

Адаптивно регуляризованный псевдообратный префильтр для передачи сигнала в многоантенных системах радиосвязи

Е. А. Бобров¹, Д. С. Миненков², Д. А. Юдаков³

Современные сотовые сети используют технологию передачи сигнала с множеством антенн на стороне базовой станции и пользователя. В работе исследуется адаптивный префильтр, использующий особую регуляризацию на основе сингулярного разложения матриц. Проводится теоретический анализ, оценка производительности и сравнение с другими методами на симуляциях в модели канала "Quadriga".

Ключевые слова: Телекоммуникации, технология ММО, оптимизация, сингулярное разложение, отношение сигнал-интерференция-шум, спектральная эффективность

1. Введение

В системах радиосвязи с множеством передающих и принимающих антенн (multiple-input multiple-output, ММО) префильтр⁴ является важной частью обработки сигнала на нисходящей линии связи, поскольку позволяет фокусировать энергию передаваемого сигнала на меньших областях (именно там, где располагаются приемники), обеспечивая большую спектральную эффективность при меньшей передаваемой мощности [1, 2].

¹Бобров Евгений Александрович — к.ф.-м.н. (2025), аспирантура кафедры Математических Методов Прогнозирования факультета ВМК МГУ, e-mail: eugenbobrov@ya.ru.

Bobrov Evgeny Aleksandrovich — PhD, Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics, Department of Mathematical Methods of Forecasting.

²Миненков Дмитрий Сергеевич — н.с. мех.-мат. ф-та МГУ; старший преподаватель КНТ МФТИ, e-mail: minenkov.ds@gmail.com.

Minenkov Dmitry Sergeevich — research fellow, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics; senior teacher, Moscow Institute of Physics and Technology, Department of Mathematics and Mathematical Methods in Physics

³Юдаков Даниил Андреевич — аспирант каф. математической теории интеллектуальных систем мех.-мат. ф-та МГУ, e-mail: d.yudakov43@gmail.com.

Yudakov Daniil Andreevich — graduate student, Lomonosov Moscow State University, Faculty of Mechanics and Mathematics, Chair of Mathematical Theory of Intellectual Systems.

⁴Линейный префильтр — это линейное преобразование вектора передаваемых символов, производимое на передатчике. В англоязычной литературе встречается под разными названиями, в т.ч. Precoding, Beamforming, Transmitting filter, etc.

Различные линейные префильтры позволяют либо максимизировать подачу энергии к пользователю, как в случае передачи с сопряжённым префильтром (далее – СФ, в англоязычной литературе известен как Maximum Ratio Transmission, MRT), либо полностью устранять межпользовательские помехи ⁵, как в методе псевдообратного префильтра (далее – ПФ, в англоязычной литературе – Zero-Forcing, ZF) [3].

В случае алгоритмов регуляризованного псевдообратного префильтра (РПФ, Regularized Zero-Forcing, RZF) достигается баланс между максимизацией мощности сигнала и минимизацией межпользовательской интерференции [4, 5, 6, 7, 8], при этом сохраняется относительно низкая вычислительная сложность по сравнению с нелинейными префильтрами. Обычно рассматривается “скалярная” регуляризация, то есть регуляризация с помощью скалярного множителя при единичной матрице соответствующего размера. В данной работе исследуется явная эвристическая формула для диагональной регуляризации, которая даёт лучшие результаты по сравнению со скалярным РПФ при той же асимптотической сложности алгоритма. Исследование подкреплено теоретическим обоснованием и тестированием в симуляторе (программном пакете) “Quadriga” [9].

Существуют и другие подходы построения линейного префильтра, такие как префильтр с учётом постфильтра (detection-aware precoding) [10], блочно-диагональный префильтр [11], а также префильтры с распределённой мощностью [12]. Также существуют различные нелинейные методы построения префильтра, такие как кодирование с учётом известного помехового сигнала (Dirty Paper Coding, DPC) и векторная перестановка (Vector Perturbation, VP), однако они значительно сложнее в реализации [13], и при использовании множества антенн предпочтение отдаётся линейным методам префильтра.

Существуют качественные обзоры методов префильтра [14, 15], а также статьи, рассматривающие различные аспекты и модификации этих алгоритмов (см., например, [16, 17, 18]). В частности, в работе [18] представлены различные варианты РПФ в случае нескольких базовых станций, позволяющие уменьшить межсотовые помехи.

Большинство работ, для упрощения анализа, не уделяют должного внимания пользователям с несколькими антеннами. В нашей работе рассматривается одна базовая станция (с большим количеством передающих антенн), обслуживающая пользователей также с несколькими антеннами каждый, при этом каждому пользователю передаётся меньшее число каналов данных, чем число антенн. Это объясняется тем, что на

⁵Помехи на приемнике, возникающие в результате одновременной передачи сигналов разным пользователям, в литературе часто называется “интерференцией”. Составляет значительную (часто БОльшую) часть всех помех.

практике каналы между различными антеннами одного пользователя часто пространственно коррелированы [19]. В результате матрицы каналов пользователя плохо обусловлены (или даже вырождены), и невозможно эффективно передавать данные с помощью максимального числа потоков.

Для решения этой проблемы вместо полной матрицы канала пользователя в префильтре можно использовать векторы из её сингулярного разложения, соответствующие наибольшим сингулярным значениям [20]. В случае пользователя с одной антенной канал можно нормализовать [6], и коэффициенты нормализации соответствуют потерям мощности сигнала на пути от передатчика к приемнику (pathloss, PL), которые могут отличаться на несколько порядков: типичные значения мощностей полученного сигнала (Reference Signal Received Power, RSRP) варьируются от -130 дБм до -70 дБм. При использовании сингулярного разложения, сингулярные значения разных пользователей имеют тот же порядок, что и соответствующие потери мощности на пути, и также сильно варьируются.

Нам удалось найти простую эвристическую формулу, которая превосходит известные алгоритмы РПФ. Алгоритмы типа РПФ [4] используют скалярную регуляризацию вида λI . В работе Э. Бьёрнсона [5] доказано из общих соображений, что максимизация функции отношения сигнал-интерференция-шум (ОСИШ⁶, Signal-to-Noise-plus-Interference Ratio, SINR), включая рассматриваемую суммарную спектральную эффективность, достигается при использовании алгоритма с подходящей диагональной матрицей регуляризации. Нгуен и Ле-Нгок [21] выводят формулу адаптивного РПФ (АРПФ) для системы с одноантенными приемниками (multi-user single-input multiple-output, MU-SIMO) с одноантенными пользователями. Алгоритм АРПФ эффективно использует разные параметры регуляризации для разных пользователей, учитывая сингулярные значения передаваемых слоёв (потоков). Мотивированные этим, мы исследуем префильтр АРПФ с диагональной регуляризацией, что соответствует общей теореме из [5] и обобщает результаты [21] на случай многоантенных приемников, в котором необходимо учитывать алгоритмы постфильтра⁷ на стороне принимающих станций. Основная сложность такой постановки задачи заключается в том, что постфильтры реализуются каждой принимающей станцией из своих соображений (т.е. заведомо несогласованно) и более того не известны для передающей станции.

⁶ОСИШ отличается от обычного отношения сигнал-шум (ОСШ) тем, что часть помех, которые связаны с интерференцией и которые можно минимизировать на передатчике, вынесены в отдельное слагаемое: ОСИШ = сигнал (дБ) - (интерференция (дБ) + остальной шум (дБ)).

⁷Постфильтр – это обработка сигнала на стороне приемника, позволяющая отфильтровать шум и усилить сигнал в соответствии с используемой схемой модуляции и кодирования.

Результатом данной работы является адаптация формулы \mathbf{W}_{ARZF} для многопользовательских многоантенных систем (multi-user multiple-input multiple-output, MU-MIMO) систем, а также подробное исследование свойств префильтра АРПФ, в том числе доказательство его асимптотической оптимальности в случае многоантенных приемников.

Статья организована следующим образом. В разделе 2.1 формулируется задача, включающая модель канала и системы, показатели качества и ограничения на мощность. Модель нисходящего канала с множеством антенн упрощается с помощью сингулярного разложения матрицы канала (раздел 2.1.1) и упрощенной модели постфильтра, асимптотически приближающей стандартные алгоритмы (раздел 2.1.3); показатели качества и ограничения по мощности обсуждаются в разделах 2.1.5 и 2.1.6. В разделе 2.2 определяются известные стандартные префильтры (СФ, ПФ, РПФ), в разделе 2.3 дается определение предлагаемому алгоритму АРПФ и исследуются его свойства и асимптотическая связь со стандартными алгоритмами. Сравнение алгоритмов на численных экспериментах в симуляторе “Quadriga” представлено в разделе 3.2. Заключение дано в разделе 4. Условные обозначения приведены в таблице 1.

В статье используются следующие обозначения. Рассматривается одна базовая станция с K пользователями, число передающих антенн — T , число приёмных антенн и передаваемых символов у пользователя k обозначаются как $R_k = 1, 2, 4$ и $L_k \leq R_k$, соответственно. Общее число приёмных антенн: $R = \sum_{k=1}^K R_k$, а число передаваемых символов: $L = \sum_{k=1}^K L_k$, при этом $L_k \leq R_k \leq T$. Жирными строчными буквами обозначаются векторы (строки или столбцы). Матрицы обозначаются жирными заглавными буквами и трактуются как совокупности векторов, например, матрица канала — это набор векторов-строк: $\mathbf{H} = [\mathbf{H}_1; \dots; \mathbf{H}_R] \in \mathbb{C}^{R \times T}$, а матрица префильтра — набор векторов-столбцов: $\mathbf{W} = (\mathbf{W}_1, \dots, \mathbf{W}_L) \in \mathbb{C}^{T \times L}$. Элементы матриц обозначаются обычными строчными буквами с двумя индексами: первый индекс — строка, второй — столбец: $\mathbf{H} = \{h_{rt}\}$, $\mathbf{W} = \{w_{tl}\}$, $r = 1, \dots, R$, $t = 1, \dots, T$, $l = 1, \dots, L$. Эрмитово сопряжение обозначается как $\mathbf{H}^H := \overline{\mathbf{H}}^T$. Диагональные и блочно-диагональные матрицы записываются как $\mathbf{S}_k = \text{diag}\{s_{k,1}, \dots, s_{k,R_k}\}$ и $\mathbf{S} = \text{bdiag}\{\mathbf{S}_1, \dots, \mathbf{S}_K\}$, соответственно. Единичная матрица размера T : $\mathbf{I}_T = \text{diag}\{1, \dots, 1\} \in \mathbb{C}^{T \times T}$. След квадратной матрицы \mathbf{A} обозначается как $\text{tr} \mathbf{A} = \sum_{k=1}^K a_{kk}$, а норма Фробениуса: $\|\mathbf{H}\| = \sqrt{\sum_{r=1, t=1}^{R, T} |h_{rt}|^2}$.

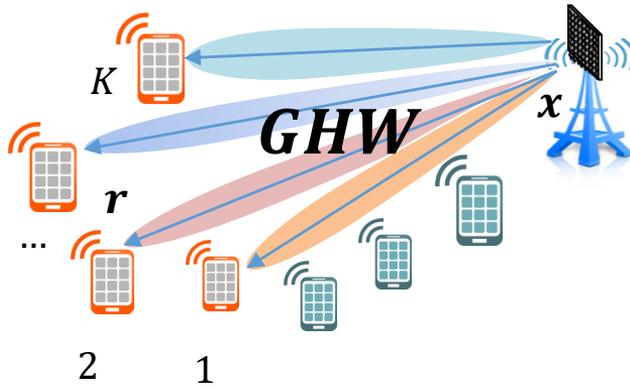


Рис. 1. Пример использования префильтра в системе с множеством антенн. Задача состоит в нахождении оптимальной матрицы префильтра \mathbf{W} для системы при заданной функции спектральной эффективности и ограничениях (25).

Таблица 1. Обозначения и символы

Символ	Обозначение
$()^H$	Эрмитово сопряжение (комплексное сопряжение + транспонирование)
\mathbf{H}, \mathbf{W}	Матрицы
\mathbf{W}_n	n -й столбец матрицы \mathbf{W}
$\mathbf{H}_m, \mathbf{W}^m$	m -я строка матриц \mathbf{H}, \mathbf{W}
h_{nm}, w_{nm}	Элемент на пересечении n -й строки и m -го столбца матриц \mathbf{H}, \mathbf{W}
$\mathbf{S} = \text{diag}(s_1, \dots, s_N)$	Диагональная матрица
K	Количество пользователей
T	Число передающих антенн
R	Общее число приёмных антенн
R_k	Число приёмных антенн у k -го пользователя
L	Общее число принимаемых символов в системе
L_k	Число принимаемых символов у k -го пользователя (ранг k -го пользователя)

2. Методы

2.1. Модель канала и системы

В соответствии с [2, 22, 12] рассматривается широкополосный канал с множеством антенн. Модель многопользовательской системы с множеством антенн описывается следующей линейной системой:

$$\mathbf{r} = \mathbf{G}(\mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{n}) = \mathbf{G}\mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{G}\mathbf{n}, \quad (1)$$

где $\mathbf{x}, \mathbf{r} \in \mathbb{C}^L$ — это соответственно *векторы передаваемого и принимаемого сигнала*, $\mathbf{H} \in \mathbb{C}^{R \times T}$ — *матрица нисходящего канала*⁸, $\mathbf{W} \in \mathbb{C}^{T \times L}$ — *матрица префильтра*, а $\mathbf{G} \in \mathbb{C}^{L \times R}$ — *блочно-диагональная матрица постфильтра*; вектор шума $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)$ предполагается комплексно-нормальным с нулевым средним и дисперсией σ^2 (см. рис. 1). Линейный префильтр и постфильтр реализуются простыми матричными умножениями.

Постоянные T, R, L обозначают количество передающих антенн, общее количество приёмных антенн и количество передаваемых символов (количество потоков), соответственно. Обычно выполняется соотношение $L \leq R \leq T$. Каждая из матриц $\mathbf{G}, \mathbf{H}, \mathbf{W}$ раскладывается по K *пользователям*:

$$\mathbf{G} = \text{bdiag}\{\mathbf{G}_1, \dots, \mathbf{G}_K\}, \quad \mathbf{H} = [\mathbf{H}_1; \dots; \mathbf{H}_K], \quad \mathbf{W} = (\mathbf{W}_1, \dots, \mathbf{W}_K),$$

как показано на рис. 2, где $\mathbf{G}_k \in \mathbb{C}^{L \times R_k}$, $\mathbf{H}_k \in \mathbb{C}^{R_k \times T}$, $\mathbf{W}_k \in \mathbb{C}^{T \times L_k}$ — матрицы, соответствующие k -му пользователю.

2.1.1. Сингулярное разложение канала

Удобно [20] представлять матрицу канала пользователя k с помощью её сингулярного разложения:

$$\begin{aligned} \mathbf{H}_k &= \mathbf{U}_k^H \mathbf{S}_k \mathbf{V}_k, & \mathbf{U}_k \mathbf{U}_k^H &= \mathbf{U}_k^H \mathbf{U}_k = \mathbf{I}_{R_k}, \\ \mathbf{S}_k &= \text{diag}\{s_1, \dots, s_{R_k}\}, & \mathbf{V}_k \mathbf{V}_k^H &= \mathbf{I}_{R_k}. \end{aligned} \quad (2)$$

⁸Здесь и далее предполагается, что передающая сторона имеет идеальную информацию о состоянии всех нисходящих каналах. Это предположение оправдано для систем с разделением восходящей и нисходящей передач по времени, в которых можно использовать принцип взаимности и проводить непосредственные измерения соответствующих каналов. Каждый пользователь, в свою очередь, обладает только информацией о собственном канале с учетом примененного префильтра, но не знает информации о канале других пользователей.

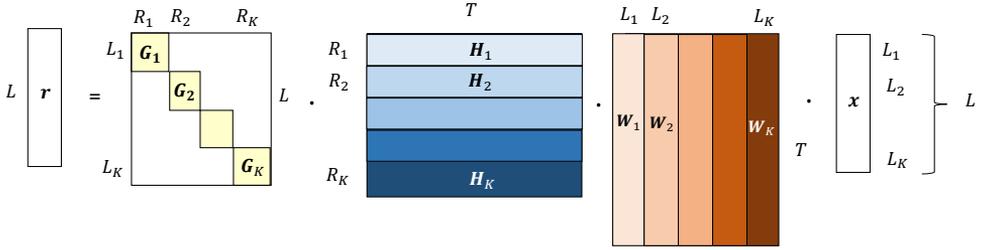


Рис. 2. Пример системы передачи в матричной форме. Многопользовательский префильтр \mathbf{W} позволяет одновременно передавать различную информацию разным пользователям.

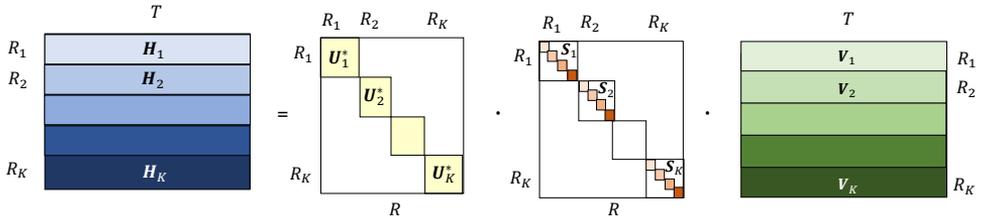


Рис. 3. Основное разложение матрицы канала.

Здесь *матрица канала* пользователя k , $\mathbf{H}_k \in \mathbb{C}^{R_k \times T}$, содержит по строкам векторы канала $\mathbf{H}_i \in \mathbb{C}^T$; сингулярные значения $\mathbf{S}_k \in \mathbb{C}^{R_k \times R_k}$ отсортированы по убыванию; $\mathbf{U}_k \in \mathbb{C}^{R_k \times R_k}$ — унитарная матрица левых сингулярных векторов; матрица $\mathbf{V}_k \in \mathbb{C}^{R_k \times T}$ состоит из *правых сингулярных векторов* — векторов-строк.

Объединяя все каналы пользователей, получаем следующее *разложение матрицы канала*: $\mathbf{H} = \mathbf{U}^H \mathbf{S} \mathbf{V}$ (лемма 1 и рис. 3), где каждая из матриц разложения — $\mathbf{U}^H \in \mathbb{C}^{R \times R}$, $\mathbf{S} \in \mathbb{C}^{R \times R}$, $\mathbf{V} \in \mathbb{C}^{R \times T}$ — состоит из K подматриц $\mathbf{U}_k^H \in \mathbb{C}^{R_k \times R_k}$, $\mathbf{S}_k \in \mathbb{C}^{R_k \times R_k}$, $\mathbf{V}_k \in \mathbb{C}^{R_k \times T}$, содержащих векторы $\mathbf{U}_l^H \in \mathbb{C}^{R_k}$, $s_l \in \mathbb{R}$, $\mathbf{V}_l \in \mathbb{C}^T$.

Лемма 1 (Основное разложение). *Для линейной системы $\mathbf{r} = \mathbf{G}(\mathbf{H}\mathbf{W}\mathbf{x} + \mathbf{n})$ существует разложение матрицы канала в виде $\mathbf{H} = \mathbf{U}^H \mathbf{S} \mathbf{V}$, где $\mathbf{S} = \text{diag}\{\mathbf{S}_1, \dots, \mathbf{S}_K\} \in \mathbb{R}_+^{R \times R}$ — диагональная матрица сингулярных значений,*

$\mathbf{V} = [\mathbf{V}_1, \dots, \mathbf{V}_K] \in \mathbb{C}^{R \times T}$ — конкатенация матриц сингулярных векторов для каналов отдельных пользователей,

а $\mathbf{U}^H = \text{bdiag}\{\mathbf{U}_1^H, \dots, \mathbf{U}_K^H\} \in \mathbb{C}^{R \times R}$ — блочно-диагональная унитарная матрица.

Отметим, что это представление не является строгим сингулярным разложением всей матрицы \mathbf{H} , так как векторы $\mathbf{V}_{k,j}, \mathbf{V}_{l,i}$, соответствующие различным пользователям $k \neq l$, в общем случае не ортогональны. Тем не менее, данное представление обладает важными свойствами: матрица $\mathbf{S} = \text{diag}(\mathbf{S}_k) \in \mathbb{C}^{R \times R}$ является *диагональной*, а $\mathbf{U} = \text{bdiag}(\mathbf{U}_k) \in \mathbb{C}^{R \times R}$ — *блочной-диагональной унитарной матрицей*. Это позволяет компенсировать множитель $\mathbf{U}^H \mathbf{S}$ на стороне приёмника (причём каждый пользователь обрабатывает только собственную часть $\mathbf{U}_k^H \mathbf{S}_k$). Таким образом, на стороне передачи достаточно инвертировать только матрицу \mathbf{V} , которая значительно проще, чем \mathbf{H} : во-первых, её строки имеют единичную норму, а во-вторых, она является естественным объектом для задачи адаптации ранга [23]. Мы активно используем лемму 1 в дальнейшем.

2.1.2. Идея исследуемого алгоритма префильтра

Кратко сформулируем суть предлагаемого алгоритма. Интуитивная идея исследуемого метода префильтра заключается в том, чтобы обеспечить максимальную мощность сигнала при одновременном устранении так называемых межпользовательских помех. Известно, что максимальную мощность сигнала обеспечивает метод сопряжённого префильтра: $\mathbf{W}_{MRT}(\mathbf{V}) = \mathbf{V}^H$, а полное устранение помех достигается с помощью псевдообратного префильтра: $\mathbf{W}_{ZF}(\mathbf{V}) = \mathbf{V}^H(\mathbf{V}\mathbf{V}^H)^{-1}$ [3]. Однако у обоих методов имеются недостатки, которые могут быть устранены с помощью регуляризованного псевдообратного префильтра.

$$\mathbf{W}_{RZF}(\mathbf{V}) = \mathbf{V}^H(\mathbf{V}\mathbf{V}^H + \lambda \mathbf{I})^{-1},$$

где параметр регуляризации $\lambda > 0$ зависит от уровня шума и среднего значения потерь на пути сигнала [4].

Если у пользователей существенно различаются потери на пути, предпочтительнее использовать регуляризацию с диагональной матрицей. Нескалярную (матричную) регуляризацию предложил Э. Бьёрнсон в работе [5], где она была получена с помощью метода двойственности. Показано, что оптимальная регуляризация имеет диагональную форму (в общем случае с различными элементами). Однако полученный результат не носил конструктивного характера и не давал явной формулы или процедуры для построения оптимальной регуляризации.

В нашей работе мы исследуем явную эвристическую формулу для диагональной регуляризации — алгоритм адаптивного регуляризованного псевдообратного префильтра (АРПФ, для англоязычной литературы мы бы перевели как Adaptive Regularized Zero-Forcing, ARZF) для многопользовательской системы с множеством антенн.

$$\mathbf{W}_{ARZF} = \mathbf{V}^H(\mathbf{V}\mathbf{V}^H + \lambda \mathbf{S}^{-2})^{-1}.$$

Теоретическое обоснование формулы представлено далее в разделе 2.2. Численные эксперименты с использованием модели канала “Quadriga” [9] показывают, что рассматриваемый метод АРПФ превосходит известные скалярные алгоритмы РПФ (см. раздел 3.2).

2.1.3. Число передаваемых символов и сопряжённый пост-фильтр

В предыдущем разделе была введено понятие сингулярного разложения (см. уравнение 2). Обычно передатчик передаёт каждому пользователю несколько символов, это число (ранг пользователя) меньше числа приёмных антенн пользователя ($L_k \leq R_k$). В таком случае естественно выбирать для передачи первые L_k векторов из матрицы \mathbf{V}_k , соответствующих L_k наибольшим сингулярным числам из матрицы \mathbf{S}_k .

Обозначим через $\tilde{\mathbf{S}}_k \in \mathbb{C}^{L_k \times L_k}$ L_k наибольших сингулярных чисел из \mathbf{S}_k , через $\tilde{\mathbf{U}}_k^H \in \mathbb{C}^{R_k \times L_k}$ и $\tilde{\mathbf{V}}_k \in \mathbb{C}^{L_k \times T}$ — соответствующие им первые левые и правые сингулярные векторы:

$$\begin{aligned}\tilde{\mathbf{S}}_k &= \text{diag}\{s_{k,1}, \dots, s_{k,L_k}\}, \\ \tilde{\mathbf{U}}_k^H &= (\mathbf{u}_{k,1}^H, \dots, \mathbf{u}_{k,L_k}^H), \\ \tilde{\mathbf{V}}_k &= [\mathbf{v}_{k,1}; \dots; \mathbf{v}_{k,L_k}],\end{aligned}\tag{3}$$

где волнистая черта $\tilde{}$ означает сокращённое представление по сингулярным векторам, то есть $\text{rank}(\tilde{\mathbf{V}}_k) = L_k \leq R_k = \text{rank}(\mathbf{V}_k)$.

Числа L_k (и, соответственно, выбор $\tilde{\mathbf{V}}_k$) определяются в задаче выбора ранга, которая вместе с планировщиком решается до этапа выбора префильтра. За дополнительной информацией по адаптации ранга см., например, [23]. Далее считаем, что L_k и $\tilde{\mathbf{V}}_k$ уже выбраны.

После выбора префильтра и передачи на стороне пользователя k необходимо выбрать матрицу постфильтра $\mathbf{G}_k \in \mathbb{C}^{L_k \times R_k}$, которая учитывает ранг пользователя L_k . Способ, которым пользователь выполняет процедуру постфильтра, существенно влияет на итоговую производительность системы, и разные алгоритмы постфильтра требуют различных оптимальных префильтров (см. [4], где префильтр является функцией матрицы постфильтра). Оптимально было бы согласованно выбирать префильтр и постфильтр, но это затруднительно из-за распределённого характера беспроводной связи. Однако существуют идеи по настройке префильтра на передатчике при предполагаемом способе постфильтра на приёмной стороне [10]. В данной работе мы не рассматриваем такой подход, хотя он может быть использован для дальнейшего улучшения.

Для проведения аналитических выкладок мы будем предполагать использование *сопряжённого постфильтра* (который достаточно хорошо

приближает реальные алгоритмы и при этом сильно упрощает задачу, см. [24]) следующего вида:

$$\mathbf{G}_k^C = \tilde{\mathbf{S}}_k^{-1} \tilde{\mathbf{U}}_k \in \mathbb{C}^{L_k \times R_k} \iff \mathbf{G}^C := \tilde{\mathbf{S}}^{-1} \tilde{\mathbf{U}} \in \mathbb{C}^{L \times R}, \quad (4)$$

где матрица $\tilde{\mathbf{U}}_k \in \mathbb{C}^{L_k \times R_k}$ содержит первые L_k сингулярных векторов, а диагональная матрица $\tilde{\mathbf{S}}_k \in \mathbb{C}^{L_k \times L_k}$ — соответствующие сингулярные числа. Блочно-диагональная матрица $\tilde{\mathbf{U}}$ составлена из блоков $\tilde{\mathbf{U}}_k$, а диагональная матрица $\tilde{\mathbf{S}}^{-1}$ из блоков $\tilde{\mathbf{S}}_k^{-1}$. Таким образом, матрица постфильтра \mathbf{G}^C также блочно-диагональна: $\mathbf{G}^C := \tilde{\mathbf{S}}^{-1} \tilde{\mathbf{U}} \in \mathbb{C}^{L \times R}$.

Теорема 1. *Сопряжённый постфильтр удаляет неиспользуемые сингулярные векторы: $\mathbf{G}^C \mathbf{H} = \tilde{\mathbf{V}}$, и уравнение модели (1) принимает вид:*

$$\mathbf{r} = \tilde{\mathbf{V}} \mathbf{W} \mathbf{x} + \tilde{\mathbf{n}}, \quad \tilde{\mathbf{n}} := \tilde{\mathbf{S}}^{-1} \tilde{\mathbf{U}} \mathbf{n}. \quad (5)$$

Доказательство. Используя лемму (1), получаем:

$$\mathbf{G}_k^C \mathbf{H}_k = \tilde{\mathbf{S}}_k^{-1} \tilde{\mathbf{U}}_k \mathbf{U}_k^H \mathbf{S}_k \mathbf{V}_k = \tilde{\mathbf{S}}_k^{-1} [\mathbf{I}_{L_k} \quad \mathbf{0}] \mathbf{S}_k \mathbf{V}_k = \tilde{\mathbf{S}}_k^{-1} \tilde{\mathbf{S}}_k \tilde{\mathbf{V}}_k = \tilde{\mathbf{V}}_k, \quad (6)$$

что при объединении по всем пользователям $k = 1, \dots, K$ даёт равенство (5). \square

Следствие 1. *При предположении о гауссовом независимом шуме $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)$, распределение эффективного шума $\tilde{\mathbf{n}} \in \mathbb{C}^{L \times 1}$ (который появляется при сопряжённом постфильтре) описывается как: $\tilde{\mathbf{n}} \sim \mathcal{CN}(0, \sigma^2 \tilde{\mathbf{S}}^{-2})$.*

Доказательство.

$$\begin{aligned} \mathbb{E}[\tilde{\mathbf{n}} \tilde{\mathbf{n}}^H] &= \mathbb{E}[\tilde{\mathbf{S}}^{-1} \tilde{\mathbf{U}} \mathbf{n} \mathbf{n}^H \tilde{\mathbf{U}}^H \tilde{\mathbf{S}}^{-1}] = \tilde{\mathbf{S}}^{-1} \tilde{\mathbf{U}} \mathbb{E}[\mathbf{n} \mathbf{n}^H] \tilde{\mathbf{U}}^H \tilde{\mathbf{S}}^{-1} \\ &= \sigma^2 \tilde{\mathbf{S}}^{-1} \mathbf{I}_L \tilde{\mathbf{S}}^{-1} = \sigma^2 \tilde{\mathbf{S}}^{-2}. \end{aligned} \quad (7)$$

\square

Замечание 1. *Следствие 1 играет ключевую роль в нашей работе, поскольку оно даёт основание использовать матрицу $\tilde{\mathbf{S}}^{-2} \in \mathbb{C}^{L \times L}$ в регуляризационном слагаемом префильтра с целью корректного учёта эффективного шума $\tilde{\mathbf{S}}^{-1} \tilde{\mathbf{U}} \mathbf{n}$.*

Замечание 2. *Формулируемая теорема существенно упрощает исходную задачу, уменьшает её размерность и позволяет унифицировать обозначения. В частности, мы можем работать с рангом пользователей размерности L_k и L , вместо пространства антенн пользователя.*

Заметим также, что достаточно выполнять лишь частичное сингулярное разложение канала $\mathbf{H}_k \in \mathbb{C}^{R_k \times T}$, сохраняя только первые L_k сингулярных значений и векторов:

$$\mathbf{H}_k \approx \tilde{\mathbf{U}}_k^H \tilde{\mathbf{S}}_k \tilde{\mathbf{V}}_k. \quad (8)$$

В связи с этим далее по тексту мы будем опускать волнистую черту и обозначать $\mathbf{U}_k, \mathbf{S}_k, \mathbf{V}_k$ вместо $\tilde{\mathbf{U}}_k, \tilde{\mathbf{S}}_k, \tilde{\mathbf{V}}_k$ соответственно.

Замечание 3. Введённый сопряжённый постфильтр является “идеальным” и не может быть реализован на практике. Однако можно показать, что реалистичные алгоритмы постфильтра, такие как постфильтр среднеквадратичной ошибки (Minimum Mean Square Error, MMSE) или постфильтр с подавлением помех (Interference Rejection Combining, IRC) [25], зачастую демонстрируют схожее поведение. В численных экспериментах для сравнения различных префильтров используется постфильтр среднеквадратичной ошибки.

2.1.4. Постфильтр для минимизации среднеквадратичной ошибки

Целью процедуры постфильтра сигнала является псевдообращение произведения матриц канала и префильтра. Наиболее распространённой формой постфильтра \mathbf{G} является матрица минимизации среднеквадратичной ошибки [26]. В статистике и обработке сигналов постфильтр среднеквадратичной ошибки — это метод оценки, минимизирующий функцию среднеквадратичной ошибки, которая, в свою очередь, служит общей мерой качества оценки предсказанных значений зависимой переменной.

Определение постфильтра среднеквадратичной ошибки реализуется следующим образом:

$$\mathbf{G}_k^{MMSE}(\mathbf{A}_k) = \mathbf{A}_k^H (\mathbf{A}_k \mathbf{A}_k^H + \sigma^2 \mathbf{I})^{-1}, \quad \mathbf{A}_k = \mathbf{H}_k \mathbf{W}_k. \quad (9)$$

Параметр P — мощность базовой станции, а σ^2 — мощность шума в системе. Метод постфильтра среднеквадратичной ошибки предполагает устранение шума при допущении его одинаковости для всех символов: $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)$, что может нарушаться на практике.

Лемма 2. Для системы (1): $\mathbf{r}_k = \mathbf{G}_k \mathbf{H}_k \mathbf{W} \mathbf{x} + \mathbf{G}_k \mathbf{n}_k \in \mathbb{C}^{L_k}$ с распределением шума $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)$ и префильтром $\mathbf{W} = \mathbf{H}^+ = \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1}$, матрица (9): \mathbf{G}^{MMSE} минимизирует норму в квадрате: $\mathbb{E}_{\mathbf{n}, \mathbf{x}} \|\mathbf{r}_k - \mathbf{x}\|^2$, $k = 1 \dots K$.

Доказательство. Подставим выражение для системы $\mathbf{r}_k = \mathbf{G}_k \mathbf{H}_k \mathbf{W} \mathbf{x} + \mathbf{G}_k \mathbf{n}_k$ в функцию потерь:

$$\mathbb{E}_{\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)} \|\mathbf{r}_k - \mathbf{x}\|^2 = \mathbb{E}_{\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)} \|\mathbf{G}_k \mathbf{H}_k \mathbf{W} \mathbf{x} - \mathbf{x} + \mathbf{G}_k \mathbf{n}_k\|^2.$$

Второе слагаемое обнуляется благодаря математическому ожиданию шума с нулевым средним. Третье слагаемое также исчезает.

Раскроем скобки, используя формулу квадрата нормы суммы:

$$\begin{aligned} \mathbb{E}_{\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)} \|(\mathbf{G}_k \mathbf{H}_k \mathbf{W} - \mathbf{I})\mathbf{x} + \mathbf{G}_k \mathbf{n}_k\|^2 &= \|(\mathbf{G}_k \mathbf{H}_k \mathbf{W} - \mathbf{I})\mathbf{x}\|^2 + \\ &+ 2\mathbb{E}_{\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)} \Re\{\langle \mathbf{G}_k \mathbf{H}_k \mathbf{W} \mathbf{x} - \mathbf{x}, \mathbf{G}_k \mathbf{n}_k \rangle\} + \mathbb{E}_{\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_R)} \|\mathbf{G}_k \mathbf{n}_k\|^2 = \\ &= \|(\mathbf{G}_k \mathbf{H}_k \mathbf{W} - \mathbf{I})\mathbf{x}\|^2 + \sigma^2 \|\mathbf{G}_k\|^2. \end{aligned} \quad (10)$$

Теперь возьмём математическое ожидание по символам \mathbf{x} с условием $\mathbb{E}_{\mathbf{x} \sim \mathcal{CN}(0, \mathbf{I}_L)} \mathbf{x} \mathbf{x}^H = \mathbf{I}$:

$$\begin{aligned} \mathbb{E}_{\mathbf{x} \sim \mathcal{CN}(0, \mathbf{I}_L)} \|(\mathbf{G}_k \mathbf{H}_k \mathbf{W} - \mathbf{I})\mathbf{x}\|^2 + \sigma^2 \|\mathbf{G}_k\|^2 &= \\ &= \|\mathbf{G}_k \mathbf{H}_k \mathbf{W} - \mathbf{I}\|^2 + \sigma^2 \|\mathbf{G}_k\|^2 \rightarrow \min_{\mathbf{G}_k}. \end{aligned} \quad (11)$$

Если выполнены условия: $\mathbf{H}_k \mathbf{W} = \mathbf{H}_k \mathbf{W}_k$, где $\mathbf{W} = \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1}$, то функция принимает вид:

$$\|\mathbf{G}_k \mathbf{H}_k \mathbf{W}_k - \mathbf{I}\|^2 + \sigma^2 \|\mathbf{G}_k\|^2 \rightarrow \min_{\mathbf{G}_k} \quad (12)$$

Вычислим градиент функции (12) и приравняем его к нулю:

$$\begin{aligned} \nabla_{\mathbf{G}_k} \{ \|\mathbf{G}_k \mathbf{H}_k \mathbf{W}_k - \mathbf{I}\|^2 + \sigma^2 \|\mathbf{G}_k\|^2 \} &= \\ 2(\mathbf{G}_k \mathbf{H}_k \mathbf{W}_k - \mathbf{I})(\mathbf{H}_k \mathbf{W}_k)^H + 2\sigma^2 \mathbf{G}_k &= \\ = 2\mathbf{G}_k (\mathbf{H}_k \mathbf{W}_k) (\mathbf{H}_k \mathbf{W}_k)^H - 2(\mathbf{H}_k \mathbf{W}_k)^H + 2\sigma^2 \mathbf{G}_k &= 0 \end{aligned} \quad (13)$$

Переносим слагаемые с \mathbf{G}_k в левую часть:

$$\mathbf{G}_k ((\mathbf{H}_k \mathbf{W}_k) (\mathbf{H}_k \mathbf{W}_k)^H + \sigma^2 \mathbf{I}) = (\mathbf{H}_k \mathbf{W}_k)^H \quad (14)$$

В итоге выражаем из уравнения \mathbf{G}_k :

$$\widehat{\mathbf{G}}_k^{MMSE} = \mathbf{G}_k = (\mathbf{H}_k \mathbf{W}_k)^H ((\mathbf{H}_k \mathbf{W}_k) (\mathbf{H}_k \mathbf{W}_k)^H + \sigma^2 \mathbf{I})^{-1} \quad (15)$$

Получаем желаемое решение (9). \square

2.1.5. Оценки качества

Для оценки качества методов префильтра используются следующие функционалы. Эти функции зависят не от конкретных передаваемых символов $\mathbf{x} \in \mathbb{C}^{L \times 1}$, а от их распределения [5]. Таким образом, получаем общую

характеристику для всех возможных символов, которые могут быть переданы с использованием данной матрицы префильтра.

Рассмотрим сквозную нумерацию символов $l = 1, \dots, L$ для всех пользователей, и определим индексную функцию $k = k(l)$, которая возвращает номер пользователя, получающего символ l . Функционал отношения сигнал-интерференция-шум (ОСИШ) для l -го символа пользователя $k = k(l)$ определяется как:

$$SINR_l(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_l, \sigma^2) := \frac{|\mathbf{G}_l \mathbf{H}_k \mathbf{W}_l|^2}{\sum_{i \neq l}^L |\mathbf{G}_l \mathbf{H}_k \mathbf{W}_i|^2 + \sigma^2 \|\mathbf{G}_l\|^2}. \quad (16)$$

Формула (16) показывает соотношение между полезной частью сигнала и помехами. Она зависит от всей матрицы префильтра $\mathbf{W} \in \mathbb{C}^{T \times L}$, где вектор $\mathbf{W}_l \in \mathbb{C}^{T \times 1}$ соответствует префильтру для l -го символа, матрицы канала $\mathbf{H}_k \in \mathbb{C}^{R_k \times T}$ пользователя k , вектора постфильтра $\mathbf{G}_l \in \mathbb{C}^{1 \times R_k}$, и уровня шума после постфильтра: $\mathbb{E}[\mathbf{G}_l \mathbf{n}] = \sigma^2 \|\mathbf{G}_l\|^2$. Формула (16) может быть эффективно вычислена для всех L слоёв при помощи матричных операций и суммирования.

Формулу можно упростить, используя теорему 1:

Следствие 2. Для сингулярного разложения (2) и предположения сопряжённого постфильтра (4), ОСИШ выражается как:

$$SINR_l^C(\mathbf{W}, \mathbf{V}_l, s_l, \sigma^2) = SINR_l(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_l^C, \sigma^2) = \frac{|\mathbf{V}_l \mathbf{W}_l|^2}{\sum_{i \neq l}^L |\mathbf{V}_l \mathbf{W}_i|^2 + \sigma^2 / s_l^2}. \quad (17)$$

Важным критерием производительности сети является *спектральная эффективность* (Spectral Efficiency, SE) пользователя k , отражающая максимальную скорость передачи информации на заданной полосе частот. При передаче одного символа ($L_k = 1$), она ограничена энтропией Шеннона и выражается через ОСИШ следующим образом:

$$SE(SINR) := \log_2(1 + SINR). \quad (18)$$

Это является теоретическим пределом передаваемой информации. Современные схемы модуляции и кодирования, в сочетании с механизмами гибридного автоматического запроса повторной передачи (Hybrid Automatic Repeat Request, HARQ) и управлением блочной ошибки (Block Error Rate, BLER), позволяют достичь скоростей, близких к этому пределу. Обратите внимание, что ОСИШ здесь используется в линейной шкале, а не в децибелах (дБ).

При передаче нескольких символов спектральная эффективность пользователя k в общем случае не является суммой спектральных эффективностей разных символов, так как используется общий транспортный

блок, к которому применяются общие алгоритмы кодирования и модуляции. Обычно вводится *эффективный* ОСИШ как функция от ОСИШ по символам:

$$SINR_k^{eff} = f(SINR_1, \dots, SINR_{L_k}),$$

и тогда по формуле (18) получаем:

$$SE_k(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_k, \sigma^2) = L_k \left(SINR_k^{eff}(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_k, \sigma^2) \right). \quad (19)$$

Существует множество подходов к оценке эффективного ОСИШ для разных схем квадратурно-амплитудной модуляции (КАМ) КАМ64 и КАМ256 (см. [27]); в численных экспериментах мы используем модель КАМ256.

Также рассмотрим приближённую формулу с использованием геометрического среднего по ОСИШ символов. Это эквивалентно обычному среднему по символам в дБ. Данное эвристическое приближение будет использоваться в градиентной оптимизации:

$$SINR_k^{eff}(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_k, \sigma^2) \approx \left(\prod_{l \in \mathcal{L}_k} SINR_l(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_l, \sigma^2) \right)^{\frac{1}{L_k}}. \quad (20)$$

Наиболее общий вид задачи — многокритериальная оптимизация вектора (SE_1, \dots, SE_K) . Для такой постановки может быть проведён анализ по Парето (см. [28, 29]), который, однако, не даёт уникального решения. Поэтому часто задача сводится к однокритериальной оптимизации: $J = J(SE_1, \dots, SE_K) \rightarrow \max$ или $J = J(SINR_1, \dots, SINR_K)$ [5]. Один из возможных вариантов — максимизация суммы спектральных эффективностей:

$$J^{SE}(\mathbf{W}) := SE(\mathbf{W}, \mathbf{H}, \mathbf{G}, \sigma^2) = \sum_{k=1}^K SE_k(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_k, \sigma^2). \quad (21)$$

Такой критерий является естественным, поскольку скорости передачи данных по пользователям являются аддитивными. Возможны и другие цели, например, оптимизация качества самых плохих пользователей. Подробный анализ таких постановок приведён в [12, разд. 7], включая:

$$SE_{\min} = \min_k SE_k \rightarrow \max \quad \text{или} \quad SINR_{\min} = \min_{1 \leq j \leq L} SINR_j \rightarrow \max. \quad (22)$$

Наконец, введём *однопользовательский* ОСИШ (Single User SINR, SUSINR) для пользователя k (эти метрики будут использоваться для оценки результатов на тестах).

$$SUSINR_k(\mathbf{S}_k, \sigma^2, P) := \frac{P}{L_k \sigma^2} \left(\prod_{l \in \mathcal{L}_k} s_l^2 \right)^{\frac{1}{L_k}}. \quad (23)$$

2.1.6. Постановка задачи и ограничения по мощности

Прежде всего, предполагается, что полная матрица канала \mathbf{H} , число пользователей K и их ранги L_k — известные заранее величины. Это означает, что задачи планировщика — каких пользователей обслуживать, и выбора ранга — какой ранг назначается каждому пользователю — уже решены. В реальных сетях такая последовательность шагов стандартна. Проблемы планировщика и выбора ранга представляют собой сложные задачи управления радиоресурсами и выходят за рамки данного исследования (см., например, [30] и ссылки внутри для задачи планировщика, [23], [10] — для задачи выбора ранга).

Эти алгоритмы влияют на свойства матрицы \mathbf{H} , например, планировщик может выбирать только пользователей с достаточно малыми корреляциями: $\|\mathbf{C}\| = \|\mathbf{V}\mathbf{V}^H - \mathbf{I}_R\| \leq \varepsilon$. Мы учитываем это и рассматриваем сценарии с малой корреляцией каналов пользователей.

Далее, используем модель канала в виде (1), что, в частности, означает точные измерения канала. Для упрощения задачи предполагаем, что стратегия постфильтра $\mathbf{G} = \mathbf{G}(\mathbf{H}, \mathbf{W})$ — известная функция, причём применение сопряжённого постфильтра (4) упрощает модель канала до вида (5). На основе этой модели вычисляется *ОСИШ* передаваемых символов по формуле (17) и эффективный *ОСИШ* пользователей по приближённой формуле (20).

Обозначим полную мощность системы через P и предполагаем, что вектор передаваемых символов нормирован: $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_L$. Тогда ограничения по мощности на префильтр могут быть следующими:

$$\|\mathbf{W}\|^2 \leq P, \quad \text{или} \quad \|(w_{t1}, \dots, w_{tL})\|^2 \leq P/T, \quad t = 1, \dots, T, \quad (24)$$

что соответствует, соответственно, полному ограничению по мощности и более реалистичному ограничению по мощности на каждую антенну [12].

Цель состоит в нахождении матрицы префильтра, максимизирующей суммарную спектральную эффективность (19) при выполнении ограничений по мощности (24), то есть:

$$\begin{aligned} J^{SE}(\mathbf{W}) &:= SE^C(\mathbf{W}), \\ \mathbf{W} &= \operatorname{argmax}_{\mathbf{W}} J^{SE}(\mathbf{W}), \\ \text{при условии: } &\|\mathbf{W}\|^2 \leq P, \end{aligned} \quad (25)$$

Даже после всех вышеуказанных упрощений, сформулированная задача остаётся слишком сложной для аналитического решения. Кроме того, она не является выпуклой или вогнутой, поэтому может иметь множество существенно различных локальных максимумов. Поэтому стратегия

данного исследования заключается в изучении эвристической формулы, превосходящей известные алгоритмы по результатам достоверного моделирования.

После задания конкретной формы префильтра $\mathbf{W} := \mu \mathbf{W}$ можно всегда выполнить ограничение по мощности, нормировкой на коэффициент μ , например, для случая (24):

$$\mu = \frac{\sqrt{P}}{\|\mathbf{W}\|} \quad \text{или} \quad \mu = \frac{\sqrt{P/T}}{\max_{t=1, \dots, T} \{\|(w_{t1}, \dots, w_{tL})\|\}}. \quad (26)$$

В моделировании мы используем более реалистичное ограничение по мощности на антенну.

Ниже сформулирован эвристический алгоритм, идея которого мотивирована упрощениями, приведёнными в следствии 1 и замечании 1. Теоретическим обоснованием служит модельная задача минимизации средней квадратичной ошибки, см., например, [4]:

$$\begin{aligned} J^{MSE}(\mathbf{W}) &:= \mathbb{E}_{\mathbf{x}, \mathbf{n}}[\|\mathbf{r}(\mathbf{W}) - \mathbf{x}\|^2], \\ \mathbf{W} &= \operatorname{argmin}_{\mathbf{W}} J^{MSE}(\mathbf{W}), \\ \text{при условии: } &\|\mathbf{W}\|^2 \leq P, \end{aligned} \quad (27)$$

где $\mathbf{r}(\mathbf{W})$ определяется моделью канала (1) или (5).

2.2. Базовые префильтры

В данном разделе приведены известные эталонные алгоритмы префильтра, а также представлено предлагаемое решение.

2.2.1. Сопряжённый префильтр

Алгоритм сопряжённого префильтра (Maximum Ratio Transmission, MRT) использует веса одного пользователя, взятые как сопряжённо-транспонированная матрица \mathbf{V}^H из сингулярного разложения. Такой подход приводит к максимизации мощности полезного сигнала одного пользователя, полностью игнорируя межпользовательские помехи. Метод сопряжённого префильтра особенно эффективен в условиях, когда уровень шума превышает уровень межпользовательских помех [3]:

$$\mathbf{W}_{MRT}(\mathbf{V}) = \mu \mathbf{V}^H, \quad (28)$$

где нормировочный множитель μ определяется из условия ограничения мощности (26).

Этот алгоритм приводит к формированию взаимно мешающих каналов, исходя из модели (5):

$$\mathbf{r} = \mathbf{V}\mathbf{W}\mathbf{x} + \mathbf{S}^{-1}\mathbf{U}\mathbf{n} = \mu\mathbf{V}\mathbf{V}^H\mathbf{x} + \mathbf{S}^{-1}\mathbf{U}\mathbf{n}.$$

2.2.2. Псевдообратный префильтр

Следующая модификация алгоритма префильтра выполняет декорреляцию передаваемых символов с помощью обратной корреляционной матрицы векторов канала. Такая конструкция префильтра направляет лучи сигнала к пользователям, не создавая между ними взаимных помех. В отличие от метода сопряжённого префильтра, метод зануления помех псевдообратного префильтра (Zero-Forcing, ZF) предпочтителен в условиях, когда межпользовательские помехи превышают уровень шума. В этом случае качество спектральной эффективности улучшается за счёт устранения этих помех [3]:

$$\mathbf{W}_{ZF}(\mathbf{V}) = \mu \mathbf{V}^\dagger = \mu \mathbf{V}^H (\mathbf{V} \mathbf{V}^H)^{-1}. \quad (29)$$

Соответствующая модель приёмника будет следующей:

$$\mathbf{r} = \mathbf{V} \mathbf{W} \mathbf{x} + \mathbf{S}^{-1} \mathbf{U} \mathbf{n} = \mu \mathbf{V} \mathbf{V}^H (\mathbf{V} \mathbf{V}^H)^{-1} \mathbf{x} + \mathbf{S}^{-1} \mathbf{U} \mathbf{n} = \mu \mathbf{x} + \mathbf{S}^{-1} \mathbf{U} \mathbf{n}.$$

Обозначим $\mathbf{F} = \mathbf{U} \mathbf{H} = \mathbf{S} \mathbf{V}$.

Теорема 2. *Предположим, что матрица канала имеет вид $\mathbf{H} = \mathbf{U}^H \mathbf{S} \mathbf{V}$ (см. Лемму 1). Тогда для метода зануления помех выполняется следующее равенство:*

$$\mathbf{W}_{ZF}(\mathbf{F}) \mathbf{S} = \mathbf{W}_{ZF}(\mathbf{V}). \quad (30)$$

Доказательство.

$$\begin{aligned} \mathbf{W}_{ZF}(\mathbf{F}) \mathbf{S} &= \mathbf{F}^H (\mathbf{F} \mathbf{F}^H)^{-1} \mathbf{S} = \mathbf{V}^H \mathbf{S} (\mathbf{S} \mathbf{V} \mathbf{V}^H \mathbf{S})^{-1} \mathbf{S} = \\ &= \mathbf{V}^H \mathbf{S} \mathbf{S}^{-1} (\mathbf{V} \mathbf{V}^H)^{-1} \mathbf{S}^{-1} \mathbf{S} = \mathbf{V}^H (\mathbf{V} \mathbf{V}^H)^{-1} = \mathbf{W}_{ZF}(\mathbf{V}). \end{aligned} \quad (31)$$

□

2.2.3. Регуляризованный псевдообратный фильтр

В геометрическом смысле в методе ПФ (29) лучи направляются не строго на пользователей, а с отклонением, что уменьшает полезный сигнал. Следующая модификация корректирует направление лучей, допуская некоторое межпользовательское вмешательство, что значительно увеличивает полезную нагрузку.

В практическом смысле, в методе ПФ может не существовать правой обратной матрицы канала или матрица $\mathbf{V} \mathbf{V}^H$ может быть плохо обусловленной, что ухудшает работу метода ПФ. Существует множество практических решений этой проблемы, основанных на регуляризации:

$$\mathbf{W}_{RZF}(\mathbf{V}) = \mu \mathbf{V}^H (\mathbf{V} \mathbf{V}^H + \lambda \mathbf{I})^{-1}. \quad (32)$$

Метод регуляризованного псевдообратного префилтра (Regularized Zero-Forcing, RZF) является самым распространённым в практических системах, и именно его мы используем в качестве основного эталонного метода. В качестве базовой настройки мы берём аналитическую форму регуляризации: $\lambda = \frac{L\sigma^2}{P}$ [8].

Этот метод не устраняет полностью межпользовательские и многослойные помехи, а допускает их в разумных пределах для увеличения мощности полезного сигнала. Он представляет собой компромисс между методами сопряжённого префилтра и псевдообратного префилтра [5], балансируя между максимизацией сигнала и минимизацией интерференции, поэтому параметр регуляризации должен быть выбран в зависимости от уровня шума.

Метод РПФ обладает следующими предельными свойствами [3]: при $\sigma^2 \rightarrow \infty$ он переходит в $\mathbf{W}_{MRT} = \mu \mathbf{V}^H$ — оптимальный в условиях *низкого ОСИШ*. При $\sigma^2 = 0$ формула совпадает с псевдообратным префилтром: $\mathbf{W}_{ZF} = \mu \mathbf{V}^H (\mathbf{V} \mathbf{V}^H)^{-1}$, которое оптимально при *высоком ОСИШ*.

Префилтр на основе ненормализованной матрицы канала [6], в случае когда количество передаваемых символов L меньше количества антенн приёмника, можно записать в следующем виде:

$$\mathbf{W}_{RZF}(\mathbf{F}) = \mu \mathbf{F}^H (\mathbf{F} \mathbf{F}^H + \lambda \mathbf{I})^{-1}, \quad (33)$$

где параметр $\mathbf{F} = \mathbf{S} \mathbf{V}$, а $\lambda = \frac{L\sigma^2}{P}$ [8] учитывает уровень шума $E[\mathbf{nn}^H] = \sigma^2 \mathbf{I}_R$. На практике это эквивалентно использованию ненормализованной матрицы канала \mathbf{H} , которая в нашей модели заменяется матрицей \mathbf{F} после разбиения канала пользователя на несколько потоков.

Сформулируем следующий известный факт о методе РПФ (32).

Теорема 3. *Рассмотрим разложение канала $\mathbf{H} = \mathbf{U}^H \mathbf{S} \mathbf{V}$ из Леммы 1. Префилтр $\mathbf{W}_{RZF}(\mathbf{V})$ с любым параметром $\lambda > 0$ является решением следующей задачи оптимизации:*

$$\mathbf{W}_{RZF}(\mathbf{V}) = \underset{\mathbf{W}}{\operatorname{argmin}} J(\mathbf{W}), \quad J(\mathbf{W}) = \|\mathbf{V} \mathbf{W} - \mathbf{I}\|_2^2 + \lambda \|\mathbf{W}\|_2^2. \quad (34)$$

Доказательство. Вычисляя градиент и приравнивая его к нулю, получаем:

$$\nabla J(\mathbf{W}) = 2\mathbf{V}^H (\mathbf{V} \mathbf{W} - \mathbf{I}) + 2\lambda \mathbf{W} = 0 \quad \Leftrightarrow \quad (\mathbf{V}^H \mathbf{V} + \lambda \mathbf{I}) \mathbf{W} = \mathbf{V}^H \quad (35)$$

$$\Leftrightarrow \quad \mathbf{W} = (\mathbf{V}^H \mathbf{V} + \lambda \mathbf{I})^{-1} \mathbf{V}^H = \mathbf{V}^H (\mathbf{V} \mathbf{V}^H + \lambda \mathbf{I})^{-1}.$$

Последнее равенство доказывается умножением на $(\mathbf{V} \mathbf{V}^H + \lambda \mathbf{I})$ справа и на $(\mathbf{V}^H \mathbf{V} + \lambda \mathbf{I})$ слева. \square

Замечание 4. Алгоритм $\mathbf{W}_{RZF}(\mathbf{V})$ также является решением задачи оптимизации с ограничением (27) при допущении $\mathbf{G}\mathbf{H} = \mathbf{V}$ (см. [4]):

$$\mathbf{W}_{RZF}(\mathbf{V}) = \underset{\mathbf{W}}{\operatorname{argmin}} \mathbb{E}_{\mathbf{x}, \mathbf{n}} [\|\mathbf{V}\mathbf{W}\mathbf{x} - \mathbf{x} + \mathbf{n}\|^2], \quad : \|\mathbf{W}\|^2 \leq P. \quad (36)$$

Эта задача сводится к (34) с множителем Лагранжа $\lambda = L\sigma^2/P$. Иными словами, $\mathbf{W}_{RZF}(\mathbf{V})$ — это частный случай префильтра Винера [4, Eq. (34), (35)], при допущении, что ковариационные матрицы сигнала и шума равны: $\mathbf{r}_{\mathbf{x}} := \mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_L$ и $\mathbf{r}_{\mathbf{n}} := \mathbb{E}[\mathbf{n}\mathbf{n}^H] = \sigma^2\mathbf{I}_R$, а также при специальном выборе постфильтра: $\mathbf{G} = \mathbf{U}$.

2.2.4. Префильтр Винера с занулением помех

В работе [4] авторы рассматривают регуляризованный псевдообратный префильтр Винера (РПФВ, Wiener Regularized Zero-Forcing, WRZF), который обеспечивает оптимум в задаче (36) для произвольных ковариационных матриц символов и шума $\mathbf{r}_{\mathbf{x}}$ и $\mathbf{r}_{\mathbf{n}}$ при известной матрице постфильтра \mathbf{G} . Мы рассматриваем детектирование сопряжённым префильтром (4), т.е. $\mathbf{G} = \mathbf{G}^C$, и применяем префильтр Винера к нормализованному каналу \mathbf{V} при $\mathbf{r}_{\mathbf{x}} = \mathbf{I}_L$ и соответствующей ковариации шума $\mathbf{r}_{\mathbf{n}} = \mathbf{S}^{-2}$.

Алгоритм

$$\mathbf{W}_{WRZF}(\mathbf{V}, \mathbf{S}) = \mu \mathbf{V}^H (\mathbf{V}\mathbf{V}^H + \lambda \mathbf{I})^{-1}, \quad \lambda = \frac{\sigma^2}{P} \operatorname{tr}(\mathbf{S}^{-2}) \quad (37)$$

является решением задачи оптимизации с ограничением (27) при допущении сопряжённого постфильтра ($\mathbf{G} = \mathbf{G}^C$, тогда $\mathbf{G}^C\mathbf{H} = \mathbf{V}$ согласно Теореме 1):

$$\mathbf{W}_{WRZF}(\mathbf{V}, \mathbf{S}) = \underset{\mathbf{W}}{\operatorname{argmin}} \mathbb{E}_{\mathbf{x}, \mathbf{n}} [\|\mathbf{V}\mathbf{W}\mathbf{x} - \mathbf{x} + \mathbf{n}\|^2], \quad : \|\mathbf{W}\|^2 \leq P, \quad (38)$$

где $\mathbb{E}[\mathbf{n}\mathbf{n}^H] = \mathbf{S}^{-2}$, $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_L$.

2.3. Предлагаемые методы префильтра

2.3.1. Определение адаптивного РПФ

Алгоритмы \mathbf{W}_{RZF} и \mathbf{W}_{WRZF} (а также префильтр Винера в более общем случае) являются алгоритмами построения префильтра со *скалярной регуляризацией*. Учитывая эффективный шум из следствия 1, рассматривается алгоритм *адаптивного РПФ (АРПФ)* с *диагональной матрицей регуляризации*:

$$\mathbf{W}_{ARZF}(\mathbf{V}) = \mu \mathbf{V}^H (\mathbf{V}\mathbf{V}^H + \lambda \mathbf{S}^{-2})^{-1}, \quad \lambda = \frac{\sigma^2 L}{P}. \quad (39)$$

АРПФ позволяет использовать различную регуляризацию для разных пользователей и слоёв в зависимости от их сингулярных значений, включая потери пути до пользователя.

Используя $\mathbf{r} = \tilde{\mathbf{V}}\mathbf{W}\mathbf{x} + \tilde{\mathbf{S}}^{-1}\tilde{\mathbf{U}}\mathbf{n}$, запишем квадрат ошибки между принятыми и переданными символами \mathbf{r} и \mathbf{x} :

$$\begin{aligned}\|\mathbf{r} - \mathbf{x}\|^2 &= \|\tilde{\mathbf{V}}\mathbf{W}\mathbf{x} - \mathbf{x} + \tilde{\mathbf{S}}^{-1}\tilde{\mathbf{U}}\mathbf{n}\|^2 \\ &\Rightarrow \mathbb{E}_{\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})} \|(\tilde{\mathbf{V}}\mathbf{W} - \mathbf{I})\mathbf{x} + \tilde{\mathbf{S}}^{-1}\tilde{\mathbf{U}}\mathbf{n}\|^2 = \|(\tilde{\mathbf{V}}\mathbf{W} - \mathbf{I})\mathbf{x}\|^2 \\ &\Rightarrow \mathbb{E}_{\mathbf{x} \sim \mathcal{CN}(0, \mathbf{I})} \|(\tilde{\mathbf{V}}\mathbf{W} - \mathbf{I})\mathbf{x}\|^2 = \|\tilde{\mathbf{V}}\mathbf{W} - \mathbf{I}\|^2.\end{aligned}\quad (40)$$

Введём в определение нормы обратную ковариационную матрицу шума $\tilde{\mathbf{S}}$, и получим следующую функцию взвешенного метода наименьших квадратов (41):

$$\text{MSE}_S(\mathbf{W}) = \|\tilde{\mathbf{V}}\mathbf{W} - \mathbf{I}\|_{\tilde{\mathbf{S}}}^2 + \lambda \|\mathbf{W}\|_2^2 = \|\tilde{\mathbf{S}}(\tilde{\mathbf{V}}\mathbf{W} - \mathbf{I})\|_2^2 + \lambda \|\mathbf{W}\|_2^2. \quad (41)$$

Таким образом, АРПФ является решением задачи оптимизации (41) (см. теорему 4).

Теорема 4. *Рассмотрим разложение канала $\mathbf{H} = \mathbf{U}^H \mathbf{S} \mathbf{V}$ из леммы 1. Префильтр (39) является решением следующей задачи оптимизации (взвешенная среднеквадратичная ошибка с регуляризацией):*

$$\mathbf{W}_{ARZF}(\mathbf{V}) = \underset{\mathbf{W}}{\text{argmin}} J_S(\mathbf{W}), \quad J_S(\mathbf{W}) := \|\mathbf{S}(\mathbf{V}\mathbf{W} - \mathbf{I})\|^2 + \lambda \|\mathbf{W}\|^2. \quad (42)$$

Доказательство. Вычислим градиент и приравняем его к нулю:

$$\begin{aligned}\nabla J_S(\mathbf{W}) = 2\mathbf{V}^H \mathbf{S}(\mathbf{S}\mathbf{V}\mathbf{W} - \mathbf{S}) + 2\lambda \mathbf{W} = 0 &\Leftrightarrow (\mathbf{V}^H \mathbf{S}^2 \mathbf{V} + \lambda \mathbf{I})\mathbf{W} = \mathbf{V}^H \mathbf{S}^2, \\ \mathbf{W} &= (\mathbf{V}^H \mathbf{S}^2 \mathbf{V} + \lambda \mathbf{I})^{-1} \mathbf{V}^H \mathbf{S}^2 = \mathbf{V}^H (\mathbf{V}\mathbf{V}^H + \lambda \mathbf{S}^{-2})^{-1}.\end{aligned}\quad (43)$$

Последнее равенство можно доказать, умножив с правой стороны на матрицу $(\mathbf{V}\mathbf{V}^H + \lambda \mathbf{S}^{-2})$ и с левой — на $(\mathbf{V}^H \mathbf{S}^2 \mathbf{V} + \lambda \mathbf{I})$. \square

Замечание 5. *Задача оптимизации (42) не является стандартной и не может быть записана в форме, аналогичной (36). Тем не менее, такая постановка задачи является более точным приближением к задаче максимизации суммарной спектральной эффективности (25), так как АРПФ обеспечивает более высокую суммарную спектральную эффективность по сравнению с РПФ и префильтром Винера. Подробные результаты моделирования приведены ниже.*

Замечание 6. Рассматриваемый алгоритм $\mathbf{W}_{ARZF}(\mathbf{V})$ и теорема 4 являются основным результатом данной работы. Алгоритмы типа префильтра Винера [4] используют скалярную регуляризацию вида $\lambda\mathbf{I}$, поэтому АРПФ не относится к этому классу. Как показано в [5], максимум функции отношения сигнал/шум+помехи (ОСИШ) пользователя (включая суммарную спектральную эффективность (19)) достигается алгоритмом с подходящей диагональной регуляризацией, а АРПФ является эвристикой такого вида.

Возможная интерпретация функции $J_{\mathbf{S}}(\mathbf{W})$ (34) заключается в следующем. Второй член $\lambda\|\mathbf{W}\|^2$ представляет собой стандартную регуляризацию по шуму, а первый член — норма $\|\mathbf{S}(\mathbf{V}\mathbf{W} - \mathbf{I})\|^2$, взвешенная по матрице \mathbf{S} , — придаёт больший вес слоям с большими сингулярными значениями. Таким образом, функция оптимизируется более точно для слоёв с более высоким качеством сигнала по сравнению с теми, у которых сигнал слабее. Иными словами, вектора префильтра для слоёв с высокими сингулярными значениями приближаются к псевдообратному префильтру, а для слоёв с низкими — к передаче с сопряжённым префильтром, то есть *адаптивный регуляризованный префильтр обеспечивает адаптивную регуляризацию*. В следующем разделе будет показано, что такой подход приводит к равномерному увеличению спектральной эффективности по сравнению с базовым методом с единичными весами.

Рассмотрим связь адаптивного регуляризационного префильтра с другими алгоритмами. Во-первых, видно, что параметр регуляризации в регуляризационном префильтре Винера представляет собой (арифметическое) среднее значение регуляризации АРПФ:

$$\frac{\sigma^2}{P} \text{tr}(\mathbf{S}^{-2}) = \frac{\sigma^2 L}{P} \cdot \frac{1}{L} \sum_{l=1}^L s_l^{-2}.$$

В случае, когда потери мощности в канале всех пользователей примерно одинаковы $s_l \approx s, l = 1, \dots, L$, АРПФ и префильтр Винера дают схожие результаты. Во-вторых, связь между префильтрами \mathbf{W}_{RZF} и \mathbf{W}_{ARZF} формулируется следующим образом:

Теорема 5. Пусть матрица канала имеет вид $\mathbf{H} = \mathbf{U}^H \mathbf{S} \mathbf{V}$ (см. Лемму 1), и обозначим $\mathbf{F} = \mathbf{U} \mathbf{H} = \mathbf{S} \mathbf{V}$. Тогда для АРПФ выполняется следующее соотношение:

$$\mathbf{W}_{ARZF}(\mathbf{V}) = \mathbf{W}_{RZF}(\mathbf{F}) \mathbf{S}. \quad (44)$$

Доказательство.

$$\begin{aligned}
 \mathbf{W}_{RZF}(\mathbf{F})\mathbf{S} &= \mathbf{F}^{\mathbf{H}}(\mathbf{F}\mathbf{F}^{\mathbf{H}} + \lambda\mathbf{I})^{-1}\mathbf{S} = \mathbf{V}^{\mathbf{H}}\mathbf{S}(\underbrace{\mathbf{S}\mathbf{V}\mathbf{V}^{\mathbf{H}}\mathbf{S} + \lambda\mathbf{I}}_{\mathbf{B}})^{-1}\mathbf{S} = \\
 &= \mathbf{V}^{\mathbf{H}}\mathbf{S}\mathbf{S}^{-1}(\mathbf{V}\mathbf{V}^{\mathbf{H}} + \mathbf{S}^{-1}\lambda\mathbf{I}\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\mathbf{S} = \\
 &= \mathbf{V}^{\mathbf{H}}(\underbrace{\mathbf{V}\mathbf{V}^{\mathbf{H}} + \lambda\mathbf{S}^{-2}}_{\mathbf{A}})^{-1} = \mathbf{W}_{ARZF}(\mathbf{V}). \quad (45)
 \end{aligned}$$

□

Замечание 7. В уравнении (45) акцент сделан на матрицах $\mathbf{A} = \mathbf{V}\mathbf{V}^{\mathbf{H}} + \lambda\mathbf{S}^{-2}$ и $\mathbf{B} = \mathbf{S}\mathbf{V}\mathbf{V}^{\mathbf{H}}\mathbf{S} + \lambda\mathbf{I}$, чтобы использовать их в следующем утверждении 6.

Замечание 8. Формулу АРПФ (39) можно также получить с помощью РСА-разложения [31], что и утверждается в Теореме 4.

Правая матрица \mathbf{S} в формуле (44) может интерпретироваться как особый вид алгоритма *распределения мощности* (см. интересное исследование в [12, разд. 7]), который распределяет общую мощность передачи между слоями. На практике предпочтительнее использовать $\mathbf{W}_{RZF}(\mathbf{V})$, а не $\mathbf{W}_{RZF}(\mathbf{F})$, так как нормы строк матрицы $\mathbf{F}\mathbf{F}^{\mathbf{H}} + \lambda\mathbf{I}$ могут существенно отличаться (вплоть до нескольких порядков), что приводит к несбалансированному распределению мощности между символами (в качестве альтернативы можно применить корректное распределение мощности для $\mathbf{W}_{RZF}(\mathbf{F})$, как это вытекает из теоремы 4). С другой стороны, параметр регуляризации в $\mathbf{W}_{RZF}(\mathbf{F})$ является более естественным и точным. Исследуемый префильтер $\mathbf{W}_{ARZF}(\mathbf{V})$ объединяет преимущества этих двух подходов и реализует их обобщение.

В данной работе мы доказываем теоретические оценки числа обусловленности обратных ковариационных матриц методов АРПФ и стандартного РПФ, что важно для систем с фиксированной точностью вычислений.

Теорема 6. Пусть $\mathbf{V}^{\mathbf{H}}\mathbf{V}(\varepsilon) = \mathbf{I} + O(\varepsilon)$, $\varepsilon \rightarrow 0$, и заданы матрицы $\mathbf{A} = \mathbf{V}^{\mathbf{H}}\mathbf{V}(\varepsilon) + \lambda\mathbf{S}^{-2} \rightarrow \mathbf{I} + \lambda\mathbf{S}^{-2}$ и $\mathbf{B} = \mathbf{S}\mathbf{V}^{\mathbf{H}}\mathbf{V}(\varepsilon)\mathbf{S} + \lambda\mathbf{I} \rightarrow \mathbf{S}^2 + \lambda\mathbf{I}$, которые инвертируются в соответствующих префильтрах $\mathbf{W}_{ARZF} = \mathbf{V}^{\mathbf{H}}\mathbf{A}^{-1}$, и $\mathbf{W}_{RZF} = \mathbf{S}\mathbf{V}^{\mathbf{H}}\mathbf{B}^{-1}$. Тогда числа обусловленности матриц \mathbf{A} и \mathbf{B} равны соответственно:

$$\chi(\mathbf{A}) = \frac{\lambda s_{\min}^{-2} + 1}{\lambda s_{\max}^{-2} + 1}, \quad \chi(\mathbf{B}) = \frac{\lambda + s_{\max}^2}{\lambda + s_{\min}^2} \quad (46)$$

и между ними выполняется соотношение:

$$1) \chi(\mathbf{A}) < \chi(\mathbf{B}), \text{ если } \lambda < s_{\min}^2 < s_{\max}^2,$$

2) $\chi(\mathbf{A}) > \chi(\mathbf{B})$, если $s_{\min}^2 < s_{\max}^2 < \lambda$,

где s_{\min} и s_{\max} — минимальные и максимальные элементы диагональной матрицы \mathbf{S} , а $\lambda = \frac{\sigma^2 L}{P}$.

Доказательство. Предположение о том, что матрица сингулярных векторов пользователей близка к унитарной, справедливо при отборе пользователей с низкой корреляцией: $\mathbf{V}^H \mathbf{V}(\varepsilon) = \mathbf{I} + O(\varepsilon)$, где $\varepsilon \rightarrow 0$. Тогда исследуемые матрицы приближаются к диагональным:

$$\mathbf{A} = \mathbf{V}^H \mathbf{V}(\varepsilon) + \lambda \mathbf{S}^{-2} = \mathbf{I} + O(\varepsilon) + \lambda \mathbf{S}^{-2} \rightarrow \mathbf{I} + \lambda \mathbf{S}^{-2}.$$

$$\mathbf{B} = \mathbf{S} \mathbf{V}^H \mathbf{V}(\varepsilon) \mathbf{S} + \lambda \mathbf{I} = \mathbf{S}(\mathbf{I} + O(\varepsilon))\mathbf{S} + \lambda \mathbf{I} = \mathbf{S}^2 + \lambda \mathbf{I} + O(\varepsilon)\mathbf{S}^2 \rightarrow \mathbf{S}^2 + \lambda \mathbf{I}.$$

Их функции обусловленности соответствуют выражению (46).

Сравнивая эти функции, можно выделить переходные режимы, в которых предпочтительнее использовать ту или иную формулу. Если $\lambda < s_{\min}^2 < s_{\max}^2$, то $\chi(\mathbf{A}) < \chi(\mathbf{B})$, и лучше использовать первую формулу. Если $s_{\min}^2 < s_{\max}^2 < \lambda$, то $\chi(\mathbf{A}) > \chi(\mathbf{B})$, и предпочтительна вторая. \square

Замечание 9. В случае $s_{\min}^2 < \lambda < s_{\max}^2$ нельзя сделать однозначный вывод — требуется дополнительный анализ. Однако отметим, что в реальных сетях используется только случай 1) $\lambda < s_{\min}^2 < s_{\max}^2$. Сингулярные векторы, чьи сингулярные значения меньше мощности шума, в системе не применяются, поэтому условие $\chi(\mathbf{A}) < \chi(\mathbf{B})$ всегда выполняется.

Математическая формулировка $\mathbf{W}_{ARZF} = \mathbf{V}^H \mathbf{A}^{-1}$ улучшает обусловленность системы при низком и среднем уровне шума λ , что делает алгоритм более точным при реализации в условиях ограниченной разрядной сетки по сравнению с другой формулировкой метода $\mathbf{W}_{RZF} = \mathbf{S} \mathbf{V}^H \mathbf{B}^{-1}$. Экспериментальное сравнение значения обусловленности показано на рис. 4.

2.3.2. Асимптотические свойства алгоритма АРПФ

В данном разделе получена асимптотическая оценка функции ОСИШ (16) для префильтра АРПФ (39) при допущении, что мощность шума на каждом устройстве значительно меньше мощности принимаемого сигнала.

Используя разложение по формуле Неймана [32], формулируем следующую лемму

Лемма 3 (Разложение обратной суммы матриц по малому параметру с остаточным членом в форме O). *Рассмотрим обратимые комплексные матрицы \mathbf{M} и \mathbf{n} одинакового размера и ранга. Для любого $0 < \varepsilon \ll 1$ и $\det \mathbf{M} \neq 0$ справедливо следующее тождество:*

$$(\mathbf{M} + \varepsilon \mathbf{n})^{-1} = \mathbf{M}^{-1} - \varepsilon \mathbf{M}^{-1} \mathbf{n} \mathbf{M}^{-1} + O(\varepsilon^2) = \mathbf{M}^{-1} + O(\varepsilon).$$

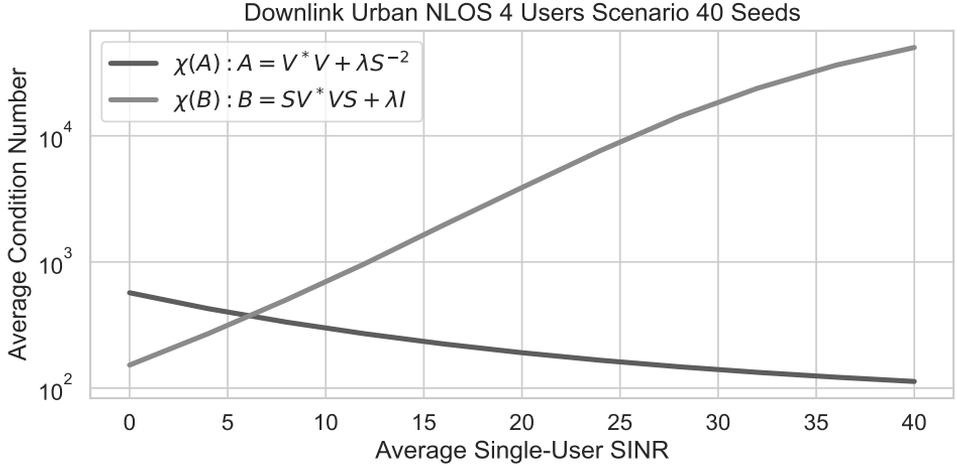


Рис. 4. Числа обусловленности $\chi(\mathbf{A})$ и $\chi(\mathbf{B})$ матриц \mathbf{A} и \mathbf{B} .

Доказательство. Рассмотрим разложение матричной функции $\mathbf{F}(\varepsilon)$, зависящей от вещественного параметра $\varepsilon > 0$:

$$\mathbf{F}(\varepsilon) = \mathbf{F}(0) + \mathbf{F}'(0)\varepsilon + \mathcal{O}(\varepsilon^2), \quad \text{где } \mathbf{F}(\varepsilon) = (\mathbf{M} + \varepsilon\mathbf{n})^{-1}.$$

Производная от обратной матрицы приведена в [33, 2.2]:

$$\mathbf{F}'(\varepsilon) = -(\mathbf{M} + \varepsilon\mathbf{n})^{-1}\mathbf{n}(\mathbf{M} + \varepsilon\mathbf{n})^{-1},$$

$$\Rightarrow \mathbf{F}'(0) = -\mathbf{M}^{-1}\mathbf{n}\mathbf{M}^{-1}$$

$$\Rightarrow \mathbf{F}(\varepsilon) = (\mathbf{M} + \varepsilon\mathbf{n})^{-1} = \mathbf{M}^{-1} - \varepsilon\mathbf{M}^{-1}\mathbf{n}\mathbf{M}^{-1} + \mathcal{O}(\varepsilon^2) = \mathbf{M}^{-1} + \mathcal{O}(\varepsilon).$$

□

Следующая теорема 7 описывает асимптотику АРПФ:

Теорема 7. Пусть матрица $\mathbf{V}\mathbf{V}^H$ имеет полный ранг, равный L . Тогда выполняются следующие асимптотики:

$$\begin{aligned} \mathbf{W}_{ARZF} &= \mathbf{V}^H ((\mathbf{V}\mathbf{V}^H)^{-1} - \lambda(\mathbf{V}\mathbf{V}^H)^{-1}\mathbf{S}^{-2}(\mathbf{V}\mathbf{V}^H)^{-1} + \mathcal{O}(\lambda^2)) = \\ &= \mathbf{W}_{ZF} - \lambda\mathbf{W}_{ZF}\mathbf{S}^{-2}(\mathbf{V}\mathbf{V}^H)^{-1} + \mathcal{O}(\lambda^2), \quad \text{при } \lambda \rightarrow 0+, \end{aligned} \quad (47)$$

$$\begin{aligned} \mathbf{W}_{ARZF} &= \lambda^{-1}\mathbf{V}^H (\mathbf{S}^2 - \lambda^{-1}\mathbf{S}^2(\mathbf{V}\mathbf{V}^H)\mathbf{S}^2 + \mathcal{O}(\lambda^{-2})) = \\ &= \lambda^{-1}\mathbf{W}_{MRT}\mathbf{S}^2 - \lambda^{-2}\mathbf{W}_{MRT}\mathbf{S}^2(\mathbf{V}\mathbf{V}^H)\mathbf{S}^2 + \mathcal{O}(\lambda^{-3}), \quad \text{при } \lambda \rightarrow +\infty. \end{aligned} \quad (48)$$

Доказательство. Для получения асимптотики при $\lambda \rightarrow 0+$ (47) используем Лемму 3 с $\mathbf{M} = \mathbf{V}\mathbf{V}^H$ и $\mathbf{n} = \mathbf{S}^{-2}$, а асимптотика при $\lambda \rightarrow +\infty$ (48) следует из применения Леммы 3 с $\mathbf{M} = \mathbf{S}^{-2}$, $\mathbf{n} = \mathbf{V}\mathbf{V}^H$, $\varepsilon = \lambda^{-1}$. \square

Свойство при малом шуме в Теореме 7 при $\lambda \rightarrow 0+$ означает, что алгоритм АРПФ стремится к ПФ. Другая асимптотика при $\lambda \rightarrow +\infty$ означает, что при преобладании шума над сигналом АРПФ обслуживает только UE с наилучшим каналом благодаря множителю \mathbf{S}^2 .

Лемма 4 (Модель системы с учётом префильтра АРПФ). *Для префильтра $\mathbf{W} = \mathbf{W}_{АРПФ}\mathbf{P}$, постфильтра $\mathbf{G} = \mathbf{P}^{-1}\mathbf{G}^C$ и матрицы корреляции $\mathbf{C} = \mathbf{V}\mathbf{V}^H - \mathbf{I} = \mathcal{O}(\lambda)$ при $\lambda \rightarrow 0$ для системы (1) справедливо:*

$$\mathbf{G}\mathbf{H}\mathbf{W} = \mathbf{I} - \lambda\mathbf{S}^{-2} + \mathcal{O}(\lambda^2). \quad (49)$$

Доказательство. Используем Формулу 47:

$$\begin{aligned} \mathbf{V}\mathbf{W}_{АРЗФ} &= \mathbf{V}\mathbf{V}^H(\mathbf{V}\mathbf{V}^H + \lambda\mathbf{S}^{-2})^{-1} = \mathbf{I} - \lambda\mathbf{S}^{-2}(\mathbf{V}\mathbf{V}^H)^{-1} + \mathcal{O}(\lambda^2) = \\ &= \mathbf{I} - \lambda(\mathbf{I} + \mathcal{O}(\lambda))^{-1} + \mathcal{O}(\lambda^2) = \mathbf{I} - \lambda\mathbf{S}^{-2} + \mathcal{O}(\lambda^2). \end{aligned} \quad (50)$$

$$\begin{aligned} \mathbf{G}\mathbf{H}\mathbf{W} &= \mathbf{P}^{-1}\underbrace{\mathbf{G}^C\mathbf{H}}_{=\mathbf{V}}\mathbf{W}_{АРЗФ}\mathbf{P} = \mathbf{P}^{-1}\mathbf{V}\mathbf{W}_{АРЗФ}\mathbf{P} = \\ &= \mathbf{P}^{-1}\mathbf{I}\mathbf{P} - \lambda\mathbf{P}^{-1}\mathbf{S}^{-2}\mathbf{P} + \mathcal{O}(\lambda^2) = \mathbf{I} - \lambda\mathbf{S}^{-2} + \mathcal{O}(\lambda^2). \end{aligned}$$

\square

Замечание 10. *В реальных сетях множество пользователей выбирается с помощью алгоритма планировщика, а число символов каждого пользователя – алгоритмом адаптации ранга [23]. Оба подхода обеспечивают выполнение $\mathbf{C} = \mathcal{O}(\lambda)$.*

Используя Лемму 4, получаем следующую теорему:

Теорема 8 (Функция ОСИШ с учётом префильтра АРПФ). *Для префильтра $\mathbf{W} = \mathbf{W}_{АРЗФ}\mathbf{P}$, постфильтра $\mathbf{G} = \mathbf{P}^{-1}\mathbf{G}^C$ и корреляционной матрицы $\mathbf{C} = \mathbf{V}\mathbf{V}^H - \mathbf{I} = \mathcal{O}(\lambda)$ функция ОСИШ (16) при условии, что мощность шума намного меньше мощности сигнала $\lambda = \frac{\sigma^2}{P} \rightarrow 0$ и при предположении $P \sim p_l, \forall l = 1, \dots, L$ имеет следующий асимптотический вид:*

$$\text{SINR}_l(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_l^C, \sigma^2) = \underbrace{\frac{p_l s_l^2}{\sigma^2}}_{\mathcal{O}(\lambda^{-1})} - 2 \underbrace{\frac{p_l}{P}}_{\mathcal{O}(1)} + \underbrace{\mathcal{O}\left(\frac{\sigma^2}{P}\right)}_{\mathcal{O}(\lambda)}, \quad \lambda \rightarrow 0. \quad (51)$$

Доказательство.

$$\begin{aligned}
 \text{SINR}_l(\mathbf{W}, \mathbf{H}_k, \mathbf{G}_l^C, \sigma^2) &:= \frac{1 - 2\lambda s_l^{-2} + \mathcal{O}(\lambda^2)}{\mathcal{O}(\lambda^4) + \frac{\sigma^2}{p_l s_l^2}} = \\
 &= \frac{p_l s_l^2}{\sigma^2} (1 - 2\lambda s_l^{-2} + \mathcal{O}(\lambda^2)) = \\
 &= \underbrace{\frac{p_l s_l^2}{\sigma^2}}_{\mathcal{O}(\lambda^{-1})} - 2 \underbrace{\frac{p_l}{P}}_{\mathcal{O}(1)} + \underbrace{\mathcal{O}\left(\frac{\sigma^2}{P}\right)}_{\mathcal{O}(\lambda)}.
 \end{aligned}$$

□

2.3.3. Оптимальная регуляризация на основе градиентного метода

Алгоритм: Оптимальная регуляризация префилтра \mathbf{R} (OPT)

Вход: Канал \mathbf{H} и его разложение $\mathbf{H} = \mathbf{U}^H \mathbf{S} \mathbf{V}$ по Лемме 1, мощность станции P , шум σ^2 , число итераций N

- 1) Определим функцию префилтра $\mathbf{W}(\mathbf{R})$ по формуле (54)
- 2) Определим функцию постфилтра $\mathbf{G}(\mathbf{R}) = \mathbf{G}^{MMSE}(\mathbf{W}(\mathbf{R}))$ по формуле (9)
- 3) Определим целевую функцию $J^{SE}(\mathbf{R}) = SE(\mathbf{W}(\mathbf{R}), \mathbf{H}, \mathbf{G}(\mathbf{R}), \sigma^2)$ из (16), (18), (21)
- 4) Задаём точность завершения по градиенту: $\varepsilon_g = 10^{-5}$
- 5) Задаём точность завершения по аргументу и значению функции: $\varepsilon_c = 10^{-9}$
- 6) Инициализируем регуляризацию: $\mathbf{R}_0 = \frac{L\sigma^2}{P} \mathbf{S}^{-2}$

7) Для $n = 0$ до $N - 1$ выполнять:

а) Вычисляем градиент $\nabla_{\mathbf{R}} J^{SE}(\mathbf{R})|_{\mathbf{R}=\mathbf{R}_n}$

б) Если выполняются условия сходимости по ε_g или ε_c :

– Вернуть $\mathbf{W}_{OPT} = \mathbf{W}(\mathbf{R}_n)$

в) Иначе:

– Вычисляем направление L-BFGS [34] по градиенту: $\mathbf{D}_n = \mathbf{D}(\nabla_{\mathbf{R}} J^{SE}(\mathbf{R})|_{\mathbf{R}=\mathbf{R}_n})$

– Оптимизируем длину шага: $\alpha_n = \arg \max_{\alpha} J^{SE}(\mathbf{R}_n + \alpha \mathbf{D}_n)$

– Обновляем регуляризацию: $\mathbf{R}_{n+1} = \mathbf{R}_n + \alpha_n \mathbf{D}_n$

8) Вернуть $\mathbf{W}_{OPT} = \mathbf{W}(\mathbf{R}_N)$

Таблица 2. Оптимальная регуляризация префилтра \mathbf{R} (OPT)

Рассматриваемый алгоритм минимизирует квадратичную оптимизационную задачу (42), однако остаётся вопрос: насколько он хорош относительно функции суммарной спектральной эффективности (21)?

В работе [5] доказано, что оптимальное решение задачи $\max f(SINR_1, \dots, SINR_K)$ при ограничении по суммарной мощности имеет форму:

$$\mathbf{W}_{OPT}(\mathbf{V}) = \mu \mathbf{V}^H (\mathbf{V} \mathbf{V}^H + \mathbf{r})^{-1}, \quad \mathbf{r} = \text{diag}(r_1, \dots, r_L). \quad (52)$$

Хотя аналитически определить конкретные значения \mathbf{r} затруднительно, структура решения остаётся полезной.

В связи с этим мы сравниваем эффективность алгоритма АРПФ с градиентным поиском по суммарной функции *спектральной эффективности* (19). Для этого формулируется итеративное решение с дифференцируемыми вложенными функциями. Предлагаемое параметрическое решение сохраняет структуру базового алгоритма РПФ, но оптимизирует целевую функцию *спектральной эффективности*, что приводит к заметному улучшению качества. Мы решаем задачу гладкой оптимизации с ограничениями: требуется найти локальный максимум спектральной эффективности (21):

$$\begin{aligned} & \underset{\mathbf{R}=\text{diag}\{r_1, \dots, r_L\}}{\text{максимизировать}} \quad SE(\mathbf{W}(\mathbf{R}), \mathbf{H}, \mathbf{G}^{\text{MMSE}}(\mathbf{W}), \sigma^2) \\ & \text{при условиях} \quad |(w_{t1}, \dots, w_{tL})|^2 \leq \frac{P}{T}, \quad t = 1, \dots, T \end{aligned} \quad (53)$$

Параметрическое решение использует формулу РПФ в следующем виде:

$$\mathbf{W}(\mathbf{V}, \mathbf{r}) = \mu(\widehat{\mathbf{W}})\widehat{\mathbf{W}}(\mathbf{V}, \mathbf{r}), \quad \widehat{\mathbf{W}}(\mathbf{r}) = \mathbf{V}^H(\mathbf{V}\mathbf{V}^H + \mathbf{r})^{-1}, \quad (54)$$

$$\mu(\widehat{\mathbf{W}}) = \frac{\sqrt{P/T}}{\max_t \|(\hat{w}_{t1}, \dots, \hat{w}_{tL})\|} \quad (55)$$

Ограничение на максимальную мощность антенн реализуется через масштабирование матрицы префильтра скаляром μ , что позволяет удовлетворить ограничению мощности, сохранив геометрию и желаемые свойства префильтра.

В дальнейших экспериментах будет оптимизироваться дифференцируемая вещественная диагональная матрица $\mathbf{r} \in \mathbb{R}^{L \times L}$, что составляет одно из основных достижений работы. Сам префильтр \mathbf{W} изучен в соответствующей статье [35].

Процесс постфильтра участвует в расчёте градиента и, следовательно, является неотъемлемой частью спектральной эффективности. Дифференцируемыми переменными выступают диагональные элементы матрицы регуляризации, по которым строится префильтр, затем вычисляется постфильтр. Оба результата используются при расчёте градиента. Регуляризация участвует во всех этих операциях как внутренняя переменная сложной композиционной функции, и её градиент может быть найден по правилу дифференцирования сложной функции с использованием алгоритма обратного распространения ошибки, как это реализовано, например, в библиотеке PyTorch.

Замечание 11. *В рассмотренной градиентной оптимизации префильтр строится с учётом конкретного постфильтра (в данном случае постфильтра среднеквадратичной ошибки), который предполагается реализуемым на стороне пользователя. Эта идея активно обсуждается в современной литературе (см., например, [10]) и может быть применена для улучшения практически любой схемы префильтра с помощью итеративной процедуры.*

3. Результаты моделирования и обсуждение

3.1. Настройка в “Quadriga”

В этом разделе описано, как получены данные с помощью “Quadriga” — открытого ПО для генерации реалистичного радиоканала. Рассматривается сценарий Urban Non-Line-of-Sight (3GPP_38.901_RMa_NLOS) [22].

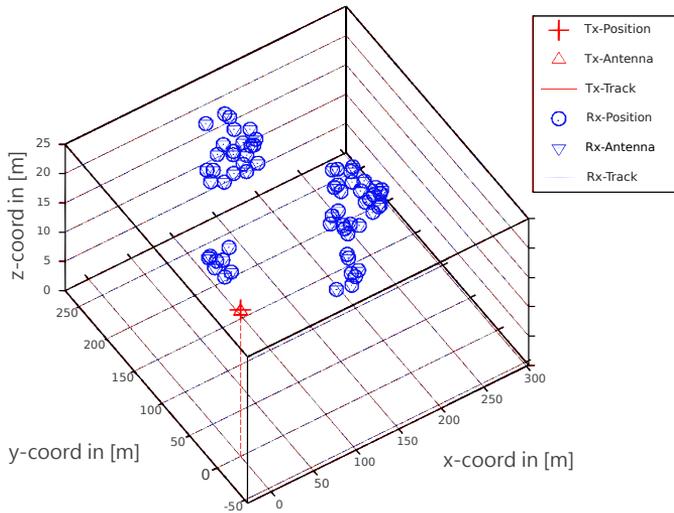


Рис. 5. Пример случайной генерации пользователей в городской среде.

Параметр	Значение
Базовая станция	
Число станций	1
Координаты (x,y,z), м	(0, 0, 25)
Антенн по оси y/z	8 / 4
Расст. между антеннами (y/z)	0.5λ / 1.7λ
Модель антенны	3gpp-macro
Ширина луча (азимут/место), град	60 / 10
Отн. перед/зад лепестков, дБ	20
Всего антенн	64
Приёмник	
Антенн по оси x	2
Расст. между антеннами (x)	0.5λ
Модель антенны	полуволновый диполь
Всего антенн	4
Симуляция Quadriga	
Центр. частота	3,5 ГГц
Выборка на метр	1
Задержка прямого пути	1
Без сферических волн	1
Генератор каналов	
Затухание в тени (sigma)	0
Деление кластеров	Ложь
Полоса пропускания	100 МГц
Число поднесущих	42

Таблица 3. Параметры генерации канала в “Quadriga”

На рис. 5 показан пример расположения пользователей: они размещаются либо в кластере одного из зданий, либо на земле рядом с ним.

Пользователи обозначены синими кружками, а базовая станция — красным. Расстояния между антеннами станции и пользователями малые по сравнению с расстояниями между станцией и пользователями, поэтому они изображены отдельными кружками, каждый содержащим несколько антенн.

Общий алгоритм моделирования для каждого случайного зерна:

- 1) Генерируем случайную среду вокруг базовой станции;
- 2) Случайно размещаем пользователей рядом с базовой станцией;
- 3) Выбираем релевантных пользователей по корреляции.

Далее процедура описана более подробно.

Во-первых, фиксируем положение базовой станции в точке $[0, 0, 25]$.

Во-вторых, выбираем случайные позиции пользователей вокруг станции:

- 1) Сэмплируем до 8 центров кластеров (x_c, y_c) в секторе 120° радиусом до 2000 м от станции. Каждый кластер соответствует части здания;
- 2) Задаём случайную высоту кластера $z_c = 1.5 + (3 \cdot U(\{1, \dots, 10\}) - 1)$ м, выбирая этаж равномерно;
- 3) Для каждого пользователя назначаем кластер $c(u)$ и сэмплируем (x_u, y_u) внутри окружности радиуса 60 м вокруг центра кластера;
- 4) Сэмплируем высоту пользователя: 80% размещаются около этажа кластера $z_u = z_{c(u)} + 3U(\{-1, 0, 1\})$, и 20% — на улице $z_u = 1.5$ м.

В-третьих, сгенерировав матрицы канала для $K_{\max} = 64$ пользователей, выполняем отбор, моделируя работу планировщика. Выбираем $K < K_{\max}$ пользователей, у которых корреляция не слишком высокая: $\text{corr}_{i,j} = |\mathbf{V}_{i,1}^H \mathbf{V}_{j,1}|^2 \leq 0.3$. Число потоков для каждого пользователя $L_k = 2$ (политика выбора ранга).

Рассматриваем два сценария:

- 1) Разные потери мощности на пути (path loss, PL): введён случайный множитель $\text{PL} \in [-10 \text{ дБ}, +10 \text{ дБ}]$ — диапазон, характерный для близких пользователей (в реальности разброс внутри базовой станции до 60 дБ);
- 2) Почти равные потери мощности на пути: $s_{i,1} \sim s_{j,1}$ (это поведение по умолчанию у “Quadriga”).

Результаты усреднены по 40 моделям каналов и выборкам пользователей:

- $\mathbf{H} \in \mathbb{C}^{K \times R_k \times T}$;
- $T = 64$ антенн базовой станции;
- $K = 4$ пользователей;
- $R = 16$ антенн на стороне пользователя;
- $L = 8$ передаваемых символов.

Несущая частота выбирается случайно в пределах полосы. Антенная решётка базовой станции: 8×4 (8 по оси по оси x и 4 по оси y), антенная решётка на пользователе: 2×1 . Каждая позиция антенной решётки имеет две кросс-поляризованные антенны. Алгоритм генерации даёт реалистичную городскую среду. Подробные параметры приведены в Табл. 3 приложения.

3.2. Результаты

Сравниваем адаптивный регуляризованный псевдообратный префильтр с опорными алгоритмами (сопряжённый префильтр (MRT), псевдообратный префильтр (ZF), регуляризованный псевдообратный префильтр (RZF), регуляризованный псевдообратный префильтр Винера (WRZF)) и оптимальным решением ОПТ. На рис. 6, табл. 4 и рис. 7, табл. 5 представлены средняя (J_SE) и минимальная спектральная эффективность для сценария с разными потерями мощности на путях. Аналогичные данные для сценария с равными потерями мощности на путях — на рис. 8/табл. 6 и рис. 9/табл. 7.

Префильтр SU SINR	$W_{MRT}(V)$	$W_{ZF}(V)$	$W_{ZF}(F)$	$W_{RZF}(V)$	$W_{RZF}(F)$	$W_{WRZF}(V)$	$W_{ARZF}(V)$	$W_{OPT}(V)$
0	1.45	2.27	0.02	2.37	2.46	1.48	3.91	4.07
4	1.75	3.13	0.05	3.19	2.79	1.84	4.91	5.10
8	2.08	4.16	0.11	4.19	3.10	2.33	6.03	6.32
12	2.41	5.29	0.26	5.31	3.37	3.02	7.17	7.63
16	2.71	6.51	0.54	6.52	3.61	4.02	8.23	9.01
20	2.99	7.79	1.03	7.79	3.66	5.48	9.26	10.42
24	3.24	9.05	1.75	9.05	3.76	7.29	10.23	11.74
28	3.47	10.40	2.77	10.40	4.08	9.58	11.24	13.09
32	3.65	11.65	3.95	11.65	4.73	11.60	12.19	14.17
36	3.76	12.83	5.16	12.83	5.55	13.10	13.13	15.24
40	3.83	13.90	6.34	13.90	6.52	14.21	14.05	16.17

Таблица 4. Таблица соответствует рис. 6 и показывает среднюю спектральную эффективность (SE) различных префильтров в *городском многолучевом (NLOS) сценарии с разными потерями пути*. Исследуемый алгоритм — $\mathbf{W}_{ARZF}(\mathbf{V})$. Оптимальная регуляризация $\mathbf{W}_{OPT}(\mathbf{V})$ была получена с помощью оптимизации методом L-BFGS.

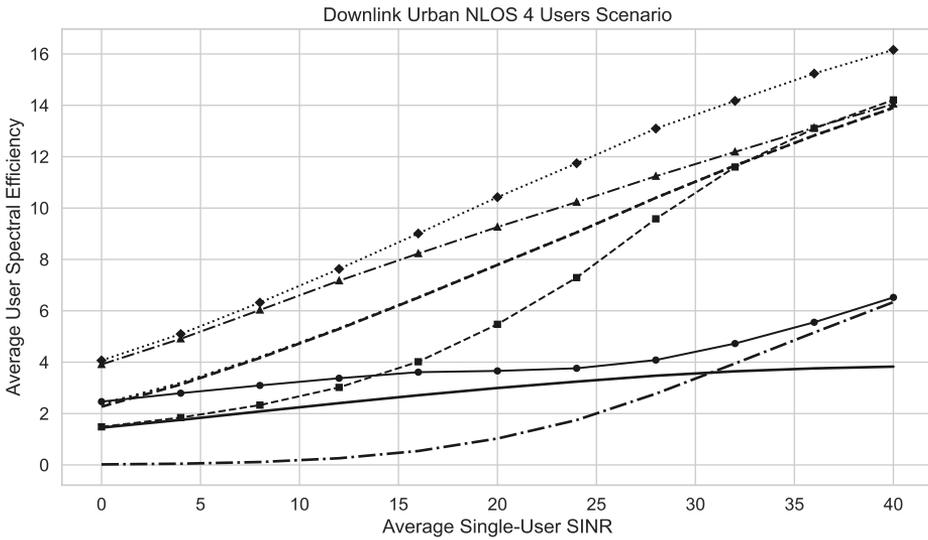


Рис. 6. Средняя спектральная эффективность (SE) различных префильтров в городской многолучевом (*NLOS*) сценарии с разными потерями пути. Зелёная линия совпадает с жёлтой. Матрица $\mathbf{F} = \mathbf{S}\mathbf{V}$. Значения приведены в табл. 4.

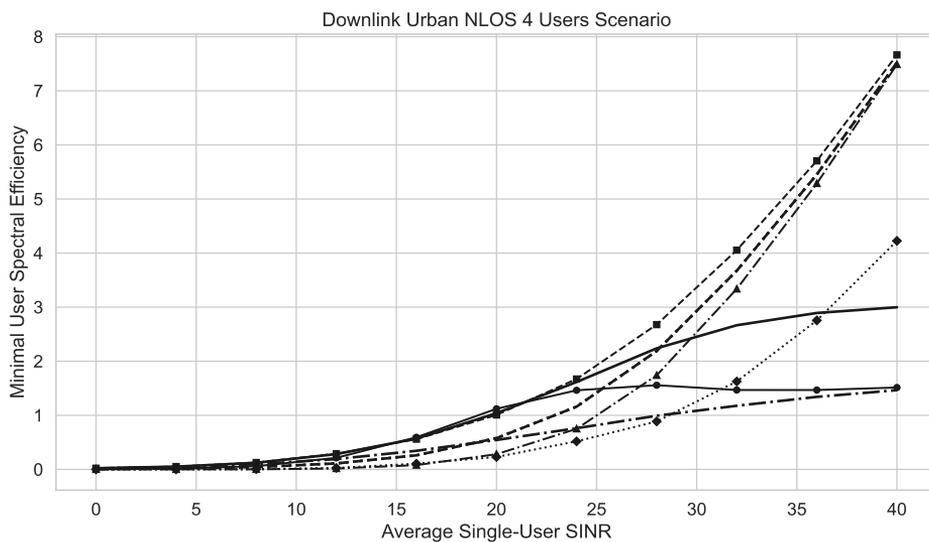


Рис. 7. Минимальная спектральная эффективность (SE) среди пользователей для различных префильтров в *городском многолучевом (NLOS) сценарии с разными потерями пути*. Зелёная линия совпадает с жёлтой. Матрица $\mathbf{F} = \mathbf{S}\mathbf{V}$. Значения приведены в табл. 5.

Префильтр SU SINR	$W_{MRT}(V)$	$W_{ZF}(V)$	$W_{ZF}(F)$	$W_{RZF}(V)$	$W_{RZF}(F)$	$W_{WRZF}(V)$	$W_{ARZF}(V)$	$W_{OPT}(V)$
0	0.02	0.01	0.02	0.01	0.00	0.02	0.00	0.00
4	0.05	0.02	0.04	0.02	0.01	0.05	0.00	0.00
8	0.12	0.04	0.10	0.04	0.06	0.13	0.00	0.00
12	0.28	0.11	0.19	0.11	0.21	0.29	0.02	0.03
16	0.57	0.26	0.34	0.26	0.59	0.56	0.08	0.11
20	1.04	0.58	0.54	0.58	1.12	1.01	0.28	0.23
24	1.62	1.16	0.76	1.16	1.46	1.67	0.75	0.52
28	2.24	2.19	0.99	2.19	1.56	2.68	1.74	0.89
32	2.67	3.67	1.18	3.67	1.47	4.05	3.34	1.63
36	2.89	5.45	1.34	5.45	1.47	5.70	5.29	2.74
40	3.00	7.52	1.47	7.52	1.51	7.67	7.49	4.25

Таблица 5. Таблица соответствует рис. 7 и показывает минимальную спектральную эффективность (SE) различных префильтров в *городском многолучевом (NLOS) сценарии с разными потерями пути*. Исследуемый алгоритм — $\mathbf{W}_{ARZF}(\mathbf{V})$. Оптимальная регуляризация $\mathbf{W}_{OPT}(\mathbf{V})$ была получена с помощью оптимизации методом L-BFGS.

Адаптивный регуляризованный псевдообратный префильтр показывает лучшую среднюю спектральную эффективность: адаптивная регуляризация учитывает потери мощности на пути и порядок сингулярных значений каждого пользователя. В диапазоне высоких одномользовательских ОСИШ псевдообратный префильтр лучше сопряжённого префильтра. При равных потерях мощности на путях спектральная эффективность АРПФ примерно равна спектральной эффективности префильтру Винера; при разных потерях мощности на путях префильтр Винера заметно хуже, а АРПФ выигрывает.

ОРТ (чёрная линия) — градиентный поиск — даёт лучшие результаты, но слишком затратен вычислительно. Тем не менее он полезен как верхняя оценка: он показывает потенциал улучшения АРПФ. Анализ минимальной спектральной эффективности показывает, что АРПФ улучшает среднюю спектральную эффективность, но слабее для самых слабых пользователей, особенно при низком однопользовательском ОСИШ. Такая разница между средней и минимальной сапектральной эффективностью известна [12].

На рис. 8/табл. 6 и рис. 9/табл. 7 представлены аналогичные результаты для равных потерь мощности на пути : АРПФ превосходит по средней спектральной эффективности, но не по минимальной; при этом он совпадает с префильтром Винера, когда потери мощности на пути одинаковые.

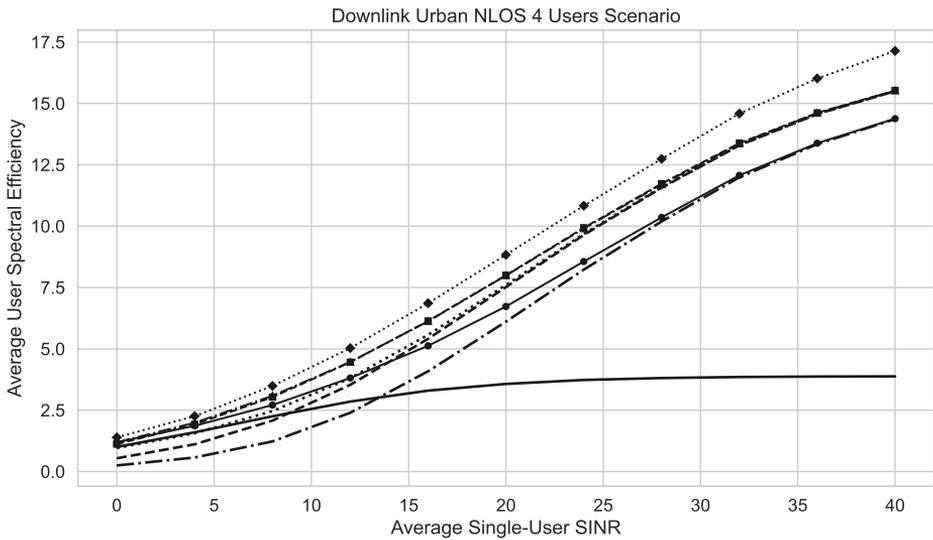


Рис. 8. Средняя спектральная эффективность (SE) различных префильтров в городском многолучевом (NLOS) сценарии с равными потерями пути (см. табл. 6). Красная линия $\mathbf{W}_{ARZF}(\mathbf{V})$ совпадает с синей линией $\mathbf{W}_{WRZF}(\mathbf{V})$, что говорит о равенстве работы алгоритмов при одинаковых потерях.

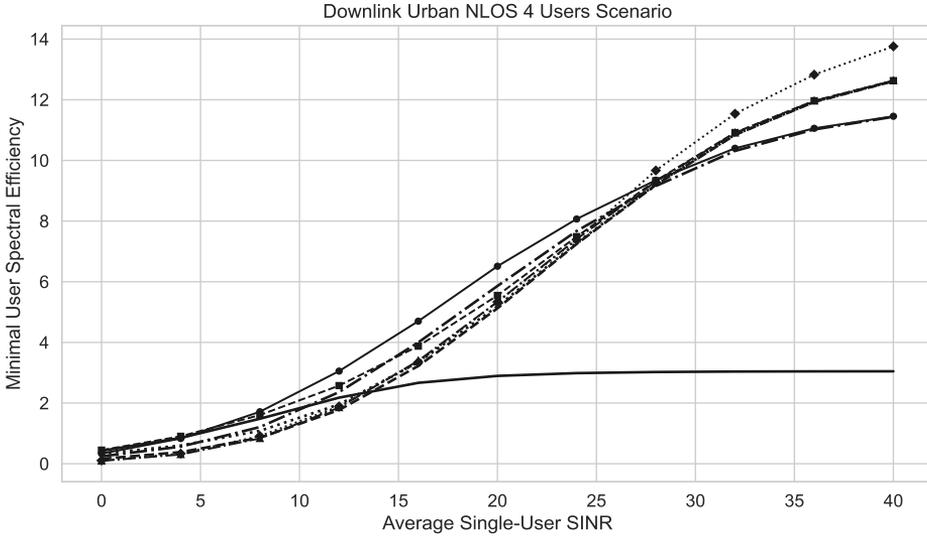


Рис. 9. Минимальная спектральная эффективность (SE) среди пользователей для различных префильтров в *городском многолучевом (NLOS) сценарии с равными потерями пути* (см. табл. 7).

Префильтр $SUSINR$	$W_{MRT}(V)$	$W_{ZF}(V)$	$W_{ZF}(F)$	$W_{RZF}(V)$	$W_{RZF}(F)$	$W_{WRZF}(V)$	$W_{ARZF}(V)$	$W_{OPT}(V)$
0	1.00	0.54	0.25	0.96	1.19	1.12	1.19	1.39
4	1.60	1.11	0.57	1.56	1.86	1.92	1.99	2.26
8	2.25	2.08	1.23	2.46	2.71	3.03	3.09	3.49
12	2.85	3.54	2.40	3.81	3.81	4.46	4.47	5.03
16	3.30	5.41	4.09	5.57	5.12	6.14	6.12	6.86
20	3.57	7.52	6.11	7.60	6.73	8.00	7.99	8.83
24	3.73	9.64	8.23	9.69	8.56	9.94	9.90	10.83
28	3.81	11.56	10.19	11.59	10.36	11.73	11.70	12.74
32	3.85	13.29	11.99	13.31	12.07	13.38	13.36	14.59
36	3.87	14.57	13.34	14.58	13.38	14.62	14.60	16.02
40	3.88	15.50	14.36	15.51	14.38	15.53	15.52	17.15

Таблица 6. Таблица соответствует рис. 8 и показывает среднюю спектральную эффективность различных префильтров в *городском многолучевом (NLOS) сценарии*. Исследуемый алгоритм — $W_{ARZF}(V)$. Оптимальная регуляризация $W_{OPT}(V)$ была настроена с помощью алгоритма оптимизации L-BFGS.

Префильтр <i>SUSINR</i>	$W_{MRT}(V)$	$W_{ZF}(V)$	$W_{ZF}(F)$	$W_{RZF}(V)$	$W_{RZF}(F)$	$W_{WRZF}(V)$	$W_{ARZF}(V)$	$W_{OPT}(V)$
0	0.42	0.16	0.24	0.33	0.33	0.45	0.09	0.11
4	0.85	0.38	0.56	0.60	0.83	0.91	0.30	0.33
8	1.48	0.86	1.21	1.08	1.71	1.60	0.83	0.91
12	2.18	1.77	2.36	1.96	3.05	2.58	1.85	1.89
16	2.67	3.22	4.00	3.34	4.70	3.88	3.40	3.35
20	2.90	5.12	5.86	5.19	6.51	5.55	5.34	5.33
24	2.99	7.25	7.67	7.28	8.07	7.49	7.41	7.38
28	3.02	9.19	9.16	9.21	9.35	9.33	9.30	9.67
32	3.04	10.84	10.31	10.85	10.40	10.92	10.91	11.54
36	3.04	11.93	11.02	11.93	11.06	11.97	11.96	12.83
40	3.05	12.61	11.44	12.62	11.45	12.63	12.63	13.74

Таблица 7. Таблица соответствует рис. 9 и показывает минимальную спектральную эффективность различных префильтров в *городском многолучевом (NLOS) сценарии*. Исследуемый алгоритм — $\mathbf{W}_{ARZF}(\mathbf{V})$. Оптимальная регуляризация $\mathbf{W}_{OPT}(\mathbf{V})$ была настроена с помощью алгоритма оптимизации L-BFGS.

4. Заключение

Ключевая задача статьи – оптимизация метода построения префильтра в многопользовательской системе с множеством антенн, в том числе на каждом приемном устройстве. В статье мы анализируем производительность различных линейных методов (сопряжённый префильтр, псевдообратный префильтр, регуляризованный псевдообратный префильтр, регуляризованный псевдообратный префильтр Винера) и вводим простую эвристическую модель адаптивного регуляризованного псевдообратного префильтра. Аналитически изучаем связь адаптивного регуляризованного псевдообратного префильтра с регуляризованным псевдообратным префильтром, а также их асимптотическое соотношение. Получаем оценку ОСИШ при низком шуме. Проведенное тестирование на реалистичном канале в системе “Quadriga” показывает стабильное улучшение средней (по принимающим устройствам) спектральной эффективности по сравнению с эталонными методами. Важно отметить, что при этом минимальная спектральная эффективность (рассчитываемая на подмножестве принимающих устройств с самым плохим априорным ОСШ) практически не меняется. Результаты для предложенного префильтра также сравниваются с нелинейным методом ОРТ, который градиентным алгоритмом оптимизирует целевую функцию и таким образом дает хорошую оценку для доступной верхней границы решения.

Доступность данных и материалов

Матрицы каналов, сгенерированные в “Quadriga”, и полный Python-код доступны на GitHub: <https://github.com/eugenbobrov/Adaptive-Regularized-Zero-Forcing-Beamforming-in-Massive-MIMO-with-Multi-Antenna-Users>.

Adaptive Regularized Zero-Forcing Beamforming in Massive MIMO with Multi-Antenna Users Bobrov E.A., Minenkov D.S., Yudakov D.A.

Modern cellular networks utilize massive MIMO technology with multiple antennas. This work investigates the adaptive regularized Zero-Forcing method employing special regularization based on SVD. We conduct theoretical analysis, performance evaluation, and comparison with other methods through simulations using the Quadriga channel model.

Keywords: Telecommunications, MIMO, optimization, singular value decomposition (SVD), signal-to-interference-and-noise ratio (SINR), spectral efficiency

References

- [1] Ngo, Hien Quoc and Larsson, Erik G and Marzetta, Thomas L, “Energy and spectral efficiency of very large multiuser MIMO systems”, *IEEE Transactions on Communications*, **61**:4 (2013), 1436–1449.
- [2] Andrews, Jeffrey G and Buzzi, Stefano and Choi, Wan and Hanly, Stephen V and Lozano, Angel and Soong, Anthony CK and Zhang, Jianzhong Charlie, “What will 5G be?”, *IEEE Journal on selected areas in communications*, **32**:6 (2014), 1065–1082.
- [3] Parfait, Tebe, Kuang, Yujun, Jerry, Kponyo, “Performance analysis and comparison of ZF and MRT based downlink massive MIMO systems”, *2014 Sixth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2014, 383–388.
- [4] Joham, Michael, Utschick, Wolfgang, Nosssek, Josef A., “Linear transmit processing in MIMO communications systems”, *IEEE Transactions on Signal Processing*, **53**:8 (2005), 2700–2712.
- [5] Björnson, Emil, Bengtsson, Mats, Ottersten, Björn, “Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]”, *IEEE Signal Processing Magazine*, **31**:4 (2014), 142–148.

- [6] Nguyen, Long D., Tuan, Hoang Duong, Duong, Trung Q., Poor, H. Vincent, “Multi-user regularized zero-forcing beamforming”, *IEEE Transactions on Signal Processing*, **67**:11 (2019), 2839–2853.
- [7] Zhang, Jiankang, Chen, Sheng, Maunder, Robert G., Zhang, Rong, Hanzo, Lajos, “Regularized zero-forcing precoding-aided adaptive coding and modulation for large-scale antenna array-based air-to-air communications”, *IEEE Journal on Selected Areas in Communications*, **36**:9 (2018), 2087–2103.
- [8] Peel, Christian B., Hochwald, Bertrand M., Swindlehurst, A. Lee, “A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization”, *IEEE Transactions on Communications*, **53**:1 (2005), 195–202.
- [9] Jaeckel, Stephan, Raschkowski, Leszek, Börner, Kai, Thiele, Lars, “QuADRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials”, *IEEE Transactions on Antennas and Propagation*, **62**:6 (2014), 3242–3256.
- [10] Shi, Shuying, Schubert, Martin, Boche, Holger, “Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization”, *IEEE Transactions on Signal Processing*, **55**:11 (2007), 5436–5446.
- [11] Caire, Giuseppe, Shamai, Shlomo, “On the achievable throughput of a multiantenna Gaussian broadcast channel”, *IEEE Transactions on Information Theory*, **49**:7 (2003), 1691–1706.
- [12] Björnson, Emil and Hoydis, Jakob and Sanguinetti, Luca, “Massive MIMO networks: Spectral, energy, and hardware efficiency”, *Foundations and Trends in Signal Processing*, **11**:3-4 (2017), 154–655.
- [13] Tran, Le-Nam and Juntti, Markku and Bengtsson, Mats and Ottersten, Bjorn, “Beamformer designs for MISO broadcast channels with zero-forcing dirty paper coding”, *IEEE transactions on wireless communications*, **12**:3 (2013), 1173–1185.
- [14] Fatema, Nusrat and Hua, Guang and Xiang, Yong and Peng, Dezhong and Natgunanathan, Iynkaran, “Massive MIMO linear precoding: A survey”, *IEEE systems journal*, **12**:4 (2017), 3920–3931.
- [15] Dhakal, Sunil, “High rate signal processing schemes for correlated channels in 5G networks”, 2019.

- [16] Bogale, Tadilo Endeshaw, Vandendorpe, Luc, “Sum MSE optimization for downlink multiuser MIMO systems with per antenna power constraint: Downlink-uplink duality approach”, *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, 2011, 2035–2039.
- [17] Wiesel, Ami, Eldar, Yonina C., Shamai, Shlomo, “Zero-forcing precoding and generalized inverses”, *IEEE Transactions on Signal Processing*, **56:9** (2008), 4409–4418.
- [18] Hoydis, Jakob and Ten Brink, Stephan and Debbah, M erouane, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?”, *IEEE Journal on selected Areas in Communications*, **31:2** (2013), 160–171.
- [19] Bj ornson, Emil and Jorswieck, Eduard and Ottersten, Bjorn, “Impact of spatial correlation and precoding design in OSTBC MIMO systems”, *IEEE Transactions on Wireless Communications*, **9:11** (2010), 3578–3589.
- [20] Sun, Liang, McKay, Matthew R., “Eigen-based transceivers for the MIMO broadcast channel with semi-orthogonal user selection”, *IEEE Transactions on Signal Processing*, **58:10** (2010), 5246–5261.
- [21] Nguyen, Duy HN, Le-Ngoc, Tho, “MMSE precoding for multiuser MISO downlink transmission with non-homogeneous user SNR conditions”, *EURASIP Journal on Advances in Signal Processing*, **2014:1** (2014), 1–12.
- [22] Tse, David, Viswanath, Pramod, “Fundamentals of wireless communication”, 2005.
- [23] Mahmood, Nurul H., Berardinelli, Gilberto, Tavares, Fernando ML, Lauridsen, Mads, Mogensen, Preben, Pajukoski, Kari, “An efficient rank adaptation algorithm for cellular MIMO systems with IRC receivers”, *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, 2014, 1–5.
- [24] Bobrov, Evgeny, Chinyaev, Boris, Kuznetsov, Viktor, Minenkov, Dmitrii, Yudakov, Daniil, “Power allocation algorithms for massive MIMO systems with multi-antenna users”, *Wireless Networks*, 2023, 1–22.
- [25] Ren, Bin, Wang, Yingmin, Sun, Shaohui, Zhang, Yawen, Dai, Xiaoming, Niu, Kai, “Low-complexity MMSE-IRC algorithm for uplink massive MIMO systems”, *Electronics Letters*, **53:14** (2017), 972–974.

- [26] Mehana, Ahmed Hesham, Nosratinia, Aria, “Diversity of MMSE MIMO receivers”, *IEEE Transactions on Information Theory*, **58**:11 (2012), 6788–6805.
- [27] Lagen, Sandra, Wanuga, Kevin, Elkotby, Hussain, Goyal, Sanjay, Patriciello, Natale, Giupponi, Lorenza, “New radio physical layer abstraction for system-level simulations of 5G networks”, *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, 2020, 1–7.
- [28] Kuhn, Harold W., Tucker, Albert W., “Nonlinear programming”, *Traces and Emergence of Nonlinear Programming*, 2014, 247–258.
- [29] Koopmans, Tjalling, “Activity analysis of production and allocation”, 1951.
- [30] Tran, Le-Nam, “An iterative precoder design for successive zero-forcing precoded systems”, *IEEE Communications Letters*, **16**:1 (2011), 16–18.
- [31] Pearson, Karl, “On lines and planes of closest fit to systems of points in space”, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **2**:11 (1901), 559–572.
- [32] Zhu, Dengkui, Li, Boyu, Liang, Ping, “On the matrix inversion approximation based on Neumann series in massive MIMO systems”, *2015 IEEE International Conference on Communications (ICC)*, 2015, 1763–1769.
- [33] Petersen, K. B., Pedersen, M. S., “The Matrix Cookbook”, 2008 <http://www2.imm.dtu.dk/pubdb/p.php?3274>.
- [34] Liu, Dong C., Nocedal, Jorge, “On the limited memory BFGS method for large scale optimization”, *Mathematical Programming*, **45**:1 (1989), 503–528.
- [35] Bobrov, Evgeny, Kropotov, Dmitry, Troshin, Sergey, ZaeV, Danila, *arXiv:2107.13440*, 2021.