constraint [Pan-Scu-Pal-Fac’10]:

\[ \mathcal{G} : \]

\[
\max_{p_q \geq 0} \quad r_q(p_q, p_{-q}) \triangleq \sum_{k=1}^{N} \log \left(1 + \sin r_q(k)\right) \\
\text{subject to} \quad \sum_{k=1}^{N} p_q(k) \leq P_q \\
\sum_{q=1}^{Q} \left| H_{qp}^{(P,S)}(k) \right|^2 p_q(k) \leq I_p(k), \quad p = 1, \ldots, P, \quad k = 1, \ldots, N.
\]

• This is a GNEP with shared constraints

• It can be “rewritten” as a VI problem (caveat: only the variational solutions with common multipliers are considered).
Formulation of the Game with Coupling Constraint via Pricing

• Consider a game with pricing:

\[ G_\lambda : \max_{p,q \geq 0} r_q(p_q, p-q) - \sum_{p=1}^{P} \sum_{k=1}^{N} \lambda_{p,k} \left| H_{qp}^{(P,S)}(k) \right|^2 p_q(k) \]

subject to \[ \sum_{k=1}^{N} p_q(k) \leq P_q \]

where the prices \( \lambda_{p,k} \geq 0 \) are chosen such that

\[ (CC) : 0 \leq \lambda_{p,k} \perp I_p(k) - \sum_{q=1}^{Q} \left| H_{qp}^{(P,S)}(k) \right|^2 p_q(k) \geq 0 \quad \forall p, \forall k \]

• Parametric NEP & CC outside: classical GT fails.

• We will now rewrite the game \( G \triangleq G_\lambda \cup (CC) \) as a VI problem.
• **Theorem (Game as a VI) [Pan-Scu-Pal-Fac’10]:** The game $\mathcal{G}$ is equivalent to the VI$(\mathcal{K}, F)$ where

\[
\mathcal{K} \triangleq \left\{ p \in \mathbb{R}_{+}^{NQ} : \sum_{k=1}^{N} p_q(k) \leq P_q \text{ and } \sum_{q=1}^{Q} \left| H_{qp}^{(P,S)}(k) \right|^2 p_q(k) \leq I_p(k), \forall q, p, k \right\}
\]

and

\[
F(p) \triangleq \begin{pmatrix}
-\nabla_{p_1} r_1(p) \\
\vdots \\
-\nabla_{p_Q}(p)
\end{pmatrix}
\]

where

- $\mathcal{G}$ is the game
- $\lambda = \text{price vector}$
- $p = \text{primal variables}$
- $\lambda = \text{multipliers}$
- $p = \text{powers of the SUs}$
- $\lambda = \text{multipliers}$

\[\text{Primal (arule) } \rightarrow (p, \lambda)\]

\[\text{Primal (arule) } \rightarrow (p, \lambda)\]
Results from the VI Framework: Solution Analysis

• Using the VI framework we can study existence/uniqueness of the solution and devise distributed algorithms.

• **Theorem (Existence and Uniqueness of the NE of $G$)** [Pan-Scu-Pal-Fac’10]:
  
  – *The VI($\mathcal{K}, F$) always admits a solution $p^{VI}$ (the NE of $G$)*
  – *If $\Upsilon \succ 0$, then $p^{VI}$ is unique.*

• We can now devise algorithms based on the VI, **but they will not be distributed due to the coupling in $\mathcal{K}$**.

• But we really want distributed algorithms...
Toward Distributed Algs: A NCP Reformulation

• Let’s introduce now the interference violation function $\Phi(\lambda): \mathbb{R}_+^{PN} \ni \lambda \mapsto \mathbb{R}^{PN}$

\[
\Phi(\lambda) \mapsto \left( \left( I_p(k) - \sum_{q=1}^{Q} \left| H_{qp}^{(P,S)}(k) \right|^2 p^*_q(k; \lambda) \right)_{k=1}^{N} \right)_{p=1}^{P}
\]

where $p^*(\lambda)$ is a Nash equilibrium of $G_\lambda$ for a given $\lambda$

• Theorem (Game as NCP) [Pan-Scu-Pal-Fac’10]: If $\Upsilon \succ 0$, then $G$ is equivalent to the NCP in the price $\lambda$

\[
\text{NCP}(\Phi): \quad 0 \leq \lambda \perp \Phi(\lambda) \geq 0 \iff \text{VI}(\mathbb{R}_+^{PN}, \Phi).
\]

• The NCP reformulation is instrumental to devise distributed algorithms.
Distributed Algorithms based on NCP

Algorithm 7: Projection algorithm with constant step-size

(S.0) Choose any $\lambda^{(0)} \geq 0$, and the step size $\tau > 0$, and set $n = 0$

(S.1) If $\lambda^{(n)}$ satisfies a suitable termination criterion: STOP

(S.2) Given $\lambda^{(n)}$, compute $p^*(\lambda^{(n)})$ as the NE solution of the NEP $\mathcal{G}_{\lambda}$ with fixed prices $\lambda = \lambda^{(n)}$

(S.3) Update the price vectors:

$$\lambda^{(n+1)} = \left[\lambda^{(n)} - \tau \Phi \left(\lambda^{(n)}\right)\right]^+ \quad (2)$$

(S.4) Set $n \leftarrow n + 1$; go to (S.1)

• **Theorem (Global convergence) [Pan-Scu-Pal-Fac’10]:** If $\Upsilon \succ 0$ and the step-size $\tau$ is sufficiently small, then the sequence $\{\lambda^{(n)}\}_{n=0}^{\infty}$ generated by the algorithm converges to a solution of the NCP($\Phi$).
• **Distributed implementation (limited signaling):**
  
  – **Inner loop:** The NE $p^*(\lambda^{(n)})$ of $\mathcal{G}_{\lambda^{(n)}}$ can be computed using the asynchronous IWFA (convergence is guaranteed under $\mathbf{Y} \succ 0$)
  
  – **Outer Loop** (*Spectrum leasing CR model*): at the iteration $n$, the PUs measure the interference violation $\Phi(\lambda^{(n)})$, update the prices $\lambda^{(n+1)}$ via the projection (2), and broadcast $\lambda^{(n+1)}$ to the SUs who play the game $\mathcal{G}_{\lambda^{(n+1)}}$
  
  – **Outer Loop** (*Common CR model*): the SUs update the prices as well by estimating the interference violation $\Phi(\lambda^{(n)})$ via consensus algorithms.

• Several other algorithms have been considered that differ in the trade-off between SUs/PUs signaling, computational complexity, convergence conditions [Pan-Scu-Pal-Fac,’10].
Convergence of Outer Loop

Outer loop convergence speed

worst-case interference violation

- Projection Algorithm

Daniel P. Palomar & Gesualdo Scutari
Convergence of Inner Loop

Inner loop convergence

- Simultaneous IWFA with pricing
- Sequential IWFA with pricing

Rates of the secondary users

Iteration index

User #1
User #5
User #15
Summary
Summary

• **Theory**: We have developed a fairly general mathematical framework based on VI suitable to study general equilibrium problems and design distributed algorithms.

• **Applications**: Using VI as a framework we have considered and solved a variety of game formulations in communications such as wireless ad-hoc networks, cognitive radio systems, network flow control problems, robust networks.

• **Take-home message**: Variational Inequality theory is a perfect mathematical framework for the analysis and design of fairly general equilibrium problems, both in theory and practice.
References - GT

- **Books and Book Chapters**


### Special Issues and Magazines:


• Journal Papers:


References - VI

• Books:

• Journal Papers (mainly in SP and Comm.)
End of Talk

Thank you!!

For more information visit:

http://www.ece.ust.hk/~palomar

http://www.sens.buffalo.edu/~gesualdo/