

# Synthèse bibliographique : Implémentation GPU sur Julia d'un simulateur de marché endogène

TARET Baptiste encadré par THOMAS Béatrice

14 Octobre 2024

## 1 Contexte

La transition énergétique vers une économie décarbonée repose sur une intégration massive des énergies renouvelables, telles que l'énergie éolienne et solaire, dans les réseaux électriques [1]. Ces sources d'énergie sont par nature fluctuantes mais aussi nettement décentralisées. En effet, à terme, ces énergies sont vouées à être intégrées au sein d'un nombre croissant de foyers. Par ailleurs, la volonté de respecter l'environnement conduit à une intensification du déploiement de solutions électriques, ce qui accroît les flux d'énergie.

Cela induit une complexité grandissante dans la gestion des flux électriques, notamment à travers les réseaux de distribution. Dans un tel modèle, les agents des réseaux électriques deviennent à la fois capables de consommer et de produire, ce sont des prosumers (Figure 1). Le flux d'énergie devient bidirectionnel, le consommateur ne se contente plus uniquement de prélever de l'énergie au réseau, il peut aussi devenir producteur et injecter de l'électricité dans le réseau.

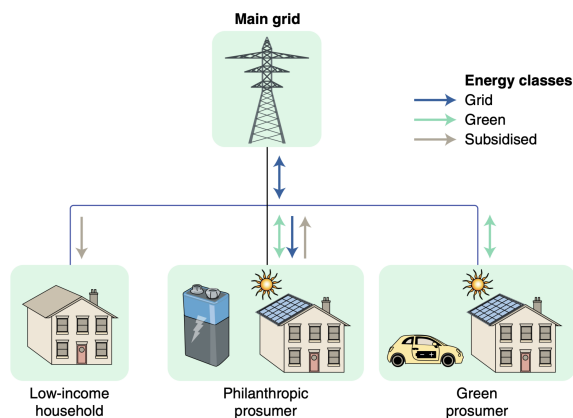


FIGURE 1 – Schéma du comportement des agents vis-à-vis du réseau électrique [2]

Parallèlement, l'émergence des smart grids, ou réseaux électriques intelligents, ouvre la voie à une gestion plus flexible et réactive de ces systèmes grâce à l'intégration des technologies de l'information et de la communication (TIC) au sein des infrastructures énergétiques. Les smart grids permettent de mieux répondre aux fluctuations de la production renouvelable et à l'évolution des comportements de consommation, notamment via la gestion des flexibilités apportées par les prosumers. Alors, deux données caractérisent ces nouveaux opérateurs : le paramètre besoin (a) et le critère de coût/bénéfice (b) (Figure 2). L'analyse de l'équilibre de ce marché se limite à ces deux paramètres. Cependant, cet équilibre n'est réalisable uniquement si les lois physiques (lois de Kirchhoff, conservation de la puissance active) sont respectées. L'objectif est donc de répondre au problème de l'offre et de la demande de manière optimale par l'intermédiaire d'une configuration du réseau électrique physiquement réalisable. De plus, ce résultat d'analyse doit être obtenu en temps réel pour pouvoir adapter l'état du réseau en conséquence. Dans ce cadre d'étude, les contraintes réseau ne sont prises en compte (chutes de tension, puissance réactive, ...) et ce sont les prosumers seuls qui régissent les échanges, le marché est désigné endogène.

## 2 Modèle mathématique du marché endogène

Pour répondre à ce problème, un cadre mathématique formel prenant en compte à la fois les lois physiques du réseau électrique et l'enjeu de l'offre et de la demande est érigé. L'étude se place sur l'ensemble  $\Omega$  des prosumers. Alors, une fonction coût  $f_n$  est définie pour chaque agent  $n$  et prend en argument la quantité totale d'énergie échangée par l'opérateur  $n$ ,

notée  $p_n$ . Alors, le but est de minimiser la fonction  $f$  [3] :

$$f = \sum_{n \in \Omega} f_n(p_n)$$

Où, la variable  $p_n$  est positive lorsque l'agent  $n$  injecte de l'énergie dans le réseau et négative lorsqu'il en prélève.

Cependant, cette optimisation se fait sous contrainte. Le réseau électrique impose que l'énergie totale échangée est nulle :

$$\sum_{n \in \Omega} p_n = 0$$

En outre, chaque agent a des besoins énergétiques qui sont caractérisés par la définition d'un intervalle de validité de  $p_n$ .

$$\underline{p}_n \leq p_n \leq \overline{p}_n, \quad n \in \Omega$$

Avec  $\underline{p}_n$  la borne inférieure de l'énergie que souhaite échanger l'opérateur  $n$  et  $\overline{p}_n$  sa borne supérieure.

Ce modèle mathématique, représentant l'état du réseau électrique à chaque instant, étant posé, il est maintenant possible de résoudre le problème informatiquement. La résolution de ce type de problème avec un grand nombre d'agents est complexe. En effet, la fonction  $f$  n'est pas systématiquement convexe, ce qui pose alors un problème d'unicité et d'existence d'une solution optimale. Pour pallier cela, la fonction  $f$  est approximée par une fonction convexe dont l'existence et l'unicité du minimum sont assurées. La résolution est alors réalisée sur cette fonction auxiliaire [4].

## 3 Gestion du réseau électrique

### 3.1 Modèle de traitement centralisé

Dans les marchés énergétiques traditionnels, même si la production et la consommation d'énergie sont en partie décentralisées, sa production est elle centralisée. Un opérateur central régule la distribution de l'énergie en fonction des besoins des consommateurs et des capacités des producteurs, comme illustré en figure 2. Cette approche devient de plus en plus inefficace avec l'augmentation des ressources énergétiques distribuées (DER) et la multiplication des agents (producteurs, consommateurs,

prosumers). En effet, cette expansion du nombre d'acteurs provoque une élévation du nombre d'interactions et donc du volume de données à traiter par l'opérateur central. Cela engendre un encombrement du réseau de communication. L'opérateur central doit collecter, traiter et redistribuer les données de l'ensemble des acteurs en temps réel, entraînant des retards et des inefficacités dans la prise de décision. Les algorithmes permettant de résoudre les problèmes de distribution de l'énergie du réseau ont une complexité en temps qui croît considérablement avec le nombre de données d'entrées. Ce problème se résout uniquement par un renforcement du réseau, très gourmand en matières premières [5]. En outre, la gestion centralisée de l'énergie pose une difficulté en termes de confidentialité et de sécurité [6]. Lorsque les données ne sont plus locales, elles risquent d'être interceptées par des tiers.

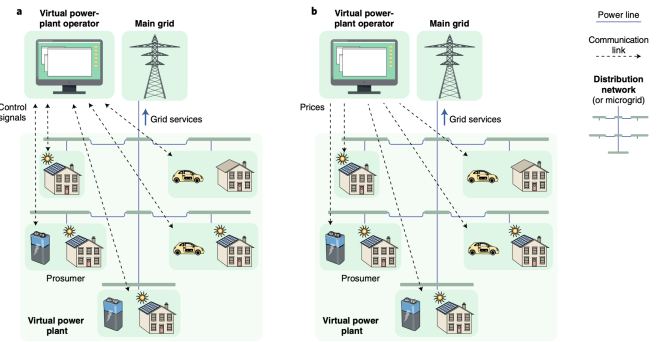


FIGURE 2 – Schéma du fonctionnement d'un réseau smart grid centralisé [2]

### 3.2 Modèle de traitement décentralisé

Pour surpasser les obstacles posés par une centralisation de la gestion du réseau électrique, il faut s'affranchir de l'échange d'informations visant uniquement à traiter les données des agents par un opérateur central. Dans ce but, il est nécessaire de décentraliser les moyens de calcul pour que les données soient traitées au plus près de l'agent à qui elles appartiennent. Ainsi, l'aller-retour servant uniquement à l'exploitation de l'information est supprimé. Cela induit un gain de temps sur le transit de données et supprime le problème de confidentialité que pose la centralisation de la gestion de l'énergie.

Cependant, le traitement des besoins de chaque prosumer doit prendre en compte l'état du réseau et les besoins des autres agents. Cela implique la né-

cessité de créer un réseau de communication entre chaque agent pour qu'ils puissent échanger et répondre eux-mêmes aux problèmes de l'offre et de la demande. Dans ce cadre, chaque agent doit négocier simultanément avec tout autre pair en fonction de sa propre fonction de coût et des objectifs globaux [3]. Ce système de commerce multi-bilatéral coordonné s'appelle le marché pair-à-pair (Figure 3). Dans ce modèle, ce n'est plus un opérateur central qui récolte toutes les données et qui résout le problème, mais c'est chaque agent qui résout de son côté une partie du problème global. Alors, les agents (consommateurs, producteurs et prosumers) échangent directement de l'énergie entre eux sans intermédiaire central, permettant ainsi une flexibilité et une réactivité accrues face aux variations de la production renouvelable et à la demande. Dans ce modèle, chaque transaction doit bénéficier mutuellement aux utilisateurs impliqués.

Ce système de stockage et de transmission d'informations qui fonctionne sans organe central de contrôle est inspiré de la blockchain. Ainsi, il garantit la confidentialité puisque chaque transaction est validée et enregistrée de manière transparente par l'ensemble des participants, sans besoin d'intermédiaire [7].

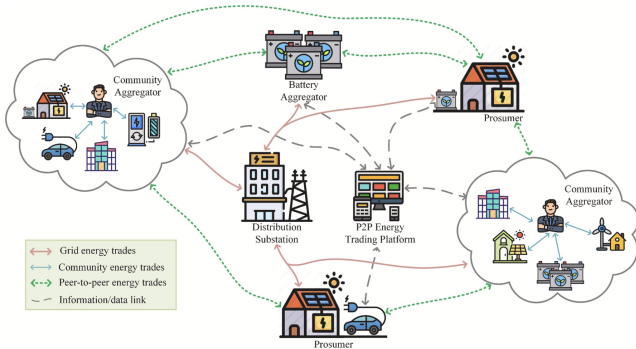


FIGURE 3 – Schéma d'un réseau smart grid décentralisé basé sur le fonctionnement d'un marché pair-à-pair [8]

Dans un tel schéma l'agent  $n$  se voit attribué une puissance fonction des puissances échangées avec les autres agents du réseau formant un ensemble noté  $\omega_n$ . Ainsi, on a :

$$p_n = \sum_{m \in \omega_n} p_{nm}$$

Où,  $p_{nm}$  désigne la puissance échangé de l'agent

$n$  vers l'agent  $m$ .

Derrière ce marché pair-à-pair présenté ci-dessus, de nombreuses communications différentes sont créées (flèches grises sur la figure). Pour un souci de lisibilité, toutes les interconnexions ne sont pas représentées, mais elles sont sous-entendues. En réalité, les relations sont identiques à celles présentées en figure 4. Le nombre de liens évolue quadratiquement avec le nombre d'agents. En effet, si on note  $n$  le nombre d'agents, du fait d'échanges multi-bilatéraux d'informations, il y a  $C = \frac{n(n-1)}{2}$  connexions.

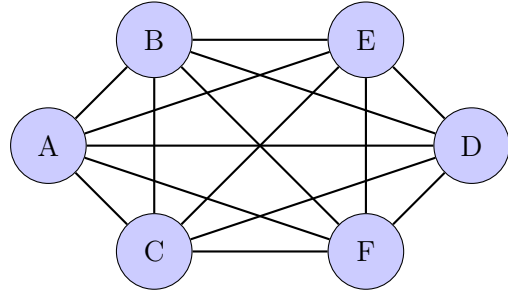


FIGURE 4 – Schéma des inter-connexions dans un marché pair-à-pair

Ainsi, cette multiplication quadratique du nombre d'interactions entre les agents complique la gestion et l'optimisation des transactions. Ces méthodes décentralisées de résolution risquent donc d'être ralenties par des limites de vitesse de transmission de l'information. Alors, ce n'est plus uniquement le traitement des données qui risque d'influencer les retards dans la gestion du réseau. Or, la résolution de manière décentralisée du problème d'équilibrage d'un marché endogène vise non seulement à préserver la confidentialité, mais aussi à être plus rapide et fiable qu'une gestion traditionnelle du réseau. Pour ce faire, le processus d'optimisation n'attend pas d'avoir l'ensemble des données de tous les utilisateurs à chaque itération. Cela est rendu possible par une implémentation asynchrone des algorithmes, faite dans le but de rendre le système plus résilient aux variations de retard [9].

## 4 Résolution du problème

Si l'on considère un nombre croissant d'agents intervenant dans le marché pair-à-pair, leur influence sur le réseau ne peut plus être négligée. Il est donc important de considérer les contraintes réseau lors

de la résolution du marché. Cette prise en compte peut être faite via l'utilisation d'un Power Flow ou d'un Optimal Power Flow. Il s'agit de réaliser un algorithme permettant de minimiser la fonction coût, notée  $f$ , introduite en deuxième partie. Ce processus vise à trouver les paramètres optimaux respectant les contraintes physiques du réseau électrique. Cela fait alors appel aux deux concepts précédemment introduits : le Power Flow (PF) et l'Optimal Power Flow (OPF).

Le PF est une technique qui consiste à analyser le comportement d'un réseau électrique en déterminant les valeurs des variables électriques telles que les tensions, les courants, les puissances actives et réactives aux différents nœuds (ou bus) du réseau. Le principal objectif de l'analyse de flux de puissance est de comprendre comment l'électricité circule à travers le réseau sous certaines conditions de charge et de génération.

L'OPF est une extension du PF classique qui intègre l'optimisation dans le processus d'analyse. Le but de l'OPF est de déterminer les valeurs optimales des variables de contrôle du réseau (telles que les tensions, les puissances actives et réactives des générateurs, et les positions des transformateurs) afin de minimiser un objectif donné tout en respectant les contraintes techniques et physiques du système. Il s'agit de minimiser les coûts de l'énergie, d'optimiser la gestion des ressources en fonction des besoins des agents tout en respectant les contraintes opérationnelles [10]. L'organisation globale de ce processus est le suivant :

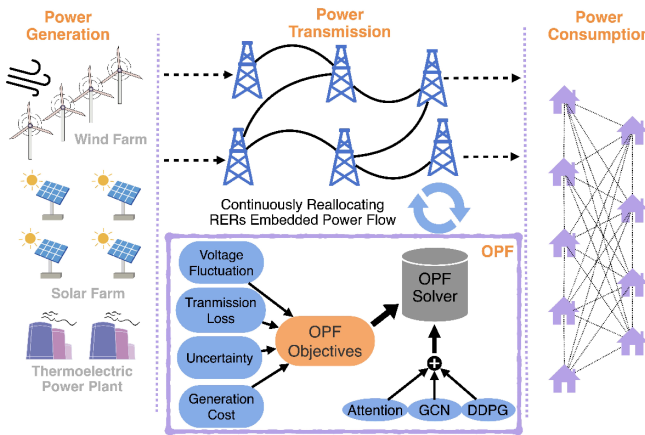


FIGURE 5 – Schéma du fonctionnement de l'OPF [11]

Ainsi, le marché endogène est un marché pair-à-

pair avec des contraintes sur le PF, c'est-à-dire un OPF où la fonction coût à minimiser est un marché pair-à-pair et non un marché centralisé.

## 4.1 Algorithmes d'optimisation

Pour approcher les solutions de ces équations non linéaires, il convient d'utiliser des méthodes numériques itératives. Les problèmes considérés sont approximatés de telle manière qu'ils puissent être considérés convexes. Ainsi, il existe un unique minimum local qui est donc assurément le minimum global, et la convergence vers cette solution est garantie ; elle sera notée  $P_{opt}$ . Le vecteur contenant les puissances de tous les agents sera noté  $P = (p_n)_{n \in \Omega}$ .

### 4.1.1 Méthode de Newton-Raphson

La méthode de Newton-Raphson est une méthode itérative utilisée pour résoudre des systèmes d'équations non linéaires, en particulier dans l'analyse des flux de puissance des systèmes électriques. Elle repose sur une approximation linéaire des équations grâce à un développement de Taylor. Cette approche fait donc appel au gradient d'une fonction et consiste à chercher le point d'annulation de ce gradient en linéarisant son expression par l'intermédiaire d'un développement limité au premier ordre. Cela permet d'obtenir le minimum de cette fonction (endroit où son gradient est nul). Cette méthode est largement utilisée pour sa rapidité de convergence dans les systèmes bien conditionnés, mais elle est sensible aux choix des valeurs initiales et peut échouer si ces valeurs sont mal choisies [12]. L'expression du schéma de récurrence pour une fonction coût  $f$  est le suivant :

$$P^{k+1} = P^k - H(P^k)^{-1} \nabla f(P^k)$$

Où,  $H$  est la matrice Hessienne,  $P$  le vecteur puissance à la  $k$ -ième itération et avec  $\lim_{k \rightarrow \infty} P^k = P_{opt}$ .

### 4.1.2 Méthode de Gauss-Seidel

L'algorithme de Gauss-Seidel est une approche itérative plus simple que Newton-Raphson, utilisée pour les systèmes d'équations linéaires.

Elle est basée sur la mise à jour successive des valeurs des variables en utilisant les valeurs déjà calculées, permettant ainsi une approche plus progressive

Critère	ADMM	Gauss-Seidel	Newton-Raphson
Vitesse de convergence	Modeste à lente	Linéaire	Quadratique
Nombre d'itérations	Variable (en fonction du problème)	Élevé	Faible
Complexité computationnelle	Modérée (calcul distribué)	Faible (mise à jour successive)	Élevée (calcul de la Jacobienne)
Adaptabilité	Très adaptable pour le calcul distribué	Simple et robuste	Sensible aux valeurs initiales

TABLE 1 – Comparaison des algorithmes Newton-Raphson, Gauss-Seidel et ADMM.

et moins coûteuse en calculs. Cependant, elle tend à converger plus lentement que Newton-Raphson [12]. Cette approche consiste à résoudre une équation du type  $AP = b$  d'inconnue  $P \in \mathbb{R}^{|\Omega|}$  où  $A \in M_{|\Omega|}(\mathbb{R})$  et  $b \in \mathbb{R}^{|\Omega|}$  par itérations successives dans le but de diagonaliser le système d'équations. La formule finale est la suivante :

$$p_i^{k+1} = \frac{1}{a_{kk}} \left( b_k - \sum_{j=1}^{i-1} a_{ij} p_j^{k+1} - \sum_{j=i+1}^{|\Omega|} a_{ij} p_j^k \right),$$

$$i \in \llbracket 1, |\Omega| \rrbracket$$

Où,  $P^k = (p_1^k, p_2^k, \dots, p_{|\Omega|}^k)$  est le vecteur puissance à la  $k$ -ième itération et avec  $\lim_{k \rightarrow \infty} P^k = P_{opt}$ .

#### 4.1.3 Algorithme des Directions Alternées des Multiplicateurs (ADMM)

Les deux méthodes précédentes ne sont pas utilisables telles quelles pour résoudre un problème d'optimisation sous contraintes décentralisé. Pour pallier cette limite, l'algorithme des Directions Alternées des Multiplicateurs (ADMM) est employé. L'ADMM consiste à remplacer le problème d'optimisation sous contrainte par de nouveaux problèmes d'optimisation sans contraintes. Ces nouveaux problèmes proviennent du problème initial auquel a été ajouté un terme de pénalité ; ils peuvent ensuite être résolus avec les méthodes de Newton-Raphson ou de Gauss-Seidel [13].

Pour cela le Lagrangien augmenté  $\mathcal{L}$  est défini de la manière suivante :

$$\mathcal{L}(p, y) = \sum_{n \in \Omega} \mathcal{L}_n(p_n, y) = \sum_{n \in \Omega} (f_n(p_n) + y^T \cdot p_n)$$

Où,  $y \in \mathbb{R}^{|\Omega|}$  le multiplicateur de Lagrange et  $\mathcal{L}_n$  le Lagrangien augmenté associé à la fonction coût de l'agent  $n$ ,  $f_n$ .

De telle sorte que résoudre problème originel défini en deuxième partie revient à trouver le couple :

$$p_n^{k+1} = \operatorname{argmin}_{p_n \in [\underline{p}_n; \bar{p}_n]} L_i(p_n, y^k)$$

$$y^{k+1} = y^k + \alpha^k p_n^{k+1}.$$

Où, "argmin" désigne l'argument du minimum et  $\alpha$  le paramètre de pénalité.

Cette approche est particulièrement adaptée dans notre cas d'optimisation décentralisée, car elle permet de résoudre des sous-problèmes en parallèle, très efficace dans des environnements distribués. Ainsi, l'ADMM permet à chaque agent d'optimiser sa propre fonction coût tout en intégrant les contraintes du réseau.

La table 1 propose une synthèse des caractéristiques de ces trois algorithmes d'optimisation de problèmes convexes sous contraintes. Il apparaît en vert ce qui concerne la décentralisation et en bleu ce qui se rapporte à la résolution.

## 4.2 Simulation du marché décentralisé

En prenant en considération les méthodes de résolution présentées, le problème d'optimisation est transformé : chaque agent  $n$  doit minimiser sa propre fonction coût  $f_n$ . Le problème revient donc à itérer de façon à minimiser la fonction  $f$  définie comme suit à l'itération  $k$  :

$$f^k = \sum_{n \in \Omega} f_n(p_n^k)$$

Où,



$$p_n^k = \operatorname{argmin}_{p_n \in [\underline{p}_n; \overline{p}_n]} (f_n(p_n) + (y^{k-1})^T p_n)$$

Et,

$$y^k = y^{k-1} + \alpha^{k-1} p_n^k$$

Ce problème de minimisation est ensuite résolu à l'aide d'une des deux méthodes d'optimisation présentées en partie 4.1.

Ensuite, pour développer un modèle de gestion du réseau décentralisé, il est nécessaire de simuler son fonctionnement afin de vérifier son bon fonctionnement. Ce processus est centralisé (spatialement) et consiste à répartir chaque agent simulé sur une architecture de calcul qui lui est réservée. L'objectif est de simuler cette gestion du réseau le plus rapidement possible. Alors, pour rendre le processus de résolution efficace, il est indispensable d'optimiser la répartition des calculs sur l'architecture de calcul de chaque agent.

Une architecture de calcul est composée de CPU (Central Processing Unit) et de GPU (Graphics Processing Unit).

Le CPU, avec ses quelques cœurs optimisés pour des tâches séquentielles, excelle dans la gestion des processus logiques et de contrôle. En revanche, le GPU, avec des milliers de cœurs optimisés pour le calcul parallèle, est idéal pour les calculs massivement parallèles nécessaires dans la simulation de grands systèmes électriques, comme les études de flux de puissance. La figure 6 expose les différences d'architecture entre ces deux unités.

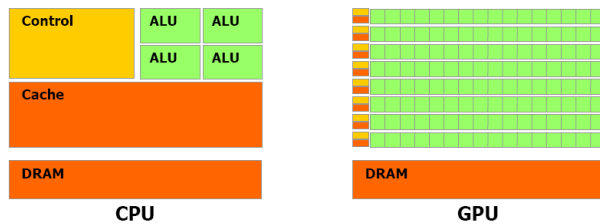


FIGURE 6 – Schéma de la structure d'un CPU et d'un GPU [14]

Les algorithmes présentés précédemment peuvent tirer parti de la parallélisation GPU pour améliorer considérablement les performances. Par exemple, l'étude sur l'accélération des flux de puissance sur GPU montre que la répartition des tâches entre le

CPU et le GPU permet un gain de temps substantiel en déléguant les calculs intensifs au GPU, comme la formation de la matrice Jacobienne et la mise à jour des tensions, tout en laissant au CPU les tâches de contrôle et de communication [15]. Ainsi, l'approche hybride exploite la capacité du GPU à gérer des calculs intensifs en parallèle tout en laissant le CPU coordonner les processus, garantissant une simulation rapide et scalable adaptée aux environnements en temps réel [16]. L'objectif étant de traiter les informations plus rapidement qu'elles n'arrivent afin d'être en mesure de réaliser un grand nombre de fois des simulations pour pouvoir régler un paramètre.

## Références

- [1] RTE. *Bilan prévisionnel : Futurs énergétiques 2050*. <https://assets.rte-france.com/prod/public/2023-10/2023-10-02-bilan-previsionnel-2023-principaux-resultats.pdf>.
- [2] Thomas MORSTYN et al. « Using Peer-to-Peer Energy-Trading Platforms to Incentivize Prosumers to Form Federated Power Plants ». In : *Nature Energy* 3.2 (fév. 2018), p. 94-101. ISSN : 2058-7546. DOI : 10.1038/s41560-017-0075-y.
- [3] F. MORET et al. « Negotiation Algorithms for Peer-to-Peer Electricity Markets : Computational Properties ». In : *2018 Power Systems Computation Conference (PSCC)*. Juin 2018, p. 1-7. DOI : 10.23919/PSCC.2018.8442914.
- [4] Hongye WANG et al. *On Computational Issues of Market-Based Optimal Power Flow / IEEE Journals & Magazine / IEEE Xplore*. <https://ieeexplore.ieee.org/abstract/document/4282060>.
- [5] Béatrice THOMAS et al. « Adéquation Algorithme Architecture pour le calcul d'un AC-Power Flow ». In : *Symposium du Génie Electrique*. 2023.
- [6] Tatiana CHERNOVA et Elena GRYAZINA. « Peer-to-Peer Market with Network Constraints, User Preferences and Network Charges ». In : *International Journal of Electrical Power & Energy Systems* 131 (oct. 2021), p. 106981. ISSN : 0142-0615. DOI : 10.1016/j.ijepes.2021.106981.

- [7] Thomas BAROCHE et al. *Exogenous Cost Allocation in Peer-to-Peer Electricity Markets / IEEE Journals & Magazine / IEEE Xplore*. <https://ieeexplore.ieee.org/abstract/document/8630697>.
- [8] Sachinkumar SUTHAR, S. Hari Charan CHERUKURI et Naran M. PINDORIYA. « Peer-to-Peer Energy Trading in Smart Grid : Frameworks, Implementation Methodologies, and Demonstration Projects ». In : *Electric Power Systems Research* 214 (jan. 2023), p. 108907. ISSN : 0378-7796. DOI : 10.1016/j.epsr.2022.108907.
- [9] Alyssia DONG et al. *Asynchronous Algorithm of an Endogenous Peer-to-Peer Electricity Market / IEEE Conference Publication / IEEE Xplore*. <https://ieeexplore.ieee.org/abstract/document/9495009>.
- [10] Allen J WOOD, Bruce F WOLLENBERG et Gerald B SHEBLÉ. *Power generation, operation, and control*. John Wiley & Sons, 2013.
- [11] Jinhao LI et al. « Deep reinforcement learning for optimal power flow with renewables using spatial-temporal graph information ». In : *arXiv preprint* (2021).
- [12] H ABAALI, Elhachmi TALBI et Rachid SKOURI. « Comparison of newton Raphson and gauss seidel methods for power flow analysis ». In : *World Acad. Sci. Eng. Technol. Int. J. Energy Power Eng* 12.9 (2018), p. 627-633.
- [13] Stephen BOYD et al. « Distributed optimization and statistical learning via the alternating direction method of multipliers ». In : *Foundations and Trends® in Machine learning* 3.1 (2011), p. 1-122.
- [14] Guillaume SALL. « Quelques algorithmes rapides pour la finance quantitative ». Thèse de doct. Déc. 2017.
- [15] Jaideep SINGH et Ipseeta ARUNI. « Accelerating power flow studies on graphics processing unit ». In : *2010 Annual IEEE India Conference (INDICON)*. IEEE. 2010, p. 1-5.
- [16] Xue LI et Fangxing LI. « GPU-based power flow analysis with Chebyshev preconditioner and conjugate gradient method ». In : *Electric Power Systems Research* 116 (2014), p. 87-93.