

Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»
Физтех-школа Прикладной Математики и Информатики
Кафедра интеллектуальных систем

Направление подготовки / специальность: 03.04.01 Прикладные математика и физика
(магистратура)

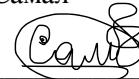
Направленность (профиль) подготовки: Математические и информационные
технологии

РАСПРЕДЕЛЕННЫЕ МЕТОДЫ ОПТИМИЗАЦИИ ДЛЯ ОЦЕНКИ СОСТОЯНИЯ ЭНЕРГЕТИЧЕСКОЙ СЕТИ

(магистерская диссертация)

Студент:

Кубентаева Самал



(подпись студента)

Научный руководитель:

Гасников Александр Владимирович,
д-р физ.-мат. наук, доц.



(подпись научного руководителя)

Консультант (при наличии):

(подпись консультанта)

Москва 2020

1 Аннотация

Популярность распределенных алгоритмов оптимизации для решения задачи Оценки Состояния Энергетической Системы (PSSE) растет. Преимуществом этих методов оптимизации является отсутствие центрального координатора, что позволяет использовать параллельные вычисления, а следовательно, приводит к ускорению работы методов. Так же благодаря основной идее разбиения исходной задачи на множество локальных подсистем, приводит к уменьшению потоков связи внутри сети, что повышает уровень кибербезопасности передаваемых данных.

В этой работе предложен новый подход для ускорения существующих алгоритмов для задачи PSSE. Используя последние результаты в теории оптимизации были получены оценки скорости сходимости нового метода. Для проверки полученных оценок, были проведены эксперименты на энергетической системе со 140 шинами и продемонстрирована состоятельность нового подхода относительно стандартных распределенных методов оптимизации. Так же было проведено исследование на чувствительность метода к изменению параметров алгоритма.

Содержание

| | | |
|----------|---|-----------|
| 1 | Аннотация | 1 |
| 2 | Введение | 3 |
| 3 | Энергетическая система | 7 |
| 3.1 | Моделирование энергетической сети | 7 |
| 4 | Постановка задачи | 12 |
| 4.1 | Задача оценки состояния энергосети | 12 |
| 4.2 | Постановка задачи | 13 |
| 5 | Методы | 16 |
| 5.1 | Общая форма согласованной оптимизации | 16 |
| 5.2 | ADMM | 19 |
| 5.3 | Новая формулировка ADMM | 20 |
| 6 | Ускоренные методы | 27 |
| 6.1 | Ускоренный градиентный метод | 28 |
| 6.2 | Ускоренный ADMM | 29 |
| 7 | Эксперименты | 31 |
| 7.1 | Пример 1 | 32 |
| 7.2 | Пример 2 | 33 |
| 7.3 | Пример 3 | 34 |
| 8 | Выводы | 35 |
| | Список литературы | 38 |

2 Введение

Начиная со второй половины XX-го века энергетические системы имели схожие этапы развития во всем мире. Так в России электростанции строились вблизи источников топлива или рек. Преимущественно заводы строились в удалении от крупных городов. Однако, Россия имеет исключительный случай ТЭЦ, который расположен в черте города и находится вблизи потребителей. Энергоснабжение потребителей осуществляется по магистральным и распределенным линиям энергопередач. Магистральные линии электропередач предназначены для передачи энергии от станции к крупным узловым подстанциям. Распределенные линии электропередач предназначены для энергоснабжения крупных городов и предприятий.

На протяжении многих лет *Централизованная энергетическая система* была устойчивой и надежной архитектурой энергетической сети. Однако, в конце XX-го века эффект "экономии масштаба" снизился по сравнению с 1950-ми годами, и нефтяной кризис 1970-х годов повысил интерес стран-импортеров энергоресурсов к улучшению эффективности и разработке новых энергетических ресурсов.

Большое влияние на развитие технологий по производству электроэнергии в Европе и США в 1970-1980гг оказала разработка газотурбинных, газопоршневых и парогазовых технологий. Таким образом повысив эффективность электростанций с низкой установленной мощностью (от 10 кВт до 10 МВт) и позволив уменьшить экономические затраты на производстве. Это привело к популярности *методов распределенной генерации мощностей*.

Раньше, в Централизованных энергетических системах применялись централизованные алгоритмы оптимизации, которые решали следующие задачи: задачу оптимального потока мощности, для минимизации веса обмениваемых данных в сети, и задачу оценки состояния энергетической системы, для восстановления параметров измеряемых величин в сети (состояние энергосистемы). В заданных задачах учитывается исходная информация о сети, такая как сопротивление линий, тепловые пределы линий, установленная мощность станции, параметры нагрузки и геометрия сети.

Идея централизованных методов оптимизации на энергетической сети, заключается в том, чтобы произвести общий сбор данных со всех участков системы на главном-серверном узле, на котором происходит пересчет параметров и далее происходит обратная передача полученных значений к узлам сети. Данный подход был эффективен при небольшом объеме системы и ранее его недостатки критически не проявлялись. Однако, с увеличением нагрузок и изменением масштаба и сложности взаимодействий участков сети, вызвало замедление работы энергосистемы при решении крупномасштабной оптимизационной задачи. Что привело к неточности и неактуальности полученного решения, а, следовательно, к ненадежной работе энергосистемы.

Эти проблемы привели к новому подходу, а именно к *распределенным методам оптимизации*. Распределенные методы кардинально отличаются от централизованных методов тем, что они разделяют вычислительную нагрузку в сети между локальными "агентами" на которых производится пересчет параметров с учетом состояний соседних узлов и обменом данных между ними. В качестве "агента" в энергосистеме может выступать либо некоторая подсистема исходной сети, либо шина, в зависимости от постановки задачи.

Распределенные методы имеют ряд серьёзных преимуществ перед централизованными. В силу особенности методов, каждый "агент" в системе учитывает состояния и значения параметров только своих соседних "агентов". Данный подход повышает безопасность хранимых и передаваемых данных, так как сокращает количество взаимодействий внутри системы. Так же это снижает загруженность каналов связи между "агентами" и приводит к повышению устойчивости к сбоям в энергосистеме. Ещё одним превосходством является способность к параллельной реализации, что приводит к значительному ускорению работы методов. Кроме того, распределенный подход гарантирует конфиденциальность передаваемых данных, заданных ограничений, функций затрат и полученных измерений.

Так же распределенные методы называют децентрализованными, так как они не имеют центрального узла-координатора, который производит пересчет параметров метода. При децентрализованном подходе, система разбивается на небольшое число подсистем ("агентов"), которые в ходе работы метода обмени-

ваются данными, полученными при решении локальной задачи. В задачах оценки состояния энергосистемы и оптимального потока мощности основным фактором точности полученного решения является, достижение минимального зазора между решениями централизованного и распределенного методов. Так как при минимизации зазора мы получим наиболее точное и полное решение для исходной системы.

В данной работе мы рассмотрим задачу оценки состояния энергетической сети (PSSE). PSSE является нелинейной задачей оптимизации, где все измерения собраны из системы SCADA. С ростом масштабов энергетической сети с Фазометрическими блоками (PMU), которые измеряют высокочастотные значения фазовых, что приводит к увеличению нагрузки на централизованных блоках вычислений. Для облегчения коммуникационной и вычислительной нагрузки было предложено децентрализованное решение [21]. При оценке состояний в системе, мы будем рассматривать локальную оценку состояний сети, которую получит каждый агент. После чего будет производиться настройка граничных состояний между соседними агентами. Таким образом будет получено единое согласованное решение для всей системы с учетом топологии соединений узлов в сети.

В работе мы будем рассматривать метод ADMM (Alternating Direction Method of Multipliers) [3], который является одним из распределенных методов оптимизации. Также мы получим его модификацию для задачи PSSE. Впервые, данный метод был представлен в работах Гловинского-Маррокко и Мерсира-Габая [13]. Позже Габай и Экштейн-Бертсекас получили теоретические оценки на сходимость метода [6]. Так же исследование метода АДММ представлено в статье [18]. Известно, что метод был получен совмещением двух методов. Первый это метод множителей Лагранжа, от которого АДММ унаследовал устойчивость при сходимости. Второй метод Двойственной декомпозиции [2], который повлиял на способность метода к распределенным вычислениям.

Решение задачи поиска оптимального потока мощности (OPF) с помощью метода АДММ было предложено в работе [7], что доказало важность распределенных методов оптимизации для решения проблем в энергетике. В данной работе рассмотрена задача PSSE для решения которой будут использованы метод АДММ

и его ускоренный вариант [9]. Практические и теоретические результаты доказали осуществимость децентрализованного подхода на электросети и продемонстрировали высокую точность и быструю сходимость методов.

Данные об энергетической сети взяты из базы данных MATPOWER для электросетей. Так же на основе реальных данных были смоделированы синтетические данные, которые обладают наихудшей геометрией соединений в сети. Так же на этих данных были проведены эксперименты.

В следующей главе будет представлено описание энергетических систем. Далее, на основе описаний электросети будет сформулирована общая задача оптимизации для оценки состояния системы. В главах 5 и 6 описаны методы АДММ и ускоренного АДММ в контексте заданной задачи и доказана теорема о скорости сходимости метода АДММ. После чего будет продемонстрирована работа этих методов на трех примерах.

3 Энергетическая система

3.1 Моделирование энергетической сети

В данном разделе будет кратко описана энергетическая система и уравнения, которым она подчиняется, так же будут введены некоторые обозначения. В работах [24], [12] дано более подробное и детальное описание физических процессов и структур энергосистем. Так же точное описание систем MATPOWER, которая используется в этой работе, можно найти в работе [23].

Мы будем рассматривать энергетическую сеть в виде связного графа $\mathcal{G} = \{\mathcal{B}, \mathcal{E}\}$:

- \mathcal{B} множество $N = |\mathcal{B}|$ шин (узлов) в системе;
- \mathcal{E} множество линий энергопередач $L = |\mathcal{E}|$.

Система содержит узлы трех типов: узлы-нагрузки, генераторы и slack-узел. Узел-нагрузки обладает известными величинами активной и реактивной мощностей. Для генераторов (узлы-генераторы) известно напряжение и активная мощность. Slack bus или, иначе, его называют балансирующим узлом, обладает известными значениями фазы и напряжения. Кроме того, последний тип узла может быть только один в энергетической сети.

- \mathbf{Y} матрица проводимости в системе \mathcal{G} ;
- Каждый узел $n \in \mathcal{B}$ в сети обладает параметрами: *напряжения* $\hat{V}_n = |V_n|e^{j\theta_n}$, где θ_n фаза на шине n , *тока* $\hat{I}_n = |I_n|e^{j\phi_n}$ и *мощности* $\hat{S}_n = P_n + jQ_n$, где P_n активная и Q_n реактивная мощности;
- Каждая линия $(m, n) \in \mathcal{E}$ передает *поток мощности* от узла m к узлу n : $S_{mn} = P_{mn} + jQ_{mn}$, где P_{mn} поток активной мощности и Q_{mn} поток реактивной мощности.

Введем следующее векторное обозначение для всех параметров в сети: $\mathbf{v} = [V_1 \dots V_N]^H$, $\mathbf{i} = [I_1 \dots I_N]^H$, $\mathbf{p} = [P_1 \dots P_N]^T$, $\mathbf{q} = [Q_1 \dots Q_N]^T$. Обозначим за \mathcal{L}

множество данных снятых с датчиков (для получения $|V_n|^2$, величин активной P_n и реактивной мощности Q_n , $n \in \mathcal{B}$).

Далее составим явные соотношения между \mathbf{v} и $|V_n|^2$, P_n , Q_n , P_{nm} , Q_{nm} . Тогда матрица проводимости в системе $\mathbf{Y} \in \mathbb{C}^{N \times N}$:

$$Y_{mn} = \begin{cases} -y_{mn}, & (m, n) \in \mathcal{E} \\ \hat{y}_{nn} + \sum_{k \in \mathcal{B}_n}, & m = n \\ 0, & \text{otherwise,} \end{cases}$$

где y_{mn} обозначает проводимость линии (m, n) , \hat{y}_{nn} проводимость узла $n \in \mathcal{B}$. Так же обозначим $\mathcal{B}_n = \{k | (n, k) \in \mathcal{E}\}$ множество ближайших соседних узлов для шины n . Матрица \mathbf{Y} является симметричной и не является Эрмитовой.

Следующее соотношение объединит закон Кирхгофа на сети и закон Ома между узловыми напряжениями и токами:

$$\mathbf{i} = \mathbf{Y}\mathbf{v}. \quad (1)$$

Заменим выражение для генерации мощности: $P_n + jQ_n = V_n I_n^*$, $\forall n \in \mathcal{B}$ в (1). Тогда получим, что

$$\mathbf{p} + j\mathbf{q} = \text{diag}(\mathbf{v})\mathbf{i}^* = \text{diag}(\mathbf{v})\mathbf{Y}^*\mathbf{v}^* \quad (2)$$

Далее получим ток на линии (m, n) из закона Ома и Кирхгофа:

$$I_{mn} = \hat{y}_{mn}V_m + y_{mn}(V_m - V_n), \quad \forall (m, n) \in \mathcal{E}, \quad (3)$$

где \hat{y}_{mn} проводимость на узле m , которая имеет линию (m, n) .

Из полученного уравнения, перепишем поток мощности для линии (m, n)

$$P_{mn} + jQ_{mn} = V_m I_{mn}^* = (\hat{y}_{mn}^* + y_{mn}^*)|V_m|^2 - y_{mn}^* V_m V_n^*, \quad \forall (m, n) \in \mathcal{E}. \quad (4)$$

Можем заметить, что в формулах (2) и (4) \mathbf{v} имеет квадратичную связь с мощностью генерации и потоком мощности.

Теперь определим матрицу $\mathbf{H} \in \mathbb{C}^{B \times B}$ для каждого параметра. Чтобы определить \mathbf{H}_{V_n} нам понадобится следующее выражение:

$$|V_n|^2 = V_n V_n^* = \mathbf{v}^H e_n e_n^T \mathbf{v}, \quad (5)$$

где $e_n \in \mathbb{R}^N$ является n -м базисным вектором в пространстве \mathbb{R}^N . Следовательно,

$$\mathbf{H}_{V_n} = e_n e_n^T. \quad (6)$$

Далее определим выражения, которые помогут составить \mathbf{H} для остальных параметров:

$$Y_n = e_n e_n^T Y, \quad \forall n \in \mathcal{B} \quad (7)$$

$$Y_{mn} = (\hat{y}_{mn} + y_{mn}) e_n e_n^T - y_{mn} e_n e_n^T, \quad \forall (m, n) \in \mathcal{E}. \quad (8)$$

Если приравнять действительные и мнимые части в уравнениях (2) и (3), то получим

$$\begin{aligned} \mathbf{H}_{P_n} &= \frac{1}{2}(Y_n + Y_n^H), & \mathbf{H}_{Q_n} &= \frac{j}{2}(Y_n - Y_n^H), \\ \mathbf{H}_{P_{mn}} &= \frac{1}{2}(Y_{mn} + Y_{mn}^H), & \mathbf{H}_{Q_{mn}} &= \frac{j}{2}(Y_{mn} - Y_{mn}^H). \end{aligned} \quad (9)$$

Далее переопределим \mathbf{H} выражения в соответствии с прямоугольным представлением напряжения $\hat{\mathbf{v}} \in \mathbb{R}^{2B}$:

$$\hat{\mathbf{v}} = \begin{bmatrix} \text{Re}(\mathbf{v}) \\ \text{Im}(\mathbf{v}) \end{bmatrix} \quad (10)$$

Тогда все матрицы из формулы (9) будут иметь следующую прямоугольную форму:

$$\hat{\mathbf{H}} = \begin{bmatrix} \text{Re}(\mathbf{H}) & -\text{Im}(\mathbf{H}) \\ \text{Im}(\mathbf{H}) & \text{Re}(\mathbf{H}) \end{bmatrix} \quad (11)$$

В общем виде уравнение баланса мощности в сети имеет вид:

$$P_k = \sum_{j=1}^N |V_k||V_j|(G_{kj} \cos(\theta_k - \theta_j) + B_{kj} \sin(\theta_k - \theta_j))$$

$$Q_k = \sum_{j=1}^N |V_k||V_j|(G_{kj} \sin(\theta_k - \theta_j) - B_{kj} \cos(\theta_k - \theta_j)),$$

где G_{kj} действительная часть элемента в матрице проводимости \mathbf{Y} , B_{kj} соответственно мнимая часть этого элемента.

Далее перейдем к DC модели, для этого введем следующие допущения для высоковольтных энергетических систем:

- линии энергопередач работают без потерь, так как сопротивления линий и емкости зарядных устройств ничтожно малы

$$y_s = \frac{1}{r_s + jx_s} \approx \frac{1}{jx_s}, \quad b_c \approx 0;$$

- разность фаз на соседних узлах мала (обычно не превышает 10-15 градусов, крайне редко составляет 30 градусов)

$$\sin(\theta_f - \theta_t - \theta_{shift}) \approx \theta_f - \theta_t - \theta_{shift};$$

- величины напряжений близки к 1.0

$$v_i = e^{j\theta_i}.$$

В силу свойств сети, где активная мощность значительно больше реактивной: $P_{mn} \gg Q_{mn}$. Приходим к выводу, что для составления задачи оптимизации следует учитывать значения активной мощности. Также мы получим что все мнимые значения в матрице \mathbf{H} будут равны 0. Следовательно, получим следующие урав-

нения:

$$\forall (m, n) \in \mathcal{E} : P_{mn} = h_{mn}(\theta_m - \theta_n) \quad (12)$$

$$\forall n \in \mathcal{B} : P_n = \sum_{m:(m,n) \in \mathcal{E}} P_{mn}, \quad (13)$$

где h_{mn} есть соответствующий элемент матрицы $\mathbf{H}_{P_{mn}}$. Уравнение (13) можно записать в виде $\mathbf{P} = \mathbf{H}\boldsymbol{\theta}$, где $\mathbf{P} \doteq (P_i | i \in \mathcal{V})$ вектор мощности генерации/потребления, $\boldsymbol{\theta} \doteq (\theta_i | i \in \mathcal{V})$ вектор фаз напряжения и \mathbf{H} взвешенная матрица Лапласа. Далее в статье мы обозначим $v_i \doteq \theta_i$, чтобы подчеркнуть что имеем дело с переменной.

4 Постановка задачи

4.1 Задача оценки состояния энергосети

Задача оценки состояния энергетической сети (PSSE) является одной из основных проблем в Энергетических системах. Крупномасштабная система имеет большое количество вычислительных единиц на узлах сети, каждая из которых имеет свою точность и погрешность измерений. Последнее является основным источником проблем в сети. Небольшая неточность в рамках одного узла при рассмотрении некоторого локального участка не несет никакой проблемы. Однако, при рассмотрении результатов относительно всей сети, эти локальные ошибки имеют свойство аккумулировать, что приводит к ненадежности всей системы. Мы не можем однозначно определить состояния на всей системе, но можем получить профиль состояний в некотором небольшой участке. Основная цель состоит в получении общего допустимого набора состояний в системе с учетом всех измерений на узлах.

Как было сказано в начале работы, растет спрос на распределенные методы, которые улучшат координацию центров управления в энергетических сетях. Эти методы для решения задачи PSSE будут распределять вычислительную нагрузку между несколькими процессорами (центрами управления в локальных участках). Распределенные методы значительно сократят время выполнения программы и гарантируют конфиденциальность передаваемых данных. Каждый процесс будет протекать автономно путем обмена информацией между соседними процессорами.

Существуют распределенные методы с иерархической структурой взаимодействия процессоров для статистической формулировки PSSE [8], так же были разработаны различные вариации в работе [25]. Распределенные схемы включают в себя блочно-якобиевские итерации [17] или алгоритм аппроксимации с заданными условиями оптимальности [4] или метод разделения матриц для подсчета инвертированной матрицы в методе Гаусса-Ньютона [1]. В большинстве из этих методов вводится предположение о независимости локальных областей. Иначе, сходимость методов не гарантируется. В работе [15] отвергли наблюдаемость век-

тора состояний всей сети в локальных областях, так как это замедлит работу алгоритма. Также были проведены исследования относительно распределенного решения задачи оценки состояний в работе [21].

4.2 Постановка задачи

Так как мы решаем задачу PSSE распределенными методами оптимизации, разобьем систему на \mathcal{K} областей. Без ограничения общности каждую область можно рассматривать как область независимого системного оператора, причем, обладающую отдельным центром управления. Данный центр может быть органом балансировки, отдельной подстанцией, городом или центром распределения электроэнергии [22].

Как было описано в главе 3 мы снимаем показания с датчиков, где \mathcal{Z} множество зашумленных данных в системе \mathcal{G} . Каждый узел системы $n \in \mathcal{B}$ имеет свое измерение $\mathbf{z}_n \in \mathcal{Z}$. Так же каждая область $k \in \mathcal{K}$ собирает M_k измерений в этой области. Тогда мы можем написать соотношение между измерением и вектором состояния в каждой области подчиняющейся линейной модели:

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{v}_k + \boldsymbol{\varepsilon}_k, \quad \forall k \in \mathcal{K}, \quad (14)$$

где $\boldsymbol{\varepsilon}_k \sim \mathcal{N}(0, \sigma_k^2)$ Гауссовы шум на данных снятых с датчиков, \mathbf{v}_k вектор состояния в области k связанный с реальными данными \mathbf{z}_k посредством матрицы \mathbf{H}_k (соответствует определенной части общей матрицы \mathbf{H} , которая отвечает за область k).

Тогда решение задачи оценки состояния энергосистемы в локальном участке k равносильно решению задачи оптимизации:

$$\min_{\mathbf{v}_k \in \mathcal{X}_k} f_k(\mathbf{v}_k) = \min_{\mathbf{v}_k \in \mathcal{X}_k} \frac{1}{2} \|\mathbf{z}_k - \mathbf{H}_k \mathbf{v}_k\|^2, \quad (15)$$

где в выпуклом множестве \mathcal{X}_k находится предварительная информацию о шинах в сетке \mathcal{G} . Например, нулевое потребление мощности или короткое замыкание, подробное изложение смотрите в [19].

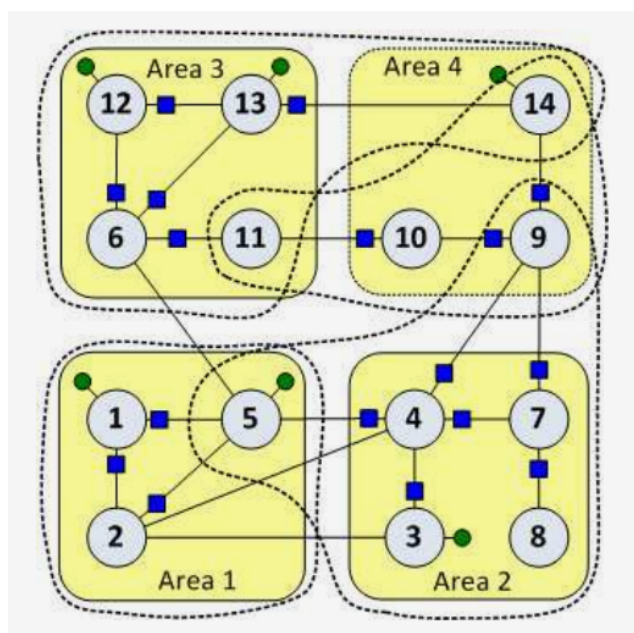


Рис. 1: Энергетическая система из 14-шин IEEE разделенная на 4 области, пунктирными линиями выделены узлы принадлежащие соответствующим областям. Измерения напряжения на узле-PMU выделены зелеными кружками и показания тока на линии синими квадратами.

Вернемся к структуре и правилам разбиения энергосистемы на области. Как показано на рисунке 1 мы разделили систему на 4 области. Причем, это разбиение подчиняется следующим правилам:

- одна область может иметь общие вершины с не более чем двумя соседними областями;
- каждая шина может принадлежать не более чем двум областям;
- каждая область будет содержать хотя бы одну шину и пересекаться хотя бы с одной соседней областью.

Мы рассматриваем два типа датчиков, присутствующих в сетке: датчики измерения тока на линии и датчики измерения напряжений на узлах. Причем, если линия соединяет шины, принадлежащие различным областям, то измерения, снятые с этого датчика, будут принадлежать одной из двух областей. Например, как показано на картинке 1, область 2 будет иметь доступ к значениям на уз-

лах $\{3, 4, 5, 7, 8, 9\}$, так как ей принадлежат датчики установленные на линиях $(4, 5), (4, 9), (7, 9)$.

Таким образом в соответствии с правилами разбиения сети, каждая область имеет общий узел с соседней областью. Тогда две области будут восстанавливать значение на общей шине, и с каждой новой итерацией будут настраиваться под значения друг друга. Поэтому, чтобы решить задачу PSSE на полной сетке шин, нам нужно ввести условия для соединений между областями. Например, четвертая область восстанавливает состояния на трех узлах $\{9, 10, 14\}$, но также собирает показания на линии $(10, 11)$. Следовательно, вектор состояния \mathbf{v}_4 содержит показания и на вершине 11, которая контролируется третьей областью. Обозначим напряжение общей шины областей k и l через $\mathbf{v}_k[l]$ и $\mathbf{v}_l[k]$. И введем новую переменную \mathbf{u}_{kl} , которая будем иметь значение консенсусной переменной с условием

$$\mathbf{v}_k[l] = \mathbf{v}_l[k] = \mathbf{u}_{kl}, \quad \forall k, l \in \mathcal{K}. \quad (16)$$

Таким образом в нашем примере $\mathbf{v}_3[4] = \mathbf{v}_4[3] = \mathbf{u}_{34}$, где \mathbf{u}_{34} состояние на узле 11.

Теперь можем записать общую задачу оптимизации для всей сети, которую мы будем решать с помощью методов ADMM и AADMM:

$$\begin{aligned} & \min_{\{\mathbf{v}_k \in \mathcal{X}_k\}} \sum_{k=1}^K f_k(\mathbf{v}_k) \\ \text{s.t.} \quad & \mathbf{v}_k[l] = \mathbf{u}_{kl} \quad \forall l \in \mathcal{B}_k \quad \forall k \in \mathcal{K}, \\ & \mathbf{v}_l[k] = \mathbf{u}_{kl} \quad \forall k \in \mathcal{B}_l \quad \forall l \in \mathcal{K}, \end{aligned} \quad (17)$$

где \mathcal{B}_k - это множество областей, имеющие общее состояние с областью k , \mathcal{B}_l соответственно с областью l .

5 Методы

В этой работе мы используем два эффективных метода решения задачи (17), которые активно используются в энергетических системах, см. работу [21]. Первый метод – это классический ADMM, который описан в статье [3]. Второй – ускоренный ADMM из статьи [9]. Мы применяем оба метода к задаче PSSE и доказываем теорему о скорости сходимости для метода ADMM применительно к нашей конкретной задаче.

5.1 Общая форма согласованной оптимизации

Задача согласованной оптимизации решает задачу вида:

$$\begin{aligned} \min \sum_{i=1}^N f_i(x_i) \\ \text{s.t. } x_i - \tilde{z}_i = 0, \quad i = 1, \dots, N, \end{aligned} \quad (18)$$

где $x_i \in \mathbb{R}^{n_i}$, $i = 1, \dots, N$ локальные переменные, функции $f_1(x_1) + \dots + f_N(x_N)$ разделены по переменным x_i . Для дальнейшего объявления переменных в (18) потребуются ввести определение и способ разделения данных.

Введем глобальную переменную $z \in \mathbb{R}^n$, где каждая локальная переменная содержит некоторый набор компонент глобальной переменной. То есть каждой компоненте локальной переменной соответствует некоторая компонента глобальной переменной, где отображение индексов можно записать как $g = G(i, j)$ (локальной компоненте $(x_i)_j$ соответствует компонента z_g).

Если $G(i, j) = j$ для $\forall i$, то все локальные переменные заменяются глобальными, и мы получим копию. Однако, это не дает пользы для уменьшения размерности ($x_i = z$), поэтому интересен только вариант, при котором $n_i \ll n$. Получим что локальный вектор содержит маленькое количество глобальных компонент z . Глобальная переменная z представляет полный вектор признаков (т.е. вектор параметров модели или независимых переменных в данных).

Различные подмножества данных разбиты между N процессорами. Тогда x_i можно рассматривать как подвектор z , соответствующий ненулевым признакам. И данный вектор обрабатывается на i -ом процессоре.

Переобозначим $\tilde{z}_i \in \mathbb{R}^{n_i}$, где $(\tilde{z}_i)_j = z_{G(i,j)}$. Интуитивно понятно, что \tilde{z}_i это представление о том, какой должна быть локальная переменная x_i . Тогда условие связи (согласования) можно представить в виде $x_i - \tilde{z}_i = 0$, $i = 1, \dots, N$, соответственно в задаче (18).

$$(x_i)_j = z_{G(i,j)}, \quad i = 1, \dots, N, \quad j = 1, \dots, n_i$$

Далее разберем пример на рисунке 2. Слева находятся локальные переменные

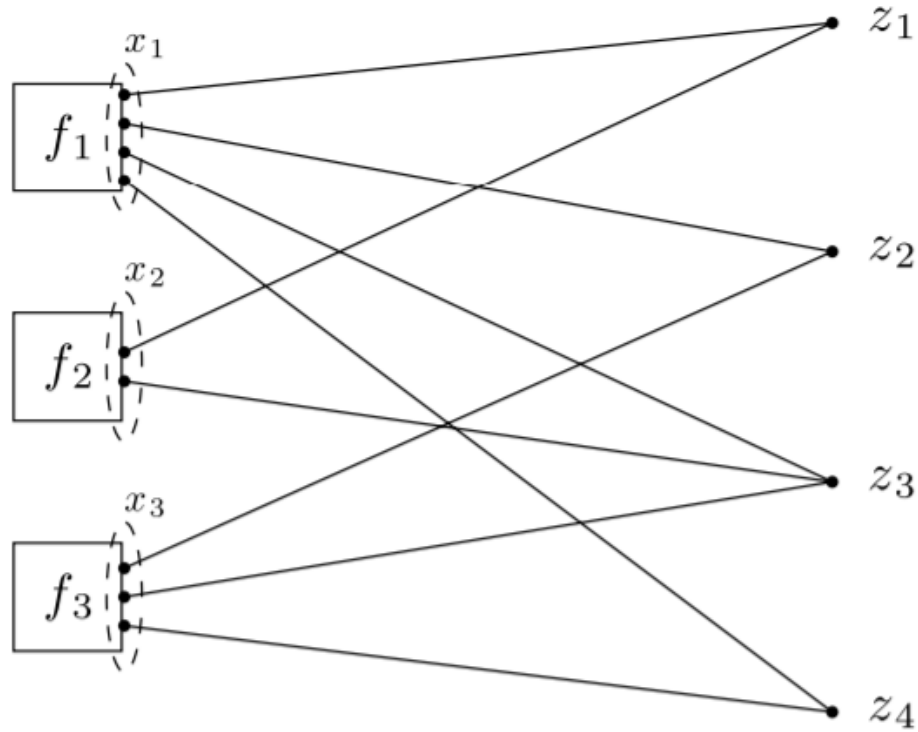


Рис. 2: General form consensus optimization

x_1, x_2, x_3 , а справа глобальная переменная z . $N = 3$ количество подмножеств глобальной переменной. Глобальная переменная размерности $n = 4$ и локальные переменные размерности $n_1 = 4, n_2 = 2, n_3 = 3$. Целевые функции и глобальные переменные образуют двудольный граф, каждое ребро которого является согласо-

ваным ограничением между компонентами локальной переменной и глобальной переменной.

Тогда для общей задачи (18) модифицированная функция Лагранжан будет выглядеть следующим образом

$$L_\rho(x, y, z) = \sum_{i=1}^N (f_i(x_i) + y_i^\top (x_i - \tilde{z}_i) + (\rho/2) \|x_i - \tilde{z}_i\|_2^2), \quad (19)$$

где $y_i \in R^{n_i}$ двойственная переменная. Таким образом мы получим схему итераций ADMM вида:

$$\begin{aligned} x_i(k+1) &:= \operatorname{argmin}_{x_i} (f_i(x_i) + y_i^\top(k) x_i + (\rho/2) \|x_i - \tilde{z}_i(k)\|_2^2) \\ z(k+1) &:= \operatorname{argmin}_z \left(\sum_{i=1}^m (-y_i^\top(k) \tilde{z}_i + (\rho/2) \|x_i(k+1) - \tilde{z}_i\|_2^2) \right) \\ y_i(k+1) &:= y_i(k) + \rho(x_i(k+1) - \tilde{z}_i(k+1)). \end{aligned} \quad (20)$$

Как видно по схеме итераций обновление переменных x_i и y_i можно проводить независимо на всех процессорах, то есть параллельно для всех i .

На шаге обновления z происходит взаимодействие между компонентами, для установления связи. Это возможно благодаря полной разделимости L_ρ по своим компонентам.

$$z_g(k+1) := \frac{\sum_{G(i,j)=g} ((x_i(k+1))_j + (1/\rho)(y_i(k))_j)}{\sum_{G(i,j)=g} 1},$$

таким образом z_g находится путем усреднения всех компонент $x_i(k+1) + (1/\rho)y_i(k)$, которые соответствуют глобальному индексу g . При применении аргумента схожего типа для общей задачи согласованной оптимизации, после первой итерации мы получим:

$$\sum_{G(i,j)=g} (y_i(k))_j = 0,$$

то есть для любого заданного индекса глобальной переменной g , сумма компонент двойственной переменной будет равна нулю. Тогда обновление z можно записать

в более простой форме

$$z_g(k+1) := \frac{1}{k_g} \sum_{G(i,j)=g} (x_i(k+1))_j$$

где k_g количество компонент в локальной переменной, соответствующих глобальной переменной z_g . Другими словами, обновление z является локальным усреднением для каждой компоненты z_g , а не глобальным усреднением. Т.е. это можно сравнить с тем, что для усреднения компоненты z_g будут использоваться только те компоненты процессов, которые имеют представление о признаке z_g . Таким образом для обновления одной компоненты z_g будут учитываться соответствующие значения локальных компонент, непосредственно оценивающих эту компоненту z_g .

5.2 ADMM

В этом разделе мы рассмотрим метод ADMM (метод множителей переменного направления) для решения задачи (17). Классическое описание метода можно посмотреть в [3] и [11]. ADMM решает задачи вида

$$\min f(\mathbf{x}) + g(\mathbf{y}) \tag{21}$$

$$\text{s.t. } A\mathbf{x} + B\mathbf{y} = \mathbf{b}, \tag{22}$$

где $X \subseteq \mathbb{R}^p$ и $Y \subseteq \mathbb{R}^q$ замкнутое выпуклое множество, вектор $\mathbf{b} \in \mathbb{R}^m$ и даны матрицы $A \in \mathbb{R}^{m \times p}$ и $B \in \mathbb{R}^{m \times q}$. Более того, будем предполагать, что решение $(\mathbf{x}^*, \mathbf{y}^*)$ существует вместе с произвольным двойственным множителем $\lambda^* \in \mathbb{R}^m$ связанным с линейным ограничением $A\mathbf{x} + B\mathbf{y} = \mathbf{b}$.

ADMM обновляет как прямые, так и двойственные переменные $(\mathbf{x}_t, \mathbf{y}_t, \lambda_t)$ в ходе алгоритма. Учитывая $(\mathbf{y}_{t-1}, \lambda_{t-1}) \in Y \times \mathbb{R}^m$ и параметр $\rho > 0$, алгоритм

вычисляет $(\mathbf{x}_t, \mathbf{y}_t, \lambda_t)$ по формулам

$$\begin{aligned}\mathbf{x}_t &= \operatorname{argmin}_{\mathbf{x} \in X} f(\mathbf{x}) + \langle \lambda_{t-1}, A\mathbf{x} + B\mathbf{y}_{t-1} - \mathbf{b} \rangle + \frac{\rho}{2} \|A\mathbf{x} + B\mathbf{y}_{t-1} - \mathbf{b}\|^2, \\ \mathbf{y}_t &= \operatorname{argmin}_{\mathbf{y} \in Y} g(\mathbf{y}) + \langle \lambda_{t-1}, A\mathbf{x}_t + B\mathbf{y} - \mathbf{b} \rangle + \frac{\rho}{2} \|A\mathbf{x}_t + B\mathbf{y} - \mathbf{b}\|^2 \\ \lambda_t &= \lambda_{t-1} + \rho(A\mathbf{x}_t + B\mathbf{y}_t - \mathbf{b}).\end{aligned}\tag{23}$$

Будем обозначать $\mathbf{z}_t \equiv (\mathbf{x}_t, \mathbf{y}_t, \lambda_t)$, $\mathbf{z} \equiv (\mathbf{x}, \mathbf{y}, \lambda)$ и Q зазор двойственности

$$\begin{aligned}Q(\mathbf{z}_t, \mathbf{z}) &= f(\mathbf{x}_t) + g(\mathbf{y}_t) + \langle \lambda, A\mathbf{x}_t + B\mathbf{y}_t - \mathbf{b} \rangle \\ &\quad - (f(\mathbf{x}) + g(\mathbf{y}) + \langle \lambda, A\mathbf{x} + B\mathbf{y} - \mathbf{b} \rangle).\end{aligned}\tag{24}$$

Кроме того, обозначим произвольную пару прямо-двойственных решений задачи (18) за $\mathbf{z}^* = (\mathbf{x}^*, \mathbf{y}^*, \lambda^*)$.

Теорема 1. Пусть $\mathbf{z}_t, t = 1, \dots, k$ последовательность сгенерированная алгоритмом ADMM с некоторым $\rho > 0$, определим $\bar{\mathbf{z}}_k = \sum_{t=1}^k \mathbf{z}_t / k$. Тогда получим

$$\begin{aligned}f(\bar{\mathbf{x}}_k) + g(\bar{\mathbf{y}}_k) - f(\mathbf{x}^*) - g(\mathbf{y}^*) &\leq \frac{1}{2k} \left(\frac{1}{\rho} \|\lambda_0\|_2^2 + \|B(\mathbf{y}_0 - \mathbf{y}^*)\|_2^2 \right), \\ \|A\bar{\mathbf{x}}_k + B\bar{\mathbf{y}}_k - \mathbf{b}\|_2 &\leq \frac{1}{k} \left(\frac{2}{\rho} \|\lambda_0 - \lambda^*\|_2 + \|B(\mathbf{y}_0 - \mathbf{y}^*)\|_2 \right).\end{aligned}\tag{25}$$

5.3 Новая формулировка ADMM

В этой части работы мы переформулируем задачу PSSE для применения метода ADMM. Напомним вид полученной ранее задачи PSSE.

$$\begin{aligned}\min_{\{\mathbf{v}_k \in \mathcal{X}_k\}} & \sum_{k=1}^K f_k(\mathbf{v}_k) \\ \text{s.t.} & \quad \mathbf{v}_k[l] = \mathbf{u}_{kl} \quad \forall l \in \mathcal{B}_k, \\ & \quad \mathbf{v}_l[k] = \mathbf{u}_{kl} \quad \forall k \in \mathcal{B}_l,\end{aligned}\tag{26}$$

где \mathcal{B}_k множество шин напрямую соединенных с шиной k , Применим метод ADMM для решения этой задачи. Как было показано в предыдущих двух секциях, ADMM

решает выпуклые задачи следующего вида

$$\begin{aligned} & \min_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} f(\mathbf{x}) + g(\mathbf{y}) \\ & \text{s.t. } A\mathbf{x} + B\mathbf{y} = \mathbf{b}, \end{aligned} \quad (27)$$

для известных матриц и вектора (A, B, b) правильной размерности. При назначении множителя Лагранжа λ для условия связи и полученное решение (\mathbf{x}, \mathbf{z}) было найдено по следующей схеме при некотором $\mu > 0$

$$\begin{aligned} \mathbf{x}^{i+1} &:= \arg \min_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) + \frac{\mu}{2} \|A\mathbf{x} + B\mathbf{z}^i - c + \lambda^i\|_2^2 \\ \mathbf{z}^{i+1} &:= \arg \min_{\mathbf{z} \in \mathcal{Z}} g(\mathbf{z}) + \frac{\mu}{2} \|A\mathbf{x}^{i+1} + B\mathbf{z} - c + \lambda^i\|_2^2 \\ \lambda^{i+1} &:= \lambda^i + A\mathbf{x}^{i+1} + B\mathbf{z}^{i+1} - c. \end{aligned} \quad (28)$$

Для применения метода ADMM к задаче (17) определим переменные $\{\mathbf{v}_k\}$ как \mathbf{x} и $\{\mathbf{u}_{kl}\}$ как \mathbf{z} с $g(\mathbf{z}) = 0$. Введем множители Лагранжа λ_{kl} . Заметим, что λ_{kl} и λ_{lk} соответствуют различным ограничениям $\mathbf{v}_k[l] = \mathbf{u}_{kl}$ и $\mathbf{v}_l[k] = \mathbf{u}_{lk}$ соответственно.

В соответствии с первым уравнением в (28) векторы состояния в каждой области $\{\mathbf{v}_k\}$ могут обновляться локально, т.е.

$$\mathbf{v}_k^{i+1} := \arg \min_{\mathbf{v}_k \in \mathcal{X}_k} f_k(\mathbf{v}_k) + \frac{\mu}{2} \sum_{l \in \mathcal{B}_k} \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^i + \lambda_{k,l}^i\|_2^2. \quad (29)$$

Мы предполагаем, что каждое состояние сети принадлежит не более чем двум областям. Тогда из второго уравнения в (28) следует, что вспомогательные переменные \mathbf{u}_{kl} могут быть найдены по формуле

$$\mathbf{u}_{kl}^{i+1} := \frac{1}{2} (\mathbf{v}_k^{i+1}[l] + \mathbf{v}_l^{i+1}[k] + \lambda_{kl}^i + \lambda_{lk}^i) \quad (30)$$

в то время как два связанных множителя обновляются следующим образом

$$\begin{aligned} \lambda_{kl}^{i+1} &:= \lambda_{kl}^i + (\mathbf{v}_k^{i+1}[l] - \mathbf{u}_{kl}^{i+1}) \\ \lambda_{lk}^{i+1} &:= \lambda_{lk}^i + (\mathbf{v}_l^{i+1}[k] - \mathbf{u}_{kl}^{i+1}). \end{aligned} \quad (31)$$

Сложим уравнения из (31) и объединим его с (30), получим $\lambda_{kl}^{i+1} = -\lambda_{lk}^{i+1}$ на всех итерациях i , если множители были инициализированы в нуле. Следовательно, вспомогательная переменная \mathbf{u}_{kl} в конечном итоге является средним значением значений общих состояний, то есть

$$\mathbf{u}_{kl}^{i+1} = \frac{1}{2} (\mathbf{v}_k^{i+1}[l] + \mathbf{v}_l^{i+1}[k]). \quad (32)$$

Далее переопределим вид нашей задачи, для удобства доказательства теоремы сходимости в нашей задаче. Определим $\mathbf{r} = A\mathbf{x} + B\mathbf{z} - \mathbf{b}$, тогда

$$\begin{aligned} \mathbf{y}^T \mathbf{r} + (\rho/2) \|\mathbf{r}\|_2^2 &= (\rho/2) \|\mathbf{r} + (1/\rho)\mathbf{y}\|_2^2 - (1/2\rho) \|\mathbf{y}\|_2^2 \\ &= (\rho/2) \|\mathbf{r} + \mathbf{u}\|_2^2 - (\rho/2) \|\mathbf{u}\|_2^2, \end{aligned} \quad (33)$$

Тогда можно переписать схему ADMM через $\mathbf{y}_{kl} = \mu\lambda_{kl}$. Получим следующую схему

$$\begin{aligned} \mathbf{v}_k^i &:= \arg \min_{\mathbf{v}_k \in \mathcal{X}_k} f_k(\mathbf{v}_k) + \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^{i-1}, \mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1} \rangle + \frac{\mu}{2} \sum_{l \in \mathcal{B}_k} \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}\|_2^2, \\ \mathbf{u}_{kl}^i &:= \frac{1}{2} \sum_{l \in \mathcal{B}_k} \mathbf{v}_k^i[l] = \frac{1}{2} (\mathbf{v}_k^i[l] + \mathbf{v}_l^i[k]), \\ \mathbf{y}_{kl}^i &:= \mathbf{y}_{kl}^{i-1} + \mu (\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^i). \end{aligned} \quad (34)$$

Тогда зазор двойственности будет иметь вид

$$Q(\mathbf{t}_k^i, \mathbf{t}_k) = f_k(\mathbf{v}_k^i) + \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^i, \mathbf{v}_k^i[l] - \mathbf{u}_{kl}^i \rangle - f_k(\mathbf{v}_k) - \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^i, \mathbf{v}_k[l] - \mathbf{u}_{kl} \rangle, \quad (35)$$

где $\mathbf{t}_k^i \equiv (\mathbf{v}_k^i, \mathbf{u}_{kl}^i, \mathbf{y}_{kl}^i)$ и $\mathbf{t}_k \equiv (\mathbf{v}_k, \mathbf{u}_{kl}, \mathbf{y}_{kl})$. Далее я попробую переписать теорему 1 и доказать ее для основной задачи на основе статьи [16].

Теорема 2. Пусть $\mathbf{v}_k^i, \mathbf{u}_{kl}^i, \mathbf{y}_{kl}^i, i = 1, \dots, N$ это последовательность сгенерированная алгоритмом 1 с $\mu > 0$. Определим $\bar{\mathbf{v}}_k = \sum_{i=1}^N \mathbf{v}_k^i / N$ и схожим образом

$\bar{\mathbf{u}}_{kl}, \bar{\mathbf{y}}_{kl}$. Тогда получим

$$\begin{aligned} f_k(\bar{\mathbf{v}}_k) - f_k(\mathbf{v}_k^*) &\leq \frac{\mu}{2N} \sum_{l \in \mathcal{B}_k} \|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2^2, \\ \|\bar{\mathbf{v}}_k[l] - \bar{\mathbf{u}}_{kl}\|_2 &\leq \frac{1}{N} \left(\frac{2}{\mu} \|\mathbf{y}_{kl}^*\|_2 + \|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2 \right). \end{aligned} \quad (36)$$

Доказательство.

1) В силу условия оптимальности для

$$\mathbf{v}_k^i = \min_{\mathbf{v}_k \in K} f(\mathbf{v}_k) + \sum_{l \in \mathcal{B}_k} \mathbf{y}_{kl}^{i-1} (\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}) + \sum_{l \in \mathcal{B}_k} \frac{\mu}{2} \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}\|_2^2,$$

мы получим

$$\begin{aligned} &f_k(\mathbf{v}_k^i) - f_k(\mathbf{v}_k) + \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^{i-1}, \mathbf{v}_k^i[l] - \mathbf{u}_{kl}^{i-1} \rangle - \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^{i-1}, \mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1} \rangle \\ &+ \frac{\mu}{2} \sum_{l \in \mathcal{B}_k} \|\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 - \frac{\mu}{2} \sum_{l \in \mathcal{B}_k} \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 \\ &\geq f_k(\mathbf{v}_k^i) - f_k(\mathbf{v}_k) + \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^{i-1}, \mathbf{v}_k^i[l] - \mathbf{v}_k[l] \rangle + \frac{\mu}{2} \sum_{l \in \mathcal{B}_k} \langle 2(\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^{i-1}), \mathbf{v}_k^i[l] - \mathbf{v}_k[l] \rangle \\ &= f_k(\mathbf{v}_k^i) - f_k(\mathbf{v}_k) + \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^{i-1} + \mu(\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^{i-1}), \mathbf{v}_k^i[l] - \mathbf{v}_k[l] \rangle \leq 0, \end{aligned} \quad (37)$$

далее учтем, что $\mathbf{y}_{kl}^i = \mathbf{y}_{kl}^{i-1} + \mu(\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^i)$ и получим следующее:

$$f_k(\mathbf{v}_k^i) - f_k(\mathbf{v}_k) \leq - \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^i + \mu(\mathbf{u}_{kl}^i - \mathbf{u}_{kl}^{i-1}), \mathbf{v}_k^i[l] - \mathbf{v}_k[l] \rangle.$$

2) Возьмем полученное соотношение и определение функции зазора (35), получим

$$\begin{aligned}
Q(\mathbf{t}_k^i, \mathbf{t}_k) &\leq - \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^i + \mu(\mathbf{u}_{kl}^i - \mathbf{u}_{kl}^{i-1}), \mathbf{v}_k^i[l] - \mathbf{v}_k[l] \rangle - \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^i, \mathbf{v}_k[l] - \mathbf{v}_{kl} \rangle \\
&+ \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}, \mathbf{v}_k^i[l] - \mathbf{u}_{kl}^i \rangle = \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}^i, \mathbf{u}_{kl}^i - \mathbf{v}_k^i[l] \rangle + \sum_{l \in \mathcal{B}_k} \mu \langle \mathbf{u}_{kl}^{i-1} - \mathbf{u}_{kl}^i, \mathbf{v}_k^i[l] - \mathbf{v}_k[l] \rangle \\
&+ \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}, \mathbf{v}_k^i[l] - \mathbf{u}_{kl}^i \rangle = \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl} - \mathbf{y}_{kl}^i, \frac{1}{\mu}(\mathbf{y}_{kl}^i - \mathbf{y}_{kl}^{i-1}) \rangle \\
&+ \sum_{l \in \mathcal{B}_k} \mu \langle \mathbf{u}_{kl}^{i-1} - \mathbf{u}_{kl}^i, \mathbf{v}_k^i[l] - \mathbf{v}_k[l] \rangle. \tag{38}
\end{aligned}$$

3) Заметим, что

$$\begin{aligned}
2\langle \mathbf{y}_{kl} - \mathbf{y}_{kl}^i, \mathbf{y}_{kl}^i - \mathbf{y}_{kl}^{i-1} \rangle &= \|\mathbf{y}_{kl} - \mathbf{y}_{kl}^{i-1}\|_2^2 - \|\mathbf{y}_{kl} - \mathbf{y}_{kl}^i\|_2^2 - \|\mathbf{y}_{kl}^{i-1} - \mathbf{y}_{kl}^i\|_2^2; \\
2\langle \mathbf{u}_{kl}^{i-1} - \mathbf{u}_{kl}^i, \mathbf{v}_k^i[l] - \mathbf{v}_k[l] \rangle &= \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 - \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^i\|_2^2 \\
&+ \|\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^i\|_2^2 - \|\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 = \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 - \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^i\|_2^2 \\
&+ \frac{1}{\mu^2} \|\mathbf{y}_{kl}^i - \mathbf{y}_{kl}^{i-1}\|_2^2 - \|\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^{i-1}\|_2^2.
\end{aligned}$$

В результате, получим

$$\begin{aligned}
Q(\mathbf{t}_k^i, \mathbf{t}_k) &\leq \frac{1}{2\mu} \sum_{l \in \mathcal{B}_k} (\|\mathbf{y}_{kl} - \mathbf{y}_{kl}^{i-1}\|_2^2 - \|\mathbf{y}_{kl} - \mathbf{y}_{kl}^i\|_2^2) - \frac{1}{2\mu} \sum_{l \in \mathcal{B}_k} (\|\mathbf{y}_{kl}^{i-1} - \mathbf{y}_{kl}^i\|_2^2) \\
&+ \frac{\mu}{2} \sum_{l \in \mathcal{B}_k} (\|\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 - \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^i\|_2^2) \\
&+ \frac{\mu}{2} \sum_{l \in \mathcal{B}_k} \left(\frac{1}{\mu^2} \|\mathbf{y}_{kl}^i - \mathbf{y}_{kl}^{i-1}\|_2^2 - \|\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 \right) \\
&= \frac{1}{2\mu} \sum_{l \in \mathcal{B}_k} (\|\mathbf{y}_{kl} - \mathbf{y}_{kl}^{i-1}\|_2^2 - \|\mathbf{y}_{kl} - \mathbf{y}_{kl}^i\|_2^2) \\
&+ \frac{\mu}{2} \sum_{l \in \mathcal{B}_k} (\|\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 - \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^i\|_2^2 - \|\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^{i-1}\|_2^2).
\end{aligned}$$

4) Полученное неравенство суммируем по всем $i = 1, \dots, N$, тогда

$$\begin{aligned} \sum_{i=1}^N Q(\mathbf{t}_k^i, \mathbf{t}_k) &\leq \frac{1}{2\mu} (\|\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}\|_2^2 - \|\mathbf{y}_{kl}^N - \mathbf{y}_{kl}\|_2^2) \\ &\quad + \frac{\mu}{2} (\|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}\|_2^2 - \|\mathbf{u}_{kl}^N - \mathbf{u}_{kl}\|_2^2) \geq 0. \end{aligned} \quad (39)$$

5) Заменяем \mathbf{t}_k на $\mathbf{t}_k^* = (\mathbf{v}_k^*, \mathbf{u}_{kl}^*, \mathbf{y}_{kl}^*)$ в (39) и учтем, что $Q(\bar{\mathbf{t}}_k, \mathbf{t}_k^*) \geq 0$, получим

$$\|\mathbf{y}_{kl}^N - \mathbf{y}_{kl}^*\|_2^2 \leq \|\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^*\|_2^2 + \mu^2 \|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2^2.$$

Следовательно

$$\begin{aligned} \|\mathbf{y}_{kl}^N - \mathbf{y}_{kl}^0\|_2 &\leq \|\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^*\|_2 + \|\mathbf{y}_{kl}^N - \mathbf{y}_{kl}^*\|_2 \\ &\leq 2\|\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^*\|_2 + \mu\|\mathbf{v}_k^0 - \mathbf{u}_{kl}^*\|_2. \end{aligned} \quad (40)$$

6) Пусть $\mathbf{t}_k = (\mathbf{v}_k^*, \mathbf{u}_{kl}^*, \mathbf{y}_{kl})$ в (39), тогда заметим что

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N Q(\mathbf{t}_k^i, (\mathbf{v}_k^*, \mathbf{u}_{kl}^*, \mathbf{y}_{kl})) &\geq Q(\bar{\mathbf{t}}_k, (\mathbf{v}_k^*, \mathbf{u}_{kl}^*, \mathbf{y}_{kl})) \\ &= f_k(\bar{\mathbf{v}}_k) - f_k(\mathbf{v}_k^*) + \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}, \bar{\mathbf{v}}_k[l] - \bar{\mathbf{u}}_{kl} \rangle, \end{aligned}$$

и

$$\frac{1}{2} (\|\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}\|_2^2 - \|\mathbf{y}_{kl}^N - \mathbf{y}_{kl}\|_2^2) = \frac{1}{2} (\|\mathbf{y}_{kl}^0\|_2^2 - \|\mathbf{y}_{kl}^N\|_2^2) - \langle \mathbf{y}_{kl}, \mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^N \rangle.$$

Следовательно, мы получим

$$\begin{aligned} &f_k(\bar{\mathbf{v}}_k) - f_k(\mathbf{v}_k^*) + \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}, \bar{\mathbf{v}}_k[l] - \bar{\mathbf{u}}_{kl} \rangle \leq \\ &\leq \frac{1}{N} \sum_{l \in \mathcal{B}_k} \frac{1}{\mu} \left[\frac{1}{2} (\|\mathbf{y}_{kl}^0\|_2^2 - \|\mathbf{y}_{kl}^N\|_2^2) - \langle \mathbf{y}_{kl}, \mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^N \rangle \right] \\ &+ \frac{\mu}{2N} \sum_{l \in \mathcal{B}_k} \left[\|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2^2 - \|\mathbf{u}_{kl}^N - \mathbf{u}_{kl}^*\|_2^2 \right], \end{aligned}$$

7) Следовательно,

$$\begin{aligned}
& f_k(\bar{\mathbf{v}}_k) - f_k(\mathbf{v}_k^*) + \sum_{l \in \mathcal{B}_k} \langle \mathbf{y}_{kl}, \bar{\mathbf{v}}_k[l] - \bar{\mathbf{u}}_{kl} + \frac{1}{\mu N} (\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^N) \rangle \\
& \leq \frac{1}{2\mu N} \sum_{l \in \mathcal{B}_k} \left(\|\mathbf{y}_{kl}^0\|_2^2 - \|\mathbf{y}_{kl}^N\|_2^2 \right) + \frac{\mu}{2N} \sum_{l \in \mathcal{B}_k} \left(\|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2^2 - \|\mathbf{u}_{kl}^N - \mathbf{u}_{kl}^*\|_2^2 \right) \\
& \leq \frac{1}{2N} \left(\frac{1}{\mu} \sum_{l \in \mathcal{B}_k} \|\mathbf{y}_{kl}^0\|_2^2 + \mu \sum_{l \in \mathcal{B}_k} \|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2^2 \right) = \frac{\mu}{2N} \sum_{l \in \mathcal{B}_k} \|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2^2,
\end{aligned}$$

для всех \mathbf{u}_{kl} , $l \in \mathcal{B}_k$, $k = 1, \dots, K$. В последнем уравнении, мы учли что при инициализации $\mathbf{y}_{kl}^0 = 0$. Далее, из соотношения следует, $\bar{\mathbf{v}}_k[l] - \bar{\mathbf{v}}_k + \frac{1}{\mu N} (\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^k) = 0$, и первое уравнение в (36) выполнено. Также, из последнего и (40) следует второе уравнение (36).

$$\begin{aligned}
\|\bar{\mathbf{v}}_k[l] - \bar{\mathbf{u}}_{kl}\|_2 & \leq \frac{1}{\mu N} \left[2\|\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^*\|_2 + \mu\|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2 \right] \\
& = \frac{1}{N} \left[\frac{2}{\mu} \|\mathbf{y}_{kl}^0 - \mathbf{y}_{kl}^*\|_2 + \|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2 \right] \\
& = \frac{1}{N} \left[\frac{2}{\mu} \|\mathbf{y}_{kl}^*\|_2 + \|\mathbf{u}_{kl}^0 - \mathbf{u}_{kl}^*\|_2 \right].
\end{aligned}$$

Далее покажем схему итераций метода ADMM. Алгоритм 1 показывает схему классического ADMM применительно к нашей задаче (17). На Шаге 4 алгоритма отдельно и независимо решаются задачи минимизации локальных областей \mathcal{K} . Далее, чтобы получить правильное решение для всей сети, на шаге 5 значения в общих узлах для каждой пары пересекающихся областей пересчитываются заново. На этом шаге учитываются оба значения $\mathbf{v}_k[l]$ и $\mathbf{v}_l[k]$, полученные в этом узле для двух различных областей k и l , и устанавливается среднее значение \mathbf{u}_{kl} . Этот шаг регулирует значения для связывания областей и получения единого сбалансированного решения для всей сети. Кроме того, вычисляется множитель Лагранжа \mathbf{y}_{kl} . Далее мы проверяем критерий завершения на шаге 7, для которого вводим первичный остаток $\mathbf{r}_k^i = \mathbf{v}_k[l]^i - \mathbf{u}_{kl}^i$ для первичной осуществимости и двойственный остаток $\mathbf{s}_k^i = \mathbf{u}_{kl}^i - \mathbf{u}_{kl}^{i-1}$ для двойственной осуществимости и перезапускаем

Algorithm 1 Alternating Direction method of Multipliers

1: **Initialize:**

матрица H , доступные измерения \mathbf{z} , параметр μ , прямая и двойственная точность $\varepsilon_{\text{primal}} = 10^{-3}$, $\varepsilon_{\text{dual}} = 10^{-4}$, начальная точка $\mathbf{u}_{kl}^0 = 0$, $\mathbf{y}_{kl}^0 = 0$ для каждого $k, l \in K$

2: $i \leftarrow 0$

3: **repeat**

4:

$$\begin{aligned} \mathbf{v}_k^i &= \min_{\mathbf{v}_k \in K} f(\mathbf{v}_k) + \sum_{l \in B_k} \mathbf{y}_{kl}^{i-1} (\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}) \\ &\quad + \sum_{l \in B_k} \frac{\mu}{2} \|\mathbf{v}_k[l] - \mathbf{u}_{kl}^{i-1}\|_2^2 \end{aligned}$$

5: $\mathbf{u}_{kl}^i = \frac{1}{2} (\mathbf{v}_k^i[l] + \mathbf{v}_l^i[k])$

6: $\mathbf{y}_{kl}^i = \mathbf{y}_{kl}^{i-1} + \mu (\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^i)$

7: **until** $\|\mathbf{r}_k^i\|_2^2 < \varepsilon_{\text{primal}}$ И $\|\mathbf{s}_k^i\|_2^2 < \varepsilon_{\text{dual}}$

8: $i \leftarrow i + 1$

НОВЫЙ ЦИКЛ.

6 Ускоренные методы

Мы предлагаем применить для нашей задачи ускоренный вариант ADMM из [9]. Этот метод также описан в [5], [10] и других. Этот метод использует идею ускоренного метода из статьи [20] (см. [14] для ускоренных переменных процедур), в которой представлена схема минимизации первого порядка с глобальной сходимостью $O(1/N^2)$ для класса гладких целевых функций.

Ускоренный метод Нестерова - это модификация градиентного спуска, который ускоряется ступеню оверрелаксации.

В работе [9] предположим, что обе функции из общей проблемы (27) сильно выпуклые. Но в нашем случае $g(\mathbf{y})$ является нулевой функцией, поэтому не сильно выпуклой. Поэтому в теории метод был применен для рассматриваемой задачи не может дать никакого ускорения, как было написано в Теореме 2 [9]. Однако

при правильном подборе параметров метод может быть значительно ускорен. Таким образом, мы получаем эффективное решение проблемы PSSE. Заметим, что на практике мы наблюдали для обоих методов ADMM и ускоренного ADMM линейную скорость сходимости! Это гораздо лучше, чем можно объяснить наиболее известными теоретическими результатами для таких методов [27, 16]. Эти практические аспекты являются одной из основных причин использования подхода ADMM и его ускоренных вариантов вместо не дополненных и не чередующихся процедур градиентного типа и ускоренных вариантов.

6.1 Ускоренный градиентный метод

В этой части работы мы рассмотрим ускоренный градиентный метод, который является одним из базовых ускоренных методов оптимизации. Этот метод является одним из широко используемых методов первого порядка, обладающий теоретической обоснованностью, является простым в реализации и эффективным на практике.

Базовая задача оптимизации:

$$\min_{\mathbf{x}} f(\mathbf{x}), \quad (41)$$

где функция $f(\mathbf{x})$ является выпуклой.

Ускоренный градиентный спуск основан на обычном методе градиентного спуска, который имеет следующую схему:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{1}{L} \nabla f(\mathbf{x}_k). \quad (42)$$

Ускоренный метод добавляет шаг экстраполяции для получения точки \mathbf{y}_k , после чего выполняется шаг градиентного спуска. Данный новый шаг оверрелоксации помогает уменьшить колебания градиентного спуска около локального минимума, поэтому метод ускоряется из-за правильного направления спуска.

Algorithm 2 Ускоренный градиентный спуск

1: **Initialize:**

$$\mathbf{x}_0 = \mathbf{x}_{-1}$$

2: **for** $k = 0, 1, \dots$ **do**

3: $\mathbf{y}_k = \mathbf{x}_k + \beta(\mathbf{x}_k - \mathbf{x}_{k-1}),$

4: $\mathbf{x}_{k+1} = \mathbf{y}_k - \frac{1}{L}\nabla f(\mathbf{y}_k).$

Градиентный спуск (42) имеет скорость сходимости $O(1/k)$ для выпуклых задач и $O((1 - \frac{\mu}{L})^k)$ для μ -сильно выпуклых задач. Для сравнения, ускоренный метод улучшает скорость сходимости до $O(1/k^2)$ и $O((1 - \sqrt{\frac{\mu}{L}})^k)$ соответственно.

6.2 Ускоренный ADMM

В этой главе мы рассмотрим модификацию метода ADMM, которая называется ускоренным методом ADMM. Обзор по многим ускоренным методам оптимизации можно найти в книге [26].

Как видно по схеме ускоренного алгоритма, были добавлены шаги оверрелаксации для снижения колебаний при сходимости методов.

Algorithm 3 Accelerated ADMM

1: **Initialize:**

matrix H , system state vector \mathbf{z} , parameter ρ , tolerances

$\varepsilon_{primal} = 10^{-3}$, $\varepsilon_{dual} = 10^{-4}$, initial points

$\mathbf{u}_{kl}^0 = \mathbf{0}$, $\hat{\mathbf{u}}_{kl}^0 = \mathbf{0}$, $\mathbf{y}_{kl}^0 = \mathbf{0}$, $\hat{\mathbf{y}}_{kl}^0 = \mathbf{0}$ for each $k, l \in K$

2: $i \leftarrow 0$

3: **repeat**

4:

$$\begin{aligned} \mathbf{v}_k^i &= \min_{\mathbf{v}_k \in K} f(\mathbf{v}_k) + \sum_{l \in B_k} \hat{\mathbf{y}}_{kl}^{i-1} (\mathbf{v}_k[l] - \hat{\mathbf{u}}_{kl}^{i-1}) \\ &\quad + \sum_{l \in B_k} \frac{\rho}{2} \|\mathbf{v}_k[l] - \hat{\mathbf{u}}_{kl}^{i-1}\|_2^2 \end{aligned}$$

5: $\mathbf{u}_{kl}^i = \frac{1}{2} (\mathbf{v}_k^i[l] + \mathbf{v}_l^i[k])$

6: $\mathbf{y}_{kl}^i = \hat{\mathbf{y}}_{kl}^{i-1} + \rho(\mathbf{v}_k^i[l] - \mathbf{u}_{kl}^i)$

7: $\alpha_i = \frac{1 + \sqrt{1 + \alpha_{i-1}^2}}{2}$;

8: $\hat{\mathbf{u}}_{kl}^i = \mathbf{u}_{kl}^i + \frac{\alpha_{i-1} - 1}{\alpha_i} (\mathbf{u}_{kl}^i - \mathbf{u}_{kl}^{i-1})$;

9: $\hat{\mathbf{y}}_{kl}^i = \mathbf{y}_{kl}^i + \frac{\alpha_{i-1} - 1}{\alpha_i} (\mathbf{y}_{kl}^i - \mathbf{y}_{kl}^{i-1})$,

10: **until** $\|\mathbf{r}^i\|_2^2 < \varepsilon_{primal}$ and $\|\mathbf{s}^i\|_2^2 < \varepsilon_{dual}$

11: $i \leftarrow i + 1 = 0$

7 Эксперименты

В этом разделе опишем проведенные эксперименты для энергетической системы.

Мы используем данные из базы данных MATPOWER для энергетической системы из 14-шин. Весь код написан на Python 3 с использованием библиотеки CVXPY и Pypower, для работы с энергетическими системами.

Для составления первой системы проведем следующий ряд шагов:

1. Установим 22 сенсора на узлы и линии электропередач на системе из 14-шин;
2. Возьмем 10 одинаковых копий системы, полученной на шаге 1;
3. Соединим последовательно каждую копию друг с другом по цепочке, чтобы получить систему из $14 * 10 = 140$ шин соединенных между собой последовательно;
4. Добавим соединяющую линию между первой копией и последней копией 14-басной системы, чтобы получить кольцевую топологию сети;
5. Делим каждый 14-шинный блок на 4 области, согласно правилам описанным в предыдущей главе.

Таким образом мы получим систему из 140 шин разделенную на 40 областей, на которой будут проверяться распределенные методы оптимизации из предыдущей главы. На рисунке 3 показана схема соединений и вид 14-шинной системы.

Данный синтетический пример имеет право на существование, так как является близким по своей структуре к реальной жизни. В реальном мире энергетические системы представляют из себя сложную структуру взаимодействий между заводом по производству электроэнергии и потребителями (другими заводами потребителями или городскими районами). Главным фактором такой системы является ее разреженная топология, которая представляет из себя основную

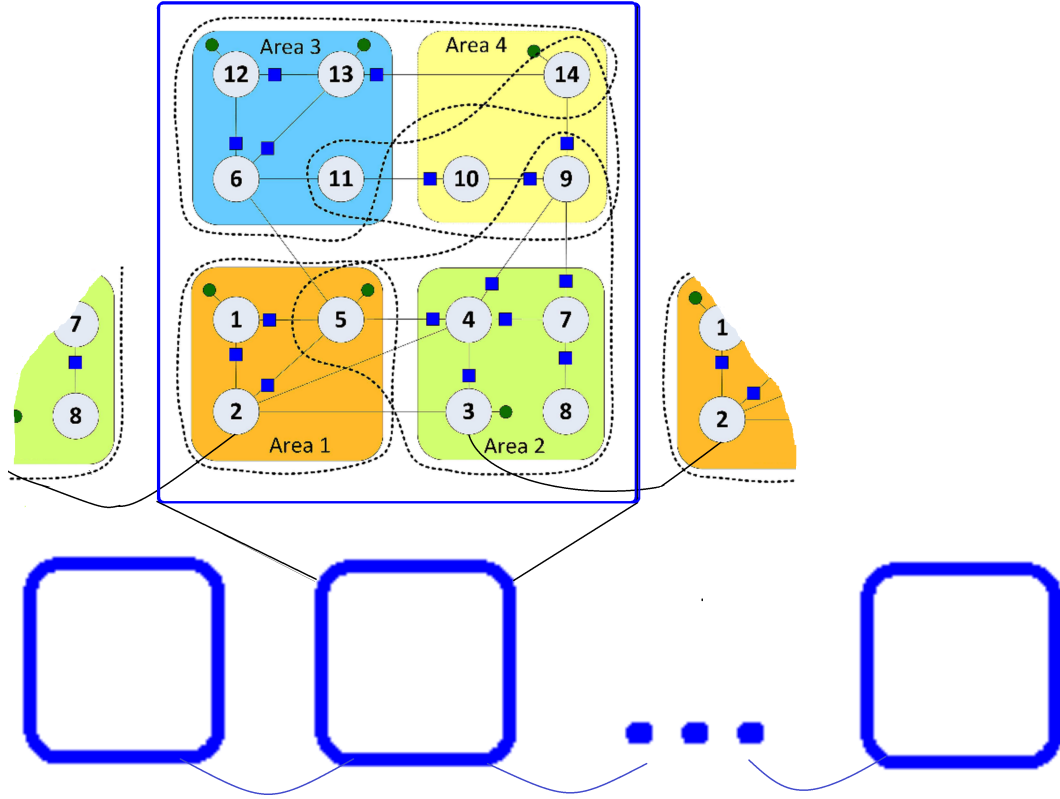


Рис. 3: General form consensus optimization

проблему для распределенных методов решения задач оптимизации на энергетических сетях. Таким образом этот пример будет явно показывать эффективность распределенных методов на энергосети с заданной структурой соединений.

7.1 Пример 1

Будем запускать методы со следующими заданными параметрами: $\mu = 1$, $\varepsilon_{dual} = 10^{-4}$, $\varepsilon_{primal} = 10^{-3}$, $\rho = 2$, где ε_{dual} это заданная двойственная точность для проверки условия $\mathbf{u}_{kl}^i = \mathbf{u}_{kl}^{i+1}$ и ε_{primal} это заданная прямая точность для осуществимости условия $\mathbf{v}_k[l] = \mathbf{u}_{kl}$, которые проверяются на всех итерациях.

Как видно на рисунке 4, ADMM и A-ADMM оба метода достигли заданной точности по прямым и двойственным переменным. Таким образом оба метода являются хорошими способами для решения задачи восстановления состояний на энергетической сети. Как видно из графиков, ускоренный метод лучше стандарт-

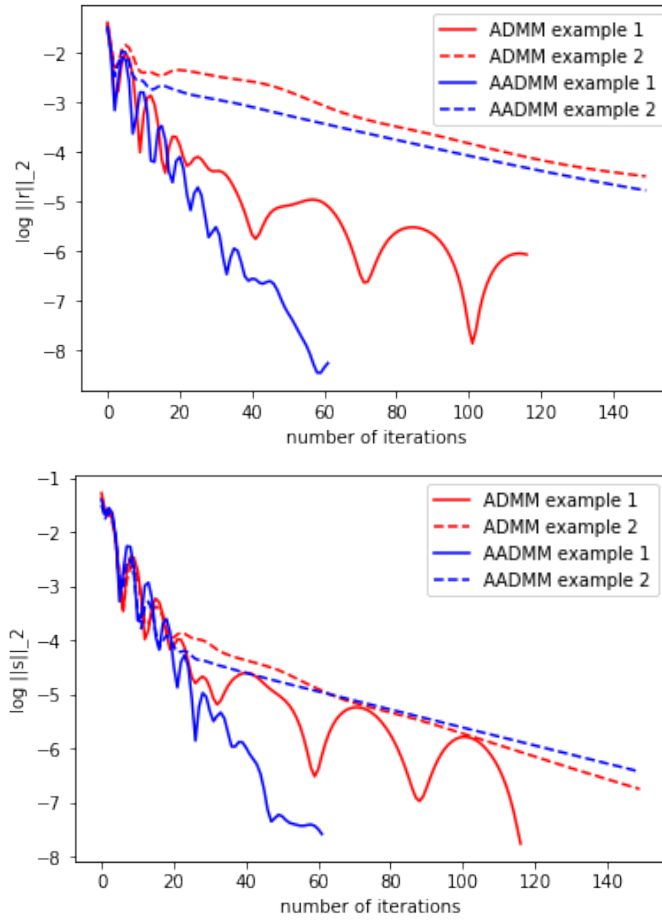


Рис. 4: ходимость по прямым и двойственным переменным в примерах 1 и 2

ного ADMM и даже на такой небольшой схеме метод показал себя значительно быстрее по количеству итераций.

7.2 Пример 2

В этом примере будем показана чувствительность методов к топологии энергетической сети. Для этого изменим схему сети из первого примера, а именно, удалим ребро соединяющее первую и последнюю группу по 14-басов. Этот шаг сделают систему в два раза более разреженной, что является плохим фактором для распределенных методов оптимизации. Будем запускать методы с теми же параметрами, что и в примере 1.

Как видно по рисунку 4 для второго эксперимента, оба метода резко ухудшили свою сходимость, что доказывает чувствительность методов к степени разреженности сети. Но ускоренный метод все так же превосходит по скорости сходимости классический метод ADMM.

7.3 Пример 3

В этом примере мы исследуем чувствительность ускоренного метода к изменению параметра ρ . Для этого мы используем схему сети из первого примера и будем изменять данный параметр. Как видно на рисунке 5, ускоренный метод

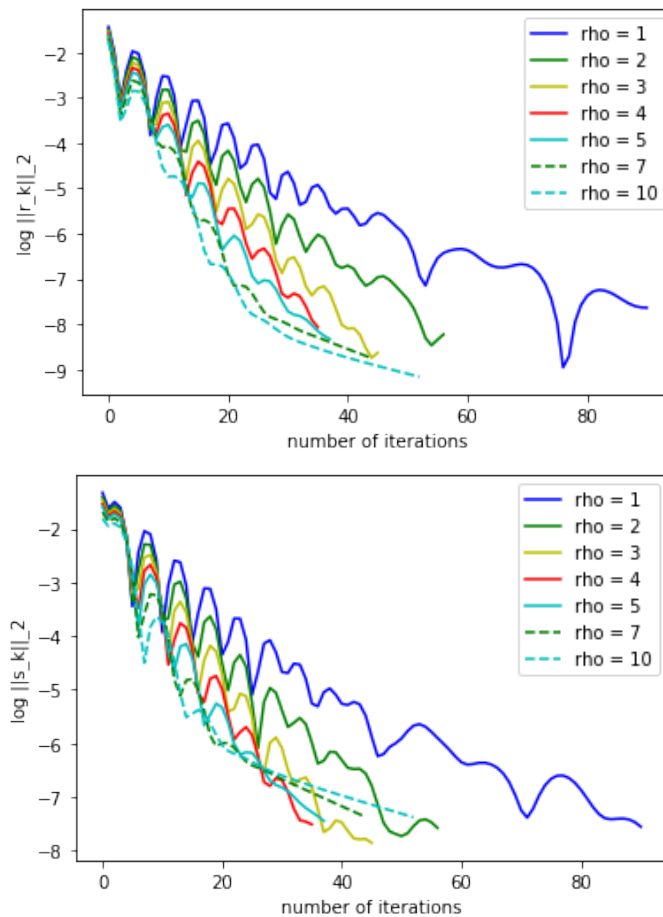


Рис. 5: Сходимость по прямым и двойственным переменным в примере 3

и правда чувствителен к изменению параметра ρ , однако, по достижению $\rho = 7$ метод перестает улучшаться.

8 Выводы

В данной работе были представлены современные методы распределенной оптимизации для решения задачи оценки состояний энергетической системы. Была доказана теорема для скорости сходимости метода ADMM для задачи оценки состояния PSSE, которая подтвердилась практическими примерами.

Таким образом мы получили, что распределенные методы ADMM и A-ADMM являются хорошими оптимизационными методами для решения энергетических задач. Оба метода сошлись до заданной точности на систематической системе из 140-шин с разреженной топологией соединений в сети, что является наилучшим типом сети для распределенных методов оптимизации. Так же была исследована чувствительность методов к изменению степени разреженности сети, которая подтвердила заданные предположения. И был проведен эксперимент по зависимости скорости сходимости ускоренного метода от изменения параметра ρ , что привело к выводу, что метод является улучшаемым при правильном подборе параметра. То есть при увеличении параметра ρ метод увеличивает вес условия связи для областей сети, тем самым увеличивая степень чувствительности к топологии соединений областей, что привело к улучшению сходимости метода.

В дальнейших работах над этой темой, планируется усложнить и расширить задачу на нелинейный (AC) случай энергетической сети. Так же исследовать работу других распределенных методов и получить численные эксперименты на реальных данных и доказать теоретические оценки на сходимость рассматриваемых методов.

Список литературы

- [1] A. Minot Y. M. Lu, Li N. A distributed Gauss-Newton method for power system state estimation // IEEE Trans. Power Syst. — 2016. — Vol. 31. — P. 3804–3815.
- [2] Benders J. F. Partitioning procedures for solving mixed-variables programming problems. — 2005. — Vol. 2(1), no. 1. — P. 3–19. — Computational Management Science.
- [3] Boyd Stephen. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. — 2010. — Vol. 3(1). — P. 1–122. — Foundations and Trends® in Machine Learning.
- [4] D. M. Falcao F. F. Wu, Murphy L. Parallel and distributed state estimation // IEEE Trans. Power Syst. — 1995. — Vol. 10. — P. 724–730.
- [5] Distributed Synthesis Using Accelerated ADMM / M. Ahmadi, M. Cubuktepe, U. Topcu, T. Tanaka // Annual American Control Conference (ACC). — 2018.
- [6] Eckstein Jonathan, Bertsekas Dimitri P. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. — 1992. — Technical report.
- [7] Erseghe Tomaso. Distributed optimal power flow using ADMM. — 2014. — Vol. 29(5). — P. 2370–2380. — IEEE Transactions on Power Systems.
- [8] F. C. Schweppe J. Wildes, Rom D. Power system static state estimation: Parts I, II, and III // IEEE Trans. Power App. Syst. — 1970. — Vol. 80. — P. 120–135.
- [9] Fast alternating direction optimization methods / T. Goldstein, B. ODonoghue, S. Setzer, R Baraniuk. — 2014. — Vol. 7. — P. 1588–1623. — SIAM Journal on Imaging Sciences.
- [10] França G., Robinson D. P., Vidal R. ADMM and Accelerated ADMM as Continuous Dynamical Systems // arXiv preprint. — 2018.

- [11] G. B. Giannakis Q. Ling G. Mateos I. D. Schizas, Zhu H. Decentralized learning for wireless communications and networking // Splitting Methods in Communication, Imaging, Science, and Engineering,. — 2016. — P. 461–497.
- [12] G. B. Giannakis V. Kekatos N. Gatsis S.-J. Kim H. Zhu, Wollenberg B. Monitoring and optimization for power grids: A signal processing perspective // IEEE Signal Process. — 2013. — Vol. 30. — P. 107–128.
- [13] Gabay Daniel, Mercier Bertrand. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. — 1976. — Vol. 2(1). — P. 17–40. — Computers and Mathematics with Applications.
- [14] Guminov S., Dvurechensky P., Gasnikov A. Accelerated Alternating Minimization // arXiv preprint arXiv:1906.03622. — 2019.
- [15] L. Xie D.-H. Choi, Kar S. Cooperative distributed state estimation: Local observability relaxed // Proc. IEEE Power Energy Society General Meeting. — 2011.
- [16] Lan G. Lectures on Optimization Methods for Machine Learning. — Georgia Institute of Technology, Atlanta, GA : H. Milton Stewart School of Industrial and Systems Engineering, 2019.
- [17] Lin S.-Y., Lin C.-H. An implementable distributed state estimator and distributed bad data processing schemes for electric power systems // IEEE Trans. Power Syst. — 1994. — Vol. 9. — P. 1277–1284.
- [18] Lions P. L., Mercier B. Splitting Algorithms for the Sum of Two Nonlinear Operators. — 2005. — Vol. 16(6). — P. 964–979. — SIAM Journal on Numerical Analysis.
- [19] Monticelli A. Electric power system state estimation // Proceedings of the IEEE. — 2000. — Vol. 88. — P. 262 – 282.
- [20] Nesterov Y. A method of solving a convex programming problem with convergence rate $O(1/k^2)$ // Soviet Mathematics Doklady. — 1983. — Vol. 27. — P. 372–376.

- [21] PSSE Redux: Convex Relaxation, Decentralized, Robust, and Dynamic Approaches / V. Kekatos, G. Wang, H. Zhu, G. B. Giannakis // arXiv: 1708.03981. — 2017. — P. 16 – 20.
- [22] Phadke A. G., Thorp J. S. Synchronized Phasor Measurements and Their Applications // NY: Springer. — 2008.
- [23] Ray D. Zimmerman Carlos E. Murillo-Sanchez. Matpower User’s Manual. — 2019.
- [24] Roald L., Andersson G. Chance-Constrained. AC Optimal Power Flow: Reformulations and Efficient Algorithms. — 2017. — Vol. 33. — P. 2906 – 2918. — IEEE Transactions on Power Systems.
- [25] S. Iwamoto M. Kusano, Quintana V. H. Hierarchical state estimation using a fast rectangular-coordinate method // IEEE Trans.Power Syst. — 1989. — Vol. 4. — P. 870–880.
- [26] Zhouchen Lin Huan Li Cong Fang. Accelerated Optimization for Machine Learning First-Order Algorithms // Springer Nature Singapore Pte Ltd. — 2020.
- [27] An accelerated linearized alternating direction method of multipliers / Y. Ouyang, Y. Chen, G. Lan, E. Pasiliao // SIAM Journal on Imaging Sciences. — 2015. — Vol. 8, no. 1. — P. 644–681.