

OTTIMIZZAZIONE NELLE SMART GRID:
SOLUZIONE CENTRALIZZATA E DISTRIBUITA

RELATORE: Ch.mo Prof. Tomaso Erseghe

LAUREANDO: Alessandro Biral

A.A. 2012-2013



UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
TESI DI LAUREA

OTTIMIZZAZIONE NELLE
SMART GRID: SOLUZIONE
CENTRALIZZATA E DISTRIBUITA

RELATORE: *Ch.mo Prof. Tomaso Erseghe*

LAUREANDO: *Alessandro Biral*

Padova, 10 dicembre 2012

Sommario

Negli ultimi anni numerosi fattori, fra cui la rapida diffusione delle fonti di energia rinnovabile e lo sviluppo del concetto di produzione distribuita, hanno favorito la nascita delle smart-grid. Con tale termine vengono identificate le cosiddette reti elettriche “intelligenti”, in grado di operare perseguendo opportuni obiettivi di efficienza. Fra questi, uno dei più significativi è rappresentato dall’ottimizzazione dei flussi di potenza che circolano in rete. La presente tesi studia proprio il problema dell’Optimal Power Flow, proponendo due diversi strumenti per la sua soluzione: l’Interior Point Method e l’Alternating Direction Multiplier Method. Entrambi sono algoritmi iterativi, ma mentre il primo sfrutta un approccio di tipo centralizzato, il secondo ne adotta uno di tipo distribuito. Dopo aver spiegato in che modo essi possono fungere da solutori per il problema dell’Optimal Power Flow, vengono analizzate le loro prestazioni sia dal punto di vista dell’efficacia che dell’efficienza risolutiva. Dal confronto di tali performance, vengono infine stabiliti quali possano essere i vantaggi e gli svantaggi derivanti dall’utilizzo di un approccio rispetto all’altro.

Indice

1	Introduzione	1
1.1	Contenuto della tesi	4
2	Optimal Power Flow	5
2.1	La rete elettrica	5
2.2	Modello matematico associato al problema dell'OPF	10
3	Algoritmo Centralizzato	13
3.1	Punti di ottimo globali e locali	14
3.2	Problemi di ottimizzazione convessi	16
3.3	La funzione duale di Lagrange	19
3.3.1	Proprietà della funzione duale di Lagrange	20
3.4	Il problema duale di Lagrange	21
3.5	Dualità forte	22
3.5.1	Condizione di complementarità	26
3.6	Condizioni di Karush-Khun-Tucker	27
3.7	Interior Point Method	29
3.7.1	Il barrier method	31
3.7.2	Applicazione allo studio dell'Optimal Power Flow	34
4	Algoritmo Distribuito	41
4.1	ADMM scalabile	42
4.2	Applicazione dell'ADMM allo studio dell'Optimal Power Flow	46
4.2.1	Soluzione locale dell'OPF per il nodo PCC	50
4.2.2	Soluzione locale dell'OPF per i nodi PEIs	52
4.2.3	Soluzione locale dell'OPF per i Loads	56

INDICE

5	Risultati delle simulazioni	61
5.1	Performance dell'IPM	62
5.2	Performance dell'ADMM	67
5.3	Confronto fra gli algoritmi	72
6	Sviluppi futuri	79
A	Strategia di aggiornamento delle variabili nell'ADMM	81
	Bibliografia	83

Capitolo 1

Introduzione

Le reti elettriche nascono storicamente con lo scopo di raccogliere grandi quantità di energia e distribuirle ininterrottamente ad una moltitudine di clienti-consumatori. Ciò può essere realizzato semplicemente ricorrendo a centrali di produzione di grandi dimensioni, flussi di potenza unidirezionali e punti di estrazione assolutamente passivi. Tale visione classica è però in fase di superamento ed oggi il processo di distribuzione di energia è sempre più mutuato da logiche di ottimizzazione delle risorse. L'unico modo per realizzare un simile cambiamento è dotare la rete stessa di capacità elaborative, che le permettano di operare scelte in maniera autonoma ed intelligente. È in questo ambito che si può inquadrare un nuovo, rivoluzionario concetto: quello di smart-grid.

Con il termine smart-grid si indica una rete elettrica in grado di autogestirsi seguendo opportuni obiettivi di efficienza. In pratica, ad essa viene chiesto di monitorare costantemente le condizioni elettriche dei suoi nodi in modo da evitare eventuali sprechi energetici, sovraccarichi o cadute di tensione. Tutto ciò ha assunto un'importanza sempre maggiore in questi ultimi anni, in cui la crescita demografica ha portato ad un conseguente aumento della richiesta di energia da parte della popolazione mondiale. In tale contesto, le smart-grid rappresentano quindi non solo uno strumento per il risparmio economico, ma anche un mezzo per tutelare il nostro pianeta e le sue risorse naturali. In quest'ottica può essere interpretata anche la rapida diffusione delle fonti di energia rinnovabile occorsa negli ultimi anni. Essa ha aperto le porte ad un nuovo concetto: quello di *produzione distribuita* dell'energia. Con tale formula si indica la pratica moderna di decentrare la generazione di energia elettrica, grazie allo sfruttamento di pannel-

1. INTRODUZIONE

li fotovoltaici, pale eoliche o impianti analoghi disposti a macchia sul territorio. Così facendo, si consente a ciascun cittadino di rivestire due ruoli in antitesi fra loro: quello classico di consumatore e quello nuovo di produttore di energia e gli si associa dunque l'epiteto di *prosumer* (producer + consumer). L'energia che egli genera può essere impiegata direttamente in ambito locale per sopperire alle richieste di potenza dell'utente stesso, ma può anche essere resa disponibile al resto della popolazione attraverso la sua immissione nella rete di distribuzione elettrica. Tale pratica non è tuttavia esente da complicazioni. Molte fonti energetiche rinnovabili, come ad esempio eolico e fotovoltaico, posseggono infatti delle caratteristiche innate di aleatorietà che ne modificano il comportamento a seconda del periodo dell'anno in cui ci si trova o addirittura all'interno delle 24 ore giornaliere. Compito della smart-grid è quindi fare in modo che il resto della rete si adatti a questi cambiamenti, grazie ad un controllo dinamico dei flussi di potenza.

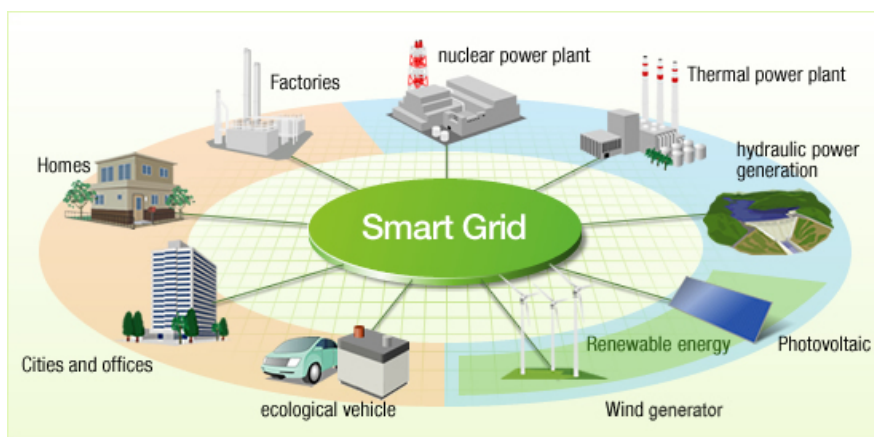


Figura 1.1: Esempio di smart-grid.

Tale scopo può essere perseguito, da un punto di vista pratico, utilizzando sensori e controllori posizionati ai vari snodi della rete e, da un punto di vista teorico, risolvendo il problema della *Power Flow Optimization*. Come vedremo meglio nel prosieguo della trattazione, quest'ultimo può essere formulato matematicamente sfruttando un generico modello di ottimizzazione del tipo:

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && h(x) = 0 \\ & && \underline{g} \leq g(x) \leq \bar{g} \end{aligned} \tag{1.1}$$

dove x rappresenta la variabile da ottimizzare, $f(x)$ è la funzione obiettivo e le relazioni $h(x) = 0$ e $\underline{g} \leq g(x) \leq \bar{g}$ riassumono i vincoli da rispettare. Risolvere il problema dell' Optimal Power Flow significa dunque passare attraverso la soluzione di un modello analogo a (1.1), studiato già a partire dagli anni '60. Solo nel 1977, però, fu pubblicata da H. H. Happ quella che può essere reputata la prima indagine organica riguardante le tecniche di trasmissione ottima di potenza [1]. Fra queste si possono annoverare il *Linear Programming Method*, il *Newton-Raphson Method*, il *Quadratic Programming Method*, il *Nonlinear Programming Method*, l'*Interior Point Method* e alcuni *Artificial Intelligence Methods*. Nel presente elaborato verrà analizzato in particolare l'Interior Point Method (IPM). Esso fu illustrato per la prima volta da N. Karmarkar nel 1984 [2] e negli anni successivi subì numerosi adeguamenti, che ne hanno migliorato le performance. La sua implementazione attuale è consultabile nell'articolo [5], da cui abbiamo preso spunto per redarre il terzo capitolo della tesi.

Nel seguito verrà proposto anche un altro solutore per il problema dell'Optimal Power Flow, che prende il nome di *Alternating Direction Multiplier Method* (ADMM). Esso non è stato citato in precedenza in quanto non rientra fra le tecniche classiche di ottimizzazione delle reti elettriche. A differenza di queste, l'ADMM sfrutta una logica di tipo distribuito, in cui il modello (1.1) viene risolto frazionandolo in vari sottoproblemi di dimensione inferiore. A partire dalla soluzione di questi ultimi è possibile ricostruire quella globale per l'intera rete sfruttando un opportuno scambio di informazioni fra i nodi che la compongono. Tale criterio verrà esposto più dettagliatamente all'interno del quarto capitolo e nel quinto le sue prestazioni saranno confrontate con quelle dell'IPM.

Ci teniamo a precisare tuttavia che il primo obiettivo del presente lavoro rimane quello di far capire l'importanza di adottare tecniche di Power Flow Optimization (indipendentemente dall'algoritmo che si decide di utilizzare per implementarle) a cui è auspicabile affiancare un sapiente impiego delle fonti di energia rinnovabili. Il binomio smart-grid ed energie rinnovabili può infatti rappresentare il futuro in una società come la nostra in cui si professa l'ecosostenibilità, ma dove ancora oggi più dell'80% del fabbisogno energetico mondiale è soddisfatto attraverso l'impiego di risorse fossili.

1.1 Contenuto della tesi

Nel *Capitolo 2* vengono elencate le principali grandezze che caratterizzano una qualsiasi rete elettrica, sia dal punto di vista quantitativo che qualitativo. Sfruttando tali informazioni viene inoltre descritto il problema dell'Optimal Power Flow (OPF), chiarendo in che modo può essere formulato attraverso l'utilizzo di un modello di programmazione non lineare.

Nel *Capitolo 3* viene presentato l'Interior Point Method, fornendo le basi teoriche che ne permettono la comprensione. In particolare, vengono analizzati concetti tipici del settore della ricerca operativa quali ad esempio il rilassamento Lagrangiano, la dualità debole e quella forte. In seguito, viene discusso in che modo l'Interior Point Method può essere riadattato per risolvere il problema della Power Flow Optimization.

Il *Capitolo 4* ha una struttura molto simile al precedente. Inizialmente infatti viene condotta un'analisi matematica propedeutica alla comprensione dell'Alternating Direction Multiplier Method. Svolta questa, viene indicato come applicare tale algoritmo al contesto della Power Flow Optimization.

Nella prima parte del *Capitolo 5* vengono presentati i risultati dell'ottimizzazione di tre diverse reti, ricavati attraverso l'implementazione dell'IPM e dell'ADMM. In particolare, si dà ragione dell'efficacia con cui questi ultimi minimizzano le perdite di rete e della complessità computazionale con cui riescono a farlo. Nella seconda parte tali parametri vengono confrontati e si cerca quindi di stabilire in che misura un solutore è preferibile rispetto all'altro.

Il *Capitolo 6*, infine, contiene degli spunti per dei possibili sviluppi futuri delle tematiche trattate all'interno dell'elaborato.

Capitolo 2

Optimal Power Flow

L'iniezione ottimale di potenza all'interno di una qualsiasi rete elettrica, con l'obiettivo di minimizzare una ben precisa funzione obiettivo, va sotto il nome di *Optimal Power Flow* (OPF) e consiste sostanzialmente nella soluzione di un problema non lineare caratterizzato da vincoli non convessi. Questi ultimi possono essere di natura diversa e rappresentare effettivi limiti fisici o semplici restrizioni dovute ad una ricerca di efficienza nella gestione della rete stessa.

I primi studi in merito al problema dell' Optimal Power Flow risalgono ai primi anni '60 [6]-[7] e hanno trovato nuovo slancio nel passato più recente grazie alla progressiva diffusione del concetto di produzione distribuita dell'energia, a cui abbiamo fatto cenno nel capitolo introduttivo. Le tecniche risolutive che ne sono scaturite includono, fra le altre, il Nonlinear Programming (NLP), il Newton-Raphson Method, il Semi Definite Programming (SDP) e l' Interior Point Method (IPM), che sarà analizzato dettagliatamente nel prossimo capitolo. Per comprenderne il funzionamento è però prima necessario definire le principali grandezze con cui avremo a che fare nel seguito e che caratterizzano matematicamente la nostra rete elettrica.

2.1 La rete elettrica

Come qualsiasi tipo di rete, anche quella elettrica può essere schematizzata attraverso un insieme di nodi, collegati tra loro grazie ad una moltitudine di rami o link. Questi ultimi hanno il compito di veicolare la corrente e, a seconda della loro numerosità e disposizione, danno origine ad una struttura topologica più o

2. OPTIMAL POWER FLOW

meno magliata, che ci accingiamo a studiare dal punto di vista elettrico. A tal proposito, si indichi con V_n la tensione del nodo n -esimo e con I_n la corrente dello stesso, dove $n \in \{1, \dots, N\}$ ed N corrisponde al numero totale di nodi. Fra il vettore delle tensioni $\mathbf{V} = [V_1, \dots, V_N]^T$ e quello delle correnti $\mathbf{I} = [I_1, \dots, I_N]^T$ esiste una relazione di proporzionalità riassunta dall'equazione:

$$\mathbf{I} = \mathbf{Y} \mathbf{V} \quad (2.1)$$

dove con \mathbf{Y} viene indicata la *matrice di ammettenza di rete*, la quale ha dimensione $N \times N$. Alla formula (2.1) si perviene introducendo la matrice di incidenza \mathbf{A} , che riassume in se' la struttura topologica del grafo ed ha dimensione $L \times N$, dove L corrisponde al numero di rami della rete. La generica riga i , con $i \in \{1, \dots, L\}$, formalizza matematicamente la relazione di adiacenza espressa dal link i -esimo. Per questo motivo, al suo interno vi saranno solo due elementi non nulli: in particolare, si avrà $A_{i,n} = 1$ se si ipotizza che l' i -esimo ramo entri nel nodo n e $A_{i,m} = -1$ se si suppone che esso esca dal nodo m . Le informazioni racchiuse in \mathbf{A} sono indispensabili per esprimere in forma matriciale i vincoli imposti dalla *legge di Kirchhoff* alle correnti. Quest'ultima afferma che, in qualsiasi istante, all'interno della nostra rete è nulla la somma algebrica delle correnti dei lati che compongono un qualsiasi insieme di taglio. Matematicamente ciò significa che:

$$I_n = \sum_{i \in \mathcal{N}_n} \pm I_{link,i} \quad (2.2)$$

dove \mathcal{N}_n rappresenta l'insieme dei lati afferenti al nodo n e $I_{link,i}$ la corrente che attraversa il link i -esimo. Introducendo il vettore $\mathbf{I}_{link} = [I_{link,1}, \dots, I_{link,L}]^T$ è possibile tradurre in forma matriciale l'equazione (2.2) come segue:

$$\mathbf{I} = \mathbf{A}^T \mathbf{I}_{link} \quad (2.3)$$

Ricordando inoltre che ciascun ramo è caratterizzato da una specifica impedenza di linea $Z_{link,i}$, possiamo mettere in relazione la corrente del generico link i con le tensioni dei nodi m ed n da esso collegati:

$$V_m - V_n = Z_{link,i} I_{link,i} \quad (2.4)$$

In forma matriciale si ha:

$$\mathbf{A} \mathbf{V} = \mathit{diag}(\mathbf{Z}_{link}) \mathbf{I}_{link} \quad (2.5)$$

con $diag(\mathbf{Z}_{link})$ matrice diagonale di dimensione $L \times L$, nella cui diagonale principale sono presenti i valori di tutte le impedenze di linea, $Z_{link,i}$ con $i \in \{1, \dots, L\}$. Mettendo insieme (2.3) e (2.5) si ottiene:

$$\mathbf{I} = \mathbf{A}^T [diag(\mathbf{Z}_{link})]^{-1} \mathbf{A} \mathbf{V} = \mathbf{Y} \mathbf{V} \quad (2.6)$$

e risulta così dimostrata l'espressione (2.1).

Un'altra grandezza molto importante per il prosieguo della nostra trattazione è rappresentata dalla potenza complessa associata a ciascuna terminazione di rete. Per il generico nodo n , $n \in \{1, \dots, N\}$, essa viene definita come il prodotto fra la tensione e il complesso-coniugato della corrente ad esso associati:

$$S_n = V_n I_n^* \quad (2.7)$$

È utile distinguere componente attiva e reattiva, definite rispettivamente come la parte reale e immaginaria di S_n , ovvero:

$$S_n = \Re[S_n] + j\Im[S_n] = P_n + jQ_n \quad (2.8)$$

Come vedremo più avanti, infatti, alcuni dei vincoli in gioco nella ricerca dell'Optimal Power Flow verranno applicati esclusivamente ad uno dei due termini.

Ciascun nodo di rete appartiene ad una delle tre seguenti categorie:

- **Power Common Couplings(PCCs)**: nodi modellati come generatori di tensione ideali, con tensione nominale a circuito aperto pari a V_{PCC} , ossia:

$$V_n = V_{PCC,n}, \quad n \in \mathcal{I}_{PCC} \quad (2.9)$$

dove l'insieme \mathcal{I}_{PCC} contiene l'indice dei nodi con caratteristiche assimilabili a quelle di un PCC. Generalmente esiste un unico nodo PCC e nelle smart-grid viene naturale identificarlo con quello che si allaccia alla rete di distribuzione esterna. Da qui in avanti si può quindi supporre senza perdita di generalità che $N_{PCC} = \# \text{ nodi PCC} = 1$ e che il nodo PCC sia sempre quello con indice 1 ($\mathcal{I}_{PCC} = \{1\}$).

- **Power Electric Interfaces(PEIs)**: nodi modellati come generatori di corrente, con corrente nominale a circuito chiuso pari a I_{PEI} , ossia:

$$I_n = I_{PEI,n}, \quad n \in \mathcal{I}_{PEI} \quad (2.10)$$

dove \mathcal{I}_{PEI} è un insieme di cardinalità $N_{PEI} = \# \text{ nodi PEI}$, contenente l'indice dei nodi con caratteristiche assimilabili a quelle di un PEI.

2. OPTIMAL POWER FLOW

- **Loads:** nodi modellati come singole impedenze Z_n , di modo che:

$$I_n = -\frac{V_n}{Z_n}, \quad n \in \mathcal{I}_{Load} \quad (2.11)$$

dove \mathcal{I}_{Load} è l'insieme che raccoglie gli indici dei nodi con caratteristiche di carico¹.

Chiaramente si avrà: $N = N_{PCC} + N_{PEI} + N_{Load}$. In generale, non è pensabile che le potenze di lavoro dei vari nodi possano essere fissate indipendentemente le une dalle altre in quanto deve essere rispettato il principio di conservazione dell'energia e devono essere bilanciate le perdite di linea. Da qui la necessità di introdurre un PCC che sia in grado di iniettare in rete la potenza necessaria a soddisfare il bilancio energetico.

La classificazione appena presentata è la medesima adottata all'interno dei file dell'IEEE che abbiamo utilizzato per condurre la parte simulativa. Qui di seguito è riportato a titolo di esempio il file relativo alla rete da 14 nodi.

Esempio 2.1. : Rete IEEE da 14 nodi

```
08/19/93 UW ARCHIVE          100.0 1962 W IEEE 14 Bus Test Case
BUS DATA FOLLOWS                14 ITEMS
 1 Bus 1   HV  1  1  3 1.060   0.0   0.0   0.0  232.4  -16.9   0.0 1.060   0.0   0.0  0.0   0.0   0
 2 Bus 2   HV  1  1  2 1.045  -4.98  21.7  12.7  40.0  42.4   0.0 1.045  50.0 -40.0  0.0  0.0   0
 3 Bus 3   HV  1  1  2 1.010 -12.72  94.2  19.0   0.0  23.4   0.0 1.010  40.0   0.0  0.0  0.0   0
 4 Bus 4   HV  1  1  0 1.019 -10.33  47.8  -3.9   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.0   0
 5 Bus 5   HV  1  1  0 1.020  -8.78   7.6   1.6   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.0   0
 6 Bus 6   LV  1  1  2 1.070 -14.22  11.2   7.5   0.0  12.2   0.0 1.070  24.0  -6.0  0.0  0.0   0
 7 Bus 7   ZV  1  1  0 1.062 -13.37   0.0   0.0   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.0   0
 8 Bus 8   TV  1  1  2 1.090 -13.36   0.0   0.0   0.0  17.4   0.0 1.090  24.0  -6.0  0.0  0.0   0
 9 Bus 9   LV  1  1  0 1.056 -14.94  29.5  16.6   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.19  0
10 Bus 10  LV  1  1  0 1.051 -15.10   9.0   5.8   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.0   0
11 Bus 11  LV  1  1  0 1.057 -14.79   3.5   1.8   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.0   0
12 Bus 12  LV  1  1  0 1.055 -15.07   6.1   1.6   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.0   0
13 Bus 13  LV  1  1  0 1.050 -15.16  13.5   5.8   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.0   0
14 Bus 14  LV  1  1  0 1.036 -16.04  14.9   5.0   0.0   0.0   0.0 0.0   0.0  0.0  0.0  0.0   0
-999
BRANCH DATA FOLLOWS                20 ITEMS
 1  2  1  1  1 0 0.01938  0.05917  0.0528  0  0  0  0  0 0.0  0.0 0.0  0.0  0.0  0.0
 1  5  1  1  1 0 0.05403  0.22304  0.0492  0  0  0  0  0 0.0  0.0 0.0  0.0  0.0  0.0
```

¹I Loads possono essere descritti anche tramite la potenza complessa ad essi associata; si ha infatti $S_n^* = \frac{|V_n|^2}{Z_n}$.

2.1 LA RETE ELETTRICA

```

2  3  1  1  1  0  0.04699  0.19797  0.0438  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
2  4  1  1  1  0  0.05811  0.17632  0.0340  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
2  5  1  1  1  0  0.05695  0.17388  0.0346  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
3  4  1  1  1  0  0.06701  0.17103  0.0128  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
4  5  1  1  1  0  0.01335  0.04211  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
4  7  1  1  1  0  0.0  0.20912  0.0  0  0  0  0  0  0.978  0.0  0.0  0.0  0.0  0.0
4  9  1  1  1  0  0.0  0.55618  0.0  0  0  0  0  0  0.969  0.0  0.0  0.0  0.0  0.0
5  6  1  1  1  0  0.0  0.25202  0.0  0  0  0  0  0  0.932  0.0  0.0  0.0  0.0  0.0
6  11  1  1  1  0  0.09498  0.19890  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
6  12  1  1  1  0  0.12291  0.25581  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
6  13  1  1  1  0  0.06615  0.13027  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
7  8  1  1  1  0  0.0  0.17615  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
7  9  1  1  1  0  0.0  0.11001  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
9  10  1  1  1  0  0.03181  0.08450  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
9  14  1  1  1  0  0.12711  0.27038  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
10 11  1  1  1  0  0.08205  0.19207  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
12 13  1  1  1  0  0.22092  0.19988  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
13 14  1  1  1  0  0.17093  0.34802  0.0  0  0  0  0  0  0.0  0.0  0.0  0.0  0.0  0.0
-999
LOSS ZONES FOLLOWS 1 ITEMS
1 IEEE 14 BUS
-99
INTERCHANGE DATA FOLLOWS 1 ITEMS
1 2 Bus 2 HV 0.0 999.99 IEEE14 IEEE 14 Bus Test Case
-9
TIE LINES FOLLOWS 0 ITEMS
-999
END OF DATA

```

Come è facile notare, tale file è suddiviso in due sezioni ben distinte: nella prima sono contenute le informazioni legate ai nodi di rete, come ad esempio la loro tensione e potenza di riferimento, nella seconda, invece, sono raccolti i dati sulle adiacenze di rete che ci consentono di risalire alla struttura della matrice di ammettenza \mathbf{Y} . Una descrizione più accurata dei contenuti del file è disponibile all'interno di un ulteriore documento messo a disposizione dall'IEEE (si veda [3]).

2.2 Modello matematico associato al problema dell'OPF

È convenzione distinguere i vincoli di rete in funzione del tipo di nodo a cui ci si riferisce. In quest'ottica risulta molto utile la classificazione introdotta nel paragrafo precedente. Si richiede infatti che la tensione del nodo PCC sia fissata a priori, che i Loads presentino una predeterminata potenza complessa di lavoro e che la potenza attiva delle PEIs sia compresa fra un valore minimo e massimo ben precisi. Tali limitazioni vengono introdotte all'interno del nostro modello di rete per rendere esplicite le restrizioni dovute alle leggi fisiche che governano la produzione e trasmissione di energia elettrica, nonché a quelle che regolano il funzionamento delle apparecchiature. Dal punto di vista matematico ciò si traduce nella necessità da parte dei diversi nodi di rete di rispettare alcuni vincoli di uguaglianza, o disuguaglianza, secondo il seguente schema:

- Nodo PCC $\longrightarrow V_n = V_{PCC}$, con $n = \mathcal{I}_{PCC}$
- Nodi PEIs $\longrightarrow \underline{P}_{PEI,n} \leq P_n \leq \bar{P}_{PEI,n}$, con $n = \mathcal{I}_{PEI}$
- Nodi Loads $\longrightarrow S_n = S_{Load,n}$, con $n = \mathcal{I}_{Load}$.

Dal momento che sia i vincoli sulle potenze dei nodi PEI e Loads che, ovviamente, quello sulla tensione del PCC possono essere espressi in funzione del solo vettore delle tensioni di rete ($\mathbf{V} = [V_1, \dots, V_N]^T$), il nostro scopo principale sarà quello di determinare quali valori associare a \mathbf{V} affinché le condizioni sopra elencate siano rispettate. Ciò significa trovare i punti ammissibili del problema, che in generale potranno essere più di uno, solo uno o anche nessuno. Gli ultimi due casi, però, sono di scarso interesse perchè, in assenza di punti ammissibili, si dice che il problema non ha soluzione e l'unico modo per uscire dall'impasse è quello di rilassare i vincoli in gioco, mentre quando si ha a che fare con un unico punto ammissibile la parte di ottimizzazione dell'iniezione di potenza in rete non può essere affrontata e si perde dunque quello che è lo scopo principale dell'Optimal Power Flow.

Una volta imposti i vincoli, il passo successivo da compiere è determinare la cosiddetta *funzione obiettivo* del nostro modello. In un'ottica di risparmio energetico, lo scopo finale del nostro lavoro non può che essere quello di assicurare

il funzionamento della rete con il livello più basso possibile di produzione energetica. Per questo motivo, la nostra funzione obiettivo sarà rappresentata dalla minimizzazione della potenza attiva generata dal nodo PCC. Individuato anche quest'ultimo tassello, abbiamo finalmente a disposizione tutti gli elementi per delineare il modello matematico associato al problema dell'Optimal Power Flow nella sua interezza. Esso può essere formulato come segue:

$$\begin{aligned}
 & \text{minimize} && \Re[S_1] \\
 & \text{subject to} && V_1 = V_{PCC} \\
 & && S_n = S_{Load,n}, && n \in \mathcal{I}_{Load} \\
 & && \underline{P}_{PEI,n} \leq \Re[S_n] \leq \overline{P}_{PEI,n}, && n \in \mathcal{I}_{PEI}
 \end{aligned} \tag{2.12}$$

Ricordando la definizione di potenza complessa data da (2.7), si ha:

$$\begin{aligned}
 & \text{minimize} && \Re[V_1 I_1^*] \\
 & \text{subject to} && V_1 = V_{PCC} \\
 & && V_n I_n^* = S_{Load,n}, && n \in \mathcal{I}_{Load} \\
 & && \underline{P}_{PEI,n} \leq \Re[V_n I_n^*] \leq \overline{P}_{PEI,n}, && n \in \mathcal{I}_{PEI}
 \end{aligned} \tag{2.13}$$

Infine, passando alla notazione matriciale ed integrando quanto appena scritto con l'equazione (2.1), si riesce ad esprimere il problema dell'OPF in funzione del solo vettore delle tensioni, che risulta essere l'unica variabile da ottimizzare. Si ottiene infatti:

$$\begin{aligned}
 & \text{minimize} && \Re[\mathbf{V}^T \mathbf{e}_1 (\mathbf{Y}_1 \mathbf{V})^*] \\
 & \text{subject to} && \mathbf{e}_1^T \mathbf{V} = V_{PCC} \\
 & && \mathbf{V}^T \mathbf{e}_n (\mathbf{Y}_n \mathbf{V})^* = S_{Load,n}, && n \in \mathcal{I}_{Load} \\
 & && \underline{P}_{PEI,n} \leq \Re[\mathbf{V}^T \mathbf{e}_n (\mathbf{Y}_n \mathbf{V})^*] \leq \overline{P}_{PEI,n}, && n \in \mathcal{I}_{PEI}
 \end{aligned} \tag{2.14}$$

dove \mathbf{e}_n rappresenta un versore di lunghezza N il cui unico elemento non nullo e pari ad 1 si trova in corrispondenza dell'indice n , ossia:

$$\mathbf{e}_n = [e_{n,1}, \dots, e_{n,n-1}, e_{n,n}, e_{n,n+1}, \dots, e_{n,N}] = [0, \dots, 0, 1, 0, \dots, 0];$$

mentre con \mathbf{Y}_n si indica la riga n -esima della matrice di ammettenza di rete $\mathbf{Y} = \mathbf{A}^T [\text{diag}(\mathbf{Z}_{link})]^{-1} \mathbf{A}$, definita in (2.6). La difficoltà nella risoluzione di questo modello risiede nella non linearità delle equazioni con cui abbiamo a che fare, che lo rendono assimilabile ad un problema di *Programmazione Non Lineare* (PNL). Questo tipo di problema è molto studiato all'interno della comunità scientifica, ma gli attuali solutori non sono in grado di garantire l'ottimalità della

2. *OPTIMAL POWER FLOW*

soluzione trovata, se non sotto particolari ipotesi, come ad esempio la convessità del problema stesso. Nel prossimo capitolo cercheremo di analizzare più a fondo le caratteristiche del modello (2.14), proponendone anche una tecnica risolutiva che va sotto il nome di Interior Point Method.

Capitolo 3

Algoritmo Centralizzato

Nel capitolo precedente abbiamo visto che il problema dell'Optimal Power Flow può essere formalizzato attraverso un modello di programmazione non lineare, il quale ha in generale la seguente struttura:

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && g_i(x) \leq 0 \quad i = 1, \dots, m \\ & && h_j(x) = 0, \quad j = 1, \dots, p \end{aligned} \tag{3.1}$$

dove almeno una fra le funzioni f , g_i ed h_j risulta essere non lineare. Si noti che, per come sono stati definiti, i vincoli di (3.1) ammettono esclusivamente lo 0 come termine di paragone a destra dei simboli di uguaglianza e disuguaglianza. Ciò però non implica alcuna perdita di generalità dacchè è possibile ricondursi ad una notazione di questo tipo semplicemente sottraendo ad entrambi i membri un eventuale termine di destra non nullo indesiderato.

Come già anticipato, ad oggi non esistono algoritmi efficienti che siano in grado di garantire l'ottimalità della soluzione trovata per il problema (3.1) nella sua formulazione più generica. In effetti, esso può risultare addirittura intrattabile anche quando si ha a che fare con un numero limitato di variabili e vincoli. Per questo motivo, spesso gli approcci adottati nelle tecniche risolutive prevedono una sorta di compromesso fra bontà della soluzione trovata ed efficienza nel determinarla. Per comprendere in cosa consiste questo trade-off, è però necessario fornire preliminarmente alcune nozioni di base sui problemi di ottimizzazione.

3.1 Punti di ottimo globali e locali

Con riferimento al generico modello di ottimizzazione (3.1), la variabile $x \in \mathbb{R}^n$ è detta *variabile di ottimizzazione*, mentre $f : \mathbb{R}^n \rightarrow \mathbb{R}$ prende il nome di *funzione obiettivo* o *funzione di costo*. Le relazioni $g_i(x) \leq 0$ corrispondono ai vincoli di disuguaglianza e le corrispondenti funzioni $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ sono dette *funzioni dei vincoli di disuguaglianza*. Allo stesso modo le equazioni $h_j = 0$ rappresentano i vincoli di uguaglianza e le funzioni $h_j : \mathbb{R}^n \rightarrow \mathbb{R}$ sono dette *funzioni dei vincoli di uguaglianza*. Nel loro complesso le funzioni appena elencate definiscono il dominio \mathcal{D} del problema di ottimizzazione attraverso la seguente relazione:

$$\mathcal{D} = \mathbf{dom}(f) \cap \bigcap_{i=1}^m \mathbf{dom}(g_i) \cap \bigcap_{j=1}^p \mathbf{dom}(h_j). \quad (3.2)$$

Un qualsiasi punto $\tilde{x} \in \mathcal{D}$ è detto *ammissibile* se soddisfa le condizioni $g_i(\tilde{x}) \leq 0$, $i = 1, \dots, m$ e $h_j(\tilde{x}) = 0$, $j = 1, \dots, p$. A sua volta il modello di partenza è detto ammissibile se possiede almeno un punto con queste caratteristiche.

Il nostro obiettivo finale è chiaramente quello di individuare all'interno dell'insieme dei punti ammissibili quello, o quelli, che minimizzano la funzione costo $f(x)$ e che indicheremo con x^* . In corrispondenza a questi punti il modello (3.1) assumerà il valore ottimale p^* , che prende il nome di *ottimo globale*. Esso viene definito dalla seguente relazione:

$$p^* = \inf\{f(x) \mid g_i(x) \leq 0, i = 1, \dots, m; h_j(x) = 0, j = 1, \dots, p\}. \quad (3.3)$$

Nella notazione da noi adottata si ammette che p^* possa assumere tutti i valori compresi fra $-\infty$ e $+\infty$, estremi inclusi. In particolare, se si dovesse avere $p^* = +\infty$, si concluderà che il problema è irrisolvibile (o inammissibile), seguendo la convenzione standard che l'estremo inferiore di un insieme vuoto sia $+\infty$. D'altro canto, se si dovesse ottenere $p^* = -\infty$, si potrebbe asserire che il problema è illimitato inferiormente e non ci sarebbero restrizioni al processo di minimizzazione della funzione obiettivo.

Un ulteriore concetto su cui è utile soffermarsi è quello di *ottimo locale*: $\tilde{x} \in \mathbb{R}^n$ viene detto punto di ottimo locale se esiste $M > 0$ tale che:

$$f(\tilde{x}) = \inf\{f(z) \mid g_i(z) \leq 0, i = 1, \dots, m; h_j(z) = 0, j = 1, \dots, p; \|z - \tilde{x}\|_2 \leq M\} \quad (3.4)$$

o, in altre parole, se \tilde{x} risolve il modello:

$$\begin{aligned} & \text{minimize} && f(z) \\ & \text{subject to} && g_i(z) \leq 0 && i = 1, \dots, m \\ & && h_j(z) = 0, && j = 1, \dots, p \\ & && \|z - \tilde{x}\|_2 \leq M \end{aligned} \tag{3.5}$$

Ciò significa che \tilde{x} minimizza la funzione obiettivo unicamente in un intorno di \tilde{x} stesso. In Figura 3.1 viene data prova, anche visiva, di ciò che si intende con punto di ottimo globale e locale¹.

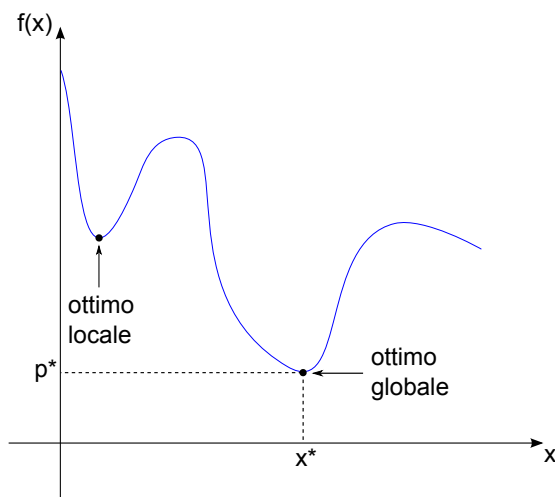


Figura 3.1: Ottimo locale e globale della funzione $f(x)$

Una delle principali carenze degli algoritmi connessi alla Programmazione Non Lineare consiste nel fatto che spesso essi si fermano in prossimità di punti di ottimo locale, senza raggiungere quelli globali da noi auspicati. È proprio questo il compromesso a cui ci si riferiva nella parte introduttiva del capitolo, quando si accennava alle tecniche di risoluzione dei modelli di PNL. Tale compromesso viene però evitato qualora si abbia a che fare con una particolare famiglia di problemi di ottimizzazione, ovvero quelli convessi.

¹Per semplicità di raffigurazione, in Figura 3.1 si è deciso di rappresentare il caso monodimensionale nel quale $x \in \mathbf{R}$. L'estensione al caso multidimensionale, in cui $x \in \mathbf{R}^n$, può essere attuata con opportuni accorgimenti, ma non avrebbe forse la stessa efficacia visiva del contesto monodimensionale.

3.2 Problemi di ottimizzazione convessi

In matematica una funzione $f : A \rightarrow \mathbb{R}$ viene detta convessa nel suo dominio A se:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad x, y \in A, \quad \lambda \in [0, 1] \quad (3.6)$$

Inoltre, se l'uguaglianza vale esclusivamente per $x = y$ oppure per $\lambda = 0$ o $\lambda = 1$, si parla di funzione strettamente convessa. In Figura 3.2 è rappresentata una curva associata ad una funzione f convessa.

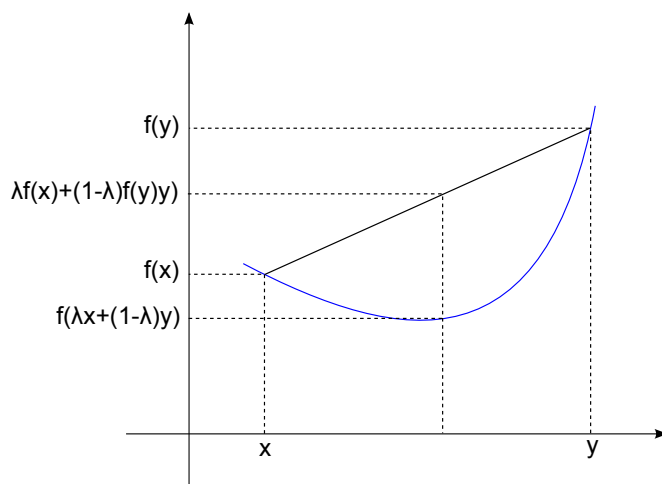


Figura 3.2: Esempio di funzione convessa f

La definizione appena fornita può essere estesa al caso multivariabile come segue: la funzione $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è convessa con parametro $M > 0$ se per ogni coppia di punti $x, y \in A$ del dominio si ha:

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq M \|x - y\|_2^2. \quad (3.7)$$

Chiaramente ciò presuppone che f sia differenziabile. Infine, una funzione f a valori reali viene detta concava se $-f$ risulta convessa. L'introduzione del concetto di convessità ci permette di definire una nuova tipologia di problemi, ai quali viene associato proprio l'epiteto di convessi.

Rispetto al generico modello (3.1), tre sono le caratteristiche che contraddistinguono un problema convesso:

- la funzione obiettivo $f(x)$ è convessa;

- le funzioni $g_i(x)$ associate ai vincoli di disuguaglianza sono convesse;
- le funzioni $h_j(x)$ associate ai vincoli d'uguaglianza sono lineari (o affini).

Esso può dunque essere formulato come segue:

$$\begin{aligned}
 & \text{minimize} && f(x) \\
 & \text{subject to} && g_i(x) \leq 0 \quad i = 1, \dots, m \\
 & && a_j^T x - b_j = 0, \quad j = 1, \dots, p
 \end{aligned} \tag{3.8}$$

Le caratteristiche distintive di (3.8) consentono di individuare alcune importanti proprietà che rendono un qualsiasi problema convesso decisamente più semplice da trattare matematicamente rispetto al generico modello di Programmazione Non Lineare (3.1).

Proposizione 3.1. *L'insieme dei punti ammissibili di un problema convesso è anch'esso convesso.*

Dimostrazione: La dimostrazione discende direttamente dal fatto che l'insieme dei punti ammissibili di un problema convesso è dato dall'intersezione del suo dominio \mathcal{D} , che è convesso, con m insiemi convessi del tipo $\{x | g_i(x) \leq 0\}$ e p iper-piani individuati dalle relazioni $\{x | a_j^T x - b_j = 0\}$. Questi ultimi sono chiaramente riconducibili ai vincoli di uguaglianza, individuati da funzioni lineari nel caso convesso. □

Più nello specifico, la prova appena fornita ci permette di concludere che anche l'insieme dei punti di ottimo di un problema convesso risulta convesso. Inoltre, se la funzione obiettivo $f(x)$ è strettamente convessa, sicuramente tale insieme contiene al massimo un punto.

Un'altra proprietà fondamentale dei problemi convessi è la seguente:

Proposizione 3.2. *In un problema convesso qualsiasi punto di ottimo locale è automaticamente un punto di ottimo globale.*

Dimostrazione: Si supponga che \tilde{x} sia un punto di ottimo locale per il problema di ottimizzazione convesso. In quanto tale, esso dovrà soddisfare i vincoli del modello (3.8) e in più dovrà verificare il seguente requisito:

$$f(\tilde{x}) = \inf \{ f(z) | g_i(z) \leq 0, i = 1, \dots, m; a_j^T z - b_j = 0, j = 1, \dots, p; \|z - \tilde{x}\|_2 \leq M \} \tag{3.9}$$

3. ALGORITMO CENTRALIZZATO

per qualche $M > 0$. Facciamo ora l'ipotesi che \tilde{x} non sia un punto di ottimo globale e che esista dunque un ulteriore valore y ammissibile tale che $f(y) < f(\tilde{x})$. Evidentemente si avrà $\|y - \tilde{x}\|_2 > M$, altrimenti, a causa di (3.9), si otterrebbe $f(\tilde{x}) \leq f(y)$, in contraddizione con quanto appena detto.

Si consideri adesso una nuova variabile v , definita dall'espressione:

$$v = (1 - \theta)\tilde{x} + \theta y, \quad \theta = \frac{M}{2\|y - \tilde{x}\|_2}$$

e si determini la differenza $v - \tilde{x}$ come segue:

$$v - \tilde{x} = (1 - \theta)\tilde{x} + \theta y - \tilde{x} = \tilde{x} - \theta\tilde{x} + \theta y - \tilde{x} = \theta(y - \tilde{x}) \quad (3.10)$$

Calcolandone la norma euclidea si ottiene:

$$\|v - \tilde{x}\|_2 = \|\theta\|_2 \|y - \tilde{x}\|_2 = \frac{M}{2\|y - \tilde{x}\|_2} \|y - \tilde{x}\|_2 = \frac{M}{2} < M. \quad (3.11)$$

Grazie alla convessità dell'insieme ammissibile², possiamo affermare che v stesso è un punto ammissibile per il problema di partenza. In più, essendo la funzione obiettivo f convessa, possiamo concludere che:

$$f(v) = f((1 - \theta)\tilde{x} + \theta y) \leq (1 - \theta)f(\tilde{x}) + \theta f(y) < f(\tilde{x}), \quad (3.12)$$

dove nell'ultima disuguaglianza si è sfruttato il fatto che $f(y) < f(\tilde{x})$ per ipotesi. Ciò però va in palese contrasto con quanto stabilito da (3.9). Siamo dunque giunti ad un assurdo, che ci porta a dire che non esiste alcun punto y ammissibile tale da verificare la relazione $f(y) < f(\tilde{x})$. Conseguentemente è lecito associare ad \tilde{x} l'appellativo di ottimo globale. \square

Si capisce a questo punto per quale motivo la convergenza degli algoritmi risolutivi presenti difficoltà minori nel caso in cui si abbia a che fare con problemi convessi. In tale circostanza, infatti, la limitazione dovuta all'individuazione dei punti di ottimo locali anzichè globali a cui si faceva riferimento nel paragrafo precedente, non ha più ragion d'essere.

Fino a questo momento abbiamo fornito la terminologia tipica dei modelli di Programmazione Non Lineare, ponendo particolare enfasi nell'analisi dei cosiddetti problemi convessi. Non abbiamo però ancora fornito alcuna delucidazione su quelle che sono ad oggi le tecniche risolutive maggiormente adottate in tale contesto. Sarà nostra premura farlo al più presto, non prima però di aver introdotto un nuovo, fondamentale concetto: quello di *funzione duale di Lagrange*.

²Si veda a tal proposito la **Proposizione 3.1**.

3.3 La funzione duale di Lagrange

Si faccia come sempre riferimento al modello (3.1), si assuma che il suo dominio

$$\mathcal{D} = \mathbf{dom}(f) \cap \bigcap_{i=1}^m \mathbf{dom}(g_i) \cap \bigcap_{j=1}^p \mathbf{dom}(h_j)$$

sia non vuoto e che esista un punto di ottimo globale p^* ad esso associato³.

L'idea che sta alla base della dualità Lagrangiana è di eliminare i vincoli dalla formulazione del modello (3.1) e di inserirli direttamente all'interno della funzione obiettivo, attraverso l'utilizzo di opportune somme pesate. Tale procedimento va sotto il nome di *rilassamento Lagrangiano* e porta alla definizione del cosiddetto *Lagrangiano* $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$, dato da:

$$L(x, \lambda, \nu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \nu_j h_j(x) \quad (3.13)$$

dove λ_i e ν_j sono chiamati *moltiplicatori di Lagrange* associati rispettivamente ai vincoli di disuguaglianza e di uguaglianza. Si noti come, a causa dell'eliminazione dei vincoli da (3.1), la regione ammissibile del problema rilassato comprenda in generale quella del problema originale. Per questa ragione non è detto che una soluzione ottima per il problema rilassato sia ammissibile per quello di partenza, ne' che il valore di tale soluzione coincida per i due modelli.

Ciononostante, l'impiego di tecniche di rilassamento⁴, se utilizzate con le dovute cautele, risulta essere estremamente utile nella soluzione di problemi di ottimizzazione per diversi motivi:

1. permette di avere una stima sul valore della soluzione ottima p^* ;
2. consente di accelerare la convergenza degli algoritmi di tipo enumerativo;
3. serve a prevedere quale sarebbe l'impatto di una certa scelta, ad esempio rimozione di alcuni vincoli, sul costo della soluzione ottima.

Nel caso specifico del rilassamento Lagrangiano, per perseguire tali obiettivi è necessario fare un ulteriore sforzo teorico ed introdurre la cosiddetta *funzione duale di Lagrange*:

$$d : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}.$$

³Si veda equazione (3.3) per la definizione di ottimo globale.

⁴Il rilassamento Lagrangiano non è l'unico conosciuto: esistono anche il *rilassamento per eliminazione* e quello *surrogato*.

3. ALGORITMO CENTRALIZZATO

Essa corrisponde al minimo del Lagrangiano rispetto alla variabile x ed è quindi data da:

$$d(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \nu_j h_j(x) \right) \quad (3.14)$$

Dal momento che la funzione duale di Lagrange rappresenta l'estremo inferiore di una famiglia di funzioni affini della coppia (λ, ν) , essa risulta concava indipendentemente dalla convessità o meno del problema di partenza.

3.3.1 Proprietà della funzione duale di Lagrange

È importante tenere a mente che la funzione duale di Lagrange è stata concepita per approssimare nella maniera più fedele possibile il generico modello di ottimizzazione (3.1). Ad essa viene quindi richiesto, in prima battuta, di fornirci informazioni su tale problema. A tal proposito torna utile il seguente risultato, che mette in relazione la funzione d con il valore della soluzione ottima di (3.1) da noi cercato.

Proposizione 3.3. $\forall \lambda \geq 0$ e $\forall \nu$ si ha:

$$d(\lambda, \nu) \leq p^*. \quad (3.15)$$

Dimostrazione: Supponiamo che \tilde{x} sia un punto ammissibile per il problema (3.1), ossia che siano verificate le condizioni $g_i(\tilde{x}) \leq 0$ e $h_j(\tilde{x}) = 0$. Avendo ipotizzato $\lambda \geq 0$, si ottiene:

$$\sum_{i=1}^m \lambda_i g_i(\tilde{x}) + \sum_{j=1}^p \nu_j h_j(\tilde{x}) \leq 0$$

infatti ogni termine appartenente alla prima sommatoria risulta essere non positivo e ciascun elemento della seconda risulta nullo. Si ha quindi:

$$L(\tilde{x}, \lambda, \nu) = f(\tilde{x}) + \sum_{i=1}^m \lambda_i g_i(\tilde{x}) + \sum_{j=1}^p \nu_j h_j(\tilde{x}) \leq f(\tilde{x})$$

e di conseguenza:

$$d(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \leq L(\tilde{x}, \lambda, \nu) \leq f(\tilde{x}) \quad (3.16)$$

Dal momento che la (3.16) è verificata per ogni punto \tilde{x} ammissibile, la proprietà (3.15) è dimostrata. \square

Abbiamo perciò individuato un lower bound per il valore di p^* , che risulta però banale nel caso in cui si ottenga $d(\lambda, \nu) = -\infty$. Tale evenienza è comunque scongiurata qualora si abbia $\lambda \geq 0$ e $(\lambda, \nu) \in \mathbf{dom}(d)$, dove con $\mathbf{dom}(d)$ viene indicato il dominio della funzione duale di Lagrange.

3.4 Il problema duale di Lagrange

Abbiamo appena affermato che per ciascuna coppia (λ, ν) con $\lambda \geq 0$, la funzione duale di Lagrange fornisce un lower bound sulla soluzione ottima del modello (3.1). Dal momento che tale lower bound dipende dai parametri λ e μ , è lecito chiedersi a questo punto quale sia il miglior lower bound che possiamo ottenere dalla funzione duale di Lagrange. Ciò significa domandarsi quali valori di λ e ν risolvono il seguente modello di ottimizzazione:

$$\begin{aligned} & \text{maximize} && d(\lambda, \nu) \\ & \text{subject to} && \lambda \geq 0 \end{aligned} \tag{3.17}$$

Il problema (3.17) viene chiamato *problema duale di Lagrange* associato al modello (3.1) originale. Quest'ultimo viene talvolta ribattezzato *problema primale*; inoltre i coefficienti λ^* e ν^* sono detti moltiplicatori di Lagrange ottimi se risolvono (3.17). In corrispondenza alla coppia (λ^*, ν^*) viene raggiunto il valore massimo della funzione $d(\lambda, \nu)$, che indicheremo con d^* e che coincide con il miglior lower-bound di p^* ottenibile attraverso il rilassamento Lagrangiano.

Si presti attenzione al fatto che il problema duale di Lagrange è di tipo convesso, indipendentemente dalla natura del problema primale. Questo perchè la funzione obiettivo $d(\lambda, \nu)$, che deve essere massimizzata, è concava e il vincolo $\lambda \geq 0$ è convesso. Una delle proprietà più semplici ma anche più importanti riconducibili al problema duale di Lagrange è la seguente:

Teorema 3.1. *Sia p^* la soluzione ottima di un qualsiasi problema primale e d^* quella del corrispondente problema duale, allora vale sicuramente la seguente disuguaglianza:*

$$d^* \leq p^*. \tag{3.18}$$

Dimostrazione: È sufficiente osservare che, in corrispondenza al punto d^* , ci troviamo sotto le ipotesi della **Proposizione 3.3** per concludere la dimostrazione. □

Questa proprietà è detta *dualità debole*. Essa vale anche quando p^* e d^* assumono valori infiniti. Per esempio, se il problema primale è illimitato inferiormente, cioè $p^* = -\infty$, si ha sicuramente $d^* = -\infty$ e quindi il problema duale di Lagrange è inammissibile (o irrisolvibile). Invertendo le parti, possiamo affermare che se il problema duale è illimitato superiormente, ossia $d^* = +\infty$, si deve per forza avere $p^* = +\infty$ e dunque il problema primale risulta inammissibile (o irrisolvibile).

È solito riferirsi alla differenza $p^* - d^*$ con l'epiteto di *marginale duale ottimo* dato che offre una misura di quanto si discosta il valore ottimo di (3.1) rispetto al suo miglior lower-bound ottenibile attraverso la funzione duale di Lagrange.

Proposizione 3.4. *Il margine duale ottimo $p^* - d^*$ è sempre non negativo.*

Dimostrazione: Discende immediatamente dal **Teorema 3.1**. □

L'approssimazione fornita da d^* può essere talvolta utilizzata per trovare un lower-bound a problemi di difficile soluzione, per i quali non interessa determinare il valore esatto di p^* , ma solo una sua stima. Come già anticipato in precedenza, infatti, la struttura del modello (3.17) è quella tipica di un problema convesso, certamente più facile da trattare rispetto al generico problema primale (3.1) di partenza.

Si faccia attenzione al fatto che l'espressione (3.18) non prevede che d^* sia strettamente inferiore a p^* , ma lascia aperta la possibilità che le due grandezze coincidano. Tale situazione è chiaramente quella da noi auspicata, in quanto ci permette di determinare il valore ottimo p^* senza alcun margine d'errore, semplicemente ricorrendo ad un rilassamento Lagrangiano. Vediamo quindi quali sono le condizioni che devono essere verificate affinché si abbia effettivamente $d^* = p^*$.

3.5 Dualità forte

Se l'uguaglianza:

$$d^* = p^* \tag{3.19}$$

è verificata, si dice che sussiste un relazione di *dualità forte* fra problema primale e duale. Alternativamente, si può affermare che se il margine duale ottimo fra i modelli (3.1) e (3.17) è nullo, vale per loro la dualità forte. Ovviamente quest'ultima non è sempre verificata a priori, ma se il problema di partenza è convesso, ovvero assume la forma (3.8), la relazione (3.19) spesso sussiste. Un semplice

risultato che ci consente di assicurare la dualità forte è costituito dal teorema di Slater:

Teorema 3.2. *La dualità forte è verificata quando il problema primale è convesso ed è soddisfatta la condizione di Slater, la quale impone che esista un valore $\tilde{x} \in \mathbf{relint}(\mathcal{D})$ ⁵ tale che:*

$$g_i(\tilde{x}) < 0, \quad i = 1, \dots, m; \quad A\tilde{x} - b = 0 \quad (3.20)$$

ovvero che \tilde{x} sia strettamente ammissibile.

Dimostrazione: Senza perdita di generalità facciamo due assunzioni ulteriori rispetto a quelle previste dall'enunciato del teorema, ossia che \mathcal{D} abbia interno non vuoto in \mathbf{R}^n e che quindi $\mathbf{relint}(\mathcal{D}) = \mathbf{int}(\mathcal{D})$ e inoltre che il rango della matrice A sia pieno ($\mathbf{rank}(A) = p$). Si supponga inoltre che p^* sia finito, visto che l'unico altro caso contemplato dalle ipotesi è che $p^* = -\infty$, gestibile attraverso la semplice dualità debole, la quale assicura che anche $d^* = -\infty$.

Definiamo innanzitutto l'insieme:

$$\mathcal{A} = \{(u, v, t) : \exists x \in \mathcal{D} | g_i(x) \leq u_i, i = 1, \dots, m, h_j(x) = v_j, j = 1, \dots, p, f(x) \leq t\} \quad (3.21)$$

che è convesso essendo $g_i(x)$ sono funzioni convesse. La soluzione ottima del nostro problema di partenza (3.1) è rappresentata da:

$$p^* = \inf\{t | (0, 0, t) \in \mathcal{A}\} \quad (3.22)$$

Si noti che se $\lambda \geq 0$:

$$d(\lambda, \nu) = \inf\{(\lambda, \nu, 1)^T(u, v, t) | (u, v, t) \in \mathcal{A}\}. \quad (3.23)$$

Dati $\lambda \geq 0$ e ν , se $d(\lambda, \nu)$ è finito allora:

$$(\lambda, \nu, 1)^T(u, v, t) \geq d(\lambda, \nu) \quad (3.24)$$

costituisce un iper-piano non verticale a supporto di \mathcal{A} . In particolare, dato che $(0, 0, p^*) \in \mathbf{bd}(\mathcal{A})$ ⁶, si ha:

$$p^* = (\lambda, \nu, 1)^T(u, v, t) \geq d(\lambda, \nu) \quad (3.25)$$

⁵Con la simbologia $\mathbf{relint}(\mathcal{D})$ si indica il più piccolo insieme di punti interni di \mathcal{D} relativamente ad H , dove \mathcal{D} è il dominio del problema di ottimizzazione in analisi, definito da (3.2), mentre H corrisponde al più piccolo sottospazio affine di \mathbf{R}^n contenente \mathcal{D} .

⁶Con la scrittura $\mathbf{bd}(\mathcal{A})$ si indica la frontiera dell'insieme \mathcal{A} stesso.

3. ALGORITMO CENTRALIZZATO

che corrisponde evidentemente a quanto affermato dalla dualità debole. Rispetto a quest'ultima, però, la formulazione (3.25) ci permette di fare un ragionamento ulteriore e di asserire che la dualità forte è assicurata se esiste un iper-piano non verticale a supporto di \mathcal{A} nel suo punto di frontiera $(0, 0, p^*)$.

Tenendo bene a mente questo risultato, introduciamo un secondo insieme convesso:

$$\mathcal{B} = \{(0, 0, s) \in \mathbf{R}^m \times \mathbf{R}^p \times \mathbf{R} \mid s < p^*\} \quad (3.26)$$

L'intersezione fra \mathcal{A} e \mathcal{B} è data dall'insieme vuoto. Per provare ciò si faccia l'ipotesi assurda che $(u, v, t) \in \mathcal{A} \cap \mathcal{B}$; questo implica che $(u, v, t) \in \mathcal{B}$ e dunque che $u = 0$, $t = 0$ e $t < p^*$, ma anche che $(u, v, t) \in \mathcal{A}$ e quindi che esista x tale che $g_i(x) \leq 0$, per $i = 1, \dots, m$, $Ax - b = 0$ e $f(x) \leq t$. Mettendo insieme i risultati appena trovati si ottiene la disuguaglianza:

$$f(x) \leq t < p^*$$

che è palesemente inammissibile essendo p^* il valore ottimo del problema primale.

Il ‘*separating hyperplane theorem*’ (si veda [4]) assicura che esiste una terna $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ ed uno scalare α tali che:

$$(u, v, t) \in \mathcal{A} \Rightarrow \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \geq \alpha \quad (3.27)$$

e

$$(u, v, t) \in \mathcal{B} \Rightarrow \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \leq \alpha. \quad (3.28)$$

Da (3.27) concludiamo che $\tilde{\lambda} \geq 0$ e $\mu \geq 0$, in caso contrario infatti la somma $\tilde{\lambda}^T u + \mu t$ sarebbe inferiormente limitata all'interno dell'insieme \mathcal{A} , in disaccordo con l'espressione (3.27) stessa. La condizione (3.28), da canto suo, implica semplicemente che $\mu t \leq \alpha$ per ogni $t < p^*$ e quindi $\mu p^* \leq \alpha$.

Aggregando i precedenti risultati, si deduce che:

$$\sum_{i=1}^m \tilde{\lambda}_i g_i(x) + \tilde{\nu}^T (Ax - b) + \mu f(x) \geq \alpha \geq \mu p^*. \quad (3.29)$$

Distinguiamo ora due casi:

Caso 1: $\mu > 0$.

Dividiamo (3.29) per μ , ottenendo:

$$L \left(x, \frac{\tilde{\lambda}}{\mu}, \frac{\tilde{\nu}}{\mu} \right) \geq p^* \quad (3.30)$$

per ogni $x \in \mathcal{D}$. Minimizzando rispetto ad x abbiamo:

$$d(\lambda, \nu) \geq p^* \tag{3.31}$$

dove:

$$\lambda = \frac{\tilde{\lambda}}{\mu}, \quad \nu = \frac{\tilde{\nu}}{\mu}.$$

La dualità debole impone invece la disuguaglianza $d(\lambda, \nu) \leq p^*$ che, confrontata con (3.31), restituisce $d(\lambda, \nu) = p^*$ dimostrando di fatto la dualità forte almeno quando $\mu > 0$.

Caso 2: $\mu = 0$.

Rielaborando la (3.29), si trova che, per ogni $x \in \mathcal{D}$,

$$\sum_{i=1}^m \tilde{\lambda}_i g_i(x) + \tilde{\nu}^T (Ax - b) \geq 0. \tag{3.32}$$

Considerando un punto \tilde{x} che soddisfa la condizione di Slater, la formula (3.32) assume un aspetto ancora più semplice, ossia:

$$\sum_{i=1}^m \tilde{\lambda}_i g_i(\tilde{x}) \geq 0. \tag{3.33}$$

Ora, dovendo per ipotesi essere verificate le relazioni $g_i(\tilde{x}) < 0$ e $\tilde{\lambda}_i \geq 0$, possiamo concludere che $\tilde{\lambda}_i = 0$. In precedenza avevamo visto inoltre che $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$, ma avendo stabilito che $\tilde{\lambda} = \mu = 0$, non può che essere $\tilde{\nu} \neq 0$.

Dalla disequazione (3.32) e dai ragionamenti appena condotti possiamo dedurre che per ogni $x \in \mathcal{D}$ si ha $\tilde{\nu}^T (Ax - b) \geq 0$. Tuttavia, il punto \tilde{x} soddisfa per ipotesi l'uguaglianza $\tilde{\nu}^T (Ax - b) = 0$ e, dato che $\tilde{x} \in \mathbf{int}(\mathcal{D})$, ciò implica che ci sono punti in \mathcal{D} per i quali $\tilde{\nu}^T (Ax - b) < 0$, a meno che non si abbia $A^T \tilde{\nu} = 0$. Quest'evenienza è però esclusa visto che contraddirebbe l'assunzione che la matrice A abbia rango pieno e ciò conclude la dimostrazione.

□

La condizione di Slater non solo implica che la dualità forte è conseguita per i problemi convessi, ma anche che il valore duale ottimo d^* è raggiunto qualora $d^* > -\infty$.

3.5.1 Condizione di complementarità

Si prenda ancora una volta come riferimento il modello (3.1) e si definisca il suo problema duale come visto in (3.17). Si supponga che fra i due sussista una relazione di dualità forte e che x^* e (λ^*, ν^*) siano i punti di ottimo del problema primale e duale rispettivamente. Ciò significa che:

$$p^* = f(x^*) = d(\lambda^*, \nu^*) = \inf_{x \in \mathcal{D}} \left(f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) + \sum_{j=1}^p \nu_j^* h_j(x) \right) \quad (3.34)$$

dove l'ultima uguaglianza discende dalla definizione della funzione duale di Lagrange.

Da quella di estremo inferiore, possiamo invece dedurre che:

$$\inf_{x \in \mathcal{D}} \left(f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) + \sum_{j=1}^p \nu_j^* h_j(x) \right) \leq f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{j=1}^p \nu_j^* h_j(x^*) \quad (3.35)$$

Infine, dato che $\lambda_i^* \geq 0, g_i(x^*) \leq 0, \forall i = 1, \dots, m$ e $h_j(x^*) = 0, \forall j = 1, \dots, p$, possiamo affermare che:

$$f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{j=1}^p \nu_j^* h_j(x^*) \leq f(x^*). \quad (3.36)$$

Mettendo a confronto le formule (3.34), (3.35) e (3.36), ci si accorge subito che le due disequazioni precedenti devono essere trasformate in relazioni di uguaglianza. Ciò ci permette di trarre alcune interessanti conclusioni. Prima di tutto, il fatto che la disuguaglianza (3.35) sia in realtà un'uguaglianza assicura che x^* minimizza $L(x, \lambda^*, \nu^*)$ rispetto alla variabile x . Questo non significa che il Lagrangiano non possa avere altri punti di minimo, ma semplicemente che x^* lo è sicuramente.

Un altro importante risultato è che:

$$\sum_{i=1}^m \lambda_i^* g_i(x^*) = 0. \quad (3.37)$$

Lo si può dedurre imponendo l'uguaglianza nella formula (3.36) e ricordando che $h_j(x^*) = 0, \forall j = 1, \dots, p$. Dal momento che ciascun termine della sommatoria (3.37) è non nullo si può inoltre concludere che:

$$\lambda_i^* g_i(x^*) = 0, \quad \forall i = 1, \dots, m \quad (3.38)$$

Questa proprietà è nota come ‘*condizione di complementarità*’ ed è verificata per qualsiasi coppia di ottimo primale p^* e duale (λ^*, ν^*) quando vale la dualità forte. La condizione (3.38) può essere espressa in forma alternativa come segue:

$$\lambda_i^* > 0 \Rightarrow g_i(x^*) = 0, \quad g_i(x^*) < 0 \Rightarrow \lambda_i^* = 0. \quad (3.39)$$

In parole povere, ciò significa che in corrispondenza del punto di ottimo l’ i -esimo moltiplicatore di Lagrange associato ai vincoli di disuguaglianza è nullo, a meno che tale vincolo non sia verificato con l’uguaglianza stretta.

3.6 Condizioni di Karush-Khun-Tucker

Si assuma ora che le funzioni $f, g_i, i = 1, \dots, m$ e $h_j, j = 1, \dots, p$ in (3.1) siano differenziabili, ma non si faccia alcuna ipotesi sulla loro convessità. Come fatto in passato, si indichino con x^* e (λ^*, ν^*) i punti di ottimo primale e duale rispettivamente e si ipotizzi che in corrispondenza ad essi il margine duale ottimo sia nullo (cioè che valga la dualità forte). Nell’ultima parte del paragrafo precedente abbiamo visto che x^* minimizza il Lagrangiano $L(x, \lambda^*, \nu^*)$ rispetto alla variabile x , perciò quest’ultimo deve avere gradiente nullo in x^* stesso:

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^p \nu_j^* \nabla h_j(x^*) = 0. \quad (3.40)$$

Questo, insieme ad altri risultati di cui abbiamo già dato prova in precedenza, consentono di formalizzare le cosiddette *condizioni di Karush-Khun-Tucker* (KKT), riportate qui di seguito:

$$\begin{aligned} \nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^p \nu_j^* \nabla h_j(x^*) &= 0 \\ g_i(x^*) &\leq 0, & i = 1, \dots, m \\ h_j(x^*) &= 0, & j = 1, \dots, p \\ \lambda_i^* &\geq 0, & i = 1, \dots, m \\ \lambda_i^* g_i(x^*) &= 0, & i = 1, \dots, m. \end{aligned} \quad (3.41)$$

Rielaborando quanto appena trovato, possiamo dire che una qualunque coppia di punti di ottimo primale e duale di un qualsiasi problema di ottimizzazione, caratterizzato da funzioni differenziabili e per il quale vale la dualità forte, deve per forza soddisfare le condizioni di Karush-Kuhn-Tucker. Queste ultime rappresentano dunque delle condizioni necessarie per l’ottimalità di x^* e (λ^*, ν^*) . Affinchè

3. ALGORITMO CENTRALIZZATO

esse risultino anche sufficienti è indispensabile che sia verificata un'altra ipotesi, ovvero la convessità del problema (3.1) di partenza.

Teorema 3.3. *Si consideri il modello di ottimizzazione (3.1). Se le funzioni f e $g_i, i = 1, \dots, m$ sono convesse, le $h_j, j = 1, \dots, p$ sono affini e i punti $\tilde{x}, \tilde{\lambda}$ e $\tilde{\nu}$ verificano le condizioni:*

$$\begin{aligned} \nabla f(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(\tilde{x}) + \sum_{j=1}^p \tilde{\nu}_j \nabla h_j(\tilde{x}) &= 0 \\ g_i(\tilde{x}) &\leq 0, & i = 1, \dots, m \\ h_j(\tilde{x}) &= 0, & j = 1, \dots, p \\ \tilde{\lambda}_i &\geq 0, & i = 1, \dots, m \\ \tilde{\lambda}_i g_i(\tilde{x}) &= 0, & i = 1, \dots, m. \end{aligned} \quad (3.42)$$

allora \tilde{x} e $(\tilde{\lambda}, \tilde{\nu})$ sono rispettivamente punti di ottimo primale e duale, con margine duale relativo nullo.

Dimostrazione: I vincoli $g_i(\tilde{x}) \leq 0, i = 1, \dots, m$ e $h_j(\tilde{x}) = 0, j = 1, \dots, p$ assicurano innanzitutto l'ammissibilità di \tilde{x} per il problema (3.1). Dal momento che $\tilde{\lambda}_i \geq 0, i = 1, \dots, m$, il Lagrangiano $L(x, \tilde{\lambda}, \tilde{\nu})$ risulta convesso in funzione di x ; inoltre l'ultima condizione di KKT ($\nabla f(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i \nabla g_i(\tilde{x}) + \sum_{j=1}^p \tilde{\nu}_j \nabla h_j(\tilde{x}) = 0$) impone che il gradiente di $L(x, \tilde{\lambda}, \tilde{\nu})$ rispetto ad x svanisca in prossimità di \tilde{x} . Da queste due considerazioni si deduce immediatamente che \tilde{x} minimizza il Lagrangiano $L(x, \tilde{\lambda}, \tilde{\nu})$, perciò:

$$d(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) = f(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i g_i(\tilde{x}) + \sum_{j=1}^p \tilde{\nu}_j h_j(\tilde{x}) = f(\tilde{x}) \quad (3.43)$$

dove nell'ultima uguaglianza si è sfruttato il fatto che $h_j(\tilde{x}) = 0, j = 1, \dots, p$ e $\tilde{\lambda}_i g_i(\tilde{x}) = 0, i = 1, \dots, m$. Da qui la prova che \tilde{x} e $(\tilde{\lambda}, \tilde{\nu})$ hanno margine duale nullo e che quindi sono punti di ottimo primale e duale rispettivamente. \square

Quindi, riassumendo, si ha che in presenza di un problema di ottimizzazione convesso, con funzione obiettivo e funzioni relative ai vincoli differenziabili e per il quale è verificata la condizione di Slater, i punti che soddisfano le condizioni di Karush-Kuhn-Tucker sono certamente punti di ottimo globali.

Anche nel caso di problemi non convessi e che non soddisfano la condizione di Slater, le condizioni di Karush-Kuhn-Tucker rappresentano comunque un importante parametro per la determinazione dell'ottimo. In questo caso però si parla

di ottimo locale (o meglio ottimi locali) e non più di quello globale. In particolare si dimostra che i punti candidati ad essere minimi locali di (3.1) possono essere determinati risolvendo le condizioni di KKT. Definiamo dunque i punti di Karush-Kuhn-Tucker del modello (3.1) come quei punti \bar{x} per i quali esiste una coppia $(\bar{\lambda}, \bar{\nu})$ in corrispondenza alla quale:

$$\begin{aligned}
 \nabla f(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla g_i(\bar{x}) + \sum_{j=1}^p \bar{\nu}_j \nabla h_j(\bar{x}) &= 0 \\
 g_i(\bar{x}) &\leq 0, & i = 1, \dots, m \\
 h_j(\bar{x}) &= 0, & j = 1, \dots, p \\
 \bar{\lambda}_i &\geq 0, & i = 1, \dots, m \\
 \bar{\lambda}_i g_i(\bar{x}) &= 0, & i = 1, \dots, m.
 \end{aligned} \tag{3.44}$$

I possibili candidati ad essere minimi locali sono proprio i punti di Karush-Kuhn-Tucker. Fra questi, il punto a cui può essere associato l'appellativo di ottimo globale è quello in corrispondenza al quale la funzione obiettivo f assume il valore più piccolo.

Abbiamo appena visto come le condizioni di Karush-Kuhn-Tucker giochino un ruolo fondamentale nel campo dell'ottimizzazione. In alcuni casi specifici, infatti, è possibile darne una soluzione analitica, che consente di determinare senza ulteriori elucubrazioni matematiche il punto di minimo del modello (3.1). Più in generale, anche quando ciò non è fattibile, molti degli algoritmi di ottimizzazione conosciuti sfruttano comunque le condizioni di KKT e anzi possono essere interpretati come metodi iterativi per la loro soluzione. Fra questi, uno dei più conosciuti è l'*Interior Point Method*[10]-[11], che ci accingiamo a presentare nel prossimo paragrafo.

3.7 Interior Point Method

Si consideri un generico problema di PNL dato da:

$$\begin{aligned}
 \text{minimize} \quad & f(x) \\
 \text{subject to} \quad & \underline{g}_i \leq g_i(x) \leq \bar{g}_i, \quad i = 1, \dots, m \\
 & h_j(x) = 0, \quad j = 1, \dots, p
 \end{aligned} \tag{3.45}$$

o, nella forma compatta:

$$\begin{aligned}
 \text{minimize} \quad & f(x) \\
 \text{subject to} \quad & \underline{g} \leq g(x) \leq \bar{g} \\
 & h(x) = 0
 \end{aligned} \tag{3.46}$$

3. ALGORITMO CENTRALIZZATO

Utilizzando delle variabili ausiliarie $u_i, i = 1, \dots, m$ ed $l_i, i = 1, \dots, m$ associate ai vincoli di disuguaglianza, è possibile rielaborare il modello (3.45) come segue:

$$\begin{aligned}
 & \text{minimize} && f(x) \\
 & \text{subject to} && g_i(x) + u_i - \bar{g}_i = 0, \quad i = 1, \dots, m \\
 & && g_i(x) - l_i - \underline{g}_i = 0, \quad i = 1, \dots, m \\
 & && h_j(x) = 0, \quad j = 1, \dots, p \\
 & && u_i \geq 0 \quad i = 1, \dots, m \\
 & && l_i \geq 0 \quad i = 1, \dots, m
 \end{aligned} \tag{3.47}$$

Il nostro primo obiettivo è quello di eliminare i nuovi vincoli di disuguaglianza $u \geq 0$ e $l \geq 0$, tenendone conto esclusivamente all'interno della funzione obiettivo. Per fare ciò definiamo la funzione $I_+ : \mathbb{R} \rightarrow \mathbb{R}$:

$$I_+(x) = \begin{cases} 0, & x \geq 0 \\ +\infty, & x < 0 \end{cases} \tag{3.48}$$

e sfruttiamola all'interno di (3.47) in questo modo:

$$\begin{aligned}
 & \text{minimize} && f(x) + \sum_{i=1}^m I_+(u_i) + \sum_{i=1}^m I_+(l_i) \\
 & \text{subject to} && g_i(x) + u_i - \bar{g}_i = 0, \quad i = 1, \dots, m \\
 & && g_i(x) - l_i - \underline{g}_i = 0, \quad i = 1, \dots, m \\
 & && h_j(x) = 0, \quad j = 1, \dots, p
 \end{aligned} \tag{3.49}$$

I modelli (3.47) e (3.49) coincidono perfettamente perchè quando $u_i \geq 0, \forall i = 1, \dots, m$ e $l_i \geq 0, \forall i = 1, \dots, m$, ovvero quando i vincoli di disuguaglianza di (3.47) sono verificati, di fatto non viene aggiunto nulla alla funzione costo di (3.49), mentre quando ciò non accade quest'ultima viene forzata ad assumere il valore $+\infty$ e ciò è corretto in quanto corrisponde a rendere il problema (3.49) inammissibile.

A questo punto il nostro scopo iniziale di generare un modello privo di vincoli di disuguaglianza partendo da (3.47) sembra raggiunto. L'inconveniente sta però nel fatto che la nuova funzione costo non è in generale differenziabile a causa dei termini aggiuntivi $\sum_{i=1}^m I_+(u_i)$ e $\sum_{i=1}^m I_+(l_i)$ e ciò non ci piace. Introduciamo quindi una nuova tecnica che va sotto il nome di '*barrier method*'.

3.7.1 Il barrier method

L'idea che sta alla base del cosiddetto *barrier method* è quella di approssimare la funzione I_+ nel seguente modo:

$$I_+(x) \simeq -\mu \ln(x) \tag{3.50}$$

dove $\mu > 0$ è un parametro che influenza l'accuratezza dell'approssimazione ottenuta. Chiaramente, più μ decresce e più l'approssimazione sarà buona.

Sostituendo (3.50) in (3.49), si ottiene:

$$\begin{aligned} \text{minimize} \quad & f(x) - \mu \sum_{i=1}^m \ln(u_i) - \mu \sum_{i=1}^m \ln(l_i) \\ \text{subject to} \quad & g_i(x) + u_i - \bar{g}_i = 0, & i = 1, \dots, m \\ & g_i(x) - l_i - \underline{g}_i = 0, & i = 1, \dots, m \\ & h_j(x) = 0, & j = 1, \dots, p \end{aligned} \tag{3.51}$$

e finalmente la funzione costo è tornata ad essere differenziabile. A questo punto è possibile effettuare il rilassamento Lagrangiano di (3.51) ed estrarne quindi il Lagrangiano associato, che risulta essere pari a:

$$\begin{aligned} L(x, l, u; y, z, w) = & f(x) - \mu \sum_{i=1}^m \ln(l_i) - \mu \sum_{i=1}^m \ln(u_i) - y^T h(x) \\ & - z^T (g(x) - l - \underline{g}) - w^T (g(x) + u - \bar{g}) \end{aligned} \tag{3.52}$$

Ricordando quanto stabilito nel paragrafo 3.6, risulta palese come la determinazione dei punti di ottimo (siano essi locali o globali) associati a (3.51) debba passare attraverso il calcolo delle condizioni di Karush-Kuhn-Tucker. Nel contesto in analisi esse assumono la seguente struttura⁷:

$$\begin{aligned} L_x &= \nabla L(x, l, u; y, z, w) = \nabla f(x) - \sum_{j=1}^p y_j \nabla h_j(x) - \sum_{i=1}^m (z_i + w_i) \nabla g_i(x) = 0 \\ L_l &= -\frac{\mu}{l_i} + z_i = 0 \Rightarrow L_l = z_i l_i - \mu = 0 \\ L_u &= -\frac{\mu}{u_i} - w_i = 0 \Rightarrow L_u = w_i u_i + \mu = 0 \\ L_y &= h_j(x) =, \\ L_z &= g_i(x) - l_i - \underline{g}_i = 0 \\ L_w &= g_i(x) + u_i - \bar{g}_i = 0. \end{aligned} \tag{3.53}$$

⁷Da qui fino alla fine del paragrafo si sottointenderà che $i \in \{1, \dots, m\}$ e $j \in \{1, \dots, p\}$ per non appesantire eccessivamente la notazione.

3. ALGORITMO CENTRALIZZATO

Per determinare i punti che soddisfano le condizioni (3.53), l'Interior Point Method prevede l'utilizzo della tecnica di Newton-Raphson, la quale può essere così declinata:

Fase 1: Adottando una tecnica di linearizzazione, si approssima il valore di una qualsivoglia funzione $f(x)$, scegliendo x_n come punto iniziale, attraverso l'espressione:

$$f(x) \simeq f(x_n) + \nabla f(x_n) \Delta x \quad (3.54)$$

da cui si ricava:

$$\Delta x = -(\nabla f(x_n))^{-1} f(x_n). \quad (3.55)$$

Fase 2: Viene aggiornato il punto iniziale per la linearizzazione x_n attraverso la relazione:

$$x_{n+1} = x_n + \Delta x \quad (3.56)$$

dove Δx è quello calcolato in (3.55).

Fase 3: Le prime due fasi vengono reiterate fino a quando non si giunge a convergenza.

Sfruttando il metodo di Newton-Raphson per calcolare una soluzione numerica delle equazioni di Karush-Kuhn-Tucker, le espressioni analoghe a (3.54) da adottare sono:

$$\begin{aligned} L_x + H \Delta x - \sum_{j=1}^p \Delta y_j \nabla h_j(x) - \sum_{i=1}^m (\Delta z_i + \Delta w_i) \nabla g_i(x) &= 0 \\ L_l + l_i \Delta z_i + z_i \Delta l_i &= 0 \\ L_u + u_i \Delta w_i + w_i \Delta u_i &= 0 \\ L_y + (\nabla h_j(x))^T \Delta x &= 0 \\ L_z + (\nabla g_i(x))^T \Delta x - \Delta l_i &= 0 \\ L_w + (\nabla g_i(x))^T \Delta x + \Delta u_i &= 0. \end{aligned} \quad (3.57)$$

dove:

$$H = \nabla^2 f(x) - \sum_{j=1}^p y_j \nabla^2 h_j(x) - \sum_{i=1}^m (z_i + w_i) \nabla^2 g_i(x). \quad (3.58)$$

Invertendo le formule precedenti si ottiene:

$$\begin{aligned} \Delta l_i &= L_z + (\nabla g_i(x))^T \Delta x \\ \Delta u_i &= -L_w - (\nabla h_j(x))^T \Delta x \\ \Delta z_i &= -\frac{L_l + z_i \Delta l_i}{l_i} = -\frac{L_l + z_i (L_z + (\nabla g_i(x))^T \Delta x)}{l_i} \\ \Delta w_i &= -\frac{L_u + w_i \Delta u_i}{u_i} = -\frac{L_u - w_i (L_w + (\nabla g_i(x))^T \Delta x)}{u_i} \end{aligned} \quad (3.59)$$

inoltre gli offset Δx e Δy sono identificati dalla soluzione del sistema:

$$\begin{bmatrix} H' & -\nabla h(x) \\ -(\nabla h(x))^T & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} L'_x \\ L_y \end{bmatrix} \quad (3.60)$$

e quindi da:

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} H' & -\nabla h(x) \\ -(\nabla h(x))^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} L'_x \\ L_y \end{bmatrix} \quad (3.61)$$

dove:

$$H' = H + \sum_{i=1}^m \left(\frac{z_i}{l_i} - \frac{w_i}{u_i} \right) \nabla g_i(x) (\nabla g_i(x))^T \quad (3.62)$$

mentre:

$$L'_x = -L_x - \sum_{i=1}^m \left(\frac{L_l + z_i L_z}{l_i} + \frac{L_u - w_i L_w}{u_i} \right) \nabla g_i(x). \quad (3.63)$$

A questo punto sia le variabili primali (x, l, u) che quelle duali (y, z, w) possono essere aggiornate sfruttando i corrispettivi offset calcolati in (3.59) e (3.61). Nel fare ciò è però opportuno sottolineare che, per come è stato determinato, il problema (3.51) richiede che le variabili ausiliarie l e u e i moltiplicatori z e w risultino sempre non negativi. L'unico modo per assicurare il rispetto di questo vincolo è affinare la fase di aggiornamento ricorrendo ai seguenti coefficienti:

$$\begin{aligned} \alpha_p &= \min \left\{ 0.9995 \min_i \left(\frac{-l_i}{\Delta l_i}, \Delta l_i < 0; \frac{-u_i}{\Delta u_i}, \Delta u_i < 0 \right), 1 \right\} \\ \alpha_d &= \min \left\{ 0.9995 \min_i \left(\frac{-z_i}{\Delta z_i}, \Delta z_i < 0; \frac{-w_i}{\Delta w_i}, \Delta w_i > 0 \right), 1 \right\}. \end{aligned} \quad (3.64)$$

In particolare, α_p e α_d modificano il criterio di aggiornamento nel seguente modo:

$$\begin{aligned} x &= x + \alpha_p \Delta x & y &= y + \alpha_d \Delta y \\ l &= l + \alpha_p \Delta l & z &= z + \alpha_d \Delta z \\ u &= u + \alpha_p \Delta u & w &= w + \alpha_d \Delta w \end{aligned} \quad (3.65)$$

È finalmente possibile procedere con la fase iterativa del metodo di Newton-Raphson, adottando come nuovo punto iniziale quello determinato dai valori calcolati in (3.65). Nel nostro caso specifico l'iterazione viene conclusa non appena le condizioni di Karush-Kuhn-Tucker in (3.53) sono verificate con sufficiente precisione.

3.7.2 Applicazione allo studio dell'Optimal Power Flow

Ora che sono stati esaminati i vari step che compongono l'Interior Point Method, è finalmente possibile esportare tale tecnica all'interno del contesto dell'Optimal Power Flow. Per fare ciò è necessario considerare il modello (2.14) come punto di partenza. Esso presenta la seguente struttura:

$$\begin{aligned}
 & \text{minimize} && \Re [\mathbf{V}^T \mathbf{e}_1(\mathbf{Y}_1 \mathbf{V})^*] \\
 & \text{subject to} && \mathbf{e}_1^T \mathbf{V} = V_{PCC} \\
 & && \mathbf{V}^T \mathbf{e}_n(\mathbf{Y}_n \mathbf{V})^* = S_{Load,n}, && n \in \mathcal{I}_{Load} \\
 & && \underline{P}_{PEI,n} \leq \Re [\mathbf{V}^T \mathbf{e}_n(\mathbf{Y}_n \mathbf{V})^*] \leq \overline{P}_{PEI,n}, && n \in \mathcal{I}_{PEI}
 \end{aligned} \tag{3.66}$$

ed è stato formalizzato grazie allo studio di una generica rete elettrica effettuato nel capitolo 2. Qui di seguito si ricorda solo in cosa consistono le principali grandezze che lo caratterizzano:

- N = numero di nodi di cui è composta la rete;
- \mathbf{V} = vettore contenente le tensioni dei nodi di rete;
- \mathbf{Y} = matrice di ammettenza di rete;
- \mathbf{e}_i = versore di lunghezza N costituito da un unico elemento non nullo e pari ad 1 in posizione i ;
- $S_i(\mathbf{V})$ = potenza complessa relativa al nodo i , $i = 1, \dots, N$;
- \mathcal{I}_{PEI} = insieme contenente gli indici dei nodi generatori (PEIs);
- \mathcal{I}_{Load} = insieme contenente gli indici dei Loads.

Seguendo i dettami visti nel paragrafo precedente, si capisce come l'implementazione dell'Interior Point Method passi attraverso i seguenti step:

1. Definire le condizioni di Karush-Kuhn-Tucker associate a (3.66), ovvero:

$$\begin{aligned}
 L_{\mathbf{V}} &= \Re[S_1(\mathbf{V}) - y \nabla(\mathbf{e}_1^T \mathbf{V}) - \sum_{i \in \mathcal{I}_{Load}} v_i \nabla S_i(\mathbf{V}) \\
 &\quad - \sum_{i \in \mathcal{I}_{PEI}} (z_i + w_i) \nabla \Re[S_i(\mathbf{V})] = 0 \\
 L_l &= z_i l_i - \mu = 0, & i \in \mathcal{I}_{PEI} \\
 L_u &= w_i u_i + \mu = 0, & i \in \mathcal{I}_{PEI} \\
 L_y &= \mathbf{e}_1^T \mathbf{V} = V_{PCC} \\
 L_v &= S_i(\mathbf{V}) - S_{Load,i} = 0, & i \in \mathcal{I}_{Load} \\
 L_z &= \Re[S_i(\mathbf{V})] - l_i - \underline{P}_{PEI,i} = 0, & i \in \mathcal{I}_{PEI} \\
 L_w &= \Re[S_i(\mathbf{V})] + u_i - \underline{P}_{PEI,i} = 0, & i \in \mathcal{I}_{PEI}
 \end{aligned} \tag{3.67}$$

dove si è sfruttato il fatto che per convenzione si associa al nodo 1 il ruolo del PCC.

2. Linearizzare le equazioni (3.67):

$$\begin{aligned}
 L_{\mathbf{V}} + H \Delta \mathbf{V} - \nabla(\mathbf{e}_1^T \mathbf{V}) \Delta y - \sum_{i \in \mathcal{I}_{Load}} \nabla S_i(\mathbf{V}) \Delta v_i \\
 - \sum_{i \in \mathcal{I}_{PEI}} \nabla S_i(\mathbf{V}) \nabla \Re[S_i(\mathbf{V})] (\Delta z_i + \Delta w_i) = 0 \\
 L_l + z_i \Delta l_i + l_i \Delta z_i = 0, & i \in \mathcal{I}_{PEI} \\
 L_u + w_i \Delta u_i + u_i \Delta w_i = 0, & i \in \mathcal{I}_{PEI} \\
 L_y + \nabla(\mathbf{e}_1^T \mathbf{V}) \Delta \mathbf{V} = 0 \\
 L_z + (\nabla S_i(\mathbf{V}))^T \Delta \mathbf{V} = 0, & i \in \mathcal{I}_{Load} \\
 L_z + (\nabla \Re[S_i(\mathbf{V})])^T \Delta \mathbf{V} - \Delta l_i = 0, & i \in \mathcal{I}_{PEI} \\
 L_w + (\nabla \Re[S_i(\mathbf{V})])^T \Delta \mathbf{V} + \Delta u_i = 0, & i \in \mathcal{I}_{PEI}
 \end{aligned} \tag{3.68}$$

dove :

$$\begin{aligned}
 H &= \nabla^2 \Re[S_1(\mathbf{V})] - y \nabla^2(\mathbf{e}_1^T \mathbf{V}) - \sum_{i \in \mathcal{I}_{Load}} v_i \nabla^2 S_i(\mathbf{V}) \\
 &\quad - \sum_{i \in \mathcal{I}_{PEI}} \nabla S_i(\mathbf{V}) (z_i + w_i) \nabla^2 \Re[S_i(\mathbf{V})]
 \end{aligned} \tag{3.69}$$

3. Invertire le uguaglianze (3.68) per determinare degli offset relativi a variabili

3. ALGORITMO CENTRALIZZATO

primali e duali:

$$\begin{aligned}
\Delta l_i &= L_z + (\nabla \mathfrak{R}[S_i(\mathbf{V})])^T \Delta \mathbf{V}, & i \in \mathcal{I}_{PEI} \\
\Delta u_i &= -L_w - (\nabla \mathfrak{R}[S_i(\mathbf{V})])^T \Delta \mathbf{V}, & i \in \mathcal{I}_{PEI} \\
\Delta z_i &= -\frac{L_l + z_i \Delta l_i}{l_i}, & i \in \mathcal{I}_{PEI} \\
\Delta w_i &= -\frac{L_u + w_i \Delta u_i}{u_i}, & i \in \mathcal{I}_{PEI}
\end{aligned} \tag{3.70}$$

e ancora:

$$\begin{bmatrix} \Delta \mathbf{V} \\ \Delta y \\ \Delta \mathbf{v} \end{bmatrix} = \begin{bmatrix} H' & -\nabla(\mathbf{e}_1^T \mathbf{V}) & -\nabla S_i(\mathbf{V}) \\ -(\nabla(\mathbf{e}_1^T \mathbf{V}))^T & 0 & 0 \\ -(\nabla S_i(\mathbf{V}))^T & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} L'_V \\ L_y \\ L_v \end{bmatrix}. \tag{3.71}$$

In (3.71) si ha $i \in \mathcal{I}_{Load}$ e inoltre le quantità H' e L'_V sono date rispettivamente da:

$$H' = H + \sum_{i \in \mathcal{I}_{PEI}} \left(\frac{z_i}{l_i} - \frac{w_i}{u_i} \right) \nabla \mathfrak{R}[S_i(\mathbf{V})] (\nabla \mathfrak{R}[S_i(\mathbf{V})])^T \tag{3.72}$$

e da:

$$L'_V = -L_V - \sum_{i \in \mathcal{I}_{PEI}} \left(\frac{L_l + z_i L_z}{l_i} - \frac{L_u - w_i L_w}{u_i} \right) \nabla \mathfrak{R}[S_i(\mathbf{V})]. \tag{3.73}$$

4. Definire i moltiplicatori α_p e α_d esattamente come è stato fatto in (3.64).

5. Aggiornare variabili primali e duali secondo le relazioni:

$$\begin{aligned}
\mathbf{V} &= \mathbf{V} + \alpha_p \Delta \mathbf{V} \\
\mathbf{l} &= \mathbf{l} + \alpha_p \Delta \mathbf{l} \\
\mathbf{u} &= \mathbf{u} + \alpha_p \Delta \mathbf{u} \\
y &= y + \alpha_d \Delta y \\
\mathbf{v} &= \mathbf{v} + \alpha_d \Delta \mathbf{v} \\
\mathbf{z} &= \mathbf{z} + \alpha_d \Delta \mathbf{z} \\
\mathbf{w} &= \mathbf{w} + \alpha_d \Delta \mathbf{w}.
\end{aligned} \tag{3.74}$$

6. Se tali valori soddisfano le equazioni (3.67) con sufficiente precisione concludere l'elaborazione, altrimenti ripetere i punti adottando come nuovo stato di partenza quello risultante da (3.74).

Viste le tante variabili in gioco, la convergenza dell'algoritmo appena presentato è tutt'altro che scontata[8]. Da questo punto di vista uno dei parametri più delicati è rappresentato dal coefficiente μ . Come già anticipato, esso determina l'accuratezza con cui i vincoli di disuguaglianza del modello (3.66) sono rispettati. L'idea per favorire la convergenza è quella di rilassare, almeno inizialmente, tali vincoli associando a μ un valore sufficientemente grande. In questo modo, durante le prime iterazioni, l'IPM è in grado di avvicinarsi al punto di ottimo facendo leva su un momentaneo ampliamento del dominio di ammissibilità del problema. In una fase successiva sarà chiaramente necessario inasprire nuovamente i vincoli di disuguaglianza, ricorrendo ad una progressiva riduzione del valore di μ . Ciò che bisogna fare nella pratica per concretizzare tale ragionamento è aggiornare ad ogni ciclo dell'IPM il parametro μ attenendosi alla relazione:

$$\mu = 0.2 \left(\frac{GAP}{2N_{PEI}} \right) = 0.2 \left(\frac{\mathbf{l}^T \mathbf{z} - \mathbf{u}^T \mathbf{w}}{2N_{PEI}} \right) \quad (3.75)$$

dove la grandezza $GAP = \mathbf{l}^T \mathbf{z} - \mathbf{u}^T \mathbf{w}$ è chiamata *complementary gap*.

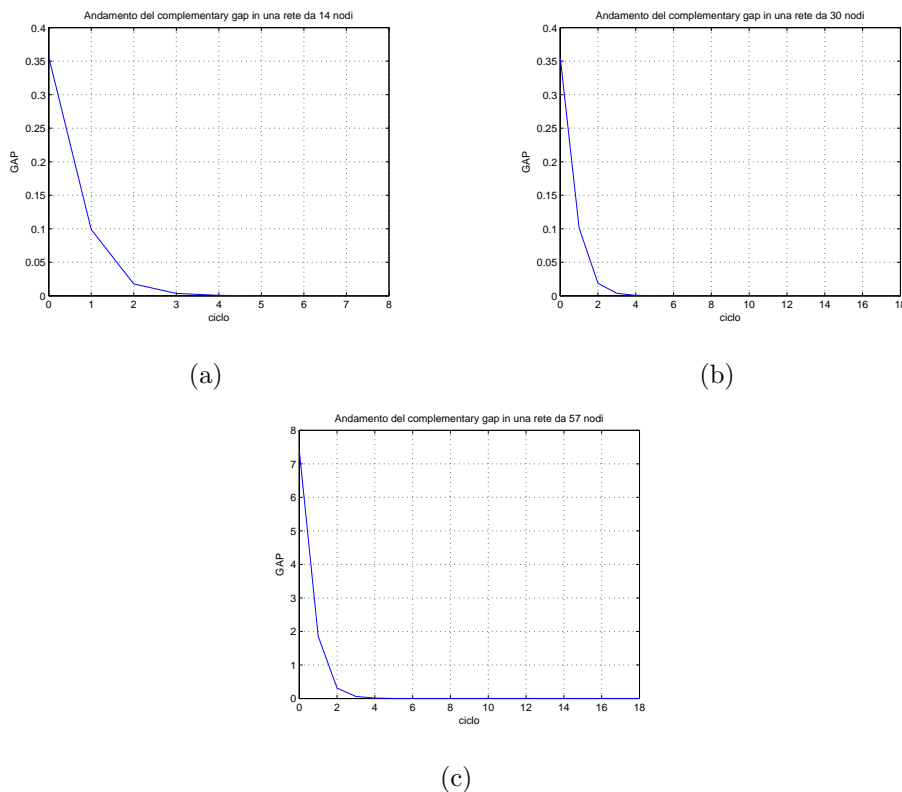


Figura 3.3: Progressivo annullamento del *complementary gap* in una rete da: (a) 14 nodi, (b) 30 nodi, (c) 57 nodi.

3. ALGORITMO CENTRALIZZATO

Dai grafici 3.3(a), 3.3(b) e 3.3(c)⁸ si può apprezzare come il complementary gap abbia un andamento monotono decrescente nel tempo, che lo porta in poche iterazioni ad annullarsi. La relazione di proporzionalità diretta vigente fra GAP e μ assicura quindi per quest'ultimo un trend simile, che coincide esattamente con quello da noi auspicato in precedenza. Si noti inoltre come le figure testimonino indirettamente il progressivo avvicinamento nel tempo fra il valore delle variabili primali \mathbf{l} e \mathbf{u} e di quelle duali \mathbf{u} e \mathbf{w} .

Un altro espediente spesso utilizzato per favorire la convergenza dell'algoritmo è rappresentato dalla *tecnica del backtracking*. Essa risulta efficace soprattutto quando la funzione costo ha una derivata molto grande in prossimità del punto di ottimo a cui vogliamo tendere con l'IPM. Una situazione di questo genere è raffigurata nell'immagine sottostante, che prenderemo come riferimento per spiegare la logica che sta dietro al backtracking.

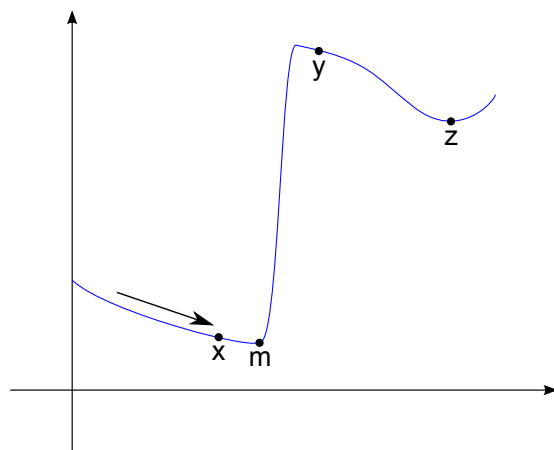


Figura 3.4: Situazione critica per la convergenza dell'Interior Point Method, in cui risulta efficace l'impiego del *backtracking*.

Si ipotizzi dunque che, nel suo processo di minimizzazione, l'IPM si stia avvicinando al punto di ottimo m seguendo la direzione indicata dalla freccia in figura e che all'iterazione i esso si trovi in corrispondenza del punto x . Al ciclo successivo ($i + 1$) gli offset calcolati tramite (3.70) potrebbero essere tali da far superare il valore minimo m e portare quindi il sistema in uno stato decisamente

⁸Le immagini in questione sono state ricavate implementando l'Interior Point Method per risolvere il problema dell'Optimal Power Flow associato a tre reti di dimensione diversa. Esse sono fornite dall'IEEE e verranno analizzate con maggior accuratezza all'interno del capitolo 5.

distante da quello ottimale. Il rischio maggiore è che il peggioramento sia così marcato da farci finire, ad esempio, nel punto y . In questo caso, infatti, le iterazioni dell'IPM successive alla $i + 1$ -esima saranno effettuate in direzione di z e non più di m . L'unico modo per ovviare a questo problema è evidentemente quello di agire a livello degli offset (3.70) ed è proprio ciò che fa il backtracking. Nello specifico, quest'ultimo prevede che l'aggiornamento delle variabili tramite (3.74) sia effettuato solo se le condizioni di KKT sono verificate con una miglior approssimazione utilizzando le variabili così aggiornate rispetto a quanto non accadeva con quelle di partenza. Se ciò non si verifica ad ogni offset viene applicato un fattore di scala $0 < \gamma < 1$ e si ridefinisce di conseguenza l'aggiornamento delle variabili come di seguito:

$$\begin{aligned}
 \mathbf{V} &= \mathbf{V} + \gamma\alpha_p\Delta\mathbf{V} \\
 \mathbf{l} &= \mathbf{l} + \gamma\alpha_p\Delta\mathbf{l} \\
 \mathbf{u} &= \mathbf{u} + \gamma\alpha_p\Delta\mathbf{u} \\
 y &= y + \gamma\alpha_d\Delta y \\
 \mathbf{v} &= \mathbf{v} + \gamma\alpha_d\Delta\mathbf{v} \\
 \mathbf{z} &= \mathbf{z} + \gamma\alpha_d\Delta\mathbf{z} \\
 \mathbf{w} &= \mathbf{w} + \gamma\alpha_d\Delta\mathbf{w}.
 \end{aligned} \tag{3.76}$$

Se, anche successivamente a questa modifica, l'approssimazione risulta peggiore di quella che si aveva inizialmente, è sufficiente ripetere il procedimento rimpicciolendo γ .

Per concludere si faccia attenzione ad un particolare: nel paragrafo 3.6 abbiamo stabilito che risolvere le equazioni di Karush-Kuhn-Tucker equivale a determinare un punto di ottimo globale solo nel caso in cui il problema di partenza sia di tipo convesso. In generale tale condizione non è verificata per il modello di ottimizzazione relativo all'Optimal Power Flow da noi presentato, sebbene le reti IEEE utilizzate per le simulazioni siano notoriamente quasi convesse. È quindi importante sottolineare come la soluzione determinata in seguito all'implementazione dell'Interior Point Method identifichi un punto di ottimo che in generale è locale e non globale. Come se non bastasse, il suo valore e la stessa capacità di convergenza dell'algoritmo dipendono fortemente da come viene scelto lo stato di partenza per l'applicazione del metodo di Newton-Raphson. Per tenere conto di questo particolare, la soluzione da noi adottata in fase di implementazione è stata di impostare come punto di partenza uno stato che sia effettivamente ammissibile

3. *ALGORITMO CENTRALIZZATO*

per la rete. Per fare ciò è stato sufficiente risolvere preliminarmente all'applicazione dell'IPM un semplice problema di ottimizzazione con funzione obiettivo nulla e con soli vincoli di uguaglianza riassuntivi delle limitazioni fisiche della rete.

Capitolo 4

Algoritmo Distribuito

Nel capitolo precedente è stato analizzato l'Interior Point Method, presentato come uno dei possibili solutori per il problema di minimizzazione delle perdite caratteristico di una qualsiasi smart grid. È facile rendersi conto come tale algoritmo adotti un approccio di tipo centralizzato: ad ogni iterazione, infatti, è un controllore esterno che, conoscendo le informazioni relative a ciascun nodo della rete, cerca di sfruttarle per pervenire ad una soluzione che sia ottima per l'intero sistema.

L'algoritmo che ci accingiamo a presentare in questo capitolo, invece, utilizza un paradigma profondamente diverso e di tipo distribuito. In questo nuovo approccio, infatti, il processo di ottimizzazione è affidato separatamente a ciascun nodo della rete, il quale ragiona in ambito locale cercando di identificare la strategia più efficace per se'. Nel fare ciò, però, esso ha una visione solo parziale della rete e può fare leva esclusivamente sulle informazioni riguardanti se stesso ed i nodi ad esso adiacenti. Per tale ragione la qualità della soluzione ottenuta, che è semplicemente costituita dal collage di tutte le soluzioni locali, non sarà mai superiore a quella ricavata con l'Interior Point Method ¹. Al contrario di quest'ultima, però, la soluzione distribuita può essere calcolata ricorrendo a risorse hardware con capacità elaborative anche molto scarse. L'algoritmo che presenteremo qui di seguito, infatti, opera frazionando il problema di partenza in vari sottoproblemi di dimensione inferiore e quindi più facili da risolvere analiticamente. Il nostro scopo sarà dunque quello di stabilire in che misura vantaggi e svantaggi dei due

¹Ciò è vero a patto che il problema di minimizzazione di partenza sia lo stesso per i 2 solutori e che valga altrettanto per il punto iniziale da cui hanno origine le due elaborazioni.

approcci si compensano e sotto quali circostanze risulta preferibile uno rispetto all'altro. Ma procediamo con ordine e presentiamo l'algoritmo distribuito da noi adottato.

4.1 ADMM scalabile

Il punto di partenza considerato per l'elaborazione distribuita va sotto il nome di Alternating Direction Multiplier Method (ADMM) [12]-[13]. Si tratta ancora una volta di un algoritmo iterativo. Inizialmente ideato per risolvere problemi di minimizzazione convessi con l'utilizzo della parallelizzazione, tale solutore è stato riadattato e reso scalabile per rispondere alle esigenze del nostro specifico contesto.

Consideriamo quindi la solita rete formata da N nodi, $\mathcal{N} = \{1, \dots, N\}$, in cui \mathcal{N}_i indica l'insieme dei nodi adiacenti all' i -esimo. Assumiamo per convenzione che \mathcal{N}_i contenga i stesso ed indichiamo con \mathcal{N}_i/i il medesimo insieme privato però del nodo i . Sia inoltre \mathbf{s}_i l'informazione sensibile (stato) al nodo i . Il nostro problema può essere descritto a partire dal seguente modello generico:

$$\begin{aligned} & \text{minimize } f(\mathbf{s}_{\mathcal{N}}) = \sum_{i \in \mathcal{N}} f_i(\mathbf{s}_{\mathcal{N}_i}) \\ & \text{with respect to } \mathbf{s}_{\mathcal{N}} \\ & \text{subject to } \mathbf{s}_{\mathcal{N}_i} \in \mathcal{C}_i, \forall i \in \mathcal{N} \end{aligned} \tag{4.1}$$

dove le funzioni f_i dipendono esclusivamente dai corrispettivi stati locali $\mathbf{s}_{\mathcal{N}_i}$ e l'insieme \mathcal{C}_i è definito dai vincoli in gioco. Per semplificare la notazione d'ora in avanti indicheremo con $\mathbf{x} = \mathbf{s}_{\mathcal{N}}$ il vettore colonna che raccoglie tutti gli stati. Inoltre $[\mathbf{x}]_i$ corrisponderà al sottoinsieme di \mathbf{x} relativo al nodo i , ossia $[\mathbf{x}]_i = \mathbf{s}_i$. La formulazione classica dell' Alternating Direction Multiplier Method prevede l'introduzione di alcune variabili ausiliarie:

- $\mathbf{x}_i, i \in \mathcal{N}$: copia locale del vettore degli stati;
- $\mathbf{z}_{i,j}, i \in \mathcal{N}, j \in \mathcal{N}_i$: ulteriore copia locale di \mathbf{x} associata ad una coppia di nodi adiacenti.

Sfruttando questo accorgimento è possibile riformulare il modello di partenza come segue:

$$\begin{aligned}
 & \text{minimize } \sum_{i \in \mathcal{N}} f_i(\mathbf{x}_i) \\
 & \text{with respect to } \{\mathbf{x}_i, \mathbf{z}_{i,j}\} \text{ for } i \in \mathcal{N}, j \in \mathcal{N}_i \\
 & \text{subject to } \mathbf{x}_i \in \mathcal{C}_i, \forall i \in \mathcal{N} \\
 & \mathbf{x}_i = \mathbf{z}_{i,j} = \mathbf{z}_{j,i}, \forall i \in \mathcal{N}, j \in \mathcal{N}_i
 \end{aligned} \tag{4.2}$$

Così come è presentato, però, il problema non consente di fornire una formulazione scalabile dell'ADMM, poichè i vincoli riguardanti le variabili ausiliarie sono applicati all'intero insieme degli stati e ciò comporta che l'informazione ad esso relativa sia diffusa e resa disponibile a ciascun nodo della rete. L'idea è dunque quella di forzare lo stato \mathbf{s}_i ad essere identico solo fra nodi adiacenti $j \in \mathcal{N}_i$ che potenzialmente saranno gli unici a sfruttare tale informazione, ossia $\mathbf{x}_i = \mathbf{x}_j = \mathbf{z}_{i,j} = \mathbf{z}_{j,i}$. In quest'ottica introduciamo la matrice di proiezione \mathbf{P}_i ², la quale restituisce $[\mathbf{x}]_i$ se moltiplicata per il vettore degli stati \mathbf{x} ($[\mathbf{x}]_i = \mathbf{P}_i \mathbf{x}$). Essa può essere definita utilizzando il prodotto di Kronecker:

$$\mathbf{P}_i = \text{diag}(\mathbf{e}_i \otimes \mathbf{1}), \tag{4.3}$$

con \mathbf{e}_i vettore colonna di lunghezza N avente un 1 in posizione i e tutti gli altri elementi nulli e $\mathbf{1}$ vettore unario, ossia con tutti gli elementi pari ad 1, di dimensione S , corrispondente alla lunghezza di ciascuno stato. Quindi la matrice di protezione congiunta per gli stati i e j è:

$$\mathbf{P}_{i,j} = \begin{cases} \mathbf{P}_i, & i = j \\ \mathbf{P}_i + \mathbf{P}_j, & i \neq j \end{cases} \tag{4.4}$$

e concorre alla riformulazione del nostro problema nel seguente modo:

$$\begin{aligned}
 & \text{minimize } \sum_{i \in \mathcal{N}} f_i(\mathbf{x}_i) \\
 & \text{with respect to } \{\mathbf{x}_i, \mathbf{z}_{i,j}\} \text{ for } i \in \mathcal{N}, j \in \mathcal{N}_i \\
 & \text{subject to } \mathbf{x}_i \in \mathcal{C}_i, \forall i \in \mathcal{N} \\
 & \mathbf{P}_{i,j} \mathbf{x}_i = \mathbf{P}_{i,j} \mathbf{z}_{i,j} = \mathbf{P}_{i,j} \mathbf{z}_{j,i}, \forall i \in \mathcal{N}, j \in \mathcal{N}_i
 \end{aligned} \tag{4.5}$$

L'algoritmo ADMM può finalmente essere applicato a questo modello e ci fornisce il criterio di aggiornamento delle copie locali del vettore degli stati $\mathbf{x}_i(t)$:

$$\mathbf{x}_i(t+1) = \underset{\mathbf{x}_i \in \mathcal{C}_i}{\text{argmin}} f_i(\mathbf{x}_i) + \mathbf{x}_i^T \mathbf{M}_i \mathbf{x}_i - \mathbf{x}_i^T \mathbf{B}_i(t+1) \tag{4.6}$$

²Si tratta di una matrice binaria diagonale.

dove:

$$\mathbf{B}_i(t+1) = 2\mathbf{M}_i\mathbf{x}_i(t) + \mathbf{U}_i(t) + \Lambda_i(t) \quad (4.7)$$

e:

$$\begin{aligned} \mathbf{M}_i &= \sum_{j \in \mathcal{N}_i} c_{i,j} \mathbf{P}_{i,j} \\ \mathbf{U}_i(t) &= \sum_{j \in \mathcal{N}_i} \eta_{i,j} \mathbf{P}_{i,j} (\mathbf{x}_j(t) - \mathbf{x}_i(t)), \quad \eta_{i,j} = \frac{2c_{i,j}c_{j,i}}{c_{i,j} + c_{j,i}} \\ \Lambda_i(t) &= \begin{cases} \mathbf{0} & , t = 0 \\ \Lambda_i(t-1) + \mathbf{U}_i(t) & t \geq 1 \end{cases} \end{aligned} \quad (4.8)$$

Gli scalari $c_{i,j}$ sono costanti positive arbitrarie che concorrono alla determinazione della velocità di convergenza e t è una variabile temporale. La derivazione di (4.6), (4.7) e (4.8) è presente in Appendice A. Si può notare come, grazie alla presenza delle matrici di proiezione $\mathbf{P}_{i,j}$, l'aggiornamento di un qualsiasi stato venga realizzato sfruttando solo una parte delle informazioni di rete. In particolare, la determinazione di $\mathbf{x}_i(t)$ tiene conto esclusivamente dei dati contenuti nel vettore $\mathbf{s}_{\mathcal{N}_i}$. Presentiamo quindi una nuova, definitiva versione dell'ADMM che pone l'accento su tale caratteristica e rende perciò l'algoritmo completamente scalabile:

$$[\mathbf{x}_i(t)]_{\mathcal{N}_i} = \underset{\mathbf{x}_i \in \mathcal{C}_i}{\operatorname{argmin}} f_i(\mathbf{s}_{\mathcal{N}_i}) + \mathbf{s}_i^T (d_i \mathbf{s}_i - [\mathbf{B}_i(t)]_i) + \sum_{k \in \mathcal{N}_i/i} \mathbf{s}_k^T (c_{i,k} \mathbf{s}_k - [\mathbf{B}_i(t)]_k) \quad (4.9)$$

con:

$$\begin{aligned} d_i &= \sum_{k \in \mathcal{N}_i} c_{i,k} \\ \mathbf{B}_i(t+1) &= 2\mathbf{M}_i\mathbf{x}_i(t) + \mathbf{U}_i(t) + \Lambda_i(t) \\ [\mathbf{M}_i\mathbf{x}_i(t)]_k &= \begin{cases} d_i[\mathbf{x}_i(t)]_i, & k = i \\ c_{i,k}[\mathbf{x}_i(t)]_k, & k \in \mathcal{N}_i/i \end{cases} \\ [\mathbf{U}_i(t)]_k &= \begin{cases} \sum_{j \in \mathcal{N}_i/i} \eta_{i,j}([\mathbf{x}_j(t)]_i - [\mathbf{x}_i(t)]_i), & k = i \\ \eta_{i,k}([\mathbf{x}_k(t)]_k - [\mathbf{x}_i(t)]_k), & k \in \mathcal{N}_i/i \end{cases} \end{aligned} \quad (4.10)$$

Come già accennato in precedenza i parametri $c_{i,j}$ regolano la velocità di convergenza del nostro algoritmo e lo fanno in maniera del tutto analoga a quanto avviene in un qualsiasi scenario di consenso. Proprio per questo motivo, viene qui proposto di impostarli in accordo con quanto accade solitamente in tale contesto, ossia imponendo $c_{i,j} = \epsilon a_{i,j}$, dove $\mathbf{A} = [a_{i,j}]$ è una matrice stocastica per righe coerente con la struttura di rete, mentre ϵ è un fattore di scala. Note le

proprietà della matrice \mathbf{A} , possiamo concludere che $d_i = \epsilon$. Inoltre, aggiungendo l'ipotesi di simmetria alla matrice \mathbf{A} si ottiene $\eta_{i,j} = \epsilon a_{i,j} = c_{i,j}$. Riassumendo, le costanti presenti nell'equazione (4.9) e nel sistema (4.10) avranno per noi la seguente struttura:

- $d_i = \epsilon$
- $c_{i,k} = \epsilon a_{i,k}$
- $\eta_{i,j} = \epsilon a_{i,j} = c_{i,j}$

Esplicitando tali relazioni all'interno di (4.9) e (4.10), queste ultime cambiano così:

$$[\mathbf{x}_i(t)]_{\mathcal{N}_i} = \underset{\mathbf{s}_{\mathcal{N}_i} \in \mathcal{C}_i}{\operatorname{argmin}} f_i(\mathbf{s}_{\mathcal{N}_i}) + \epsilon \left[\mathbf{s}_i^T (\mathbf{s}_i - [\mathbf{B}_i(t)]_i) + \sum_{k \in \mathcal{N}_i/i} a_{i,k} \mathbf{s}_k^T (\mathbf{s}_k - [\mathbf{B}_i(t)]_k) \right] \quad (4.11)$$

con:

$$\begin{aligned} [\mathbf{B}_i(t)]_k &= 2[\mathbf{M}_i \mathbf{x}_i(t-1)] + [\mathbf{U}_i(t-1)]_k + [\mathbf{\Lambda}_i(t-1)]_k \\ [\mathbf{M}_i \mathbf{x}_i(t)]_k &= \begin{cases} [\mathbf{x}_i(t)]_i, & k = i \\ [\mathbf{x}_i(t)]_k, & k \in \mathcal{N}_i/i \end{cases} = [\mathbf{x}(t)]_k \\ [\mathbf{U}_i(t)]_k &= \begin{cases} \sum_{j \in \mathcal{N}_i/i} a_{i,j} ([\mathbf{x}_j(t)]_i - [\mathbf{x}_i(t)]_i), & k = i \\ [\mathbf{x}_k(t)]_k - [\mathbf{x}_i(t)]_k, & k \in \mathcal{N}_i/i \end{cases} \\ [\mathbf{\Lambda}_i(t)]_k &= \begin{cases} 0, & t = 0, k \in \mathcal{N}_i \\ [\mathbf{\Lambda}_i(t-1)]_k + [\mathbf{U}_i(t)]_k, & t \geq 1, k \in \mathcal{N}_i \end{cases} \end{aligned} \quad (4.12)$$

All'inizio del capitolo si è fatto cenno al fatto che l'ADMM scalabile è un algoritmo iterativo e in (4.11) e (4.12) si sono fornite le espressioni che, ad ogni ciclo, consentono l'aggiornamento delle diverse variabili in gioco. Non abbiamo tuttavia ancora stabilito con quali principi e con che tempistiche tali aggiornamenti devono essere realizzati. Qui di seguito si farà piena chiarezza su questi aspetti, analizzando nel dettaglio i vari step che compongono l'algoritmo e che devono essere ripetuti per ciascun nodo della rete in esame. Distinguiamo a tal proposito le azioni che vanno intraprese nel ciclo iniziale rispetto a quelle da adottare nei cicli successivi:

t = 0: Al primo passo ($t = 0$) viene anzitutto inizializzato $[\mathbf{x}_i(0)]_{\mathcal{N}_i}$, ossia il valore delle variabili di stato relative al nodo i e a tutti i suoi adiacenti, per come sono conosciute da i stesso. In seguito, il nodo i invia ai suoi adiacenti ($j \in$

\mathcal{N}_i/i) tali informazioni, ricevendo in cambio da ciascuno di essi le grandezze $[\mathbf{x}_j(0)]_i$ e $[\mathbf{x}_j(0)]_j$. In ultima istanza viene posto $[\mathbf{U}_i(0)]_{\mathcal{N}_i} = [\mathbf{\Lambda}_i(0)]_{\mathcal{N}_i} = 0$.

t > 0: In un qualsiasi istante $t > 0$ viene prima di tutto calcolato $[\mathbf{B}_i(t)]_{\mathcal{N}_i}$, attraverso la prima equazione di (4.12). Conoscendo tale informazione ed ereditando dal precedente ciclo ($t - 1$) i valori $[\mathbf{U}_i(t - 1)]_{\mathcal{N}_i}$ e $[\mathbf{\Lambda}_i(t - 1)]_{\mathcal{N}_i}$ è possibile aggiornare $[\mathbf{x}_i(t)]_{\mathcal{N}_i}$, secondo quanto previsto da (4.11). A questo punto non rimane che aggiornare anche $[\mathbf{U}_i(t)]_{\mathcal{N}_i}$ e $[\mathbf{\Lambda}_i(t)]_{\mathcal{N}_i}$ come indicato nelle ultime due equazioni di (4.11) e ripetere la procedura fino a che non si giunge a convergenza.

I passaggi appena descritti sono stati declinati prendendo come riferimento il singolo nodo i e devono quindi essere ripetuti in maniera del tutto analoga anche per gli altri $N - 1$ nodi della rete che si vuole ottimizzare. Implicitamente ciò significa che il modello di partenza (4.21) può essere risolto analizzando N sottoproblemi di natura più semplice, che in generale avranno la seguente struttura:

$$\begin{aligned} & \text{minimize } f_i(\mathbf{x}_i) + \epsilon \left[\mathbf{x}_i^T (\mathbf{x}_i - [\mathbf{B}_i(t)]_i) + \sum_{k \in \mathcal{N}_i/i} a_{i,k} \mathbf{x}_k^T (\mathbf{x}_k - [\mathbf{B}_i(t)]_k) \right] \\ & \text{subject to } \mathbf{x}_i \in \mathcal{C}_i, \forall i \in \mathcal{N} \end{aligned} \quad (4.13)$$

Si noti come fino a questo momento non sia stato proposto nessun criterio per stabilire se e quando l'algoritmo giunge a convergenza. Ciò è dovuto al fatto che tale criterio assume forme diverse in base al tipo di problema che si sta analizzando e all'obiettivo che ci si è posti. È dunque giunto il momento di fornire una nuova versione dell'Alternating Direction Multiplier Method[14], che ricalca dal punto di vista logico quella appena presentata per il caso generale, ma la cui formulazione è riadattata per tenere conto delle caratteristiche specifiche del problema dell'Optimal Power Flow.

4.2 Applicazione dell'ADMM allo studio dell'Optimal Power Flow

Nello scorso paragrafo abbiamo visto che l'Alternating Direction Multiplier Method permette di trovare la soluzione ad un modello di programmazione non lineare suddividendolo in vari sottoproblemi di natura più semplice. Per capire cosa significa tutto ciò all'interno del contesto dell'Optimal Power Flow è

necessario intanto precisare a cosa corrispondono le variabili di stato \mathbf{x}_i che lo caratterizzano. Si faccia quindi riferimento al modello (2.14); si è più volte fatto notare come in esso l'unica grandezza su cui è possibile giocare per determinare l'ottimo sia rappresentata dalle tensioni di rete. Risulta quindi immediato identificare proprio con queste ultime le variabili di stato, secondo la seguente relazione: $[\mathbf{x}_i(t)]_j = [\mathbf{V}_i(t)]_j = V_{i,j}(t)$ (scrittura che indica la tensione del nodo j così come è conosciuta dal nodo i all'istante t). In questo modo l'espressione analoga a (4.11) risulta essere:

$$\begin{aligned} & [\mathbf{V}_i(t)]_{j \in \mathcal{N}_i} = \{V_{i,j}(t)\}_{j \in \mathcal{N}_i} = \underset{\mathbf{V}_{\mathcal{N}_i}(t) \in \mathcal{C}_i}{\operatorname{argmin}} f_i(\mathbf{V}_{\mathcal{N}_i}(t)) \\ & + \epsilon \left(|V_{i,i}(t)|^2 - \Re [V_{i,i}^*(t) B_{i,i}(t)] \right) + \epsilon \sum_{k \in \mathcal{N}_i/i} a_{i,k} \left(|V_{i,k}(t)|^2 - \Re [V_{i,k}^*(t) B_{i,k}(t)] \right) \end{aligned} \quad (4.14)$$

dove:

$$\begin{aligned} B_{i,k}(t) &= 2V_{i,k}(t-1) + U_{i,k}(t-1) + \Lambda_{i,k}(t-1) \\ U_{i,k}(t) &= \begin{cases} \sum_{j \in \mathcal{N}_i/i} a_{i,j} (V_{j,i}(t) - V_{i,i}(t)), & k = i \\ V_{k,k}(t) - V_{i,k}(t), & k \in \mathcal{N}_i/i \end{cases} \\ \Lambda_{i,k}(t) &= \begin{cases} 0, & t = 0, k \in \mathcal{N}_i \\ \Lambda_{i,k}(t-1) + U_{i,k}(t), & t \geq 1, k \in \mathcal{N}_i \end{cases} \end{aligned} \quad (4.15)$$

In alternativa, si può introdurre la notazione $\tilde{B}_{i,k}(t) = \frac{B_{i,k}(t)}{2}$ in modo da semplificare la scrittura della funzione obiettivo come segue:

$$\begin{aligned} & \{V_{i,j}(t)\}_{j \in \mathcal{N}_i} = \underset{\mathbf{V}_{\mathcal{N}_i}(t) \in \mathcal{C}_i}{\operatorname{argmin}} f_i(\mathbf{V}_{\mathcal{N}_i}(t)) \\ & + \epsilon \left(|V_{i,i}(t) - \tilde{B}_{i,i}(t)|^2 + \sum_{k \in \mathcal{N}_i/i} a_{i,k} |V_{i,k}(t) - \tilde{B}_{i,k}(t)|^2 \right) = \\ & \underset{\mathbf{V}_{\mathcal{N}_i}(t) \in \mathcal{C}_i}{\operatorname{argmin}} f_i(\mathbf{V}_{\mathcal{N}_i}(t)) + \epsilon \sum_{k \in \mathcal{N}_i} a_{i,k} |V_{i,k}(t) - \tilde{B}_{i,k}(t)|^2 \end{aligned} \quad (4.16)$$

deve nell'ultima uguaglianza si è tacitamente imposto $a_{i,i} = 1$. Conseguentemente cambiano anche i criteri di aggiornamento di $\tilde{B}_{i,k}$, $U_{i,k}$ e $\Lambda_{i,k}$, che diventano:

$$\begin{aligned} \tilde{B}_{i,k}(t) &= V_{i,k}(t-1) + U_{i,k}(t-1) + \Lambda_{i,k}(t-1) \\ U_{i,k}(t) &= \begin{cases} \frac{1}{2} \sum_{j \in \mathcal{N}_i/i} a_{i,j} (V_{j,i}(t) - V_{i,i}(t)), & k = i \\ \frac{1}{2} (V_{k,k}(t) - V_{i,k}(t)), & k \in \mathcal{N}_i/i \end{cases} \\ \Lambda_{i,k}(t) &= \begin{cases} 0, & t = 0, k \in \mathcal{N}_i \\ \Lambda_{i,k}(t-1) + U_{i,k}(t), & t \geq 1, k \in \mathcal{N}_i \end{cases} \end{aligned} \quad (4.17)$$

Resta invece valida la condizione iniziale $U_{i,k}(0) = \Lambda_{i,k}(0) = 0, \forall k \in \mathcal{N}_i$.

Affinchè le espressioni (4.16) e (4.17) possano portare ad un'effettiva ottimizzazione delle risorse di rete, è necessario che siano inserite all'interno di un contesto dinamico, in cui l'aggiornamento delle variabili in ambito locale sia comunicato agli altri nodi della rete tramite un opportuno scambio di messaggi. I criteri che regolano tale procedimento sono stati analizzati nella parte finale del paragrafo precedente ed hanno portato alla caratterizzazione dei vari step che compongono un ciclo dell'ADMM. Questi ultimi vengono riproposti all'interno dell'**Algorithm 1** in una versione più compatta e riadattata al contesto dell'Optimal Power Flow.

È possibile osservare come, nella versione scalabile dell'ADMM da noi implementata, lo scambio di messaggi sia effettuato esclusivamente fra nodi adiacenti. Le uniche informazioni che devono essere condivise da tutti i nodi della rete sono infatti quelle relative ai coefficienti ϵ e $a_{i,j}$, che comunque possono essere impostate al ciclo iniziale una volta per tutte. Si noti inoltre che nell'**Algorithm 1** si è ipotizzato di far iterare il procedimento infinite volte, cosa evidentemente impossibile. Nella pratica, ciò che si fa è impostare un criterio di uscita che, una volta soddisfatto, permetta di concludere il processo di ottimizzazione assicurando un livello di accuratezza della soluzione sufficientemente buono. A regime, quest'ultima sarà rappresentata dai termini $V_{i,i}$, per $i = 1, \dots, N$.

Per il prosieguo della nostra trattazione è ora fondamentale capire come risolvere i sottoproblemi riportati nell'equazione (4.18) dell'**Algorithm 1**. A tal fine torna utile la classificazione introdotta nel paragrafo 2.1, che diversifica i nodi di rete in base alle loro caratteristiche elettriche e che ora ci consente di identificare tre solutori per l'equazione (4.18). Prima di fare ciò introduciamo però una nuova notazione che ci consentirà di semplificare i calcoli nel prosieguo della trattazione. Definiamo dunque un vettore colonna \mathbf{e} di lunghezza N_i con un unico elemento non nullo e pari ad 1 in prima posizione³. Introduciamo inoltre i vettori:

$$\mathbf{v} = \begin{bmatrix} V_{i,i} \\ [V_{i,j}]_{j \in \mathcal{N}_i/i} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} Y_{i,i}^* \\ [Y_{i,j}^*]_{j \in \mathcal{N}_i/i} \end{bmatrix} \quad (4.20)$$

³ N_i è pari al numero di adiacenti del generico nodo i , ipotizzando che i stesso lo sia.

Algorithm 1: Elaborazione richiesta dall'ADMM per ciascun nodo i

for $t = 0$ **to** ∞ **do**

if $t = 0$ **then**

 | Inizializzare il vettore locale delle tensioni $\mathbf{V}_i = [V_{i,j}]_{j \in \mathcal{N}_i}$

else

 | Definire

$$\alpha_j = \begin{cases} 1, & j = i \\ a_{i,j}, & j \in \mathcal{N}_i/i \end{cases}$$

$$\tilde{\mathbf{B}}_{i,j} = V_{i,j} + U_{i,j} + \Lambda_{i,j}$$

 | poi aggiornare il vettore \mathbf{V}_i tramite

$$\mathbf{V}_i = \underset{\mathbf{V}_{\mathcal{N}_i \in \mathcal{C}_i}}{\operatorname{argmin}} f_i(\mathbf{V}) + \epsilon \sum_{j \in \mathcal{N}_i} \alpha_j \left| \mathbf{V}_{i,j}(t) - \tilde{\mathbf{B}}_{i,j}(t) \right|^2 \quad (4.18)$$

end

 | Inviare \mathbf{V}_i ai nodi adiacenti e ricevere da essi $V_{j,i}$ e $V_{j,j}$, $j \in \mathcal{N}_i/i$

 | Calcolare i coefficienti

$$U_{i,j} = \begin{cases} \frac{1}{2} \sum_{k \in \mathcal{N}_i/i} a_{i,k} (V_{k,i} - V_{i,i}), & j = i \\ \frac{1}{2} (V_{j,j} - V_{i,j}), & j \in \mathcal{N}_i/i \end{cases} \quad (4.19)$$

if $t = 0$ **then**

 | Azzerare la memoria $\Lambda_{i,j} = 0, j \in \mathcal{N}_i$

else

 | Aggiornare la memoria $\Lambda_{i,j} = \Lambda_{i,j} + U_{i,j}, j \in \mathcal{N}_i$

end

end

Essi restituiscono:

$$\begin{aligned}
 S_i &= V_i I_i^* = \mathbf{v}^* [\mathbf{y} \mathbf{e}^T] \mathbf{v} \\
 P_i &= \Re[S_i] = \mathbf{v}^* \mathbf{M}_P \mathbf{v}, \quad \mathbf{M}_P = \frac{1}{2}(\mathbf{y} \mathbf{e}^T + \mathbf{e} \mathbf{y}^*) \\
 Q_i &= \Im[S_i] = \mathbf{v}^* \mathbf{M}_Q \mathbf{v}, \quad \mathbf{M}_Q = \frac{1}{2j}(\mathbf{y} \mathbf{e}^T - \mathbf{e} \mathbf{y}^*)
 \end{aligned} \tag{4.21}$$

Infine definiamo:

$$\mathbf{M} = \text{diag}(1, [a_{i,j}]_{j \in \mathcal{N}_i/i}), \quad \boldsymbol{\beta} = \begin{bmatrix} \tilde{B}_{i,i} \\ [\tilde{B}_{i,j}]_{j \in \mathcal{N}_i/i} \end{bmatrix} \tag{4.22}$$

4.2.1 Soluzione locale dell'OPF per il nodo PCC

Sia dato un nodo i all'interno della nostra rete e sia esso un Point of Common Coupling (PCC). Ad esso può essere associato un problema di minimizzazione riassunto dall'equazione (4.18), che risolve l'Optimal Power Flow in ambito locale coinvolgendo solo il nodo i ed i suoi adiacenti. Sfruttando la notazione appena introdotta, il problema in questione può essere riscritto in forma compatta come segue:

$$\begin{aligned}
 &\text{minimize} && f_i(\mathbf{v}^* \mathbf{M}_P \mathbf{v}) + \epsilon(\mathbf{v} - \boldsymbol{\beta})^* \mathbf{M}(\mathbf{v} - \boldsymbol{\beta}) \\
 &\text{subject to} && \mathbf{v} \in \mathcal{C}_i
 \end{aligned} \tag{4.23}$$

Nel caso del nodo PCC, la regione ammissibile \mathcal{C}_i è semplicemente definita dal vincolo:

$$\mathcal{C}_i = \{\mathbf{v} | v_1 = V_{i,i} = V_{PCC} = 1\} \tag{4.24}$$

dove l'ultima uguaglianza è verificata purchè si ipotizzi di normalizzare tutte le tensioni di rete rispetto a V_{PCC} (cosa che daremo per scontata nel prosieguo dell'analisi). Cerchiamo ora di semplificare il problema (4.23), trasformandolo in uno equivalente ma di dimensione 2. Per procedere in tal senso applichiamo prima di tutto la sostituzione:

$$\mathbf{x} = \mathbf{M}^{\frac{1}{2}} \mathbf{v}$$

ottenendo il seguente modello:

$$\begin{aligned}
 &\text{minimize} && f_i(\mathbf{x}^* [\frac{1}{2}(\mathbf{s} \mathbf{e}^T + \mathbf{e} \mathbf{s}^*)] \mathbf{x}) + \epsilon \|\mathbf{x} - \mathbf{p}\|^2 \\
 &\text{subject to} && x_1 = 1
 \end{aligned} \tag{4.25}$$

dove $\mathbf{s} = \mathbf{M}^{-\frac{1}{2}} \mathbf{y}$ e $\mathbf{p} = \mathbf{M}^{\frac{1}{2}} \boldsymbol{\beta}$. Si noti come la matrice \mathbf{M} sia per costruzione definita positiva e consenta quindi di applicare i precedenti passaggi senza dover fare altre ipotesi matematiche.

Il problema può essere ulteriormente compattato prestando attenzione al fatto che il termine legato alla potenza attiva appartiene al sottospazio individuato da \mathbf{s} ed \mathbf{e} . Applichiamo dunque la seguente trasformazione:

$$\mathbf{Q}^* \mathbf{x}$$

con \mathbf{Q} matrice ortonormale, le cui prime due colonne sono individuate da:

$$\mathbf{Q}\mathbf{e} = \mathbf{e}, \quad \mathbf{Q}\mathbf{e}_2 = \frac{\mathbf{s} - s_1\mathbf{e}}{\sqrt{K_i}} \quad (4.26)$$

e mappano quindi il sottospazio sopra citato. In (4.26) lo scalare K_i è dato da:

$$K_i = \|\mathbf{s}\|^2 - |s_1|^2 = \sum_{j \in \mathcal{N}_i/i} \frac{|Y_{i,j}|^2}{a_{i,j}} > 0. \quad (4.27)$$

Adottando tali accorgimenti il problema (4.25) assume la seguente forma equivalente:

$$\begin{aligned} &\text{minimize} && f_i(\mathbf{q}_{12}^* \mathbf{C} \mathbf{q}_{12}) + \epsilon \|\mathbf{q} - \mathbf{b}\|^2 \\ &\text{subject to} && q_1 = 1 \end{aligned} \quad (4.28)$$

dove $\mathbf{q}_{12} = [q_1, q_2]^T$, $\mathbf{b} = \mathbf{Q}^* \mathbf{M}^{\frac{1}{2}} \boldsymbol{\beta}$ e la matrice \mathbf{C} corrisponde a:

$$\mathbf{C} = \begin{bmatrix} \Re[Y_{i,i}^*] & \frac{1}{2}\sqrt{K_i} \\ \frac{1}{2}\sqrt{K_i} & 0 \end{bmatrix} \quad (4.29)$$

Si noti come la formulazione fornita in (4.28) forzi la soluzione $\mathbf{q} = \mathbf{b}$ per tutti gli elementi di \mathbf{q} stesso, ad eccezione dei primi due, consentendoci così di scrivere:

$$\mathbf{q} = \mathbf{b} + (q_1 - b_1)\mathbf{e} + (q_2 - b_2)\mathbf{e}_2 \quad (4.30)$$

Ora, adottando la trasformazione inversa $\mathbf{v} = \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} \mathbf{q}$, si ricava:

$$\mathbf{v} = \boldsymbol{\beta} + \mathbf{e}(q_1 - b_1) + \frac{\mathbf{M}^{-1} \mathbf{y} - y_1 \mathbf{e}}{\sqrt{K_i}} (q_2 - b_2) \quad (4.31)$$

o, affrancandoci dalla notazione matriciale:

$$\begin{aligned} V_{i,i} &= q_1 \\ V_{i,j} &= \beta_{i,j} + \frac{q_2 - H_i Y_{i,j}^*}{\sqrt{K_i} a_{i,j}}, \quad j \in \mathcal{N}_i/i \end{aligned} \quad (4.32)$$

con:

$$H_i = \frac{1}{\sqrt{K_i}} \sum_{j \in \mathcal{N}_i/i} Y_{i,j} \beta_{i,j}. \quad (4.33)$$

4. ALGORITMO DISTRIBUITO

Gli unici termini non noti di (4.32) sono rappresentati da q_1 e q_2 , che devono perciò essere determinati risolvendo il seguente problema di ottimizzazione ad hoc:

$$\begin{aligned} & \text{minimize} && f_i(\mathbf{q}_{12}^* \mathbf{C} \mathbf{q}_{12}) + \epsilon \|\mathbf{q}_{12} - \mathbf{b}_{12}\|^2 \\ & \text{subject to} && q_1 = 1 \end{aligned} \quad (4.34)$$

dove $\mathbf{b}_{12} = [b_1, b_2]^T$. Come anticipato all'inizio di questa sezione, la soluzione del problema (4.23) passa quindi attraverso lo studio del modello (4.34), il quale ha una dimensione inferiore, pari esattamente a 2.

Vista la semplicità del vincolo in (4.34), è possibile renderlo esplicito direttamente all'interno della funzione obiettivo. Ciò ci permette di rielaborare il modello (4.28), trasformandolo nel seguente problema di ottimizzazione convesso di dimensione 1:

$$\text{minimize} \quad \left| q_2 - H_i + \frac{\sqrt{K_i}}{2\epsilon} \right|^2 \quad (4.35)$$

nonchè di stabilire che:

$$V_{i,i} = v_1 = q_1 = 1. \quad (4.36)$$

Il punto di minimo di (4.35) è facilissimo da calcolare e coincide chiaramente con il valore di q_2 in corrispondenza al quale la funzione obiettivo si annulla, ovvero:

$$q_2 = H_i - \frac{\sqrt{K_i}}{2\epsilon}. \quad (4.37)$$

Quest'ultimo risultato consente di concludere la trattazione per il Power Common Coupling. Riassumendo, possiamo quindi affermare che, per soddisfare le richieste relative all'Optimal Power Flow in ambito locale, il nodo PCC deve soddisfare i seguenti requisiti:

$$\begin{aligned} V_{i,i} &= 1 \\ V_{i,j} &= \beta_{i,j} - \frac{Y_{i,j}^*}{2\epsilon a_{i,j}}, \quad j \in \mathcal{N}_i/i. \end{aligned} \quad (4.38)$$

4.2.2 Soluzione locale dell'OPF per i nodi PEIs

Come fatto in precedenza, si consideri il modello (4.23), dove però stavolta al nodo i vengono associate le caratteristiche di una Power Electric Interface (PEI). In questo contesto, l'insieme di ammissibilità è costituito da:

$$\mathcal{C}_i = \{ \mathbf{v} \mid \underline{P}_{PEI,i} \leq \mathbf{v}^* \mathbf{M}_P \mathbf{v} \leq \overline{P}_{PEI,i} \} \quad (4.39)$$

mentre la funzione f_i è nulla⁴. Tali precisazioni permettono di riadattare il modello (4.23) al caso in esame, ottenendo la seguente formulazione:

$$\begin{aligned} & \text{minimize} && (\mathbf{v} - \boldsymbol{\beta})^* \mathbf{M} (\mathbf{v} - \boldsymbol{\beta}) \\ & \text{subject to} && \underline{P}_{PEI,i} \leq \mathbf{v}^* \mathbf{M}_P \mathbf{v} \leq \overline{P}_{PEI,i} \end{aligned} \quad (4.40)$$

Purtroppo, la presenza di vincoli di disuguaglianza rende impossibile determinare una soluzione in forma chiusa per (4.40) e l'unico modo per uscire dall'empasse è quello ridefinire l'insieme di ammissibilità (4.39) nel seguente modo:

$$\mathcal{C}_i = \{ \mathbf{v} \mid \mathbf{v}^* \mathbf{M}_P \mathbf{v} = \overline{P}_{PEI,i} \} \quad (4.41)$$

La scelta di uguagliare la potenza attiva dei nodi PEIs a quella massima erogabile non è casuale. Si ricordi infatti che lo scopo finale del nostro lavoro è quello di assicurare il funzionamento della rete richiedendo meno energia possibile al nodo PCC. In prossimità dell'ottimo, quindi, la potenza attiva associata a quest'ultimo tenderà a diminuire sempre più e la conseguente riduzione di energia immessa in rete dovrà essere sopperita dal lavoro dei PEIs, ai quali verrà chiesto di uscire a piena potenza. Per queste ragioni, possiamo affermare che l'espedito utilizzato in (4.41) non mina in alcun modo la validità della soluzione che andremo a determinare ed il modello (4.40) può quindi essere riscritto come segue:

$$\begin{aligned} & \text{minimize} && (\mathbf{v} - \boldsymbol{\beta})^* \mathbf{M} (\mathbf{v} - \boldsymbol{\beta}) \\ & \text{subject to} && \mathbf{v}^* \mathbf{M}_P \mathbf{v} = \overline{P}_{PEI,i} \end{aligned} \quad (4.42)$$

Sfruttando la trasformazione $\mathbf{x} = \mathbf{M}^{\frac{1}{2}} \mathbf{v}$, si ottiene il problema equivalente:

$$\begin{aligned} & \text{minimize} && \|\mathbf{x} - \mathbf{p}\|^2 \\ & \text{subject to} && \mathbf{x}^* \left[\frac{1}{2} (\mathbf{s} \mathbf{e}^T + \mathbf{e} \mathbf{s}^*) \right] \mathbf{x} = \overline{P}_{PEI,i} \end{aligned} \quad (4.43)$$

con $\mathbf{s} = \mathbf{M}^{-\frac{1}{2}} \mathbf{y}$ e $\mathbf{p} = \mathbf{M}^{\frac{1}{2}} \boldsymbol{\beta}$. A questo punto è possibile adottare ancora una volta la sostituzione $\mathbf{q} = \mathbf{Q}^* \mathbf{x}$ con \mathbf{Q} matrice ortonormale che soddisfa le relazioni (4.26), in modo da ottenere:

$$\begin{aligned} & \text{minimize} && \|\mathbf{q} - \mathbf{b}\|^2 \\ & \text{subject to} && \mathbf{q}_{12}^* \mathbf{C} \mathbf{q}_{12} = \overline{P}_{PEI,i} \end{aligned} \quad (4.44)$$

⁴Si ricorda infatti che, nella sua formulazione (2.13), il problema dell'Optimal Power Flow prevede di minimizzare la potenza attiva relativa al solo nodo PCC.

4. ALGORITMO DISTRIBUITO

dove si ricorda che $\mathbf{q}_{12} = [q_1, q_2]^T$, $\mathbf{b} = \mathbf{Q}^* \mathbf{M}^{\frac{1}{2}} \boldsymbol{\beta}$ e:

$$\mathbf{C} = \begin{bmatrix} \Re [Y_{i,i}^*] & \frac{1}{2} \sqrt{K_i} \\ \frac{1}{2} \sqrt{K_i} & 0 \end{bmatrix}.$$

Come nel caso del PCC, da (4.44) possiamo concludere che:

$$\mathbf{q} = \mathbf{b} + (q_1 - b_1) \mathbf{e} + (q_2 - b_2) \mathbf{e}_2 \quad (4.45)$$

da cui, adottando la trasformazione inversa $\mathbf{v} = \mathbf{M}^{-\frac{1}{2}} \mathbf{Q} \mathbf{q}$, si ricava:

$$\mathbf{v} = \boldsymbol{\beta} + \mathbf{e}(q_1 - b_1) + \frac{\mathbf{M}^{-1} \mathbf{y} - y_1 \mathbf{e}}{\sqrt{K_i}} (q_2 - b_2) \quad (4.46)$$

che equivale a dire:

$$\begin{aligned} V_{i,i} &= q_1 \\ V_{i,j} &= \beta_{i,j} + \frac{q_2 - H_i Y_{i,j}^*}{\sqrt{K_i}} \frac{1}{a_{i,j}}, \quad j \in \mathcal{N}_i/i. \end{aligned} \quad (4.47)$$

Non rimane che determinare i valori di q_1 e q_2 andando a studiare il modello di ottimizzazione:

$$\begin{aligned} &\text{minimize} \quad \|\mathbf{q}_{12} - \mathbf{b}_{12}\|^2 \\ &\text{subject to} \quad \mathbf{q}_{12}^* \mathbf{C} \mathbf{q}_{12} = \bar{P}_{PEI,i} \end{aligned} \quad (4.48)$$

Il problema può essere affrontato attraverso la sua proiezione nello spazio individuato dagli autovettori della matrice \mathbf{C} . Si riscriva allora quest'ultima nella forma:

$$\mathbf{C} = \mathbf{T} \begin{bmatrix} \sigma_+ & 0 \\ 0 & \sigma_- \end{bmatrix} \mathbf{T}^T \quad (4.49)$$

in cui sono messi in evidenza i suoi autovalori $\sigma_+ > 0$ e $\sigma_- < 0$, dati da:

$$\sigma_{\pm} = \frac{\Re [Y_{i,i}] \pm \sqrt{\Re [Y_{i,i}]^2 + K_i}}{2} \quad (4.50)$$

e la matrice di trasformazione \mathbf{T} è definita come:

$$\mathbf{T} = \begin{bmatrix} \frac{2\sigma_+}{\sqrt{4\sigma_+^2 + K_i}} & \frac{2\sigma_-}{\sqrt{4\sigma_-^2 + K_i}} \\ \frac{\sqrt{K_i}}{\sqrt{4\sigma_+^2 + K_i}} & \frac{\sqrt{K_i}}{\sqrt{4\sigma_-^2 + K_i}} \end{bmatrix} = \frac{1}{\sqrt{4\sigma_+^2 + K_i}} \begin{bmatrix} 2\sigma_+ & -\sqrt{K_i} \\ \sqrt{K_i} & 2\sigma_+ \end{bmatrix}. \quad (4.51)$$

A questo punto siamo finalmente in grado di rielaborare il modello (4.48) mappandolo nella variabile $\mathbf{z} = \mathbf{T}^T \mathbf{q}_{12}$ come qui di seguito suggerito:

$$\begin{aligned} &\text{minimize} \quad \|\mathbf{z} - \mathbf{g}\|^2 \\ &\text{subject to} \quad |z_1|^2 \sigma_+ + |z_2|^2 \sigma_- = \bar{P}_{PEI,i} \end{aligned} \quad (4.52)$$

dove $\mathbf{g} = \mathbf{T}^T \mathbf{b}_{12}$. Esprimendo tutte le grandezze presenti in (4.52) in notazione polare, attraverso l'utilizzo di modulo e fase, si ha:

$$\begin{aligned} & \text{minimize} && |z_1|^2 - 2|z_1||g_1| \cos(\varphi_{z_1} - \varphi_{g_1}) + |z_2|^2 - 2|z_2||g_2| \cos(\varphi_{z_2} - \varphi_{g_2}) \\ & \text{subject to} && |z_1| = \sqrt{\frac{\bar{P}_{PEI,i} - \sigma_- |z_2|^2}{\sigma_+}} \end{aligned} \quad (4.53)$$

da cui banalmente si ha $\varphi_{z_1} = \varphi_{g_1}$ e $\varphi_{z_2} = \varphi_{g_2}$. Perciò:

$$\begin{aligned} z_1 &= e^{j\varphi_{g_1}} \sqrt{\frac{\bar{P}_{PEI,i} - \sigma_- |z_2|^2}{\sigma_+}} \\ z_2 &= e^{j\varphi_{g_2}} |z_2| \end{aligned} \quad (4.54)$$

Affinchè le relazioni (4.54) siano verificate, è chiaramente necessario che $|z_2| \geq 0$ e $|z_2|^2 \geq \sqrt{\frac{\bar{P}_{PEI,i}}{\sigma_-}}$. Una volta sostituito il vincolo $|z_1| = \sqrt{\frac{\bar{P}_{PEI,i} - \sigma_- |z_2|^2}{\sigma_+}}$ all'interno della funzione obiettivo del problema (4.54), possiamo quindi affermare che il valore assoluto $|z_2|$ è identificato dalla soluzione di:

$$|z_2| = \underset{x \geq 0, x^2 \geq -k_3}{\operatorname{argmin}} f(x), \quad f(x) = x^2 - 2k_1x - k_2\sqrt{x^2 + k_3} \quad (4.55)$$

dove:

$$k_1 = \frac{|g_2|}{1 + \frac{|\sigma_-|}{\sigma_+}}, \quad k_2 = \frac{2|g_1|}{1 + \frac{|\sigma_-|}{\sigma_+}} \sqrt{\frac{|\sigma_-|}{\sigma_+}}, \quad k_3 = \frac{\bar{P}_{PEI,i}}{|\sigma_-|}. \quad (4.56)$$

Il punto di minimo in (4.55) può essere calcolato semplicemente andando a vedere dove la derivata prima di $f(x)$ si annulla, ossia imponendo:

$$f'(x) = 2(x - k_1) - \frac{k_2x}{\sqrt{x^2 + k_3}} = 0. \quad (4.57)$$

Con alcuni semplici passaggi algebrici, dalla formula (4.57) è possibile approdare alla seguente equazione di quarto grado:

$$x^4 - 2k_1x^3 + \left(k_1^2 - \frac{1}{4}k_2^2 + k_3\right)x^2 - 2k_1k_3x + k_1^2k_3 = 0 \quad (4.58)$$

la quale è risolvibile in forma chiusa. Per quanto detto in precedenza, le radici di (4.58) che non soddisfano le condizioni:

$$x \geq 0, \quad x^2 \geq -k_3$$

non sono ammissibili per il nostro problema di partenza e devono quindi essere scartate. Se anche in seguito a questa scrematura si dovesse ottenere più di

una soluzione ammissibile, è necessario scegliere quella che minimizza la funzione $f(x)$. Una volta determinato tale valore, le espressioni (4.54) forniscono la chiave per calcolare entrambi gli elementi del vettore $\mathbf{z} = [z_1, z_2]^T$. Infine, applicando la trasformazione inversa $\mathbf{q}_{12} = [\mathbf{U}^T]^{-1} \mathbf{z}$, si riescono ad estrapolare i coefficienti q_1 e q_2 che risolvono il problema (4.48) e che vanno quindi sostituiti in (4.47). In questo modo si conclude anche l'analisi relativa alle Power Electric Interfaces.

4.2.3 Soluzione locale dell'OPF per i Loads

Si prenda sempre come riferimento il modello (4.23), ma stavolta si supponga che i abbia le caratteristiche di un Load. In un frangente simile, l'insieme di ammissibilità \mathcal{C}_i è dato da:

$$\mathcal{C}_i = \{\mathbf{v} | \mathbf{v}^* \mathbf{M}_P \mathbf{v} = P_{Load,i}; \mathbf{v}^* \mathbf{M}_Q \mathbf{v} = Q_{Load,i}\} \quad (4.59)$$

mentre la funzione f_i è nulla. Il problema (4.23) di partenza può quindi essere rielaborato nel seguente modo:

$$\begin{aligned} & \text{minimize} && (\mathbf{v} - \boldsymbol{\beta})^* \mathbf{M} (\mathbf{v} - \boldsymbol{\beta}) \\ & \text{subject to} && \mathbf{v}^* \mathbf{M}_P \mathbf{v} = P_{Load,i} \\ & && \mathbf{v}^* \mathbf{M}_Q \mathbf{v} = Q_{Load,i} \end{aligned} \quad (4.60)$$

Applicando la sostituzione $\mathbf{x} = \mathbf{M}^{\frac{1}{2}} \mathbf{v}$, si ottiene poi:

$$\begin{aligned} & \text{minimize} && \|\mathbf{x} - \mathbf{p}\|^2 \\ & \text{subject to} && \mathbf{x}^* \left[\frac{1}{2} (\mathbf{s} \mathbf{e}^T + \mathbf{e} \mathbf{s}^*) \right] \mathbf{x} = P_{Load,i} \\ & && \mathbf{x}^* \left[\frac{1}{2j} (\mathbf{s} \mathbf{e}^T - \mathbf{e} \mathbf{s}^*) \right] \mathbf{x} = Q_{Load,i} \end{aligned} \quad (4.61)$$

con $\mathbf{s} = \mathbf{M}^{-\frac{1}{2}} \mathbf{y}$ e $\mathbf{p} = \mathbf{M}^{\frac{1}{2}} \boldsymbol{\beta}$. Sfruttando la solita trasformazione $\mathbf{q} = \mathbf{Q}^* \mathbf{x}$, già commentata nel §4.2.1, si giunge a formulare il modello equivalente:

$$\begin{aligned} & \text{minimize} && \|\mathbf{q} - \mathbf{b}\|^2 \\ & \text{subject to} && \mathbf{q}_{12}^* \mathbf{C} \mathbf{q}_{12} = P_{Load,i} \\ & && \mathbf{q}_{12}^* \mathbf{D} \mathbf{q}_{12} = Q_{Load,i} \end{aligned} \quad (4.62)$$

dove l'unico elemento di novità rispetto a (4.44) è rappresentato dalla matrice:

$$\mathbf{D} = \begin{bmatrix} \Im [Y_{i,i}^*] & -\frac{1}{2j} \sqrt{K_i} \\ \frac{1}{2j} \sqrt{K_i} & 0 \end{bmatrix}. \quad (4.63)$$

Ancora una volta la formulazione (4.62) assicura che:

$$\begin{aligned} V_{i,i} &= q_1 \\ V_{i,j} &= \beta_{i,j} + \frac{q_2 - H_i Y_{i,j}^*}{\sqrt{K_i} a_{i,j}}, \quad j \in \mathcal{N}_i/i. \end{aligned} \quad (4.64)$$

dove i termini q_1 e q_2 sono da determinare come soluzione del seguente modello di ottimizzazione:

$$\begin{aligned} \text{minimize} \quad & \|\mathbf{q}_{12} - \mathbf{b}_{12}\|^2 \\ \text{subject to} \quad & |q_1|^2 Y_{i,i}^* + \sqrt{K_i} q_1 q_2 = P_{Load,i} + jQ_{Load,i} \end{aligned} \quad (4.65)$$

in cui il vincolo complesso riassume sia la richiesta sulla potenza attiva che su quella reattiva del nodo i . Si prendano ora in considerazione due casi distinti:

$P_{Load,i} + jQ_{Load,i} = \mathbf{0}$: Il nodo i è detto inattivo e le possibili soluzioni di (4.65) sono due:

$$\begin{cases} q_1 = 0 \\ q_2 = 0 \end{cases} \quad \begin{cases} q_1 = \frac{\sqrt{K_i} \beta_{i,i} - Y_{i,i}^* H_i}{K_i + |Y_{i,i}|^2} \sqrt{K_i} \\ q_2 = -\frac{Y_{i,i}}{\sqrt{K_i}} q_1 \end{cases} \quad (4.66)$$

La coppia che a noi interessa è ovviamente quella che minimizza la norma quadratica $\|\mathbf{q}_{12} - \mathbf{b}_{12}\|^2$.

$P_{Load,i} + jQ_{Load,i} \neq \mathbf{0}$: Il nodo i risulta attivo ed il vincolo di uguaglianza in (4.65) impone:

$$q_2 = \left(\frac{P_{Load,i} - jQ_{Load,i}}{\sqrt{K_i} |q_1|^2} - \frac{Y_{i,i}}{\sqrt{K_i}} \right) q_1, \quad q_1 \neq 0. \quad (4.67)$$

Rendendo esplicita tale relazione all'interno della funzione obiettivo di (4.65), si ricava:

$$\begin{aligned} \text{minimize} \quad & |q_1|^2 \frac{K_i + |Y_{i,i}|^2}{K_i} + \frac{1}{|q_1|^2} \frac{|P_{Load,i} + jQ_{Load,i}|^2}{K_i} \\ & - 2\Re \left[q_1^* \left(\frac{\sqrt{K_i} \beta_{i,i} - Y_{i,i}^* H_i}{\sqrt{K_i}} + \frac{1}{|q_1|^2} \frac{(P_{Load,i} + jQ_{Load,i}) H_i}{\sqrt{K_i}} \right) \right] \end{aligned} \quad (4.68)$$

in cui $q_1 \in \mathbb{C}$ e $q_1 \neq 0$. Passando alla rappresentazione polare di q_1 , la (4.68) ci consente di stabilirne immediatamente la fase, che coincide con:

$$\varphi_{q_1} = \arg(q_1) = \arg \left(\frac{\sqrt{K_i} \beta_{i,i} - Y_{i,i}^* H_i}{\sqrt{K_i}} + \frac{1}{|q_1|^2} \frac{(P_{Load,i} + jQ_{Load,i}) H_i}{\sqrt{K_i}} \right). \quad (4.69)$$

4. ALGORITMO DISTRIBUITO

Sfruttando tale informazione e ricorrendo a qualche semplice artificio matematico, possiamo affermare che il valore assoluto $|q_1|$ è a sua volta identificato dalla soluzione di:

$$|q_1| = \operatorname{argmin}_{x>0} f(x), \quad f(x) = x^2 + \frac{k_1}{x^2} - \sqrt{k_2x^2 + \frac{k_3}{x^2} + k_4} \quad (4.70)$$

dove:

$$\begin{aligned} k_1 &= \frac{|P_{Load,i} + jQ_{Load,i}|^2}{K_i + |Y_{i,i}|^2} \\ k_2 &= 4K_i \frac{|\sqrt{K_i}\beta_{i,i} - Y_{i,i}^*H_i|^2}{(K_i + |Y_{i,i}|^2)^2} \\ k_3 &= 4K_i \frac{|P_{Load,i} + jQ_{Load,i}|^2|H_i|^2}{(K_i + |Y_{i,i}|^2)^2} \\ k_4 &= 8K_i \frac{\Re [(\sqrt{K_i}\beta_{i,i} - Y_{i,i}^*H_i)^*(P_{Load,i} + jQ_{Load,i})H_i]}{(K_i + |Y_{i,i}|^2)^2}. \end{aligned} \quad (4.71)$$

Di nuovo, il problema (4.70) può essere risolto imponendo:

$$f'(x) = 2x - \frac{2k_1}{x^3} - \frac{k_2x - \frac{k_3}{x^3}}{\sqrt{k_2x^2 + \frac{k_3}{x^2} + k_4}} = 0 \quad (4.72)$$

da cui si ottiene l'equazione di sesto ordine:

$$\begin{aligned} k_2y^6 + \left(k_4 - \frac{1}{4}k_2^2\right)y^5 + (k_3 - 2k_1k_2)y^4 + \left(\frac{1}{2}k_2k_3 - 2k_1k_4\right)y^3 \\ + (k_1^2k_2 - 2k_1k_3)y^2 + \left(k_1^2k_4 - \frac{1}{4}k_3^2\right)y + k_1^2k_3 = 0 \end{aligned} \quad (4.73)$$

dove $y = x^2$. Dalle soluzioni ricavate in seguito all'inversione $x = \sqrt{y}$, devono essere eliminate quelle con componenti immaginarie non nulle (visto che $x = |q_1|$) e quelle che soddisfano $x \leq 0$ (visto che è esplicitamente richiesto in (4.70)). Fatto ciò, dalle radici rimanenti si estrae quella che minimizza la funzione $f(x)$, la quale ci permette finalmente di ricavare $|q_1|$. Da qui determinare q_1 e q_2 è un gioco da ragazzi.

In conclusione si può osservare come l'analisi appena condotta sia stata utile per capire come valorizzare gli elementi del vettore $\mathbf{q}_{12} = [q_1, q_2]^T$ sia quando si ha a che fare con un carico attivo che con uno inattivo. Partendo da questi dati, è possibile inoltre calcolare le tensioni $V_{i,i}$ e $V_{i,j}, j \in \mathcal{N}_i/i$ associate al carico i semplicemente sfruttando (4.64). Ciò conclude lo studio dell'Optimal Power Flow in ambito locale quando si ha a che fare con un Load.

Ora che abbiamo in mano tutti gli strumenti matematici che ci consentono di risolvere i sottoproblemi (4.18), l'unico argomento rimasto in sospeso è quello relativo al criterio di uscita da adottare per interrompere l'elaborazione dell'**Algorithm 1**. Quest'ultimo può essere declinato in varie forme frutto di altrettanti ragionamenti. In fase simulativa noi abbiamo scelto di interrompere l'elaborazione al ciclo n -esimo se, a tale iterazione, si fosse riscontrata una variazione non significativa della potenza attiva del PCC rispetto a quella ottenuta al ciclo $n - 1$.

Capitolo 5

Risultati delle simulazioni

Nei capitoli precedenti si sono fornite le basi teoriche per la comprensione di due algoritmi: l'Interior Point Method e l'Alternating Direction Multiplier Method. Fra le pieghe dell'analisi condotta, sono state inoltre evidenziate le caratteristiche che tali solutori devono soddisfare per poter essere applicati al contesto dell'Optimal Power Flow. Ciò che ancora non è stato affrontato è invece uno studio prestazionale delle due tecniche, del quale ci accingiamo a parlare qui di seguito.

Per la valutazione delle performance, l'ambiente computazionale da noi scelto è stato Matlab. Esso mette a disposizione alcuni script prodotti dall'IEEE, che formalizzano matematicamente le principali caratteristiche di una qualsiasi rete elettrica e quindi anche di una smart-grid. Ci siamo focalizzati in particolare su tre file (*'ieee14cdf.txt'*, *'ieee30cdf.txt'* e *'ieee57cdf.txt'*), che riassumono al loro interno la struttura di altrettante reti: la prima da 14, la seconda da 30 e l'ultima da 57 nodi. Oltre ai semplici aspetti topologici, al loro interno è presente anche una classificazione dei nodi in base alle loro proprietà elettriche, da cui è possibile stabilire chi svolge il ruolo di Power Common Coupling, chi quello di Power Electric Interface e chi quello di Load. Infine, pure le informazioni riguardanti i vincoli di rete sono ivi estraibili e permettono di dare piena forma al modello di ottimizzazione non lineare, così come è stato introdotto in (2.14). Per risolvere quest'ultimo, sono state predisposte due funzioni distinte che implementano l'IPM e l'ADMM seguendo fedelmente le indicazioni riportate nei capitoli precedenti. I grafici risultanti dal loro impiego saranno presentati a breve e permetteranno di mettere a confronto le prestazioni dei due solutori. Ma iniziamo analizzando le performance dell'algoritmo centralizzato.

5.1 Performance dell'IPM

Nel campo della ricerca operativa, il primo e più importante elemento da considerare per determinare la bontà di un qualsiasi solutore è rappresentato dall'efficacia con cui esso riesce ad ottimizzare la funzione di costo. Nel nostro particolare contesto quest'ultima è data dalla minimizzazione della potenza attiva associata al nodo PCC. Risulta quindi interessante fornire alcune informazioni in merito, valutando quali effetti produce l'impiego dell'IPM su tale grandezza. A tal proposito si faccia riferimento ai tre grafici riportati in Figura 5.1, Figura 5.2 e Figura 5.3. Essi sono relativi alle tre diverse reti IEEE precedentemente citate, ma riportano tutti il valore della potenza attiva del PCC normalizzata¹ all'uscita da ciascun ciclo dell'Interior Point Method. Si può apprezzare come le tre immagini presentino un andamento molto simile: una prima fase oscillatoria in cui l'algoritmo cerca di determinare il punto di ottimo della smart-grid di riferimento ed una seconda più lineare in cui viene perfezionato l'assestamento attorno a tale valore.

Si presti attenzione al fatto che, durante la ricerca dell'ottimo, la rete attraversa degli stati in cui la potenza attiva del PCC risulta inferiore rispetto a quella assunta a convergenza e ciò potrebbe far pensare ad un errore nella determinazione del minimo. La spiegazione di tale comportamento è però un'altra e risiede nel fatto che, in corrispondenza di tali stati, i vincoli di rete imposti dal modello (2.14) non sono rispettati. Nulla vieta infatti all'Interior Point Method di attraversare punti al di fuori dell'insieme di ammissibilità di (2.14) durante il processo di ottimizzazione, ma questi ultimi non possono chiaramente essere scelti come soluzione del nostro problema e devono essere abbandonati immediatamente. È proprio questo ciò che accade nella fase transitoria di Figura 5.1, Figura 5.2 e Figura 5.3.

Un dato importante che è possibile estrarre dai grafici sopra riportati è quello relativo alla potenza di funzionamento del nodo PCC a regime. Ragionando sempre in termini di grandezze normalizzate, essa risulta pari a 2.0679 per la rete da 14 nodi, 2.3226 per quella da 30 nodi e 3.9094 per quella da 57 nodi. È interessante stabilire in che misura tali valori contribuiscano a soddisfare la richiesta di energia complessiva (P_{tot}) all'interno delle relative smart-grid.

¹La normalizzazione è fatta rispetto alla tensione di riferimento del PCC, che qui in Italia è di 230V; quindi per risalire alla reale potenza di ciascun nodo è necessario moltiplicare per 230².

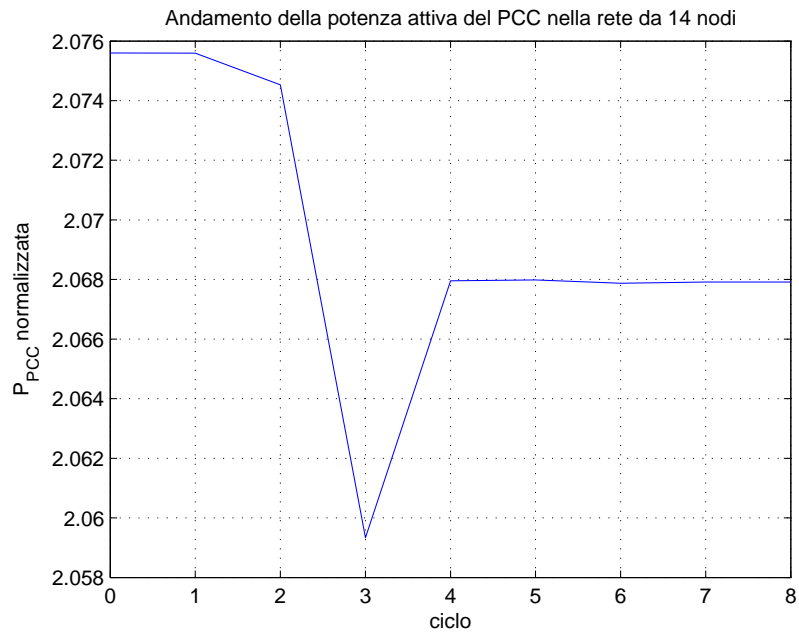


Figura 5.1: Processo di assestamento della potenza attiva del PCC nella rete da 14 nodi in seguito all'applicazione dell'Interior Point Method.

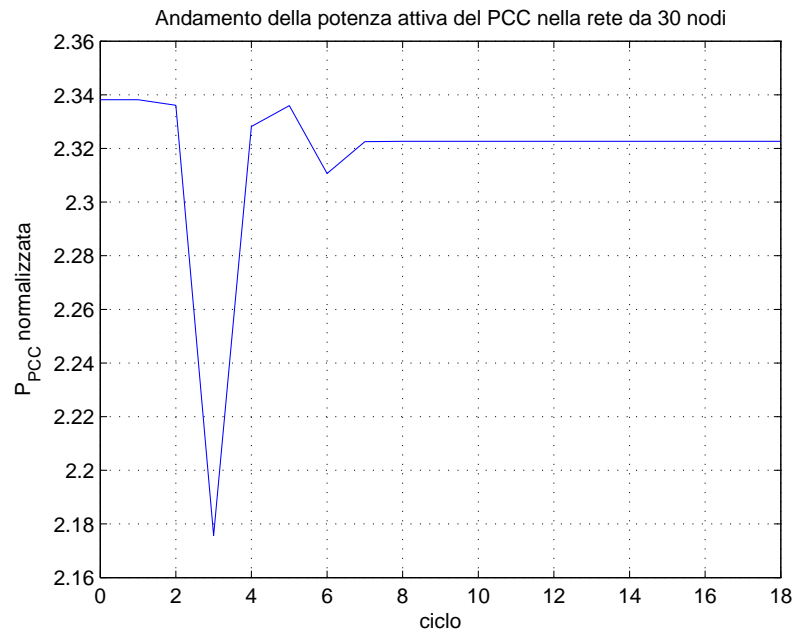


Figura 5.2: Processo di assestamento della potenza attiva del PCC nella rete da 30 nodi in seguito all'applicazione dell'Interior Point Method.

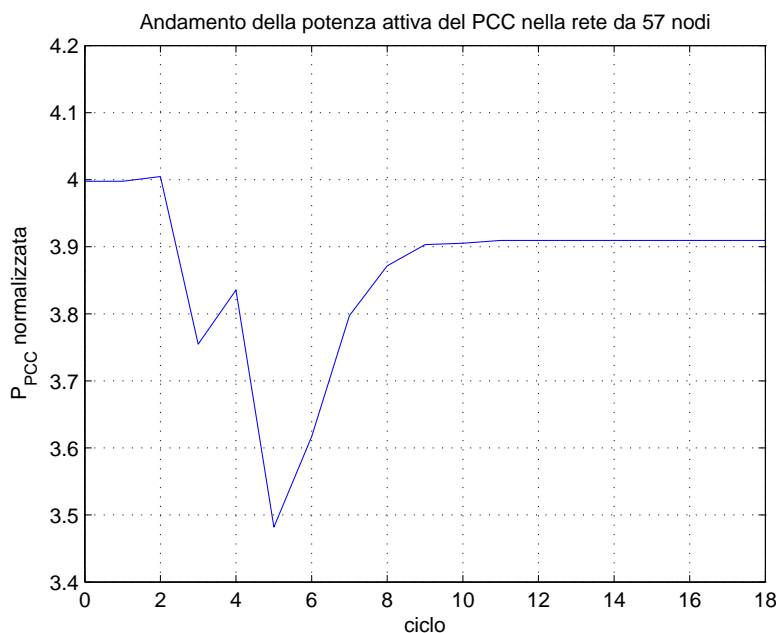


Figura 5.3: Processo di assestamento della potenza attiva del PCC nella rete da 57 nodi in seguito all'applicazione dell'Interior Point Method.

Con tale obiettivo in mente, si prenda in considerazione la grandezza P_{tot} : essa è pari a 20.43 per la rete da 14 nodi, 18.175 per quella da 30 nodi e 61.998 per quella da 57 nodi. Calcolando il rapporto fra P_{PCC} e P_{tot} si ottiene proprio la percentuale di potenza fornita dal nodo PCC rispetto a quella complessivamente richiesta dalla rete. Si ha in particolare:

- Rete con 14 nodi $\longrightarrow \frac{P_{PCC}}{P_{tot}} = \frac{2.0679}{20.43} = 0.10122 = 10.122\%$;
- Rete con 30 nodi $\longrightarrow \frac{P_{PCC}}{P_{tot}} = \frac{2.3226}{18.175} = 0.12779 = 12.779\%$;
- Rete con 57 nodi $\longrightarrow \frac{P_{PCC}}{P_{tot}} = \frac{3.9094}{61.998} = 0.06306 = 6.306\%$.

Il resto della potenza è chiaramente prodotta dai generatori locali (i PEIs), che nel nostro contesto sono associati a fonti di energia rinnovabili. Proprio per questo motivo è auspicabile che ad essi venga richiesta la maggior parte dell'energia di cui necessita la rete.

Osservando le immagini 5.1, 5.2 e 5.3, ci si accorge inoltre che il rapporto P_{PCC}/P_{tot} non subisce una variazione considerevole fra prima e dopo l'utilizzo dell'Interior Point Method. Come anticipato nel capitolo 3, ciò è dovuto al fatto che, per giungere a convergenza, l'algoritmo necessita di partire da una condizione iniziale sufficientemente vicina a quella di ottimo, soprattutto quando esso viene utilizzato per risolvere un problema di ottimizzazione non convesso, come nel nostro caso. È questa una delle principali carenze del solutore centralizzato. Il suo punto di forza maggiore sta invece nella velocità di convergenza: per la rete da 14 nodi sono infatti sufficienti 8 cicli dell'algoritmo per giungere a regime e anche per quelle da 30 e 57 nodi non è necessario affrontarne più di 18. Ciò si traduce in un tempo di elaborazione (T_{el}) di circa 0.1s nel primo caso, 0.5s nel secondo e 1.7s nel terzo.

Per comprovare l'avvenuta convergenza dell'algoritmo, un'ultima grandezza che può essere presa in considerazione è la potenza reattiva associata alle Power Electric Interfaces delle tre reti in esame (Figura 5.4, Figura 5.5 e Figura 5.6).

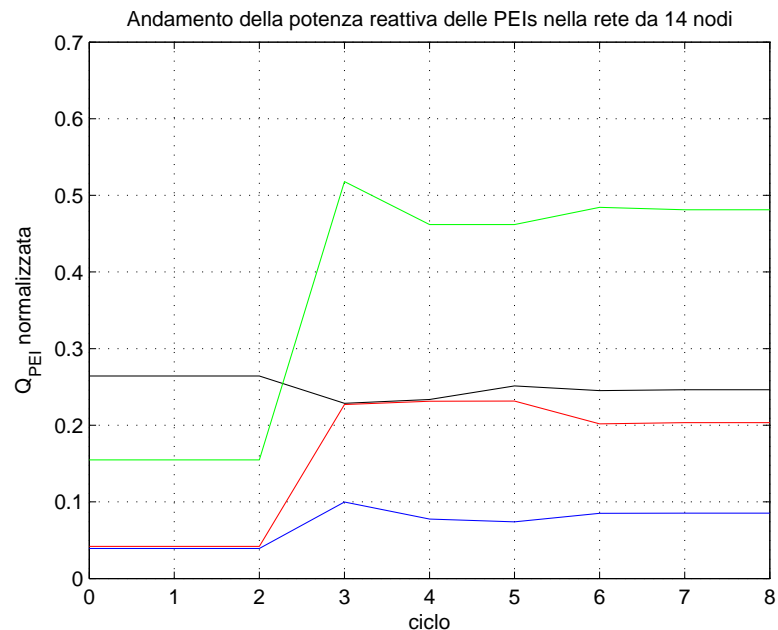


Figura 5.4: Processo di assestamento della potenza reattiva delle PEIs nella rete da 14 nodi in seguito all'applicazione dell'Interior Point Method.

5. RISULTATI DELLE SIMULAZIONI

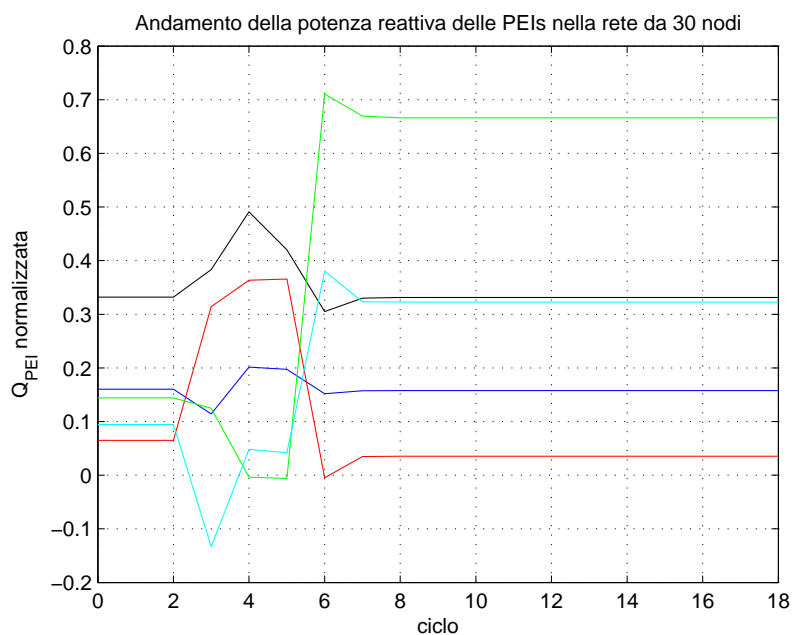


Figura 5.5: Processo di assestamento della potenza reattiva delle PEIs nella rete da 30 nodi in seguito all'applicazione dell'Interior Point Method.

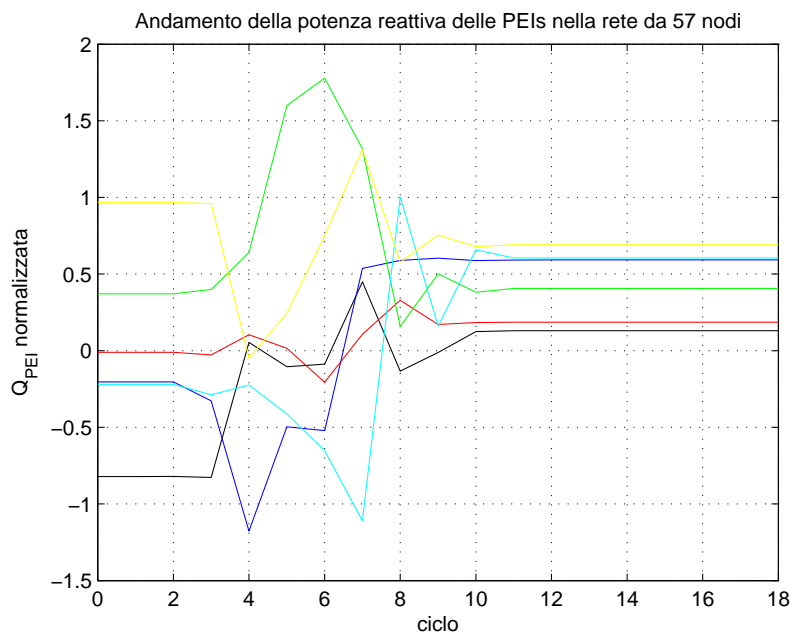


Figura 5.6: Processo di assestamento della potenza reattiva delle PEIs nella rete da 57 nodi in seguito all'applicazione dell'Interior Point Method.

Essa non risulta vincolata, almeno direttamente, dalle richieste del modello relativo all'Optimal Power Flow. Il suo andamento, come detto ci permette quindi di avere una conferma del fatto che l'algoritmo sia effettivamente giunto a regime. Dopo un breve transitorio, infatti, è possibile osservare come tutte le curve si assestino attorno ad un valore ben preciso e vi rimangano fisse fino all'ultima iterazione dell'IPM. La tabella sottostante riassume tutte le principali grandezze esaminate nel presente paragrafo per valutare le prestazioni dell'Interior Point Method.

	rete con 14 nodi	rete con 30 nodi	rete con 57 nodi
P_{PCC}	2.0679	2.3226	3.9094
P_{tot}	20.43	18.175	61.998
P_{PCC}/P_{tot}	10.122%	12.779%	6.306%
# iterazioni	8	18	18
T_{el}	0.1s	0.5s	1.7s

Tabella 5.1: Informazioni sulle performance dell'Interior Point Method applicato a tre reti di dimensioni diverse.

5.2 Performance dell'ADMM

Anche per l'Alternating Direction Multiplier Method i grafici più importanti sono quelli riportati in Figura 5.7, Figura 5.8 e Figura 5.9. Essi contengono l'andamento della funzione di costo durante il processo di ottimizzazione indotto dall'algoritmo e sono stati generati a partire dalle solite tre reti dell'IEEE. Ciò che salta subito all'occhio è la lunga durata associata al transitorio in tutti e tre i casi. Prima di arrivare ad una situazione di regime, nella quale il valore della P_{PCC} non varia più apprezzabilmente, sono infatti necessarie circa 3000 iterazioni per la rete da 14 nodi, circa 4000 per quella da 30 nodi e circa 10000 per quella da 57 nodi. Ciò si traduce in un tempo di elaborazione (T_{el}) di 15s nel primo caso, di 50s nel secondo e di 210s nel terzo. Tale fatto può essere spiegato ricordando che, ad ogni ciclo, l'algoritmo distribuito ragiona in ambito locale per giungere ad una soluzione che sia valida per l'intera rete.

In questo processo, però, le decisioni repute migliori per un singolo nodo possono rivelarsi controproducenti per un qualsiasi altro nodo e portare ad un

5. RISULTATI DELLE SIMULAZIONI

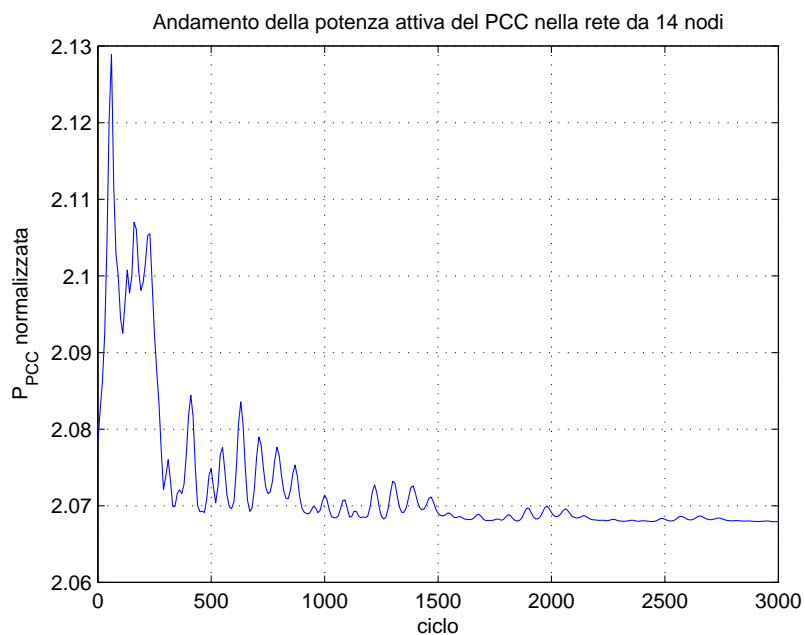


Figura 5.7: Processo di assestamento della potenza attiva del PCC nella rete da 14 nodi in seguito all'applicazione dell'Alternating Direction Multiplier Method.

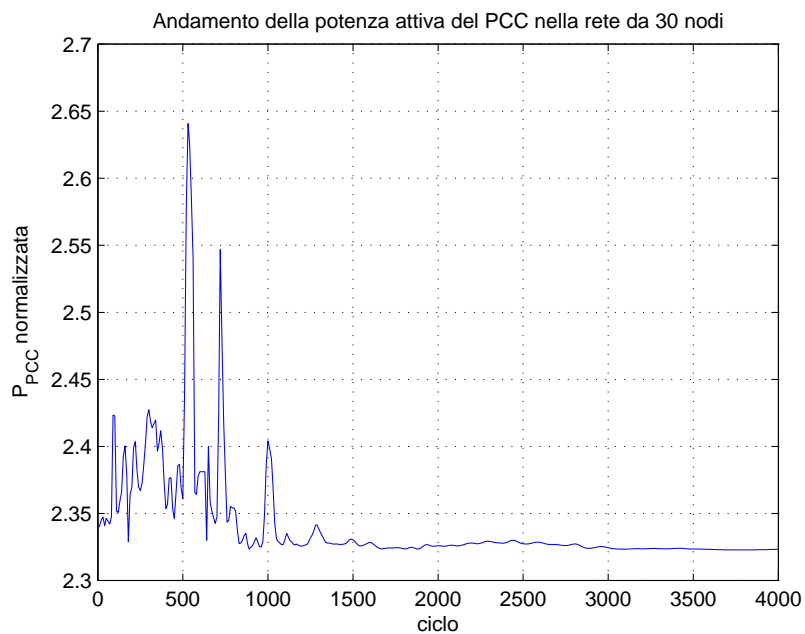


Figura 5.8: Processo di assestamento della potenza attiva del PCC nella rete da 30 nodi in seguito all'applicazione dell'Alternating Direction Multiplier Method.

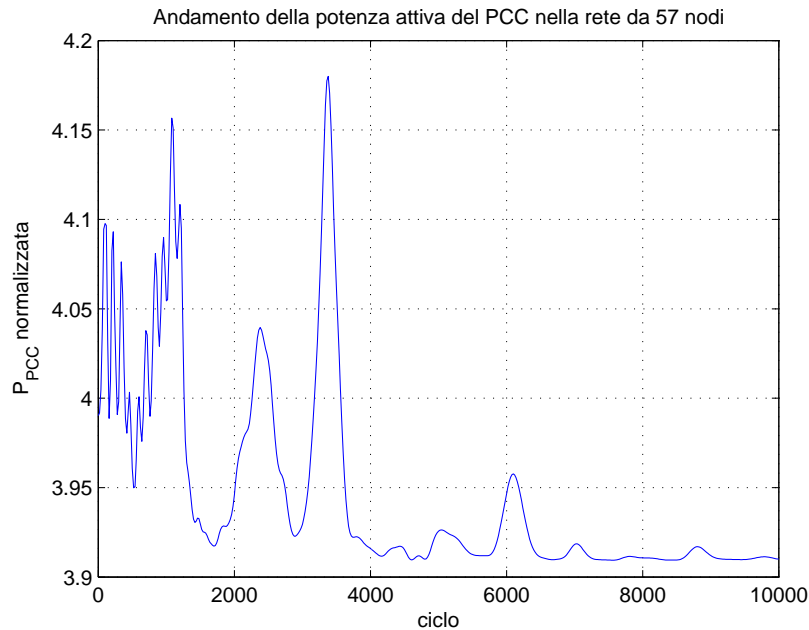


Figura 5.9: Processo di assestamento della potenza attiva del PCC nella rete da 57 nodi in seguito all'applicazione dell'Alternating Direction Multiplier Method.

temporaneo allontanamento dalla situazione di ottimo. La necessità di bilanciare le richieste di tutti, arrivando infine ad una sorta di “accordo fra le parti”, è quindi il motivo che porta ad un incremento nel numero di iterazioni. Inoltre, per come è stato implementato da noi l'ADMM, esso non utilizza alcuna tecnica di parallelizzazione che permetterebbe l'elaborazione simultanea di ciascun nodo di rete all'interno di ogni ciclo e comporterebbe quindi un netto accorciamento nella durata di quest'ultimo. D'altro canto, uno dei vantaggi dell'algorithm distribuito è rappresentato dal fatto che, durante la fase di ottimizzazione, gli stati attraversati sono sempre ammissibili per il modello (2.14). Questo perchè, ad ogni iterazione, ciascun nodo assicura il rispetto dei vincoli di rete ad esso associati.

Analogamente a quanto era stato fatto in precedenza per l'IPM, in Figura 5.10, Figura 5.11 e Figura 5.12 vengono riportate le curve che rappresentano il progressivo assestamento del valore della potenza reattiva delle Power Electric Interfaces nella rete da 14, 30 e 57 nodi rispettivamente. Da un rapido confronto di queste immagini con le corrispettive figure 5.4, 5.5 e 5.6 si può apprezzare come il valore raggiunto a regime da IPM e ADMM sia ancora una volta praticamente coincidente.

5. RISULTATI DELLE SIMULAZIONI

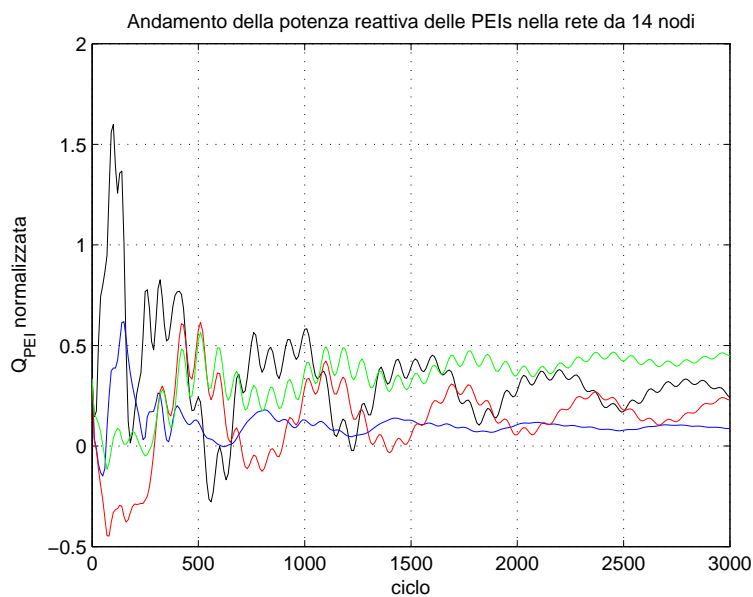


Figura 5.10: Processo di assestamento della potenza reattiva delle PEIs nella rete da 14 nodi in seguito all'applicazione dell'Alternating Direction Multiplier Method.

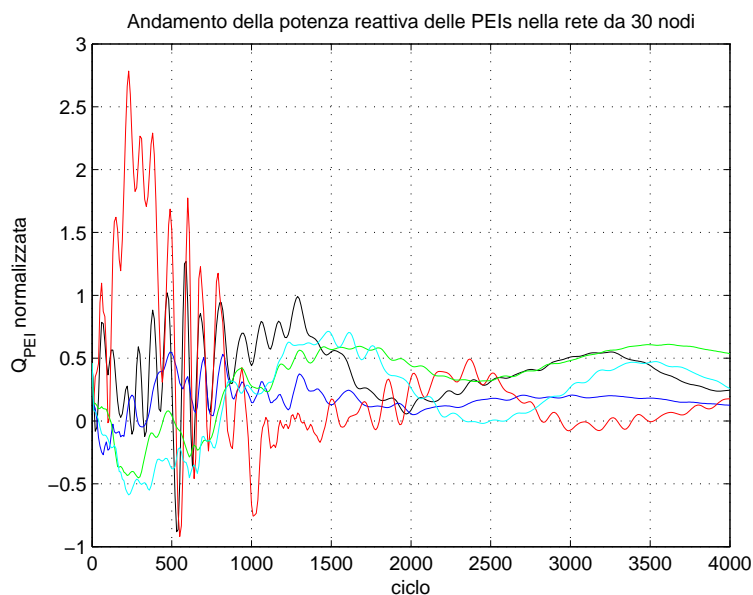


Figura 5.11: Processo di assestamento della potenza reattiva delle PEIs nella rete da 30 nodi in seguito all'applicazione dell'Alternating Direction Multiplier Method.

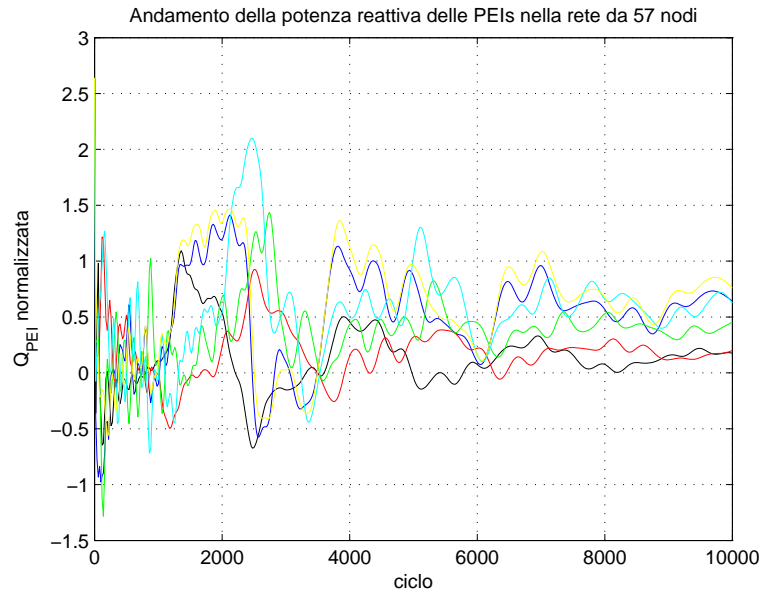


Figura 5.12: Processo di assestamento della potenza reattiva delle PEIs nella rete da 57 nodi in seguito all'applicazione dell'Alternating Direction Multiplier Method.

Per completare lo studio sulle performance dell'ADMM, non rimane che dare uno sguardo alla potenza attiva del nodo PCC (P_{PCC}) a regime e calcolare in che misura essa provveda alle richieste energetiche globali (P_{tot}) della rete. Osservando le figure 5.7, 5.8 e 5.9, si può subito apprezzare come, una volta giunti a convergenza, si abbia $P_{PCC} = 2.0679$ per la rete da 14 nodi, $P_{PCC} = 2.3233$ per quella da 30 e $P_{PCC} = 3.91$ per quella da 57. I dati sulle potenze attive rimangono invece rispettivamente: $P_{tot} = 20.43$, $P_{tot} = 18.175$ e $P_{tot} = 61.998$. Con un ragionamento analogo a quello visto nel paragrafo precedente, si ha dunque:

- Rete con 14 nodi $\longrightarrow \frac{P_{PCC}}{P_{tot}} = \frac{2.0679}{20.43} = 0.10122 = 10.122\%$;
- Rete con 30 nodi $\longrightarrow \frac{P_{PCC}}{P_{tot}} = \frac{2.3233}{18.175} = 0.12783 = 12.783\%$;
- Rete con 57 nodi $\longrightarrow \frac{P_{PCC}}{P_{tot}} = \frac{3.91}{61.998} = 0.06307 = 6.307\%$.

5. RISULTATI DELLE SIMULAZIONI

Anche per l'Alternating Direction Multiplier Method, i principali indicatori prestazionali sopra analizzati sono riassunti nella tabella qui di seguito riportata.

	rete con 14 nodi	rete con 30 nodi	rete con 57 nodi
P_{PCC}	2.0679	2.3233	3.91
P_{tot}	20.43	18.175	61.998
P_{PCC}/P_{tot}	10.122%	12.783%	6.307%
# iterazioni	3000	4000	10000
T_{el}	15s	50s	210s

Tabella 5.2: Informazioni sulle performance dell'Alternating Direction Multiplier Method applicato a tre reti di dimensioni diverse.

5.3 Confronto fra gli algoritmi

I grafici 5.13, 5.14 e 5.15 mettono a confronto le traiettorie seguite dall'IPM e dall'ADMM nella loro ricerca del punto di ottimo.

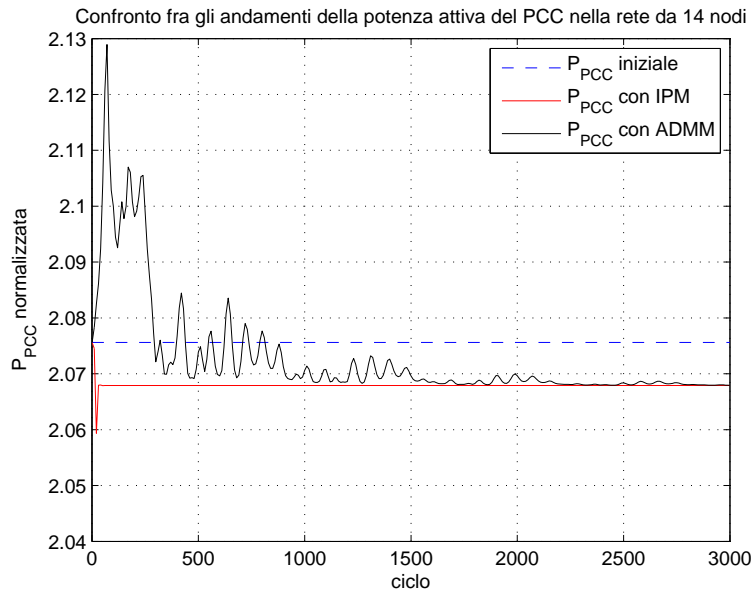


Figura 5.13: Processo di assestamento della potenza attiva del PCC nella rete da 14 nodi.

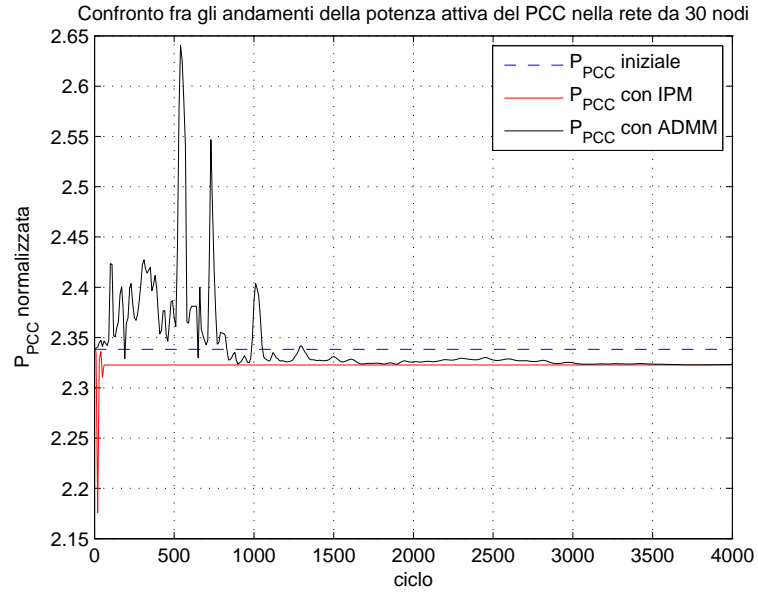


Figura 5.14: Processo di assestamento della potenza attiva del PCC nella rete da 30 nodi.

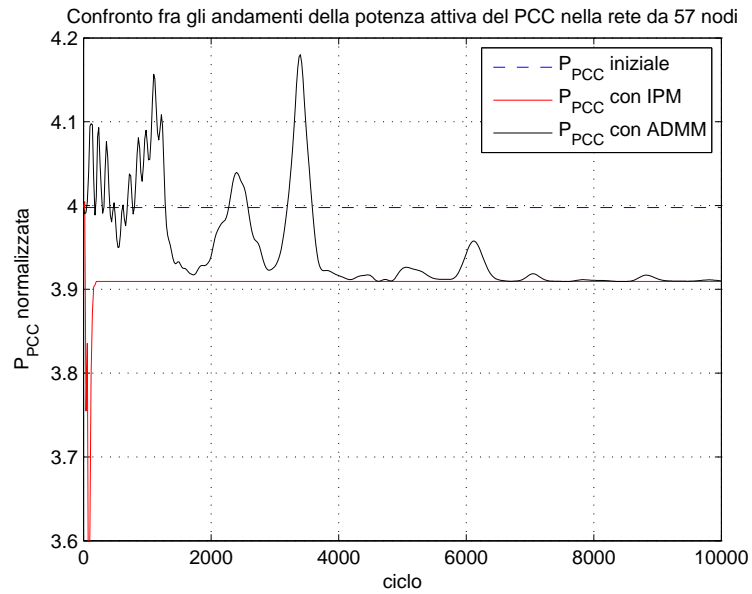


Figura 5.15: Processo di assestamento della potenza attiva del PCC nella rete da 57 nodi.

5. RISULTATI DELLE SIMULAZIONI

Osservando tali andamenti e comparando i dati presenti in Tabella 5.1 con quelli riportati in Tabella 5.2, si può apprezzare come, una volta giunti a regime, per tutte e tre le reti proposte la potenza attiva associata al PCC sia pressochè la stessa sia in seguito all'applicazione dell'IPM che a quella dell'ADMM. Scavando più a fondo, si può verificare che, a convergenza raggiunta, anche il vettore \mathbf{V} contenente le tensioni di rete risulta identico nei due casi. Tale fatto è comprovato dall'osservazione di Figura 5.16, Figura 5.17 e Figura 5.18, in cui sono riportati i valori delle tensioni relative a ciascun nodo per le tre reti in esame in seguito all'applicazione sia dell'IPM che dell'ADMM. Il grado di ottimizzazione ottenuto con i due algoritmi è quindi identico² e, valutando esclusivamente questo parametro, nessuno dei due solutori risulta preferibile rispetto all'altro. Mettendo a confronto Interior Point Method e Alternating Direction Multiplier Method, un altro criterio da tenere in debita considerazione è quello relativo alla velocità con cui i due algoritmi giungono a convergenza.

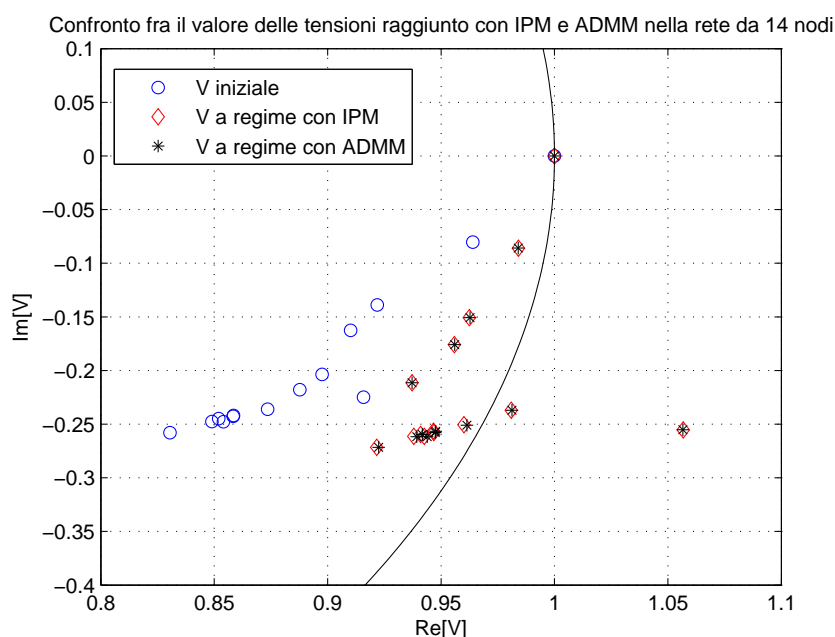


Figura 5.16: Valore delle tensioni iniziali e a regime nella rete da 14 nodi.

²La convergenza di risultati ottenuti con IPM e ADMM può essere verificata anche confrontando i grafici relativi alla potenza reattiva dei nodi PEIs, che avevamo riportato nei due paragrafi precedenti.

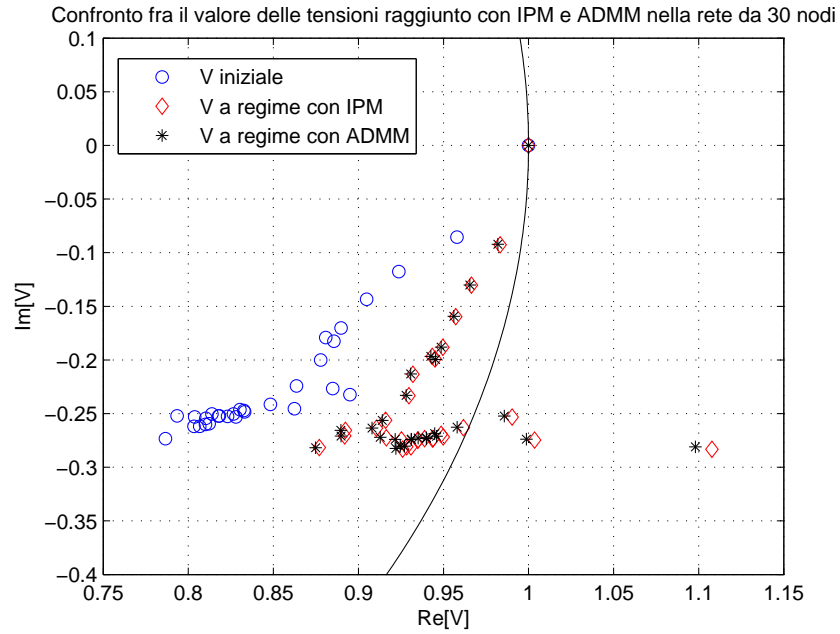


Figura 5.17: Valore delle tensioni iniziali e a regime nella rete da 30 nodi.

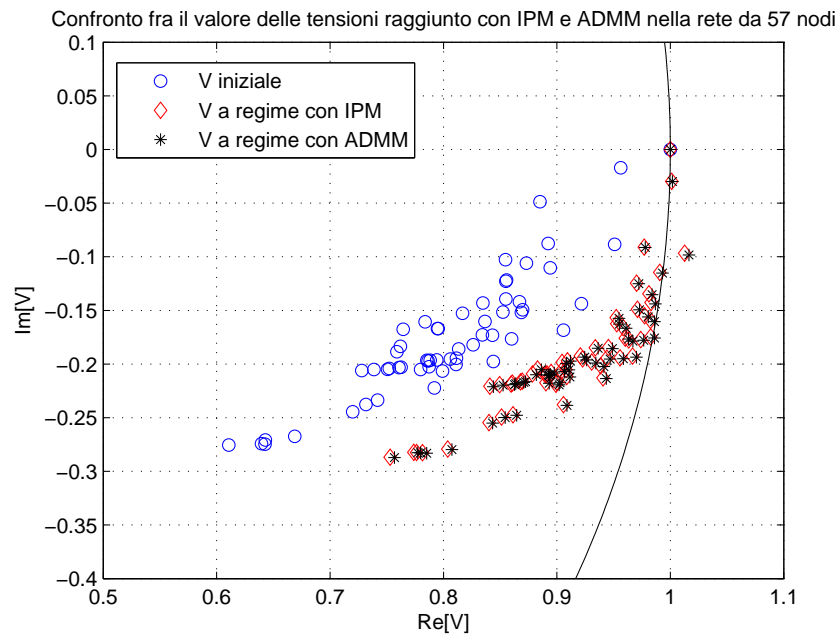


Figura 5.18: Valore delle tensioni iniziali e a regime nella rete da 57 nodi.

Da questo punto di vista il primo solutore si rivela essere decisamente più appetibile del secondo. Nel paragrafo precedente abbiamo infatti visto come il tempo di convergenza (T_{el}) dell'IPM sia di tre ordini di grandezza inferiore rispetto a quello dell'ADMM, in tutte e tre le reti da noi considerate. Vista la sua portata, tale risultato oscura completamente i vantaggi associati all'utilizzo di un algoritmo di tipo distribuito, che chiariremo qui di seguito.

Nel capitolo 4 abbiamo stabilito come, ad ogni ciclo dell'Alternating Direction Multiplier Method, ciascun nodo collabori al processo di determinazione dell'ottimo attraverso un'elaborazione condotta in locale. Si è inoltre visto che, per eseguire tale elaborazione, ciascun nodo necessita di conoscere esclusivamente alcune informazioni relative ai propri adiacenti. Per questo motivo, ogni scambio di messaggi avviene esclusivamente fra nodi fisicamente vicini in rete e può essere effettuato sfruttando sapientemente gli stessi link elettrici grazie all'impiego della *Power Line Communication* (PLC). Nel caso si voglia implementare l'Interior Point Method, invece, è indispensabile prevedere l'utilizzo di un controllore che raccolga e processi i dati associati a tutti i nodi della rete. Per fare ciò le soluzioni attuabili sono due:

1. Costruire una rete a stella a supporto di quella elettrica ed adibita esclusivamente allo scambio di informazioni sensibili per la determinazione dell'ottimo. Il nodo accentratore sarà chiaramente rappresentato proprio dal controllore citato in precedenza. Tale espediente risulta però dispendioso dal punto di vista economico sia perchè devono essere dedicate risorse alla costruzione della nuova rete informativa, sia perchè quest'ultima deve essere costantemente mantenuta per assicurarne il funzionamento.
2. Connettere direttamente alla rete elettrica il controllore centralizzato e sfruttare anche stavolta i link elettrici per far arrivare a quest'ultimo i dati. Con questo secondo stratagemma si ravvisa però una notevole inefficienza nella trasmissione delle informazioni provenienti da nodi fisicamente remoti rispetto al controllore. Per giungere a destinazione, infatti, esse devono percorrere lunghe distanze attraversando anche molteplici nodi che semplicemente le scarteranno.

A prescindere dalla soluzione adottata, inoltre, un qualsiasi controllore centralizzato deve essere dotato di capacità elaborative e memoria ragguardevoli, visto

l'ingente numero di dati che è costretto a processare. Per sfruttare una tecnica distribuita come quella dell'ADMM, invece, è sì necessario fare ricorso ad una maggior quantità di controllori ma essi possono presentare caratteristiche decisamente meno performanti.

Capitolo 6

Sviluppi futuri

Rileggendo con occhio critico i risultati del capitolo precedente si può intuire come la tecnica dell'IPM, applicata al problema dell'OPF, abbia raggiunto oggi una maturità tale da fornire performance decisamente ragguardevoli. Calato nel medesimo contesto, l'ADMM risulta invece ancora acerbo, almeno nella forma da noi presentata. Come visto, tale “inadeguatezza” è relativa in particolare alla sua complessità computazionale. In questa prima parte del capitolo si presentano quindi alcuni spunti per un suo sviluppo futuro.

Il primo accorgimento che si può pensare di adottare è di reimplementare l'algoritmo distribuito imponendo però che le elaborazioni locali relative a ciascun nodo siano effettuate in concomitanza l'una con l'altra. Un simile espediente è utilizzabile perchè, come più volte anticipato, ad ogni ciclo dell'ADMM ciascun nodo può tranquillamente svolgere i calcoli di propria competenza in maniera del tutto indipendente rispetto agli altri. In effetti, le uniche informazioni di cui egli necessita sono quelle ricevute prima di iniziare l'elaborazione in locale dai nodi adiacenti. Ricorrere al “calcolo parallelo” può così rappresentare un utile strumento per ridurre la durata di un generico ciclo dell'ADMM. Si supponga infatti che il tempo di elaborazione associato ad un qualsiasi nodo sia rappresentabile attraverso una variabile pseudo-aleatoria che assume con probabilità 1 il valore t (espresso in secondi). Con la versione da noi implementata per condurre le simulazioni del capitolo 5, l'algoritmo impiegherebbe $N \cdot t$ secondi¹ per portare a termine tutte le elaborazioni locali. Sfruttando la parallelizzazione, tale obiettivo verrebbe invece raggiunto in soli t secondi vista la contemporaneità con cui i

¹ N = numero di nodi della rete.

calcoli potrebbero essere effettuati.

Anche ricorrendo ad una simile strategia, rimane però il problema relativo al numero elevato di iterazioni richieste dall'ADMM per arrivare a convergenza. L'idea che potrebbe migliorare tale comportamento è quella di non procedere più nell'elaborazione ragionando nodo per nodo, ma di aggregare questi ultimi in cluster (come viene fatto in [14]). Ciò comporta chiaramente di rivedere la struttura dei sottoproblemi generati dall'ADMM per riassumere localmente le richieste dell'Optimal Power Flow. Dato che essi saranno associati non più ad un unico nodo ma ad una loro moltitudine, presumibilmente si riveleranno essere di più difficile soluzione. Come contraltare, però, l'elaborazione effettuata a livello di cluster potrebbe sfruttare più dati e quindi più gradi di libertà nella ricerca dell'ottimo e ciò potrebbe portare auspicabilmente ad un'accelerazione nei tempi di convergenza dell'algoritmo.

Oltre a quelli appena citati per migliorare l'Alternating Direction Multiplier Method, la tesi offre ulteriori spunti per possibili sviluppi futuri. In primo luogo sarebbe interessante indagare quale comportamento presentano i due algoritmi da noi analizzati con reti di dimensioni maggiori rispetto a quelle considerate nelle simulazioni. Ampliando in tal senso il loro bacino di applicazione si potrebbe infatti condurre uno studio più approfondito sulla complessità computazionale di IPM e ADMM, stabilendo in che modo essi reagiscono di fronte ad un aumento del numero di variabili e di vincoli in gioco nel problema di ottimizzazione. In seconda battuta si potrebbe pensare di modificare lo stesso modello di Power Flow Optimization di partenza, inglobando al suo interno logiche di natura diverse da quelle da noi considerate. In effetti, nella stesura dell'elaborato, noi ci siamo focalizzati esclusivamente su un criterio di minimizzazione degli sprechi energetici, ma all'interno di una smart-grid si possono individuare numerosi altri obiettivi. Uno dei più ricorrenti è ad esempio rappresentato dalla riduzione dei costi di approvvigionamento di ciascun utente che si allaccia alla rete. Infine, lo studio affrontato in questa tesi potrebbe essere arricchito da un'analisi delle prestazioni conseguibili con algoritmi diversi dall'Interior Point Method e dall'Alternating Direction Multiplier Method (si veda ad esempio [9]). Una volta stabilito il migliore fra tutti, si avrebbe in mano un importante tassello per lo sviluppo futuro delle smart-grid e quindi dell'intera rete elettrica mondiale.

Appendice A

Strategia di aggiornamento delle variabili nell'ADMM

Si prenda come riferimento il modello (4.5) e se ne calcoli il Lagrangiano associato:

$$L(\mathbf{x}_i, \mathbf{z}_{i,j}, \lambda_{i,j}) = \sum_{i \in \mathcal{N}} f_i(\mathbf{x}_i) + \sum_{j \in \mathcal{N}_i} 2\lambda_{i,j}^T \mathbf{P}_{i,j}(\mathbf{x}_i - \mathbf{z}_{i,j}) + c_{i,j}(\mathbf{x}_i - \mathbf{z}_{i,j})^T \mathbf{P}_{i,j}(\mathbf{x}_i - \mathbf{z}_{i,j}) \quad (\text{A.1})$$

dove i coefficienti $\lambda_{i,j}$ identificano i moltiplicatori di Lagrange. Ciò che l'Alternating Direction Multiplier Method si ripromette di fare è determinare quali aggiornamenti sequenziali devono essere intrapresi per minimizzare tale Lagrangiano. Essi vengono così declinati:

$$\begin{aligned} \{\mathbf{x}_i(t+1)\} &= \operatorname{argmin}_{\mathbf{x}_i \in \mathcal{C}_i} L(\mathbf{x}_i(t), \mathbf{z}_{i,j}(t), \lambda_{i,j}(t)) \\ \{\mathbf{z}_{i,j}(t+1)\} &= \operatorname{argmin}_{\mathbf{z}_{i,j} | \mathbf{P}_{i,j} \mathbf{z}_{i,j} = \mathbf{P}_{i,j} \mathbf{z}_{j,i}} L(\mathbf{x}_i(t+1), \mathbf{z}_{i,j}(t), \lambda_{i,j}(t)) \\ \lambda_{i,j}(t+1) &= \lambda_{i,j}(t) + c_{i,j}(\mathbf{x}_i(t+1) - \mathbf{z}_{i,j}(t+1)) \end{aligned} \quad (\text{A.2})$$

per $t \geq 0$. Gli stati iniziali $\lambda_{i,j}(0)$ e $\mathbf{z}_{i,j}(0)$ sono impostati arbitrariamente $\forall i = 1, \dots, N; j \in \mathcal{N}_i$. Estraeendo da (A.1) le informazioni relative ad un singolo nodo i e rimuovendo alcuni termini costanti ininfluenti nella determinazione del punto di minimo, si può riscrivere la prima delle equazioni in (A.2) nel seguente modo:

$$\mathbf{x}_i(t+1) = \operatorname{argmin}_{\mathbf{x}_i \in \mathcal{C}_i} f_i(\mathbf{x}_i) + \sum_{j \in \mathcal{N}_i} \mathbf{x}_i^T \mathbf{P}_{i,j} (2\lambda_{i,j}(t) + c_{i,j} \mathbf{x}_i - 2c_{i,j} \mathbf{z}_{i,j}(t)) \quad (\text{A.3})$$

Calcolando invece il gradiente relativo alla seconda espressione di (A.2), imponendolo pari a 0 e risolvendo rispetto a $\mathbf{z}_{i,j}$ si ottiene:

$$\mathbf{P}_{i,j} \mathbf{z}_{i,j}(t+1) = \mathbf{P}_{i,j} \frac{c_{i,j} \mathbf{x}_i(t+1) + c_{j,i} \mathbf{x}_j(t+1)}{c_{i,j} + c_{j,i}} + \mathbf{P}_{i,j} \frac{\lambda_{i,j}(t) + \lambda_{j,i}(t)}{c_{i,j} + c_{j,i}} \quad (\text{A.4})$$

Combinando (A.4) con l'ultima uguaglianza di (A.2) si ha:

$$\mathbf{P}_{i,j}\lambda_{i,j}(t+1) + \mathbf{P}_{i,j}\lambda_{j,i}(t+1) = 0 \quad (\text{A.5})$$

per $t \geq 0$. Perciò l'aggiornamento di $\mathbf{z}_{i,j}$ si semplifica in:

$$\mathbf{P}_{i,j}\mathbf{z}_{i,j}(t+1) = \mathbf{P}_{i,j} \frac{c_{i,j}\mathbf{x}_i(t+1) + c_{j,i}\mathbf{x}_j(t+1)}{c_{i,j} + c_{j,i}} \quad (\text{A.6})$$

che è valida per $t \geq 1$. Quindi in questo intervallo temporale $\mathbf{z}_{i,j}(t)$ è una funzione istantanea di $\mathbf{x}_i(t)$ e non necessita di alcuna memoria sul valore passato degli stati. Per rendere valida la (A.6) anche per $t = 0$ è sufficiente imporre $\lambda_{i,j}(0) = 0$. Sostituendo infine l'espressione (A.6) all'interno dell'aggiornamento di $\mathbf{x}_i(t)$ riportato in (A.3), si perviene finalmente alla formula (4.6). Inoltre (4.7) e (4.8) sono ricavabili a partire dal seguente sistema di equazioni:

$$\begin{aligned} \mathbf{M}_i &= \sum_{j \in \mathcal{N}_i} c_{i,j} \mathbf{P}_{i,j} \\ \mathbf{U}_i(t) &= 2 \sum_{j \in \mathcal{N}_i} c_{i,j} \mathbf{P}_{i,j} (\mathbf{z}_{i,j}(t) - 2\mathbf{M}_i \mathbf{x}_i(t)) \\ \mathbf{\Lambda}_i(t) &= -2 \sum_{j \in \mathcal{N}_i} \mathbf{P}_{i,j} \lambda_{i,j}(t) \end{aligned} \quad (\text{A.7})$$

In particolare, l'espressione di $\mathbf{U}_i(t)$ in (4.8) è ottenuta dalla sostituzione di (A.6) in (A.7) e da qualche semplice arrangiamento matematico, mentre l'espressione di $\mathbf{\Lambda}_i(t)$ è ottenuta dallo sfruttamento della terza relazione di (A.2) all'interno di (A.7).

Bibliografia

- [1] H. H. Happ, “Optimal Power Dispatch - A Comprehensive Survey”, *IEEE Transactions on Power Apparatus Systems*, vol. PAS-96, pp.841-854, 1977.
- [2] N. Karmarkar, “A New Polynomial-time Algorithm for Linear Programming”, *Combinatorica* vol. 4, pp. 373-395, 1984.
- [3] “Partial Description of the IEEE Common Data Format for the Exchange of Solved Load Flow Data”, disponibile all’indirizzo <http://www.ee.washington.edu/research/pstca/formats/cdf.txt>.
- [4] S. Boyd & L. Vandenberghe, *Convex Optimization*, Cambridge, 2005, pp.22-26.
- [5] Q. Jiang & G. Geng, “A reduced-space interior point method for transient stability constrained optimal power flow”, *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1232-1240, 2010.
- [6] W. Tinney & C. Hart, “Power flow solution by Newton’s method”, *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-86, no. 11, pp. 1449-1460, 1967.
- [7] H. Dommel & W. Tinney, “Optimal power flow solutions” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-87, no. 10, pp. 1866-1876, 1968.
- [8] H. Wei, H. Sasaki & J. Kubokawa, “An interior point nonlinear programming for optimal power flow problems with a novel data structure”, *IEEE Transactions on Power Systems*, vol. 13, no. 3, 1998.

BIBLIOGRAFIA

- [9] X. Bai, H. Wei, K. Fujisawa & Y. Wang, “Semidefinite programming for optimal power flow problems”, *Int. Jour. Electrical Power and Energy Systems*, vol. 30, no. 6-7, pp. 383-392, 2008.
- [10] Y. C. Wu, A. S. Debs & R. E. Marsten, “A direct nonlinear predictor-corrector primal-dual interior point algorithm for optimal power flows”, *IEEE Transactions on Power Systems*, vol. 9, no. 2, pp. 870-877, 1998.
- [11] G. L. Torres & V. U. Quintana, “An interior-point method for nonlinear optimal power flow using voltage rectangular coordinates”, *IEEE Transactions on Power Systems*, vol. 13, no. 4, 1998.
- [12] T. Erseghe, “A distributed and scalable processing method based upon ADMM”, *Signal Processing Letters, IEEE*, vol. 19, no. 9, pp. 563-566, 2012.
- [13] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers”, *Found. Trends Mach. Learn.*, vol. 3, pp. 1-122, 2011.
- [14] B. H. Kim, “A comparison of distributed optimal power flow algorithms”, *IEEE Transactions on Power Systems*, vol. 15, no. 2, 2000.

Ringraziamenti

Voglio prima di tutto ringraziare i miei genitori, ai quali dedico questa tesi, che mi sono sempre stati vicini in questi anni di università, festeggiando ogni mio piccolo successo e confortandomi nei momenti di tristezza. Grazie soprattutto per aver sopportato i lati spigolosi del mio carattere, dimostrando in più di un'occasione una pazienza ed una fiducia incrollabili. Grazie per non avermi fatto mancare nulla in tutti questi anni e per avermi permesso di raggiungere oggi un importante traguardo per il mio futuro. So di avervi dimostrato la mia gratitudine molte meno volte di quelle che avreste meritato, ma spero che leggendo queste poche righe possiate capire quanto vi voglio bene.

Ringrazio poi il professor Tomaso Erseghe, che reputo una delle persone più brillanti conosciute durante la mia carriera universitaria: grazie per tutti gli insegnamenti che mi ha dato, sia quelli strettamente legati allo svolgimento della tesi, sia quelli che esulano da essa.

Un enorme ringraziamento va anche a Francesco (Spazzola), Francesco (Sonda) e Chethan. Grazie per l'amicizia sincera che ci lega, per tutti i consigli che mi date, ma anche per tutti quelli che mi chiedete e che dimostrano la vostra fiducia in me. Grazie per i pomeriggi trascorsi a "studiare" insieme, per le innumerevoli serate passate a parlare e divertirsi in compagnia e per tutte le esperienze che avete voluto condividere con me. Grazie soprattutto perchè, se gli anni delle superiori in primis e dell'università poi hanno avuto un sapore per me speciale, lo devo in gran parte a voi.

Un sincero grazie anche a mia sorella, che ha sempre tentato di farmi ragionare nei momenti più difficili della mia vita. Mi dispiace non dimostrarti molto spesso il mio affetto, ma voglio che tu sappia che tengo molto a te.

Grazie all'inimitabile Gio per essere stato una presenza costante nella mia infanzia e in tutti gli anni a seguire. Grazie anche a mia cugina Marina e a mia zia Elisa, che hanno sempre fatto il tifo per me e che non hanno mai mancato di farmi sentire il loro appoggio.

Grazie di cuore alla Maria che, da quando ho memoria, ha sempre fatto parte della mia vita e che mi ha cresciuto dimostrandomi sempre grande affetto. Se mi sono trasformato nella persona che sono oggi è anche grazie ai tuoi insegnamenti.

Grazie a tutti i componenti della compagnia del mare perchè, anche se non ci vediamo per 10 mesi all'anno, riuscite a rendere ogni vacanza a Bibione un'espe-

BIBLIOGRAFIA

rienza a suo modo indimenticabile.

Grazie ai miei compagni di università ed in particolare a Claudio, Silvia e Leonardo, con cui ho condiviso la maggior parte del percorso di studi e ad Alex che ho incontrato solo negli ultimi due anni di magistrale. Grazie per avermi aiutato nel difficile cammino che ho scelto.

Ci tengo a ringraziare anche tutti gli apprendisti della Telecom, con cui sto condividendo la prima, bellissima esperienza lavorativa della mia vita. Grazie perchè se ogni mattina inizia con il sorriso in bocca è certamente anche merito vostro.

Un grazie particolare va alla persona con cui posso dire di aver condiviso tutti o quasi gli istanti della mia vita e che anche in questo momento è qui al mio fianco. Grazie a mio fratello-gemello Andrea perchè ci sei sempre stato per me, perchè ci lega un rapporto che poche persone al mondo possono dire di aver sperimentato e che io sono contento di poter condividere ogni giorno con te.

Infine rivolgo un ringraziamento a tutte le persone che, in un modo o nell'altro, hanno fatto parte della mia esistenza ed hanno avuto un significato per essa. Grazie per esserci state e per aver riempito i miei primi 25 anni di vita di ricordi felici.