

Федеральное государственное автономное образовательное учреждение  
высшего образования «Национальный исследовательский университет  
«Высшая школа экономики»

На правах рукописи

Алтухов Дмитрий Игоревич

**ОПТИМИЗАЦИОННЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ  
ФУНКЦИОНАЛЬНЫХ СЕТЕЙ ГОЛОВНОГО МОЗГА  
ПО МАГНИТОЭНЦЕФАЛОГРАФИЧЕСКИМ  
ДАННЫМ.**

ДИССЕРТАЦИЯ  
на соискание учёной степени  
кандидата компьютерных наук

Научный руководитель:  
PhD, профессор  
Осадчий Алексей Евгеньевич

Москва — 2020

## Оглавление

	Стр.
<b>Введение . . . . .</b>	<b>5</b>
<b>Глава 1. Взаимодействие нейронных ансамблей и методы его изучения в электрофизиологии. . . . .</b>	<b>18</b>
1.1 Биологические механизмы обмена информацией между нейронными популяциями. . . . .	19
1.1.1 Осцилляции и функциональная интеграция. . . . .	21
1.1.2 Взаимодействие через когерентность (гипотеза СТС). . . . .	22
1.1.3 Формальное определение когерентности . . . . .	24
1.2 Модель неинвазивных МЭГ/ЭЭГ измерений. . . . .	29
1.2.1 Уравнения Maxwella в квазистатическом приближении. .	29
1.2.2 Первичные и вторичные токи. . . . .	31
1.2.3 Модель токового диполя. . . . .	32
1.2.4 Поиск решений для кусочно постоянного профиля проводимости. . . . .	34
1.2.5 Случай сферически симметричного проводника. . . . .	36
1.2.6 Матрица прямой модели и поля чувствительности для ЭЭГ и МЭГ . . . . .	38
1.2.7 Шум сенсоров и шум мозга. Вызванная и индуцированная активность. . . . .	41
1.3 Методы оценки активации нейронных популяций. . . . .	44
1.3.1 Сканирующие алгоритмы . . . . .	47
1.3.2 Алгоритмы пространственной фильтрации . . . . .	50
1.3.3 Методы глобальной оптимизации (методы минимальной нормы) . . . . .	58
1.4 Использование методов решения обратной задачи для оценки фазовой синхронности. . . . .	69
1.4.1 DICS . . . . .	70
1.4.2 iDICS . . . . .	71
1.4.3 GCS . . . . .	72

<b>Глава 2. Метод оценки синхронизации с произвольными фазовыми задержками. . . . .</b>	<b>74</b>
2.1 Протечка сигнала и ее удаление ортогональным проецированием. . . . .	74
2.1.1 Порождающая модель матрицы кросс-спектральной плотности и метод ImCoh . . . . .	76
2.1.2 Оператор проецирования от протечки сигнала. . . . .	78
2.1.3 Произведение Кронекера и векторизация . . . . .	78
2.1.4 Построение оператора проецирования . . . . .	81
2.1.5 Модели со свободной ориентацией диполя . . . . .	84
2.2 PSIICOS . . . . .	90
2.2.1 PSIICOS как метод оптимальной фильтрации. . . . .	91
2.2.2 Пространственная смешенность оценки и нормализация весов. . . . .	95
2.2.3 Действие фильтра на мнимую и действительную часть кросс-спектра. PSIICOS Unbiased. . . . .	96
2.2.4 Регуляризация. . . . .	98
2.2.5 Нормализация кросс-спектральных коэффициентов в пространстве источников. PSIICOS Normalized. . . . .	99
2.3 GO-PSIICOS . . . . .	101
2.3.1 Оценка методом смешанной нормы. . . . .	101
2.3.2 Оптимизация функционала с нормой $L_{21}$ . . . . .	102
2.3.3 Оптимизация методом ISTA. . . . .	104
2.3.4 Контроль сходимости схемы ISTA. . . . .	108
2.3.5 Ускорение схемы ISTA . . . . .	113
2.3.6 Метод смешанной нормы с итеративным перевзвешиванием. . . . .	116
2.3.7 Схема метода GO-PSIICOS . . . . .	118
<b>Глава 3. Анализ свойств метода оценки синхронизации с нулевым фазовым сдвигом. . . . .</b>	<b>119</b>
3.1 Исследование воспроизводимости решения для метода PSIICOS. .	119
3.2 Симуляции методом Монте-Карло. . . . .	120
3.3 Симуляции с тремя сетями, перекрывающимися во времени. . . .	123
3.4 Сравнение на реальных данных. . . . .	130

3.5	Действие проекции на компоненты кросс-спектра в пространстве источников. . . . .	135
3.6	Выбор ранга проекции . . . . .	136
3.7	Влияние неточностей прямой модели на качество решений . . . . .	139
3.8	Сравнение стратегий нормализации весов фильтра для метода PSIICOS. . . . .	142
3.9	Численное исследование метода нормализации кросс-спектральных коэффициентов в пространстве источников (PSIICOS Normalized). . . . .	145
3.10	Демонстрация работы метода GO-PSIICOS на симуляционных данных. . . . .	148
<b>Заключение . . . . .</b>		150
<b>Список сокращений и условных обозначений . . . . .</b>		152
<b>Список литературы . . . . .</b>		153
<b>Список рисунков . . . . .</b>		162
<b>Приложение А. Реализация алгоритма RAP-MUSIC на языке Python . . . . .</b>		165
<b>Приложение Б. Распределение индуцированной активности по коре в реальных данных . . . . .</b>		166

## Введение

Начиная с зарождения первых цивилизаций и до сегодняшнего момента вопросы о том, как устроены сознание и разум в человеке и других живых существах, продолжают быть для нас ключевыми. Так сложилось, что их осмысление происходило сперва в рамках религиозной, а затем и философской парадигмы, которые в западноевропейской традиции значительно перекликались и дополняли друг друга.

Научный подход к вопросам познания сформировался лишь относительно недавно в рамках группы дисциплин, включающей в себя нейробиологию, нейрофизиологию, когнитивную психологию и электрофизиологию. В рамках этих дисциплин ключевым для понимания когнитивной функции человека и животных в самом широком смысле становится устройство центральной нервной системы и, в частности, головного мозга.

Интересно, что механизмы познания связывались с функцией головного мозга не всегда даже в рамках материалистического описания. Так например, Аристотель считал источником мысли сердце, а мозгу отводил лишь роль радиатора, охлаждающего кровь. Тем не менее, уже в эпоху классической античности Гален сформировал идею о том, что именно мозг является источником мысли, а значит и тем инструментом, с помощью которого реализуется познание. На укоренение идеи о том, что изучение работы центральной нервной системы и её высшего отдела — коры больших полушарий — способно предложить ответ на фундаментальный вопрос “что представляет собой человеческий интеллект”, ушло еще более полутора тысяч лет. На сегодняшний день удовлетворительного ответа на этот вопрос по-прежнему нет и появится он, вероятно, не скоро. Однако в ходе долгого и непростого движения к этой Ultima Thule наше понимание более прикладных вещей, вращающихся около нейрофизиологии, несравненно обогатилось. С практической точки зрения трудно переоценить значение понимания работы ЦНС для медицины, не ограничиваясь, однако, лишь ею.

Сегодня, вместе со всеобщим размыванием междисциплинарных границ, нейронауки все больше оказываются связанными с более техническими и инженерно-прикладными дисциплинами. Так, в 1943 году вдохновленные архитектурой нейронных ансамблей живого мозга Маккаллок и Питтс создают

первую вычислительную модель нейронной сети, породив тем самым столь популярный сегодня класс алгоритмов машинного обучения [1]. Все большую популярность приобретают сегодня мозг-компьютерные интерфейсы, позволяющие формировать управляющую команду на основе электромагнитной активности мозга напрямую, что открывает совершенно новые перспективы для интеграции человека с машиной.

В этой связи развитие методов, связанных с изучением строения и работы мозга а также декодирование порождаемых им сигналов представляет сегодня чрезвычайный интерес. Вместе с тем, за последние сто лет благодаря резкому скачку в развитии электроники, физики и компьютерных наук набор инструментов в руках ученого-нейрофизиолога существенно обогатился. На сегодняшний день существующие методы с точки зрения необходимости хирургического вмешательства для проведения измерений можно разделить на инвазивные и неинвазивные.

К первой группе относится интракраниальная энцефалография — метод, в котором электрические потенциалы записываются напрямую с коры больших полушарий. Недостатки и преимущества такого подхода очевидны. К первым прежде всего относится необходимость хирургического вмешательства для проведения измерений, что существенно ограничивает возможности исследователя-нейрофизиолога в получении данных для исследования. На практике осуществление таких измерений на человеке возможно лишь для пациентов, прошедших операцию на мозге в связи с каким-либо неврологическим заболеванием, как правило эпилепсией. В ходе операции для мониторинга активности мозга после хирургического вмешательства на кору головного мозга устанавливаются электроды, регистрирующие электрическую активность. Ясно, что количество таких данных, как и возможность проведения каких-либо сложных когнитивных экспериментов на пациентах, прошедших операцию на мозге, весьма ограничены. При этом качество электрического сигнала, записанного в непосредственной близости от его источника, несравненно выше того, что можно получить, записывая электроэнцефалограмму с поверхности кожи головы.

Неинвазивные методы, с другой стороны, представляют собой намного более гибкий инструмент для исследований головного мозга человека в силу отсутствия необходимости проведения операции. Для изучения анатомической организации мозга а также в качестве вспомогательного инструмента при анализе активности нейронных популяций коры используются методы

структурной нейровизуализации, такие как магнитно-резонансная томография (МРТ), компьютерная томография (КТ) и диффузионная тензорная визуализация (ДТВ). Они позволяют неинвазивно получать статические трехмерные изображения тканей головного мозга. Для изучения динамической активности нейронов используются функциональные методы нейровизуализации, а именно — функциональная магнитно-резонансная томография (фМРТ), позитронно-эмиссионная томография (ПЭТ), электроэнцефалография (ЭЭГ), а также магнитная энцефалография (МЭГ).

При этом лишь последние два метода измеряют электрическую активность мозга непосредственно, тогда как фМРТ и ПЭТ меряют локальный кровоток, который меняется сравнительно медленно, существенно ограничивая временное разрешение этих методов. Так, для ЭЭГ и МЭГ временное разрешение оказывается равным  $\approx 1\text{мс}$ , а методы, измеряющие локальный кровоток, позволяют разрешить лишь процессы с характерными временами порядка одной секунды и медленнее. Вместе с тем, осцилляторные электрофизиологические процессы, порождаемые тканями головного мозга, имеют характерные времена от 0.1 секунды и быстрее [2]. Таким образом, среди всех имеющихся на сегодняшний день инструментов анализа, только ЭЭГ и МЭГ позволяют осуществлять неинвазивные записи сравнительно быстрой электрофизиологической активности головного мозга, что делает их незаменимым инструментом при изучении осцилляций их синхронизации в головном мозге человека.

Способность порождать осцилляции или ритмическую токовую активность является существенной чертой, присущей работе нейронных популяций. Природа возникающих ритмов, а также их функциональное назначение на сегодняшний день остаются предметом изучения, и нет единой, принятой всеми точки зрения на этот счет. Однако, широко принимается гипотеза, согласно которой осцилляции, порождаемые различными нейронными популяциями, служат механизмом, позволяющим различным функционально-специфичным областям мозга избирательно осуществлять обмен информацией друг с другом. Иными словами, предполагается, что осцилляции ответственны за процессы *функциональной интеграции*.

Согласно существующим представлениям, функциональная интеграция нейронных ансамблей осуществляется за счет синхронизации порождаемых этими ансамблями осцилляций. При этом области коры, в которых ритмическая

активность синхронизована, получают возможность эффективнее передавать информацию, а десинхронизованные области, напротив, перестают обмениваться сигналами. Такое представления об организации эффективных каналов передачи информации между нейронными ансамблями за счет синхронизации получило в литературе название «взаимодействие через когерентность» (в английском варианте *communication through coherence*, СТС) [3]. Иными словами, синхронизация осцилляций являются тем механизмом, который позволяет динамически связывать в сети функционально специфичные области мозга для выполнения определенной когнитивной задачи. Изучение таких сетей, возникающих и распадающихся в процессе решения мозгом определенных когнитивных задач, является сегодня одной из центральных тем в изучении мозговой активности, как в норме, так и при патологии [4–7], [8; 9]

С точки зрения исследования таких сетей выделяют понятие *функциональной коннективности*, понимая под этим статистические закономерности в одновременной активации (в самом широком смысле) различных областей мозга. При этом вывод о том, что эти области мозга работали синхронно, делается на основании вычисления определенной метрики, отражающей степень сходства измеренных (или математически восстановленных) в этих областях сигналов. Такие метрики называются *мерами коннективности*.

Многое в области изучения функциональной коннективности было сделано с использованием технологии фМРТ, однако отмеченное выше ограничение фМРТ в виде плохого временного разрешения делает электрофизиологические методы измерений незаменимыми при анализе коннективности. Особое место при этом занимает магнитная энцефалография, которая в сочетании с методами восстановления сигнала на коре головного мозга в силу более высокой точности прямой модели по сравнению с ЭЭГ предоставляет в руки исследователя уникальное сочетание менее чем сантиметрового разрешения по пространству и миллисекундного разрешения по времени [10–12].

Вообще, оценка коннективности на основании неинвазивных электрофизиологических данных, представляет собой сложную инженерную задачу, на решение которой научное сообщество уже потратило немало сил. За последние несколько десятилетий было разработано и опробовано множество методов оценки функциональной коннективности от стандартных подходов, включающих меры синхронизации сигналов во временной и частотной области (таких как корреляция и когеренция), до более изощренных, зачастую нелинейных

мер коннективности [13–31]. Ни одна из предложенных мер, обладая своими достоинствами и недостатками, не является, однако, универсальной в силу сохраняющихся технических затруднений [32; 33].

Одной из наиболее существенных проблем, возникающих при оценке функциональной коннективности является так называемая *протечка сигнала*, объясняемая тем, что обратная задача для ЭЭГ/МЭГ является некорректной. Практически это означает, что имея ограниченный набор измерений нельзя однозначно восстановить конфигурацию источников, породивших сигнал. Из этого, в свою очередь, следует невозможность полностью размешать сигналы, записанные сенсорами, — в каждый из восстановленных сигналов неизбежно будут подмешаны сигналы от остальных источников. Следовательно, все восстановленные сигналы будут в какой-то степени похожи друг на друга, даже если исходные сигналы не демонстрировали никаких признаков синхронизации. А значит и меры коннективности, будучи мерами сходства сигналов, будут демонстрировать завышенные значения. Возникает проблема различия истинной синхронности и той, которая порождена фундаментальными ограничениями неинвазивной электрофизиологии.

Впервые попытка решения этой проблемы была предпринята в 2004 году в статье Г. Нолте [19], в которой авторы предлагают использовать в качестве меры коннективности величину, называемую мнимой частью когеренции. Для этого каждый сигнал сначала необходимо перевести в частотную область, затем для каждой пары сигналов посчитать функцию когерентности, и наконец, взять от полученной величины её мнимую часть. Идея такого метода оценки коннективности заключается в том, что мнимая часть когеренции имеет ненулевое значение лишь для сигналов с ненулевой разностью фаз, тогда как эффект протечки сигнала всегда проявляется в виде ложной синхронизации с нулевой фазовой задержкой, давая тем самым вклад лишь в действительную часть когеренции. Действительно, такой подход существенно повышает устойчивость метода к протечке сигнала. Тем не менее, так как функция когерентности нормируется на оцененные мощности сигналов (которые, будучи чисто действительными величинами, подвержены влиянию протечки сигнала) итоговые оценки коннективности по мнимой части когеренции также, пусть и в меньшей степени, испорчены эффектом протечки.

На эту деталь в 2007 году обратил внимание Стэм в своей статье [26]. Стэм предложил использовать для оценки синхронизации вместо мнимой части

когеренции среднее значение знака разности фаз двух сигналов. Такая мера оказывается очень похожей на мнимую часть когеренции, однако нормировка (скрытая в операции взятия знака мнимой части) теперь производится лишь на чисто мнимые величины, которые не зависят от протечки сигнала. Стэм назвал свою меру индексом фазовой задержки (phase lag index, PLI).

Следующей ступенью эволюции в цепочке методов, основанных на идеи мнимой части когеренции стала мера, называемая взвешенным индексом фазовой задержки (weighted phase lag index, wPLI). Ее описал Винк с соавторами в своей статье 2011 года. Мотивацией к разработке новой меры коннективности послужил тот факт, что мера PLI оказалась слишком неустойчивой по отношению к шуму. Основной недостаток индекса фазовой задержки, как и его преимущество перед мнимой частью когеренции, кроется в операции взятия знака. Дело в том, что для шумовых источников случайно меняющийся знак разности фаз оказывает слишком большое влияние на измерения. Чтобы избавиться от этого недостатка, Винк предложил взвешивать знак разности фаз на амплитуду мнимой части соответствующих кросс-спектральных коэффициентов. Таким образом, вклад от шумовых источников малой амплитуды оказывается малым, что делает меру более устойчивой.

Семейство мер коннективности, основанных на мнимой части когеренции, не исчерпывается обозначенными тремя подходами. Аналогичная идея, но под немного другим углом, была применена в статье [25] 2012 года. В ней в качестве меры синхронности авторы используют корреляцию огибающих двух узкополосных сигналов. Проблема протечки сигнала в статье решена следующим образом: на коре восстанавливаются два временных ряда, затем один из них проецируется ортогонально второму, после чего вычисляются огибающие и рассчитывается коэффициент корреляции между ними. Такой подход, основанный на ортогонализации временных рядов, оказывается эквивалентным взятию мнимой части соответствующего кросс-спектрального коэффициента.

Все изложенные методики, основанные на мнимой части когеренции, имеют, однако, один существенный недостаток, а именно — все они не чувствительны к синхронизации с нулевой фазовой задержкой. Как уже отмечалось выше, операция взятия мнимой части когеренции эквивалентна удалению из данных профилей синхронизации с нулевой фазовой задержкой. Практически это означает не только невозможность детектирования сетей, синхронизированных с нулевой фазой, но и плохое отношение сигнал / шум (ОСШ) для

сетей, для которых фазовая задержка мала. Более того, чем ближе эта фазовая задержка к нулю, тем хуже ОСШ для отдельно взятой пары источников. И наоборот, чем разность фаз двух сигналов ближе к  $\pi/2$ , тем выше значение ОСШ.

Ясно, что такое неравномерное распределение детекторных характеристик метода по фазовым задержкам ограничивает возможности исследователя. Этот факт усугубляется тем, что синхронизация с нулевой фазой, по всей видимости, является широко представленным явлением в организации осцилляторной мозговой активности, [34–36], которое может быть объяснено наличием общего входа для двух узлов сети, либо их двунаправленным взаимодействием, [37].

По этой причине на сегодняшний день в неинвазивной электрофизиологии имеется острая потребность в появлении инструмента измерения коннективности, который, с одной стороны, будет устойчив к эффекту протечки сигнала, а с другой — будет способен обнаруживать сети для всего спектра фазовых задержек.

Попытка создать такой метод была предпринята в 2015 году в статье [38]. В ней авторы использовали принципиально иной метод борьбы с эффектом протечки сигнала. Идея этого метода состоит в использовании информации о взаимном расположении источников сигнала и сенсоров для конструкции особых пространственных фильтров, которые позволяют очистить один источник от сигнала, пришедшего от другого источника для последующего измерения какого-либо индекса синхронности. Авторы в качестве такого индекса предложили использовать корреляцию огибающих сигналов. Более детально структура предложенного метода такова. Во-первых, по сигналам на сенсорах восстанавливаются сигналы на источниках. Далее фиксируется один из источников на коре. Все остальные источники пространственно фильтруются от активности, протекшей от фиксированного источника. Далее меряется корреляция огибающих между фиксированным источником и всеми остальными. Чтобы получить значение коннективности для каждой пары источников нужно повторить процедуру, выбирая в качестве фиксированного источника каждый из оставшихся. Наконец, так как полученная матрица коннективностей будет вообще говоря асимметричной, значения коннективности для пар  $(i,j)$  и  $(j,i)$  усредняются. Такую эвристику авторы статьи назвали методом геометрической поправки (geometric correction scheme, GCS).

Метод GCS концептуально явился серьезным продвижением вперед, так как теперь появилась возможность детектировать сети малыми сдвигами фаз оставаясь (по крайней мере, в теории) вне влияния эффекта протечки сигнала. В действительности, однако, такой метод коррекции лишь частично нивелирует этот эффект, так как он не учитывает протечку от третьих источников при оценке коннективности. В качестве примера можно рассмотреть ситуацию, когда имеется три мощных источника, никакие два из которых не были синхронизированы. В такой постановке несмотря на отсутствие синхронностей метод GCS будет давать высокие значения коннективности для всех трех пар связей, так как хотя для каждой пары коррекция очистит сигналы от протечки друг в друга, сигнал от третьего источника, протекая в каждый источник из пары, создаст общую компоненту в восстановленных источниках. В результате коннективность, которую мы измерим для исходно не синхронных источников, после геометрической коррекции для пары источников фактически будет отражать степень протечки от третьего источника в каждый сигнал из пары. Ясно, что если третий сигнал лежит близко к первым двум, эффект протечки будет весьма существенным. В результате для большого количества активных источников даже очищенный сигнал оказывается крайне загрязненным, что существенно ограничивает применимость GCS к практическим задачам.

*Таким образом, до сих пор не существует метода оценки коннективностей, позволяющего надежно детектировать сети с малыми фазовыми задержками и при этом свободного от эффекта протечки сигнала*

Современная практика использования мер коннективности в нейрофизиологических исследованиях в подавляющем большинстве случаев следует одной из двух возможных схем. Первая схема предполагает изучение нейрофизиологического эффекта в *пространстве сенсоров*, т.е. выбранная исследователем мера коннективности применяется непосредственно к сигналам, записанным электродами. Второй вариант предполагает переход в пространство источников — сначала оцениваются возможные источники записанной электрофизиологической активности на коре, а затем к этим источникам применяется та или иная мера коннективности.

Очевидным образом, первый вариант позволяет дать лишь весьма грубую оценку локализации узлов восстановленных сетей, поэтому часто используется лишь как первое приближение к результату. Более интересным, хотя и более сложным концептуально и более трудоемким с точки зрения вычислительных

ресурсов, является второй вариант, в котором сначала оценивается сигнал на источниках, а потом считается мера коннективности.

Оценка источников в неинвазивной электрофизиологии является некорректной обратной задачей [39]: ее решение не определено однозначно. Иными словами, любые электрофизиологические измерения, сделанные ограниченным набором сенсоров, можно объяснить бесконечным количеством конфигураций источников электромагнитной активности, расположенных на коре. Подавляющее большинство таких конфигураций при этом будет абсурдным с точки зрения физиологии. Среди бесконечного набора решений необходимо выбрать то, которое с одной стороны хорошо объясняет наблюдения, а с другой — соответствует имеющимся представлениям о физиологии мозга.

Поэтому решение обратной задачи в неинвазивной электрофизиологии всегда требует внедрения в модель дополнительной априорной информации о структуре решения. Не в каждом методе решения обратной задачи можно явно указать тот момент, в который делается дополнительное предположение о структуре решения, однако большая часть таких методов (например, [40—43]) может быть описана в терминах Тихоновской регуляризации [44], позволяющей свести задачу поиска решения к минимизации функционала, состоящего из двух членов: первый — насколько хорошо решение объясняет измеренный сигнал, второй — насколько оно соответствует тому классу решений, который мы считаем «физиологичным». При этом, формализация понятия «физиологичный» может включать в себя широкий спектр различных предположений о структуре решения — от естественного требования непрерывности по пространству и времени (как в MNE, [40]) до информации об анатомическом строении мозга.

Оценка источников в такой постановке происходит оптимально с точки зрения минимизации выбранного функционала, однако в задаче поиска синхронных осцилляций оценка источников не является самоцелью. Оптимальность этой оценки не гарантирует оптимальности оценки достаточных статистик синхронности в пространстве источников.

Двухступенчатая процедура оценки коннективностей в общем случае дает субоптимальные результаты с точки зрения оценивания соответствующих статистик. Оптимальное оценивание по наблюдениям статистик синхронности в пространстве источников, требует рассмотрения порождающих моделей с формулировкой априорных посылок для сетей вместо таковых для источников. В

частности, желательно было бы находить такие решения обратной задачи, которые объясняют измерения *минимальным набором сетей*. Мотивация такого подхода кроется в известном принципе бритвы Оккама — объяснение наблюдаемых данных должно быть максимально простым. Известно, что решения такого вида, то есть те, в которых число отдельных структурных элементов, объясняющих данные, минимально, реализуются при помощи спарсной регуляризации. Как вводить такую регуляризацию в рамках двухступенчатой процедуры, однако, не совсем понятно.

Имея в виду все вышесказанное, можно заключить, что на сегодняшний день процедура оценки коннективности по неинвазивным электрофизиологическим данным с одной стороны все еще является плохо разработанной и нуждается в улучшениях (неслучайна регулярная публикация новых методологических статей по теме оценки коннективностей), а с другой является ключевым инструментом для современной нейрофизиологии, следуя за смещением акцента в изучении мозга от активации его отдельных областей к взаимодействию между ними.

**Целью** данной работы, таким образом, является разработка метода оценки коннективностей, который

- позволяет оценить фазовую синхронность в условиях взаимной протечки сигналов
- чувствителен к сетям с малыми фазовыми задержками
- оптимален с точки зрения оценки достаточной статистики для коннективности
- способен учитывать априорную информацию об организации фазовых синхронностей
- не чувствителен к протечке сигнала на уровне сетей

а также его валидация в применении к симуляционным МЭГ-данным.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. Разработать методику очистки сигнала от объемной проводимости
2. Исследовать свойства методики очистки сигнала для сетей с малым и большим фазовым сдвигом; сравнить с методиками, описанными в литературе.
3. Разработать методологию оптимального оценивания достаточной статистики для фазовых синхронностей.

4. Реализовать алгоритм оценивания, позволяющий использовать спарсную регуляризацию
5. Разработать код для численного решения задачи невыпуклой оптимизации
6. Разработать методику визуализации найденных сетей.
7. Разработать методологию генерации данных, симулирующих мозговую активность
8. Сравнить детекторные характеристики разработанного метода на симуляционных данных по стандартным метрикам (AUC-ROC, AUC-Pre-Rec)

**Научная новизна:**

1. Впервые задача оценки коннективностей была рассмотрена в пространстве пар источников.
2. Впервые была продемонстрирована возможность очистки сигнала от эффекта протечки при помощи операции ортогональной проекции.
3. Впервые был сформулирован критерий оптимальной очистки от протечки сигнала для оценки фазовой синхронизации.
4. Впервые задача оценки матрицы кросс-спектральной плотности в пространстве источников была решена методом глобальной оптимизации.

**Теоретическая и практическая значимость.** Теоретическая значимость определяется тем, что впервые предложен подход к борьбе с проблемой протечки сигнала через векторизацию порождающей модели кросс-спектра, а также с помощью методов оптимальной фильтрации и глобальной оптимизации.

Практическая значимость состоит в том, что разработанный набор алгоритмов предоставляет новый инструмент в руки исследователя-электрофизиолога. Этот инструмент позволяет изучать взаимодействия корковых структур, ранее доступные для изучения только инвазивными методами.

**Методология и методы исследования.** Исследования основаны на теории обратных задач, теории оценивания, методах цифровой обработке сигналов, оптимальной фильтрации, глобальной оптимизации невыпуклых функций, а также на работах по методам оценки фазовой связности в электрофизиологии.

**Основные положения, выносимые на защиту:**

1. Разработан метод, позволяющий обнаруживать связанные по фазе источники с оклонулевыми фазовыми задержками по неинвазивным

МЭГ-записям. Суть метода заключается в построении пространственного фильтра, действующего в пространстве векторизованных матриц кросс-спектральной плотности мощности, который позволяет подавить вклад членов, ответственных за эффект протечки сигнала, маскирующий информацию о взаимодействии с оклонулевой фазой.

2. Показана оптимальность предложенного фильтра с точки зрения удаления вклада трех источников в оценку фазовой связности для фиксированной сети.
3. На основе метода глобальной оптимизации IrMxNE и предложенного фильтра разработан метод, устойчивый к протечке сигнала на уровне сетей и на уровне источников. Первое обеспечивается спарсными свойствами метода IrMxNE, второе — свойствами разработанного фильтра.

**Достоверность** полученных результатов обеспечивается теоретическими выкладками, результатами численного моделирования, сравнении с другими методами оценки фазовой связности, а также валидацией разработанного метода на реальных данных.

### Апробация работы.

Основные результаты работы докладывались на:

1. Международная конференция “Methodological problems of cortex regions functional synchronisation assessment based on MEG/EEG data”,  
Тема: *Globally-optimized power and shift invariant imaging of coherent sources (GO-PSIICOS)*  
Москва, Россия, 2015.
2. Международная конференция “Brain Connectivity Workshop 2015”,  
Тема: *GO-PSIICOS (Globally-Optimized Power and Shift Invariant Imaging of Coherent Sources),*  
Сан Диего, США, 2015.
3. Международная конференция “Biomag 2016”,  
Тема: *Power and shift invariant imaging of coherent sources by MEG data,*  
Сеул, Южная Корея, 2016.
4. Международная конференция “Biomag 2018”,  
Тема: *Oblique projection for phase shift invariant imaging of coherent sources,*  
Филадельфия, США, 2018.

5. Международная конференция “Biomag 2018”,

Тема: *NeuroPycon: A python package for efficient multi-modal brain network analysis,*

Филадельфия, США, 2018.

**Личный вклад.** Все представленные в диссертации результаты получены лично автором. При подготовке статей и докладов автор опирался на помошь соавторов и научного руководителя.

**Публикации.** Основные результаты по теме диссертации изложены в 5 печатных изданиях, 5 из которых изданы в журналах, рекомендованных ВАК, 0 — в тезисах докладов.

Полный объём диссертации составляет 167 страниц, включая 25 рисунков и 0 таблиц. Список литературы содержит 80 наименований.

## Глава 1. Взаимодействие нейронных ансамблей и методы его изучения в электрофизиологии.

Чтобы обрисовать контекст, в котором задача оценки коннективностей имеет смысл, необходимо сперва описать основные принципы работы мозга, механизмы передачи информации между нейронными популяциями и то, каким образом мы можем получить информацию о таком обмене информацией средствами неинвазивной электрофизиологии.

Согласно последним оценкам [45], в мозге здорового взрослого человека присутствует порядка 86 миллиардов нейронов. Каждый отдельно взятый нейрон представляет собой сравнительно простой биохимический механизм, принимающий электрические импульсы от других нейронов на вход и способный генерировать выходной импульс в случае, если входной сигнал превышает определенный порог. Тем не менее, этого элементарного функционала отдельно взятых нейронов оказывается достаточно для обеспечения сложнейшей работы головного мозга в целом. Мозг, таким образом, представляет собой сложную систему, составленную из сравнительно просто устроенных элементов.

По-прежнему неизвестно, за счет чего достигается качественный скачок в поведении системы, состоящей из совокупности нервных клеток, на пути от простой биоэлектрической сети, узлы которой способны обмениваться электрическими импульсами, к системе, обладающей интеллектом и сознанием собственного “я”. Каким-то образом оказывается возможен переход количества в качество, — увеличение числа сравнительно простых элементов системы, таких как нейрон, до невероятных 86 миллиардов порождает удивительную сложность и богатство поведенческих сценариев.

Тем не менее есть надежда, что ключ к пониманию работы мозга по крайней мере частично лежит в изучении работы отдельных его частей и способов обмена информацией между ними.

## 1.1 Биологические механизмы обмена информацией между нейронными популяциями.

Разберем сперва, какие именно механизмы позволяют нейронам обмениваться электрическим сигналом на примере пирамидальных нейронов, из которых по большей части и состоит кора больших полушарий.

Пирамидальный нейрон, как и любой другой, имеет аксон и разветвленную сеть отростков, называемых дендритами. Активность нейрона характеризуется генерацией так называемых потенциалов действия или спайков в ответ на внешний электрохимический вход.

Потенциалы действия, распространяясь по аксону, достигают синапса. Пришедший импульс играет роль триггера, провоцируя выделение нейромедиаторов в синаптическую щель. Нейромедиаторы достигают окончания дендрита следующего нейрона, вызывая в нем увеличение, либо уменьшение трансмембранныго потенциала — в зависимости от того, являлся нейрон, сгенерировавший потенциал действия, тормозным или возбуждающим. Трансмембранный потенциал определяется как разность потенциалов между внешней и внутренней поверхностью мембранны клетки и зависит от молярных концентраций ионов по обе стороны мембранны. В такой схеме обмена информацией между двумя нейронами первый нейрон, породивший потенциал действия, называется пресинаптическим, а второй, принимающий сигнал — постсинаптическим.

При этом, к каждому нейрону подведено множество аксонов. Трансмембранный потенциал отдельно взятого нейрона аккумулирует воздействия от нейронов, с которыми у него имеются синаптические связи. Будучи равным  $\approx -70$  микровольт в покое, трансмембранный потенциал изменяется в результате потенциалов действия, приходящих от пресинаптических нейронов и, в при достижении порога деполяризации мембранны в  $\approx -30$  микровольт постсинаптический нейрон сам производит потенциал действия. Таким образом, нейроны выполняют роль сумматоров в иерархической сети передачи сигнала.

Однако прежде чем трансмембранный потенциал клетки “почувствует” изменения, пришедший импульс должен через дендритные отростки достичь мембранны. Электрический потенциал, распространяющийся по дендриту от синапса к мемbrane носит название *постсинаптического потенциала*. При этом,

само распространение электрической активности по дендриту есть не что иное как движение ионов, то есть электрический ток. На макроуровне совокупность нейронов и окружающая их спинномозговая жидкость представляют собой объемный проводник, из-за чего распространение токов вдоль дендритных отростков, называемых *первичными токами*, вызывает появление *вторичных* объемных токов в мозге.

Вследствие такой электрохимической проницаемости трансмембранные потенциалы локальных нейронных популяций оказываются связаны между собой. В силу того, что именно трансмембранный потенциал управляет генерацией потенциала действия, а значит и передачей информации от нейрона к нейрону, оказывается, что отдельные нейроны в информационном смысле не изолированы друг от друга, и можно говорить о передаче информации не только от нейрона к нейрону, но и от популяции нейронов к популяции.

Ключевой особенностью для работы групп пирамидальных нейронов является особая организация их дендритных отростков, а именно — наличие одного выделенного дендрита, значительно превосходящего по размерам остальные. Такой дендрит называется апикальным, а все остальные в рамках такой дихотомии — базальными.

Пространственно апикальные дендриты пирамидальных нейронов в коре больших полушарий организованы весьма особым образом — если говорить о локальной популяции, апикальные дендриты в ней ориентированы параллельно друг другу. Вследствие этого первичные токи пирамидальных нейронов внутри популяции суммируются, делая возможным их детектирование в неинвазивной электрофизиологии. Действительно, электрическая активность отдельного нейрона слишком слаба для регистрации за пределами черепной коробки существующими сегодня методами, в то время как суммарная активность группы из десятков тысяч нейронов в силу коллинеарности их апикальных дендритов производит достаточный ток для регистрации за пределами головы.

Таким образом, обмен информацией между нейронами происходит посредством двух основных механизмов — за счет распространения по аксону потенциала действия и за счет распространения постсинаптических потенциалов, при чем оба эти механизма тесно связаны между собой.

Одним из эпифеноменов такой взаимосвязи является появление в паттернах суммарного постсинаптического потенциала ритмической осцилляторной активности на сравнительно низких частотах. Широко известен, например,

альфа-ритм, традиционно относимый к полосе частот 8 – 12 Гц, с регистрацией которого в 1924 году немецким физиологом Гансом Бергером ведет отсчет современная электрофизиология.

Стоит отметить, что выделение отдельных полос частот из общего спектра электрической активности групп нейронов неслучайно и связано с определенными функциональными атрибутами, различными физиологически обусловленными причинами их возникновения а также с их пространственной специфичностью. Так, уже упомянутый альфа-ритм в покое наиболее сильно представлен в затылочных долях мозга и считается ритмом покоя зрительной коры в отсутствие визуального раздражения. К другим функционально-специфичным и широко освещенным в литературе полосам частот относят дельта-, тета-, бета- и гамма-ритмы.

### **1.1.1 Осцилляции и функциональная интеграция.**

Роль осцилляций в передаче информации между нейронными ансамблями на сегодняшний день все еще остается плохо изученной, однако согласно принятой в научном сообществе гипотезе взаимодействия через когерентность [3], именно осцилляции постсинаптического потенциала являются медиатором коммуникации между отдельными областями коры.

Такое взаимодействие предполагается необходимым в осуществлении так называемой *функциональной интеграции*. Суть понятия функциональной интеграции состоит в следующем.

Известно, что существуют определенные зоны мозга, активность в которых функционально специфична. Так например было показано, что электрическая стимуляция определенных областей прецентральной зоны провоцирует неконтролируемые движения а иногда и активирует целые двигательные программы. При этом в зависимости от конкретного места стимуляции действуют различные части тела. Эти области прецентральной зоны принято называть первичной моторной корой. Оказывается, что в первичной моторной коре представлена подробная карта тела человека. Аналогичная ситуация наблюдается и для постцентральной зоны — стимуляция соответствующих областей вызывает тактильные ощущения в различных частях тела. Источник

тактильного ощущения на теле при этом также зависит от конкретного места стимуляции. В соответствии с этим выделяют соматосенсорную кору. Такие функционально-специфичные зоны коры существуют и для других отделов мозга. Тем не менее, не всякую активность удается так хорошо локализовать на коре.

Так или иначе, субъективные ощущения подсказывают нам, что в реальной жизни деятельность мозга никогда не сводится к какой-либо одной отдельно взятой функции, как, скажем, движение указательным пальцем. Скорее наоборот, большое количество параллельных, но связанных друг с другом процессов осуществляется мозгом одновременно. При этом для нормальной работы такие процессы должны взаимодействовать друг с другом. Для примера возьмем задачу в которой испытуемому предлагали зажечь спичку, а затем проделать тоже самое, но с предварительной анестезией кончиков пальцев. Было показано, что потеря чувствительности ведет к весьма значительному увеличению времени на выполнение той же задачи. Из этого можно заключить, что информация о тактильных ощущениях определенным образом *интегрируется* мозгом с двигательной программой при выполнении задачи зажигания спички.

Термин *функциональная интеграция* используется для обозначения такого явления взаимодействия функционально-специфичных отделов мозга для решения сложной когнитивной задачи, как в задаче со спичкой. Предполагается, что раз существуют зоны мозга, ответственные за осуществление той или иной функции, и активные во время осуществления этой функции, то совместное осуществление нескольких таких функций, требующих обмена информацией между собой должно вызывать совместную активацию соответствующих областей мозга и обмен информацией между ними.

### **1.1.2 Взаимодействие через когерентность (гипотеза СТС).**

Теперь, введя понятие функциональной интеграции, мы можем вернуться к гипотезе *взаимодействия через когерентность*. Согласно этой гипотезе, функциональная интеграция реализуется в мозге посредством установления когерентных осцилляций постсинаптических потенциалов во взаимодействующих областях коры.

Обоснование гипотезы состоит в следующем. Вероятность того, что нейрон произведет потенциал действия тем больше, чем выше величина локального постсинаптического потенциала. В силу того, что этот потенциал осциллирует, нейрон способен реагировать на входящие потенциалы действия не всегда, а только в те моменты, когда он был уже достаточно возбужден, т.е. тогда, когда локальный постсинаптический потенциал находился в максимуме. Рассмотрим теперь два нейрона, находящиеся в двух различных нейронных популяциях. Назовем первый из них нейроном А, а второй — нейроном В. Для успешной передачи информации от нейрона А к нейрону В необходимо, чтобы потенциал популяции нейрона В имел достаточно высокое значение в тот момент, когда спайк от нейрона А достигнет цели. Кроме того, чтобы нейрон А был способен произвести спайк, локальный потенциал его популяции в момент генерации также должен быть достаточно высок. Иными словами, осцилляции постсинаптического потенциала регулируют моменты возможного приема и передачи информации между нейроном А и нейроном В. Таким образом, возникает два временных окна, ритмически появляющихся и пропадающих во времени — первое окно соответствует промежуткам времени, когда нейрон А может передать сигнал, а второе — тем моментам времени, когда нейрон В может его принять. Если считать, что потенциал действия распространяется от аксона нейрона А к дендриту нейрона В каждый раз за одно и то же время, то для установления стабильного канала передачи информации от нейрона А к нейрону В необходимо, чтобы эти два окна открывались и закрывались через одинаковые промежутки времени, и чтобы окна приема информации было сдвинуто относительно окна передачи на величину, равную времени распространения сигнала от нейрона А к нейрону В. Последнее справедливо в том случае, если осцилляции двух нейронных популяций *когерентны*, то есть имеют одинаковую частоту и, возможно, некоторый ненулевой фазовый сдвиг. Таким образом, согласно гипотезе взаимодействия через когерентность, синхронизация осцилляций нейронных популяций способствует обмену информацией в виде потенциалов действия между ними, в то время как рассинхронизация осцилляций ведет к блокированию обмена сигналами.

### 1.1.3 Формальное определение когерентности

Дадим более формальное определение когерентности абстрактных сигналов. Пусть имеется два случайных процесса  $s_1, s_2$ :

$$s_1 = s_1(t, \omega), s_2 = s_2(t, \omega) \quad (1.1)$$

где переменная  $t$  — время, а  $\omega$  — случайная величина из вероятностного пространства  $\Omega$ . Под сигналами  $s_1(t), s_2(t)$  будем понимать конкретные реализации случайных процессов  $s_1(t, \omega), s_2(t, \omega)$ , фиксируя величину  $\omega = \omega_0$ .

Наиболее естественным способом проверить, является ли один случайный процесс сдвинутой во времени копией другого, является подсчет кросс-ковариации, которая определяется как:

$$C_{s_1, s_2}(t_1, t_2) = cov(s_1(t_1), s_2(t_2)) \stackrel{\text{def}}{=} \mathbf{E} \left\{ (s_1(t_1, \omega) - \mu_1(t_1))(s_2(t_2, \omega) - \mu_2(t_2)) \right\} \quad (1.2)$$

где  $\mu_i(t_i) = E(s_i(t_i)), i = 1, 2$  — среднее по ансамблю функции случайной величины величины  $s_i(t_i, \omega)$ .

Для простоты изложения будем считать, что сигналы имеют нулевое среднее, имея в виду, что мы всегда можем вычесть среднее по ансамблю из каждой реализации случайного процесса. Рассмотрим сначала случай стационарных в широком смысле процессов, а потом обобщим рассуждения на случай нестационарности. Стационарными в широком смысле называются такие случайные процессы, для которых среднее не зависит от времени, а автоковариация зависит только от сдвигов по времени. Кроме того, мы потребуем от процессов *совместной* стационарности в широком смысле, то есть мы будем считать, что кросс-ковариация, также зависит только от сдвига по времени  $\tau = t_2 - t_1$ :

$$C_{s_1, s_2}(t_1, t_1 + \tau) = C_{s_1, s_2}(\tau) \quad (1.3)$$

Введем дополнительные понятия, которые понадобятся нам для определения когерентности. Сперва вспомним понятие энергии сигнала. Энергией сигнала на промежутке  $\left[-\frac{T}{2}, \frac{T}{2}\right]$  называют величину

$$E_T \stackrel{\text{def}}{=} \mathbf{E} \left\{ \int_{-T/2}^{+T/2} |s(t, \omega)|^2 dt \right\} \quad (1.4)$$

Полная энергия сигнала получается взятием предела от  $E_T$ :

$$E \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} E_T \quad (1.5)$$

Для стационарных процессов, однако, величина полной энергии, вообще говоря, не определена, поэтому вводят понятие средней мощности сигнала:

$$P \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \frac{E_T}{T} \quad (1.6)$$

Введем операцию преобразования Фурье на отрезке  $\left[-\frac{T}{2}, \frac{T}{2}\right]$  для процесса  $s$  следующим образом:

$$F_s^T(f, \omega) \stackrel{\text{def}}{=} \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} s(t, \omega) e^{-ift} dt \quad (1.7)$$

Тогда, пользуясь теоремой Парсеваля, для средней мощности сигнала будем иметь

$$P = \lim_{T \rightarrow \infty} \mathbf{E} \left\{ \int_{-\infty}^{+\infty} |F_s^T(f, \omega)|^2 df \right\} \quad (1.8)$$

Далее, определим спектральную плотность мощности сигнала:

$$S_{ss}(f) \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \mathbf{E} \left\{ |F_s^T(f, \omega)|^2 \right\} \quad (1.9)$$

Спектральная плотность мощности характеризует распределение мощности сигнала по частотам. По аналогии с 1.9 вводят понятие кросс-спектральной плотности мощности

$$S_{s_1, s_2}(f) \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \mathbf{E} \left\{ F_{s_1}^T(f, \omega)^* F_{s_2}^T(f, \omega) \right\} \quad (1.10)$$

Введя основные понятия, мы готовы теперь определить когеренцию и дать интуитивное представление о том, почему величина когерентности отражает степень синхронизации сигналов.

Для частотного образа функции кросс-ковариации случайных процессов, стационарных в широком смысле, справедлива теорема Винера-Хинчина, согласно которой кросс-ковариация и кросс-спектральная плотность мощности связаны друг с другом через преобразование Фурье:

$$S_{s_1, s_2}(f) = \int_{-\infty}^{+\infty} C_{s_1, s_2}(\tau) e^{-if\tau} d\tau \quad (1.11)$$

Или, используя 1.10:

$$\lim_{T \rightarrow \infty} \mathbf{E} \left\{ F_{s_1}^T(f, \omega)^* F_{s_2}^T(f, \omega) \right\} = \int_{-\infty}^{+\infty} C_{s_1, s_2}(\tau) e^{-if\tau} d\tau \quad (1.12)$$

Рассмотрим подробнее, как устроено произведение  $F_{s_1}^T(f, \omega)^* F_{s_2}^T(f, \omega)$ . Пользуясь экспоненциальным представлением комплексного числа, для коэффициента Фурье на частоте  $f$  будем иметь:

$$F_{s_j}^T(f) = A_j(f) * e^{i\varphi_j}, j = \overline{1, 2} \quad (1.13)$$

где  $A_i(f)$  — амплитуда синусоиды на частоте  $f$  в соответствующем разложении Фурье, а  $\varphi_i$  — ее фаза. Отметим, что частотное представление позволяет в явном виде получить доступ к фазе сигнала.

Используя 1.13, можем переписать выражение для произведения фурье-образов на отрезке  $\left[-\frac{T}{2}, \frac{T}{2}\right]$  как:

$$F_{s_1}^T(f, \omega)^* F_{s_2}^T(f, \omega) = A_1(f, \omega) A_2(f, \omega) e^{i(\varphi_2(f, \omega) - \varphi_1(f, \omega))} \quad (1.14)$$

Из соотношения выше видно, что на частоте  $f$  фаза произведения фурье-образов реализаций случайного процесса совпадает с разностью фаз соответствующих комплексных экспонент для индивидуальных преобразований Фурье. Применяя далее операцию усреднения по ансамблю, получим значение кросс-спектральной плотности мощности на частоте  $f$ :

$$S_{s_1, s_2}(f) = \mathbf{E} \left\{ A_1(f, \omega) A_2(f, \omega) e^{i(\varphi_2(f, \omega) - \varphi_1(f, \omega))} \right\} \quad (1.15)$$

Таким образом, мы получили, что значение кросс-спектральной плотности мощности равно среднему по ансамблю от комплексного числа, аргумент которого равен разности фаз гармоник индивидуальных сигналов с частотой

$f$ , а модуль — произведению амплитуд этих гармоник. Имея в виду, что поле комплексных чисел изоморфно двумерному векторному пространству, можем понимать операцию усреднения в уравнении 1.15 как усреднение векторов на плоскости, повернутых относительно направления оси абсцисс на угол, равный разности фаз сигналов. Нетрудно понять, что средний вектор при фиксированных амплитудах будет иметь тем большую норму, чем более воспроизведима соответствующая разность фаз по реализациям случайного процесса. Воспроизводимая по реализациям разность фаз, в свою очередь, свидетельствует о синхронизации случайных процессов, т.е. о наличии некой линейной связи между ними. Таким образом, длина результирующего вектора может служить мерой синхронности случайных процессов. Вместе с тем, чтобы исключить влияние амплитуды на получающееся значение, применяют нормировку на среднее по ансамблю значение мощностей сигналов на той же частоте. Получающаяся нормированная величина и есть по определению когерентность двух сигналов:

$$G_{s_1,s_2}(f) \stackrel{\text{def}}{=} \frac{S_{s_1,s_2}(f)}{\sqrt{S_{s_1,s_1}(f)S_{s_2,s_2}(f)}} \quad (1.16)$$

Или, используя определение кросс-спектральной плотности мощности:

$$G_{s_1,s_2}(f) = \lim_{T \rightarrow \infty} \frac{\mathbf{E}\left\{F_{s_1}^T(f)^* F_{s_2}^T(f)\right\}}{\sqrt{\mathbf{E}\left\{F_{s_1}^T(f)^* F_{s_1}^T(f)\right\} \mathbf{E}\left\{F_{s_2}^T(f)^* F_{s_2}^T(f)\right\}}} \quad (1.17)$$

Определенная таким образом функция когерентности может быть использована лишь для стационарных в широком смысле сигналов. В то же время предположение о стационарности зачастую противоречит природе изучаемых процессов. В частности, нестационарны по своей природе сигналы, возникающие в электрофизиологии. Тем не менее, возможно обобщить определение когерентности и на случай нестационарности. Обобщение достигается путем использования вместо преобразования Фурье одного из *частотно-временных преобразований*. В частности, можно использовать преобразование Фурье с окном, вейвлет-преобразование или узкополосную фильтрацию с центральной частотой  $f$  с последующим извлечением аналитического сигнала.

В каждом из этих методов возникает характерная ширина временного окна (возможно различная для разных частот  $f$ ). Предполагая, что статистические свойства случайного процесса меняются со временем медленно, можем

повторить рассуждения выше для сужения стационарных процессов на временное окно частотно-временного преобразования за исключением перехода, для которого мы использовали теорему Винера-Хинчина. Теорема Винера-Хинчина, однако, также допускает обобщение на случай нестационарных процессов [46].

Таким образом, для нестационарных процессов будем иметь:

$$G_{s_1, s_2}(f, t) \stackrel{\text{def}}{=} \frac{\mathbf{E}\left\{\mathcal{F}_{s_1}(f, t, \omega)^* \mathcal{F}_{s_2}(f, t, \omega)\right\}}{\sqrt{\mathbf{E}\left\{\mathcal{F}_{s_1}(f, t, \omega)^* \mathcal{F}_{s_1}(f, t, \omega)\right\} \mathbf{E}\left\{\mathcal{F}_{s_2}(f, t, \omega)^* \mathcal{F}_{s_2}(f, t, \omega)\right\}}} \quad (1.18)$$

где  $\mathcal{F}_{s_i}(f, t, \omega)$  — частотно-временной образ сигнала  $s_i$  ( $i = \overline{1, 2}$ )

Отметим одно интересное свойство функции когерентности. Как в стационарном, так и в нестационарном случае когерентность является комплекснозначной функцией частоты  $f$ . При этом значения этой функции, согласно 1.15, пропорциональны средневзвешенной комплексной экспоненте с аргументом, зависящим от разности фаз. При этом веса равны произведениям амплитуд исходных сигналов на заданной частоте. Если в иллюстративных целях предположить, что амплитуды и фазы независимы, получим для когерентности следующее выражение:

$$G_{s_1, s_2}(f, t) = \frac{\mathbf{E}\left\{A_1(f, t, \omega) A_2(f, t, \omega)\right\}}{\sqrt{\mathbf{E}\left\{A_1(f, t, \omega)^2\right\} \mathbf{E}\left\{A_2(f, t, \omega)^2\right\}}} \mathbf{E}\left\{e^{i\Delta\varphi(f, t, \omega)}\right\}, \quad (1.19)$$

где  $\Delta\varphi = \varphi_2 - \varphi_1$ . Рассмотрим подробнее множитель  $\mathbf{E}\left\{e^{i\Delta\varphi(f, t, \omega)}\right\}$ . В случае, если разность фаз  $\Delta\varphi$  от реализации к реализации будет иметь воспроизводимые близкие к нулю значения, средняя комплексная экспонента также будет иметь близкий к нулю аргумент, т.е. действительная часть результирующей комплексной экспоненты будет иметь высокое значение по сравнению с мнимой частью. Иными словами, если случайные процессы были синхронизированы с нулевой фазой, их функция когерентности будет иметь почти действительные значения. С другой стороны, если разность фаз воспроизведено близка к  $\pi/2$ , итоговое значение когерентности будет близко к мнимому числу. Это наблюдение понадобится нам в дальнейшем, так как одной из наиболее популярных сегодня методик оценки коннективности на неинвазивных данных является так называемая мнимая часть когерентности.

## 1.2 Модель неинвазивных МЭГ/ЭЭГ измерений.

В предыдущих разделах мы рассмотрели, каким образом установление когерентных осцилляций между постсинаптическими потенциалами нейронных популяций позволяет осуществлять функциональную интеграцию, а также сделали первый шаг на пути к оценке когерентности на основании измерений электрической активности мозга — дали формальное определение функции когерентности двух случайных процессов.

В определении когерентности 1.16 фактически содержится рецепт ее вычисления — нам лишь необходимо оценить значения соответствующих математических ожиданий из данных.

Однако в неинвазивной электрофизиологии мы не имеем прямого доступа к локальным осцилляциям потенциалов в нейронных популяциях, так как активность мозга регистрируется при помощи сенсоров, находящихся за пределами головы и следовательно отражает некую суммарную активацию нейронных ансамблей.

Как было отмечено выше, источником сигнала, снимаемого при помощи ЭЭГ или МЭГ служат первичные токи, т.е. токи, распространяющиеся вдоль апикальных дендритов пирамидальных нейронов коры. Первичные токи вызывают появление вторичных токов в объемном проводнике, которым является мозг. В силу параллельной ориентации апикальных дендритов, электромагнитные поля, порождаемые первичными токами, накладываются, генерируя суммарное электрическое поле достаточной силы для регистрации за пределами головы.

Чтобы понять, как соотносятся суммарные колебания постсинаптического потенциала, порождаемые нейронными популяциями, с регистрируемым сигналом, начнем с рассмотрения уравнений Максвелла.

### 1.2.1 Уравнения Максвелла в квазистатическом приближении.

В общем виде для проводящей среды уравнения Максвелла можно записать в виде

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (1.20)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (1.21)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (1.22)$$

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}), \quad (1.23)$$

где  $\mathbf{E}$  — напряженность электрического поля,  $\mathbf{B}$  — магнитная индукция,  $\rho$  — объемная плотность стороннего электрического заряда,  $\mathbf{J}$  — плотность электрического тока,  $\mu_0$  — магнитная постоянная,  $\epsilon_0$  — электрическая постоянная.

При этом для мозга магнитная проницаемость  $\mu = \mu_0$ . В немагнитной среде без дисперсии  $\mathbf{J}$  является суммой омических и поляризационных токов:

$$\mathbf{J} = \sigma \mathbf{E} + \frac{\partial \mathbf{P}}{\partial t}, \quad (1.24)$$

где  $\mathbf{P} = (\epsilon - \epsilon_0)\mathbf{E}$  — вектор поляризации, а  $\epsilon$  — диэлектрическая проницаемость среды.

В электрофизиологии принято рассматривать уравнения Максвелла квазистатическом приближении, которое заключается в том, что в уравнениях выше мы можем опустить производные напряженности электрического поля и магнитной индукции по времени ( $\frac{\partial \mathbf{E}}{\partial t}$  и  $\frac{\partial \mathbf{B}}{\partial t}$ ). Дело в том, что в нейромагнетизме мы как правило имеем дело с частотами колебаний  $\leq 100$  Гц; на клеточном уровне электрохимические процессы по большей части происходят на частотах менее 1000 кГц. Таким образом, для электрического процесса, проходящего на частоте  $f$ , будем иметь

$$\mathbf{E} = \mathbf{E}_0 e^{i2\pi ft} \quad (1.25)$$

Рассматривая теперь уравнения 1.23, 1.24, будем иметь:

$$\nabla \times \mathbf{B} = \mu_0(\sigma E + (\epsilon - \epsilon_0) \frac{\partial \mathbf{E}}{\partial t} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}), \quad (1.26)$$

Квазистатическое приближение уравнений Максвелла справедливо в том случае, если члены в уравнении выше, содержащие производные по времени, малы по сравнению с омическими токами:  $|\epsilon \frac{\partial \mathbf{E}}{\partial t}| \ll |\sigma \mathbf{E}|$ , то есть если  $2\pi f \epsilon / \sigma \ll 1$ .

Полагая  $\sigma = 0.3 \text{ Ом}^{-1} \cdot \text{м}^{-1}$  (значение для тканей мозга),  $\epsilon = 10^5 \epsilon_0$ , а  $f = 100$  Гц, получим  $2\pi f \epsilon / \sigma = 2 \times 10^{-3} \ll 1$ .

Кроме того, значение  $\frac{\partial \mathbf{B}}{\partial t}$  также должно быть мало. Из уравнений 1.21 и 1.23 будем иметь:

$$\nabla \times \nabla \times \mathbf{E} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = -\mu_o \frac{\partial}{\partial t}(\sigma \mathbf{E} + \epsilon \frac{\partial \mathbf{E}}{\partial t}) = -i2\pi f \mu_0 (\sigma + i2\pi f \epsilon) \mathbf{E} \quad (1.27)$$

Для решений этого уравнения характерный пространственный масштаб имеет величину

$$\lambda_c = |2\pi f \mu_0 \sigma (1 + 2\pi f \epsilon / \sigma)|^{-1/2}. \quad (1.28)$$

Для значений параметров, указанных выше,  $\lambda_c = 65\text{м}$  и значительно пре-восходит диаметр головы. Следовательно, можем пренебречь вкладом члена  $\frac{\partial \mathbf{B}}{\partial t}$  в значение величины  $\mathbf{E}$ .

В квазистатическом приближении  $\nabla \times \mathbf{E} = 0$ , а значит напряженность электрического поля представима в виде скалярного потенциала

$$\mathbf{E} = -\nabla V \quad (1.29)$$

Окончательно в квазистатическом приближении система уравнений Максвелла перепишется в виде:

$$\Delta V = -\frac{\rho}{\epsilon_0} \quad (1.30)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (1.31)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad (1.32)$$

### 1.2.2 Первичные и вторичные токи.

Рассмотрим теперь, как первичные и вторичные токи входят в уравнения Максвелла. Из уравнений выше видно, что объемная плотность тока  $\mathbf{J}(\mathbf{r})$  представима в виде суммы двух компонент. Объемные омические токи  $\mathbf{J}^v(\mathbf{r}) = \sigma(\mathbf{r})\mathbf{E}(\mathbf{r})$  по своей физической природе пассивны и появляются в следствие действия макроскопического электрического поля на носители заряда

в проводящей среде. Эти токи мы будем называть *вторичными*. Остальные члены уравнения будем относить к *первичным* токам  $\mathbf{J}^p$ :

$$\mathbf{J}(\mathbf{r}) = \mathbf{J}^p(\mathbf{r}) + \sigma(\mathbf{r})\mathbf{E}(\mathbf{r}) = \mathbf{J}^p(\mathbf{r}) - \sigma(\mathbf{r})\nabla V(\mathbf{r}) \quad (1.33)$$

Такое определение не имело бы смысла без указания соответствующего пространственного масштаба. Здесь  $\sigma(\mathbf{r})$  — макроскопическая проводимость; детали, относящиеся к клеточному масштабу, мы оставляем без внимания. Иными словами, весь объем мозга моделируется как однородный объемный проводник. Разделение в уравнении 1.33 иллюстрирует тот факт, что нейрональная активность порождает первичные токи по большей части вблизи клетки, тогда как вторичные токи протекают повсюду в объеме проводящей среды. Таким образом, найдя источники первичных токов, мы найдем источники нейрональной активности мозга.

Необходимо дополнительно отметить, что источники токов  $\mathbf{J}^p$  должны рассматриваться как батарейки, помещенные в макроскопический проводник. При этом, хотя преобразование градиентов концентраций в ток происходит в основном благодаря диффузии, первичные токи во многом определяются структурой проводимости на клеточном уровне. В частности, клеточные мембранны, будучи хорошим проводником, определяют направление как внутриклеточных, так и межклеточных токов.

Перепишем уравнения Максвелла с учетом разделения на первичные и вторичные токи:

$$\Delta V = -\frac{\rho}{\epsilon_0} \quad (1.34)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (1.35)$$

$$\nabla \times \mathbf{B} = \mu_0(\mathbf{J}^p - \sigma\nabla V), \quad (1.36)$$

### 1.2.3 Модель токового диполя.

Итак, для определения источников нейрональной активности на коре необходимо локализовать источники первичных токов.

Теоретическая оценка снизу минимального количества синапсов пирамидальных нейронов, необходимого для создания измеримого за пределами

головы сигнала (для МЭГ такой сигнал соответствует плотности тока в 10 нА/м<sup>2</sup>) составляет примерно один миллион синапсов, что соответствует размеру участка коры, порядка одного квадратного миллиметра с учетом того, что только небольшая часть синапсов для каждого отдельно взятого нейрона активна в каждый фиксированный момент времени [39]. Более реалистичные оценки дают значения на минимальную активную площадь коры от 40 м<sup>2</sup> до 200 м<sup>2</sup>. При этом источник и сток заряженных частиц разнесены на расстояние порядка толщины коры больших полушарий, которая для человека составляет 2–4 мм, [47]. По отношению к характерному расстоянию до сенсоров, которое в случае МЭГ для ближайших к сенсорам точек коры составляет порядка нескольких сантиметров, размеры локальных источников суммарного тока принято считать точечными. Для моделирования таких точечных источников тока на коре вводят понятие *токового диполя*.

Под токовым диполем понимают такой источник тока, у которого сток и исток разнесены на пренебрежимо малое расстояние. Токовый диполь характеризуется тремя координатами, задающими его положение в пространстве, а также дипольным моментом — вектором, характеризующим величину и направление тока.

Токовый диполь  $\mathbf{Q}$  в точке  $\mathbf{r}_Q$  можно представить себе как сосредоточенный в одной точке источник первичного тока:

$$\mathbf{J}^p(\mathbf{r}) \stackrel{\text{def}}{=} \mathbf{Q}\delta(\mathbf{r} - \mathbf{r}_Q), \quad (1.37)$$

где  $\delta(\mathbf{r})$  — дельта-функция Дирака. В МЭГ/ЭЭГ приложениях токовые диполи используют в качестве эквивалентных источников первичных токов для участков коры по площади иногда доходящих до нескольких квадратных сантиметров.

В силу линейности уравнений Максвелла активность распределенного по коре набора токовых диполей отобразится на сенсорах как суперпозиция активностей отдельных диполей, поэтому для построения модели данных, записываемых в МЭГ/ЭЭГ парадигме, необходимо рассмотреть, как выглядит за пределами головы электромагнитная активность отдельного токового диполя.

### 1.2.4 Поиск решений для кусочно постоянного профиля проводимости.

Для определения магнитного поля за пределами головы проинтегрируем уравнение 1.36:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_G (\mathbf{J}^p(\mathbf{r}') - \sigma(\mathbf{r}') \nabla V(\mathbf{r}')) \times \frac{\mathbf{R}}{R^3} dv' \quad (1.38)$$

Здесь  $\mathbf{R} = \mathbf{r} - \mathbf{r}'$  — расстояние от источника тока до точки, в которой мы хотим определить значение вектора магнитной индукции. При этом мы предполагаем, что за пределами головы источники первичных токов отсутствуют, а проводимость равна нулю, поэтому объемный интеграл в уравнении 1.38 берется не по бесконечной области, а лишь по объему головы  $G$ .

Заметим, что  $\sigma(\mathbf{r}')$  вообще говоря различна для разных точек головы. Тем не менее, в объеме головы можно выделить области однородной проводимости, такие как кожа головы, череп, цереброспинальная жидкость, серое и белое вещество. Будем считать, что в объеме головы содержится  $m$  таких областей, которые мы обозначим как  $G_i$ . Тогда уравнение 1.38 перепишется в виде:

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi} \sum_{i=1}^m \sigma_i \int_{G_i} \nabla V(\mathbf{r}') \times \frac{\mathbf{R}}{R^3} dv', \quad (1.39)$$

где

$$\mathbf{B}_0(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_G \mathbf{J}^p(\mathbf{r}') \times \frac{\mathbf{R}}{R^3} dv' \quad (1.40)$$

представляет собой магнитное поле, порождаемое только первичными токами.

Далее, используя векторное равенство  $\nabla \times (V \nabla(1/R)) = \nabla V \times \nabla(1/R)$  и  $\int_G \nabla \times \mathbf{u} dv = \int_S \mathbf{S} \times \mathbf{u}$  получим

$$\int_{G_i} \nabla' V \times \frac{\mathbf{R}}{R^3} dv' = \int_{G_i} \nabla' V \times \nabla' \frac{1}{R} dv' = - \int_{\partial G_i} V \nabla' \frac{1}{R} \times \mathbf{dS}' \quad (1.41)$$

Здесь  $d\mathbf{S}' = \mathbf{n}(\mathbf{r}') dS'$ ;  $\mathbf{n}(\mathbf{r}')$  — вектор внешней нормали к границе  $i$ -той области  $\partial G_i$ . Комбинируя слагаемые на каждой границе  $S_{ij}$ , получим

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) - \frac{\mu_0}{4\pi} \sum_{i,j}' (\sigma_i - \sigma_j) \int_{S_{ij}} V(\mathbf{r}') \frac{\mathbf{R}}{R^3} \times \mathbf{dS}'_{ij} \quad (1.42)$$

Из этого уравнения легко увидеть, что при расчете магнитного поля объемные токи могут быть заменены эквивалентными поверхностными токами  $-(\sigma_i - \sigma_j)V(\mathbf{r}')\mathbf{n}_{ij}(\mathbf{r}')$  на каждой из границ между областями однородной проводимости  $S_{ij}$ .

Чтобы получить  $\mathbf{B}$  из уравнения 1.42, необходимо знать значения  $V$  на  $S_{ij}$ . Пользуясь теоремой Грина  $\int_G (\varphi \nabla^2 \psi - \psi \nabla^2 \varphi) dv = \int_S (\varphi \nabla \psi - \psi \nabla \varphi) \cdot d\mathbf{S}$ , получим:

$$\begin{aligned} \sum_{i=1}^m \sigma_i \int_{G_i} \left[ \frac{1}{R} \nabla'^2 V - V \nabla'^2 \frac{1}{R} \right] dv' &= \\ = \sum'_{ij} \int_{S_{ij}} \left[ \sigma_i \left[ \frac{1}{R} \nabla'_i V - V \nabla' \frac{1}{R} \right] - \sigma_j \left[ \frac{1}{R} \nabla'_j V - V \nabla' \frac{1}{R} \right] \right] \cdot d\mathbf{S}'_{ij} \quad (1.43) \end{aligned}$$

Здесь  $\nabla'_i$  обозначает взятие градиента на границе внутри области с индексом  $i$ . Так как плотность тока  $\sigma \nabla V \cdot \mathbf{n}$  непрерывна на  $S_{ij}$ , правая часть уравнения 1.43 приводится к

$$- \sum'_{ij} (\sigma_i - \sigma_j) \int_{S_{ij}} V \nabla' \frac{1}{R} \cdot d\mathbf{S}'_{ij} \quad (1.44)$$

Для преобразования левой части уравнения сначала применим оператор дивергенции к уравнению 1.33:

$$\nabla \cdot \mathbf{J}^p = \nabla \cdot \mathbf{J} + \nabla \cdot (\sigma \nabla V) \quad (1.45)$$

Далее, используя уравнение 1.32 и тот факт, что дивергенция ротора равна нулю, получим, что для квазистатического приближения  $\nabla \cdot \mathbf{J} = 0$ , а значит:

$$\nabla \cdot (\sigma \nabla V) = \nabla \cdot \mathbf{J}^p \quad (1.46)$$

Теперь, используя для левой части уравнения 1.43 уравнение 1.46, а также равенство  $\nabla'^2 \frac{1}{R} = -4\pi\delta(\mathbf{R})$ , получим:

$$V(\mathbf{r})\sigma(\mathbf{r}) = \sigma_0 V_0(\mathbf{r}) - \frac{1}{4\pi} \sum'_{ij} (\sigma_i - \sigma_j) \int_{S_{ij}} V(\mathbf{r}') \frac{\mathbf{R}}{R^3} \cdot d\mathbf{S}'_{ij} \quad (1.47)$$

где

$$V_0(\mathbf{r}) = -\frac{1}{4\pi\sigma_0} \int_G \mathbf{J}^p(\mathbf{r}') \cdot \nabla' \frac{1}{R} dv' = \frac{1}{4\pi\sigma_0} \int_G \frac{\nabla' \cdot \mathbf{J}^p}{R} dv' \quad (1.48)$$

Чтобы получить интегральное уравнение для величины  $V$  на поверхностях  $S_{ij}$ , устремим  $\mathbf{r}$  к точке на поверхности  $S_{ij}$ . Используя предельное значение, [48]

$$\lim_{\mathbf{r}_0 \rightarrow \mathbf{r} \in S_{ij}} \int_{S_{ij}} V(\mathbf{r}') \frac{\mathbf{r}_0 - \mathbf{r}'}{|\mathbf{r}_0 - \mathbf{r}'|^3} \cdot d\mathbf{S}_{ij} = -2\pi V(\mathbf{r}) + \int_{S_{ij}} V(r') \frac{\mathbf{R}}{R^3} \cdot d\mathbf{S}'_{ij} \quad (1.49)$$

получим итегральное уравнение для  $V(\mathbf{r})$ :

$$(\sigma_i + \sigma_j)V(\mathbf{r}) = 2\sigma_0 V_0(\mathbf{r}) + \frac{1}{2\pi} \sum'_{ij} \int_{S_{ij}} V(\mathbf{r}') d\Omega_r(\mathbf{r}') \quad (1.50)$$

где  $\mathbf{r} \in S_{ij}$ , а  $d\Omega_r(\mathbf{r}') = -|\mathbf{r} - \mathbf{r}'|^{-3}(\mathbf{r} - \mathbf{r}') \cdot d\mathbf{S}'_{ij}$  —телесный угол, под которым дифференциальный элемент поверхности  $d\mathbf{S}'_{ij}$  в точке  $\mathbf{r}'$  виден из точки  $\mathbf{r}$ .

Соотношение выше представляет собой линейное интегральное уравнение на величину потенциала электрического поля  $\sigma(\mathbf{r})$  на поверхностях разрыва функции  $\sigma(\mathbf{r})$ . Одной из таких поверхностей является, в частности, поверхность кожи головы, на которой мы меряем разность потенциалов при записи ЭЭГ. Таким образом, уравнение 1.50 позволяет нам получить модель данных для ЭЭГ измерений. Кроме того, восстановив значения потенциала  $V$  на поверхностях  $S_{ij}$  и подставив их в уравнение 1.42, можем найти значение вектора магнитной индукции за пределами головы, вследствие чего получим модель измерений для МЭГ.

В общем случае уравнения 1.42 и 1.50 нужно решать численно, однако оказывается, что для уравнения 1.42 описывающего поведение вектора магнитной индукции, возможно дополнительное упрощение, позволяющее получить аналитическое выражение для неизвестной величины  $\mathbf{B}(\mathbf{r})$ . Суть упрощающего предположения состоит в том, что объем головы приближенно полагается сферой.

### 1.2.5 Случай сферически симметричного проводника.

Рассмотрим уравнение 1.42 для сферически-симметричного проводника. Для простоты предположим, что начало координат совпадает с центром симметрии сферы, аппроксимирующей объем головы. Во-первых, заметим, что

радиальная компонента вектора магнитной индукции  $B_r = \mathbf{B}(\mathbf{r})\mathbf{e}_r = \mathbf{B}(\mathbf{r})\frac{\mathbf{r}}{|\mathbf{r}|}$  не чувствительна к вторичным токам, так как

$$(\mathbf{r} - \mathbf{r}') \times n(\mathbf{r}') \cdot \mathbf{e}_r = (\mathbf{r} - \mathbf{r}') \times \frac{\mathbf{r}'}{|\mathbf{r}'|} \cdot \frac{\mathbf{r}}{|\mathbf{r}|} = 0 \quad (1.51)$$

Следовательно, радиальная компонента вектора магнитной индукции может быть рассчитана из уравнения 1.38 простой заменой  $\mathbf{J}$  на  $\mathbf{J}^p$ :

$$B_r = -\frac{\mu_0}{4\pi} \int \mathbf{J}^p(\mathbf{r}') \times \frac{\mathbf{R}}{R^3} \cdot \mathbf{e}_r dv' \quad (1.52)$$

Для источников, имеющих вид токового диполя соотношение для определения вектора магнитной индукции за пределами головы приобретает совсем простой вид:

$$B_r = -\frac{\mu_0}{4\pi} \frac{\mathbf{Q} \times \mathbf{r}_Q \cdot \mathbf{e}_r}{|\mathbf{r} - \mathbf{r}_Q|^3} \quad (1.53)$$

Кроме того, так как  $\nabla \mathbf{B} = 0$  за пределами проводника, значение величины  $\mathbf{B}$  может быть получено из величины скалярного потенциала  $\mathbf{B} = -\mu_0 \nabla U$ . Так как  $\nabla \cdot \mathbf{B} = 0$ ,  $U$  — гармоническая функция и следовательно может быть однозначно восстановлена по своей нормальной производной на границе проводника при условии, что она обращается в ноль на бесконечности. Для сферы нормальная производная  $\partial U / \partial \mathbf{r} = -B_r / \mu_0$ . Проинтегрировав это соотношение, получим:

$$U(\mathbf{r}) = \frac{1}{\mu_0} \int_{t=1}^{\infty} B_r(t\mathbf{r}) dt \quad (1.54)$$

Таким образом, в случае сферической симметрии проводника, зная радиальную компоненту вектора магнитной индукции мы можем восстановить скалярный потенциал  $U$ , а значит и весь вектор  $\mathbf{B}$ . Иными словами, для сферически симметричного проводника магнитное поле за его пределами целиком определяется его радиальной компонентой  $B_r$ .

Отметим одно важное следствие соотношения 1.52. В случае когда первичный ток имеет радиальное направление,  $\mathbf{J}^p = J^p(\mathbf{r}')\mathbf{r}'/|\mathbf{r}'|$ ,  $B_r$  становится равным нулю. Из  $B_r = 0$  следует  $U = 0$ , а значит и  $\mathbf{B} = 0$ . Это соотношение остается справедливым для любого осесимметричного тока в осесимметричном проводнике. Кроме того, если источник первичного тока находится в центре

сферы  $B_r$  также будет равно нулю в силу  $\mathbf{r}_Q = 0$ . Вообще, чем ближе источник первичного тока к центру сферы, тем меньше вклад от него в суммарное магнитное поле за пределами проводника.

Таким образом, для сферически симметричного проводника МЭГ-сенсоры чувствительны только к тангенциальным компонентам первичных токов, а также нечувствительны к глубоким источникам. Для ЭЭГ, тем не менее, такое заключение оказывается неверным.

### 1.2.6 Матрица прямой модели и поля чувствительности для ЭЭГ и МЭГ

Из уравнений 1.42 и 1.50, видно, что связь первичных токов с величиной электрического потенциала и вектором магнитной индукции линейна. Следовательно, сигнал записанный отдельным электродом ЭЭГ или МЭГ также должен зависеть от первичных токов по линейному закону:

$$b_i = \int \mathcal{L}_i^b(\mathbf{r}) \cdot \mathbf{J}^p(\mathbf{r}) dv, \quad (1.55)$$

$$V_i = \int \mathcal{L}_i^E(\mathbf{r}) \cdot \mathbf{J}^p(\mathbf{r}) dv \quad (1.56)$$

величины  $\mathcal{L}_i^b, \mathcal{L}_i^E$  при этом называются полями чувствительности  $i$ -го сенсора при измерении магнитного поля в случае МЭГ и разности потенциалов для ЭЭГ. Чувствительность сенсора к первичному току в точке пространства  $\mathbf{r}$ , вообще говоря, зависит от направления первичного тока в этой точке. Вместе с тем, из уравнений 1.42, 1.50 следует, что связь между чувствительностями к разным ориентациям первичного тока для фиксированной точки  $\mathbf{r}$  линейна. Следовательно, чтобы определить вклад в измерения на фиксированном сенсоре от произвольно ориентированного первичного тока в фиксированной точке пространства  $\mathbf{r}$ , достаточно знать лишь чувствительности сенсора к току в этой точке вдоль трех независимых направлений. Математически это соответствует тому факту, что поле чувствительности в точке  $\mathbf{r}$  является трехмерным вектором, каждая из координат которого определяет чувствительность сенсора к току, текущему вдоль соответствующего направления, а вклад в измерения от

первичного тока в точке  $\mathbf{r}$  входит в интеграл в уравнениях выше как скалярное произведение вектора первичного тока на вектор поля чувствительности. Дополнительно, поля чувствительности для МЭГ и ЭЭГ зависят от профиля проводимости  $\sigma = \sigma(\mathbf{r})$  в объеме головы а также от конструкции сенсора.

При моделировании первичных токов токовыми диполями будем иметь (см. 1.37):

$$\mathbf{J}^p = \sum_{d=1}^D \mathbf{Q}_d \delta(\mathbf{r} - \mathbf{r}_d) \quad (1.57)$$

где  $D$  — общее число активных токовых диполей;  $\mathbf{r}_d$  — координаты токового диполя с индексом  $d$ ;

Уравнения 1.56 тогда преобразуются к виду

$$B_i = \sum_{d=1}^D \mathcal{L}_i^b(\mathbf{r}_{Q_d}) \cdot \mathbf{Q}_d \quad (1.58)$$

$$V_i = \sum_{d=1}^D \mathcal{L}_i^E(\mathbf{r}_{Q_d}) \cdot \mathbf{Q}_d \quad (1.59)$$

Таким образом, и в ЭЭГ и в МЭГ модальности регистрации электрической активности мозга для отдельного сенсора регистрируемый сигнал есть взвешенная сумма активаций отдельных токовых диполей. При этом веса в этой сумме определяются чувствительностью сенсора к активации в точке  $\mathbf{r}_d$ .

Отметим также, что так как вклад в итоговое измерение на сенсоре от отдельного диполя входит в сумму как скалярное произведение, каждый из произвольно ориентированных токовых диполей может быть эквивалентно представлен в виде линейной комбинации трех ортогональных токовых диполей  $Q_d^x, Q_d^y, Q_d^z$ . Вычислив значения чувствительностей сенсоров по направлениям этих диполей  $\mathcal{L}_{i,x}, \mathcal{L}_{i,y}, \mathcal{L}_{i,z}$ , получим

$$B_i = \sum_{d=1, k \in \{x, y, z\}}^D \mathcal{L}_{i,k}^b(\mathbf{r}_{Q_d}) \cdot \mathbf{Q}_d^k \quad (1.60)$$

$$V_i = \sum_{d=1, k \in \{x, y, z\}}^D \mathcal{L}_{i,k}^E(\mathbf{r}_{Q_d}) \cdot \mathbf{Q}_d^k \quad (1.61)$$

Или, в матричной форме для всего массива сенсоров:

$$\mathcal{B} = \mathcal{G}^b \mathcal{Q} \quad (1.62)$$

$$\mathcal{V} = \mathcal{G}^E \mathcal{Q} \quad (1.63)$$

где  $\mathcal{B}$  — вектор измерений радиальной компоненты магнитного поля на каждом из МЭГ сенсоров,  $\mathcal{V}$  — аналогичный вектор для ЭЭГ,  $\mathcal{Q} = (Q_1^x, Q_1^y, Q_1^z, \dots, Q_D^x, Q_D^y, Q_D^z)^T$ , и, если измерения производятся при помощи  $N$  сенсоров,

$$\mathcal{G}^b \stackrel{\text{def}}{=} \begin{bmatrix} \mathcal{L}_{1,x}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{1,y}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{1,z}^b(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{1,x}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{1,y}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{1,z}^b(\mathbf{r}_{Q_D}) \\ \mathcal{L}_{2,x}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{2,y}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{2,z}^b(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{2,x}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{2,y}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{2,z}^b(\mathbf{r}_{Q_D}) \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \mathcal{L}_{N,x}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{N,y}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{N,z}^b(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{N,x}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{N,y}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{N,z}^b(\mathbf{r}_{Q_D}) \end{bmatrix} \quad (1.64)$$

$$\mathcal{G}^E \stackrel{\text{def}}{=} \begin{bmatrix} \mathcal{L}_{1,x}^E(\mathbf{r}_{Q_1}), & \mathcal{L}_{1,y}^E(\mathbf{r}_{Q_1}), & \mathcal{L}_{1,z}^E(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{1,x}^E(\mathbf{r}_{Q_D}), & \mathcal{L}_{1,y}^E(\mathbf{r}_{Q_D}), & \mathcal{L}_{1,z}^E(\mathbf{r}_{Q_D}) \\ \mathcal{L}_{2,x}^E(\mathbf{r}_{Q_1}), & \mathcal{L}_{2,y}^E(\mathbf{r}_{Q_1}), & \mathcal{L}_{2,z}^E(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{2,x}^E(\mathbf{r}_{Q_D}), & \mathcal{L}_{2,y}^E(\mathbf{r}_{Q_D}), & \mathcal{L}_{2,z}^E(\mathbf{r}_{Q_D}) \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \mathcal{L}_{N,x}^E(\mathbf{r}_{Q_1}), & \mathcal{L}_{N,y}^E(\mathbf{r}_{Q_1}), & \mathcal{L}_{N,z}^E(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{N,x}^E(\mathbf{r}_{Q_D}), & \mathcal{L}_{N,y}^E(\mathbf{r}_{Q_D}), & \mathcal{L}_{N,z}^E(\mathbf{r}_{Q_D}) \end{bmatrix} \quad (1.65)$$

В такой формулировке матрицы  $\mathcal{G}^b, \mathcal{G}^E$  называются *матрицами прямой модели* для МЭГ и ЭЭГ соответственно; при этом столбцы этих матриц принято называть *топографиями*. Три топографии в каждой точке обхема мозга  $\mathbf{r}$  образуют трехмерное линейное пространство в силу линейности поля чувствительности при фиксированном  $\mathbf{r}$ .

Как уже было отмечено выше, в случае сферической симметрии вклад от радиальных источников в магнитное поле за пределами головы равен нулю, и следовательно для МЭГ линейное пространство, натянутое на три топографии, взятые в одной точке, является на самом деле двумерным. Таким образом, принимая модель сферической симметрии для МЭГ, можем переписать матрицу прямой модели в виде

$$\mathcal{G}_s^b = \begin{bmatrix} \mathcal{L}_{1,x'}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{1,y'}^b(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{1,x'}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{1,y'}^b(\mathbf{r}_{Q_D}), \\ \mathcal{L}_{2,x'}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{2,y'}^b(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{2,x'}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{2,y'}^b(\mathbf{r}_{Q_D}), \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \mathcal{L}_{N,x'}^b(\mathbf{r}_{Q_1}), & \mathcal{L}_{N,y'}^b(\mathbf{r}_{Q_1}), & \cdots & \mathcal{L}_{N,x'}^b(\mathbf{r}_{Q_D}), & \mathcal{L}_{N,y'}^b(\mathbf{r}_{Q_D}) \end{bmatrix} \quad (1.66)$$

Штрихами в уравнении выше отмечены новые компоненты полей чувствительности, так как для сокращения размерности локальных подпространств топографий с 3 до 2 они должны быть, вообще говоря, пересчитаны для других направлений первичных токов.

### **1.2.7 Шум сенсоров и шум мозга. Вызванная и индуцированная активность.**

Рассуждения, в ходе которых мы определили матрицы прямых моделей для ЭЭГ и МЭГ, а также построили порождающую модель данных при неинвазивных изменениях, справедливы в том случае, если условия проведения измерений были идеальны, то есть проходили в условиях отсутствия шума.

Однако при проведении реальных экспериментов, предполагающих применение инструментария неинвазивной электрофизиологии, приходится считаться с шумовым загрязнением сигнала, сопровождающим любые попытки получить количественную оценку тех или иных параметров нейрональной активации.

При этом источники шума для ЭЭГ/МЭГ имеют различную природу и как правило представлены совместно. Для построения корректной модели данных в присутствии шумовых источников необходимо определить само понятие шумового сигнала для ЭЭГ/МЭГ измерений, классифицировать возможные источники шума и промоделировать каждый из возможных источников.

Шумом мы будем называть ту часть записанного при помощи ЭЭГ/МЭГ сигнала, которая не имеет отношения к изучаемому явлению. В частности, шумом мы будем называть спонтанную активность мозга, которая напрямую не связана с мозговой функцией, изучаемой в эксперименте.

Данное определение предполагает наличие некоторого рецепта, следуя которому мы могли бы отделить шум от полезного сигнала. Разумеется, когда речь заходит об изучении активности мозга, весьма затруднительно однозначно указать, какая часть записанного сигнала является следствием манипуляций экспериментатора с испытуемым (например, предъявление стимула) или биомаркером изучаемого заболевания. Тем не менее, использование методов теории

оценивания позволяет с определенной долей уверенности говорить о взаимосвязи определенных статистик, полученных на основании измерений, с изучаемым физиологическим явлением.

Рассмотрим теперь различные возможные причины появления шума в записанных данных. Во-первых, к источникам шума стоит отнести шум самих сенсоров — неточности измерений сигнала, носящие случайных характер и некоррелированные между сенсорами. Такой шум будет иметь диагональную матрицу ковариации по сенсорам, которую методами калибровки можно привести к единичной, вследствие чего его принято называть *пространственно некоррелированным* или *пространственно белым*.

Во-вторых к шумовому сигналу мы будем относить любой сигнал, записанный массивом сенсоров совместно и имеющий источник за пределами объема головного мозга испытуемого. Такие сигналы как правило называют *артефактными*. К артефактным сигналам относят источники электромагнитных помех вблизи регистрирующего устройства (например, линия тока бытовой электросети, дающая помеху на частоте 50 или 60 Герц в зависимости от стандарта, принятого в стране, в которой производились измерения) а также биологические источники электрохимической активности, такие как биение сердца, движения глаз, сокращение мышц тела (в особенности, головы и лица). Артефактный сигнал будет с необходимостью коррелирован между сенсорами.

В-третьих, к шумовому сигналу мы будем относить так называемый *шум мозга*, т.е. сигнал, производимый непосредственно нейронными ансамблями мозга испытуемого, но не относящийся непосредственно к физиологическому явлению, изучаемому в эксперименте. При этом шумовая компонента мозгового сигнала определяется по-разному в зависимости от процедуры обработки данных эксперимента и как правило связана с различными техниками усреднения по повторениям. Так, в случае изучения вызванных потенциалов сигналы выравниваются по метке предъявления стимула и усредняются во временной области. В результате такой процедуры сигнал, не привязанный по фазе к моменту предъявления стимула, считается шумовым и затухает с увеличением числа повторений, тогда как активность, привязанная по фазе к стимулу, подчеркивается в сравнении с ослабленным шумовым фоном. Получающуюся после усреднения по ансамблю повторений временную последовательность называют в этом случае *вызваным ответом*.

С другой стороны, при изучении осцилляций вместо усреднения временных последовательностей как таковых часто пользуются усреднением абсолютных значений их частотно-временных образов, имея в виду, что ответ мозга на предъявление стимула может возникать в виде осцилляций, которые не привязаны по фазе к моменту предъявления стимула, но привязаны по амплитуде, то есть усиливаются или ослабляются в ответ на стимул. Итоговая временная последовательность при такой обработке называется *индуцированной активностью*. В качестве мозгового шума здесь выступает активность, амплитуда которой не привязана по фазе к стимулу. При подсчете индуцированной активности из отдельных временных последовательностей для каждого повторения часто вычтывают вызванный ответ, чтобы разделить активность, также привязанную к стимулу по фазе от активности, привязанной только по амплитуде.

Похожая стратегия была использована нами при определении функции когерентности двух случайных процессов при подсчете матрицы кросс-спектральной плотности сигнала (1.10) с тем отличием, что в случае матрицы кросс-спектральной плотности мы усредняли не модули коэффициентов частотно-временного преобразования, а сами эти коэффициенты. Дополнительно отметим, что при подсчете кросс-спектральной плотности мы также вычитали вызванный ответ в момент вычисления кросс-ковариации (см. 1.2). При изучении когерентности сигналов, полученных при неинвазивных измерениях электрической активности мозга дополнительным источником шума, препятствующего обнаружению сетей служит *эффект протечки сигнала*, который мы определим и подробно обсудим в главе 2.

Нетрудно понять, что шум мозга, также как и артефактный сигнал, будет коррелирован между сенсорами. Отличие состоит в том, что для мозгового шума структура матрицы кросс-ковариации между сенсорами неслучайна и существенно зависит от структуры матрицы прямой модели, посчитанной нами на этапе решения прямой задачи для ЭЭГ/МЭГ. Знание о пространственной структуре мозгового шума, как мы увидим ниже, может быть использовано для ослабления соответствующей шумовой компоненты в данных при помощи методов пространственной фильтрации.

### 1.3 Методы оценки активации нейронных популяций.

Задачу оценки величин, положений и ориентаций токовых диполей  $\mathcal{Q}$  на основании измерений  $\mathcal{B}, \mathcal{V}$  в электрофизиологии принято называть *обратной задачей* МЭГ/ЭЭГ. Как и в случае оценки кросс-спектральных коэффициентов, решение обратной задачи МЭГ/ЭЭГ не единственно. Для выбора какого-либо одного решения  $\mathcal{Q}$  по коре применяют различные эвристики, ограничивающие выбор из бесконечного множества возможных вариантов. Как правило, получающееся решение отвечает тому или иному критерию оптимальности в соответствии с используемой эвристикой, при условии, что предположения модели выполняются. В этом разделе мы рассмотрим основные методы решения обратной задачи для поиска активных токовых диполей на коре, а затем перейдем к рассмотрению методик для оценки кросс-спектральных коэффициентов с учетом свободной ориентации.

Все существующие методики решения обратной задачи МЭГ/ЭЭГ можно условно разделить на два класса. К первому классу относятся алгоритмы, основанные на поиске заранее заданного числа эквивалентных токовых диполей, объясняющих измерения наилучшим образом. При этом оцениваются положения и ориентации этих диполей. К этому классу относится необходимый нам для дальнейшего изложения алгоритм MUSIC (multiple signal classification), [49; 50], а также его модификация RAP-MUSIC (recursively applied and projected MUSIC), [51], которые условно можно назвать сканирующими.

Второй класс алгоритмов, называемых в литературе «имиджинговыми», ставит задачу отыскания активности, распределенной сразу по всей коре. Ко второй группе методов можно условно отнести методы оптимальной пространственной фильтрации, которые восстанавливают сигнал отдельно для каждой выбранной точки коры в соответствии с неким локальным критерием оптимальности, а также методы, основанные на выборе решения с минимальной нормой.

Суть подхода оптимальной фильтрации состоит в том, что для фиксированной точки внутри объема мозга ставится задача нахождения пространственного фильтра, оптимизирующего определенную характеристику сигнала, восстанавливаемого при помощи этого фильтра. В качестве такой характеристики может выступать, например, отношение сигнал/шум, или же мы можем

руководствоваться критерием минимизации протечки сигнала от других источников в точку, в которой мы хотим восстановить активность. Здесь важно отметить, что конкретный вид решения, полученного в результате оптимизации выбранного функционала качества будет зависеть также от предполагаемой пространственной структуры шума.

Отметим, что структура восстановленной после применения совокупности найденных фильтров активации на коре при таком подходе, вообще говоря, субоптимальна с точки зрения объяснения сигнала, измеренного сенсорами (так как мы оптимизировали другой функционал качества). Проблема недопределенности системы уравнений при этом в некотором смысле остается за скобками, так как для каждой точки коры решение восстанавливается индивидуально — без учета вклада в решение активаций, восстановленных в других точках коры. Таким образом, для алгоритмов оптимальной фильтрации найденное решение является оптимальным в локальном, но не в глобальном смысле.

Задача отыскания активаций, наилучшим образом объясняющих измерения, (т.е. оптимальных в глобальном смысле) ставится для другого подкласса имиджинговых методов. При этом, как уже было отмечено выше, в силу недопределенности системы линейных уравнений, связывающих активации на коре с сигналом на сенсорах, существует бесконечное множество конфигураций источников, идеально объясняющих померенный сигнал. Тем не менее, среди таких решений в силу зашумленности измерений а также неточностей при построении прямой модели реальное распределение активаций (в выделенных точках) коры не содержится, так как эти «идеальные» с точки зрения объяснения измерений решения объясняют в том числе и шумовую компоненту, которая зачастую оказывается больше или сравнима по амплитуде с истинной активацией.

Итак, при решении обратной задачи методами глобальной оптимизации существует две проблемы: бесконечное множество возможных решений и зашумленность измерений. Чтобы справиться с первой проблемой, для выбора из бесконечного множества решений пользуются критерием минимальности нормы решения. Иными словами, среди всех возможных конфигураций первичных токов в объеме (или на поверхности коры) мозга в качестве решения выбирается такая конфигурация, норма которой минимальна среди всех возможных. Условие минимальности нормы в некотором смысле является следствие принципа бритвы Оккама: мы ищем наиболее простое решение, удовлетворяющее

наблюдениям. Какие решения при этом считать простыми — неочевидный вопрос. Ответ на него зависит от выбора конкретного вида нормы, которую мы хотим минимизировать. Наиболее популярными вариантами являются  $L_2$ - и  $L_1$ -нормы. Наиболее простой пример минимальной  $L_2$ -нормы решения соответствует случаю, когда проблему зашумленности данных мы оставляем без внимания. Тогда решение, соответствующее минимуму  $L_2$ -нормы, получается применением оператора, соответствующего псевдообратной матрице, взятой для матрицы прямой модели. Для  $L_1$ -нормы ситуация несколько сложнее, так как решение не может быть получено в явном виде, и требуется численная оптимизация соответствующего функционала, сводящаяся к задаче линейного программирования.

Рассмотрим теперь, каким образом решается проблема зашумленности данных. Здесь вновь существует два подхода. Первый из них используется значительно реже и состоит в удалении шумовой компоненты из данных посредством сокращенного сингулярного разложения матрицы прямой модели. Такой подход, например, использовали авторы, метода Minimum Current Estimate (MCE) [41], порождающего решения с минимальной  $L_1$ -нормой.

Другой, более популярный подход состоит в использовании тихоновской регуляризации. В рамках этого подхода норма решения и глобальная ошибка в объяснении измерений минимизируются совместно, как части одного общего функционала качества, позволяя тем самым соблюсти баланс между простотой решения и тем, насколько хорошо оно объясняет измерения, (в том числе, содержащийся в них шум). Соотношение между «простотой» решения и величиной ошибки при таком подходе можно регулировать настраивая величину метапараметра, называемого параметром регуляризации. Меняя значение параметра регуляризации мы стремимся найти такое значение, при котором полученное решение объясняет только «полезную» часть сигнала, записанного сенсорами и полностью игнорирует шумовую компоненту.

Отметим, что с точки зрения байесовской статистики тихоновская регуляризация эквивалентна нахождению такого решения обратной задачи, которое соответствует точке максимума апостериорной плотности вероятности. Минимизируемая норма решения в такой интерпретации задается априорной плотностью распределения вероятности.

Среди методов, основанных на тихоновской регуляризации, отметим прежде всего Minimum Norm Estimate (MNE) [40], минимизирующий  $L_2$ -норму

решения, и его вариацию — метод dSPM, нормирующий величину восстановленного значения первичного тока в каждой точке на оцененную величину шума в ней же.

### 1.3.1 Сканирующие алгоритмы

Наиболее естественным алгоритма поиска фиксированного числа эквивалентных токовых диполей (dipole fitting), объясняющих данные наилучшим способом, является оптимизация их положений и ориентаций методом наименьших квадратов. Такой подход, однако, обладает существенными недостатками, к которым относится невыпуклость целевой функции при такой оптимизации, что приводит к застреванию алгоритма в локальных минимумах.

## MUSIC

Чтобы обойти эту проблему, Мошер и Лихи предложили использовать для поиска активных токовых диполей алгоритм MUSIC [49; 50], разработанный и использовавшийся ранее в радиопеленгации и сонарах.

Рассмотрим подробно суть метода в применении к данным ЭЭГ/МЭГ. Начнем с рассмотрения порождающей модели сигнала на сенсорах, как мы это уже делали для оценки фазовой синхронности (см. 2.1). На этот раз, однако, заложим в модель возможность свободной ориентации диполей. От каждого активного токового диполя на коре будем иметь вклад на сенсорах вида:

$$\mathbf{x}_k(t) = \begin{bmatrix} | & | & | \\ \mathbf{g}_k^1 & \mathbf{g}_k^2 & \mathbf{g}_k^3 \\ | & | & | \end{bmatrix} \begin{pmatrix} s_{k,1}(t) \\ s_{k,2}(t) \\ s_{k,3}(t) \end{pmatrix} \quad (1.67)$$

, где  $k$  — индекс токового диполя  $s_{k,i}$  — компоненты соответствующего дипольного момента,  $\mathbf{g}_k^i$  — вектора-топографии  $k$ -го токового диполя для трех

ориентаций тока ( $i = 1, 2, 3$ ). Тогда вклад от всех активных токовых диполей будет виден на сенсорах как

$$\mathbf{x}(t) = \mathbf{Gs}(t) + \boldsymbol{\omega}(t), \quad (1.68)$$

Здесь вновь  $\mathbf{s}(t)$  —  $3n$ -мерный вектор-столбец активаций источников,  $\mathbf{x}(t)$  —  $m$ -мерный вектор-столбец сигналов на сенсорах,  $t$  — время, а  $\mathbf{G}$  —  $m \times 3n$  матрица линейного отображения пространства источников пространство сенсоров.

Уравнение 1.68 задает соответствие между пространством источников и пространством сенсоров для каждого временного среза  $t$ . При условии, что было записано  $T$  таких срезов, можем переписать уравнение 1.68 в матричной форме:

$$\mathbf{X} = \mathbf{GS} + \boldsymbol{\Omega} \quad (1.69)$$

Заметим, что столбцы матрицы  $\mathbf{X}$  порождаются линейными комбинациями векторов-топографий активных токовых диполей, а значит все возможные конфигурации наблюдаемого сигнала живут внутри некоторого линейного подпространства линейной оболочки этих топографий. Это линейное подпространство называется *подпространством сигнала*. При этом количество активных токовых диполей  $r$  задает размерность этого подпространства. Чтобы выделить его, применим к матрице  $\mathbf{X}$  сингулярное разложение и зафиксируем первые  $r$  левых собственных векторов:

$$\mathbf{U}, \mathbf{S}, \mathbf{V} = svd(\mathbf{X}) \quad (1.70)$$

$$\mathbf{U} = \begin{bmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_m \\ | & & | \end{bmatrix} \quad (1.71)$$

$$\mathbf{U}_r = \begin{bmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_r \\ | & & | \end{bmatrix} \quad (1.72)$$

Матрица  $\mathbf{U}_r$  называется *матрицей подпространства сигнала*; ее столбцы задают ортонормальный базис этого подпространства. Чтобы найти все активные токовые диполи, для каждого источника на коре вычисляют *корреляцию подпространств* между подпространством сигнала и линейной оболочкой трех (двух в случае МЭГ) топографий, соответствующих данному источнику. Корреляция подпространств  $C_k$  вычисляется как максимальное собственное число

произведения матрицы левых собственных векторов для топографий  $k$ -го источника и матрицы подпространства сигнала:

$$\mathbf{U}_{k,g}, \mathbf{S}_{k,g}, \mathbf{V}_{k,g} = svd \left( \begin{bmatrix} | & | & | \\ \mathbf{g}_k^1 & \mathbf{g}_k^2 & \mathbf{g}_k^3 \\ | & | & | \end{bmatrix} \right) \quad (1.73)$$

$$C_k = \lambda_{max}(\mathbf{U}_{k,g} \mathbf{U}_r^T) \quad (1.74)$$

Активными считаются такие токовые диполи, для которых величина  $C_k$  выше некоторого заранее порога (авторы статьи [49] рекомендуют для порога значение 0.95).

Таким образом, для нахождения всех активных токовых диполей необходимо «просканировать» объем или поверхность мозга на предмет источников, для которых корреляция подпространств с подпространством сигнала превышает некоторое пороговое значение.

Одним из существенных недостатков подхода MUSIC является потенциальная сложность в разделении нескольких одновременно активных токовых диполей, а также в выделении ситуации, когда реальный источник имел не фокальную, а распределенную структуру. В этом случае в распределении в объеме или по поверхности коры величин  $C_k$  будут присутствовать локальные максимумы, анализ которых требует дополнительных усилий.

## RAP-MUSIC

Чтобы справиться с этой проблемой, Мошер и Лихи предложили обобщение метода MUSIC, которое они назвали Recursively applied and projected MUSIC или RAP-MUSIC [51].

Идея метода заключается в последовательном применении MUSIC-скана для нахождения диполя с максимальным значением  $C$  с последующей проекцией матрицы данных и топографий оставшихся источников ортогонально подпространству топографий диполя с максимальным  $C_k$ , с передующей итерацией.

Процедура продолжается до тех пор, пока максимальное по источникам значение корреляции подпространств на новом шаге не опустится ниже порогового значения.

Описание алгоритма RAP-MUSIC на языке Python приведено в приложении А.

Такая процедура позволяет исключить вклад в измерения от уже найденных диполей, последовательно объясняя оставшуюся в данных дисперсию. Пороговый критерий останова алгоритма гарантирует отсутствие ложноположительных срабатываний.

Отдельного упоминания для алгоритмов MUSIC и RAP-MUSIC заслуживает выбор ранга подпространства сигнала. С одной стороны, этот ранг должен совпадать с количеством реально присутствующих в данных диполей. С другой стороны, *a priori*, до применения алгоритма это число не известно. Хорошая новость заключается в том, что переоценка ранга подпространства сигнала практически не влияет на качество работы этих двух алгоритмов, поэтому авторы статей [49; 51] в качестве общей рекомендации советуют выбирать значения этого ранга больше реально ожидаемых.

Оба алгоритма, будучи достаточно простыми и вычислительно эффективными, являются удобным инструментом для решения обратной задачи в первом приближении, что может быть полезно во-первых для примерного понимания картины распределения источников, а во-вторых в качестве этапа предобработки при решении ОЗ для сужения количества вариантов поиска с последующим анализом более изощренными, но не столь вычислительно быстрыми алгоритмами.

### 1.3.2 Алгоритмы пространственной фильтрации

Другой класс алгоритмов решения обратной задачи основан на методах оптимальной пространственной фильтрации сигнала. Различные критерии оптимальности порождают при этом различные алгоритмы. В общем виде задача построения оптимальных пространственных фильтров для  $k$ -го источника сводится к нахождению матрицы  $\mathbf{A}_k$ , при умножении которой на матрицу данных

$\mathbf{X}$ , получается матрица  $\mathbf{S}_k$  размером  $3 \times T$  временных рядов для трех ортогональных направлений тока в точке  $k$ , отвечающая некоторому критерию оптимальности:

$$f(\mathbf{A}_k \mathbf{X}) \rightarrow \text{opt} \quad (1.75)$$

Проведем обзор существующих критериев оптимальности и получающихся из них методов на основании статьи [52].

### Пространственный фильтр, оптимизирующий отношение сигнал-шум в заданной точке

Начнем рассмотрение семейства методов пространственной фильтрации с метода, максимизирующего отношение сигнал-шум в заданной точке.

Сначала определим отношение сигнал-шум в терминах матриц ковариаций шума и данных. Пусть  $\mathbf{R}_{X,k} = \mathbf{XX}^T$  — матрица ковариации сигнала на сенсорах, приходящего от  $k$ -го источника (предполагается что мы уже вычли среднее из временных рядов в матрице  $\mathbf{X}$ ), а  $\mathbf{R}_n$  — матрица ковариации шума на сенсорах. Тогда отношение сигнал-шум на сенсорах определяется как

$$\rho = \frac{\text{tr}(\mathbf{R}_{X,k})}{\text{tr}(\mathbf{R}_n)} \quad (1.76)$$

Тогда для отфильтрованного при помощи фильтра  $\mathbf{A}_k$  (матрица размером  $m \times 3$ )  $k$ -го источника ОСШ будет выглядеть как

$$\rho_k = \frac{\text{tr}(\mathbf{A}_k \mathbf{R}_{X,k} \mathbf{A}_k^T)}{\text{tr}(\mathbf{A}_k \mathbf{R}_n \mathbf{A}_k^T)} \quad (1.77)$$

Тогда задача построения пространственного фильтра  $\mathbf{A}_k$ , который максимизирует отношение сигнал-шум для  $k$ -го отфильтрованного источника формулируется как

$$\mathbf{A}_k = \underset{\mathbf{A}_k}{\operatorname{argmax}}(\rho_k) = \underset{\mathbf{A}_k}{\operatorname{argmax}} \left( \frac{\text{tr}(\mathbf{A}_k \mathbf{R}_{X,k} \mathbf{A}_k^T)}{\text{tr}(\mathbf{A}_k \mathbf{R}_n \mathbf{A}_k^T)} \right) \quad (1.78)$$

$$s.t. : \mathbf{A}_k \mathbf{A}_k^T = \mathbf{I}_3 \quad (1.79)$$

Ограничение необходимо, чтобы фильтр не менял пространственную структуру шума. Шум будем предполагать пространственно белым с дисперсией  $\sigma_n^2$ . В общем случае это предположение можно удовлетворить предварительным отбеливанием данных, применив к ним оператор  $\mathbf{R}_n^{-\frac{1}{2}}$ .

В одномерном случае, соответствующем фиксированной ориентации диполя, матрица фильтров  $\mathbf{A}_k$  вырождается в вектор-строку, а матрица ковариации для источника с индексом  $k$  приобретает вид

$$\mathbf{R}_{X,k} = \sigma_k^2 \mathbf{g}_k \mathbf{g}_k^T \quad (1.80)$$

где  $\sigma_k^2$  — дисперсия сигнала. Для матрицы ковариации шума в предположении, что он является пространственно белым, будем иметь  $\mathbf{R}_n = \mathbf{I}$ . Тогда 1.79 в одномерном случае преобразуется к виду

$$\mathbf{A}_k = \underset{\|\mathbf{A}_k\|=1}{\operatorname{argmax}} \left( \frac{\sigma_k^2 \|\mathbf{A}_k \mathbf{g}_k\|^2}{\sigma_n^2 \mathbf{A}_k \mathbf{I} \mathbf{A}_k^T} \right) = \frac{\sigma_k^2}{\sigma_n^2} \underset{\|\mathbf{A}_k\|=1}{\operatorname{argmax}} \left( \|\mathbf{A}_k \mathbf{g}_k\|^2 \right) \quad (1.81)$$

Очевидно, что среди всех векторов единичной длины максимальное значение ОСШ будет достигаться на векторе, сонаправленном с ориентацией вектора топографии  $\mathbf{g}_k$ :

$$\mathbf{A}_k = \frac{\mathbf{g}_k^T}{\|\mathbf{g}_k\|} \quad (1.82)$$

В случае свободной ориентации диполя матрица  $\mathbf{R}_{X,k}$  примет вид

$$\mathbf{R}_{X,k} = \sigma_k^2 [\mathbf{g}_k^1, \mathbf{g}_k^2, \mathbf{g}_k^3] [\mathbf{g}_k^1, \mathbf{g}_k^2, \mathbf{g}_k^3]^T \quad (1.83)$$

В этом случае любые три вектора (два в случае МЭГ), образующие ортонормированный базис подпространства, натянутого на вектора  $\mathbf{g}_k^1, \mathbf{g}_k^2, \mathbf{g}_k^3$ , могут быть использованы как компоненты оптимального фильтра. В частности, для построения фильтра  $\mathbf{A}_k$  можем использовать первые три левых сингулярных вектора матрицы  $\mathbf{g}_k$ :

$$\mathbf{U}_{k,g}, \mathbf{S}_{k,g}, \mathbf{V}_{k,g} = svd \left( \begin{bmatrix} | & | & | \\ \mathbf{g}_k^1 & \mathbf{g}_k^2 & \mathbf{g}_k^3 \\ | & | & | \end{bmatrix} \right) \quad (1.84)$$

$$\mathbf{A}_k = \begin{bmatrix} | & | & | \\ \mathbf{U}_{k,g}^1 & \mathbf{U}_{k,g}^2 & \mathbf{U}_{k,g}^3 \\ | & | & | \end{bmatrix} \quad (1.85)$$

Действительно, проекция векторов фильтра на ортогональное дополнение линейной оболочки векторов  $\mathbf{g}_k^1, \mathbf{g}_k^2, \mathbf{g}_k^3$ , при подсчете произведения  $\mathbf{A}_k \mathbf{g}_k \mathbf{g}_k^T \mathbf{A}_k^T$  будет давать нулевой вклад, а значит, так как вектора  $\mathbf{A}_k$  фиксированной длины (единичные), они должны целиком принадлежать линейной оболочке  $span(\mathbf{g}_k^1, \mathbf{g}_k^2, \mathbf{g}_k^3)$ . Вместе с тем, из ограничения  $\mathbf{A}_k \mathbf{A}_k^T = \mathbf{I}_3$  следует, что матрица  $\mathbf{A}_k$  является матрицей перехода от одного ортонормированного базиса линейной оболочки  $span(\mathbf{g}_k^1, \mathbf{g}_k^2, \mathbf{g}_k^3)$  к другому, а произведение  $\mathbf{A}_k \mathbf{g}_k \mathbf{g}_k^T \mathbf{A}_k^T$  реализует этот переход. Но след матрицы является инвариантом при переходе от одного ортонормированного базиса к другому, а значит  $tr(\mathbf{A}_k \mathbf{g}_k \mathbf{g}_k^T \mathbf{A}_k^T)$  будет одинаков для любого  $\mathbf{A}_k$ , строки которого образуют ортонормированный базис  $span(\mathbf{g}_k^1, \mathbf{g}_k^2, \mathbf{g}_k^3)$ .

Рассмотрим теперь способ получения фильтра, оценивающего истинную ориентацию диполя и восстанавливающего сигнал вдоль этой ориентации. После применения полученной матрицы фильтра  $\mathbf{A}_k$  к данным, мы будем получать три временных ряда, соответствующих трем ортогональным направлениям тока в точке  $k$ . Для получения фильтра, восстанавливающего единственный временной ряд вдоль истинной ориентации тока в  $k$ -ой точке также пользуются критерием максимизации ОСШ. В этом случае ставится задача отыскания такой ориентации фильтров  $\boldsymbol{\Theta}_k = (\theta^1, \theta^2, \theta^3)^T$ , для которой отфильтрованный сигнал  $\mathbf{S}_k = \boldsymbol{\Theta}_k^T \mathbf{A}_k \mathbf{X}$  будет обладать максимальной мощностью среди всех возможных ориентаций.

Для нахождения такой ориентации достаточно взять собственный вектор матрицы  $\mathbf{A}_k \mathbf{X} \mathbf{X}^T \mathbf{A}_k^T$ , соответствующий ее максимальному собственному числу. Тогда ориентированный фильтр будет выглядеть как

$$\mathbf{A}_k^{or} = \boldsymbol{\Theta}^T \mathbf{A}_k \quad (1.86)$$

Заметим, что процедура построения такого фильтра очень похожа на способ вычисления корреляции подпространств в алгоритме MUSIC.

## Пространственный фильтр, минимизирующий вклад третих источников в оценку

Рассмотрим теперь пространственный фильтр, оптимизационным критерием которого является задача минимизации вклада активных источников, находящихся в отличных от целевой точках коры. Этот критерий эквивалентен задаче отыскания фильтра, для которого отфильтрованный сигнал будет обладать минимальной энергией при условии, что сигнал из целевой точки фильтром не искажается.

Математически эта задача в случае фиксированной ориентации диполей запишется как

$$\mathbf{E} \left\{ \| \mathbf{a}_k \mathbf{X} \|^2 \right\} \rightarrow \min_{\mathbf{a}_k}, \text{s.t.} : \mathbf{a}_k \mathbf{g}_k = 1 \quad (1.87)$$

Выпишем соответствующий лагранжиан с учетом того, что  $\mathbf{R}_{\mathbf{X}} = \mathbf{E} \{ \mathbf{X} \mathbf{X}^T \}$  — это матрица ковариации сигнала на сенсорах:

$$f_L(\mathbf{a}_k, \alpha) = \mathbf{a}_k \mathbf{R}_{\mathbf{X}} \mathbf{a}_k^T + \alpha (\mathbf{a}_k \mathbf{g}_k - 1) \quad (1.88)$$

Далее, приравняв производную лагранжиана по  $\mathbf{a}_k^T$  к нулю, будем иметь:

$$\frac{\partial f_L}{\partial \mathbf{a}_k} = 2 \mathbf{R}_{\mathbf{X}} \mathbf{a}_k^T + \alpha \mathbf{g}_k = 0 \quad (1.89)$$

Откуда

$$\mathbf{a}_k = -\frac{\alpha}{2} \mathbf{g}_k^T \mathbf{R}_{\mathbf{X}}^{-1} \quad (1.90)$$

Используя ограничение, получим

$$1 = \mathbf{a}_k \mathbf{g}_k = -\frac{\alpha}{2} \mathbf{g}_k^T \mathbf{R}_{\mathbf{X}}^{-1} \mathbf{g}_k \implies \alpha = -\frac{2}{\mathbf{g}_k^T \mathbf{R}_{\mathbf{X}}^{-1} \mathbf{g}_k} \quad (1.91)$$

Окончательно, будем иметь

$$\mathbf{a}_k = \frac{\mathbf{g}_k^T \mathbf{R}_{\mathbf{X}}^{-1}}{\mathbf{g}_k^T \mathbf{R}_{\mathbf{X}}^{-1} \mathbf{g}_k} \quad (1.92)$$

Можно показать, что в случае, если количество активных диполей меньше количества сенсоров, а временные профили активации активных диполей

некоррелированы, такой фильтр будет выдавать нулевые значения для всех точек  $k$ , не совпадающих с положениями активных диполей. Такое поведение фильтра приводит к тому, что в случае, если точки дискретизации коры не совпадают идеально с положениями активных диполей, такие диполи будут не видны при фильтрации. Кроме того, в случае малого количества активных источников такой фильтр крайне плохо работает при наличии шума в данных, так как матрица  $\mathbf{R}_X^{-1}$  в этом случае является вырожденной (ее ранг совпадает с количеством активных диполей).

Чтобы исправить негативные эффекты этого фильтра, пользуются регуляризацией матрицы  $\mathbf{R}_X$ :

$$\mathbf{R}_X^{reg} = \mathbf{R}_X + \alpha \mathbf{I} \quad (1.93)$$

Полученная матрица используется вместо матрицы  $\mathbf{R}_X$  в уравнении 1.92. Подстраивая коэффициент регуляризации  $\alpha$ , можно управлять степенью сглаженности пространственной чувствительности фильтров  $\mathbf{a}_k$ .

Отметим, что решения для фильтров вида 1.92, содержащие матрицу  $\mathbf{R}_X^{reg}$  вместо  $\mathbf{R}_X$  получаются для видоизмененной версии исходной оптимизационной задачи 1.87, которую можно записать как

$$E\left\{ \| \mathbf{a}_k \mathbf{X} \|^2 \right\} + \alpha \| \mathbf{a}_k \|^2 \rightarrow \min_{\mathbf{a}_k}, s.t. : \mathbf{a}_k \mathbf{g}_k = 1 \quad (1.94)$$

Искомые выражения для фильтров  $\mathbf{a}_k$  вновь могут быть получены методом множителей Лагранжа. Второй член в оптимизируемом функционале 1.94 накладывает ограничение на норму искомых фильтров, не давая ей расти слишком сильно и исключая ту ситуацию, когда оптимизация приводит к выбору фильтра, лежащего в подпространстве собственных векторов матрицы  $\mathbf{R}_X$ , соответствующих малым (шумовым) собственным числам. В этом случае угол между фильтром и соответствующей топографией будет близок к  $\pi/2$ , и чтобы удовлетворить ограничению, оптимизация будет компенсировать величину проекции  $\mathbf{a}_k$  на  $\mathbf{g}_k$  за счет увеличения нормы  $\mathbf{a}_k$ .

Пространственные фильтры, использующие матрицу ковариации измерений, обычно принято называть адаптивными (так как они не только используют знания о геометрии взаимного расположения коры и массива сенсоров, но и подстраиваются под текущую активность) или бимформерами (в

дословном переводе — формирователь луча), имея в виду, что они конструируют фильтры, сфокусированные на заданных пространственные направлениях, сродни лучу прожектора.

Рассмотрим теперь ту же самую задачу в векторной постановке. Отметим сначала, что для скоррелированных источников процедура, описанная выше, будет давать ошибочные значения мощностей сигналов в источниках. Чтобы убедиться в этом, рассмотрим, как действуют фильтры на вектора прямой модели в одномерном случае в предположении отсутствия шума, а также считая, что количество активных источников  $q$  меньше количества сенсоров  $m$ . В таких предположениях матрица ковариации сигнала на сенсорах  $\mathbf{R}_X$  размером  $m \times m$  будет иметь ранг  $q$ , и следовательно будет необратимой. Тем не менее, ее можно приблизить псевдообратной матрицей. Тогда, так как  $\mathbf{R}_X = \mathbf{G}\mathbf{R}_S\mathbf{G}^T$

$$\mathbf{R}_X^{-1} \approx (\mathbf{G}\mathbf{R}_S\mathbf{G}^T)^\dagger = (\mathbf{G}^\dagger)^T \mathbf{R}_S^{-1} \mathbf{G}^\dagger = (\mathbf{G}^\dagger)^T \mathbf{R}_S^{-1} \mathbf{G}^\dagger \quad (1.95)$$

Кроме того, так как  $q < m$ , псевдообратная для матрицы  $\mathbf{G}$  будет также ее левой обратной:

$$\mathbf{G}^\dagger \mathbf{G} = ((\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T) \mathbf{G} = \mathbf{I} \quad (1.96)$$

А значит, для отдельных топографий активных диполей будет выполнено

$$\mathbf{G}^\dagger \mathbf{g}_k = \mathbf{1}_k \quad (1.97)$$

где  $\mathbf{1}_k$  обозначает вектор, у которого на  $k$ -ом месте стоит единица, а на всех остальных — нули.

Теперь, пользуясь формулой 1.92, можем расписать произведение  $\mathbf{a}_k \mathbf{g}_l$ :

$$\mathbf{a}_k \mathbf{g}_l = \frac{\mathbf{g}_k^T \mathbf{R}_X^{-1} \mathbf{g}_l}{\mathbf{g}_k^T \mathbf{R}_X^{-1} \mathbf{g}_k} = \frac{(\mathbf{G}^\dagger \mathbf{g}_k)^T \mathbf{R}_S^{-1} \mathbf{G}^\dagger \mathbf{g}_l}{(\mathbf{G}^\dagger \mathbf{g}_k)^T \mathbf{R}_S^{-1} \mathbf{G}^\dagger \mathbf{g}_k} = \frac{\mathbf{1}_k^T \mathbf{R}_S^{-1} \mathbf{1}_l}{\mathbf{1}_k^T \mathbf{R}_S^{-1} \mathbf{1}_l} = \frac{\mathbf{R}_{S,k,l}^{-1}}{\mathbf{R}_{S,k,k}^{-1}} \quad (1.98)$$

Таким образом, на незашумленных измерениях в случае, если матрица  $\mathbf{R}_S$  диагональна (а значит диагональна и  $\mathbf{R}_S^{-1}$ ), то есть источники нескоррелированы, адаптивный пространственный фильтр будет идеально восстанавливать источники, так как фильтр, направленный на  $k$ -ый источник будет полностью подавлять сигнал, приходящий от всех источников, кроме  $k$ -го. В этом случае произведение  $\mathbf{a}_k \mathbf{g}_l = \delta_k^l$ , где  $\delta_k^l$  — дельта Кронекера.

Вместе с тем, если источники обладают коррелированными временными рядами, неизбежно возникает искажение восстановленного сигнала за счет пропущи в точку, на которую направлен фильтр, сигнала от этих коррелированных источников.

В случае, если их положение заранее известно, эту ситуацию можно исправить за счет дополнительных линейных ограничений вида  $\mathbf{a}_k \mathbf{g}_l = 0$  для ограниченного заранее известного набора индексов точек коры  $l$ .

Именно такая ситуация возникает при попытке сформулировать критерий оптимальной адаптивной фильтрации в векторном случае. Действительно, для каждой точки коры мы будем иметь идеально скоррелированные источники по трем направлениям ориентации токовых диполей, так как они будут представлять собой проекции одного и того же вектора плотности тока на три разных направления. С другой стороны, это один из тех редких случаев, когда мы знаем топографии источников с коррелированными временными рядами, поэтому можем явно выписать линейные ограничения на оптимизируемый функционал. Для  $i$ -го из трех фильтров будем иметь:

$$\mathbf{E} \left\{ \|\mathbf{a}_k^i \mathbf{X}\|^2 \right\} \rightarrow \min_{\mathbf{a}_k}, s.t. : \mathbf{a}_k^i \mathbf{g}_k^j = \delta_i^j, j = 1, 2, 3 \quad (1.99)$$

В векторном виде, совокупно для трех фильтров, эта задача может быть переписана в виде:

$$\text{tr} \left\{ \mathbf{A}_k \mathbf{R}_X \mathbf{A}_k^T \right\} \rightarrow \min_{\mathbf{A}_k}, s.t. : \mathbf{A}_k \mathbf{G}_k = \mathbf{I} \quad (1.100)$$

Удобнее, однако, продолжить вывод покомпонентно. Вновь выпишем лагранжиан для 1.99 оптимизационной задачи:

$$f_L(\mathbf{a}_k, \lambda) = \mathbf{a}_k^i \mathbf{R}_X (\mathbf{a}_k^i)^T + \lambda (\mathbf{g}_k^T (\mathbf{a}_k^i)^T - \mathbf{1}_i) \quad (1.101)$$

Дифференцируем и приравниваем к нулю:

$$\frac{\partial f_L}{\partial \mathbf{a}_k^i} = 2 \mathbf{R}_X (\mathbf{a}_k^i)^T + \mathbf{g}_k \lambda^T = 0 \implies (\mathbf{a}_k^i)^T = -\frac{1}{2} \mathbf{R}_X^{-1} \mathbf{g}_k \lambda^T \quad (1.102)$$

Используя ограничение  $\mathbf{g}_k^T (\mathbf{a}_k^i)^T = \mathbf{1}_i$ , выразим  $\lambda$ :

$$\mathbf{g}_k^T (\mathbf{a}_k^i)^T = \mathbf{1}_i = -\frac{1}{2} \mathbf{g}_k^T \mathbf{R}_X^{-1} \mathbf{g}_k \lambda^T \implies \lambda^T = -2 (\mathbf{g}_k^T \mathbf{R}_X^{-1} \mathbf{g}_k)^{-1} \mathbf{1}_i \quad (1.103)$$

Окончательно, для  $i$ -го фильтра будем иметь

$$(\mathbf{a}_k^i)^T = \mathbf{R}_{\mathbf{X}}^{-1} \mathbf{g}_k (\mathbf{g}_k^T \mathbf{R}_{\mathbf{X}}^{-1} \mathbf{g}_k)^{-1} \mathbf{1}_i \quad (1.104)$$

Или, комбинируя все три фильтра в одной векторной записи и транспортируя:

$$\mathbf{A}_k = (\mathbf{G}_k^T \mathbf{R}_{\mathbf{X}}^{-1} \mathbf{G}_k)^{-1} \mathbf{G}_k^T \mathbf{R}_{\mathbf{X}}^{-1} \quad (1.105)$$

### 1.3.3 Методы глобальной оптимизации (методы минимальной нормы)

К этой группе методов мы относим такие подходы решения обратной задачи, которые основываются на минимизации глобального расхождения между измеренными данными и данными, которые получаются при отображении на сенсоры восстановленных источников. Отметим, что некоторые методы из этой группы также могут быть классифицированы как пространственные фильтры, однако нам для дальнейшего изложения удобнее причислять их к текущей группе.

Начнем рассмотрение с самого простого метода этого семейства — восстановления источников методом псевдообратной матрицы.

#### Умножение на псевдообратную матрицу

Начнем с упрощенной порождающей модели данных, предположив, что шума нет:

$$\mathbf{x}(t) = \mathbf{G}\mathbf{s}(t) \quad (1.106)$$

В случае, если количество активных источников меньше количества сенсоров, а положения (а значит и топографии) этих источников известны, система

уравнений выше является переопределенной и может быть легко решена в смысле минимизации среднеквадратичной ошибки. Решение в этом случае широко известно и может быть выписано в явном виде через псевдообратную матрицу, которая выписывается в явном виде как  $\mathbf{G}^\dagger = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T$ , поскольку  $\mathbf{G}^T \mathbf{G}$  обратима:

$$\mathbf{s} = \mathbf{G}^\dagger \mathbf{x}$$

На практике, однако, в задаче картирования мы имеем дело с противоположной ситуацией, когда количество и положение реальных источников заранее неизвестно, и мы покрываем кору равномерной сеткой точек, в каждой из которых мы пытаемся восстановить активность, возможно нулевую. В этом случае количество неизвестных компонент вектора  $\mathbf{s}$  в уравнении 1.106 оказывается много больше количества доступных измерений, и система уравнений оказывается недоопределенной, а значит имеет бесконечно много решений.

В этой ситуации представляется разумным выбрать в качестве решения системы 1.106 такое, которое обладает минимальной нормой. На языке оптимизации такую задачу можно сформулировать как

$$\frac{1}{2} \|\mathbf{s}\|^2 \rightarrow \min_{\mathbf{s}}, \text{ s.t. : } \mathbf{G}\mathbf{s} = \mathbf{x} \quad (1.107)$$

Отметим вновь, что в случае обратной задачи МЭЭГ матрица  $\mathbf{G}$  является «широкой», т.е. количество ее столбцов больше, чем количество строк. Примем здесь также, что столбцы матрицы  $\mathbf{G}$  образуют полную систему векторов, что как правило выполняется на практике при условии, что количество сенсоров достаточно мало. В этом случае матрица  $\mathbf{G}\mathbf{G}^T$  является обратимой.

Выпишем функцию Лагранжа для оптимизационной задачи 1.107:

$$f_L(\mathbf{s}, \lambda) = \frac{1}{2} \|\mathbf{s}\|^2 + \lambda (\mathbf{G}\mathbf{s} - \mathbf{x}) \quad (1.108)$$

Выпишем и приравняем частные производные лагранжиана по каждой из переменных:

$$\frac{\partial f_L(\mathbf{s}, \lambda)}{\partial \lambda} = \mathbf{G}\mathbf{s} - \mathbf{x} = 0 \quad (1.109)$$

$$\frac{\partial f_L(\mathbf{s}, \lambda)}{\partial \mathbf{s}} = \mathbf{s} + \mathbf{G}^T \lambda = 0 \quad (1.110)$$

Далее выразим  $\mathbf{s}$  из уравнения 1.110 и подставим полученное выражение в 1.109:

$$\mathbf{s} = -\mathbf{G}^T \lambda \quad (1.111)$$

$$\mathbf{G}\mathbf{s} = -\mathbf{G}\mathbf{G}^T \lambda = \mathbf{x} \quad (1.112)$$

Из 1.112 получим выражение для  $\lambda$ :

$$\lambda = -(\mathbf{G}\mathbf{G}^T)^{-1} \mathbf{x} \quad (1.113)$$

Наконец, подставив 1.113 в 1.111, получим окончательное выражение для  $\mathbf{s}$ :

$$\mathbf{s} = \mathbf{G}^T (\mathbf{G}\mathbf{G}^T)^{-1} \mathbf{x} \quad (1.114)$$

Как видим, полученное решение имеет структуру псевдообратной матрицы, полученной в явном виде для матрицы, столбцы которой образуют полную систему векторов.

Мы показали таким образом, что псевдообратная матрица в обратной задаче МЭЭГ производит решение, в точности удовлетворяющее исходной линейной системе, и при этом обладающее минимальной нормой среди бесконечного множества таких решений.

Вместе с тем, несмотря на казалось бы очевидные преимущества такого способа оценки источников в обратной задаче МЭЭГ, на практике описанный выше метод никогда не используется. Это вызвано сочетанием двух обстоятельств.

Во-первых, реальные данные, как уже неоднократно упоминалось ранее, неизбежно содержат шумовую компоненту, которую можно моделировать как некоторую добавку  $\omega$  к исходной линейной системе, с помощью которой мы будем оценивать распределение источников  $\hat{\mathbf{s}}$ :

$$\mathbf{G}\hat{\mathbf{s}} = \mathbf{x}_{noisy} \quad (1.115)$$

$$\mathbf{x}_{noisy} = \mathbf{x} + \omega \quad (1.116)$$

Отметим, что на этот раз незашумленные измерения  $\mathbf{x}$  представляют собой некую математическую абстракцию — прямого доступа к этой величине мы не имеем.

Во-вторых, хоть матрица  $\mathbf{G}\mathbf{G}^T$  и является обратимой, она как правило плохо обусловлена в силу того, что топографии (столбцы матрицы  $\mathbf{G}$ ) сильно коррелируют друг с другом. Эти два фактора приводят к тому, что вклад шумовой компоненты  $\omega$  в оценку  $\mathbf{s}$  значительно усиливается ввиду наличия больших собственных значений в матрице  $(\mathbf{G}\mathbf{G}^T)^{-1}$ :

$$\begin{aligned}\hat{\mathbf{s}} &\stackrel{def}{=} \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{x}_{noisy} = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{x} + \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\omega \\ \mathbf{s} &= \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{x} \\ \hat{\mathbf{s}} &= \mathbf{s} + \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\omega\end{aligned}$$

Иными словами, оценка  $\hat{\mathbf{s}}$  перестает адекватно отражать реальное распределение источников  $\mathbf{s}$  из-за шумовой добавки  $\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\omega$ , усиленной неустойчивым обращением матрицы.

Очевидная модификация метода, направленного на устранение этого недостатка посредством стабилизации обращения матрицы  $\mathbf{G}\mathbf{G}^T$  получила название оценки методом минимальной нормы (Minimum Norm Estimate или сокращенно MNE) [40].

### Оценка методом минимальной нормы (MNE).

Оценка методом минимальной нормы является на сегодняшний день одним из наиболее популярных алгоритмов анализа распределения источников в электрофизиологии, а также представляет собой отправную точку для множества других методов, тем или иным способом улучшающих его работу. К таким методам-модификациям относятся прежде всего алгоритмы sLORETA и dSPM, [43; 53]. Использование самого метода в приложении к электрофизиологии впервые было предложено в статье [40].

Суть алгоритма MNE, как уже было сказано, заключается в стабилизации обращения матрицы  $\mathbf{G}\mathbf{G}^T$  посредством поднятия ее диагонали. Аналогичная идея использовалась в методе DICS, который описан в разделе 1.4.1. Оценка источников методом минимальной нормы задается соотношением

$$\hat{\mathbf{s}} = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T + \alpha\mathbf{I}_m)^{-1}\mathbf{x}_{noisy}, \quad (1.117)$$

где  $\alpha$  — параметр регуляризации, а  $\mathbf{I}_m$  — единичная матрица размера  $m \times m$ , где  $m$  — количество сенсоров. При должном выборе параметра регуляризации  $\alpha$  обращение матрицы  $\mathbf{G}\mathbf{G}^T + \alpha\mathbf{I}_m$  становится теперь численно устойчивым, что ограничивает неконтролируемое усиление шумовой компоненты, наблюдаемое для оценок, полученных умножением на псевдообратную матрицу.

Отметим здесь, что формула 1.117 может быть переписана в эквивалентном виде:

$$\hat{\mathbf{s}} = (\mathbf{G}^T \mathbf{G} + \alpha \mathbf{I}_n)^{-1} \mathbf{G}^T \mathbf{x}_{noisy}, \quad (1.118)$$

что легко проверить, подставив сингулярное разложением матрицы  $\mathbf{G}$  в обе формулы 1.117, 1.118, упростив полученные выражения и сравнив результаты.

При этом исходная зашумленная система теперь решается не в точности, а приближенно; параметр  $\alpha$  же регулирует баланс между тем, насколько мала норма восстановленного решения и тем, насколько хорошо это решение объясняет данные. Количественной оценкой последнего является норма разности между решением, умноженным на прямую модель, и вектором измерений.

Более формально, оценка 1.118 может быть найдена как решение следующей задачи оптимизации:

$$\frac{1}{2} \|\mathbf{G}\hat{\mathbf{s}} - \mathbf{x}_{noisy}\|^2 + \frac{1}{2}\alpha \|\hat{\mathbf{s}}\|^2 \rightarrow \min_{\hat{\mathbf{s}}} \quad (1.119)$$

Действительно, дифференцируя функционал 1.119 по  $\mathbf{s}$ , получим:

$$\begin{aligned} \mathbf{G}^T(\mathbf{G}\hat{\mathbf{s}} - \mathbf{x}_{noisy}) + \alpha\hat{\mathbf{s}} &= 0 \\ \mathbf{G}^T\mathbf{G}\hat{\mathbf{s}} + \alpha\hat{\mathbf{s}} &= \mathbf{G}^T\mathbf{x}_{noisy} \\ \hat{\mathbf{s}} &= (\mathbf{G}^T \mathbf{G} + \alpha \mathbf{I}_n)^{-1} \mathbf{G}^T \mathbf{x}_{noisy} \end{aligned}$$

Таким образом становится понятен смысл метода минимальной нормы: с одной стороны, мы пытаемся как можно точнее решить исходную систему уравнений, в том числе и объясняя и шум в данных, а с другой мы стремимся ограничить норму решения, не давая ему слишком сильно расти в попытках угнаться за объяснением шумовой компоненты, которая требует значительного роста нормы решения в силу плохой обусловленности матрицы  $\mathbf{G}\mathbf{G}^T$ . Параметр

регуляризации  $\alpha$  при этом отвечает за баланс между этими двумя конкурирующими процессами.

Существует однако еще один способ взглянуть на метод минимальной нормы, позволяющий как по-новому переосмыслить сам метод, так и существенно обобщить его, сформировав определенный математический каркас, на основе которого может быть сформулировано множество других методов поиска источников в обратной задаче МЭЭГ. Этот способ заключается в описании обратной задачи на языке байесовской вероятности.

Начнем с рассмотрения исходной системы уравнений 1.115:

$$\mathbf{G}\hat{\mathbf{s}} = \mathbf{x}_{noisy} = \mathbf{x} + \boldsymbol{\omega} \quad (1.120)$$

При этом мы хотим найти оценку  $\hat{\mathbf{s}}$  для величины  $\mathbf{s}$ , для которой выполнено

$$\mathbf{Gs} = \mathbf{x}. \quad (1.121)$$

Предположим, что шум  $\boldsymbol{\omega}$  распределен нормально с единичной матрицей ковариации и нулевым средним:

$$\mathbf{Gs} - \mathbf{x}_{noisy} = -\boldsymbol{\omega} \sim \mathcal{N}(0, \mathbf{I}_m) \quad (1.122)$$

Предположим также, что источники  $\mathbf{s}$  некоррелированы и задаются многомерным нормальным распределением с нулевым средним и дисперсией  $\alpha$ :

$$\mathbf{s} \sim \mathcal{N}(0, \alpha \mathbf{I}_n) \quad (1.123)$$

В этом случае критерий максимума апостериорной вероятности говорит нам, что наиболее вероятным распределением (здесь пространственным, а не вероятностным) источников  $\hat{\mathbf{s}}$  с учетом наблюдаемых измерений  $\mathbf{x}_{noisy}$  будет такое, которое максимизирует апостериорную плотность вероятности для величины  $\mathbf{s}$ :

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmax}} \left( e^{-\frac{1}{2}\|\mathbf{Gs}-\mathbf{x}_{noisy}\|^2} e^{-\frac{1}{2}\alpha\|\mathbf{s}\|^2} \right) \quad (1.124)$$

Логарифмируя соотношение 1.124, получим следующую формулу для оценки  $\hat{\mathbf{s}}$ :

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmin}} \left( \frac{1}{2} \|\mathbf{Gs} - \mathbf{x}_{noisy}\|^2 + \frac{1}{2} \alpha \|\mathbf{s}\|^2 \right) \quad (1.125)$$

Как видим, критерий апостериорного максимума в точности привел нас к задаче оптимизации 1.119, сформулированной для метода MNE. На этот раз, однако, статистическая формулировка позволяет нам увидеть метод MNE в перспективе.

Во-первых, становится ясна роль параметра регуляризации  $\alpha$ : в вероятностной формулировке этот гиперпараметр вводится как дисперсия источников при условии, что дисперсия шума на сенсорах равна единице, то есть представляет собой квадрат отношения сигнал-шум.

Во-вторых, становится понятно, что выбор функционала качества для метода MNE обусловлен предположениями о распределениях вероятностей шумовой компоненты и источников. И если предположение о нормальном некоррелированном шуме можно назвать естественным, то предположение о нормальности и некоррелированности источников не кажется столь очевидным. Немедленно возникает вопрос, как изменятся свойства метода и можно ли его улучшить, если выбирать функцию распределения априорной вероятности по-другому. В этом и состоит обобщающая сила вероятностного взгляда на обратную задачу.

В действительности, аргументированный выбор функции распределения априорной вероятности предполагает наличие определенных знаний о характере распределения источников в конкретной задаче поиска активации по МЭЭГ измерениям. Источником таких знаний может служить прежде всего информация об анатомическом строении мозга испытуемого, а также знания о физиологических процессах, характерных для рассматриваемой экспериментальной парадигмы.

Вместе с тем, функция распределения априорной вероятности может быть выбрана и из общих соображений исходя из желаемых свойств решения. Дело в том, что одним из широко известных свойств решений, полученных при помощи метода MNE, является сглаживание источников по пространству. Речь идет о том, что в случае, если истинная активация распределена в пространстве “фокально”, то есть по относительно небольшому участку коры, оценка этой активации методом MNE как правило распределена по области, своим размером сильно превышающей истинную, затрудняя таким образом оценку мощности источника сигнала в случаях, когда несколько источников находятся близко

друг к другу. Кроме того сильным пространственным сглаживанием осложнена оценка фазовой связности источников, восстановленных методом MNE, в силу высокой вероятности взаимной протечки сигнала.

С другой стороны здравый смысл подсказывает нам, что в случае поиска источника, сосредоточенного в одной точке, наиболее простым решением, объясняющим наблюдения, должен быть единственный токовый диполь, расположенный в этой точке. Однако метод MNE в силу пространственного сглаживания будет всегда давать оценку, рассредоточенную пятном вблизи точки, которая действительно была активна, т.е. оценка будет содержать целую группу активных диполей. Такое положение веще наводит на мысль, что руководствуясь принципом максимальной простоты решения, мы должны пересмотреть наш критерий простоты.

При формулировке метода псевдообратной матрицы, а также метода MNE мы отталкивались от минимизации  $L_2$ -нормы решения. Очевидным решением описанной выше проблемы является модификация критерия: мы должны минимизировать количество ненулевых компонент в векторе источников  $\mathbf{s}$ . Такой критерий формализуется как минимизация  $L_0$ -“нормы” решения (в действительности  $L_0$  не является нормой, т.к. для нее нарушается свойство однородности). Так как решение в то же время должно минимизировать норму невязки  $\|\mathbf{G}\hat{\mathbf{s}} - \mathbf{x}\|_{L_2}$ , функционал качества решения по аналогии с 1.119 запишется в виде

$$\frac{1}{2} \|\mathbf{G}\hat{\mathbf{s}} - \mathbf{x}_{noisy}\|_{L_2}^2 + \alpha \|\hat{\mathbf{s}}\|_{L_0} \rightarrow \min_{\hat{\mathbf{s}}} \quad (1.126)$$

В вероятностной постановке такой оптимизационный критерий соответствует выбору в качестве априорной функции распределения многомерного распределения Бернулли в предположении, что все источники независимы. Иными словами, для каждого источника предполагается, что он активен с вероятностью  $p$  и не активен с вероятностью  $1 - p$ , а вероятность совместной активации любого набора диполей-источников равна произведению вероятностей активации каждого из них в отдельности. Для отдельно взятого источника с индексом  $i$  такая функция распределения будет выглядеть как

$$P_{s_i} = p^{(1-\delta(s_i))} (1-p)^{\delta(s_i)} \quad (1.127)$$

где  $\delta(s_i)$  представляет собой дельта-функцию. Эта формула представляет собой обобщение функции плотности вероятности для одного испытания Бернулли на случай непрерывной переменной.

Для совокупной активации с учетом независимости активаций отдельных диполей (предполагается, что активации всех диполей равновероятны с вероятностью  $p$ ) функция распределения будет выглядеть как произведение функций распределения для компонент вектора  $\mathbf{s}$ . Выпишем ее и приведем к форме функции распределения из экспоненциального семейства, чтобы ее было удобнее использовать в критерии максимума апостериори:

$$\begin{aligned}
 P_{\mathbf{s}} &= \prod_{i=1}^N p^{(1-\delta(s_i))} (1-p)^{\delta(s_i)} = \exp \left( \sum_{i=1}^N \left( \log(p^{1-\delta(s_i)}) + \log((1-p)^{\delta(s_i)}) \right) \right) = \\
 &= \exp \left( \sum_{i=1}^N (\log(p) - \delta(s_i) \log(p) + \delta(s_i) \log(1-p)) \right) = \\
 &= \exp \left( - \sum_{i=1}^N \left( \delta(s_i) \log \left( \frac{p}{1-p} \right) - \log(p) \right) \right) = \\
 &= \exp \left( - \log \left( \frac{p}{1-p} \right) \|\mathbf{s}\|_{L_0} + N \log(p) \right) \quad (1.128)
 \end{aligned}$$

Записывая для априорного распределения 1.128 критерий максимума апостериори, логарифмируя и учитывая, что  $N \log(p)$  не зависит от  $\mathbf{s}$ , получим оптимизационный критерий 1.126. Отметим, что вероятностная формулировка позволяет в явном виде получить выражение для параметра  $\alpha$  через вероятность активации одного источника:

$$\alpha = \log \left( \frac{p}{1-p} \right) \quad (1.129)$$

Проблема с задачей оптимизации 1.126 заключается в том, что она является НР-сложной, а значит на практике метод, основанный на оптимизации функционала 1.126 использовать не получится.

Тем не менее, существуют подходы к решению обратной задачи, основанные на аппроксимации  $L_0$ -нормы таким образом, чтобы приближенная задача оптимизации стала алгоритмически разрешимой.

Хронологически первым методом решения обратной задачи для МЭЭГ, основанным на этой идее стал метод минимальных токов (Minimum Current Estimate, MCE).

## Метод минимальных токов (MCE)

В своей основе суть метода заключается в аппроксимации дельта-функции, используемой в “норме”  $L_0$  из функционала 1.126 модулем, соответствующим норме  $L_1$  и решении возникающей задачи оптимизации. Выбор нормы  $L_1$  при формулировке метода через критерий максимума апостериори соответствует выбору в качестве априорной функции распределения источников распределения Лапласа.

Метод MCE в силу специфического выбора функции распределения источников позволяет получать спарсные решения, т.е. решения, сильно разреженные по пространству, демонстрируя свойства аналогичного ему метода LASSO [54], используемого в теории оценивания и машинном обучении. Как и метод LASSO, метод MCE стремится обнулить максимальное количество компонент вектора  $\mathbf{s}$ , чего и следует ожидать от метода, основанного на аппроксимации нормы  $L_0$ .

Оценка источников методом минимальных токов была предложена в статье [41], и методологически является важным алгоритмом давшим развитие ряду более совершенных методов. Тем не менее, на практике этот метод сейчас практически не используется несмотря на выгодное для определенных приложений свойство спарсности решений, что обусловлено рядом присущих ему недостатков.

Во-первых, метод является более требовательным к вычислениям по сравнению с MNE, так как формула оценки методом MCE не может быть выписана в явном виде и решение необходимо получать численно. Кроме того, оптимизируемый функционал качества больше не является гладкой функцией, что затрудняет построение эффективной численной схемы. Иными словами, расплатой за свойство пространственной разреженности решения являются сложности с поиском экстремума функционала качества.

Во-вторых, метод никак не учитывает временную динамику сигнала, так как поиск решения происходит для каждого временного среза отдельно. На практике это приводит к тому, что временные ряды, восстановленные для источников методом MCE, не являются гладкими: для соседних временных срезов могут быть обнулены абсолютно разные компоненты вектора  $\mathbf{s}$ . Здесь стоит отметить, что хотя метод MNE также не учитывает временную структуру сигнала

явно, он, тем не менее, сохраняет гладкость сигнала по времени, так как представляет собой лишь умножение сигнала на невырожденную матрицу.

Чтобы убедиться в том, что метод МСЕ нарушает временную гладкость сигнала, рассмотрим, как будет выглядеть его функционал качества для каждого среза по времени  $t$ :

$$\frac{1}{2} \|\mathbf{G}\hat{\mathbf{s}}(t) - \mathbf{x}_{noisy}(t)\|_{L_2}^2 + \alpha \sum_i |\hat{s}_i(t)| \rightarrow \min_{\hat{\mathbf{s}}} \quad (1.130)$$

Так как минимум для каждого среза ищется независимо от других, а квадрат нормы  $L_2$  и норма  $L_1$  являются аддитивными функциями временных срезов, для совокупности всех временных срезов, для которых решается обратная задача (их количество обозначим  $T$ ), эквивалентный функционал качества можно записать как

$$\frac{1}{2} \left\| \mathbf{G}\hat{\mathbf{S}} - \mathbf{X}_{noisy} \right\|_{fro}^2 + \alpha \sum_{i,j} |\hat{S}_{i,j}| \rightarrow \min_{\hat{\mathbf{S}}} \quad (1.131)$$

где  $\hat{\mathbf{S}}$  — матрица размером  $N \times T$ , содержащая по строкам временные ряды, оцененные для каждого источника, а  $\mathbf{X}_{noisy}$  — матрица размером  $M \times T$  временных рядов, записанных сенсорами.

Такая запись позволяет увидеть, что при оптимизации функционала 1.131 компоненты вектора  $\mathbf{s}$  штрафуются от временного среза к временному срезу при помощи нормы  $L_1$  точно также, как и по пространству, то есть свойство разреженности решения, сопутствующее этой норме, применяется не только к пространственной структуре восстановленного сигнала, но и к его временной динамике. В итоге решения, полученные методом МСЕ, оказываются нестабильными. Аналогично для метода МНЕ решение штрафуется и по пространству и по времени нормой  $L_2$ , соответствующей гладким решениям.

Вместе с тем, временная спарсность противоречит нашей интуиции относительно природы сигнала, так как допускает, что соседние временные срезы могут соответствовать абсолютно разному пространственному распределению источников тогда как мы ожидаем определенной гладкости во временной динамике источников, поскольку характерные времена процессов, порождающих на коре измеримый сигнал (десятки миллисекунд), сравнительно велики по отношению к типичным временам сэмплирования (десятие доли миллисекунды).

Отсюда возникает желание совместить пространственную спарсность метода МСЕ со временной гладкостью метода МНЕ. Эта идея приводит к группе

методов, основанных на комбинации норм, штрафующих матрицу решения  $\mathbf{S}$  неодинаково по времени и пространству. Методы, основанные на этой идее получили название Mixed Norm Estimate (MxNE), или оценка методом смешанной нормы. Они легли в основу нашего метода поиска сетей GO-PSICOS. Метод MxNE и его модификацию для поиска сетей мы рассмотрели в разделе [2.3](#).

#### **1.4 Использование методов решения обратной задачи для оценки фазовой синхронности.**

Как мы уже отмечали во введении, оценка коннективности в пространстве сенсоров имеет два недостатка:

- перемешивание сигнала на сенсорах искусственно увеличивает коннективность а также ведёт к неправильной оценке мощностей.
- анализ сенсоров даёт менее пространственно специфичную картину активности.

Для оценки коннективности на уровне сенсоров методы, рассмотренные в предыдущем разделе, могут быть использованы напрямую, как часть двухэтапной процедуры. Сначала при помощи одного из методов решения обратной задачи оцениваются временные ряды источников. Затем для источников применяется одна из мер коннективности.

Такая двухшаговая процедура является распространенным способом оценки коннективности на источниках. Тем не менее, она имеет один важный недостаток: так как методы решения ОЗ неизбежно ведут к протечке сигнала между источниками. И хотя часть этих методов стремится подавить протечку сигнала в каждую отдельную точку коры, с точки зрения последующей оценки коннективности такой подход является субоптимальным, т.к. в функционал качества не закладывается минимизация общей протечки сигнала для двух точек, между которыми будет померяна коннективность.

### 1.4.1 DICS

Рассмотрим теперь одну важную модификацию метода векторных бимформеров, разработанную для анализа коннективностей [55].

Суть метода состоит в построении адаптивных фильтров аналогично тому, как это делалось в разделе 1.3.2, но на этот раз в частотной области.

Рассмотрим подробно этапы построения DICS-фильтров. Начнем вновь с рассмотрения порождающей модели сигнала

$$\mathbf{X} = \mathbf{GS} + \boldsymbol{\Omega}, \quad (1.132)$$

Применим к данным частотно-временное преобразование и посчитаем кросс-спектр по аналогии с 2.2 и 2.6:

$$\mathbf{C}^{\mathbf{XX}} = \mathbf{GC}^{\mathbf{ss}}\mathbf{G}^T + \mathbf{C}^{\omega\omega} \quad (1.133)$$

Оптимизационная задача метода DICS формализуется следующим образом:

$$\text{tr} \left\{ \mathbf{A}_k \mathbf{C}^{\mathbf{XX}} \mathbf{A}_k^T \right\} + \alpha \text{tr} \left\{ \mathbf{A}_k \mathbf{A}_k^T \right\} \rightarrow \min_{\mathbf{A}_k}, s.t. : \mathbf{A}_k \mathbf{G}_k = \mathbf{I} \quad (1.134)$$

Оптимизационный критерий аналогичен критерию из раздела 1.3.2 для векторной постановки задачи адаптивной фильтрации, минимизирующей вклад от третих источников, с добавлением регуляризации матрицы кросс-спектра.

Проводя покомпонентный вывод методом множителей Лагранжа по аналогии с выводом формул 1.104, 1.105, получим выражения для фильтров в следующем виде:

$$\mathbf{A}_k = \left( \mathbf{G}_k^T (\mathbf{C}^{\mathbf{XX}} + \alpha \mathbf{I})^{-1} \mathbf{G}_k \right)^{-1} \mathbf{G}_k^T (\mathbf{C}^{\mathbf{XX}} + \alpha \mathbf{I})^{-1} \quad (1.135)$$

Для получения оценок на элементы матрицы кросс-спектральной плотности для источников умножим кросс-спектр измерений на полученные матрицы фильтров:

$$\mathbf{C}_{kl}^{loose} = \mathbf{A}_k \mathbf{C}^{\mathbf{XX}} \mathbf{A}_l^H \quad (1.136)$$

Полученная матрица  $3 \times 3$ ,  $\mathbf{C}_{kl}^{loose}$  содержит оценки элементов кросс-спектра источников по трем ортогональным направлениям. Оценочное значение элемента кросс-спектра для истинной ориентации диполя выбирается как след матрицы  $\mathbf{C}_{kl}^{loose}$ :

$$\mathbf{C}_{kl}^{\text{SS}} = \text{tr} \{ \mathbf{A}_k \mathbf{C}^{\text{XX}} \mathbf{A}_l^H \} \quad (1.137)$$

В случае, если максимальное собственное значение матрицы  $\mathbf{C}_{kl}^{loose}$  сильно превосходит два других, в качестве оценки на  $\mathbf{C}_{kl}^{\text{SS}}$  можно брать максимальное собственное число матрицы  $\mathbf{C}_{kl}^{loose}$ , что соответствует случаю, когда направления диполей в точках  $k$  и  $l$  приближенно оставались неизменными за время оценки матрицы  $\mathbf{C}^{\text{XX}}$ .

Окончательно, для двух источников  $k, l$  будем иметь оценку квадрата когерентности вида

$$\text{coh}(k, l)^2 = \frac{\mathbf{C}_{kl}^{\text{SS}}}{\sqrt{\mathbf{C}_{kk}^{\text{SS}} \mathbf{C}_{ll}^{\text{SS}}}} \quad (1.138)$$

Отметим, что такой метод оценки функциональной коннективности, хотя и является на сегодняшний день одним из самых популярных в нейронаучном сообществе, все-таки оказывается подвержен проблеме протечки сигнала (см. результаты симуляций в главе 3), а следовательно склонен порождать ложноположительные срабатывания, что приводит к сложностям в аккуратной интерпретации полученного распределения коннективностей.

Кроме того, этот метод страдает от общей для своего класса алгоритмов проблемы — качество его работы заметно снижается при наличии коррелированных источников, [56], что представляется ключевым недостатком для метода поиска когерентных сигналов, в том числе с малыми фазовыми задержками.

#### 1.4.2 iDICS

Чтобы сделать метод DICS устойчивым к протечке сигнала, мы можем воспользоваться идеей метода мнимой когерентности 2.1.1. Для этого при оценке внедиагональных элементов матрицы кросс-спектра в пространстве источников

будем умножать DICS фильтры не на весь кросс-спектр на сенсорах, а только на его мнимую часть.

Такая модификация метода DICS предоставляет сравнительно простой способ оценки мнимой части матрицы кросс-спектральной плотности в пространстве источников. Использование мнимой части, как и для метода *imcoh*, позволяет устраниить протечки сигнала при оценке коннективности.

Формально метод iDICS получается модификацией формулы 1.136: вместо полного кросс-спектра на сенсорах необходимо взять его мнимую часть:

$$\mathbf{C}_{kl}^{loose} = \mathbf{A}_k \Im(\mathbf{C}^{\mathbf{XX}}) \mathbf{A}_l^H \quad (1.139)$$

Чтобы фильтры были действительнозначными, при их построении можно пользоваться только действительной частью кросс-спектра на источниках:

$$\mathbf{A}_k = \left( \mathbf{G}_k^T (\Re(\mathbf{C}^{\mathbf{XX}}) + \alpha \mathbf{I})^{-1} \mathbf{G}_k \right)^{-1} \mathbf{G}_k^T (\Re(\mathbf{C}^{\mathbf{XX}}) + \alpha \mathbf{I})^{-1} \quad (1.140)$$

Вычисление когерентности на источниках требует нормализации оцененного внедиагонального элемента кросс-спектра на мощности взаимодействующих источников. Использование мнимой части кросс-спектра не позволяет оценить мощности источников, поэтому для оценки мощностей необходимо либо пользоваться действительной частью кросс-спектра с теми же действительнозначными фильтрами.

### 1.4.3 GCS

Оценка методом мнимой когерентности, а также методы, основанные на той же идее, имеют один существенный недостаток: они не позволяет детектировать сети с близкой к нулю фазовой задержкой. Это наблюдение заставляет искать иные способы борьбы с методом протечки сигнала.

Одним из таких методов является метод геометрической коррекции (GCS), [38]. Идея метода основана на том наблюдении, что протечка сигнала из точки Б коры в точку А пропорциональна скалярному произведению топографии точки Б на фильтр, восстанавливающий сигнал в точке А. Метод геометрической коррекции предлагает для оценки когерентности между точками А и Б удалять из сигнала источник с топографией как у точки Б.

Оценка методом GCS включает четыре шага:

1. спроектировать вектор сигнала на сенсорах ортогонально топографии точки Б.
2. на основе спроектированных данных одним из методов решения обратных задач оценить активность в точке А.
3. использовать исходный, неспроектированный вектор сигнала на сенсорах для оценки активности в точке Б.
4. посчитать какую-либо меру коннективности между активностями оцененными в точках А и Б (coh, plv и т.д.)

Фактически метод геометрической коррекции предлагает использовать вместо идеи imcoh пространственную фильтрацию, которая устраняет протечку сигнала между точками, в которых мы хотим померить коннективность.

К достоинствам этого метода относится прежде всего сохранение в сигнале информации о коннективности с нулевой фазовой задержкой. Кроме того, метод является сравнительно простым в вычислительном плане.

К недостаткам метода относится в первую очередь тот факт, что метод геометрической коррекции при проекции игнорирует протечку сигнала от всех остальных источников, кроме пары, для которой считается коннективность. Это приводит к тому, что протечка сигнала от трех источников в точки А и Б все равно вызывает ложную коннективность. Иными словами, метод GCS полностью избавляется от ложной коннективности, вызванной протечкой сигнала только в том случае, если активно всего два источника, а шум мозга отсутствует.

## Глава 2. Метод оценки синхронизации с произвольными фазовыми задержками.

### 2.1 Протечка сигнала и ее удаление ортогональным проецированием.

Рассмотрим типичную постановку эксперимента в ЭЭГ/МЭГ парадигме. Пусть, было записано  $K$  эпох гомогенной электрофизической активности продолжительностью  $\Delta t$  с помощью энцефалографа (магнитного или электрического), имеющего  $m$  сенсоров. Далее, допустим, что имеется информация об анатомии мозга испытуемого (этого достаточно использовать стандартную анатомическую модель мозга, но для приложений предпочтительнее иметь индивидуальный МРТ снимок). Последнее допущение позволяет нам рассмотреть пространство источников — сетку, состоящую из  $n$  точек на коре и аппроксимирующую поверхность мозга испытуемого (2.1). Решая так называемую прямую задачу [2], мы можем осуществить отображение  $n$  точек пространства источников в  $m$ -мерное пространство сенсоров с оператором отображения  $\mathbf{G}$ , имея в виду, что обычно  $n \gg m$ , так как для качественной аппроксимации нам нужно намного больше точек на поверхности мозга, чем сенсоров, регистрирующих мозговую активность.

Оператор  $\mathbf{G}$  — суть, решение уравнений Максвелла для точечных источников электромагнитной активности, расположенных в каждом узле сетки, определяющей конфигурацию источников, и так как уравнения Максвелла линейны, оператор  $\mathbf{G}$  также является линейным [39]. Мы также должны учесть влияние шума источники которого располагаются во внешней среде и который в первом приближении предполагается белым.

В этих предположениях можем записать порождающую модель данных на сенсорах:

$$\mathbf{x}_k(t) = \mathbf{G} \cdot \mathbf{s}_k(t) + \boldsymbol{\omega}_k(t), \quad (2.1)$$

где  $\mathbf{s}_k(t)$  —  $n$ -мерный вектор-столбец активаций источников,  $\mathbf{x}_k(t)$  —  $m$ -мерный вектор-столбец сигналов на сенсорах,  $t$  — время, а  $\mathbf{G}$  —  $m \times n$  матрица линейного

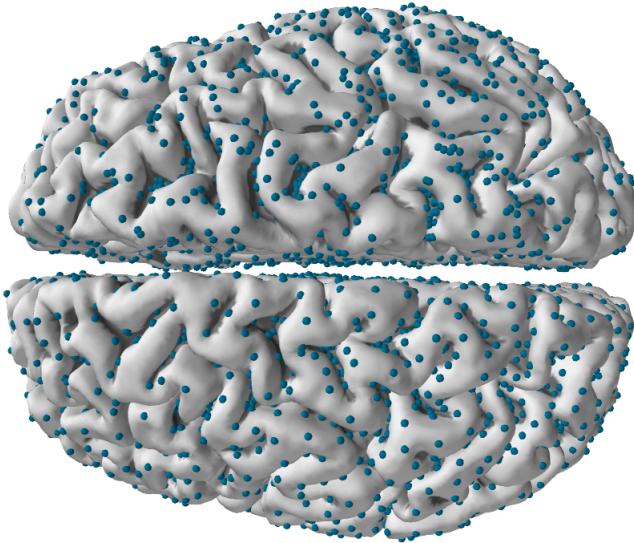


Рисунок 2.1: Пример сетки в пространстве источников, построенной на основе трехмерной анатомической модели

отображения пространства источников в пространство сенсоров. Индекс  $k$  обозначает номер эпохи. Применим к (2.1) частотное преобразование. Существует много способов произвести эту операцию, например, — вейвлет-преобразование или узкополосная фильтрация с последующим аналитическим представлением сигнала. Мы не будем углубляться в детали теории частотно-временных преобразований; для подробного изложения предмета см. [57], о применении частотно-временных преобразований к обработке электрофизиологических данных см. [58]. Мы только упомянем, что вне зависимости от особенностей выбранного метода образом частотно-временного преобразования матрицы  $\mathbf{X}$  размера  $m \times T$  будет комплексный тензор, каждый временной срез которого содержит в себе информацию о фазовом и амплитудном спектре сигнала в каждый момент времени. Чтобы упростить вывод, мы зафиксируем одну частоту имея в виду, что дальнейшие выкладки справедливы и для остальных частот преобразования. В итоге мы можем записать:

$$\hat{\mathbf{x}}_{k,f_i}(t) = \mathbf{G} \cdot \hat{\mathbf{s}}_{k,f_i}(t) + \hat{\boldsymbol{\omega}}_{k,f_i}(t), \quad (2.2)$$

где  $\hat{\mathbf{x}}, \hat{\mathbf{s}}, \hat{\boldsymbol{\omega}}$  — комплекснозначные образы  $\mathbf{x}, \mathbf{s}, \boldsymbol{\omega}$  соответственно. Для простоты далее будем опускать нижний индекс  $f_i$ :

$$\hat{\mathbf{x}}_k(t) = \mathbf{G} \cdot \hat{\mathbf{s}}_k(t) + \hat{\boldsymbol{\omega}}_k(t) \quad (2.3)$$

### 2.1.1 Порождающая модель матрицы кросс-спектральной плотности и метод ImCoh

Теперь, если мы для каждой эпохи умножим  $\hat{\mathbf{x}}_k(t)$  на его эрмитово сопряжение и усредним результат, мы получим порождающую модель кросс-спектра на сенсорах:

$$\begin{aligned} \langle \hat{\mathbf{x}}_k(t) \hat{\mathbf{x}}_k(t)^\dagger \rangle_{k=1}^K &= \langle (\mathbf{G} \cdot \hat{\mathbf{s}}_k(t) + \hat{\omega}_k(t))(\mathbf{G} \cdot \hat{\mathbf{s}}_k(t) + \hat{\omega}_k(t))^\dagger \rangle_{k=1}^K = \\ &= \mathbf{G} \cdot \langle \hat{\mathbf{s}}_k(t) \hat{\mathbf{s}}_k(t)^\dagger \rangle_{k=1}^K \cdot \mathbf{G}^T + \mathbf{G} \cdot \langle \hat{\mathbf{s}}_k(t) \hat{\omega}_k(t)^\dagger \rangle_{k=1}^K + \\ &\quad + \langle \hat{\omega}_k(t) \hat{\mathbf{s}}_k(t)^\dagger \rangle_{k=1}^K \cdot \mathbf{G}^T + \langle \hat{\omega}_k(t) \hat{\omega}_k(t)^\dagger \rangle_{k=1}^K, \end{aligned} \quad (2.4)$$

где оператор  $\langle \cdot \rangle_{k=1}^K$  означает усреднение по эпохам.

Так как шум предполагается белым, а сигнал на сенсорах  $\hat{\mathbf{s}}_k$  и шум  $\hat{\omega}_k$  взаимно независимыми, можем заметить, что второй и третий члены уравнения (2.4) равны нулю, когда число эпох достаточно велико. Также отметим, что  $\hat{\mathbf{x}}_k(t) \hat{\mathbf{x}}_k(t)^\dagger$ ,  $\hat{\mathbf{s}}_k(t) \hat{\mathbf{s}}_k(t)^\dagger$  и  $\hat{\omega}_k(t) \hat{\omega}_k(t)^\dagger$  представляют собой внешние произведения комплексно-значных векторов на свои комплексные сопряжения, и следовательно являются эрмитовыми матрицами, что означает, что значения, находящиеся на диагоналях этих матриц, принадлежат к области вещественных чисел. Это свойство, очевидно, сохраняется и после усреднения. Более того, раз элементы векторов являются комплексными числами, они могут быть представлены в виде  $\hat{\xi}(t) = r(t) \cdot e^{i\varphi(t)}$ , где  $\varphi(t)$  соответствует мгновенной фазе сигнала, а  $r(t)$  амплитуде. Следовательно,  $\langle \hat{\xi}_p \hat{\xi}_q^* \rangle = \langle r_p r_q e^{i\Delta\varphi} \rangle$  или, если для иллюстрации мы примем, что амплитуды и фазы независимы (справедливость последнего является предметом разногласий для электрофизиологических данных [20], [19]), мы получим, что  $\langle \hat{\xi}_p \hat{\xi}_q^* \rangle = \langle r_p r_q \rangle \langle e^{i\Delta\varphi} \rangle$ .

Из последнего соотношения можно сделать несколько выводов. Во-первых, элементы матрицы кросс-спектра представляют собой степень стабильности разности фаз  $\Delta\varphi$  по эпохам. Так, если разность фаз была достаточно равномерно распределена по всему возможному интервалу принимаемых значений, то среднее  $\langle e^{i\Delta\varphi} \rangle$  будет приблизительно равно нулю. С другой стороны, если разность фаз сохраняется от эпохи к эпохе, результирующий коэффициент будет отличен от нуля, что соответствует случаю установления коннективности

между сигналами. Во-вторых, если разность фаз мала, элементы взаимного спектра могут быть близки к ненулевому вещественному числу.

Введем следующее обозначение для матрицы кросс-спектральной плотности:

$$\mathbf{C}^{\mathbf{vv}} \stackrel{def}{=} \langle \hat{\mathbf{v}}_k(t) \hat{\mathbf{v}}_k(t)^\dagger \rangle_{k=1}^K, \quad (2.5)$$

Используя определение (2.5) и опуская нулевые члены, (2.4) перепишется в виде

$$\mathbf{C}^{\mathbf{xx}}(t) = \mathbf{G} \mathbf{C}^{\mathbf{ss}}(t) \mathbf{G}^T + \mathbf{C}^{\omega\omega}(t) \quad (2.6)$$

Теперь рассмотрим более детально диагональ матрицы  $\mathbf{C}^{\mathbf{ss}}$ . Как уже было сказано, элементы главной диагонали этой матрицы являются вещественными числами и они представляют собой значения мощностей сигналов пространства источников, имеющих частоту  $f_i$ . В структуре порождающей модели отражен тот факт, что после отображения оператором из пространства источников в пространство сенсоров с помощью оператора  $\mathbf{G}$  эти мощностные члены будут смешаны с истинной коннективностью, и так как исходная система уравнений была сильно недоопределенна (условие  $n \gg m$ ), разделить сигналы не представляется возможным. Математически этим и объясняется эффект протечки сигнала.

Чтобы внести дополнительную ясность, представим себе ситуацию, когда синхронизация между источниками полностью отсутствует, но при этом некоторые участки мозга активны. В такой постановке все элементы матрицы  $\mathbf{C}^{\mathbf{ss}}$ , лежащие вне главной диагонали, будут равны нулю, но для  $\mathbf{C}^{\mathbf{xx}}$  это выполниться не будет. Пары сенсоров, расположенные близко к активным участкам мозга, будут иметь большие кросс-спектральные коэффициенты, что приведет к ложному обнаружению коннективности между источниками.

В ранее упомянутом методе ImCoh (в настоящее время, вероятно, наиболее часто используемом для измерения коннективности) предлагается рассматривать только мнимые части уравнения (2.6), что уберегает нас от негативного эффекта протечки сигнала, но при этом мы также теряем действительную часть «хорошего» сигнала, несущую информацию о фазовой синхронизации.

### 2.1.2 Оператор проецирования от протечки сигнала.

Для более полного использования информации о фазовой синхронизации, содержащейся в кросс-спектре мы предлагаем другой подход к устранению эффекта протечки сигнала. Распишем произведение матриц  $\mathbf{G}\mathbf{C}^{\text{ss}}\mathbf{G}^T$  в правой части уравнения (2.6):

$$\mathbf{C}^{\text{xx}}(t) = \sum_{p=1}^n \sum_{q=1}^n \mathbf{g}_p \mathbf{g}_q^T c_{pq}^{\text{ss}}(t) + \mathbf{C}^{\omega\omega}(t), \quad (2.7)$$

где  $\mathbf{g}_p$  — столбец матрицы  $\mathbf{G}$ , который называют *топографией* источника  $p$ , поскольку он определяет, каким образом сигнал, поступающий от источника  $p$ , будет виден на сенсорах. Можно заметить, что кросс-спектр на уровне сенсоров представляет собой взвешенную сумму внешних произведений топографий, взятую с коэффициентами, являющимися элементами кросс-спектра из пространства источников.

### 2.1.3 Произведение Кронекера и векторизация

Векторизуем уравнение (2.7):

$$\text{vec}(\mathbf{C}^{\text{xx}}(t)) = \sum_{p=1}^n \sum_{q=1}^n \text{vec}(\mathbf{g}_p \mathbf{g}_q^T) c_{pq}^{\text{ss}}(t) + \text{vec}(\mathbf{C}^{\omega\omega}(t)) \quad (2.8)$$

Для упрощения записи будем использовать понятие произведения Кронекера. Произведением Кронекера матриц  $A$  и  $B$ , имеющих размеры  $p \times q$  и  $r \times s$  соответственно, называется матрица вида

$$\mathbf{A} \otimes \mathbf{B} \stackrel{\text{def}}{=} \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{bmatrix} \quad (2.9)$$

Произведение Кронекера является билинейной ассоциативной операцией:

$$\mathbf{A} \otimes (\mathbf{B} + \mathbf{C}) = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{C} \quad (2.10)$$

$$(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C} \quad (2.11)$$

$$\alpha(\mathbf{A} \otimes \mathbf{B}) = (\alpha\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (\alpha\mathbf{B}) \quad (2.12)$$

$$\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C} \quad (2.13)$$

Заметим также, что произведение Кронекера не является симметричным:  $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$ . Приведем другие полезные соотношения:

$$(\mathbf{A} \otimes \mathbf{B})^T = \mathbf{A}^T \otimes \mathbf{B}^T \quad (2.14)$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AB}) \otimes (\mathbf{CD}) \quad (2.15)$$

Отметим особенно интересное для нас свойство произведения Кронекера, связывающее его с процедурой векторизации. Для матриц  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  справедливо (доказательство см. в [59]):

$$vec(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) vec(\mathbf{B}) \quad (2.16)$$

В нашем случае приведенное выражение принимает самую простую форму:

$$vec[\mathbf{g}_p \mathbf{g}_q^T] = vec \left[ \begin{pmatrix} g_p^1 \\ \vdots \\ g_p^m \end{pmatrix} \cdot 1 \cdot \begin{pmatrix} g_q^1 & \dots & g_q^m \end{pmatrix} \right] = \begin{pmatrix} g_q^1 \\ \vdots \\ g_q^m \end{pmatrix} \otimes \begin{pmatrix} g_p^1 \\ \vdots \\ g_p^m \end{pmatrix} \cdot vec(1) = \mathbf{g}_q \otimes \mathbf{g}_p, \quad (2.17)$$

где  $\mathbf{g}_q \otimes \mathbf{g}_p$  представляет собой вектор-столбец размера  $m^2 \times 1$ . Введем понятие *коммутационной матрицы* [59]. Коммутационная матрица — это матрица, которая связывает операции транспонирования и векторизации и определяется следующим соотношением:

$$vec(\mathbf{A}) = \mathbf{K}^{(m,n)} vec(\mathbf{A}^T) \quad (2.18)$$

где  $\mathbf{A}$  — матрица размером  $m \times n$ , а матрица  $\mathbf{K}^{(m,n)}$  имеет размеры  $mn \times mn$ . Матрица  $K^{(m,n)}$  является частным случаем матрицы перестановок и может быть получена из единичной перестановкой строк.

Далее нас будет интересовать только случай квадратных матриц  $\mathbf{A}$ . Получим в явном виде выражение для коммутационной матрицы в случае квадратных матриц.

Пусть матрица  $\mathbf{A}$  имеет размеры  $n \times n$ . Обозначим соответствующую ей коммутационную матрицу как  $\mathbf{K}_n$ . Обозначим символом  $\mathbf{e}_i$  вектор-столбец длины  $n$ , у которого на  $i$ -ом месте стоит единица, а на всех остальных — нули. Тогда для векторизации  $\mathbf{A}^T$  будем иметь:

$$\begin{aligned} \text{vec}(\mathbf{A}^T) &= \text{vec}\left(\sum_{i,j} A_{i,j} \mathbf{e}_j \mathbf{e}_i^T\right) = \text{vec}\left(\sum_{i,j} (\mathbf{e}_i^T \mathbf{A} \mathbf{e}_j)(\mathbf{e}_j \mathbf{e}_i^T)\right) = \\ &= \text{vec}\left(\sum_{i,j} \mathbf{e}_j \mathbf{e}_i^T \mathbf{A} \mathbf{e}_j \mathbf{e}_i^T\right) = \sum_{i,j} \text{vec}(\mathbf{e}_j \mathbf{e}_i^T \mathbf{A} \mathbf{e}_j \mathbf{e}_i^T) = \left(\sum_{i,j} (\mathbf{e}_i \mathbf{e}_j^T) \otimes (\mathbf{e}_j \mathbf{e}_i^T)\right) \text{vec}(\mathbf{A}) \end{aligned} \quad (2.19)$$

Таким образом,

$$\mathbf{K}_n = \sum_{i,j} (\mathbf{e}_i \mathbf{e}_j^T) \otimes (\mathbf{e}_j \mathbf{e}_i^T) \quad (2.20)$$

Теперь перепишем выражение (2.8), используя новые обозначения:

$$\text{vec}(\mathbf{C}^{\mathbf{xx}}(t)) = \sum_{p=1}^n \sum_{q=1}^n \mathbf{g}_q \otimes \mathbf{g}_p \cdot c_{pq}^{\text{ss}}(t) + \text{vec}(\mathbf{C}^{\omega\omega}(t)) \quad (2.21)$$

Теперь можно видеть, что векторизованный кросс-спектр на уровне сенсоров представлен линейной комбинацией векторов  $\mathbf{g}_q \otimes \mathbf{g}_p$  в  $m^2$ -мерном векторном пространстве. Назовем эти векторы *2-тотографиями*. Нам уже известно, что эффект протечки сигнала, от которого необходимо избавиться, обусловлен 2-тотографиями особого вида  $\mathbf{g}_p \otimes \mathbf{g}_p$  (всего  $n$  векторов), конкретный вид которых нам задан через оператор  $\mathbf{G}$ . Следовательно, мы можем спроектировать выражение (2.21) ортогонально линейному пространству, натянутому на 2-тотографии источников протечки сигнала.

### 2.1.4 Построение оператора проецирования

Прежде чем мы начнем построение оператора ортогональной проекции от подпространства протечки сигнала, необходимо понять, как соотносится подпространство протечки сигнала с линейными оболочками 2-топографий действительной и мнимой частей порождающей модели кросс-спектра на сенсорах.

Во-первых, выделим в уравнении (2.21) действительную и мнимую части. Заметим, что так как матрица  $\mathbf{C}^{\text{ss}}$  является эрмитовой,  $c_{pq}^{\text{ss}} = \overline{c_{qp}^{\text{ss}}}$  (верхняя черта обозначает комплексное сопряжение):

$$\begin{aligned} \sum_{p=1}^n \sum_{q=1}^n \mathbf{g}_q \otimes \mathbf{g}_p \cdot c_{pq}^{\text{ss}}(t) &= \operatorname{Re} \left( \sum_{p,q=1}^{n,n} \mathbf{g}_q \otimes \mathbf{g}_p \cdot c_{pq}^{\text{ss}}(t) \right) + i \cdot \operatorname{Im} \left( \sum_{p,q=1}^{n,n} \mathbf{g}_q \otimes \mathbf{g}_p \cdot c_{pq}^{\text{ss}}(t) \right) = \\ &= \sum_{p \leq q}^{n,n} (\mathbf{g}_q \otimes \mathbf{g}_p + \mathbf{g}_p \otimes \mathbf{g}_q) \operatorname{Re} (c_{pq}^{\text{ss}}(t)) + i \cdot \sum_{p < q}^{n,n} (\mathbf{g}_q \otimes \mathbf{g}_p - \mathbf{g}_p \otimes \mathbf{g}_q) \operatorname{Im} (c_{pq}^{\text{ss}}(t)) \end{aligned} \quad (2.22)$$

Отметим изменения в индексах суммирования.

Для удобства обозначим линейное пространство, натянутое на 2-топографии, ответственные за протечку сигнала, как  $S_{SL}$ , а линейную оболочку 2-топографий мнимой части —  $S_{Im}$ .

Из уравнения (2.22) видно, что 2-топографии действительной и мнимой частей устроены по-разному, а именно, — векторы  $\mathbf{g}_q \otimes \mathbf{g}_p + \mathbf{g}_p \otimes \mathbf{g}_q$  являются симметричными, тогда как 2-топографии мнимых частей  $\mathbf{g}_q \otimes \mathbf{g}_p - \mathbf{g}_p \otimes \mathbf{g}_q$  антисимметричны по индексам  $p, q$ . Нас интересует, как это структурное отличие проявляется во взаимосвязи подпространств  $S_{Re}$  и  $S_{Im}$  с подпространством протечки сигнала  $S_{VC}$ . Покажем, что 2-топографии мнимой части ортогональны векторам, на которые натянуто подпространство объемной проводимости:

$$\begin{aligned} (\mathbf{g}_s \otimes \mathbf{g}_s)^T (\mathbf{g}_q \otimes \mathbf{g}_p - \mathbf{g}_p \otimes \mathbf{g}_q) &= (\mathbf{g}_s^T \otimes \mathbf{g}_s^T)(\mathbf{g}_q \otimes \mathbf{g}_p) - (\mathbf{g}_s^T \otimes \mathbf{g}_s^T)(\mathbf{g}_p \otimes \mathbf{g}_q) = \\ &= (\mathbf{g}_s^T \mathbf{g}_q \otimes \mathbf{g}_s^T \mathbf{g}_p - \mathbf{g}_s^T \mathbf{g}_p \otimes \mathbf{g}_s^T \mathbf{g}_q) \stackrel{*}{=} (\mathbf{g}_s^T \mathbf{g}_p \mathbf{g}_s^T \mathbf{g}_q - \mathbf{g}_s^T \mathbf{g}_p \mathbf{g}_s^T \mathbf{g}_q) = 0 \end{aligned} \quad (2.23)$$

Равенство  $*$  сохраняется, так как  $\mathbf{g}_s^T \mathbf{g}_q$  и  $\mathbf{g}_s^T \mathbf{g}_p$  являются скалярными величинами, и мы можем опустить операцию “ $\otimes$ ” и поменять местами множители. Из (2.23) можно видеть, что подпространство объемной проводимости

ортогонально мнимой части подпространства. Для действительной части такое соотношение не сохраняется:

$$\begin{aligned} (\mathbf{g}_s \otimes \mathbf{g}_s)^T (\mathbf{g}_q \otimes \mathbf{g}_p + \mathbf{g}_p \otimes \mathbf{g}_q) &= (\mathbf{g}_s^T \otimes \mathbf{g}_s^T)(\mathbf{g}_q \otimes \mathbf{g}_p) + (\mathbf{g}_s^T \otimes \mathbf{g}_s^T)(\mathbf{g}_p \otimes \mathbf{g}_q) = \\ &= (\mathbf{g}_s^T \mathbf{g}_q \otimes \mathbf{g}_s^T \mathbf{g}_p + \mathbf{g}_s^T \mathbf{g}_p \otimes \mathbf{g}_s^T \mathbf{g}_q) = 2\langle \mathbf{g}_s, \mathbf{g}_p \rangle \langle \mathbf{g}_s, \mathbf{g}_q \rangle \end{aligned} \quad (2.24)$$

После проведенных операций легко увидеть, что проекция ортогонально подпространству протечки сигнала оказывает влияние на подпространство действительной части кросс-спектра на источниках. Следовательно, нужно добиться того, чтобы протечка сигнала была удалена из данных насколько это возможно с учетом минимального воздействия на подпространство действительной части кросс-спектра (см.рис. 2.2). Для достижения этой цели необходимо уменьшить размерность подпространства объемной проводимости неким оптимальным способом.

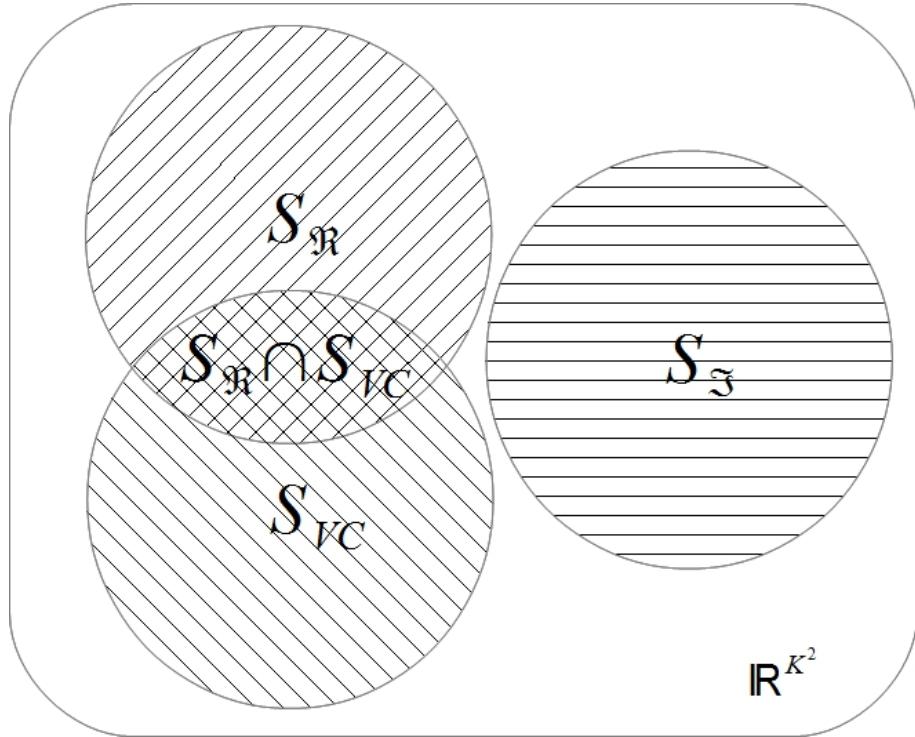


Рисунок 2.2: Взаимосвязь подпространств для кросс-спектра на уровне сенсоров

Подпространство протечки сигнала  $S_{SL}$  и подпространство действительной части кросс-спектра  $S_{\Re}$  имеют непустое пересечение. Кроме того, оба этих подпространства ортогональны подпространству мнимой части кросс-спектра  $S_{\Im}$ . Пересечение  $S_{SL}$  с  $S_{\Re}$  содержит вклад как от протечки сигнала, так и от истинно взаимодействующих источников, расположенных близко друг к другу и характеризующихся малой разностью фаз.

Имея в виду все вышеперечисленное, вернемся к построению проектора. Рассмотрим матрицу, составленную из вектор-столбцов, образующих подпространство протечки сигнала:

$$\mathbf{F} = \begin{bmatrix} & & & \\ | & | & | & | \\ \mathbf{g}_1 \otimes \mathbf{g}_1 & \mathbf{g}_2 \otimes \mathbf{g}_2 & \dots & \mathbf{g}_n \otimes \mathbf{g}_n \\ | & | & & | \end{bmatrix} \quad (2.25)$$

Произведем сингулярное разложение матрицы  $\mathbf{F}$  [60].

$$\mathbf{F} = \mathbf{U} \mathbf{S} \mathbf{V}^T = \begin{bmatrix} & & & | \\ | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_{m^2} \\ | & | & & | \end{bmatrix} \begin{pmatrix} \lambda_1 & 0 & \dots \\ 0 & \lambda_2 & \dots \\ \vdots & & \ddots \end{pmatrix} \begin{bmatrix} - & \mathbf{v}_1 & - \\ - & \mathbf{v}_2 & - \\ \dots & & \\ - & \mathbf{v}_n & - \end{bmatrix} \quad (2.26)$$

В соответствии со свойствами сингулярного разложения, первые  $r$  колонок матрицы  $\mathbf{U}$  образуют ортонормальный базис  $r$ -мерного линейного пространства, являющегося лучшим  $r$ -мерным приближением  $n$ -мерного подпространства протечки сигнала. Используем эти  $r$  векторов для построения проектора с уменьшенным рангом:

$$\mathbf{U}_r = \begin{bmatrix} & & & | \\ | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_r \\ | & | & & | \end{bmatrix}; \quad (2.27)$$

$$\mathbf{P}_r = \mathbf{I} - \mathbf{U}_r \mathbf{U}_r^T \quad (2.28)$$

Итак, мы построили оператор проекции ортогонально подпространству протечки сигнала  $\mathbf{P}_r$  с сокращенным рангом  $r$ . Умножение уравнения (2.21) на этот оператор приводит к тому, что из порождающей модели кросс-спектра на уровне сенсоров частично удаляются члены ответственные за эффект протечки сигнала; параметр  $r$  при этом определяет баланс между желаемым уровнем очистки от протечки сигнала и воздействием проектора на действительную часть кросс-спектра.

Окончательно, выражение для кросс-спектра на уровне сенсоров после проекции от протечки сигнала запишется в виде:

$$\text{vec}(\mathbf{C}^{\text{xx}})^\perp = \mathbf{P}_r \text{vec}(\mathbf{C}^{\text{xx}}) = \sum_{p=1}^n \sum_{q=1}^n \mathbf{P}_r \mathbf{g}_q \otimes \mathbf{g}_p \cdot c_{pq}^{\text{ss}}(t) + \mathbf{P}_r \text{vec}(\mathbf{C}^{\omega\omega}(t)) \quad (2.29)$$

Или, пользуясь 2.22:

$$\begin{aligned} \text{vec}(\mathbf{C}^{\text{xx}})^\perp &= \sum_{p \leq q}^{n,n} \mathbf{P}_r (\mathbf{g}_q \otimes \mathbf{g}_p + \mathbf{g}_p \otimes \mathbf{g}_q) \text{Re} (c_{pq}^{\text{ss}}(t)) + \\ &\quad + i \cdot \sum_{p < q}^{n,n} (\mathbf{g}_q \otimes \mathbf{g}_p - \mathbf{g}_p \otimes \mathbf{g}_q) \text{Im} (c_{pq}^{\text{ss}}(t)) + \mathbf{P}_r \text{vec}(\mathbf{C}^{\omega\omega}(t)) \end{aligned} \quad (2.30)$$

Отметим, что в мнимой части 2.30 оператор проецирования  $\mathbf{P}_r$  можно опустить в силу 2.23.

Элементы полученного таким образом векторизованного кросс-спектра  $\text{vec}(\mathbf{C}^{\text{xx}})^\perp$  теперь могут быть использованы для оценки коннективностей как на уровне сенсоров, так и в источниках.

### 2.1.5 Модели со свободной ориентацией диполя

С точки зрения анатомии каждая топография прямой модели  $\mathbf{G}$  представляет собой распределение электромагнитного поля, порождаемого так называемыми первичными токами, то есть токами, текущими через апикальные дендриты кортикальных пирамидальных нейронов. Так как апикальные дендриты расположены перпендикулярно к кортикальной мантии, первичные токи по отношению к кортикальной мантии также имеют нормальную ориентацию. Следовательно, точность прямой модели напрямую зависит от точности оценки вектора нормали к поверхности коры, а значит и от количества точек, используемых для аппроксимации поверхности коры. Вместе с тем, хотя современное моделирование с использованием метода магнитно-резонансной томографии позволяет получать весьма детальную реконструкцию мозга с размером аппроксимирующих сеток порядка нескольких сотен тысяч узлов, использование столь подробных сеток приводит к значительному ухудшению производительности алгоритмов, работающих в пространстве источников, вследствие высоких затрат по памяти и вычислительному времени при работе с большими массивами данных.

По этой причине использование разреженных сеток со сравнительно небольшим числом узлов стало общепринятой практикой. Как уже отмечалось выше, недостаток такого подхода состоит в том, что при прореживании сетки эффективная ориентация локальной нормали к коре становится неопределенной.

Неопределенность появляется из-за того, что после разрежения размер области на коре, соответствующей отдельно взятому узлу сетки, возрастает, и так как каждая такая область характеризуется в общем случае переменной кривизной, эффективная ориентация токового диполя, находящегося внутри отдельно взятой области аппроксимации, будет зависеть от того, в какой именно части области находился диполь в тот или иной момент записи.

Частично справиться с этим эффектом позволяют модели, со *свободной ориентацией диполя*, в которых эффективная ориентация становится дополнительным параметром, который необходимо оценить из данных [61].

Для учета свободной ориентации следует представить топографию в точке  $p$  коры в виде линейной комбинации трех ортогональных друг к другу векторов топографии, размещенных в одной точке:

$$\mathbf{g}_p = \mathbf{g}_p^x \xi + \mathbf{g}_p^y \eta + \mathbf{g}_p^z \zeta = [\mathbf{g}_p^x, \mathbf{g}_p^y, \mathbf{g}_p^z] \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} \quad (2.31)$$

В случае МЭГ измерений, для модели сферически симметричного проводника, магнитное поле от токового диполя с радиальной ориентацией равно нулю, и следовательно введенная тройка векторов может быть заменена парой диполей в плоскости, перпендикулярной радиальному направлению для каждого узла. Следовательно, уравнение (2.31) перепишется для МЭГ в виде

$$\mathbf{g}_p = \mathbf{g}_p^x \xi + \mathbf{g}_p^y \eta = [\mathbf{g}_p^x, \mathbf{g}_p^y] \begin{pmatrix} \xi \\ \eta \end{pmatrix} \quad (2.32)$$

Соответственно изменятся выражения и для 2-топографий проточки сигнала:

$$\begin{aligned}
\mathbf{g}_p \otimes \mathbf{g}_p &= (\mathbf{g}_p^x \xi + \mathbf{g}_p^y \eta) \otimes (\mathbf{g}_p^x \xi + \mathbf{g}_p^y \eta) = \\
&= \mathbf{g}_p^x \otimes \mathbf{g}_p^x \xi^2 + \mathbf{g}_p^x \otimes \mathbf{g}_p^y \xi \eta + \mathbf{g}_p^y \otimes \mathbf{g}_p^x \eta \xi + \mathbf{g}_p^y \otimes \mathbf{g}_p^y \eta^2 = \\
&= [\mathbf{g}_p^x \otimes \mathbf{g}_p^x, \mathbf{g}_p^x \otimes \mathbf{g}_p^y + \mathbf{g}_p^y \otimes \mathbf{g}_p^x, \mathbf{g}_p^y \otimes \mathbf{g}_p^y] \begin{pmatrix} \xi^2 \\ \xi \eta \\ \eta^2 \end{pmatrix} \tag{2.33}
\end{aligned}$$

Выражение в квадратных скобках здесь представляет собой матрицу, записанную по столбцам. Можно заметить, что 2-топографии протечки сигнала теперь принадлежат линейной оболочке трех векторов, а именно —  $\mathbf{g}_p^x \otimes \mathbf{g}_p^x$ ,  $\mathbf{g}_p^x \otimes \mathbf{g}_p^y + \mathbf{g}_p^y \otimes \mathbf{g}_p^x$  и  $\mathbf{g}_p^y \otimes \mathbf{g}_p^y$ . Разумеется, представленная таким образом 2-топография протечки сигнала зависит только от параметра  $\theta_p$  — угла между эффективной ориентацией топографии  $\mathbf{g}_p$  и ее  $x$ -компоненты — через  $\xi = \sin(\theta_p)$  и  $\eta = \cos(\theta_p)$ :

$$\mathbf{g}_p \otimes \mathbf{g}_p = [\mathbf{g}_p^x \otimes \mathbf{g}_p^x, \mathbf{g}_p^x \otimes \mathbf{g}_p^y + \mathbf{g}_p^y \otimes \mathbf{g}_p^x, \mathbf{g}_p^y \otimes \mathbf{g}_p^y] \begin{pmatrix} \cos(\theta_p)^2 \\ \cos(\theta_p) \sin(\theta_p) \\ \sin(\theta_p)^2 \end{pmatrix}$$

Но это однопараметрическое семейство векторов не может быть эффективно сведено к линейному векторному пространству с размерностью меньше 3 для построения соответствующего проектора, а значит в случае МЭГ для модели сферически симметричного проводника мы должны добавить все три 2-топографии источника  $p$  в матрицу  $F$ :

$$\mathbf{F} = [\mathbf{g}_1^x \otimes \mathbf{g}_1^x, \mathbf{g}_1^x \otimes \mathbf{g}_1^y + \mathbf{g}_1^y \otimes \mathbf{g}_1^x, \mathbf{g}_2^x \otimes \mathbf{g}_2^x, \dots, \mathbf{g}_n^y \otimes \mathbf{g}_n^y] \tag{2.34}$$

Рассмотрим теперь случай ЭЭГ-измерений. Так как в случае ЭЭГ линейные подпространства, натянутые на три вектора топографий в каждой точке, вообще говоря имеют разммерность 3, мы вынуждены учитывать каждый из векторов топографий при составлении матрицы  $\mathbf{F}$ . Тем не менее, расписав аналог выражения 2.33 для свободно ориентированного диполя в случае ЭЭГ, получим, что независимых векторов в подпространстве объемной проводимости для источника с индексом  $p$  будет всего 6 (вместо возможных 9):

$$\begin{aligned}
\mathbf{g}_p \otimes \mathbf{g}_p &= (\mathbf{g}_p^x \xi + \mathbf{g}_p^y \eta + \mathbf{g}_p^z \zeta) \otimes (\mathbf{g}_p^x \xi + \mathbf{g}_p^y \eta + \mathbf{g}_p^z \zeta) = \\
&= \mathbf{g}_p^x \otimes \mathbf{g}_p^x \xi^2 + \mathbf{g}_p^y \otimes \mathbf{g}_p^y \eta^2 + \mathbf{g}_p^z \otimes \mathbf{g}_p^z \zeta^2 + \\
&\quad + \mathbf{g}_p^x \otimes \mathbf{g}_p^y \xi \eta + \mathbf{g}_p^y \otimes \mathbf{g}_p^x \eta \xi + \\
&\quad + \mathbf{g}_p^x \otimes \mathbf{g}_p^z \xi \zeta + \mathbf{g}_p^z \otimes \mathbf{g}_p^x \zeta \xi + \\
&\quad + \mathbf{g}_p^y \otimes \mathbf{g}_p^z \eta \zeta + \mathbf{g}_p^z \otimes \mathbf{g}_p^y \zeta \eta = \\
&= [\mathbf{g}_p^x \otimes \mathbf{g}_p^x, \mathbf{g}_p^y \otimes \mathbf{g}_p^y, \mathbf{g}_p^z \otimes \mathbf{g}_p^z, \mathbf{g}_p^x \otimes \mathbf{g}_p^y + \mathbf{g}_p^y \otimes \mathbf{g}_p^x, \mathbf{g}_p^x \otimes \mathbf{g}_p^z + \mathbf{g}_p^z \otimes \mathbf{g}_p^x, \mathbf{g}_p^y \otimes \mathbf{g}_p^z + \mathbf{g}_p^z \otimes \mathbf{g}_p^y] \begin{pmatrix} \xi^2 \\ \eta^2 \\ \zeta^2 \\ \xi \eta \\ \xi \zeta \\ \eta \zeta \end{pmatrix}
\end{aligned}$$

Выражение в квадратных скобках вновь обозначает матрицу, записанную по столбцам, а  $\xi, \eta, \zeta$  — компоненты разложения диполя с истинной ориентацией по базису  $\mathbf{g}_p^x, \mathbf{g}_p^y, \mathbf{g}_p^z$ .

Все шесть колонок для каждого источника  $p$  матрицы в выражении выше для ЭЭГ должны быть добавлены в матрицу  $\mathbf{F}$  для последующего расчета операции проецирования.

Повторяя далее процедуру 2.26, 2.27, 2.28, получим оператор проекции от протечки сигнала в случае свободной ориентации диполя.

Рассмотрим теперь, как преобразуется уравнение 2.29 для свободной ориентации диполя в случае МЭГ и ЭЭГ. Проведем рассуждения сначала для ЭЭГ, а затем упростим для случая сферической симметрии в МЭГ.

Итак, теперь векторы  $\mathbf{g}_p, \mathbf{g}_q$  представляют собой линейную комбинацию троек дипольных векторов  $\mathbf{g}^x, \mathbf{g}^y, \mathbf{g}^z$ . Обозначим  $\xi_p, \eta_p, \zeta_p$  значения проекций вектора  $\mathbf{g}_p$  на  $\mathbf{g}_p^x, \mathbf{g}_p^y, \mathbf{g}_p^z$ , для вектора  $\mathbf{g}_q$  определим аналогичные величины  $\xi_q, \eta_q, \zeta_q$ . Тогда для 2-топографии сети  $p - q$ , пользуясь 2.16 будем иметь:

$$\begin{aligned} \mathbf{g}_p \otimes \mathbf{g}_q &= \text{vec} \left( \begin{bmatrix} | & | & | \\ \mathbf{g}_p^x & \mathbf{g}_p^y & \mathbf{g}_p^z \\ | & | & | \end{bmatrix} \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} \begin{pmatrix} \xi_q, \eta_q, \zeta_q \end{pmatrix} \begin{bmatrix} - & \mathbf{g}_q^x & - \\ - & \mathbf{g}_q^y & - \\ - & \mathbf{g}_q^z & - \end{bmatrix} \right) = \\ &= \begin{bmatrix} | & | & | \\ \mathbf{g}_q^x & \mathbf{g}_q^y & \mathbf{g}_q^z \\ | & | & | \end{bmatrix} \otimes \begin{bmatrix} | & | & | \\ \mathbf{g}_p^x & \mathbf{g}_p^y & \mathbf{g}_p^z \\ | & | & | \end{bmatrix} \cdot \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} \otimes \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} \quad (2.35) \end{aligned}$$

Пользуясь выражением 2.35 рассмотрим, как будут устроены топографии действительной и мнимой части для свободной ориентации диполя. Для компактности обозначим тройку векторов топографий для диполя с индексом  $k$  как  $\Gamma_k$ :

$$\boldsymbol{\Gamma}_k = \begin{bmatrix} | & | & | \\ \mathbf{g}_k^x & \mathbf{g}_k^y & \mathbf{g}_k^z \\ | & | & | \end{bmatrix} \quad (2.36)$$

Тогда для действительной части, пользуясь 2.20, получим:

$$\begin{aligned}
& \mathbf{g}_p \otimes \mathbf{g}_q + \mathbf{g}_q \otimes \mathbf{g}_p = \\
&= \boldsymbol{\Gamma}_q \otimes \boldsymbol{\Gamma}_p \cdot \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} \otimes \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} + \boldsymbol{\Gamma}_p \otimes \boldsymbol{\Gamma}_q \cdot \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} \otimes \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} = \\
&= (\boldsymbol{\Gamma}_q \otimes \boldsymbol{\Gamma}_p + \boldsymbol{\Gamma}_p \otimes \boldsymbol{\Gamma}_q) \cdot \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} \otimes \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} \quad (2.37)
\end{aligned}$$

В последнем равенстве мы поменяли порядок сомножителей в произведении Кронекера векторов  $(\xi, \eta, \zeta)^T$ . Здесь  $\mathbf{K}_3$  — матрица коммутации, определенная в соответствии с 2.18. В явном виде матрица  $K_3$  в соответствии с формулой 2.20 запишется как:

Аналогично для 2-топографий мнимой части будем иметь:

$$\begin{aligned}
 \mathbf{g}_p \otimes \mathbf{g}_q - \mathbf{g}_q \otimes \mathbf{g}_p &= \\
 &= \boldsymbol{\Gamma}_q \otimes \boldsymbol{\Gamma}_p \cdot \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} \otimes \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} - \boldsymbol{\Gamma}_p \otimes \boldsymbol{\Gamma}_q \cdot \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} \otimes \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} = \\
 &= (\boldsymbol{\Gamma}_q \otimes \boldsymbol{\Gamma}_p - \boldsymbol{\Gamma}_p \otimes \boldsymbol{\Gamma}_q \cdot \mathbf{K}_3) \cdot \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} \otimes \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} \quad (2.39)
 \end{aligned}$$

Для МЭГ в случае сферической симметрии соответствующие формулы будут выглядеть как

$$\mathbf{g}_p \otimes \mathbf{g}_q + \mathbf{g}_q \otimes \mathbf{g}_p = (\boldsymbol{\Gamma}_q \otimes \boldsymbol{\Gamma}_p + \boldsymbol{\Gamma}_p \otimes \boldsymbol{\Gamma}_q \cdot \mathbf{K}_2) \cdot \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} \otimes \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} \quad (2.40)$$

$$\mathbf{g}_p \otimes \mathbf{g}_q - \mathbf{g}_q \otimes \mathbf{g}_p = (\boldsymbol{\Gamma}_q \otimes \boldsymbol{\Gamma}_p - \boldsymbol{\Gamma}_p \otimes \boldsymbol{\Gamma}_q \cdot \mathbf{K}_2) \cdot \begin{pmatrix} \xi_q \\ \eta_q \\ \zeta_q \end{pmatrix} \otimes \begin{pmatrix} \xi_p \\ \eta_p \\ \zeta_p \end{pmatrix} \quad (2.41)$$

$$\mathbf{K}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.42)$$

Окончательно выражение формула 2.30 для ЭЭГ и сферически-симметричного случая МЭГ запишется как

$$\begin{aligned}
 vec(\mathbf{C}^{\text{xx}})^\perp &= \sum_{p \leq q}^{n,n} \mathbf{P}_r (\boldsymbol{\Gamma}_q \otimes \boldsymbol{\Gamma}_p + \boldsymbol{\Gamma}_p \otimes \boldsymbol{\Gamma}_q \cdot \mathbf{K}_l) \theta_q \otimes \theta_p Re(c_{pq}^{\text{ss}}(t)) + \\
 &+ i \cdot \sum_{p < q}^{n,n} (\boldsymbol{\Gamma}_q \otimes \boldsymbol{\Gamma}_p - \boldsymbol{\Gamma}_p \otimes \boldsymbol{\Gamma}_q \cdot \mathbf{K}_l) \theta_q \otimes \theta_p Im(c_{pq}^{\text{ss}}(t)) + \\
 &+ \mathbf{P}_r vec(\mathbf{C}^{\omega\omega}(t)) \quad (2.43)
 \end{aligned}$$

где  $l = 2$  для МЭГ и  $l = 3$  для ЭЭГ, а  $\theta_k$  — двухмерный или трехмерный вектор ориентации токового диполя.

Уравнение 2.43 послужит нам отправной точкой для дальнейшего анализа. Далее из него необходимо оценить комплексные коэффициенты матрицы кросс-спектральной плотности мощности в пространстве источников, а также неизвестные вектора ориентаций  $\theta_k$ .

## 2.2 PSICOS

Идея использования пространственной фильтрации в качестве метода борьбы с протечкой сигнала от совокупности всех активных источников была рассмотрена нами в разделе 1.3. Также в разделе 1.3 мы рассмотрели существующие методы решения обратной задачи для МЭЭГ.

Теперь мы можем использовать PSICOS-проекцию для оценки коннективности на уровне источников.

Процедура поиска источников состоит в следующем:

1. Находим кросс-спектр на уровне сенсоров
2. По посчитанной матрице прямой модели строим оператор проекции от протечки сигнала (PSICOS-проекция)
3. Векторизуем кросс-спектр на сенсорах и применяем к результату PSICOS-проекцию
4. По векторизованному спроектированному кросс-спектру на сенсорах ищем элементы кросс-спектра на источниках, используя фильтр, максимизирующий SNR в данной точке (см. 1.3.2).

Для оценки кросс-спектрального коэффициента в точке  $i, j$ , мы берем топографии соответствующих точек  $\mathbf{g}_i$ ,  $\mathbf{g}_j$  и строим по ним 2-топографию  $\mathbf{g}_j \otimes \mathbf{g}_i$  и нормируем к единице полученный вектор. Результат и будет искомым фильтром. Затем мы применяем данным (т.е. к векторизованному кросс-спектру) PSICOS-проекцию и применяем к спроектированным данным полученный фильтр.

Такая процедура предоставляет работающую на практике эвристику без строгого обоснования. Вместе с тем, существует более общий взгляд на операцию проекции от протечки сигнала, позволяющий получить описанный выше алгоритм более формально, а также дающий возможность сформулировать обобщения алгоритма PSICOS.

### 2.2.1 PSICOS как метод оптимальной фильтрации.

Протечка сигнала является главной проблемой при оценке коннективности при условии, что мы не хотим пренебрегать частью сигнала, содержащей информацию о коннективности с малыми задержками по фазе.

Важно понимать, что эта протечка не ограничивается проникновением друг в друга сигналов только между двумя источниками, для которых мы измеряем коннективность. Любой источник, протекающий *одновременно* в эти два приведет к появлению ложной коннективности.

Чтобы правильно оценить коннективность между парой источников, необходимо построить такой набор фильтров, который максимально подавлял бы источники, дающие большую протечку одновременно в два целевых источника. Построим такой фильтр.

В теории обратных задач вводят понятие ядра разрешения  $R(\mathbf{r}, \mathbf{r}')$  (resolution kernel, RK [62]): возьмем обратный оператор, восстанавливающий по измерениям активность в точке коры  $\mathbf{r}$ , и зафиксируем другую точку на коре,  $\mathbf{r}'$ . Функция  $R(\mathbf{r}, \mathbf{r}')$  показывает, сколько сигнала протекает из точки  $\mathbf{r}'$  в точку  $\mathbf{r}$  при оценке активности в точке  $\mathbf{r}$ . Для линейных методов решения обратной задачи функция  $R$  может быть записана как скалярное произведение:

$$R(i, k) = \mathbf{l}_i^T \mathbf{g}_k \quad (2.44)$$

Здесь  $\mathbf{l}_i$  — фильтр в направлении  $i$ -ой точки коры, а  $\mathbf{g}_j$  — топография точки коры с индексом  $k$ .

Для совокупности всех  $N$  точек коры можно определить векторнозначную функцию  $\mathbf{B}(i)$ , которая имеет смысл протечки сигнала от каждой точки коры в фиксированную точку  $i$ .

$$\mathbf{B}(i)^T = [R(i, 1), R(i, 2), \dots, R(i, N)] \quad (2.45)$$

Эта функция получила в литературе название Beam Response (BR) [62].

Зафиксируем теперь две точки коры  $i, j$ . Для этих точек степень перекрытия векторов  $\mathbf{B}(i)$  и  $\mathbf{B}(j)$  определяет, насколько мощной будет общая для двух точек компонента нежелательного сигнала. Так, если  $\mathbf{B}(i)$  и  $\mathbf{B}(j)$  имеют только неотрицательные компоненты и при этом ортогональны друг другу, общий сигнал, протекающий одновременно и в  $i$  и в  $j$  будет равен нулю. В этом случае

оценка коннективности будет полностью свободна от негативного эффекта протечки сигнала. Отсюда возникает желание построить такой набор фильтров, который минимизирует перекрытие  $\mathbf{B}(i)$  и  $\mathbf{B}(j)$  для каждой пары точек.

Для построения таких фильтров необходимо определить, как измерять перекрытие для векторов BR. В качестве меры перекрытия напрашивается скалярное произведение векторов  $\mathbf{B}(i)$ ,  $\mathbf{B}(j)$ : если оно равно нулю, т.е. вектора ортогональны, общая протечка отсутствует. Такая мера перекрытия, однако, является ошибочной в силу возможной знакопеременности компонент векторов  $\mathbf{B}(i)$ ,  $\mathbf{B}(j)$ . Например, в случае, когда есть всего два источника, а функции  $\mathbf{B}$  для них равны  $(1, -1)$  и  $(1, 1)$ , формально BR будут ортогональны, однако протечка сигнала между соответствующими источниками будет весьма значительной.

Чтобы преодолеть проблему знакопеременных компонент, будем измерять степень перекрытия как скалярное произведение векторов  $\mathbf{B}(i)$ ,  $\mathbf{B}(j)$ , покомпонентно возведённых в квадрат. Результат этого скалярного произведения, функцию  $\mu(i, j)$ , назовём *взаимной протечкой сигнала* для точек  $i, j$ :

$$\mu(i, j) \stackrel{\text{def}}{=} (\mathbf{B}(i)^{\odot 2})^T \mathbf{B}(j)^{\odot 2} \quad (2.46)$$

Здесь операция  $(\cdot)^{\odot 2}$  означает поэлементное возвведение в квадрат.

С учетом соотношений 2.44, 2.16 взаимную протечку сигнала можно переписать как

$$\begin{aligned} \mu(i, j) &= \sum_k (\mathbf{l}_i^T \mathbf{g}_k)^2 (\mathbf{l}_j^T \mathbf{g}_k)^2 = \sum_k (\mathbf{l}_i^T \mathbf{g}_k \mathbf{g}_k^T \mathbf{l}_j)^2 = (\mathbf{l}_j \otimes \mathbf{l}_i)^T \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T (\mathbf{l}_j \otimes \mathbf{l}_i), \\ \boldsymbol{\Gamma} &= (\mathbf{g}_1 \otimes \mathbf{g}_1, \dots, \mathbf{g}_N \otimes \mathbf{g}_N) \end{aligned} \quad (2.47)$$

Мы хотим найти такую пару фильтров  $\mathbf{l}_i, \mathbf{l}_j$ , которая минимизировала бы взаимную протечку сигнала для точек  $i, j$ , сохраняя при этом полезный сигнал. Для этого необходимо дополнительно наложить ограничение на длину фильтров. Одним из возможных ограничений является требование, чтобы коэффициент усиления сигнала, приходящего из точек с топографиями  $\mathbf{g}_i, \mathbf{g}_j$  равнялся единице. С учетом этого ограничения оптимизационная задача на поиск пары фильтров  $\mathbf{l}_i, \mathbf{l}_j$  запишется как

$$\frac{1}{2} (\mathbf{l}_j \otimes \mathbf{l}_i)^T \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T (\mathbf{l}_j \otimes \mathbf{l}_i) \rightarrow \min \quad (2.48)$$

$$s.t. : \mathbf{l}_i^T \mathbf{g}_i = 1, \mathbf{l}_j^T \mathbf{g}_j = 1 \quad (2.49)$$

Такая задача оптимизации не разрешается в явном виде методом множителей Лагранжа из-за произведения Кронекера в целевом функционале. Ее необходимо решать численно, например методом градиентного спуска. Тем не менее, эту задачу можно значительно упростить, обобщив понятие фильтров на пространство размерности  $M^2$ .

Отметим, что из ограничений задачи 2.49 следует, что

$$(\mathbf{l}_j \otimes \mathbf{l}_i)^T (\mathbf{g}_j \otimes \mathbf{g}_i) = 1 \quad (2.50)$$

Теперь воспользуемся тем, что ограничение, записанное в таком виде, как и целевая функция в 2.49 зависят от фильтров  $\mathbf{l}_i, \mathbf{l}_j$  только через их кронекеровское произведение. Обозначим это произведение как  $\mathbf{v}_{ij}$ . Для простоты также обозначим  $\mathbf{g}_j \otimes \mathbf{g}_i$  как  $\mathbf{q}_{ij}$ .

$$\frac{1}{2} \mathbf{v}_{ij}^T \mathbf{\Gamma} \mathbf{\Gamma}^T \mathbf{v}_{ij} \rightarrow \min \quad (2.51)$$

$$s.t. : \mathbf{v}_{ij}^T \mathbf{q}_{ij} = 1 \quad (2.52)$$

Такая задача относительно переменной  $\mathbf{v}_{ij}$  легко решается методом множителей Лагранжа. Ее лагранжиан запишется как

$$L(\mathbf{v}_{ij}, \lambda) = \frac{1}{2} \mathbf{v}_{ij}^T \mathbf{\Gamma} \mathbf{\Gamma}^T \mathbf{v}_{ij} + \lambda (\mathbf{v}_{ij}^T \mathbf{q}_{ij} - 1) \quad (2.53)$$

Дифференцируя лагранжиан по  $\mathbf{v}_{ij}$ , получим

$$\frac{\partial L(\mathbf{v}_{ij}, \lambda)}{\partial \mathbf{v}_{ij}} = \mathbf{\Gamma} \mathbf{\Gamma}^T \mathbf{v}_{ij} + \lambda \mathbf{q}_{ij} = 0 \quad (2.54)$$

Если  $\lambda \neq 0$ , уравнение 2.54 не имеет решений: столбцы матрицы  $\mathbf{\Gamma}$  являются векторизациями симметричных матриц, а значит таковым является и произведение  $\mathbf{\Gamma} \mathbf{\Gamma}^T \mathbf{v}_{ij}$ ; при этом  $\mathbf{q}_{ij}$  векторизацией симметричной матрицы не является. Это означает, что  $\lambda = 0$ , и уравнение на  $\mathbf{v}_{ij}$  выглядит как

$$\mathbf{\Gamma} \mathbf{\Gamma}^T \mathbf{v}_{ij} = 0 \quad (2.55)$$

Следовательно что вектор  $\mathbf{v}_{ij}$  должен принадлежать ортогональному дополнению линейной оболочки столбцов матрицы  $\mathbf{\Gamma}$ . Это значит, что вектор  $\mathbf{v}_{ij}$  можно получить как результат применения проекции  $\mathbf{P} = \mathbf{I} - \mathbf{\Gamma} \mathbf{\Gamma}^\dagger$  к некоторому вектору  $\tilde{\mathbf{v}}_{ij}$  из общего линейного пространства размерностью  $M^2$ .

Продифференцируем теперь лагранжиан 2.53 по  $\lambda$ :

$$\frac{\partial L(\mathbf{v}_{ij}, \lambda)}{\partial \lambda} = \mathbf{v}_{ij}^T \mathbf{q}_{ij} - 1 = 0 \quad (2.56)$$

С учетом общего вида вектора  $\mathbf{v}_{ij}$ , а также идемпотентности и симметричности оператора проекции получим

$$\tilde{\mathbf{v}}_{ij}^T \mathbf{P}^T \mathbf{q}_{ij} = \tilde{\mathbf{v}}_{ij}^T \mathbf{P} \mathbf{q}_{ij} = \tilde{\mathbf{v}}_{ij}^T \mathbf{P}^2 \mathbf{q}_{ij} = (\mathbf{P} \tilde{\mathbf{v}}_{ij})^T \mathbf{P} \mathbf{q}_{ij} = \mathbf{v}_{ij}^T \mathbf{P} \mathbf{q}_{ij} = 1 \quad (2.57)$$

Последнее уравнение задает гиперплоскость в подпространстве векторов, ортогональных столбцам  $\mathbf{\Gamma}$ . Следовательно поставленная задача оптимизации имеет бесконечное количество решений. Для выбора единственного решения можем как и раньше воспользоваться критерием минимальности нормы. Среди векторов этой гиперплоскости минимальным по норме вектором является

$$\mathbf{v}_{ij} = \frac{\mathbf{P} \mathbf{q}_{ij}}{\|\mathbf{P} \mathbf{q}_{ij}\|^2} \quad (2.58)$$

Таким образом получаем, что фильтр, минимизирующий взаимную протечку сигнала для точек коры  $i, j$  получается как результат проекции, полученной нами в разделе 2.1.4, примененной к соответствующей 2-топографии.

Отметим, что оператор проекции  $\mathbf{P}$  действует как проекция на множество, состоящее из объединения множества векторизованных антисимметричных матриц и векторизованных симметричных матриц, не принадлежащих линейной оболочке 2-топографий, соответствующих объемной проводимости. В предельном случае, когда количество точек на коре превосходит величину  $M(M+1)/2$ , равную общему количеству симметричных матриц размером  $M \times M$ , применение оператора  $\mathbf{P}$  полностью обнуляет векторизованные симметричные матрицы и оставляет антисимметричные матрицы нетронутыми. Фактически это эквивалентно взятию мнимой части кросс-спектра, так как топографии мнимой части антисимметричны и не изменяются проекцией.

Вместе с тем при слишком большом количестве источников численная размерность линейной оболочки столбцов матрицы  $\mathbf{\Gamma}$  будет существенно меньше взятого количества источников, а наличие шумовых собственных чисел в матрице  $\mathbf{\Gamma} \mathbf{\Gamma}^T$  приведет к чрезмерному удалению полезного сигнала из действительной части кросс-спектра. Поэтому для адекватного удаления взаимной протечки сигнала лучше искусственно ограничивать ранг проекции, например фиксируя шумовой уровень для собственных чисел матрицы  $\mathbf{\Gamma} \mathbf{\Gamma}^T$ . Ограничение ранга

проекции фактически эквивалентно регуляризации функционала, которую мы рассмотрим далее.

Отметим также, что полученный фильтр 2.58 оказывается невозмож но разложить на кронекеровское произведение двух фильтров, действующих в пространстве размерности  $M$ . Тем не менее, его применение к матрице кросс-спектральной плотности мощности на сенсорах дает оценку элемента кросс-спектра на источниках с желаемыми свойствами.

### 2.2.2 Пространственная смещённость оценки и нормализация весов.

Проанализируем полученные выражения для фильтров 2.58 с точки зрения пространственной смещённости оценок. Будем считать, что оценка является несмешённой, если фильтр  $\mathbf{v}_{ij}$  среди всех возможных фильтров имеет максимум на 2-топографии  $\mathbf{q}_{ij}$ .

Применим произвольный фильтр  $\mathbf{v}_{rs} = \mathbf{P}\mathbf{q}_{rs}/\|\mathbf{P}\mathbf{q}_{rs}\|^2$  к 2-топографии  $\mathbf{q}_{ij}$ :

$$\frac{\mathbf{q}_{rs}^T \mathbf{P} \mathbf{q}_{ij}}{\|\mathbf{P} \mathbf{q}_{rs}\|^2} = \frac{\|\mathbf{P} \mathbf{q}_{ij}\|}{\|\mathbf{P} \mathbf{q}_{rs}\|} \cos(\mathbf{P} \mathbf{q}_{rs}, \mathbf{P} \mathbf{q}_{ij}) \quad (2.59)$$

Как видим, применение произвольного фильтра к нашей 2-топографии может давать величину как меньше 1, так и больше. При этом целевой фильтр  $\mathbf{v}_{ij}$  на этой топографии всегда даёт усиление, равное 1, а значит оценка, получаемая при помощи фильтра 2.58, является пространственно смещённой.

Отметим, что базовая процедура оценки элементов кросс-спектра на уровне источников, описанная в разделе 2.2 также будет давать смещённую оценку, т.к. мы нормализовали фильтр  $\mathbf{g}_j \otimes \mathbf{g}_i$  без учета оператора проекции  $\mathbf{P}$ .

Для построения несмешённой оценки можно воспользоваться нормированной фильтров аналогично той, которая используется в методе sLORETA [43]. Нормированный фильтр будет выглядеть следующим образом:

$$\mathbf{v}_{ij} = \frac{\mathbf{P} \mathbf{q}_{ij}}{\|\mathbf{P} \mathbf{q}_{ij}\|} \quad (2.60)$$

При такой нормировке применение фильтра  $\mathbf{v}_{rs} = \mathbf{P}\mathbf{q}_{rs}/\|\mathbf{P}\mathbf{q}_{rs}\|$  к топографии  $\mathbf{q}_{ij}$  по сравнению с фильтром  $\mathbf{v}_{ij}$  будет выглядеть как

$$\frac{\mathbf{q}_{rs}^T \mathbf{P} \mathbf{q}_{ij}}{\|\mathbf{P} \mathbf{q}_{rs}\|} = \|\mathbf{P} \mathbf{q}_{ij}\| \cos(\mathbf{P} \mathbf{q}_{rs}, \mathbf{P} \mathbf{q}_{ij}) \leq \frac{\mathbf{q}_{ij}^T \mathbf{P} \mathbf{q}_{ij}}{\|\mathbf{P} \mathbf{q}_{ij}\|} = \|\mathbf{P} \mathbf{q}_{ij}\| \quad (2.61)$$

Таким образом, фильтры вида 2.60 дают согласно определению, принятому в теории обратных задач, пространственно несмещенную оценку.

Модификацию метода PSICOS с нормализацией фильтров вида 2.61 мы будем называть далее по тексту PSICOS Unbiased.

### 2.2.3 Действие фильтра на мнимую и действительную часть кросс-спектра. PSICOS Unbiased.

Если в случае с решением обратных задач для восстановления активности в фиксированной точке коры мы имели дело лишь с одной топографией, то теперь, при восстановлении элемента матрицы кросс-спектральной плотности в пространстве источников нам необходимо отдельно рассматривать действительную и мнимую часть кросс-спектра.

При выводе порождающей модели кросс-спектра в разделе 2.1.4 мы получили, что 2-топографии действительной и мнимой части кросс-спектра имеют разную структуру: действительная часть состоит из векторизации симметричных матриц, тогда как мнимая — из векторизаций антисимметричных.

Полученный нами фильтр 2.60 обладает симметричной структурой, а значит никак не влияет на мнимую часть кросс-спектра. При этом когда мы анализировали пространственную смещеннность оценки, мы никак не учитывали, что вместе с 2-топографией  $\mathbf{q}_{ij}$  в действительную часть спектра всегда входит источник с сопряженным элементом кросс-спектра на источниках и 2-топографией  $\mathbf{q}_{ji}$ .

Проанализируем пространственную смещеннность оценки полученным фильтром отдельно для действительной и мнимой 2-топографий элемента  $i, j$ .

2-топография действительной части запишется как  $\mathbf{q}_{ij} + \mathbf{q}_{ji}$ . При этом сам вектор  $\mathbf{q}_{ij}$  может быть расписан как

$$\mathbf{q}_{ij} = \frac{1}{2}(\mathbf{q}_{ij} + \mathbf{q}_{ji}) + \frac{1}{2}(\mathbf{q}_{ij} - \mathbf{q}_{ji}) \quad (2.62)$$

Тогда действие фильтра  $\mathbf{v}_{rs}$  на 2-топографию действительной части для точки с индексами  $i, j$  с учетом ортогональности симметричных и антисимметричных слагаемых запишется как

$$\begin{aligned} & \frac{\left(\frac{1}{2}(\mathbf{q}_{rs} + \mathbf{q}_{sr}) + \frac{1}{2}(\mathbf{q}_{rs} - \mathbf{q}_{sr})\right)^T \mathbf{P} (\mathbf{q}_{ij} + \mathbf{q}_{ji})}{\sqrt{\left(\frac{1}{2}(\mathbf{q}_{rs} + \mathbf{q}_{sr}) + \frac{1}{2}(\mathbf{q}_{rs} - \mathbf{q}_{sr})\right)^T \mathbf{P} \left(\frac{1}{2}(\mathbf{q}_{rs} + \mathbf{q}_{sr}) + \frac{1}{2}(\mathbf{q}_{rs} - \mathbf{q}_{sr})\right)}} = \\ &= \frac{(\mathbf{q}_{rs} + \mathbf{q}_{sr})^T \mathbf{P} (\mathbf{q}_{ij} + \mathbf{q}_{ji})}{\sqrt{(\mathbf{q}_{rs} + \mathbf{q}_{sr})^T \mathbf{P} (\mathbf{q}_{rs} + \mathbf{q}_{sr}) + (\mathbf{q}_{rs} - \mathbf{q}_{sr})^T (\mathbf{q}_{rs} - \mathbf{q}_{sr})}} = \\ &= \frac{\cos(\mathbf{P}(\mathbf{q}_{rs} + \mathbf{q}_{sr}), \mathbf{P}(\mathbf{q}_{ij} + \mathbf{q}_{ji})) \|\mathbf{P}(\mathbf{q}_{ij} + \mathbf{q}_{ji})\|}{\sqrt{1 + \|\mathbf{q}_{rs} - \mathbf{q}_{sr}\|^2 / \|\mathbf{P}(\mathbf{q}_{rs} + \mathbf{q}_{sr})\|^2}} \quad (2.63) \end{aligned}$$

Как видим, для действительной части максимальное значение для точки  $i, j$  не обязательно достигается на фильтре  $\mathbf{v}_{ij}$ . Получается, что при рассмотрении только действительной части оценка оказывается смещенной.

Можем получить аналогичную формулу для мнимой части:

$$\begin{aligned} & \frac{\left(\frac{1}{2}(\mathbf{q}_{rs} + \mathbf{q}_{sr}) + \frac{1}{2}(\mathbf{q}_{rs} - \mathbf{q}_{sr})\right)^T \mathbf{P} (\mathbf{q}_{ij} - \mathbf{q}_{ji})}{\sqrt{\left(\frac{1}{2}(\mathbf{q}_{rs} + \mathbf{q}_{sr}) + \frac{1}{2}(\mathbf{q}_{rs} - \mathbf{q}_{sr})\right)^T \mathbf{P} \left(\frac{1}{2}(\mathbf{q}_{rs} + \mathbf{q}_{sr}) + \frac{1}{2}(\mathbf{q}_{rs} - \mathbf{q}_{sr})\right)}} = \\ &= \frac{\cos(\mathbf{q}_{rs} - \mathbf{q}_{sr}, \mathbf{q}_{ij} - \mathbf{q}_{ji}) \|\mathbf{q}_{ij} - \mathbf{q}_{ji}\|}{\sqrt{1 + \|\mathbf{q}_{rs} - \mathbf{q}_{sr}\|^2 / \|\mathbf{P}(\mathbf{q}_{rs} + \mathbf{q}_{sr})\|^2}} \quad (2.64) \end{aligned}$$

Здесь оценка также оказывается смещенной.

Чтобы избежать этой проблемы, изменим нормировку нашего фильтра. Менять при этом будем по-разному для действительной и для мнимой части, получая таким образом два различных фильтра:

$$\mathbf{v}_{ij}^{Re} = \frac{2\mathbf{P}\mathbf{q}_{ij}}{\|\mathbf{P}(\mathbf{q}_{ij} + \mathbf{q}_{ji})\|} \quad (2.65)$$

$$\mathbf{v}_{ij}^{Im} = \frac{2\mathbf{q}_{ij}}{\|\mathbf{q}_{ij} - \mathbf{q}_{ji}\|} \quad (2.66)$$

Для такой нормировки оценка и по действительной и по мнимой частям отдельно является несмещенной. Несмещенной также будет оценка по мощности кросс-спектрального коэффициента  $i, j$ , т.е. по корню из суммы квадратов мнимой и действительной частей.

## 2.2.4 Регуляризация.

Как уже отмечалось, при достаточно большом количестве точек коры проекция, построенная без ограничения ранга проекции будет удалять слишком большую долю сигнала из действительной части.

Ограничение ранга проекции позволяет справиться с этой проблемой, однако такой подход не дает возможности однозначно выбрать решение для фильтров: после проекции мы получали бесконечное множество решений, лежащих в плоскости, ортогональной спроектированной 2-топографии для точки  $i, j$ . Выбор решения осуществлялся нами на основе ad hoc эвристики, что фильтр должен обладать минимально возможной нормой.

Здесь мы рассмотрим альтернативный способ решения этой проблемы. Способ основан на тихоновской регуляризации. Его суть состоит в том, что мы сразу закладываем в целевой функционал слагаемое, штрафующее норму фильтра. Задача оптимизации в этом случае запишется как

$$\frac{1}{2} \mathbf{v}_{ij}^T \mathbf{\Gamma} \mathbf{\Gamma}^T \mathbf{v}_{ij} + C \frac{1}{2} \|\mathbf{v}_{ij}\|^2 \rightarrow \min \quad (2.67)$$

$$s.t. : \mathbf{v}_{ij}^T \mathbf{q}_{ij} = 1 \quad (2.68)$$

Лагранжиан для этой задачи оптимизации будет выглядеть как

$$L(\mathbf{v}_{ij}, \lambda) = \frac{1}{2} \mathbf{v}_{ij}^T \mathbf{\Gamma} \mathbf{\Gamma}^T \mathbf{v}_{ij} + C \frac{1}{2} \|\mathbf{v}_{ij}\|^2 + \lambda (\mathbf{v}_{ij}^T \mathbf{q}_{ij} - 1) \quad (2.69)$$

Возьмем его производную по  $\mathbf{v}_{ij}$ :

$$\frac{\partial L(\mathbf{v}_{ij}, \lambda)}{\partial \mathbf{v}_{ij}} = \mathbf{\Gamma} \mathbf{\Gamma}^T \mathbf{v}_{ij} + C \mathbf{v}_{ij} + \lambda \mathbf{q}_{ij} = 0 \quad (2.70)$$

На этот раз благодаря регуляризации уравнение 2.70 имеет решение для ненулевого  $\lambda$ :

$$\mathbf{v}_{ij} = -\lambda (\mathbf{\Gamma} \mathbf{\Gamma}^T + C \mathbf{I})^{-1} \mathbf{q}_{ij} \quad (2.71)$$

Домножая левую и правую часть на  $\mathbf{q}_{ij}$  и учитывая ограничение  $\mathbf{v}_{ij}^T \mathbf{q}_{ij} = 1$ , получим выражение для  $\lambda$ :

$$\lambda = -\frac{1}{\mathbf{q}_{ij}^T (\mathbf{\Gamma} \mathbf{\Gamma}^T + C \mathbf{I})^{-1} \mathbf{q}_{ij}} \quad (2.72)$$

Для фильтров  $\mathbf{v}_{ij}$  будем иметь

$$\mathbf{v}_{ij} = \frac{(\mathbf{\Gamma}\mathbf{\Gamma}^T + C\mathbf{I})^{-1}\mathbf{q}_{ij}}{\mathbf{q}_{ij}^T(\mathbf{\Gamma}\mathbf{\Gamma}^T + C\mathbf{I})^{-1}\mathbf{q}_{ij}} \quad (2.73)$$

Фильтры, дающие пространственно несмещенную оценку, получим аналогично 2.2.2:

$$\mathbf{v}_{ij}^{Re} = \frac{2(\mathbf{\Gamma}\mathbf{\Gamma}^T + C\mathbf{I})^{-1}\mathbf{q}_{ij}}{\sqrt{(\mathbf{q}_{ij} + \mathbf{q}_{ji})^T(\mathbf{\Gamma}\mathbf{\Gamma}^T + C\mathbf{I})^{-1}(\mathbf{q}_{ij} + \mathbf{q}_{ji})}} \quad (2.74)$$

$$\mathbf{v}_{ij}^{Im} = \frac{2(\mathbf{\Gamma}\mathbf{\Gamma}^T + C\mathbf{I})^{-1}\mathbf{q}_{ij}}{\sqrt{(\mathbf{q}_{ij} - \mathbf{q}_{ji})^T(\mathbf{\Gamma}\mathbf{\Gamma}^T + C\mathbf{I})^{-1}(\mathbf{q}_{ij} - \mathbf{q}_{ji})}} \quad (2.75)$$

Для мнимой части в силу ортогональности симметричных и антисимметричных векторов можем переписать фильтр в виде

$$\mathbf{v}_{ij}^{Im} = \frac{2}{\sqrt{C}} \frac{\mathbf{q}_{ij}}{\|(\mathbf{q}_{ij} - \mathbf{q}_{ji})\|} \quad (2.76)$$

## 2.2.5 Нормализация кросс-спектральных коэффициентов в пространстве источников. PSICOS Normalized.

Полученные после фильтрации элементы кросс-спектра в пространстве источников нуждаются в нормализации, так как их величина зависит не только от постоянства разности фаз между двумя источниками, но и от мощности взаимодействующих источников. Это приводит к тому, что ненормализованный кросс-спектральный коэффициент для оценки функциональной коннективности дает смещение в сторону более мощных источников.

Другим негативным эффектом отсутствия нормализации является тот факт, что по абсолютной величине кросс-спектрального коэффициента нельзя сделать вывод о постоянстве разности фаз. С другой стороны нормированная величина позволяет определить порог, по которому можно судить о наличии или отсутствии фазового взаимодействия.

В чем сложность нормализации при использовании метода PSICOS? При оценке когерентности используется нормализация на мощность источников. При использовании PSICOS-проекции нормализация на мощность оказывается

невозможной, так как мощностная компонента, присутствующая в диагональной части кросс-спектра, удаляется из сигнала. Если же оценивать мощность по неспроектированным данным, мощностная компонента, сильно загрязненная протечкой сигнала, будет давать сильные искажения в оценку.

При оценке элемента кросс-спектра мы усредняли комплекснозначные величины для отдельных эпох и временных срезов. Каждая такая величина представляла собой вектор на плоскости, повернутый на разность фаз между двумя сигналами относительно оси абсцисс с длиной, пропорциональной произведению амплитуд двух сигналов. Такое усреднение приводило к тому, что вектора с постоянной разностью фаз усиливали друг друга, а вектора со случайной разностью фаз — ослабляли. Тем не менее, средний вектор все равно оказывался пропорционален амплитудам сигналов: мощные источники со случайной фазовой задержкой могли дать кросс-спектральный коэффициент больше, чем слабые источники со стабильной разностью фаз.

Для оценки мощности мы можем изменить стратегию при усреднении. Будем усреднять квадраты длин соответствующих 2-мерных векторов. Тогда полученное среднее будет давать оценку произведения мощностей взаимодействующих сигналов. Нормализация элемента кросс-спектра на источниках после применения PSICOS-проекции на корень из этой величины будет давать нормализованный коэффициент, ограниченный по модулю (кросс-спектральный коэффициент — комплексная величина) отрезком  $[0, 1]$ .

Похожая стратегия нормализации используется в методе wPLI [27]. Эта мера основана на мнимой части кросс-спектра, поэтому информация о мощности отдельных источников в ней отсутствует. Для нормализации используется среднее значение модуля мнимой части “кросс-спектра”, посчитанное для каждой эпохи отдельно.

Модификацию метода PSICOS с описанной выше нормализацией кроссспектральных коэффициентов мы будем называть PSICOS Normalized.

## 2.3 GO-PSIICOS

Применение метода проекции для очистки от протечки сигнала, рассмотренное нами до сих пор, позволяет оценить элементы матрицы кросс-спектральной плотности на источниках оптимально в локальном, но не глобальном смысле. Один из недостатков такого подхода заключается в том, что локально оптимальные методы не позволяют использовать априорную информацию о пространственном распределении сетей. В частности, они не позволяют получать “спарсные”, или разреженные по пространству решения — свойство, которого легко добиться, используя методы глобальной оптимизации.

Свойство спарсности при поиске сетей оказывается крайне желательным по следующей причине. Дело в том, что протечка сигнала при поиске сетей проявляет себя не только на уровне отдельных источников, между которыми мы хотим установить наличие фазовой связности, но и на уровне сетей. Проявляется этот эффект в том, что для двух истинно взаимодействующих точек коры А и В мы увидим, что все точки окрестности точки А взаимодействуют со всеми точками окрестности точки В. Это обстоятельство было отмечено в статье [63] и получило название “призрачные взаимодействия”.

Далее мы рассмотрим модификацию метода PSIICOS, которая, используя методы глобальной оптимизации, а именно оценку методом смешанной нормы, позволяет справиться и с этим негативным эффектом протечки сигнала — проблемой призрачных взаимодействий или ложноположительными срабатываниями второго рода.

### 2.3.1 Оценка методом смешанной нормы.

Оценка методом смешанной нормы, как уже было сказано, основана на комбинировании различных норм в регуляризационном слагаемом, штрафующих строки и столбцы матрицы  $\mathbf{S}$  по-разному. Эта идея формализуется при помощи понятия *смешанной нормы*, которая для матрицы  $\mathbf{X}$  определяется как:

$$\|\mathbf{X}\|_{p,q} \stackrel{\text{def}}{=} \left( \sum_i \left( \sum_j X_{ij}^p \right)^{\frac{q}{p}} \right)^{\frac{1}{q}} \quad (2.77)$$

Как видим, смешанная норма представляет собой последовательное вычисление норм  $L_p$  и  $L_q$ , при этом вторая норма берется от вектора-столбца значений первой нормы на строках матрицы  $\mathbf{X}$ .

Использование смешанной нормы вместо обычной в регуляризационном слагаемом функционала качества позволяет получать решения с желаемыми свойствами пространственной спарсности и временной гладкости в случае, если первая штрафующая норма порождает гладкие решения (например,  $L_2$ ), а вторая — спарсные (например,  $L_1$ ).

Такой подход однако требует дополнительных усилий при численной оптимизации полученного функционала качества, так как он, как и в случае метода МСЕ, представляет собой негладкую функцию от активаций на коре.

Свойство разреженности порождаемых решений при этом может быть дополнительно усилено в случае, если в качестве второй нормы мы используем квазинорму  $L_q$  с коэффициентом  $q$ , принимающим значения в промежутке от 0 до 1. В этом случае за дополнительную спарсность придется расплачиваться еще большими сложностями в оптимизации функционала качества, так как при использовании дробной квазинормы целевой функционал не только не является гладким, но также не является выпуклым. Тем не менее, такая оптимизация, как мы рассмотрим далее, возможна, хоть и не гарантирована от сходимости к локальным минимумам.

Но сначала рассмотрим стратегию оптимизации функционала, получающегося при использовании смешанных норм, обладающих спарсностью по пространству, для нормы  $L_{21}$ .

### 2.3.2 Оптимизация функционала с нормой $L_{21}$

При оптимизация функционала с нормами, порождающими спарсные решения, такими как норма  $L_1$ , решение в общем случае не может быть получено в явном виде, что уже отмечалось ранее. Тем не менее, в частном случае регуляризации нормой  $L_1$ , когда матрица прямой модели является единичной,

оказывается возможным выразить решение в явном виде [64] (в этом случае задача фактически сводится к удалению шума из вектора наблюдений  $\mathbf{x}$ ). Рассмотрим этот случай, так как он поможет нам построить итерационную схему для общего случая с произвольной матрицей  $\mathbf{G}$  и рассмотреть обобщение этой схемы на случай смешанной нормы, полученное в статье [65]. Запишем сначала функционал качества:

$$f(\hat{\mathbf{s}}) = \frac{1}{2} \|\hat{\mathbf{s}} - \mathbf{x}_{noisy}\|_{L_2}^2 + \alpha \|\hat{\mathbf{s}}\|_{L_1} \rightarrow \min_{\hat{\mathbf{s}}} \quad (2.78)$$

Полученный функционал, как нетрудно заметить, распадается на отдельные неотрицательные слагаемые, каждое из которых зависит только от одной компоненты вектора  $\hat{\mathbf{s}}$ , а значит минимизация всего функционала 2.78 достигается минимизацией каждого из слагаемых вида

$$y_i = \frac{1}{2}(\hat{s}_i - x_i)^2 + \alpha |\hat{s}_i| \quad (2.79)$$

Критерием оптимума выражения 2.79 является принадлежность нуля его субдифференциалу:  $0 \in \partial y_i$ . Для каждого из слагаемых  $y_i$  везде кроме  $\hat{s}_i = 0$  выражение 2.79 является дифференцируемым. Поэтому для  $\hat{s}_i \neq 0$  условие  $0 \in \partial y_i$  эквивалентно равенству нулю производной по  $\hat{s}_i$ :

$$\frac{dy_i}{d\hat{s}_i} = (\hat{s}_i - x_i) + \alpha \operatorname{sign}(\hat{s}_i) = 0 \quad (2.80)$$

С учетом ограничений на знак  $\hat{s}_i$  получим, что  $\hat{s}_i = x_i - \alpha$ , если  $x_i > \alpha$ , и  $\hat{s}_i = x_i + \alpha$ , если  $x_i < -\alpha$

В случае  $\hat{s}_i = 0$  субдифференциал 2.79 является отрезком  $[-x_i - \alpha, -x_i + \alpha]$ . Этот отрезок содержит ноль тогда и только тогда, когда  $|x_i| \leq \alpha$ . То есть, если  $|x_i| \leq \alpha$ , значение  $\hat{s}_i = 0$  является оптимальным.

Окончательно получим, что для  $\hat{s}_i$  возможно три значения в зависимости от того, к какому промежутку принадлежит  $x_i$ :

$$\hat{s}_i = h(x_i, \alpha) = \begin{cases} x_i - \alpha, & x_i > \alpha \\ 0, & |x_i| \leq \alpha \\ x_i + \alpha, & x_i < -\alpha \end{cases} \quad (2.81)$$

Отметим также, что оператор  $h(x, c)$  может быть переписан в более компактной форме:

$$h(x, c) = x \left(1 - \frac{c}{|x|}\right)^+, \text{ где } (\cdot)^+ \stackrel{\text{def}}{=} \max(\cdot, 0) \quad (2.82)$$

Таким образом, в рассмотренной нами простейшей постановке применение  $L_1$ -регуляризатора порождает решения, которые получаются из измерений по компонентным применением оператора  $h(x, c)$ , который обнуляет компоненты вектора измерений, не превосходящие по амплитуде порог  $c$ . Такой оператор известен в литературе под названием soft thresholding.

### 2.3.3 Оптимизация методом ISTA.

Теперь перейдем к рассмотрению общего случая с произвольной матрицей  $\mathbf{G}$ . Одним из способов оптимизации функционала качества в этом случае является ММ-алгоритм (majorization-minimization), позволяющий для выпуклых но негладких функций строить итерационные схемы оптимизации, использующие более простую целевую функцию. Получающийся в итоге алгоритм получил название ISTA (iterative shrinkage thresholding algorithm) (см. например [64]).

При построении ММ-алгоритма для целевой функции  $f(\mathbf{s})$  и некоторой точки  $\hat{\mathbf{s}}^{(k)}$  строится мажоризирующая функция  $g^{(k)}$ , совпадающая с функцией  $f$  в точке  $\hat{\mathbf{s}}^{(k)}$ :

$$\begin{aligned} g^{(k)}(\mathbf{s}) &\geq f(\mathbf{s}) \\ g^{(k)}(\hat{\mathbf{s}}^{(k)}) &= f(\hat{\mathbf{s}}^{(k)}) \end{aligned}$$

При этом следующее приближение к решению  $\hat{\mathbf{s}}^{(k+1)}$  находится как минимум мажоризирующей функции, построенной для точки  $\hat{\mathbf{s}}^{(k)}$ . Такая итерационная процедура обладает монотонной сходимостью к точке минимума функционала  $f$ , при условии, что этот функционал является выпуклым [66]. Конкретный выбор последовательности функций  $g^{(k)}$  при этом делается исходя из удобства минимизации.

Чтобы получить функцию, обладающую искомыми свойствами, достаточно добавить к исходной функции  $f$  неотрицательное слагаемое, которое обращается в ноль в точке  $\hat{\mathbf{s}}^{(k)}$ . В качестве такого слагаемого возьмем выражение  $\frac{1}{2}(\mathbf{s} - \hat{\mathbf{s}}^{(k)})^T(\beta\mathbf{I} - \mathbf{G}^T\mathbf{G})(\mathbf{s} - \hat{\mathbf{s}}^{(k)})$ , выбрав для  $\beta$  значение, превосходящее

максимальное собственное число матрицы  $\mathbf{G}^T \mathbf{G}$ , чтобы гарантировать положительность. Раскрыв скобки и приведя подобные слагаемые и выделив полный квадрат, получим для  $g^{(k)}$  выражение

$$g_k(\mathbf{s}) = \frac{1}{2} \left\| \hat{\mathbf{s}}^{(k)} + \frac{1}{\beta} \mathbf{G}^T (\mathbf{x}_{noisy} - \mathbf{G} \hat{\mathbf{s}}^{(k)}) - \mathbf{s} \right\|_{L_2}^2 + \frac{\alpha}{\beta} \|\mathbf{s}\|_{L_1} + C, \quad (2.83)$$

где в  $C$  сгруппированы слагаемые, которые не зависят от  $\mathbf{s}$  и следовательно не влияют на минимизацию. Отметим, что в полученной функции  $g^{(k)}$  аргумент  $\mathbf{s}$  присутствует под знаком нормы без умножения на матрицу  $\mathbf{G}$ , то есть для этой функции задача поиска минимума сводится к частному случаю единичной матрицы прямой модели, рассмотренному нами выше, и ее минимум может быть найден в явном виде как результат применения оператора  $h(x, c)$  к компонентам вектора  $\hat{\mathbf{s}}^{(k)} + \frac{1}{\beta} \mathbf{G}^T (\mathbf{x}_{noisy} - \mathbf{G} \hat{\mathbf{s}}^{(k)})$ :

$$\hat{\mathbf{s}}^{(k+1)} = h \left( \hat{\mathbf{s}}^{(k)} + \frac{1}{\beta} \mathbf{G}^T (\mathbf{x}_{noisy} - \mathbf{G} \hat{\mathbf{s}}^{(k)}), \frac{\alpha}{\beta} \right) \quad (2.84)$$

Согласно свойствам ММ-алгоритма полученная последовательность оценок  $\hat{\mathbf{s}}^{(k)}$  сходится к искомому минимуму функционала  $f$ .

Отметим, что при посторении мажоранты для функционала качества конкретный вид штрафующей нормы никак не учитывался; тип используемой нормы оказывается только при подсчете оператора  $h$  после того как, используя метод ММ, мы распутывали переменные под знаком нормы, сводя задачу к случаю единичной матрицы.

На самом деле полученный алгоритм на каждой итерации можно условно разбить на два базовых шага:

1. градиентный спуск для  $L_2$ -нормы ошибки в объяснении измерений; величина  $\beta$  при этом регулирует скорость градиентного спуска и ограничена снизу максимальным собственным значением оператора  $\mathbf{G}^T \mathbf{G}$ . Этот шаг не зависит от конкретного вида штрафующей нормы.
2. удаление шума для вновь полученной градиентным спуском точки с использованием априорного знания о сигнале, кодируемого нормой  $L_1$ .

Это наблюдение мотивирует нас обобщить оператор soft thresholding на более широкий класс функций-регуляризаторов. В общем случае при оптимизации суммы двух функций

$$f_1(\mathbf{x}) + \alpha f_2(\mathbf{x}) \rightarrow \min_{\mathbf{x}}, \quad (2.85)$$

где обе функции  $f_1, f_2$  являются выпуклыми, а функция  $f_1$  дополнительно является гладкой, вводится понятие *оператора приближения* (proximity operator или proximal operator):

$$\text{prox}_{cf_2}(\mathbf{y}) = \arg \min_{\mathbf{x}} \left( \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_{L_2}^2 + cf_2(\mathbf{x}) \right) \quad (2.86)$$

Как и для 2.78 выражение 2.86 можно трактовать как оператор удаления шума из вектора  $\mathbf{x}$  с учетом априорного знания о структуре сигнала, которая кодируется при помощи регуляризатора  $f_2$ . Отметим также, что при выборе  $f_2 = \|\cdot\|_{L_1}$  мы получим в точности функционал 2.78, который приводит к формуле оператора soft thresholding.

Используя оператор приближения, для функционалов вида

$$f(\hat{\mathbf{s}}) = \frac{1}{2} \|\mathbf{x}_{noisy} - \mathbf{G}\hat{\mathbf{s}}\|_{L_2}^2 + \alpha f_{norm}(\hat{\mathbf{s}}) \rightarrow \min_{\hat{\mathbf{s}}}. \quad (2.87)$$

можем выписать в общем виде выражение для вычисления следующего шага ММ-алгоритма с произвольным выпуклым регуляризатором  $f_{norm}$ :

$$\hat{\mathbf{s}}^{(k+1)} = \text{prox}_{\frac{\alpha}{\beta} f_{norm}} \left( \hat{\mathbf{s}}^{(k)} + \frac{1}{\beta} \mathbf{G}^T (\mathbf{x}_{noisy} - \mathbf{G}\hat{\mathbf{s}}^{(k)}) \right) \quad (2.88)$$

Получим теперь в явном виде оператор приближения для регуляризатора, использующего смешанную норму [65] с учетом того, что теперь вместо векторов  $\mathbf{x}, \mathbf{y}$  необходимо использовать матрицы  $\mathbf{S}, \mathbf{X}$ . Выпишем сначала определение:

$$\begin{aligned} \text{prox}_{c\|\cdot\|_{L_{21}}}(\tilde{\mathbf{S}}) &= \arg \min_{\mathbf{S}} \left( \frac{1}{2} \|\mathbf{S} - \tilde{\mathbf{S}}\|_{fro}^2 + c \|\mathbf{S}\|_{L_{21}} \right) = \\ &= \arg \min_{\mathbf{S}} \left( \frac{1}{2} \|\mathbf{S} - \tilde{\mathbf{S}}\|_{fro}^2 + c \sum_i \|\mathbf{S}_{i,:}\|_{L_2} \right), \end{aligned} \quad (2.89)$$

где  $\mathbf{S}_{i,:}$  обозначает  $i$ -ую строку матрицы  $\mathbf{S}$  или, что то же самое, временной ряд  $i$ -го источника.

Минимум выражения в скобках достигается в тех точках, в которых субдифференциал выражения содержит нулевой вектор.

В случае  $\mathbf{S}_i = \mathbf{0}$  блоки субдифференциала нормы  $L_{21}$ , соответствующие строкам матрицы  $\mathbf{S}$ , будут состоять из всех векторов  $\mathbf{d}$ , которые удовлетворяют соотношению

$$\|\mathbf{x}\|_{L_2} \geq \mathbf{d}^T \mathbf{x} \quad (2.90)$$

для любого вектора  $\mathbf{x}$ . Максимум выражения  $\mathbf{d}^T \mathbf{x}$  достигается, когда  $\mathbf{x}$  коллинеарен  $\mathbf{d}$ . Тогда условие 2.90 перепишется как

$$\|\mathbf{d}\|_{L_2} \leq 1, \quad (2.91)$$

т.е. соответствующие блоки субдифференциала состоят из векторов,  $L_2$ -норма которых не превосходит 1.

Теперь условие принадлежности нуля соответствующему блоку субдифференциала может быть записано как

$$\tilde{\mathbf{S}}_i + c\mathbf{d} = \mathbf{0}, \quad (2.92)$$

где  $\mathbf{d}$  — некоторый вектор, соответствующий блоку субдифференциала. Такой вектор найдётся при условии, что

$$\left\| \tilde{\mathbf{S}}_i \right\|_{L_2} \leq c, \quad (2.93)$$

т.е. решение  $\mathbf{S}_i = \mathbf{0}$  минимизирует функционал 2.89 по соответствующим компонентам тогда и только тогда, когда  $\left\| \tilde{\mathbf{S}}_i \right\|_{L_2} \leq c$ .

Если  $\|\mathbf{S}_{i,:}\| \neq 0$ , выражение в скобках является дифференцируемым по соответствующим компонентам, поэтому для поиска минимума достаточно приравнять к нулю производную по компонентам вектора  $\mathbf{S}_i$ :

$$(s_{ij} - \tilde{s}_{ij}) + c \frac{s_{ij}}{\|\mathbf{S}_{i,:}\|} = 0 \quad (2.94)$$

$$\tilde{s}_{ij} = s_{ij} \left( 1 + \frac{c}{\|\mathbf{S}_{i,:}\|} \right) \quad (2.95)$$

Зафиксировав индекс  $k$ , для каждого индекса  $l$  возведем последнее равенство в квадрат и сложим результаты, чтобы выразить норму  $\|\mathbf{S}_{i,:}\|$  через норму  $\left\| \tilde{\mathbf{S}}_{i,:} \right\|$ :

$$\left\| \tilde{\mathbf{S}}_{i,:} \right\|^2 = \|\mathbf{S}_{i,:}\|^2 \left( 1 + \frac{c}{\|\mathbf{S}_{i,:}\|} \right)^2 = (\|\mathbf{S}_{i,:}\| + c)^2 \quad (2.96)$$

Откуда

$$\|\mathbf{S}_{i,:}\| = \|\tilde{\mathbf{S}}_{i,:}\| - c \quad (2.97)$$

Подставляя 2.97 в 2.95, а также учитывая, что так как  $\|\mathbf{S}_{i,:}\| > 0$ , всегда выполнено  $\|\tilde{\mathbf{S}}_{i,:}\| > c$ , получим

$$s_{ij} = \tilde{s}_{ij} \left( 1 - \frac{c}{\|\tilde{\mathbf{S}}_{i,:}\|} \right)^+, \text{ где } (\cdot)^+ \stackrel{\text{def}}{=} \max(\cdot, 0) \quad (2.98)$$

Окончательно будем иметь

$$\text{prox}_{c\|\cdot\|_{L_{21}}}(\tilde{\mathbf{S}}) = \text{diag} \left\{ \left( 1 - \frac{c}{\|\tilde{\mathbf{S}}_{i,:}\|} \right)^+ \right\}_i \tilde{\mathbf{S}} \quad (2.99)$$

Применение оператора 2.99 к правилу обновления целевой переменной 2.88 позволяет сформулировать алгоритм ISTA, обобщенный на случай регуляризации со смешанной нормой.

### 2.3.4 Контроль сходимости схемы ISTA.

Остановка итерационной схемы ISTA возможна в соответствии с эвристикой, часто используемой на практике при оптимизации того или иного функционала, которая заключается в контроле нормы разности решения на текущем и предыдущем шаге. Если эта норма опускается ниже некоторого заранее заданного порога, алгоритм останавливается.

В некоторых случаях, однако, оказывается возможным сформулировать более обоснованный критерий останова, основанный на теории двойственности для задач оптимизации и использующий понятие *разрыва двойственности* как средство контроля оптимальности найденного решения.

Как раз такой критерий останова использовали авторы статьи [65]. Разберем подробнее суть подхода к контролю оптимальности найденного численно решения на основе теории двойственности.

Выпишем задачу оптимизации с нормой  $L_{21}$ :

$$\frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{GS}\|_{fro}^2 + \alpha \|\mathbf{S}\|_{L_{21}} \rightarrow \min_{\mathbf{S}} \quad (2.100)$$

Перепишем эту задачу в более общем виде:

$$f_1(\mathbf{S}) - f_2(\mathbf{GS}) \rightarrow \min_{\mathbf{S}} \quad (2.101)$$

Здесь функция  $f_1(\mathbf{S}) = \alpha \|\mathbf{S}\|_{L_{21}}$  является выпуклой, а функция  $f_2(\mathbf{V}) = \frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{V}\|_{fro}^2$  — вогнутой. Кроме того,  $f_1$  и  $f_2 \circ \mathbf{G}$  определены всюду на  $\mathbb{R}^{n \times m}$ . При этом относительно дифференцируемости функций  $f_1, f_2$  мы дополнительно ничего не предполагаем.

Можем переписать исходную задачу оптимизации как задачу оптимизации с ограничениями-равенствами:

$$\begin{aligned} f_1(\mathbf{S}) - f_2(\mathbf{V}) &\rightarrow \min_{\mathbf{S}, \mathbf{V}} \\ \text{s.t.: } \mathbf{V} &= \mathbf{GS} \end{aligned} \quad (2.102)$$

Для этой оптимационной задачи 2.102 можем выписать функцию Лагранжа:

$$L(\mathbf{S}, \mathbf{V}, \boldsymbol{\Lambda}) = f_1(\mathbf{S}) - f_2(\mathbf{V}) + \text{Tr}(\boldsymbol{\Lambda}^T(\mathbf{V} - \mathbf{GS})) \quad (2.103)$$

Для функции Лагранжа в теории выпуклой оптимизации двойственная функция определяется как функция от множителей Лагранжа, равная инфикуму от функции Лагранжа в каждой точке:

$$g(\boldsymbol{\Lambda}) = \inf_{\mathbf{S}, \mathbf{V}} L(\mathbf{S}, \mathbf{V}, \boldsymbol{\Lambda}) \quad (2.104)$$

Распишем, чему равна двойственная функция для Лагранжиана 2.103:

$$\begin{aligned} g(\boldsymbol{\Lambda}) &= \inf_{\mathbf{S}, \mathbf{V}} (f_1(\mathbf{S}) - f_2(\mathbf{V}) - \text{Tr}(\boldsymbol{\Lambda}^T(\mathbf{GS} - \mathbf{V}))) = \\ &= \inf_{\mathbf{S}, \mathbf{V}} ((\text{Tr}(\boldsymbol{\Lambda}^T \mathbf{V}) - f_2(\mathbf{V})) - (\text{Tr}(\boldsymbol{\Lambda}^T \mathbf{GS}) - f_1(\mathbf{S}))) = \\ &= \inf_{\mathbf{V}} (\text{Tr}(\boldsymbol{\Lambda}^T \mathbf{V}) - f_2(\mathbf{V})) - \sup_{\mathbf{S}} ((\text{Tr}(\mathbf{G}^T \boldsymbol{\Lambda})^T \mathbf{S}) - f_1(\mathbf{S})) \end{aligned} \quad (2.105)$$

Как видим, двойственная функция распалась на две сходные по форме компоненты, каждая из которых требует поиска оптимума по своей переменной.

В теории выпуклой оптимизации рассматривают выпуклую и вогнутую сопряженные функции, которые определяются для некоторой функции  $f(\mathbf{S})$  как

$$f^*(\Lambda) \stackrel{def}{=} \sup_{\mathbf{S}} (\text{Tr}(\Lambda^T \mathbf{S}) - f(\mathbf{S})) \text{ — выпуклая сопряженная} \quad (2.106)$$

$$f_*(\Lambda) \stackrel{def}{=} \inf_{\mathbf{S}} (\text{Tr}(\Lambda^T \mathbf{S}) - f(\mathbf{S})) \text{ — вогнутая сопряженная} \quad (2.107)$$

Смысл названий этих функций в том, что первая из двух является выпуклой функцией как поточечный супремум от семейства линейных функций, а вторая как поточечный инфимум является вогнутой [67].

С использованием этих определений можем переписать двойственную функцию как

$$g(\Lambda) = (f_2)_*(\Lambda) - (f_1)^*(\mathbf{G}^T \Lambda) \quad (2.108)$$

Распишем, чему равно выражение 2.108 для функционала вида 2.87. В этом случае  $f_1(\mathbf{S}) = \alpha f_{norm}(\mathbf{S})$ , а  $f_2(\mathbf{S}) = -\frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{S}\|_{fro}^2$ . Действовать будем по аналогии с выводом формулы для разрыва двойственности из статьи [68].

Воспользуемся правилами вычисления сопряжений для произведения функции на константу [67]:  $(cf)^*(\Lambda) = cf^*(\Lambda/c)$ . Получим

$$(f_1)^*(\mathbf{G}^T \Lambda) = \alpha (f_{norm})^*(\mathbf{G}^T \Lambda / \alpha) \quad (2.109)$$

Выпишем теперь в явном виде, чему равно сопряжение от  $f_2$ :

$$(f_2)^*(\Lambda) = \inf_{\mathbf{S}} (\text{Tr}(\Lambda^T \mathbf{S}) + \frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{S}\|_{fro}^2) = -\sup_{\mathbf{S}} (-\text{Tr}(\Lambda^T \mathbf{S}) - \frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{S}\|_{fro}^2)$$

Проведём замену переменной  $\mathbf{Z} = \mathbf{X}_{noisy} - \mathbf{S}$ :

$$(f_2)^*(\Lambda) = \text{Tr}(\Lambda^T \mathbf{X}_{noisy}) - \sup_{\mathbf{S}} (\text{Tr}(\Lambda^T \mathbf{Z}) - \frac{1}{2} \|\mathbf{Z}\|_{fro}^2) \quad (2.110)$$

Супремум в правой части выражения 2.110 представляет собой выпуклое сопряжение от квадрата нормы Фробениуса. Сопряжение половины квадрата нормы равно половине квадрата её двойственной нормы, [65; 67]. Нормы  $L_p$  и  $L_{p'}$  называются двойственными друг к другу, если для них выполнено  $\frac{1}{p} + \frac{1}{p'} = 1$ . Норма Фробениуса в соответствии с этим определением является двойственной к самой себе, т.к. по форме совпадает с нормой  $L_2$  для векторов, только в случае матриц берётся их векторизация. Получаем

$$(f_2)^*(\Lambda) = \text{Tr}(\Lambda^T \mathbf{X}_{noisy}) - \frac{1}{2} \|\Lambda\|_{fro}^2 \quad (2.111)$$

Тогда двойственная функция запишется как

$$g(\Lambda) = \text{Tr}(\Lambda^T \mathbf{X}_{noisy}) - \frac{1}{2} \|\Lambda\|_{fro}^2 - \alpha(f_{norm})^*(\mathbf{G}^T \Lambda / \alpha) \quad (2.112)$$

Для нормы без степени сопряженная функция равна индикаторной функции единичного шара для двойственной нормы. В случае  $f_{norm} = \|\cdot\|_{L_{21}}$ , сопряженная функция будет выглядеть как ([65]):

$$(\|\cdot\|_{L_{21}})^*(\Lambda) = I_{\{\Lambda: \|\Lambda\|_{L_{2,\infty}} \leq 1\}}(\Lambda) = \begin{cases} 0, & \|\Lambda\|_{L_{2,\infty}} \leq 1 \\ \infty, & \|\Lambda\|_{L_{2,\infty}} > 1 \end{cases} \quad (2.113)$$

Для выражения 2.112 наличие индикатора означает, что для определённых значений двойственной переменной двойственная функция будет принимать значение  $-\infty$ . На практике, чтобы разрыв двойственности принимал значения в более разумном диапазоне, можем спроектировать целевую переменную двойственной функции на множество допустимых значений — единичный шар при норме  $L_{2,\infty}$ , [65; 68]. Эта эвристика основана на том соображении, что Рассмотренная нами задача 2.102 является выпуклой и содержит только ограничения-равенства, а значит для нее выполнены условия Слейтера (при условии, что области определения функций  $f_1$  и  $f_2$  имеют непустое пересечение), и следовательно для нее имеет место сильная двойственность, т.е.

$$\inf_{\mathbf{S}, \mathbf{V}} (f_1(\mathbf{S}) - f_2(\mathbf{V})) = \sup_{\Lambda} (g(\Lambda)) = \sup_{\Lambda} ((f_2)_*(\Lambda) - (f_1)^*(\mathbf{G}^T \Lambda)) = \tilde{p}, \quad (2.114)$$

где  $\tilde{p}$  — значение исходного функционала в точке минимума. Поэтому для сходящейся последовательности значений прямой переменной значения двойственной переменной также сходятся, а значит и проекции двойственной переменной на выпуклое множество также должны сходиться.

С использованием такой эвристики получим для двойственной функции выражение

$$g(\Lambda) = \text{Tr}(\bar{\Lambda}^T \mathbf{X}_{noisy}) - \frac{1}{2} \|\bar{\Lambda}\|_{fro}^2, \quad (2.115)$$

$$\bar{\Lambda}^{(k)} = \Lambda^{(k)} / \max \left( \left\| \mathbf{G}^T \Lambda^{(k)} / \alpha \right\|_{L_{2,\infty}}, 1 \right) \quad (2.116)$$

**Получим теперь выражение для двойственной переменной через прямую.** Обозначим  $\tilde{\Lambda}$  оптимальное значение двойственной переменной. Оптимальное значение исходного функционала достигается в точке минимума

функции Лагранжа с зафиксированной  $\Lambda = \tilde{\Lambda}$ , т.е.  $\tilde{p} = \inf_{\mathbf{S}, \mathbf{V}} (L(\mathbf{S}, \mathbf{V}, \tilde{\Lambda})) = g(\tilde{\Lambda})$ . Тогда для матриц  $\tilde{\mathbf{S}}, \tilde{\mathbf{V}}$ , минимизирующих исходный функционал, с учётом  $\tilde{\mathbf{V}} = \mathbf{G}\tilde{\mathbf{S}}$  должно быть выполнено

$$f_1(\tilde{\mathbf{S}}) + (f_1)^*(\mathbf{G}^T \tilde{\Lambda}) = \text{Tr}((\mathbf{G}^T \tilde{\Lambda})^T \tilde{\mathbf{S}}) \quad (2.117)$$

$$f_2(\mathbf{G}\tilde{\mathbf{S}}) + (f_2)_*(\tilde{\Lambda}) = \text{Tr}(\tilde{\Lambda}^T \mathbf{G}\tilde{\mathbf{S}}) \quad (2.118)$$

Эти соотношения представляют условие стационарности функции лагранжа в точке минимума. В нашем случае второе из этих соотношений запишется как

$$-\frac{1}{2} \left\| \mathbf{X}_{noisy} - \mathbf{G}\tilde{\mathbf{S}} \right\|_{fro}^2 + \text{Tr}(\tilde{\Lambda}^T \mathbf{X}_{noisy}) - \frac{1}{2} \left\| \tilde{\Lambda} \right\|_{fro}^2 = \text{Tr}(\tilde{\Lambda}^T \mathbf{G}\tilde{\mathbf{S}})$$

Или эквивалентно

$$\begin{aligned} -\frac{1}{2} \left( \left\| \mathbf{X}_{noisy} - \mathbf{G}\tilde{\mathbf{S}} \right\|_{fro}^2 - 2 \text{Tr}(\tilde{\Lambda}^T (\mathbf{X}_{noisy} - \mathbf{G}\tilde{\mathbf{S}})) + \left\| \tilde{\Lambda} \right\|_{fro}^2 \right) &= \\ &= -\frac{1}{2} \left( \left\| \mathbf{X}_{noisy} - \mathbf{G}\tilde{\mathbf{S}} - \tilde{\Lambda} \right\|_{fro}^2 \right) = 0 \end{aligned}$$

А значит справедливо

$$\tilde{\Lambda} = \mathbf{X}_{noisy} - \mathbf{G}\tilde{\mathbf{S}}, \quad (2.119)$$

что задает связь между прямой и двойственной переменными. Отметим, что двойственную переменную можно интерпретировать как ошибку в объяснении измерений при помощи найденного решения  $\hat{\mathbf{S}}$ .

Формулы 2.116, 2.119 задают конкретный способ расчёта значения двойственного функционала на каждой итерации по значению прямой переменной. Разность значения оптимизируемого функционала на текущей итерации и соответствующего ему значения двойственного функционала задают разрыв двойственности, который в силу свойства сильной двойственности, выполненного для нашей задачи, должен сходиться к нулю и следовательно может быть использован в критерии останова алгоритма.

### 2.3.5 Ускорение схемы ISTA

На практике однако оказывается, что метод ISTA является не самым эффективным способом оптимизации функционала качества для спарсных норм. Было показано, что его скорость сходимости имеет порядок  $O(1/K)$ , где  $K$  — количество итераций. В 2009 авторы статьи [69] предложили ускоренную итерационную схему на основе алгоритма ISTA, которая получила название Fast ISTA или FISTA, и имеет порядок  $O(1/K^2)$ . Позднее в статьях [70] и [71] для оптимизации функционала со смешанной нормой  $L_{21}$  было предложено использовать метод блочно-координатного градиентного спуска (block-coordinate descent, BCD), которая на практике показывает еще более высокие скорости сходимости, чем алгоритм FISTA.

Суть подхода, основанного на блочно-координатном спуске заключается в следующем. Полученный в разделе 2.3.3 алгоритм ISTA на каждой итерации совершают шаг градиентного спуска с постоянным параметром  $1/\beta$ , одинаковым для всех источников. Иными словами, все компоненты матрицы оценки источников на  $k$ -ом шаге  $\hat{\mathbf{S}}_k$  обновляются одновременно. Метод блочно-координатного спуска состоит в том, чтобы обновлять элементы матрицы  $\hat{\mathbf{S}}_k$  группами. Каждая группа при этом состоит из элементов одной строки матрицы  $\hat{\mathbf{S}}_k$ . Так как структура оператора приближения для нормы  $L_{21}$  позволяет рассчитывать его для каждой такой “подитерации” отдельно, обновление переменных из каждого блока происходит как и для схемы ISTA в явном виде:

$$\hat{\mathbf{S}}_i^{(k+1)} = \text{prox}_{\alpha\mu(i)\|\cdot\|_{L_2}} \left( \hat{\mathbf{S}}_i^{(k)} + \mu(i) \mathbf{G}_i^T (\mathbf{X}_{noisy} - \mathbf{G} \hat{\mathbf{S}}^{(k)}) \right) \quad (2.120)$$

Здесь мы используем оператор приближения для нормы  $L_2$ , а не для нормы  $L_{21}$ , так как за раз мы обновляем только одну строку матрицы  $\hat{\mathbf{S}}$ , и для одной строки штрафующая норма будет как раз  $L_2$ . В качестве шага градиентного спуска  $\mu(i)$  используем  $1/\|\mathbf{G}_i^T \mathbf{G}_i\|_{L_2}$ , где  $\|\mathbf{G}_i^T \mathbf{G}_i\|_{L_2}$  — константа Липшица для градиента по  $i$ -му блоку матрицы  $\hat{\mathbf{S}}$ .

Оператор приближения  $\text{prox}_{\alpha\mu(i)\|\cdot\|_{L_2}}$  запишется аналогично формуле 2.99:

$$\text{prox}_{\alpha\mu(i)\|\cdot\|_{L_2}}(\mathbf{s}) = \left( 1 - \frac{\mu(i)\alpha}{\|\mathbf{s}\|_{L_2}} \right)^+ \mathbf{s} \quad (2.121)$$

На практике при вычислении оператора приближения 2.121 следует учитывать, что какие-то блоки матрицы  $\hat{\mathbf{S}}$  будут равны нулю. В этом случае формально оператор положительного среза  $(\cdot)^+$  будет давать ожидаемое нулевое значение, однако для практических вычислений более удобно пользоваться изменённым видом формулы 2.121:

$$\text{prox}_{\alpha \mu(i) \|\cdot\|_{L_2}}(\mathbf{s}) = \left(1 - \frac{\mu(i)\alpha}{\max(\|\mathbf{s}\|_{L_2}, \mu(i)\alpha)}\right) \mathbf{s} \quad (2.122)$$

Такая формулировка лучше для программирования алгоритма, так как позволяет избежать деления на ноль.

**Для дополнительного ускорения расчётов воспользуемся методом активного множества.** Метод блочно-координатного спуска как правило применяют поочерёдно для всех строк матрицы  $\hat{\mathbf{S}}$  в рамках одной “глобальной” итерации. Но для обратной задачи МЭЭГ количество строк  $\hat{\mathbf{S}}$  равно количеству источников на коре, поэтому такое обновление может занять продолжительное время. Вместе с тем из-за свойств спарсной нормы большая часть строк матрицы  $\hat{\mathbf{S}}$  должна оставаться нулевой. Это соображение можно использовать, чтобы обновлять строки матрицы  $\hat{\mathbf{S}}$  исходя из данных на текущей итерации.

Одним из способов это сделать является метод активного множества. Этот метод основан на выделении подгруппы “плохих” источников и применении блочно-координатного спуска только на этой подгруппе. Плохими считаются источники, топографии которых наиболее сильно коррелированы с остаточной ошибкой на текущем шаге. Если после оптимизации на этой подгруппе решение всё ещё не является оптимальным для полной задачи по критерию разрыва двойственности, активное множество обновляется и процедура повторяется но уже с частью рассчитанных значениями для строк матрицы  $\hat{\mathbf{S}}$ .

Обновление активного множества включает две процедуры:

1. удаление из множества источников, которые оказались нулевыми на текущем шаге оптимизации
2. добавление к множеству новых источников, для которых топография коррелирует с остаточной ошибкой.

Процедура выбора плохих источников неслучайна. Рассмотрим первое условие стационарности лагранжиана в точке минимума 2.117. Для  $f_1(\mathbf{A}) = \alpha \|\mathbf{A}\|_{L_1}$  оно запишется как

$$\alpha \left\| \hat{\mathbf{S}} \right\|_{L_{21}} + \alpha I_{\{\mathbf{S}: \|\mathbf{S}\|_{L_{2,\infty}} \leq 1\}}(\mathbf{G}^T \boldsymbol{\Lambda} / \alpha) = \text{Tr}((\mathbf{G}^T \boldsymbol{\Lambda})^T \hat{\mathbf{S}}). \quad (2.123)$$

Здесь  $\alpha I_{\{\mathbf{S}: \|\mathbf{S}\|_{L_{2,\infty}} \leq 1\}}(\mathbf{S})$  — индикатор единичного шара в пространстве с нормой  $L_{2,\infty}$ .

Наличие индикатора эквивалентно требованию

$$\max_i (\mathbf{G}_i^T \boldsymbol{\Lambda}) \leq \alpha,$$

или, что то же самое, для всех  $i$  должно выполняться

$$\mathbf{G}_i^T \boldsymbol{\Lambda} \leq \alpha \quad (2.124)$$

Если это условие выполнено, индикатор равен нулю, и выражение 2.123 можно расписать как

$$\alpha \left\| \hat{\mathbf{S}} \right\|_{L_{21}} = \text{Tr}((\mathbf{G}^T \boldsymbol{\Lambda})^T \hat{\mathbf{S}}) \iff \alpha \sum_i \left\| \hat{\mathbf{S}}_i \right\|_{L_2} = \sum_i (\mathbf{G}_i^T \boldsymbol{\Lambda})^T \hat{\mathbf{S}}_i \quad (2.125)$$

Правая часть в выражении 2.125 представляет собой суммы скалярных произведений. Пользуясь 2.124 и неравенством Коши-Буняковского-Шварца, получим

$$\alpha \sum_i \left\| \hat{\mathbf{S}}_i \right\|_{L_2} \leq \sum_i \left\| \mathbf{G}_i^T \boldsymbol{\Lambda} \right\|_{L_2} \left\| \hat{\mathbf{S}}_i \right\|_{L_2} \leq \alpha \sum_i \left\| \hat{\mathbf{S}}_i \right\|_{L_2} \quad (2.126)$$

Последнее выполнено тогда и только тогда, когда для каждого индекса  $i$  либо  $\left\| \hat{\mathbf{S}}_i \right\|_{L_2} = 0$  и  $\left\| \mathbf{G}_i^T \boldsymbol{\Lambda} \right\|_{L_2} \leq \alpha$ , либо  $\left\| \hat{\mathbf{S}}_i \right\|_{L_2} \neq 0$  и  $\left\| \mathbf{G}_i^T \boldsymbol{\Lambda} \right\|_{L_2} = \alpha$ . Т.е. в точке оптимума для всех ненулевых источников проекция двойственной переменной на соответствующие топографии не длиннее  $\alpha$ , а для ненулевых длина такой проекции в точности равна  $\alpha$ .

В обратную сторону утверждение также верно. Если для строк некоторой матрицы  $\mathbf{S}$  выполнено, что либо  $\left\| \mathbf{S}_i \right\|_{L_2} = 0$  и  $\left\| \mathbf{G}_i^T (\mathbf{X}_{noisy} - \mathbf{GS}) \right\|_{L_2} \leq \alpha$ , либо  $\left\| \mathbf{S}_i \right\|_{L_2} \neq 0$  и  $\left\| \mathbf{G}_i^T (\mathbf{X}_{noisy} - \mathbf{GS}) \right\|_{L_2} = \alpha$ , то эта точка является точкой минимума нашего функционала.

Это наблюдение позволяет выделить из всех источников те, которые с большей вероятностью не равны нулю, и обновлять в блочно-координатном спуске только их. Отбор индексов источников  $i$ , которые необходимо оптимизировать в первую очередь, при этом происходит по критерию  $\left\| \mathbf{G}_i^T (\mathbf{M} - \mathbf{GS}) \right\|_{L_2} > \alpha$ .

Отметим, что этот критерий весьма близок по духу к алгоритму RAP-MUSIC, 1.3.1. В алгоритме RAP-MUSIC мы искали источник, топография

которого сильнее других коррелируют с подпространством сигнала, из которого удален вклад источников, найденных на предыдущих итерациях. Здесь мы для обновления активного множества делаем то же самое за исключением выделения подпространства сигнала — корреляция топографии считается напрямую с остаточной нормой ошибки.

Сформулируем алгоритм оптимизации блочно-координатным спуском с использованием активного множества. На первом шаге необходимо инициализировать массив  $\hat{\mathbf{S}}^{(1)}$  нулями. Далее найдём набор  $A$ , состоящий из  $l$  источников, которые наиболее сильно нарушают условие  $\|\mathbf{G}_i^T(\mathbf{M} - \mathbf{GS})\|_{L_2} > \alpha$ . Для матрицы  $\mathbf{G}_A$  — ограничения  $\mathbf{G}$  на  $A$  — и  $\hat{\mathbf{S}}_A$  — ограничения  $\hat{\mathbf{S}}$  на  $A$  найдём оптимальное решение для функционала 2.100 методом блочно-координатного спуска. В качестве критерия оптимальности используем разрыв двойственности для ограниченной задачи. Строкам  $\hat{\mathbf{S}}$ , соответствующим элементам множества  $A$ , присвоим найденные значения  $\mathbf{S}_A$ . Если полученная матрица  $\hat{\mathbf{S}}$  является оптимальной по критерию разрыва двойственности, алгоритм останавливается. Если нет, дополняем множество  $A$  новыми источниками, наиболее сильно нарушающими условие  $\|\mathbf{G}_i^T(\mathbf{M} - \mathbf{GS})\|_{L_2} > \alpha$  и повторяем алгоритм для нового множества  $A$ .

### 2.3.6 Метод смешанной нормы с итеративным перевзвешиванием.

Оказывается, что на основе метода 2.3.3 можно сформулировать алгоритм, который аппроксимирует штрафующую норму  $L_0$  точнее, тем самым усиливая спарсность получаемых решений и позволяя избавиться от некоторых негативных свойств алгоритма с нормой  $L_1$ , в число которых входит смешённость оценки амплитуды источников.

Авторы статьи [68] предложили использовать в качестве штрафующей функции смешанную псевдонорму  $L_{2,0.5}$ . Эта псевдонорма обладает более крутым градиентом вблизи нуля и тем самым лучше моделирует форму индикаторной функции, которая используется при построении алгоритма с  $L_0$ -штрафом.

Тем не менее, использование псевдонормы  $L_{2,0.5}$  имеет свою цену: эта псевдонорма не является выпуклой, что усложняет задачу численной оптимизации целевого функционала.

Тем не менее, в работах [70; 72; 73] было показано, что итоговый невыпуклый функционал можно оптимизировать, разбив оптимационную процедуру на подитерации, на каждой из которых оптимизируется выпуклый функционал с нормой  $L_{21}$ . Рассмотрим более подробно суть этого подхода.

Функционал качества с нормой  $L_{2,0.5}$  может быть записан как

$$\frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{GS}\|_{fro}^2 + \alpha \sum_i \sqrt{\|\mathbf{S}_i\|_{L2}} \rightarrow \min_{\mathbf{S}}. \quad (2.127)$$

Главную проблему для оптимизации 2.127 представляет квадратный корень в штрафующей норме. Идея метода последовательного перевзвешивания заключается в том, что мы аппроксимируем корень, используя веса для строк матрицы  $\mathbf{S}$  следующим образом:

$$\hat{\mathbf{S}}^{(k)} = \arg \min_{\mathbf{S}} \frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{GS}\|_{fro}^2 + \alpha \sum_i \frac{1}{\sqrt{\|\hat{\mathbf{S}}_i^{(k-1)}\|_{L2}}} \|\mathbf{S}_i\|_{L2} \quad (2.128)$$

Для  $k$ -го шага значения  $1/\sqrt{\|\hat{\mathbf{S}}_i^{(k-1)}\|_{L2}}$  рассчитываются на основе решения, найденного на шаге  $k-1$ , а значит являются константными. Обозначим величины  $\sqrt{\|\hat{\mathbf{S}}_i^{(k-1)}\|_{L2}}$  как  $w^{(k)}$ .

Задача 2.128 может быть эквивалентно сформулирована как

$$\begin{aligned} \hat{\mathbf{S}}^{(k)} &= \arg \min_{\mathbf{S}} \frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{GW}^{(k)}\mathbf{S}\|_{fro}^2 + \alpha \sum_i \|\mathbf{S}_i\|_{L2} = \\ &= \arg \min_{\mathbf{S}} \frac{1}{2} \|\mathbf{X}_{noisy} - \mathbf{G}^{(k)}\mathbf{S}\|_{fro}^2 + \alpha \sum_i \|\mathbf{S}_i\|_{L2} \end{aligned} \quad (2.129)$$

Здесь  $\mathbf{W}^{(k)} = \text{diag}(w^{(k)})$ , а  $\mathbf{G}^{(k)} = \mathbf{GW}^{(k)}$ .

Как видим, оптимационная задача для подитерации свелась к задаче 2.100. Решение для каждой подитерации теперь может быть получено методом, описанным в разделе 2.3.3 с техниками ускорения расчета из раздела 2.3.5. Найденные последовательные приближения  $\hat{\mathbf{S}}^{(k)}$  сходятся к локальному минимуму функционала 2.127.

Так как исходный функционал не являлся выпуклым, сходимость такого метода к глобальному минимуму функционала 2.127 не гарантирована и может зависеть от выбора начального приближения для весов  $w^{(k)}$ .

### 2.3.7 Схема метода GO-PSIICOS

Globally optimized PSIICOS — это комбинация метода очистки от протечки сигнала PSIICOS и метода спарсного решения обратной задачи IrMxNE.

Процедура заключается в следующем.

1. Для каждого временного среза рассчитываем векторизованный кросс-спектр в пространстве сенсоров. Усреднение проводится по эпохам.
2. К каждому временному срезу применяем оператор PSIICOS-проекции. Совокупность спроектированных векторизованных кросс-спектров складываем в матрицу как столбцы. Полученную матрицу будем называть времененным рядом кросс-спектра.
3. В качестве прямого оператора для метода IrMxNE берем матрицу, в которой каждый столбец является 2-топографией для двух различных точек коры. Таким образом каждой паре точек на коре соответствует столбец полученной матрицы. (для учета свободной ориентации диполя в случае МЭГ — 4 столбца)
4. Для полученной матрицы прямой модели и спроектированного временного ряда кросс-спектра применяем алгоритм IrMxNE.

Такая процедура позволяет нам избавиться от обеих проблем, вызванных протечкой сигнала: PSIICOS-проекция устраниет эффект протечки сигнала на уровне отдельных точек, а спарсный регуляризатор — на уровне сетей.

## Глава 3. Анализ свойств метода оценки синхронизации с нулевым фазовым сдвигом.

В этой главе мы исследуем свойства методов, основанных на проекции от протечки сигнала, а также сравним результаты их работы с методами, являющимися на данный момент стандартом в области исследований функциональной коннективности.

Начнем с исследования свойств базовой процедуры, описанной в разделе 2.2.

### 3.1 Исследование воспроизводимости решения для метода PSICOS.

Для проверки стабильности решений мы использовали процедуру бутстрэпа, аналогичную описанной в [74]. Мы генерировали  $B$  различных кросс-спектральных временных рядов (СТ), полученных усреднением по эпохам, индексы которых берутся как сочетания с повторениями, набранные с равными вероятностями из полного набора индексов.

Затем на каждой итерации бутстрэпа небольшое число пар  $t$ , соответствующих сетям с наибольшими значениями кросс-спектральных коэффициентов, полученными в результате применения алгоритма PSICOS, группировались в несколько кластеров в соответствии с процедурой попарной пространственной кластеризации (Pairwise Spatial Clustering, [75]).

Для каждого полученного кластера мы рассчитывали среднюю сеть. Для этого мы усредняли координаты начал и концов сетей в кластере, предварительно ориентируя их таким образом, чтобы концы были направлены в одну полуплоскость. Средние сети, полученных на каждом шаге бутстрэпа, сохранились.

Для количественной оценки пространственного разброса полученных таким образом средних сетей мы определяли расстояние между парой сетей как минимум по двум возможным ориентациям сетей от суммы евклидовых расстояний между их концами.

В соответствии с определенной таким образом функцией расстояния мы определяли индекс воспроизводимости  $\eta$  как единицу, деленную на среднее по всем  $B$  средним сетям расстояние от сети до ближайшего соседа.

### 3.2 Симуляции методом Монте-Карло.

Чтобы сравнить предложенный алгоритм с другими методами, оценивающими коннективность в пространстве источников, мы сконструировали ряд реалистичных симуляций.

Для симуляций использовали поверхность коры реального испытуемого, реконструированную по МРТ-снимкам при помощи пакета программ FreeSurfer. Для аппроксимации этой поверхности трехмерной вычислительной сеткой мы использовали 15000 узлов. Для этой поверхности мы рассчитали матрицу прямой модели высокого разрешения  $\mathbf{G}^{HR}$ , в которой на каждую точку коры приходится две топографии, соответствующие модели со свободной ориентацией диполя для МЭГ в случае сферической симметрии. Эти топографии мы рассчитывали при помощи метода главных компонент для  $M \times 3$  прямой модели для отдельно взятой точки, выкидывая компоненту с наименьшим собственным числом.

Мы симулировали 100 повторений (эпох) эксперимента, в котором за стимулом следовал всплеск индуцированной активности на частоте 10 Гц. Индуцированная активность симулировалась как пары синусоид, не привязанные по фазе к стимулу, но с неслучайной разностью фаз по отношению друг к другу. Их разность фаз  $\delta\phi$  выбиралась для каждой эпохи случайно из равномерного распределения на отрезке  $[-\pi/4, \pi/4]$ .

Шум мозга мы моделировали как  $Q = 1000$  кортиальных источников, активность которых никак не связана со стимулом. Положения на коре этих источников выбирались для каждой эпохи независимо и случайно.

Для моделирования их временных профилей активации мы фильтровали реализации гауссовского случайного процесса полосовым БИХ-фильтром пятого порядка в полосах частот, соответствующих тета (4–7 Гц), альфа (8–12 Гц), бета (15–30 Гц) и гамма (30–50, 50–70 Гц) активности. Относительные вклады этих полос частот мы взвешивали таким образом, чтобы итоговая спектральная

плотность мощности имела вид  $1/f$ , соответствующий реалистичному спектру данных МЭГ. Далее общий вклад шума мозга в итоговый сигнал взвешивался, чтобы соответствовать типичным соотношениям сигнал-шум, наблюдаемым в реальных записях.

Для отображения шумовых источников на сенсоры умножали каждый срез полученных временных профилей на соответствующие столбцы матрицы  $\mathbf{G}^{HR}$  и складывали результаты.

ОСШ в симуляционных данных мы определяли как отношение фробениусских норм матрицы данных и матрицы шума, отфильтрованных в целевой полосе (8–12 Гц).

Для вычисления матрицы кросс-спектральной плотности на сенсорах, соответствующей целевой полосе, мы проводили следующую процедуру. Мы фильтровали симулированный сигнал на каждом сенсоре в диапазоне (8–12 Гц) и вычисляли аналитический сигнал. Затем для каждого временного среза мы брали внешние произведения с сопряжением комплекснозначного вектора аналитических сигналов на сенсорах и усредняли их по эпохам.

Чтобы наши симуляции были больше приближены к реальным условиям, когда настоящие источники не всегда точно ложатся на узлы вычислительной сетки, мы использовали сетку с высоким разрешением (15000 узлов) только для симуляции данных. Для оценки коннективности по симуляционным данным мы использовали в 10 раз более разреженную сетку, насчитывающую 1503 узла.

В качестве метрик качества мы использовали кривые Receiver Operating Characteristics (ROC) и Precision-Recall (PR). Метрика Precision-Recall лучше работает в ситуации, когда число истинно положительных мало по сравнению с общим количеством вариантов. ROC-кривая, или график *чувствительность* —  $1 - \text{специфичность}$  в этом случае информативна только для очень высоких значений специфичности.

Для каждой реализации метода Монте-Карло мы считали ROC-кривую и затем усредняли кривые по ансамблю из 1000 реализаций. Сравнения проводились для двух различных значений ОСШ: 1 и 0.2.

Так как целевой характеристикой предложенного метода является одинаковое качество решений для произвольных фазовых задержек между связанными источниками, мы сравнивали поведение методов для двух различных

фазовых задержек  $\delta\phi = \pi/2 - \pi/20$  и  $\delta\phi = \pi/20$  радиан. Также мы исследовали поведение нашего метода на равномерной сетке средних значений фазового сдвига.

Для первого теста мы моделировали одну пару взаимодействующих источников. Для каждой эпохи узлы этой сети мы располагали в двух случайно выбранных точках на сетке с высоким разрешением. Для оценки изменяющейся временной активности таких сетей на каждой эпохе мы умножали временной профиль активации на временное окно, соответствующее активации в течение первой половины каждой эпохи.

На каждой итерации процедуры Монте-Карло мы симулировали данные из 100 эпох с фиксированной пространственной конфигурацией целевых источников и изменяющейся от эпохи к эпохе конфигурацией источников шума мозга. Для локализации сетей мы использовали всю протяженность эпохи чтобы смоделировать ситуацию, когда мы не знаем точно, на каком временном промежутке сеть активна.

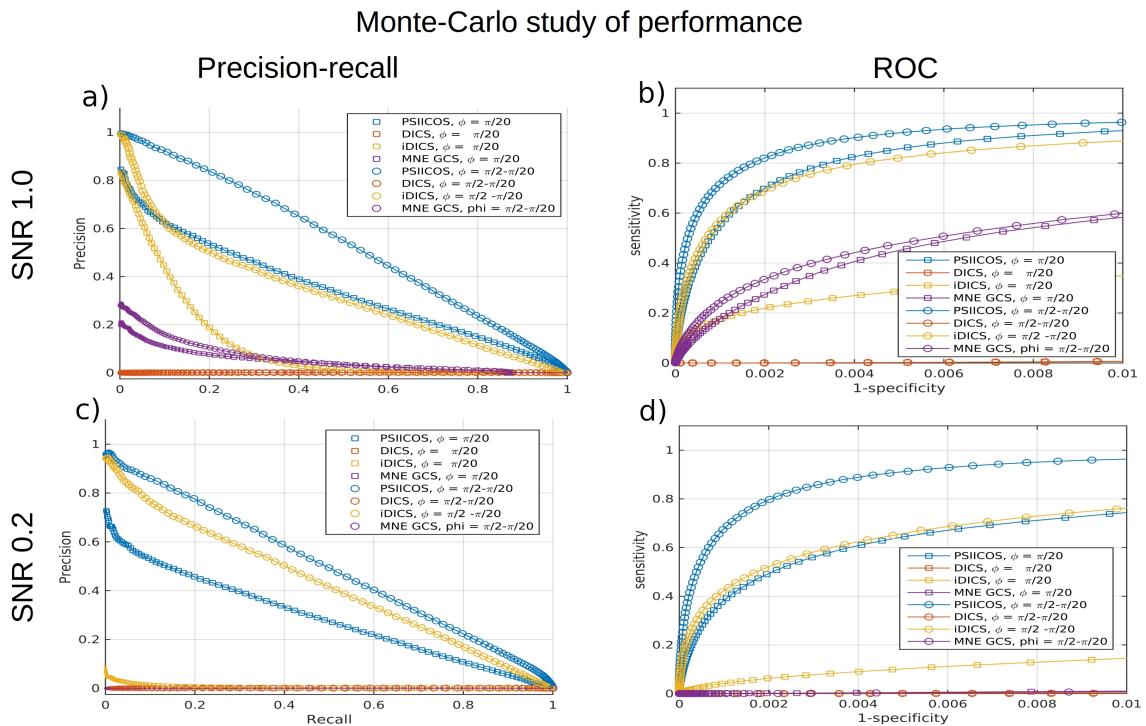


Рисунок 3.1: Сравнение Precision-Recall (панели (a), (c)) и ROC- (панели (b), (d)) кривых в задаче локализации сетей для методов PSIICOS, DICCS, iDICCS и GCS MNE для двух значений ОСШ на основе 1000 Монте-Карло итераций.

На графике 3.1 видно, что для каждого из условий симуляции качество решений алгоритма PSIIICOS стабильно превосходит другие методы и позволяет получать решения хорошего качества независимо от величины фазовой задержки.

По графику ROC-кривой для средней разницы фаз  $\delta\varphi = \pi/2 - \pi/20$  видно, что показатели iDICS сравнимы с показателями PSIIICOS. Однако для  $\delta\varphi = \pi/20$  iDICS не способен адекватно распознавать сети в силу значительного снижения ОСШ, вызванного тем, что сети с близкой к нулю фазовой задержкой дают очень маленький вклад во минимуму часть кросс-спектра.

Метод MNE GCS ведет себя лучше, чем DICS для высоких значений ОСШ. Для низкого ОСШ оба метода показывают плохие результаты. В то же время метод PSIIICOS демонстрирует адекватное качество решения для обоих значений ОСШ.

В случае низкого ОСШ большая разница в качестве решения PSIIICOS между двумя значениями фазовых задержек вероятно объясняется наличием в итерациях Монте-Карло сетей с пространственно близкими узлами. Это приводит к тому, что действительная часть членов, соответствующих истинному взаимодействию, значительно искажается проекцией. Вместе с тем, остальные методы при таких условиях для практических перестают работать в случае близких к нулю фазовых задержек.

### 3.3 Симуляции с тремя сетями, перекрывающимися во времени.

Для второго теста мы симулировали данные с тремя одновременно активными сетями, пространственная и временная структура активации которых изображены на 3.12. Главная сложность в этом тесте заключалась в том, чтобы разделить эти три сети в условиях перекрывающихся окон временной активации.

Для этих тестов мы сравнивали PSIIICOS с тремя другими методами детекции синхронных источников. Первым из этих методов был метод DICS (см. 1.4.1, [55]). При этом мы восстанавливали элементы кросс-спектра на источниках, не нормируя их на мощность, чтобы корректно сравнить результаты с PSIIICOS.

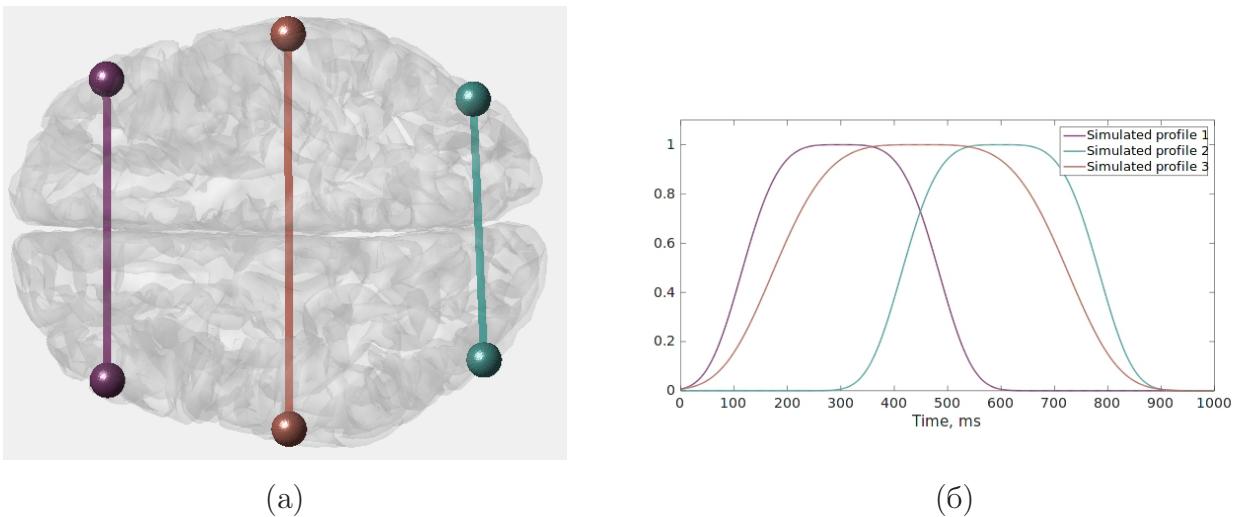


Рисунок 3.2: Три симулированные пары взаимодействующих источников. (а) — пространственное распределение симулированных источников, (б) — временные профили активации для каждой из трех сетей.

Чтобы исключить влияние протечки сигнала, мы также использовали модифицированную версию алгоритма DICS, в которой используется только мнимая часть кросс-спектра (iDICS, см. разд. 1.4.2). Алгоритм iDICS был вторым из набора методов, с которыми мы сравнивали PSICOS.

В качестве третьего метода мы взяли алгоритм GCS 1.4.3. Как и в оригинальной статье [38], для построения обратного оператора мы использовали метод MNE 1.3.3 и использовали известную матрицу ковариации шума мозга. Алгоритму PSICOS эта информация не нужна и поэтому не использовалась.

На графике 3.3 изображена пространственная структура решения для симуляционных данных в пространстве источников. На графике отображены 0.1% сетей с максимальными значениями статистики, оцененной методом PSICOS и другими тремя алгоритмами. Как видно из графика, алгоритм PSICOS сохраняет адекватное качество решения для всего спектра тестируемых условий.

Как и для симуляций в разделе 3.2, метод iDICS показывает себя практически так же хорошо, как и PSICOS для значений фазовой задержки, близких к  $\pi/2$ . Метод MNE GCS для случая высокого ОСШ достаточно надежно определяет две фронтальные сети для обоих значений фазовой задержки, но полностью перестает работать для низкого ОСШ и оклонулевой разности фаз. Для случая фазовой задержки, близкой к  $\pi/2$ , и низкого ОСШ MNE GCS обнаруживает только среднюю сеть (сеть с наибольшим индивидуальным ОСШ) и порождает большое количество ложно-положительных сетей.

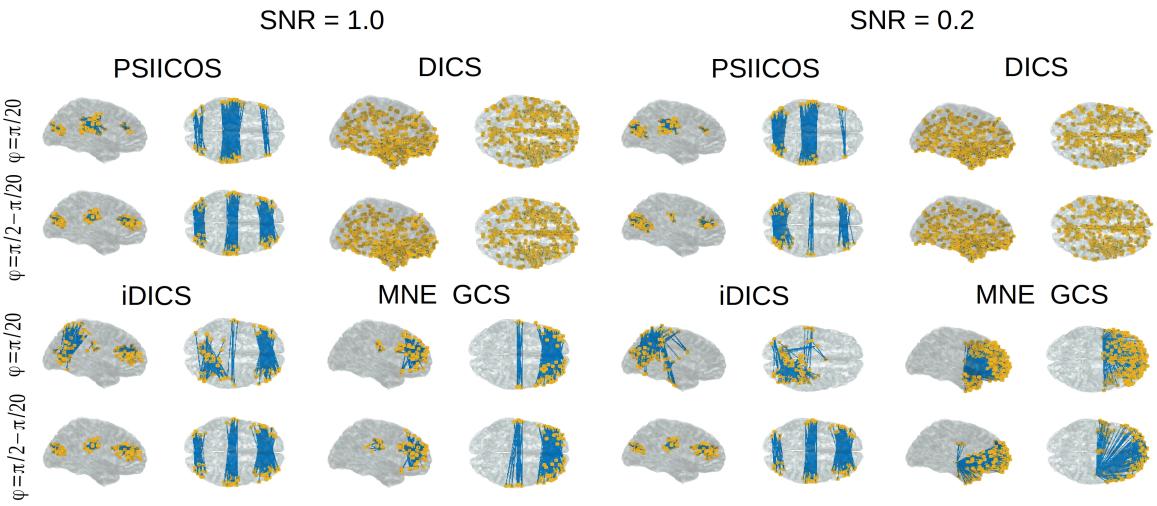


Рисунок 3.3: Пространственное распределение сетевых источников для симуляционных данных.

На графике изображены 0.1% сетей с наибольшим значением статистики. Из графиков видно, что только алгоритм PSIIICOS сохраняет качество решения для всего диапазона изучаемых значений. Метод iDICS ведет себя практически также хорошо, как и PSIIICOS, для фазовых сдвигов, близких к  $\pi/2$ , что согласуется с нашими предыдущими результатами. Метод MNE GCS для высоких значений ОСШ достаточно надежно определяет две фронтальные сети для каждого значения фазового сдвига, но полностью перестает работать для низких значений ОСШ и близких к нулю фазовых задержек. Для задержек, близких к  $\pi/2$ , при низком ОСШ этот метод находит только среднюю сеть (для нее индивидуальное ОСШ выше всего) а также порождает большое количество ложноположительных срабатываний.

Умножая временной ряд кросс-спектра на сенсорах на среднюю топографию полученных кластеров, можем восстановить временную активность сетей. На 3.4 изображены временные окна, в течение которых каждая пара взаимодействующих источников была активна. Для построения этого графика мы использовали спроектированный от протечки сигнала временной ряд кросс-спектра, к которому мы применяли наиболее простую процедуру оценки источников — фильтр, максимизирующий SNR в заданной точке 1.3.2, с той разницей, что здесь все операции проводились в пространстве 2-топографий с размерностью  $M^2$ .

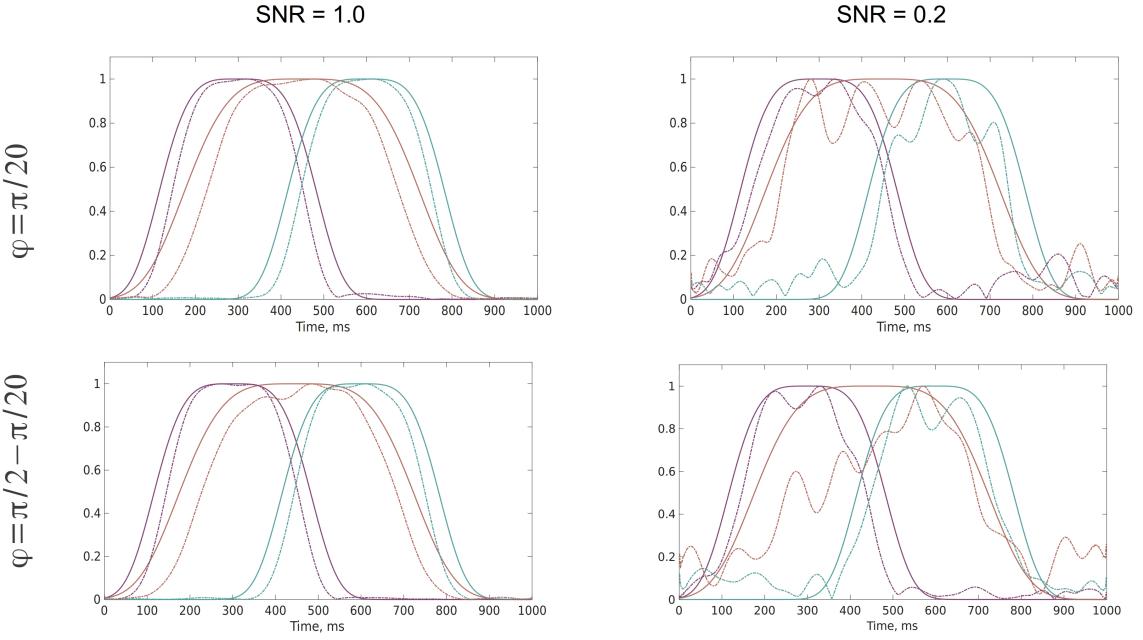


Рисунок 3.4: Временные профили активации симулированных сетей, восстановленные при помощи PSICOS, для двух значений фазового сдвига и двух значений ОСШ.

Непрерывные линии отображают истинные симулированные профили, а пунктир — восстановленные алгоритмом PSICOS.

Далее, для системного сравнения четырех рассматриваемых методов на графике 3.5 мы модифицированные кривые Precision-Recall, на которых размером маркера мы отобразили количество истинных сетей, обнаруженных алгоритмом. На левой части рисунка изображен полный график Precision-Recall; на панелях (b) и (d) показаны данные в диапазоне 0–0.15.

Структура этих графиков подтверждает выводы, основанные на визуальном анализе пространственного распределения обнаруженных сетей (см. рис. 3.3).

Использованные до сих пор метрики оценивают качество решения независимо от порога, что не годится для оценки качества работы алгоритма на реальных данных. Для применения к реальным данным мы предлагаем использовать процедуру бутстрэпа, описанную в разделе 3.1, чтобы оценить стабильность получаемых решений и использовать ее в качестве метрики качества. График 3.6 иллюстрирует результаты процедуры бутстрэпа в применении к реалистичным симуляционным данным.

### Realistic simulations, three transient networks

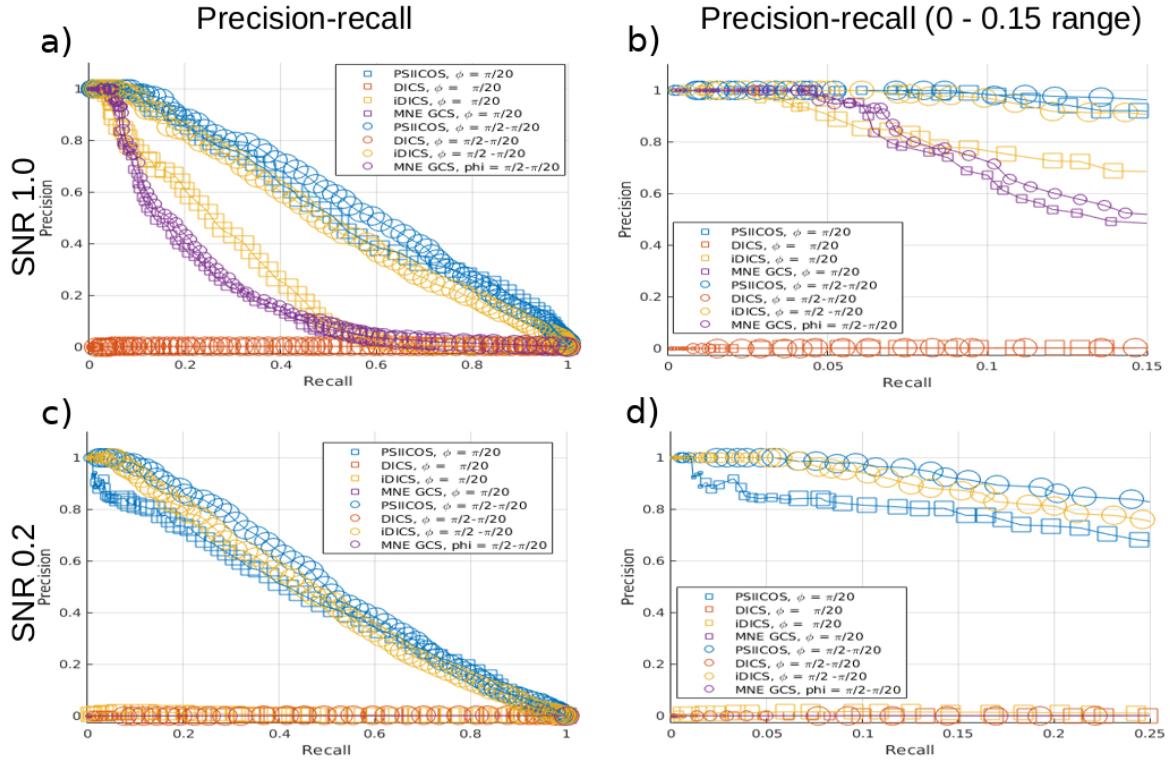


Рисунок 3.5: Сравнение PSIIICOS с остальными методами в задаче обнаружения трёх одновременно активных сетей.

На рисунках (а), (с) изображены кривые Precision-Recall для всего диапазона возможных значений; на рисунках (б), (д) изображены те же кривые, но для диапазона значений 0–0.15.

Наши симуляции показывают, что предложенный подход позволяет достичь лучшего качества решений при поиске взаимодействующих источников как при оклонулевых, так и при близких к  $\pi/2$  фазовых сдвигах. Более того, алгоритм PSIIICOS позволяет обнаруживать сети для всего спектра возможных фазовых задержек. Чтобы дополнительно проиллюстрировать это качество, мы применили предложенный метод к симуляционным данным с тремя сетями на равномерной сетке фазовых задержек в диапазоне от 0 до  $\pi/2$ . Этот анализ мы провели для двух значений ОСШ и сравнили три различных подхода: использование только мнимой части (iDICCS), использование только спроектированной действительной части и использование полной кросс-спектральной матрицы (PSIIICOS). На графике 3.7 изображена метрика качества решения в

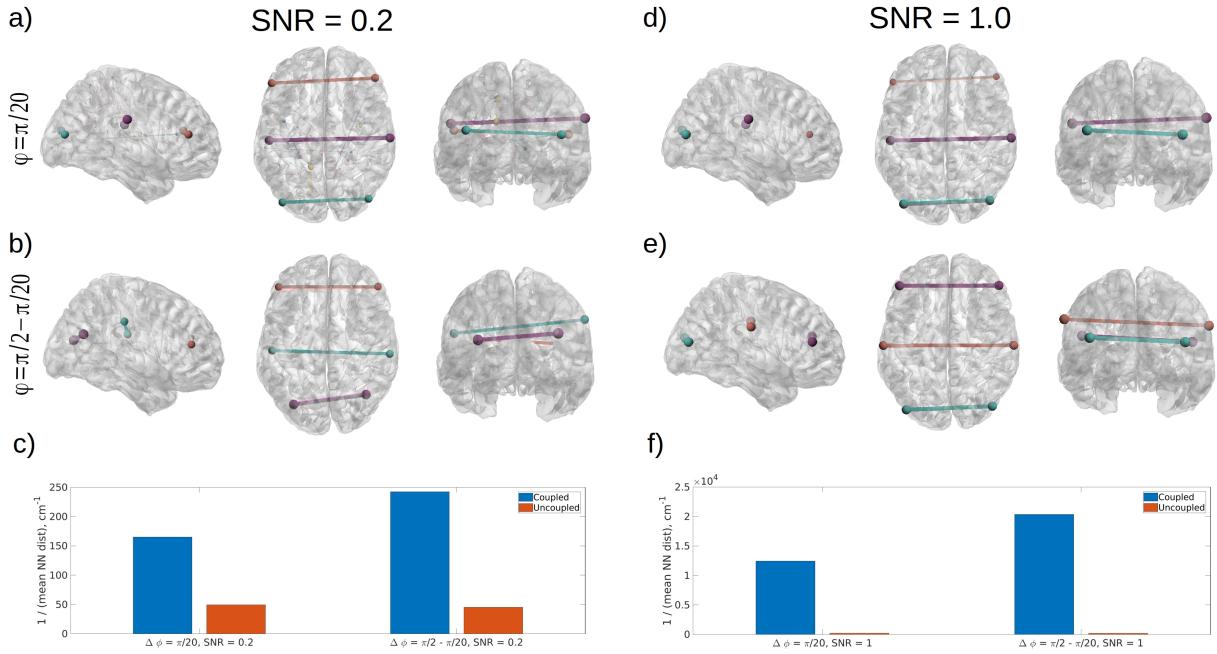


Рисунок 3.6: Исследование стабильности метода PSIICOS при помощи процедуры бутстрэпа для двух значений фазового сдвига и двух ОСШ.

(a), (b), (d), (f) — стабильные сети отображены в виде отрезков, концы которых соответствуют взаимодействующим областям коры. Насыщенность цвета отрезка и его концов прямо пропорциональна коэффициенту воспроизводимости  $\eta$ , оцененному при помощи процедуры бутстрэпа. Отрезки, концы которых оказывались ближе 1 см друг к другу, мы рисовали одним цветом. На графиках (c) и (d) изображен индекс воспроизводимости  $\eta$  для двух средних фазовых сдвигов, полученных в результате описанной процедуры бутстрэпа для двух различных значений ОСШ ( $\text{ОСШ}=0.2$  и  $\text{ОСШ}=1$  соответственно).

каждом из случаев для различных фазовых сдвигов. В качестве метрики мы использовали площадь под кривой Precision-Recall. Для значений фазы, близких к  $\pi/2$ , информация о взаимодействующих источниках содержится в основном во мнимой части кросс-спектра, что легко заметить из графика 3.7, где использование только мнимой части позволяет достичь высокого качества решения для близких к  $\pi/2$  фазовых сдвигов. Для реалистичного значения  $\text{ОСШ}=0.1$  качество решения резко ухудшается как только значение фазового сдвига покидает окрестность точки  $\pi/2$ .

С другой стороны, для близких к нулю сдвигов информация о коннективности содержится в основном в действительной части. Действительная часть

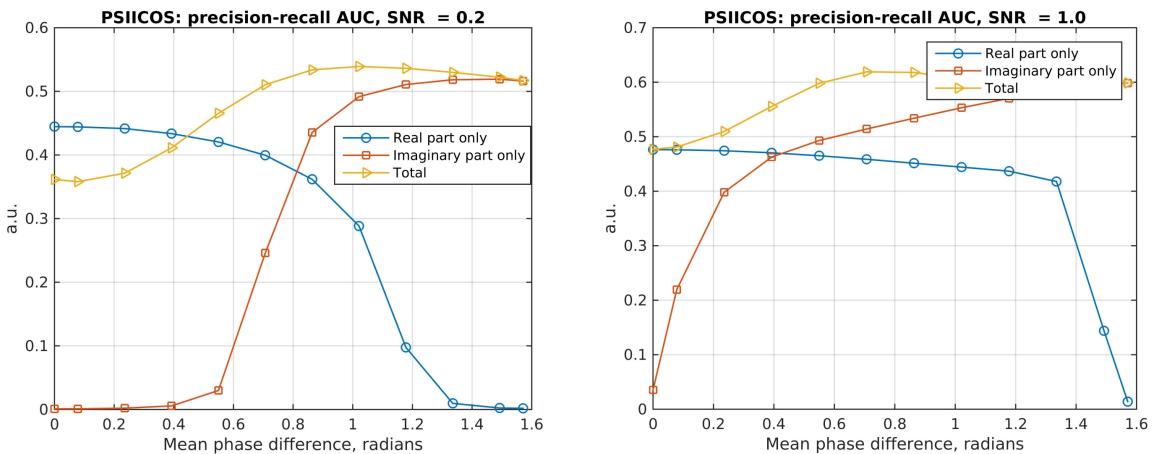


Рисунок 3.7: Качество решений алгоритма PSIIICOS как функция средней фазовой задержки

(a) — для низкого значения ОСШ и (b) — для высокого. Для близких к  $\pi/2$  средних значений фазового сдвига информация о взаимодействии источников содержится в основном во мнимой части кросс-спектра. Поэтому использование только мнимой части (что эквивалентно iDICS) позволяет достигать высокого качества решения; качество решения при этом затухает с уменьшением фазового сдвига. С другой стороны, для близких к нулю значений фазовой задержки информация о взаимодействии содержится в основном в действительной части кросс-спектра. Действительная часть загрязнена протечкой сигнала, которая может быть удалена при помощи PSIIICOS-проекции. После этого очищенная действительная часть используется для обнаружения сетей с близкими к нулю фазовыми задержками (синие кривые). Одновременное использование действительной и мнимой компоненты позволяет добиться равномерного качества решений по всему диапазону средних фазовых сдвигов (оранжевая кривая).

загрязнена протечкой сигнала, которую мы эффективно удаляем PSIIICOS-проекцией. Далее спроектированная действительная часть может быть использована для обнаружения сетей с близкими к нулю значениями фазового сдвига. Предложенный метод PSIIICOS позволяет использовать очищенную действительную и мнимую части совместно для достижения равномерного по фазовым сдвигам качества решения.

### 3.4 Сравнение на реальных данных.

Так как главной целью данного исследования является изложение новой методологической базы для оценки коннективности и валидации результатов при помощи симуляций, важно проиллюстрировать поведение метода на МЭГ данных из реального эксперимента. Тем не менее, полный групповой анализ и систематическое сравнение с другими методами лежит за пределами этого в большей степени методологического исследования. Поэтому нашей задачей было продемонстрировать, как наш метод может быть использован на реальных данных и оценить его устойчивость. Необходимо понимать, что оценка работы метода на реальных данных не способна предоставить критерий качества работы алгоритма, так как мы не знаем истинной картины взаимодействия участков коры. В то же время мы можем продемонстрировать методику возможного использования нашего алгоритма и оценить его устойчивость к изменчивой структуре данных методом бутстрэпа.

Мы применили предложенный метод для анализа реальных МЭГ-данных, полученных для исследования воображаемых вращений в статье [76]. Часть данных для одного испытуемого была выложена в свободный доступ в ходе соревнования по анализу данных на конференции Biomag 2010.

В эксперименте испытуемый определял латеральность изображенной на экране кисти руки нажатием на одну из двух соответствующих кнопок. Изображения рук были повернуты к наблюдателю на некоторый угол, и для ответа испытуемый должен был мысленно их повернуть. Авторы статьи исследовали электрическую активность коры при осуществлении воображаемых вращений.

Данные содержали 800 эпох с рандомизированными углами поворота и латеральностями. Эти 800 эпох были разделены на 5 блоков. Каждая эпоха включала следующую последовательность событий: фиксационный крест (3 сек. предъявления), изображение руки (предъявление до момента ответа испытуемого), фиксационный крест (0.5 сек.), окрашенный в красный или зеленый цвет в зависимости от правильности ответа испытуемого. В среднем испытуемые проводили в МЭГ-сканнере 70 минут.

В ходе эксперимента мозговая активность регистрировалась МЭГ-устройством системы VSM/CTF Systems, включающим 151 аксиальный градиометр. Аналоговый сигнал был отфильтрован фильтром низких частот с частотой

среза 300 Гц и сэмплирован на частоте 1200 Гц. После удаления артефактных сегментов для анализа осталось 259 эпох, соответствующих изображениям левой руки. Их длительность лежала в пределах от 1.52 сек до 2.08 сек. Мы выровняли эпохи таким образом, чтобы начало каждой эпохи соответствовало моменту 0.5 сек до предъявления изображения. Для анализа мы брали 1 секунду после предъявления стимула.

Для анализа мы использовали следующие полосы частот: тета (4–8 Гц), альфа (8–12 Гц), бета (16–24 Гц), нижняя гамма (30–60 Гц), верхняя гамма (65–85 Гц).

На первом шаге мы применяли полосовой КИХ-фильтр с нулевой фазой для отделения целевой полосы частот. Далее мы усредняли по эпохам внешние произведения временных рядов с примененным преобразованием в аналитический сигнал. Таким образом мы получали временной ряд кросс-спектра в виде трехмерного массива, который после векторизации задавался матрицей  $M^2 \times T$ , где  $M = 151$  — количество сенсоров в установке. Далее мы применяли к временному ряду кросс-спектра PSIICOS-проекцию с рангом  $R = 150$ , который мы определяли исходя из соображений, изложенных в разделе 3.5. Мы проводили анализ и с другими значениями параметра  $R$  в диапазоне от 100 до 250, и в этих пределах ранг проекции не оказывал значительного влияния на решение.

Далее мы отдельно анализировали действительную и мнимую части векторизованного временного ряда кросс-спектра. Вместо простого усреднения вдоль временной оси для восстановления пространственной структуры мы вычисляли SVD-разложение действительной и мнимой части спроектированного кросс-спектрального временного ряда. Далее мы проводили поиск сетей-источников для первых двух левых собственных векторов мнимой и действительной части на подробной сетке из 15000 узлов.

Для проверки стабильности получаемых решений мы использовали процедуру бутстрэпа, описанную в разделе 3.1 для  $m = 30$  и  $B = 100$  и отрисовывали индекс воспроизводимости  $\eta$ . Мы отображали полученные сети только в рамках стабильных, воспроизводимых на бутстрэпе пучков сетей. Сети, которые отстояли друг от друга на расстояние меньше 1 см, мы рисовали одним цветом. Также для каждой воспроизводимой компоненты мы рисовали временной профиль активации как правый собственный вектор временного ряда кросс-спектра.

На графике 3.8 мы отобразили результаты частотно-специфичного анализа воспроизводимости решения. На графике изображена величина, обратная

к расстоянию до ближайшей сети, сгруппированная по 5 полосам частот для действительной и мнимой частей. Как видно из графика, сети, обнаруженные с использованием действительной части кросс-спектра, оказываются более воспроизводимыми, чем сети, обнаруженные по мнимой части. Наиболее воспроизводимые сети по действительной части, находятся в бета и обоих гамма-диапазонах. Воспроизводимые сети по мнимой части кросс-спектра принадлежат полосам частот альфа, бета и верхняя гамма. Важно отметить, что хотя сети по действительной части в целом характеризуются большей воспроизводимостью, это оказывается не так для полосы частот альфа.

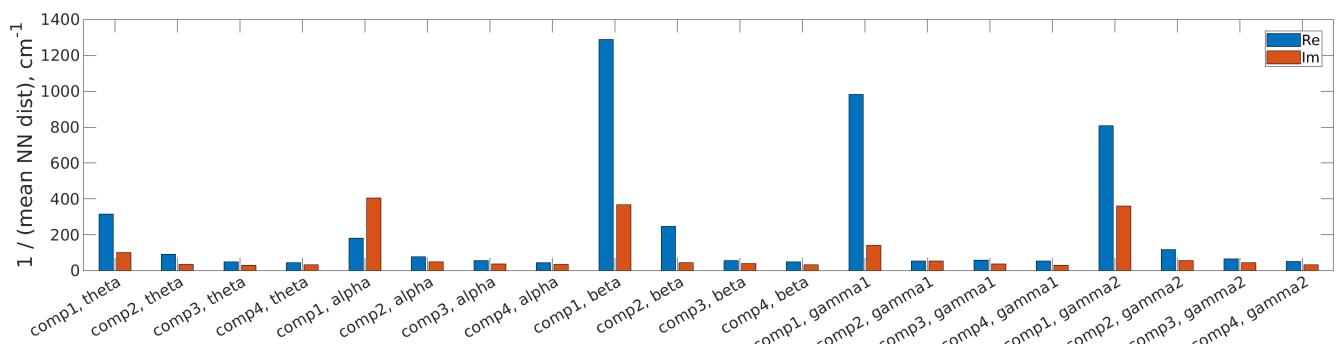


Рисунок 3.8: Полосы частот с наиболее воспроизводимыми сетями, определенными при помощи процедуры бутстрэпа.

Бутстрэп проводился последовательным выбором подмножества эпох, на котором вычислялся кросс-спектр, к которому затем применялась PSICOS-проекция. Столбики на графике отображают индекс воспроизводимости, который вычислялся как обратное к среднему расстоянию до ближайшего соседа для каждой отдельной сети, найденной по мнимой и действительной спроектированной части кросс-спектра для четырех наиболее значимых главных компонент. Результаты для действительной части представлены синим цветом, а для мнимой — желтым.

На графиках 3.9, 3.10, 3.11 изображена пространственная структура обнаруженных сетей для этих частотных диапазонов. Сети, узлы которых оказывались ближе 1 сантиметра друг к другу, мы рисовали одним цветом. Перекрывающиеся сети мы рисовали с увеличенным размером маркера и увеличенной толщиной линии. Прозрачность линии мы меняли в зависимости от плотности сетей в кластере с учетом перекрытий. Также для каждой воспроизводимой компоненты мы рисовали профиль ее временной активации как правый собственный вектор сингулярного разложения соответствующей части временного ряда кросс-спектра.

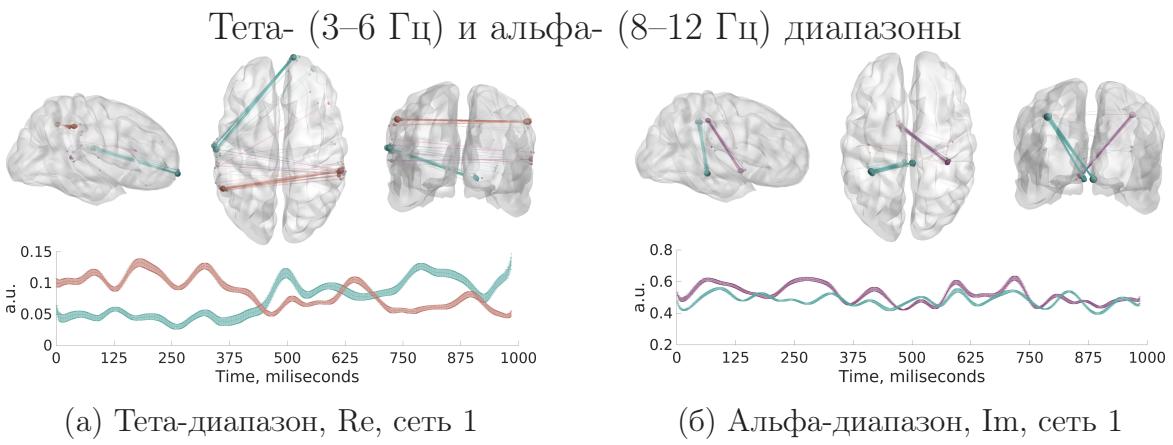


Рисунок 3.9: Пространственная структура и временная динамика наиболее воспроизводимых сетей в диапазонах частот тета (3–6 Гц) и альфа (8–12 Гц)

Анализ действительной части кросс-спектра показывает наличие следующих сетей с преимущественно малыми фазовыми задержками. В тета-диапазоне (рис. 3.9) мы видим сети, соединяющие правую орбито-фронтальную кору, вовлеченную в сенсорную интеграцию, с левым билатеральным височнотеменным узлом, про который известно, что он активен при воображении движения ([77]). Также мы видим кросс-латеральную сеть, соединяющую два теменных отдела.

В бета-диапазоне мы видим механистически правдоподобную кросс-латеральную связь между вентральным зрительным путем и зоной представления руки в моторной области правого полушария. Дополнительно мы наблюдаем взаимодействие между зонами представительства руки сенсомоторной области правого полушария и соматосенсорной области левого, рис. 3.10, что частично согласуется с наблюдениями, описанными в статье [78] и соотносится с профилями функциональной коннективности, описанными в фМРТ-исследовании воображаемых вращений [79]. Наконец, мы обнаружили взаимодействие зоны представительства руки в моторной области правого полушария с височным полюсом левого полушария в гамма-диапазоне, рис. 3.11а, а также с орбитофронтальной зоной левого полушария в верхнем гамма-диапазоне, рис. 3.11б. Правдоподобность этих наблюдений подкреплена предполагаемой ролью височного полюса левого полушария, в которую входит “... визуальное различение двумерных изображений и мнемоническая функция соотнесения и научения” ([80]).

Сети, найденные по мнимой части кросс-спектра, включают связь в альфа-диапазоне билатеральных сенсорных областей со срединной внутренней

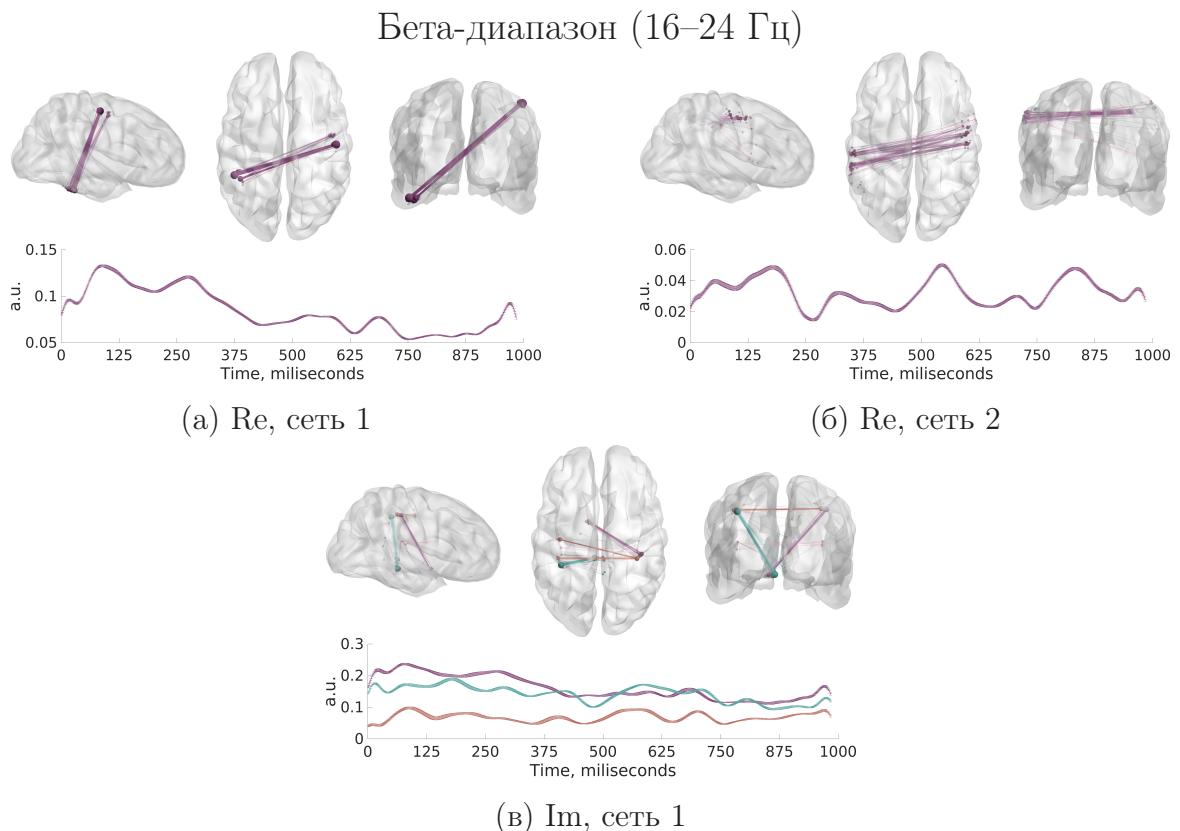


Рисунок 3.10: Пространственная структура и временная динамика наиболее воспроизводимых сетей в бета-диапазоне (16–24 Гц)

поверхностью коры, 3.96, а также кросс-латеральные связи между относящимися к руке зонами сенсорной коры в бета-диапазоне, 3.10в.

Мы также сравнили результаты нашего анализа коннективности с пространственно-специфичным распределением индуцированной мощности по коре, Б. Результаты этого сравнения показывают, что все обнаруженные сети кроме одной билатеральной сети в бета-диапазоне не демонстрируют ситуации, когда оба узла сети лежат в областях с высокой мощностью. В то же время, сенсомоторная область правого полушария, будучи узлом большинства найденных сетей, совпадает с областью наибольшей активации. Так как правая мотосенсорная кора играет ключевую роль в исходном исследовании [76] и так как мы оценивали коннективность не для одной точки с остальными частями коры, а всех точек со всеми, мы находим эти результаты обнадеживающими.

### Нижний (30–60 Гц) и верхний (65–85 Hz) гамма-диапазоны

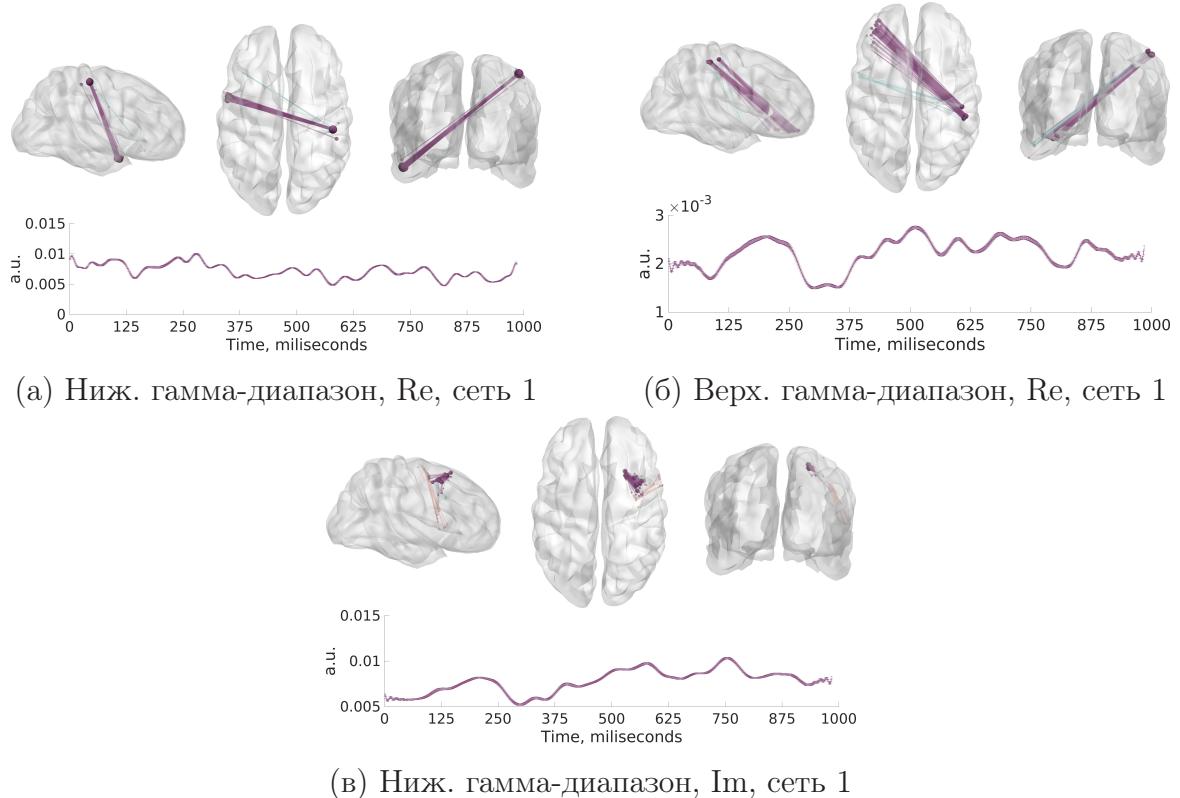


Рисунок 3.11: Пространственная структура и временная динамика наиболее воспроизводимых сетей в нижнем (30–60 Гц) и верхнем (65–85 Гц) гамма-диапазонах

### 3.5 Действие проекции на компоненты кросс-спектра в пространстве источников.

Из-за ортогональности 2-топографий действительной и мнимой частей операция проекции никак не влияет на мнимую часть кросс-спектра. На действительную часть, содержащую информацию о малых фазовых задержках, проекция напротив оказывает значительное влияние. Удаляя часть, соответствующую протечке сигнала, мы неизбежно удаляем и часть полезного сигнала. Качество работы предложенного алгоритма подавления протечки сигнала в действительной части кросс-спектра существенно зависит от следующих двух факторов:

- PSIICOS-проекция должна максимально подавлять мощность компонент в подпространстве протечки сигнала и одновременно с этим минимально изменять мощность истинной действительной части когеренции.

- Как и для всех алгоритмов решения обратной задачи PSICOS-проекция существенно зависит от точности прямой модели. Неизбежная погрешность при ее построении [51] оказывает соответствующее влияние на качество решений предложенного метода. Вместе с тем, ошибки прямой модели в разумных пределах должны вызывать лишь незначительное ухудшение качества решения.

Степень выполнения этих условий зависит от феноменологических характеристик массива сенсоров, а также от магнитных топографий имеющихся мозговых источников. В этом разделе мы опишем результаты численного моделирования, направленного на изучение этой проблемы.

Подавление мощности в подпространстве действительной части кроссспектра очевидным образом проявляется наиболее сильно для источников с кореллированными топографиями. Для исследования зависимости подавления от коэффициента корреляции между топографиями взаимодействующих источников для каждой пары индексов  $i, j$  мы вычисляли нормы до и после проекции для 2-топографий вида  $\mathbf{g}_i \otimes \mathbf{g}_i + \mathbf{g}_j \otimes \mathbf{g}_j$ ,  $\mathbf{g}_j \otimes \mathbf{g}_i + \mathbf{g}_i \otimes \mathbf{g}_j$  а также  $\mathbf{g}_j \otimes \mathbf{g}_i - \mathbf{g}_i \otimes \mathbf{g}_j$ , соответствующие подпространствам  $S^{SL}$ ,  $S^S$ ,  $S^R$ . Результаты представлены на графике 3.12.

Для каждой сети с узлами в точках  $i$  и  $j$  на коре соответствует три точки разных цветов на графике: красная, синяя и желтая. Градации синего цвета мы использовали для кодировки информации о величине нормы 2-топографии действительной части  $\mathbf{g}_j \otimes \mathbf{g}_i + \mathbf{g}_i \otimes \mathbf{g}_j$ . Градации красного цвета использовались для 2-топографий мнимой части, и желтого — для топографий протечки сигнала. Положение каждой цветной точки вдоль оси  $x$  для подграфиков (a), (b) определялось величиной коэффициента корреляции топографий  $\mathbf{g}_i$  и  $\mathbf{g}_j$ . Кроме того, мы отобразили те же самые данные на втором графике, где вместо корреляции по оси  $x$  использовалось расстояние между узлами сети  $i, j$ . Насыщенность цвета точек отражает плотность точек в соответствующей области графика.

### 3.6 Выбор ранга проекции

Мы изучили зависимость коэффициентов ослабления подпространствах протечки сигнала, а также мнимой и действительной части в зависимости от

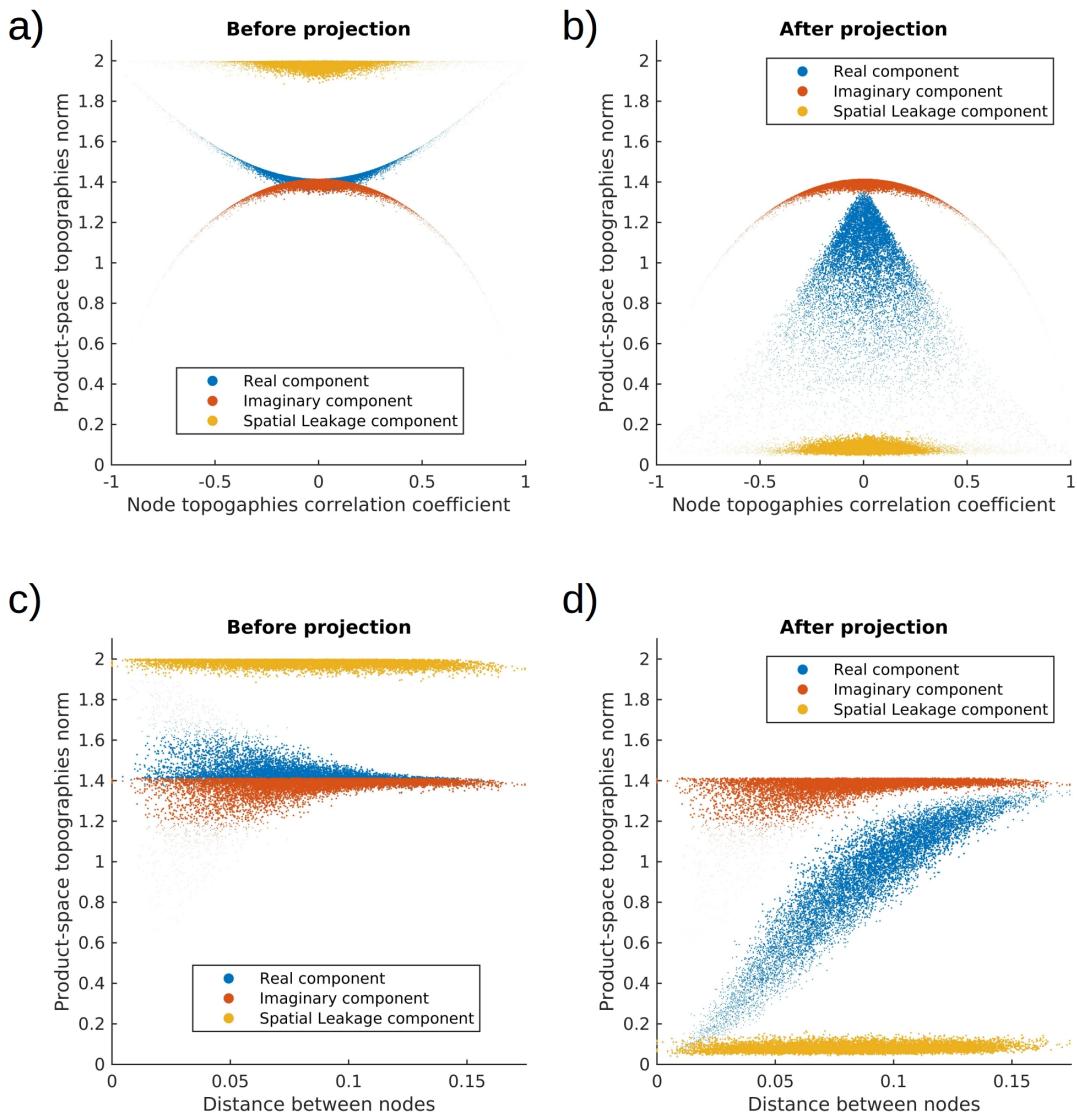


Рисунок 3.12: Нормы 2-тографий для трех подпространств кросс-спектра до и после проекции в зависимости от коэффициента корреляции топографий взаимодействующих узлов на коре (графики (а) и (б)), а также от расстояния между этими узлами (графики (с) и (д)).

До проекции (графики (а), (с)) в кросс-спектре на уровне сенсоров наибольший вклад дает мощностная компонента (обозначена желтым цветом). После проекции (графики (б), (д)) вклад мощностной компоненты уменьшается в среднем в 25 раз. Вместе с тем наблюдается неизбежное, но значительно меньшее ослабление 2-тографий действительной части, соответствующих истинной коннективности. В среднем они ослабляются в 1.6 раз.

ранга проекции. На рисунке 3.13 показано среднее затухание мощности в трех подпространствах в зависимости от ранга проекции.

На практике для ускорения расчетов при обработке МЭГ-данных часто пользуются методом главных компонент для сокращения размерности пространства сенсоров. Различные кривые на графике показывают, каким образом ведет себя PSICOS-проекция для разных рангов при таком сокращении размерности.

Чтобы получить этот график, мы использовали симуляции Монте-Карло. На каждой итерации мы случайным образом выбирали подмножество из 200 источников и вычисляли все 2-топографии из трех подпространств  $S_{SL}$ ,  $S_{\Re}$ ,  $S_{\Im}$ . Мы использовали источники с фиксированными ориентациями, ортогональными к поверхности коры. Затем мы изменяли ранг проекции и вычисляли спроектированные версии 2-топографий этих сетей. Для количественной оценки эффекта затухания мы рассчитывали среднее отношение норм исходных и спроектированных 2-топографий для каждого значения ранга проекции. Этот процесс мы повторили 100 раз и усреднили результат.

Из-за пересечения подпространств  $S_{SL}$  и  $S_{\Re}$  операция проецирования стремится подавить мощность компоненты  $SL$ , что неизбежно приводит к подавлению мощности в подпространстве  $S_{\Re}$ . Чем меньше мощность в подпространстве  $S_{\Re}$  подавляется для фиксированного коэффициента подавления мощности в  $S_{SL}$ , тем выше итоговое качество решения. Поэтому в дополнение к кривым затухания для подпространств мы также нанесли логарифм отношения коэффициентов ослабления, наблюдаемых для векторов в подпространстве  $S_{\Re}$ , к векторам в подпространстве  $S_{SL}$ . Использование SVD-разложения при формировании оператора проекции делает возможным значительно более быстрое уменьшение мощности в подпространстве протечки сигнала  $S_{SL}$ , чем в подпространстве  $S_{\Re}$ . Другими словами мы видим, что увеличение ранга проекции приводит к все большему доминированию мощности в подпространстве  $S_{\Re}$  по сравнению с подпространством  $S_{SL}$ .

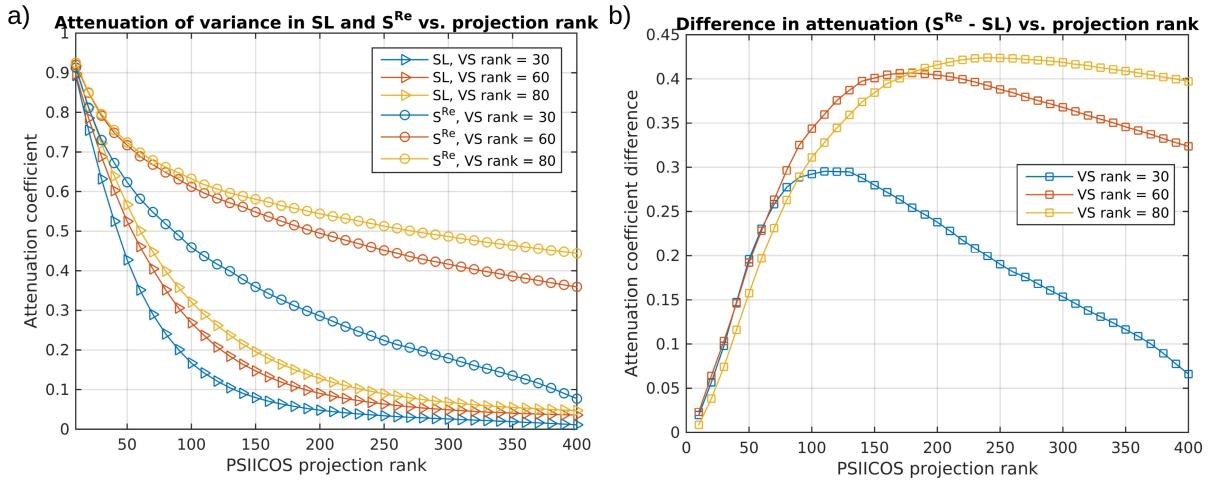


Рисунок 3.13: (а) Ослабление мощности в подпространствах  $S_{SL}$  и  $S_{Re}$  и (б) Разница в коэффициенте затухания для подпространств  $S_{Re}$  и  $S_{SL}$  как функции ранга проекции для трех различных размерностей пространства виртуальных сенсоров (VS).

Эти кривые позволяют принять обоснованное решение относительно выбора ранга проекции. Мы предлагаем выбирать такое значение ранга проекции PSIIICOS, которое соответствует наибольшей разнице в затухании мощности.

### 3.7 Влияние неточностей прямой модели на качество решений

Описанная проекция строится на основании прямой модели, которая в реальных условиях неизбежно содержит к неточностям. В этом разделе мы изучили эффективность предложенной операции проецирования при наличии структурированных и неструктурных ошибок в прямой модели.

В этом исследовании мы использовали две различные модели шума. Первая модель соответствовала пространственно неструктурному шуму и была реализована путем добавления масштабированного случайного шума к матрице прямой модели.

Чтобы получить матрицу реализации шума при оценке прямой модели, мы сэмплировали данные из стандартного нормального распределения  $N(0,1)$ . Далее мы нормировали эти данные умножением элементов матрицы шума на квадратный корень из среднего значения следа матрицы  $\mathbf{G}\mathbf{G}^T$ , где  $\mathbf{G}$  — матрица прямой модели. Затем мы добавляли эту матрицу к истинной прямой модели и корректировали количество шума с помощью параметра  $\alpha$ .

Вторая модель соответствовала более реалистичному сценарию пространственно структурированного искажения. Для генерации пространственно структурированной шумовой матрицы мы использовали модели головы, рассчитанные для  $N = 10$  испытуемых, и посчитали разницу между прямыми моделями для каждой пары субъектов. Мы вычисляли матрицу структурированного шума как среднее из этих попарных разниц. Затем мы нормировали полученную матрицу структурированного шума и добавляли ее к истинной прямой модели. Как и в предыдущем случае, мы регулировали количество шума с помощью параметра  $\alpha$ .

Потенциально неточности прямой модели могут привести к одновременному уменьшению ослабления мощности подпространства протечки сигнала и уменьшению мощности подпространства  $S_{\mathfrak{R}}$ . Мы использовали симуляции Монте-Карло для численного изучения этих эффектов. На каждой итерации мы случайным образом выбирали подмножество из 200 источников и рассчитывали средний коэффициент ослабления для трех подпространств. Мы сделали это как для случая структурированного, так и для неструктурированного шума, а также для двух разных значений ранга проекции (350 и 500). Результаты представлены на рисунке 3.14.

Мы можем видеть, что хотя неточности в прямой модели приводят к ухудшению работы предлагаемого метода проекции, для типичных уровней шума в прямой модели для MEG порядка 10 % ([51]) мы имеем почти одинаковое значение подавления мощности в подпространстве протечки сигнала и только около 20 % дополнительного уменьшения мощности в подпространстве  $S_{\mathfrak{R}}$ .

Из-за пересечения подпространств  $S_{SL}$  и  $S_{\mathfrak{R}}$  при попытке подавить компоненту протечки сигнала, алгоритм также подавляет мощность в подпространстве истинного взаимодействия  $S_{\mathfrak{R}}$ . Неточности в прямой модели увеличивают нежелательное подавление мощности в подпространстве  $S_{\mathfrak{R}}$  и уменьшают эффективность подавления мощности в подпространстве протечки сигнала. Такое ухудшение производительности с увеличением погрешности в прямой модели можно охарактеризовать соотношением коэффициентов затухания 2-топографий в подпространствах  $S_{SL}$  и  $S_{\mathfrak{R}}$ . Когда это соотношение меньше единицы, мы можем сказать, что мощность в подпространстве  $S_{SL}$  уменьшилась сильнее, чем мощность в подпространстве  $S_{\mathfrak{R}}$ . Следовательно чем меньше это соотношение, тем лучше производительность предлагаемой схемы.

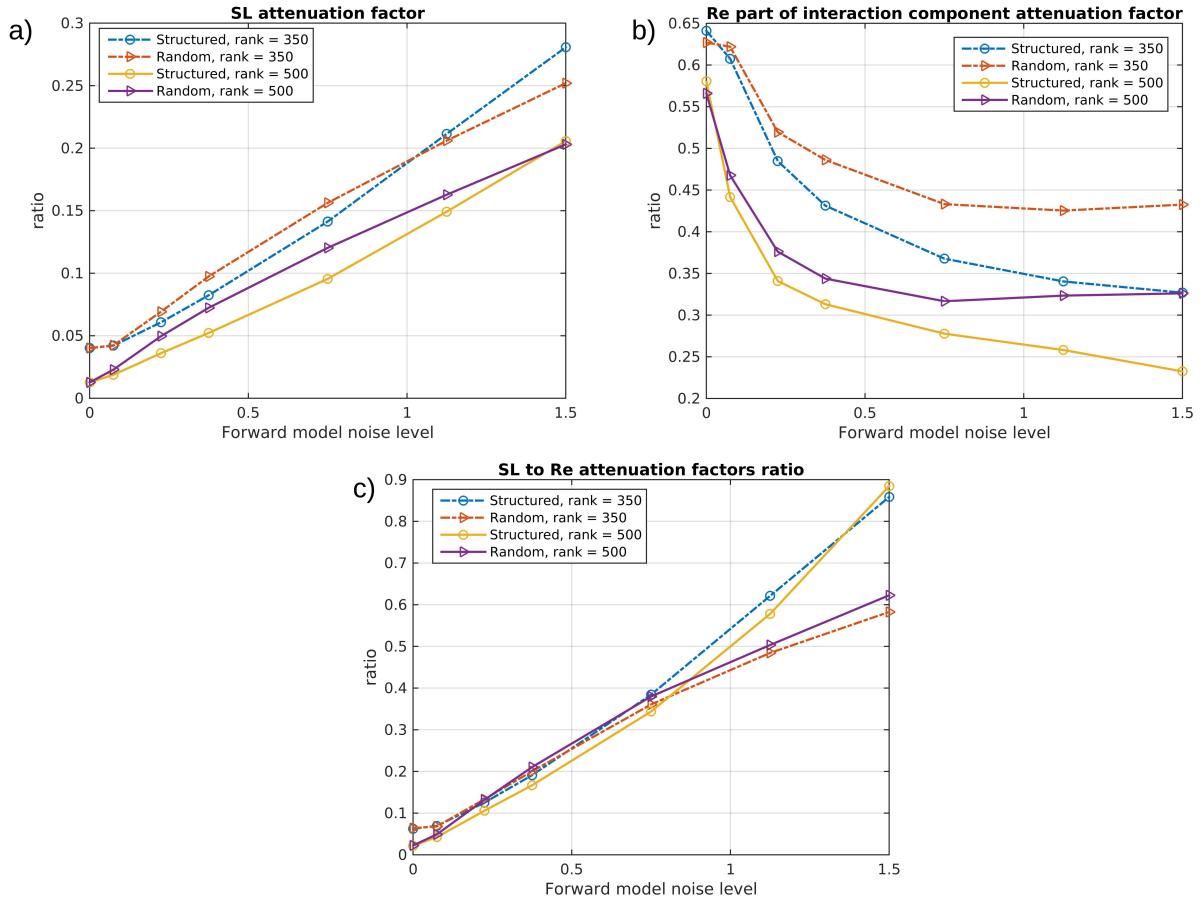


Рисунок 3.14: Влияние шума в прямой модели на качество проекции.

На панели (а) изображена зависимость коэффициента затухания от интенсивности шума  $\alpha$ . Панель (б) показывает коэффициент ослабления мощности в подпространстве  $S_{\Re}$ . На панели (с) нарисовано подавление мощности в подпространстве  $SL$  по сравнению с мощностью в  $S_{\Re}$ . Как видим, оно монотонно убывает с ростом интенсивности шума в прямой модели.

На панели (с) рисунка 3.14 мы отобразили соотношение коэффициентов ослабления мощности в подпространствах  $S_{SL}$  и  $S_{\Re}$ . Как мы видим, значение этого отношения плавно меняется с ростом интенсивности шума в прямой модели и остается низким для уровня шума 10 %.

На основании представленных численных результатов мы можем заключить, что метод является достаточно надежным для типичных уровней шума в прямой модели и может использоваться в реальных условиях. Однако, как и во многих других методах, использующих прямую модель (например, 1.4.1), усилия, направленные на получение более точных прямых моделей способны принести ощутимую отдачу и ими не следует пренебрегать.

### 3.8 Сравнение стратегий нормализации весов фильтра для метода PSICOS.

Как мы уже отмечали в разделе 2.2.2, базовая процедура оценки кросс-спектральных коэффициентов (см. раздел 2.2) в пространстве источников дает пространственно смещенную оценку. В разделе 2.2.2 мы предложили модификацию метода, основанную на другой нормализации весов фильтров по аналогии с методом sLORETA, [43]. В этом разделе мы сравнили качество работы базового алгоритма PSICOS и модифицированного метода с теоретически пространственно несмещенной оценкой, который мы будем называть PSICOS Unbiased.

Для сравнения как и ранее мы использовали симуляции Монте-Карло, описанные в разделе 3.2, со случайно выбираемыми положениями узлов сети для каждой итерации процедуры. Как и в предыдущих разделах, качество работы алгоритмов мы сравнивали на основе ROC-кривых и кривых Precision-Recall. В этом анализе мы ограничились 100 итерациями Монте-Карло.

Результаты этих симуляций для двух значений фазовой задержки ( $\text{ОСШ}=0.2$  и  $\text{ОСШ}=1$ ), а также для трех значений фазового сдвига,  $\varphi = \pi/20$ ,  $\varphi = \pi/4$  и  $\varphi = \pi/2 - \pi/20$ , представлены на рис. 3.15.

Из графика 3.15 видно, что для малых фазовых задержек как в случае высокого, так и в случае низкого значения ОСШ, метод PSICOS Unbiased показывает значительно лучшие результаты. Для остальных значений фазового сдвига качество решений обеих версий алгоритма приблизительно одинаковое с небольшим преимуществом метода PSICOS Unbiased для низкого ОСШ.

Такое преимущество PSICOS Unbiased для  $\varphi = \pi/20$  объясняется тем, что в методе PSICOS при нормализации мы не делали поправку на применение оператора проекции, который не влияет на 2-топографии мнимой части, но может значительно уменьшать норму 2-топографий действительной. В этой связи пространственная смещенность оценки методом PSICOS должна быть наиболее сильно представлена, когда фазовые задержки сетей близки к нулю, что и вызывает снижение качества детекции для случая  $\varphi = \pi/20$ .

На рисунке 3.16 представлен характерный пример, взятый с одной из итераций Монте-Карло ( $\text{ОСШ}=1$  и  $\varphi = \pi/20$ ), когда алгоритм PSICOS показал плохие результаты по сравнению с PSICOS Unbiased. На графике хорошо виден эффект пространственной смещенности метода PSICOS для более глубокого

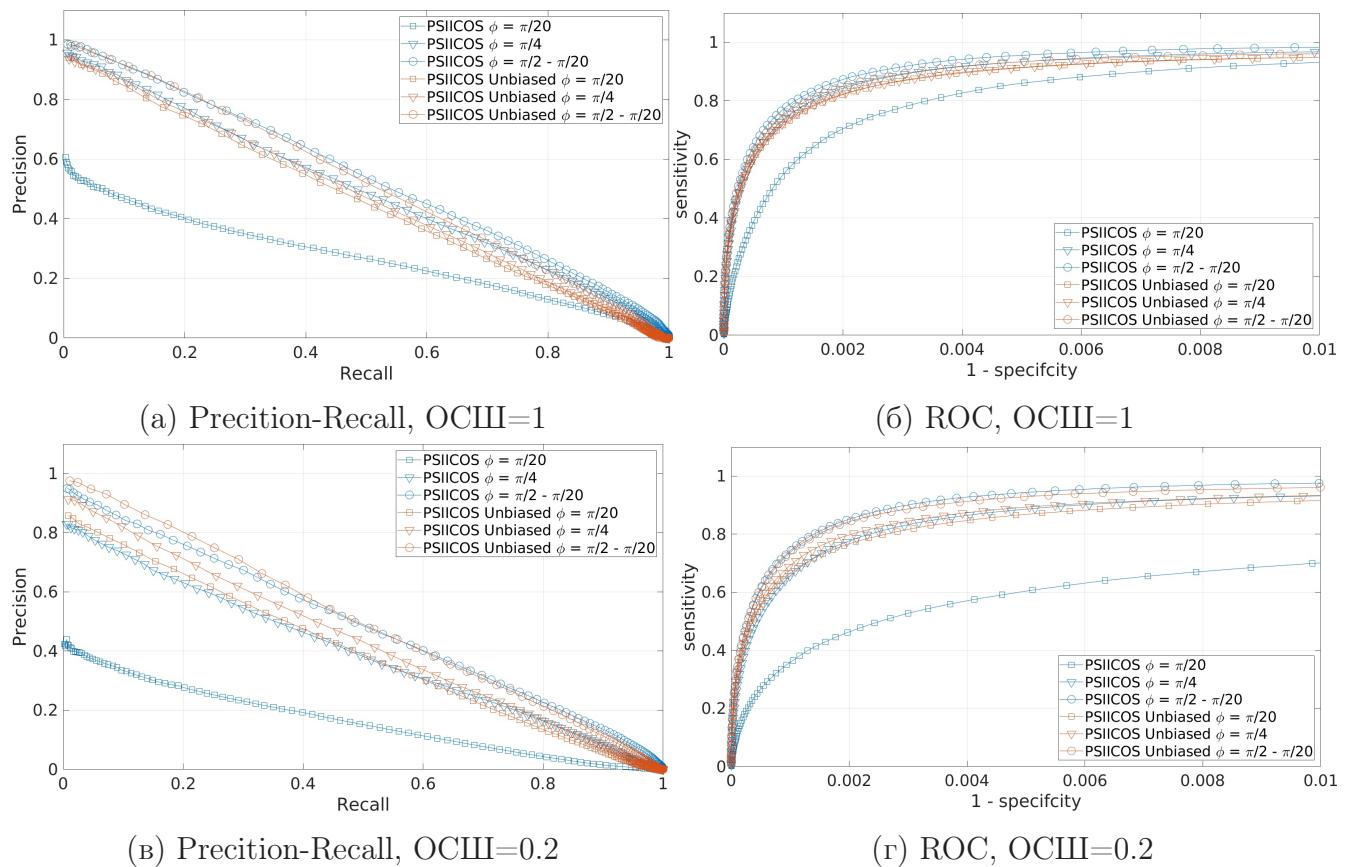


Рисунок 3.15: Сравнение методов PSIIICOS и PSIIICOS Unbiased в задаче детекции сетей со случайными положениями узлов для 100 итераций Монте-Карло.

источника в случае низкой фазовой задержки. Похожая картина наблюдалась и в других отдельно взятых итерациях Монте-Карло, в которых метод PSIIICOS показывал плохое качество решения на индивидуальной кривой Precision-Recall.

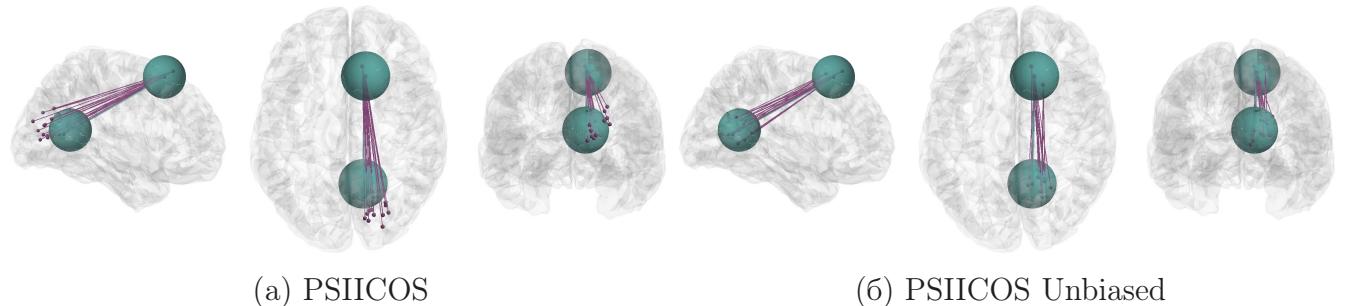


Рисунок 3.16: Пример плохого решения для алгоритма PSIIICOS по сравнению с PSIIICOS Unbiased в задаче детекции сетей со случайными положениями узлов для одной из итераций Монте-Карло и фазового сдвига  $\phi = \pi/20$ .

Фиолетовые отрезки (по 20 для каждого метода) — найденные сети; зелеными сферами отмечены зоны, внутри которых найденные сети относились к истинно положительным срабатываниям. Радиус каждой сферы равен 1.8 см.

Далее мы сравнили качество работы двух версий алгоритма PSIIICOS в задаче детекции трех сетей с фиксированными положениями. Пространственная структура и временные профили активации брались в соответствии с описанием в разделе 3.3. Как и для сетей со случайными положениями узлов, мы симулировали сети для двух значений ОСШ и трех значений фазового сдвига: ОСШ=0.2, ОСШ=16  $\phi = \pi/20, \phi = \pi/4, \phi = \pi/2 - \pi/20$ .

Так как в этих симуляциях пространственная структура шума мозга случайна (она задается как 1000 случайно разбросанных по коре генераторов, см. раздел 3.2), результаты в разных повторениях этих симуляций несколько отличаются. Чтобы уменьшить влияние этой случайности, мы усреднили итоговые кривые по 10 повторениям эксперимента.

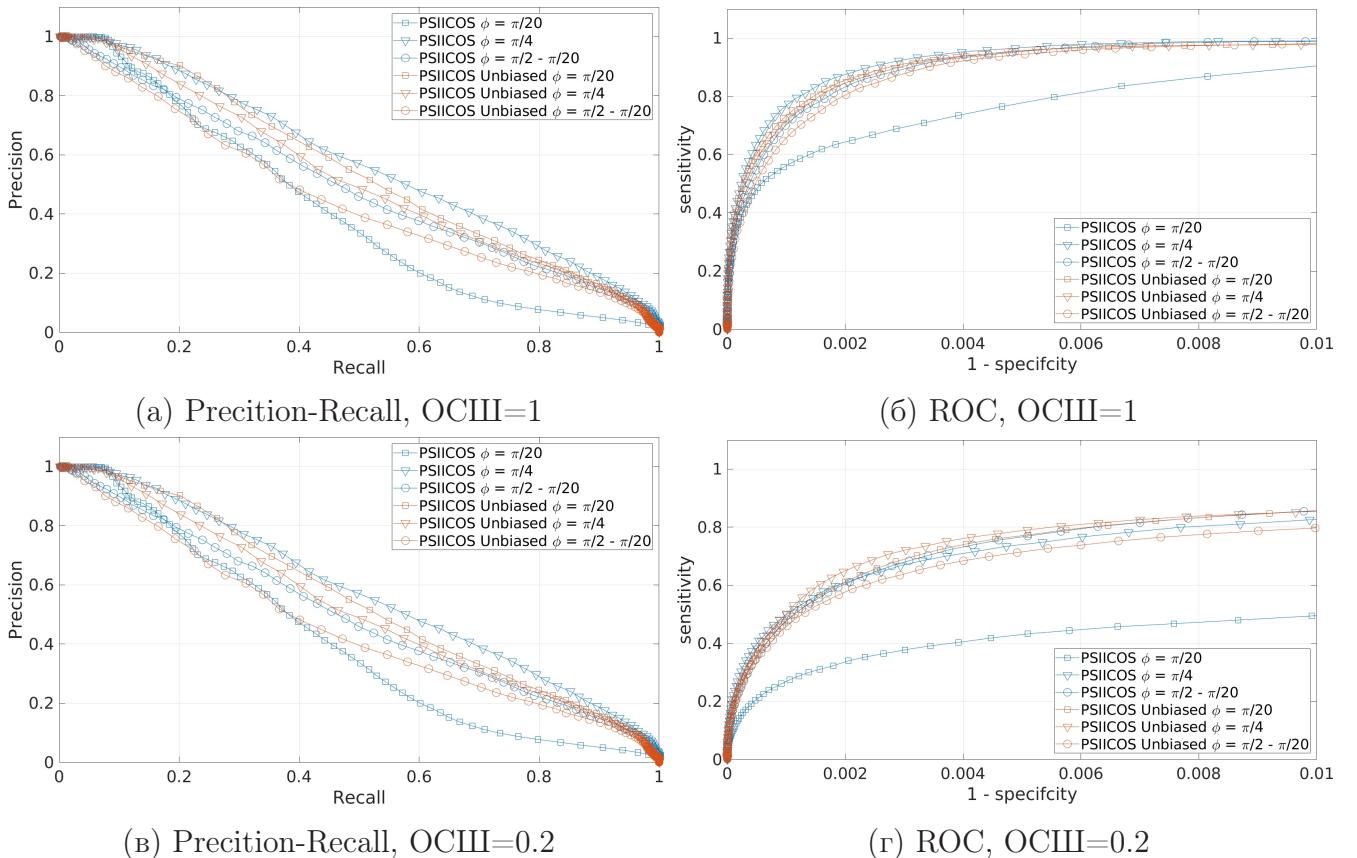


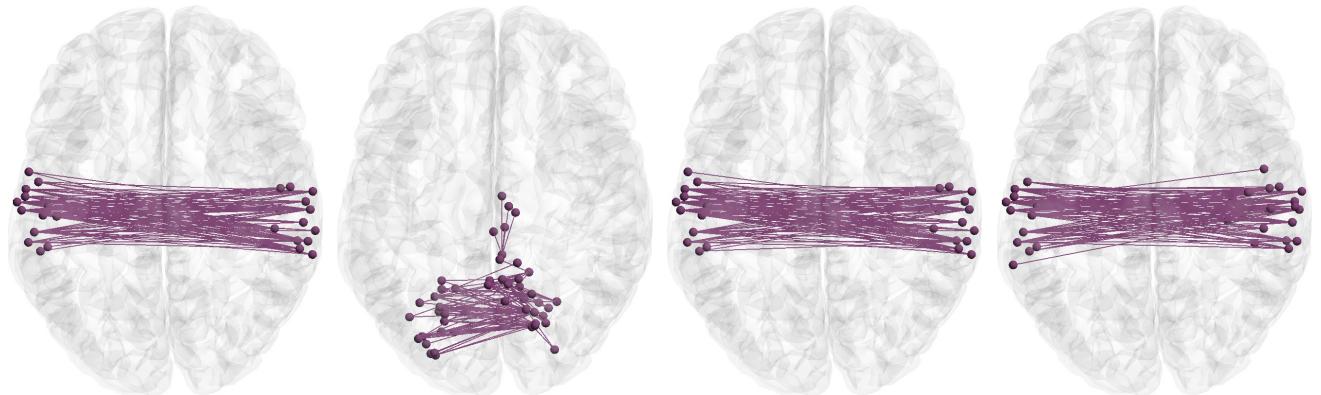
Рисунок 3.17: Сравнение методов PSIIICOS и PSIIICOS Unbiased в задаче детекции трех сетей с фиксированными положениями источников и перекрывающимися временными профилями активации, усредненные по 10 повторениям.

Результаты представлены на рис. 3.17. Из этих графиков мы вновь видим, что качество решений метода PSIIICOS для малых фазовых задержек несколько проседает по сравнению с PSIIICOS Unbiased. При этом для фазовых сдвигов  $\phi = \pi/4, \phi = \pi/2 - \pi/20$  качество решений PSIIICOS по кривым оказывается несколько лучше, чем таковое для PSIIICOS Unbiased. Это объясняется способом

построения соответствующих кривых и пространственной структурой сетей: метод PSIICOS Unbiased для этой конкретной конфигурации склонен накапливать больше сетей вокруг одной из истинных локаций, прежде чем переходить к остальным. Это приводит к выходу пучка отрезков за пределы сферы, в которую должно попасть решение, что с точки зрения нашей процедуры является ложным срабатыванием, но для реальных приложений не критично.

### 3.9 Численное исследование метода нормализации кросс-спектральных коэффициентов в пространстве источников (PSIICOS Normalized).

В разделе 2.2.5 мы обозначили проблему локализации методом PSIICOS, которая заключается в том, что элементы кросс-спектра зависят не только от фазовой связности источников, но и от их амплитуд, что не позволяет сделать однозначного вывода о наличии коннективности между парой источников.



(а)  $\alpha = 0.1$ , ОСШ=0.2 (б)  $\alpha = 2$ , ОСШ=0.2 (в)  $\alpha = 0.1$ , ОСШ=1 (г)  $\alpha = 2$ , ОСШ=1

Рисунок 3.18: Локализация сети с фиксированными положениями узлов методом PSIICOS Unbiased для фазового сдвига  $\varphi = \pi/4$  для высокого и низкого ОСШ (0.2 и 1) в случае сильной связности ( $\alpha = 0.1$ ) и отсутствия связности ( $\alpha = 2$ )

На рис. 3.18 мы проиллюстрировали этот эффект. Для построения этого графика мы использовали симуляции, аналогичные описанным в разделе 3.3, но на этот раз мы ограничились только одной центральной сетью. На рис. 3.18 мы изобразили 100 самых мощных сетей, полученных алгоритмом PSIICOS

Unbiased для низкого и высокого значения ОСШ при низком и высоком значении разброса фазы сети,  $\alpha$ . Параметр  $\alpha$  при симуляциях регулировал величину случайной добавки к постоянной разности фаз для каждой сети. Чем ближе к нулю значение  $\alpha$ , тем сильнее фазовая связность. Значение  $\alpha = 2$  соответствует полному отсутствию связности.

Из рис. 3.18 видно, что алгоритм корректно находит сети в случае высокой фазовой связности ( $\alpha = 0.1$ , панели 3.18a, 3.18b), однако при отсутствии фазовой связности ситуация не столь хороша. В этом случае при низком ОСШ (панель 3.18б) алгоритм находит шумовые сети, но так как выдаваемые им значения для сетей ненормированные, мы не можем отбросить эти сети по одному прогону алгоритма. Тем не менее, эту проблему можно решить, используя процедуру бутстрэпа и считая индекс воспроизводимости: для шумовых сетей он будет низким, а для настоящих — высоким.

Хуже обстоит дело в случае отсутствия фазовой связности и высокого значения ОСШ (панель 3.18г). В при таких параметрах алгоритм находит ошибочно находит центральную сеть из-за того, что мощность ее узлов выделяется на фоне окружающего шума, а также из-за коррелированности их временных профилей активации. В этом случае процедура бутстрала покажет воспроизводимую сеть, хотя в действительности фазовой связности нет.

В разделе 2.2.5 мы предложили модификацию метода PSICOS, которая осуществляет нормализацию кросс-спектральных коэффициентов с учетом наличия протечки сигнала. За счет нормализации мы можем справиться с ложноположительными срабатываниями, аналогичными случаю 3.18г.

Чтобы продемонстрировать это, мы симулировали данные для одной центральной сети для набора индексов фазовой связности в диапазоне от 0.1 (сильная связность) до 2 (связность отсутствует) для низкого и высокого значения ОСШ (0.2 и 1 соответственно) и для трех значений фазовых сдвигов ( $\varphi = \pi/20, \varphi = \pi/4, \varphi = \pi/2 - \pi/20$ ). Для каждой комбинации параметров мы применяли метод PSICOS Unbiased и отбирали сети с максимальным абсолютным значением оцененного элемента кросс-спектра. Для этой сети мы проводили нормализацию в соответствии с процедурой, описанной в разд. 2.2.5 и рисовали полученное значение на графике. Результаты представлены на рис. 3.19.

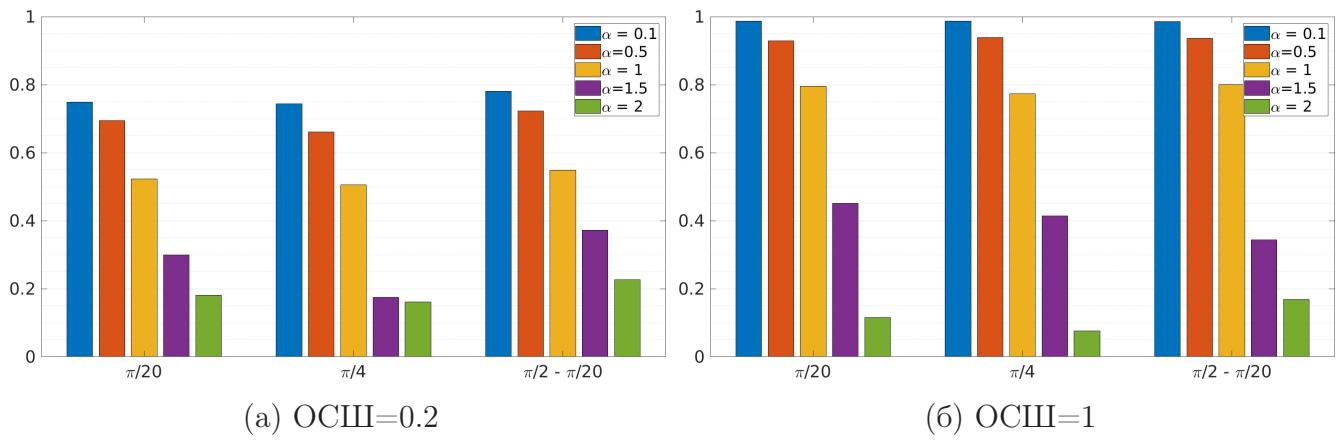


Рисунок 3.19: Нормализованный индекс фазовой связности, полученный методом PSIICOS Normalized для одной симулированной сети в условиях высокого и низкого ОСШ в зависимости от индекса истинной фазовой связности  $\alpha$  и значения фазовой задержки.

Из этого графика видно, что в случае, когда фазовая связность отсутствует ( $\alpha = 2$ ), полученные нормализованные значения не превосходят порога в 0.3. При этом нормализованный индекс монотонно убывает с ростом параметра  $\alpha$ .

Вместе с тем отметим, что более высокое значение ОСШ позволяет получать более высокие значения для нормализованного индекса для одинаковых значений  $\alpha$  в случае, когда фазовая связность есть, оставляя индекс на том же уровне или ниже, когда ее нет. Похожий эффект ослабления значений меры коннективности при понижении ОСШ отмечали авторы метода wPLI, [27].

Индекс, полученный методом PSIICOS Normalized, теоретически может быть использован и для локализации сетей, однако такая нормировка приводит к появлению пространственной смещенностии оценки (рассуждения аналогичны изложенным в разд. 2.2.2), поэтому для использования на практике разумнее сначала локализовать потенциальные сети одним из методов (PSIICOS Unbiased или GO-PSIICOS) и затем использовать PSIICOS Normalized для удаления ложноположительных срабатываний из-за мощности.

### 3.10 Демонстрация работы метода GO-PSIICOS на симуляционных данных.

В этом разделе мы продемонстрировали работу метода GO-PSIICOS, который решает задачу поиска сетей методом глобальной оптимизации и обладает свойством разреженности решений по пространству, что позволяет подавить “призрачные взаимодействия” в решении.

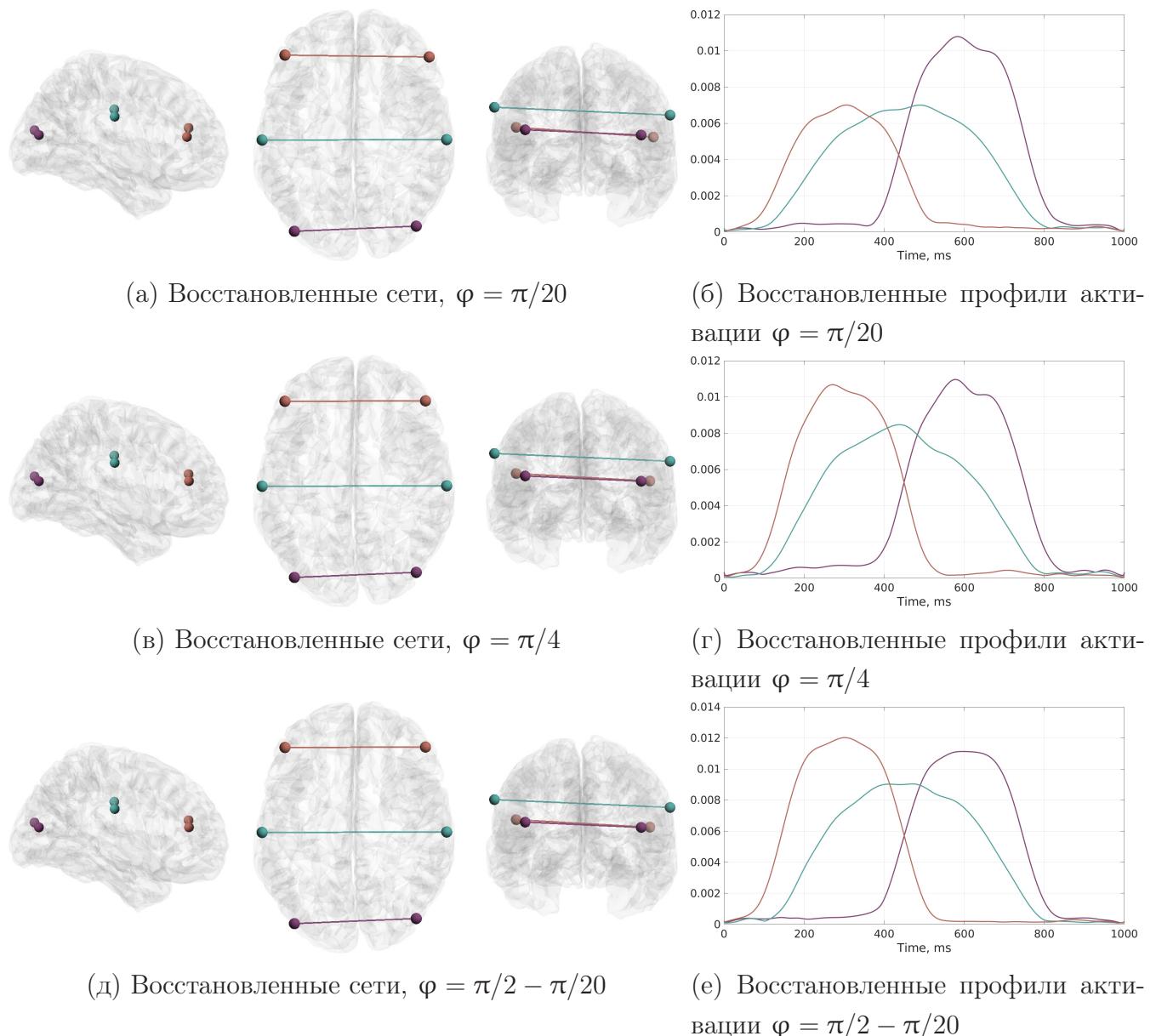


Рисунок 3.20: Применение метода GO-PSIICOS к задаче детекции трех сетей с перекрывающимися профилями активности для ОСШ=0.5 и фазовых сдвигов  $\varphi = \pi/20, \varphi = \pi/4, \varphi = \pi/2 - \pi/20$ .

Цвета сетей на панели 3.20а соответствуют цветам восстановленных профилей на панели 3.20б.

Для демонстрации работы метода мы использовали симуляции с тремя сетями с перекрывающимися профилями активности, описанные в разд. 3.3. Результаты работы метода для ОСШ=0.5 представлены на рис. 3.20.

Как видно на графике, алгоритм справился с нахождением всех трех сетей и достаточно точно восстановил их временные профили.

## Заключение

Основные результаты работы заключаются в следующем.

1. Был проведен обзор исследований изменения функциональной коннективности мозга при различных патологиях.
2. Был разработан метод очистки данных ЭЭГ и МЭГ от протечки сигнала, на основе которого было разработано семейство алгоритмов оценки фазовой синхронности, позволяющих находить сети с близкими к нулю фазовыми задержками.
3. Задача оценки фазовой синхронности в условиях протечки сигнала была сформулирована и решена как задача оптимальной фильтрации.
4. Был предложен алгоритм, позволяющий обнаруживать сети с близкими к нулю фазовыми задержками, оптимальный в глобальном смысле и позволяющий справиться с проблемой ложноположительных срабатываний второго рода, вызванных протечкой сигнала.
5. Было проведено численное исследование свойств предложенной проекции, показавшее, что разработанная методика позволяет подавить вклад подпространства протечки сигнала в оцененную на сенсорах матрицу кросс-спектральной плотности мощности.
6. Численное исследование свойств метода проекции показало, что разработанная методика позволяет находить сети с близкими к нулю фазовыми задержками в условиях неинвазивных МЭГ измерений, которые характеризуются значительной протечкой сигнала между источниками.
7. Сравнение с имеющимися на данный момент алгоритмами оценки коннективности по неинвазивным данным на основе симуляций показало значительное превосходство предложенной техники обнаружения сетей в условиях малых фазовых задержек.
8. Применение метода очистки от протечки сигнала к реальным данным позволило обнаружить физиологически правдоподобные сети, которые невозможно обнаружить другими способами.
9. Было проведено численное исследование влияния значений ранга предложенной проекции на свойства алгоритма, которое позволило получить эвристику для выбора ранга.

10. Численное исследование влияния неточностей прямой модели на качество решений предложенного алгоритма показало, что характерные для реальных записей диапазоны ошибок в оценке прямой модели слабо сказываются на качестве получаемых решений.
11. Для выполнения поставленных задач был создан пакет утилит в среде MATLAB, в который входят средства генерации тестовых данных, визуализации пространственной и временной структуры сетей и наконец программные реализации разработанных и использованных для валидации алгоритмов.
12. Наработки, полученные в ходе работы над данной диссертацией, были внедрены в пакеты программ Visbrain и Neuropycon, доступные для публичного использования.

## Список сокращений и условных обозначений

<b>ВЕМ</b>	boundary element method, метод граничных элементов
<b>ЦНС</b>	центральная нервная система
<b>ЭЭГ</b>	электроэнцефалография
<b>МЭГ</b>	магнитная электроэнцефалография
<b>МРТ</b>	магнитно-резонансная томография
<b>фМРТ</b>	функциональная магнитно-резонансная томография
<b>КТ</b>	компьютерная томография
<b>ДТВ</b>	диффузионно-тензорная визуализация
<b>ПЭТ</b>	позитронно-эмиссионная томография
<b>СТС</b>	communication through coherence; взаимодействие через когерентность
<b>PLI</b>	phase lag index; индекс фазовой задержки
<b>wPLI</b>	weighted phase lag index; взвешенный индекс фазовой задержки
<b>ОСШ</b>	отношение сигнал / шум
<b>GCS</b>	geometric correction scheme
<b>DICS</b>	dynamic imaging of coherent sources
<b>MNE</b>	minimum norm estimate
<b>MCE</b>	minimum current estimate
<b>ISTA</b>	iterative shrinkage thresholding algorithm
<b>MxNE</b>	mixed-norm estimate
<b>irMxNE</b>	iterative reweighted mixed-norm estimate
<b>PSF</b>	point spread function
<b>RK</b>	resolution kernel
<b>BR</b>	beam response
<b>ROC</b>	Receiver Operating Characteristics
<b>PR</b>	Precision-Recall
<b>БИХ-фильтр</b>	фильтр с бесконечной импульсной характеристикой
<b>КИХ</b>	фильтр с конечной импульсной характеристикой

## Список литературы

1. *McCulloch W. S.* A logical calculus of the ideas immanent in nervous activity / W. S. McCulloch, W. Pitts // The bulletin of mathematical biophysics. — 1943. — Т. 5. — С. 115—133.
2. *Buzsaki G.* Rhythms of the brain / Buzsaki G. — New York : Oxford Univ. Press, 2006.
3. *Fries P.* Rhythms for Cognition : Communication through Coherence / P. Fries. — 2015.
4. The brainweb: phase synchronization and large-scale integration / F. Varela [и др.] // Nature Reviews Neurosci. — 2001. — Т. 2, April. — С. 229—39. — URL: [http://www.nature.com/nrn/journal/v2/n4/full/nrn0401\\_229a.html](http://www.nature.com/nrn/journal/v2/n4/full/nrn0401_229a.html).
5. Fast transient networks in spontaneous human brain activity / A. P. Baker [и др.] // eLIFE. — 2014. — Т. 3. — URL: <http://elifesciences.org/content/elife/3/e01867.full.pdf>.
6. Inferring spatiotemporal network patterns from intracranial EEG data / A. Ossadtchi [и др.] // Clinical Neurophysiology. — 2010. — Т. 121, № 6. — С. 823—835. — URL: <http://www.sciencedirect.com/science/article/pii/S1388245710000350>.
7. Direct Recordings from Human Anterior Insula Reveal its Leading Role within the Error-Monitoring Network / J. Bastin [и др.]. — 2017.
8. Alterations of Intrinsic Brain Connectivity Patterns in Depression and Bipolar Disorders: A Critical Assessment of Magnetoencephalography-Based Evidence / G. Alamian [и др.] // Frontiers in Psychiatry. — 2017. — Т. 8, March. — С. 1—17. — URL: <http://journal.frontiersin.org/article/10.3389/fpsyg.2017.00041/full>.
9. Measuring alterations in oscillatory brain networks in schizophrenia with resting-state MEG: State-of-the-art and methodological challenges / G. Alamian [и др.] // Clinical Neurophysiology. — 2017. — Т. 128, № 9. — С. 1719—1736. — URL: <http://dx.doi.org/10.1016/j.clinph.2017.06.246>.

10. Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain / M. Hamalainen [и др.] // Reviews on modern physics. — 1993. — Т. 65, № 2. — URL: <http://journals.aps.org/rmp/abstract/10.1103/RevModPhys.65.413>.
11. *Baillet S.* Electromagnetic Brain Mapping / S. Baillet, J. Mosher, R. Leahy // IEEE Signal Processing Magazine. — 2001. — Нояб. — Т. 18, № 1. — С. 14—30. — URL: [http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=962275&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs\\_all.jsp%3Farnumber%3D962275](http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=962275&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D962275).
12. Good practice for conducting and reporting MEG research / J. Gross [и др.] // NeuroImage. — 2013. — Т. 65. — С. 349—363. — URL: <http://dx.doi.org/10.1016/j.neuroimage.2012.10.001>.
13. *Marzetti L.* Understanding brain connectivity from EEG data by identifying systems composed of interacting sources / L. Marzetti, C. Del, G. Nolte. — 2008.
14. *Schoffelen J.-m.* Source Connectivity Analysis With MEG and EEG / J.-m. Schoffelen, J. Gross. — 2009.
15. A symmetric multivariate leakage correction for MEG connectomes / G. L. Colclough [и др.] // NeuroImage. — 2015. — Т. 117. — С. 439—448. — URL: <http://dx.doi.org/10.1016/j.neuroimage.2015.03.071>.
16. *Kaminski M.* Directed Transfer Function is not influenced by volume conduction - inexpedient pre-processing should be avoided / M. Kaminski, K. J. Blinowska // Frontiers in Computational Neuroscience. — 2014. — Т. 8, № 61. — URL: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4050361/>.
17. *Greenblatt R. E.* Connectivity measures applied to human brain electrophysiological data / R. E. Greenblatt, M. E. Pfleger, A. E. Ossadtchi // Journal of Neuroscience Methods. — 2012. — Т. 207, № 1. — С. 1—16. — URL: <http://dx.doi.org/10.1016/j.jneumeth.2012.02.025>.
18. *Hillebrand A.* Recent developments in MEG network analysis / A. Hillebrand, C. J. Stam // Magnetoencephalography. From Signals to Dynamic Cortical Networks. — 2014. — С. 263—277. — URL: [http://link.springer.com/chapter/10.1007%2F978-3-642-33045-2\\_12](http://link.springer.com/chapter/10.1007%2F978-3-642-33045-2_12).

19. Identifying true brain interaction from EEG data using the imaginary part of coherency / G. Nolte [и др.] // Clinical Neurophysiology. — 2004. — Т. 115, № 10. — С. 2292–2307. — URL: <http://www.ncbi.nlm.nih.gov/pubmed/15351371>.
20. Measuring phase synchrony in brain signals / J.-P. Lachaux [и др.] // Human Brain Mapping. — 1999. — Т. 8, № 4. — С. 194–208. — URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0032757697&partnerID=40&md5=c404423056859241c3cb46401cafabb9> ; cited By (since 1996)852.
21. Measuring electrophysiological connectivity by power envelope correlation : a technical review on MEG methods / G. C. O'Neill [и др.]. —
22. *Brookes M. J.* Measuring functional connectivity in MEG : A multivariate approach insensitive to linear source leakage / M. J. Brookes, M. W. Woolrich, G. R. Barnes // NeuroImage. — 2012. — Т. 63, № 2. — С. 910–920. — URL: <http://dx.doi.org/10.1016/j.neuroimage.2012.03.048>.
23. Investigating the electrophysiological basis of resting state networks using magnetoencephalography / M. J. Brookes [и др.]. — 2011.
24. Frequency-dependent functional connectivity within resting-state networks : An atlas-based MEG beamformer solution / A. Hillebrand [и др.] // NeuroImage. — 2012. — Т. 59, № 4. — С. 3909–3921. — URL: <http://dx.doi.org/10.1016/j.neuroimage.2011.11.005>.
25. Large-scale cortical correlation structure of spontaneous oscillatory activity / J. F. Hipp [и др.] // Nature Neuroscience. — 2012. — Т. 15, № 6. — С. 884–890. — arXiv: [NIHMS150003](https://arxiv.org/abs/1204.1500).
26. *Stam C.* Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources / C. Stam, G. Nolte, A. Daffertshofer // Hum Brain Mapp. — 2007. — Т. 28, № 11. — С. 1178–93. — URL: <http://www.ncbi.nlm.nih.gov/pubmed/17266107>.
27. An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias / M. Vinck [и др.] // Neuroimage. — 2011. — Т. 55, № 4. — С. 1548–65. — URL: <http://www.ncbi.nlm.nih.gov/pubmed/21276857>.
28. Third order spectral analysis robust to mixing artifacts for mapping cross-frequency interactions in EEG/MEG / F. Chella [и др.]. — 2015.

29. Bispectral pairwise interacting source analysis for identifying systems of cross-frequency interacting brain sources from electroencephalographic or magnetoencephalographic signals / F. Chella [и др.]. — 2016.
30. Transfer entropy in magnetoencephalographic data: Quantifying information flow in cortical and cerebellar networks / M. Wibral [и др.] // Progress in Biophysics and Molecular Biology. — 2011. — Т. 105, № 1/2. — С. 80—97. — URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-79551689714&partnerID=40&md5=31bfe00660c19cb98378965b389b5a0c> ; cited By (since 1996)26.
31. Coupling of Regional Activations in a Human Brain During an Object and Face Affect Recognition Task / A. A. Ioannides [и др.]. — 2000.
32. How reliable are MEG resting-state connectivity metrics ? / G. L. Colclough [и др.] // NeuroImage. — 2016. — Т. 138. — С. 284—293. — URL: <http://dx.doi.org/10.1016/j.neuroimage.2016.05.070>.
33. *Bastos A. M.* A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls / A. M. Bastos, J.-m. Schoffelen. — 2016.
34. Visuomotor integration is associated with zero time-lag synchronization among cortical areas / P. Roelfsema [и др.] // Nature. — 1997. — Т. 385, № 6612. — С. 157—161. — URL: <http://www.ncbi.nlm.nih.gov/pubmed/8990118>.
35. *Singer W.* Temporal Coherence: A Versatile Code for the Definition of Relations / W. Singer // The Senses: A Comprehensive Reference. — 1999. — Т. 2. — С. 1—9.
36. *Engel A. K.* Dynamic predictions: oscillations and synchrony in top-down processing / A. K. Engel, P. Fries, W. Singer // Nat. Rev. Neurosci. — 2001. — Т. 2001, № 2. — С. 704—716.
37. *Rajagovindan R.* Decomposing Neural Synchrony: Toward an Explanation for Near-Zero Phase-Lag in Cortical Oscillatory Networks / R. Rajagovindan, M. Ding // PLoS One. — 2008. — Т. 3, № 11. — e3649. — URL: <http://www.ncbi.nlm.nih.gov/pubmed/18987745>.
38. A Geometric Correction Scheme for Spatial Leakage Effects in MEG/EEG Seed-Based Functional Connectivity Mapping / V. Wens [и др.]. — 2015.

39. Magnetoencephalography - Theory, instrumentation, and applications to noninvasive studies of the working human brain / M. Hamalainen [и др.] // Reviews of Modern Physics. — 1993. — Т. 65, № 2. — С. 413—497. — URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-26244462082&partnerID=40&md5=1a3b2aeaf6cd3c1434f29dc7b19dd159> ; cited By (since 1996)2011.
40. *Hämäläinen M. S.* Interpreting magnetic fields of the brain: minimum norm estimates / M. S. Hämäläinen, R. J. Ilmoniemi // Medical & Biological Engineering & Computing. — 1994. — Т. 32, № 1. — С. 35—42.
41. *Uutela K.* Visualization of magnetoencephalographic data using minimum current estimates / K. Uutela, M. S. Hamalainen, E. Somersalo // NeuroImage. — 1999. — Т. 10. — С. 173—180.
42. *Pascual-Marqui R.* Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain / R. Pascual-Marqui, C. Michel, D. Lehmann // International Journal of Psychophysiology. — 1994. — Т. 18, № 1. — С. 49—65. — URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-0027998259&partnerID=40&md5=e97677b75abc9f64e5f1be79b0d30c02> ; cited By (since 1996)1300.
43. *Pascual-Marqui R. D.* Standardized low resolution brain electromagnetic tomography (sLORETA): technical details / R. D. Pascual-Marqui // Methods & Findings in Experimental & Clinical Pharmacology. — 2002. — С. 1—16.
44. *Тихонов А. Н.* О некорректных задачах линейной алгебры и устойчивом методе их решения / А. Н. Тихонов // Докл. АН СССР. — 1965. — Т. 163, вып. 3. — С. 591—594.
45. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain / F. A. Azevedo [и др.] // Journal of Comparative Neurology. — 2009. — Т. 513, № 5. — С. 532—541.
46. *Lu W.* The Wiener-Khinchin Theorem for Non-wide Sense stationary Random Processes / W. Lu, N. Vaswani. — 2009. — arXiv: 0904.0602. — URL: <http://arxiv.org/abs/0904.0602>.
47. *Saladin K.* Human Anatomy / K. Saladin. — McGraw-Hill, 2011. — URL: <https://books.google.ru/books?id=DfESQgAACAAJ>.

48. *Владимиров В.* Уравнения математической физики / В. Владимиров. — Издательство "Наука", 1981.
49. *Mosher J. C.* Multiple Dipole Modeling and Localization from Spatio-Temporal MEG Data / J. C. Mosher, P. S. Lewis, R. M. Leahy // IEEE Transactions on Biomedical Engineering. — 1992. — Т. 39, № 6. — С. 541—557.
50. *Schmidt R. O.* Multiple emitter location and signal parameter estimation / R. O. Schmidt // Adaptive Antennas for Wireless Communications. — 1986. — № 3. — С. 190—194.
51. *Mosher J.* Source localization using recursively applied and projected (RAP) MUSIC / J. Mosher, R. Leahy // Signal Processing, IEEE Transactions on. — 1999. — Февр. — Т. 47, № 2. — С. 332—340.
52. *Gross J.* Linear transformations of data space in MEG / J. Gross, A. A. Ioannides // Physics in Medicine & Biology. — 1999. — Т. 44, № 2081.
53. Dynamic Statistical Parametric Mapping / A. M. Dale [и др.] // Neuron. — 2000. — Т. 26, № 1. — С. 55—67. — URL: <http://linkinghub.elsevier.com/retrieve/pii/S0896627300811381>.
54. *Tibshirani R.* Regression Selection and Shrinkage via the Lasso / R. Tibshirani. — 1996. — arXiv: [11/73273 \[1369-7412\]](https://arxiv.org/abs/11/73273). — URL: <http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.7574>.
55. Dynamic imaging of coherent sources: Studying neural interactions in the human brain. / J. Gross [и др.] // Proceedings of the National Academy of Sciences of the United States of America. — 2001. — Т. 98, № 2. — С. 694—9.
56. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. / B. D. Van Veen [и др.] // IEEE Transactions on Biomedical Engineering. — 1997. — Т. 44, № 9. — С. 867—880. — arXiv: [10.623056 \[10.1109\]](https://arxiv.org/abs/10.623056).
57. *Oppenheim A. V.* Discrete Time Signal Processing / A. V. Oppenheim, R. W. Schafer. — 3-rd. — Prentice-Hall, Inc., 1998. — С. 897.
58. *Freeman W. J.* Imaging Brain Function With EEG / W. J. Freeman, R. Q. Quiroga. — Springer New York, 2013.
59. *Magnus J. R.* Matrix Differential Calculus with Applications in Statistics and Econometrics / J. R. Magnus, H. Neudecker. — Second. — John Wiley, 1999.

60. *Golub G. H.* Matrix Computations (3rd Ed.) / G. H. Golub, C. F. Van Loan. — USA : Johns Hopkins University Press, 1996.
61. Distributed current estimates using cortical orientation constraints / F. H. Lin [и др.] // Human Brain Mapping. — 2006. — T. 27, № 1. — C. 1—13.
62. *Sekihara K.* Adaptive spatial filters for electromagnetic brain imaging / K. Sekihara, S. S. Nagarajan ; под ред. J. H. Nagel. — Berlin : Springer Berlin Heidelberg, 2007.
63. Ghost interactions in MEG/EEG source space: A note of caution on inter-areal coupling measures / J. M. Palva [и др.] // NeuroImage. — 2018.
64. *Selesnick I.* Sparse Signal Restoration / I. Selesnick // Connexions. — 2009. — C. 1—13.
65. *Gramfort A.* Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods / A. Gramfort, M. Kowalski, M. S. Hamalainen // Physics in Medicine and Biology. — 2012. — T. 57, № 7. — C. 1—7.
66. *Combettes P. L.* Signal recovery by proximal forward-backward splitting / P. L. Combettes, V. R. Wajs // Multiscale Modeling and Simulation. — 2005. — T. 4, № 4. — C. 1168—1200.
67. *Boyd S.* Convex Optimization / S. Boyd, L. Vandenberghe. — Cambridge University Press, 2004. — C. 487. — eprint: [1111.6189v1](https://arxiv.org/abs/1111.6189v1).
68. *Srohmeier D.* The iterative reweighted Mixed-Norm Estimate for spatio-temporal MEG/EEG source reconstruction / D. Srohmeier, J. Haueisen, A. Gramfort // IEEE Transactions on Medical Imaging. — 2014. — T. 35, № 10. — C. 2218—2228.
69. *Beck A.* A Fast Iterative Shrinkage-Thresholding Algorithm / A. Beck, M. Teboulle // Society for Industrial and Applied Mathematics Journal on Imaging Sciences. — 2009. — T. 2, № 1. — C. 183—202.
70. *Rakotomamonjy A.* Surveying and comparing simultaneous sparse approximation (or group-lasso) algorithms / A. Rakotomamonjy // Signal Processing. — 2011. — T. 91, № 7. — C. 1505—1526. — URL: <http://dx.doi.org/10.1016/j.sigpro.2011.01.012>.

71. *Qin Z.* Efficient block-coordinate descent algorithms for the Group Lasso / Z. Qin, K. Scheinberg, D. Goldfarb // Mathematical Programming Computation. — 2013. — Т. 5, № 2. — С. 143—169.
72. *Candès E. J.* Enhancing sparsity by reweighted l1 minimization / E. J. Candès, M. B. Wakin, S. P. Boyd // Journal of Fourier Analysis and Applications. — 2008. — Т. 14, № 5/6. — С. 877—905.
73. *Gasso G.* Recovering sparse signals with a certain family of nonconvex penalties and DC programming / G. Gasso, A. Rakotomamonjy, S. Canu // IEEE Transactions on Signal Processing. — 2009. — Т. 57, № 12. — С. 4686—4698.
74. Investigations of dipole localization accuracy in MEG using the bootstrap / F. Darvas [и др.] // NeuroImage. — 2005. — Т. 25, № 2. — С. 355—368. — URL: <http://www.sciencedirect.com/science/article/pii/S1053811904005865>.
75. Connectivity differences in brain networks / A. Zalesky [и др.] // NeuroImage. — 2012. — Т. 60, № 2. — С. 1055—1062. — URL: <http://dx.doi.org/10.1016/j.neuroimage.2012.01.068>.
76. Interactions between posterior gamma and frontal alpha/beta oscillations during imagined actions / F. P. de Lange [и др.] // Frontiers in Human Neuroscience. — 2008. — Т. 2, August. — С. 1—12. — URL: <http://journal.frontiersin.org/article/10.3389/neuro.09.007.2008/abstract>.
77. *Hanakawa T.* Motor Planning, Imagery, and Execution in the Distributed Motor Network: A Time-Course Study with Functional MRI / T. Hanakawa, M. A. Dimyan, M. Hallett // Cerebral Cortex. — 2008. — Т. 18, № 12. — С. 2775—2788. — URL: <http://dx.doi.org/10.1093/cercor/bhn036>.
78. The functional role of dorso-lateral premotor cortex during mental rotation: An event-related fMRI study separating cognitive processing steps using a novel task paradigm / C. Lamm [и др.] // NeuroImage. — 2007. — Т. 36, № 4. — С. 1374—1386. — URL: <http://www.sciencedirect.com/science/article/pii/S1053811907003205>.
79. *Striem-Amit E.* Sensorimotor-independent development of hands and tools selectivity in the visual cortex / E. Striem-Amit, G. Vannuscorps, A. Caramazza // Proceedings of the National Academy of Sciences. — 2017. — Т. 114, № 18. — С. 4787—4792. — URL: <http://www.pnas.org/content/114/18/4787.abstract>.

80. *Dupont S.* Investigating temporal pole function by functional imaging. / S. Dupont // Epileptic Disord. — 2002. — T. 2002, № 9. — S17—22.

## Список рисунков

2.1 Пример сетки в пространстве источников, построенной на основе трехмерной анатомической модели . . . . .	75
2.2 Взаимосвязь подпространств для кросс-спектра на уровне сенсоров .	82
3.1 Сравнение Precision-Recall (панели (а), (с)) и ROC- (панели (б), (д)) кривых в задаче локализации сетей для методов PSIICOS, DICS, iDICS и GCS MNE для двух значений ОСШ на основе 1000 Монте-Карло итераций. . . . .	122
3.2 Три симулированные пары взаимодействующих источников. . . . .	124
3.3 Пространственное распределение сетевых источников для симуляционных данных. . . . .	125
3.4 Временные профили активации симулированных сетей, восстановленные при помощи PSIICOS, для двух значений фазового сдвига и двух значений ОСШ. . . . .	126
3.5 Сравнение PSIICOS с остальными методами в задаче обнаружения трёх одновременно активных сетей. . . . .	127
3.6 Исследование стабильности метода PSIICOS при помощи процедуры бутстрэпа для двух значений фазового сдвига и двух ОСШ. . . . .	128
3.7 Качество решений алгоритма PSIICOS как функция средней фазовой задержки . . . . .	129
3.8 Полосы частот с наиболее воспроизводимыми сетями, определенными при помощи процедуры бутстрэпа. . . . .	132
3.9 Пространственная структура и временная динамика наиболее воспроизводимых сетей в диапазонах частот тета (3–6 Гц) и альфа (8–12 Гц) . . . . .	133
3.10 Пространственная структура и временная динамика наиболее воспроизводимых сетей в бета-диапазоне (16–24 Гц) . . . . .	134
3.11 Пространственная структура и временная динамика наиболее воспроизводимых сетей в нижнем (30–60 Гц) и верхнем (65–85 Гц) гамма-диапазонах . . . . .	135

3.12 Нормы 2-топографий для трех подпространств кросс-спектра до и после проекции в зависимости от коэффициента корреляции топографий взаимодействующих узлов на коре (графики (а) и (б)), а также от расстояния между этими узлами (графики (с) и (д)). . . . .	137
3.13 (а) Ослабление мощности в подпространствах $S_{SL}$ и $S_{Re}$ и (б) Разница в коэффициенте затухания для подпространств $S_{Re}$ и $S_{SL}$ как функции ранга проекции для трех различных размерностей пространства виртуальных сенсоров (VS). . . . .	139
3.14 Влияние шума в прямой модели на качество проекции. . . . .	141
3.15 Сравнение методов PSICOS и PSICOS Unbiased в задаче детекции сетей со случайными положениями узлов для 100 итераций Монте-Карло. . . . .	143
3.16 Пример плохого решения для алгоритма PSICOS по сравнению с PSICOS Unbiased в задаче детекции сетей со случайными положениями узлов для одной из итераций Монте-Карло и фазового сдвига $\varphi = \pi/20$ . . . . .	143
3.17 Сравнение методов PSICOS и PSICOS Unbiased в задаче детекции трех сетей с фиксированными положениями источников и перекрывающимися временными профилями активации, усредненные по 10 повторениям. . . . .	144
3.18 Локализация сети с фиксированными положениями узлов методом PSICOS Unbiased для фазового сдвига $\varphi = \pi/4$ для высокого и низкого ОСШ (0.2 и 1) в случае сильной связности ( $\alpha = 0.1$ ) и отсутствия связности ( $\alpha = 2$ ). . . . .	145
3.19 Нормализованный индекс фазовой связности, полученный методом PSICOS Normalized для одной симулированной сети в условиях высокого и низкого ОСШ в зависимости от индекса истинной фазовой связности $\alpha$ и значения фазовой задержки. . . . .	147
3.20 Применение метода GO-PSICOS к задаче детекции трех сетей с перекрывающимися профилями активности для ОСШ=0.5 и фазовых сдвигов $\varphi = \pi/20, \varphi = \pi/4, \varphi = \pi/2 - \pi/20$ . . . . .	148
Б.1 Распределение индуцированной активности в тета (3–6 Гц) и альфа (8–12 Гц) диапазонах . . . . .	167

Б.2 Распределение индуцированной активности в бета-диапазоне (16–24 Гц) . . . . .	167
Б.3 Распределение индуцированной активности в нижнем (30–60 Гц) и верхнем (65–85 Гц) гамма-диапазонах. . . . .	167

## Приложение А

### Реализация алгоритма RAP-MUSIC на языке Python

Листинг А.1: Алгоритм RAP-MUSIC

```

def RAP_MUSIC(X, G, threshold):
    """
    Параметры
    -----
    X : матрица измерений
    G : матрица прямой модели для свободной ориентации
    threshold : порог корреляции подпространств

    Возвращаем
    -----
    active_dipole_indices : индексы найденных диполей
    """

# инициализируем пустой список индексов активных диполей
active_dipole_indices = []

while True:
    # ищем корреляции подпространств для каждого источника
    C = MUSIC_scan(X, G)
    if max(C) < threshold:
        break
    k = argmax(C)
    active_dipole_indices.append(k)
    # проецируем от диполей k-го источника
    X, G = project_away_from_k(X, G, k)

return active_dipole_indices

```

## Приложение Б

### Распределение индуцированной активности по коре в реальных данных

Чтобы подтвердить наши выводы и убедиться, что полученные сети не полностью совпадают с областями коры с доминирующей активностью мы представили здесь распределение индуцированной активности по коре для целевых частотных диапазонов, см. рис. [Б.1](#), [Б.2](#), [Б.3](#). Эти распределения были получены при помощи пакета MNE-python в соответствии со следующей процедурой.

Во-первых, чтобы компенсировать разницу в продолжительности эпох из-за вариабельности времени отклика, мы обрезали каждое испытание в [-0,5, 1,024] секундном интервале, который соответствовал временному диапазону самой короткой эпохи. Чтобы смягчить граничные эффекты фильтрации, мы обратили во времени каждую эпоху и “приклеили” ее слева и справа к исходной. Эти эпохи тройной длины были затем отфильтрованы в пяти целевых частотных диапазонах КИХ-фильтром с нулевой фазой и переобрезаны до временного интервала [0,1] секунд, где 0 соответствует моменту предъявления стимула в исходной неинвертированной эпохе. Далее для пространства источников с 8194 узлами мы вычислили обратный оператор MNE со свободной ориентацией (коэффициент свободной ориентации = 0,2) и взвешиванием по глубине (коэффициент = 0,6) и применили его к отдельным испытаниям. Наконец, мы возвели значения для эпох на уровне источников в квадрат и усреднили по времени и по эпохам. Полученные значения мы нанесли на графики.

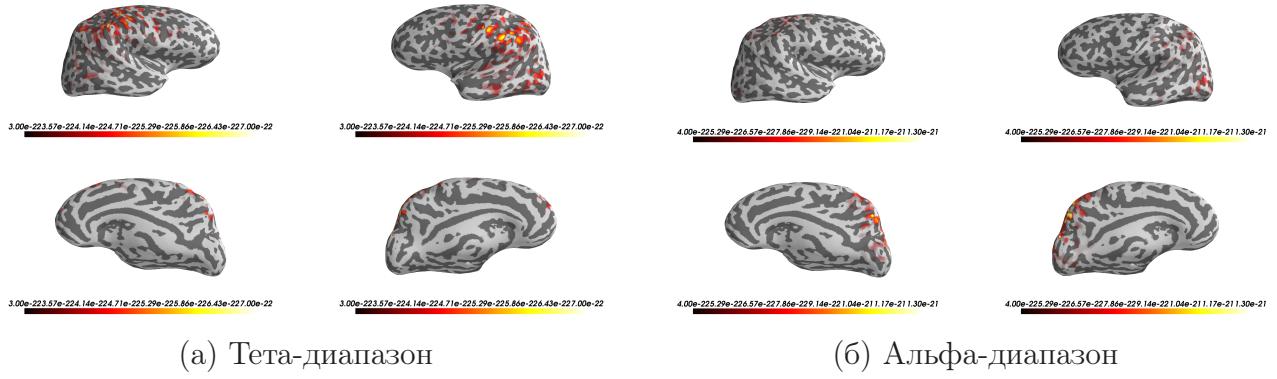


Рисунок Б.1: Распределение индуцированной активности в тета (3–6 Гц) и альфа (8–12 Гц) диапазонах

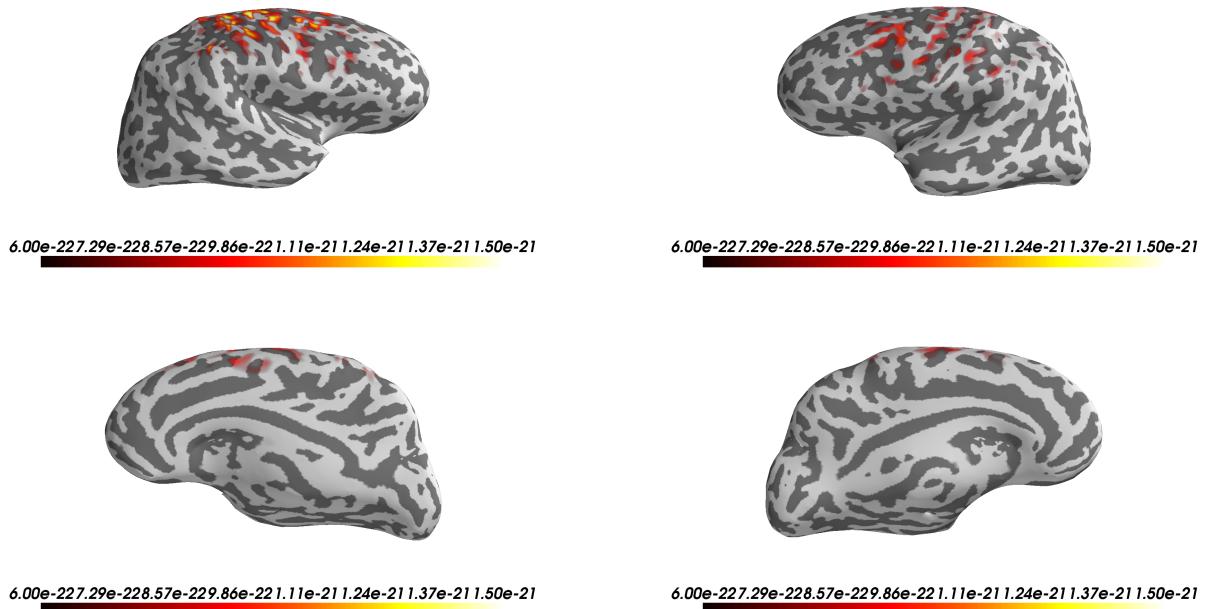


Рисунок Б.2: Распределение индуцированной активности в бета-диапазоне (16–24 Гц)

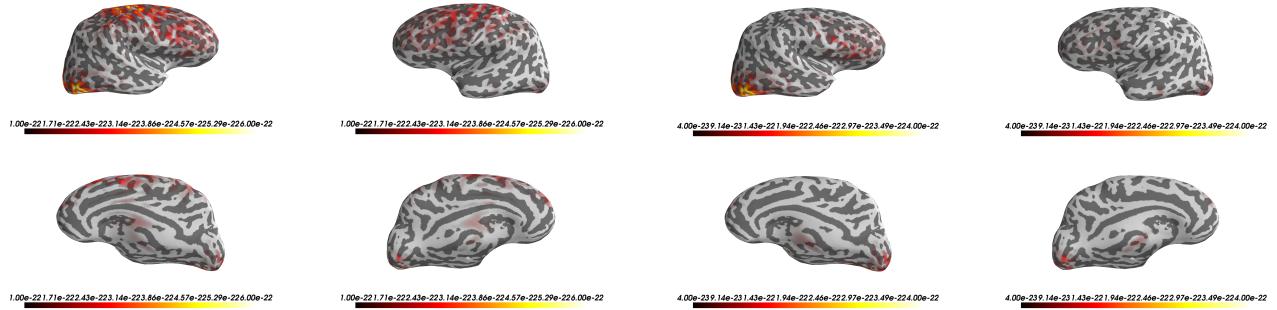


Рисунок Б.3: Распределение индуцированной активности в нижнем (30–60 Гц) и верхнем (65–85 Гц) гамма-диапазонах.