# ECE 5314: Power System Operation \& Control 

Lecture 5: Economic Dispatch

Vassilis Kekatos

R2 A. Gomez-Exposito, A. J. Conejo, C. Canizares, Electric Energy Systems: Analysis and Operation, Chapter 5

## Problem statement

- $N$ power generation units
- serving a given load $D$
- $P_{i}$ : power output of unit $i$ [MW]
- $C_{i}\left(P_{i}\right)$ : operation cost $[\$ / \mathrm{h}]$


Find the most economic dispatch (ED) of units

$$
\begin{aligned}
\min _{\left\{P_{i}\right\}_{i=1}^{N}} & \sum_{i=1}^{N} C_{i}\left(P_{i}\right) \\
\text { s.to } & \sum_{i=1}^{N} P_{i}=D \\
& P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }, i=1, \ldots, N
\end{aligned}
$$

Convex problem if $C_{i}\left(P_{i}\right)$ are convex functions

## Cost functions and energy markets




- Typical options for cost functions (linear, quadratic, piecewise linear)
- Who solves the economic dispatch?
- Fuel-cost curves or market bids
- Bilateral contracts and spot markets (2010)
- Day-ahead and real-time markets


## Dispatch with linear costs

$$
\begin{aligned}
\min _{\left\{P_{i}\right\}_{i=1}^{N}} & \sum_{i=1}^{N} a_{i} P_{i} \\
\text { s.to } & \sum_{i=1}^{N} P_{i}=D \\
& P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }, i=1, \ldots, N
\end{aligned}
$$

- solved as an LP or by simply sorting the cost coefficients
- uniqueness issues (case when $a_{2}=a_{3}$ of the example)


## Example:

$D=350 \mathrm{MW}$

|  | $P_{i}^{\min }[\mathrm{MW}]$ | $P_{i}^{\max }[\mathrm{MW}]$ | $a_{i}[\$ / \mathrm{MW}]$ |
| :---: | :---: | :---: | :---: |
| G1 | 0 | 100 | 20 |
| G2 | 0 | 400 | 25 |
| G3 | 0 | 100 | 22 |
| G4 | 0 | 200 | 18 |

## Dispatch with convex piecewise-linear costs

The range $\left[0, P_{i}^{\max }\right]$ is divided into $K_{i}$ blocks
Cost $C_{i}\left(P_{i}\right)$ expressed as $K_{i}$ pairs: $\left\{\left(p_{i, k}, a_{i, k}\right)=(\text { block size, price })\right\}_{k=1}^{K_{i}}$

- Any power in $\left[0, p_{i, 1}\right]$ has incremental cost $a_{i, 1} \$ / \mathrm{MW}$
- Any power in $\left[p_{i, 1}, p_{i, 2}\right.$ ] has incremental cost $a_{i, 2} \$ / \mathrm{MW} \ldots$


Cost $C_{i}\left(P_{i}\right)$ is convex if prices are increasing ( $a_{1}<\ldots<a_{i, K_{i}}$ )

## Rewriting convex piecewise linear costs



Express each cost function as the pointwise maximum of linear functions:

$$
C_{i}\left(P_{i}\right)=\max _{k \in\left\{1, \ldots, K_{i}\right\}}\left\{a_{i, k} P_{i}+b_{i, k}\right\}
$$

Q.5. 1 Show that $b_{i, k}=\left(a_{i, k-1}-a_{i, k}\right) p_{i, k-1}+b_{i, k-1}$.

## Using the epigraph form

dispatch with convex piecewise-linear costs

$$
\begin{aligned}
\min _{\left\{P_{i}\right\}} & \sum_{i=1}^{N} \max _{k \in\left\{1, \ldots, K_{i}\right\}}\left\{a_{i, k} P_{i}+b_{i, k}\right\} \\
\text { s.to } & \sum_{i=1}^{N} P_{i}=D \\
& 0 \leq P_{i} \leq P_{i}^{\max }, \forall i
\end{aligned}
$$

dispatch in epigraph form

$$
\begin{aligned}
\min _{\left\{P_{i}, t_{i}\right\}} & \sum_{i=1}^{N} t_{i} \\
\text { s.to } \quad & a_{i, k} P_{i}+b_{i, k} \leq t_{i}, \forall i, k \\
& \sum_{i=1}^{N} P_{i}=D \\
& 0 \leq P_{i} \leq P_{i}^{\max }, \forall i
\end{aligned}
$$

An alternative formulation

- Introduce one variable per block

$$
P_{i}=\sum_{k=1}^{K_{i}} P_{i, k}
$$

- Let $p_{i, 0}=0$ and $p_{i, K_{i}}=P_{i}^{\max }$


$$
\begin{aligned}
\min _{\left\{P_{i, k}\right\}} & \sum_{i=1}^{N} \sum_{k=1}^{K_{i}} a_{i, k} P_{i, k} \\
\text { s.to } & \sum_{i=1}^{N} \sum_{k=1}^{K_{i}} P_{i, k}=D \\
& 0 \leq P_{i, k} \leq p_{i, k}-p_{i, k-1}, \quad \forall i, k
\end{aligned}
$$

- At the optimum: $P_{i, k}^{*}=p_{i, k}-p_{i, k-1}$ if $P_{i, k+1}^{*}>0$


## Comparing the two formulations

Epigraph form

$$
\begin{array}{ll}
\min & \sum_{i=1}^{N} t_{i} \\
\text { s.to } & a_{i, k} P_{i}+b_{i, k} \leq t_{i} \forall i, k \\
& \sum_{i=1}^{N} P_{i}=D \\
& 0 \leq P_{i} \leq P_{i} \forall i
\end{array}
$$

variables: $2 N$
equality constraints: 1
inequality constraints: $\sum_{i=1}^{N} K_{i}+2 N$

Form with one variable per block

$$
\begin{aligned}
\min & \sum_{i=1}^{N} \sum_{k=1}^{K_{i}} a_{i, k} P_{i, k} \\
\text { s.to } & \sum_{i=1}^{N} \sum_{k=1}^{K_{i}} P_{i, k}=D \\
& 0 \leq P_{i, k} \leq p_{i, k}-p_{i, k-1} \forall i, k
\end{aligned}
$$

variables: $\sum_{i=1}^{N} K_{i}$
equality constraints: 1
inequality constraints: $2 \sum_{i=1}^{N} K_{i}$ solved by simply sorting prices

## Dispatch with convex quadratic costs

$$
\begin{aligned}
\min _{\left\{P_{i}\right\}_{i=1}^{N}} & \sum_{i=1}^{N}\left(a_{i} P_{i}+c_{i} P_{i}^{2}\right) \\
\text { s.to } & \sum_{i=1}^{N} P_{i}=D \\
& P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }, i=1, \ldots, N
\end{aligned}
$$

- convex quadratic costs $c_{i}>0\left[\$ /(\mathrm{MW})^{2}\right]$
- solved as a quadratic program (QP)
- no uniqueness issues


## Lagrangian function

For general (strictly) convex and differentiable $C_{i}\left(P_{i}\right)$
Introduce Lagrange multipliers:

$$
\begin{array}{rlr}
\min _{\left\{P_{i}\right\}_{i=1}^{N}} & \sum_{i=1}^{N} C_{i}\left(P_{i}\right) \\
\text { s.to } & \sum_{i=1}^{N} P_{i}=D \\
& P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }, i=1, \ldots, N & \longleftarrow \lambda \\
& \leftarrow\left\{\left(\underline{\mu}_{i}, \bar{\mu}_{i}\right)\right\}_{i=1}^{N}
\end{array}
$$

## Lagrangian function

$$
\begin{aligned}
L= & \sum_{i=1}^{N} C_{i}\left(P_{i}\right)-\lambda\left(\sum_{i=1}^{N} P_{i}-D\right) \\
& +\sum_{i=1}^{N} \underline{\mu}_{i}\left(P_{i}^{\min }-P_{i}\right)+\sum_{i=1}^{N} \bar{\mu}_{i}\left(P_{i}-P_{i}^{\max }\right)
\end{aligned}
$$

## Optimality conditions

For $i=1, \ldots, N$,

1. $\sum_{i=1}^{N} P_{i}^{*}=D ; \quad P_{i}^{\text {min }} \leq P_{i}^{*} \leq P_{i}^{\max }$
2. $\underline{\mu}_{i}^{*} \geq 0 ; \quad \bar{\mu}_{i}^{*} \geq 0$
3. $C_{i}^{\prime}\left(P_{i}^{*}\right)=\lambda^{*}-\bar{\mu}_{i}^{*}+\underline{\mu}_{i}^{*}$
4. $\underline{\mu}_{i}^{*}\left(P_{i}^{\min }-P_{i}^{*}\right)=0 ; \quad \bar{\mu}_{i}^{*}\left(P_{i}^{*}-P_{i}^{\max }\right)=0$

Conditions 3 and 4 can be equivalently written as

$$
C_{i}^{\prime}\left(P_{i}^{*}\right)= \begin{cases}\lambda^{*}, & \text { if } P_{i}^{\min }<P_{i}^{*}<P_{i}^{\max } \\ \lambda^{*}+\underline{\mu}_{i}^{*}, & \text { if } P_{i}^{*}=P_{i}^{\min } \\ \lambda^{*}-\bar{\mu}_{i}^{*}, & \text { if } P_{i}^{*}=P_{i}^{\max }\end{cases}
$$

## Optimality conditions simplified

Because $\underline{\mu}_{i}^{*}, \bar{\mu}_{i}^{*} \geq 0$, optimality conditions become

$$
\begin{aligned}
& \text { 1. } \sum_{i=1}^{N} P_{i}^{*}=D \\
& 2 .^{\prime} \begin{cases}C_{i}^{\prime}\left(P_{i}^{*}\right)=\lambda^{*}, & \text { if } P_{i}^{\min }<P_{i}^{*}<P_{i}^{\max } \\
C_{i}^{\prime}\left(P_{i}^{*}\right) \geq \lambda^{*}, & \text { if } P_{i}^{*}=P_{i}^{\min } \\
C_{i}^{\prime}\left(P_{i}^{*}\right) \leq \lambda^{*}, & \text { if } P_{i}^{*}=P_{i}^{\max }\end{cases}
\end{aligned}
$$

- Any $P_{G_{1}}^{*}, \ldots, P_{G_{N}}^{*}, \lambda^{*}$ satisfying the above conditions will be optimal
- $C_{i}^{\prime}\left(P_{i}\right)$ is the marginal or incremental cost for unit $i$
- Optimal dispatch when all units not operating at their limits have the same incremental cost


## Solving economic dispatch

- Strict convexity implies that $C_{i}^{\prime}\left(P_{i}\right)$ is increasing in $P_{i}$
- Define the increasing function $f_{i}=\left(C_{i}^{\prime}\right)^{-1}$
- Given a $\lambda$, the $\left\{P_{i}\right\}_{i=1}^{N}$ satisfying condition $2^{\prime}$ are expressed as

$$
P_{i}(\lambda)= \begin{cases}f_{i}(\lambda), & \text { if } C_{i}^{\prime}\left(P_{i}^{\min }\right)<\lambda<C_{i}^{\prime}\left(P_{i}^{\max }\right) \\ P_{i}^{\min }, & \text { if } \lambda \leq C_{i}^{\prime}\left(P_{i}^{\min }\right) \\ P_{i}^{\max }, & \text { if } \lambda \geq C_{i}^{\prime}\left(P_{i}^{\max }\right)\end{cases}
$$

- Goal: find $\lambda^{*}$ such that $\left\{P_{i}\right\}_{i=1}^{N}$ satisfy condition 1 (power balance) too
- $P_{i}(\lambda)$ and $\sum_{i=1}^{N} P_{i}(\lambda)-D$ are increasing in $\lambda$
- Bisection on $\lambda$ until $\left|\sum_{i=1}^{N} P_{i}(\lambda)-D\right| \leq \epsilon$ (tolerance)


## $\lambda$-iteration or bisection method

Given tolerance $\epsilon>0$, start with $\underline{\lambda}$ and $\bar{\lambda}$ so that $\underline{\lambda} \leq \lambda^{*} \leq \bar{\lambda}$ for example: $\underline{\lambda}=\min _{i} C_{i}^{\prime}\left(P_{i}^{\min }\right)$ and $\bar{\lambda}=\max _{i} C_{i}^{\prime}\left(P_{i}^{\max }\right)$.

1. Set $\lambda=(\underline{\lambda}+\bar{\lambda}) / 2$
2. Find $P_{i}(\lambda)$ for all $i$ as follows
2.1 If $C_{i}^{\prime}\left(P_{i}^{\text {min }}\right) \geq \lambda$, set $P_{i}(\lambda)=P_{i}^{\text {min }}$
2.2 If $C_{i}^{\prime}\left(P_{i}^{\text {max }}\right) \leq \lambda$, set $P_{i}(\lambda)=P_{i}^{\max }$
2.3 Otherwise, set $P_{i}(\lambda)$ as the solution to $C_{i}^{\prime}\left(P_{i}\right)=\lambda$
3. If $\sum_{i=1}^{N} P_{i}(\lambda)-D>\epsilon$, set $\bar{\lambda}:=\lambda$ and go to Step 1
4. If $\sum_{i=1}^{N} P_{i}(\lambda)-D<-\epsilon$, set $\underline{\lambda}:=\lambda$ and go to Step 1
5. Else $\left\{P_{i}(\lambda)\right\}_{i=1}^{N}$ is the solution within the specified tolerance

## Graphical illustration


$\lambda$-iteration method for two generators with quadratic costs

## $\lambda$-iteration with convex piecewise linear costs

Convex piecewise linear $C_{i}\left(P_{i}\right)$ is non-differentiable
KKT conditions with constraints $P_{i}^{\text {min }} \leq P_{i} \leq P_{i}^{\text {max }}$ kept implicit

$$
\text { Lagrangian function : } L\left(P_{1}, \ldots, P_{N}, \lambda\right)=\sum_{i=1}^{N} C_{i}\left(P_{i}\right)-\lambda\left(\sum_{i=1}^{N} P_{i}-D\right)
$$

Optimality conditions:

$$
\begin{aligned}
& \text { 1. } \sum_{i=1}^{N} P_{i}^{*}=D \\
& \text { 2. } P_{i}^{*} \in \arg \min _{P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }}\left\{C_{i}\left(P_{i}\right)-\lambda^{*} P_{i}\right\} \quad \forall i
\end{aligned}
$$

Step 2 of $\lambda$-iteration now becomes

$$
\min _{P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }}\left\{C_{i}\left(P_{i}\right)-\lambda P_{i}\right\}
$$

## Economic interpretation

Optimality conditions with general costs

$$
\begin{aligned}
& \text { 1. } \sum_{i=1}^{N} P_{i}^{*}=D \\
& \text { 2. } P_{i}^{*} \in \arg \min _{P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }}\left\{C_{i}\left(P_{i}\right)-\lambda^{*} P_{i}\right\} \quad \forall i
\end{aligned}
$$

Interpret multiplier $\lambda[\$ / \mathrm{MW}]$ as the price at which unit $i$ will be compensated for producing $P_{i} \mathrm{MW}$

Given $\lambda$, unit $i$ chooses $P_{i}$ so its net cost $C_{i}\left(P_{i}\right)-\lambda P_{i}$ is minimized (net revenue maximized)

Optimal multiplier $\lambda^{*}$ maximizes the total net revenue

## Sensitivity interpretation

$$
\begin{aligned}
C(D):=\min _{\left\{P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }\right\}_{i=1}^{N}} & \sum_{i=1}^{N} C_{i}\left(P_{i}\right) \\
\text { s.to } & \sum_{i=1}^{N} P_{i}=D
\end{aligned}
$$

Cost of generation as a function of demand $D$ is a perturbation function!

If the problem is convex, then function $C(D)$ is convex

If $C(D)$ is differentiable and strong duality holds, then $C^{\prime}(D)=\lambda^{*}(D)$

Marginal price $\lambda^{*}$ : extra price one would pay for slight increase in demand

## Participation factors

Consider an optimal dispatch $P$ and a small load variation $\Delta D$
Assume new dispatch does not alter which units are at $P_{i}^{\min }$ or $P_{i}^{\max }$
current $E D$
$\sum_{i} P_{i}=D$

$$
\lambda=C_{i}^{\prime}\left(P_{i}\right)
$$

perturbed $E D$
$\sum_{i} \Delta P_{i}=\Delta D$
$\Delta \lambda=C_{i}^{\prime \prime}\left(P_{i}\right) \Delta P_{i}$

Participation factors: How much each unit contributes to serve the new load?

$$
\frac{\Delta P_{i}}{\Delta D}=\frac{\Delta \lambda / C_{i}^{\prime \prime}\left(P_{i}\right)}{\sum_{k} \Delta \lambda / C_{k}^{\prime \prime}\left(P_{k}\right)}=\frac{1 / C_{i}^{\prime \prime}\left(P_{i}\right)}{\sum_{k} 1 / C_{k}^{\prime \prime}\left(P_{k}\right)} \in(0,1)
$$

constant for quadratic costs

Used for fast generation in response to load variation $P_{i}^{\text {new }}=P_{i}+\frac{\Delta P_{i}}{\Delta D} \Delta D$

## Elastic demand

Elastic demand is characterized by its utility function $U_{j}\left(D_{j}\right)[\$ / \mathrm{h}]$ : the benefit by consuming $D_{j} \mathrm{MW}$ for 1 h

$$
\begin{aligned}
\min _{\left\{P_{i}\right\}_{i=1}^{N},\left\{D_{j}\right\}_{j=1}^{M}} & \sum_{i=1}^{N} C_{i}\left(P_{i}\right)-\sum_{j=1}^{M} U_{j}\left(D_{j}\right) \\
\text { s.to } & \sum_{i=1}^{N} P_{i}=\sum_{j=1}^{M} D_{j} \\
& P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }, i=1, \ldots, N \\
& D_{j}^{\min } \leq D_{j} \leq D_{j}^{\max }, j=1, \ldots, M
\end{aligned}
$$

- demand bids $U_{j}\left(D_{j}\right)$ are submitted by utilities, load serving entities, aggregators, or industrial costumers
- convex problem for concave $U_{j}\left(D_{j}\right)$ 's (diminishing returns or buy in bulk)


## Transmission losses

Accounting for losses on transmission lines (3-5\%):

$$
\begin{align*}
\min _{\left\{P_{i}\right\}_{i=1}^{N}} & \sum_{i=1}^{N} C_{i}\left(P_{i}\right) \\
\text { s.to } & \sum_{i=1}^{N} P_{i}=P_{\text {loss }}(\mathbf{p}) \\
& P_{i}^{\min } \leq P_{i} \leq P_{i}^{\max }, i=1, \ldots, N
\end{align*}
$$

where $\mathbf{p}:=\left[\begin{array}{lll}P_{1} & \ldots & P_{N}\end{array}\right]^{\top}$ captures generation and (fixed or elastic) demand

- Losses $P_{\text {loss }}(\mathbf{p})$ is a non-linear function of $\mathbf{p}$ (non-convex problem)
- In the past, modeled as quadratic functions: $P_{\text {loss }}(\mathbf{p})=\mathbf{p}^{\top} \mathbf{B p}+\mathbf{c}^{\top} \mathbf{p}$
- Now typically calculated by (successive) linearization of power flow equations or via the optimal power flow problem


## Penalty factors

Lagrangian function (non-convex problem; KKT conditions are only necessary)
$L=\sum_{i=1}^{N} C_{i}\left(P_{i}\right)-\lambda\left(\sum_{i=1}^{N} P_{i}-P_{\text {loss }}(\mathbf{p})\right)+\sum_{i=1}^{N} \underline{\mu}_{i}\left(P_{i}^{\min }-P_{i}\right)+\sum_{i=1}^{N} \bar{\mu}_{i}\left(P_{i}-P_{i}^{\max }\right)$
Optimality conditions simplified:

$$
\begin{aligned}
& \text { 1. } \sum_{i=1}^{N} P_{i}^{*}=P_{\text {loss }}\left(\mathbf{p}^{*}\right) \\
& 2 .^{\prime} \begin{cases}C_{i}^{\prime}\left(P_{i}^{*}\right)=\lambda^{*}\left(1-\left.\frac{\partial P_{\text {loss }}(\mathbf{p})}{\partial P_{i}}\right|_{\mathbf{p}=\mathbf{p}^{*}}\right), & \text { if } P_{i}^{\min }<P_{i}^{*}<P_{i}^{\max } \\
C_{i}^{\prime}\left(P_{i}^{*}\right) \geq \lambda^{*}\left(1-\left.\frac{\partial P_{\text {loss }}(\mathbf{p})}{\partial P_{i}}\right|_{\mathbf{p}=\mathbf{p}^{*}}\right), & \text { if } P_{i}^{*}=P_{i}^{\min } \\
C_{i}^{\prime}\left(P_{i}^{*}\right) \leq \lambda^{*}\left(1-\left.\frac{\partial P_{\text {loss }}(\mathbf{p})}{\partial P_{i}}\right|_{\mathbf{p}=\mathbf{p}^{*}}\right), & \text { if } P_{i}^{*}=P_{i}^{\max }\end{cases}
\end{aligned}
$$

Marginal costs: $f_{i} C_{i}^{\prime}\left(P_{i}^{*}\right)$ for penalty factors $f_{i}=\left(1-\left.\frac{\partial P_{\text {loss }}(\mathbf{p})}{\partial P_{i}}\right|_{\mathbf{p}=\mathbf{p}^{*}}\right)^{-1}$

## Solving economic dispatch with losses

Solve economic dispatch ignoring losses to get initial dispatch $\mathbf{p}^{0}$
At iteration $t=0,1, \ldots$, do

1. Calculate $P_{\text {loss }}\left(\mathbf{p}^{t}\right)$ and $f_{i}^{t}:=\left(1-\left.\frac{\partial P_{\text {loss }}(\mathbf{p})}{\partial P_{i}}\right|_{\mathbf{p}^{t}}\right)^{-1}$ for $i=1, \ldots, N$
2. Find $\mathbf{p}^{t+1}$ via $\lambda$-iteration with costs $\left\{f_{i}^{t} C_{i}\right\}_{i}$ and $\sum_{i=1}^{N} P_{i}=P_{\text {loss }}\left(\mathbf{p}^{t}\right)$
3. If $\left\|\mathbf{p}^{t+1}-\mathbf{p}^{t}\right\|_{2}<\epsilon$ stop; otherwise go to Step 1

- heuristic manner to account for losses
- how to compute $\nabla P_{\text {loss }}(\mathbf{p})$ ?


## Computing penalty factors

Active power injections are not independent (swing bus 1 makes up for losses)

$$
\mathbf{p}=\left[\begin{array}{c}
p_{1}(\tilde{\mathbf{p}}) \\
\tilde{\mathbf{p}}
\end{array}\right] \in \mathbb{R}^{N}
$$

Active power losses: $P_{\text {loss }}(\mathbf{p})=\mathbf{p}^{\top} \mathbf{1}_{N}=p_{1}(\tilde{\mathbf{p}})+\tilde{\mathbf{p}}^{\top} \mathbf{1}_{N-1}$

Interest in finding $\nabla_{\mathbf{p}} P_{\text {loss }}(\mathbf{p})$ (assume $\frac{\partial P_{\text {loss }}(\mathbf{p})}{\partial p_{1}}=0$ )

Gradient vector for losses: $\nabla_{\tilde{\mathbf{p}}} P_{\text {loss }}(\tilde{\mathbf{p}})=\mathbf{1}+\nabla_{\tilde{\mathbf{p}}} p_{1}(\tilde{\mathbf{p}})$

Resort to power flow equations to find $\nabla_{\tilde{\mathbf{p}}} p_{1}(\tilde{\mathbf{p}})$

## Finding $\nabla_{\tilde{\mathbf{p}}} p_{1}(\tilde{\mathbf{p}})$ from power flow equations

power flow equations (dependence on voltage magnitudes has been ignored)

$$
\begin{aligned}
p_{1} & =f_{1}(\boldsymbol{\theta}) \\
\tilde{\mathbf{p}} & =\mathbf{f}_{2}(\boldsymbol{\theta})
\end{aligned}
$$

where $\boldsymbol{\theta}:=\left[\begin{array}{lll}\theta_{2} & \cdots & \theta_{N}\end{array}\right]^{\top}$ is the vector of voltage phases $\left(\theta_{1}=0\right)$

Jacobian matrix: $\frac{\partial \tilde{\mathbf{p}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}=\frac{\partial \mathbf{f}_{2}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}=\mathbf{J}(\boldsymbol{\theta}) \in \mathbb{R}^{(N-1) \times(N-1)}$

Jacobian matrix of inverse function: under technical conditions

$$
\frac{\partial \boldsymbol{\theta}(\tilde{\mathbf{p}})}{\partial \tilde{\mathbf{p}}}=\frac{\partial \mathbf{f}_{2}^{-1}(\tilde{\mathbf{p}})}{\partial \tilde{\mathbf{p}}}=[\mathbf{J}(\boldsymbol{\theta}(\tilde{\mathbf{p}}))]^{-1} \in \mathbb{R}^{(N-1) \times(N-1)}
$$

Gradient vector:
$\nabla_{\tilde{\mathbf{p}}} P_{\text {loss }}(\tilde{\mathbf{p}})=\mathbf{1}+\left(\frac{\partial \boldsymbol{\theta}(\tilde{\mathbf{p}})}{\partial \tilde{\mathbf{p}}}\right)^{\top} \nabla_{\boldsymbol{\theta}} f_{1}(\boldsymbol{\theta})=\mathbf{1}+[\mathbf{J}(\tilde{\mathbf{p}})]^{-\top} \nabla_{\boldsymbol{\theta}} f_{1}(\boldsymbol{\theta})$

