

УДК: 519.8

Компенсация собственных нелинейных помех на основе смешанного метода Ньютона

А. А. Дегтярев^a, С. А. Бахурин^b

МФТИ,

Россия, 141701, Московская обл., г. Долгопрудный, Институтский пер. 9

Университет Иннополис,

Россия, 420500, г. Иннополис, ул. Университетская, д.1

E-mail: ^a degtyarev.aa@phystech.edu, ^b bakhurin.sa@mpt.ru

Получено 27.10.2024.

Принято к публикации 01.06.2016.

В статье исследуется одно из возможных решений задачи компенсации собственных помех (SIC, Self-Interference Cancellation), возникающей при проектировании полнодуплексных (IBFD, In-band Full-Duplex) систем связи. Подавление собственных помех осуществляется в цифровой области с помощью многослойных нелинейных моделей, которые адаптируются на основе метода градиентного спуска. Наличие локальных оптимумов и седловых точек при адаптации многослойных моделей делает невозможным использование методов второго порядка ввиду знаконеопределенности матрицы Гессе.

В данной работе предложено использовать смешанный метод Ньютона (MNM, mixed Newton method), который учитывает информацию о смешанных производных второго порядка функции потерь и, как следствие, обеспечивает высокую скорость сходимости по сравнению с традиционными методами первого порядка. Использование лишь только смешанных частных производных второго порядка при построении матрицы Гессе позволяет избежать проблемы "застревания" в седловых точках при использовании смешанного метода Ньютона для адаптации многослойных нелинейных компенсаторов собственных помех при проектировании полнодуплексных систем связи.

В качестве модели собственных нелинейных помех выбрана модель Гаммерштейна с комплексными параметрами. Данный выбор обусловлен тем, что модель эффективно описывает физические свойства, лежащие в основе формирования собственных помех. Благодаря свойству голоморфности выхода модели смешанный метод Ньютона обеспечивает свойство «отталкивания» от седловых точек в ландшафте функции потерь.

В работе приводятся кривые сходимости при адаптации модели Гаммерштейна смешанным методом Ньютона, а также при помощи классических подходов на основе метода градиентного спуска. Кроме того, приводится вывод предложенного метода, а также оценка вычислительной сложности.

Ключевые слова: метод второго порядка, комплекснозначный гессиан, полнодуплексные системы связи, компенсация собственных помех

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (Госзадание), проект No. FSMG-2024-0011

© 2025 Александр Андреевич Дегтярев, Сергей Алексеевич Бахурин
Статья доступна по лицензии Creative Commons Attribution-NoDerivs 3.0 Unported License. Чтобы получить текст
лицензии, посетите вебсайт <http://creativecommons.org/licenses/by-nd/3.0/> или отправьте письмо в Creative
Commons, PO Box 1866, Mountain View, CA 94042, USA.

UDC: 519.8

Non-linear self-interference cancellation on base of mixed Newton method

A. A. Degtyarev^a, S. A. Bakhurin^b

MIPT,

Russian Federation, 141701, Moscow Region, Dolgoprudny, Institutskiy Lane 9

Innopolis University,

Russian Federation, 420500, Innopolis, st. Universitetskaya 1

E-mail: ^a degtyarev.aa@phystech.edu, ^b bakhurin.sa@mipt.ru

Received 27.10.2024.

Accepted for publication 01.06.2016.

The paper investigates a potential solution to the problem of Self-Interference Cancellation (SIC) encountered in the design of In-Band Full-Duplex (IBFD) communication systems. The suppression of self-interference is implemented in the digital domain using multilayer nonlinear models adapted via the gradient descent method. The presence of local optima and saddle points in the adaptation of multilayer models prevents the use of second-order methods due to the indefinite nature of the Hessian matrix.

This work proposes the use of the Mixed Newton Method (MNM), which incorporates information about the second-order mixed partial derivatives of the loss function, thereby enabling a faster convergence rate compared to traditional first-order methods. By constructing the Hessian matrix solely with mixed second-order partial derivatives, this approach mitigates the issue of "getting stuck" at saddle points when applying the Mixed Newton Method for adapting multilayer nonlinear self-interference compensators in full-duplex system design.

The Hammerstein model with complex parameters has been selected to represent nonlinear self-interference. This choice is motivated by the model's ability to accurately describe the underlying physical properties of self-interference formation. Due to the holomorphic property of the model output, the Mixed Newton Method provides a "repulsion" effect from saddle points in the loss landscape.

The paper presents convergence curves for the adaptation of the Hammerstein model using both the Mixed Newton Method and conventional gradient descent-based approaches. Additionally, it provides a derivation of the proposed method along with an assessment of its computational complexity.

Keywords: second-order method, complex-valued hessian, in-band full-duplex systems, self-interference cancellation

Citation: *Computer Research and Modeling*, 2025, vol. 16, no. 5, pp. 1–16 (Russian).

This work was supported by the Ministry of Science and Higher Education of the Russian Federation (Goszadaniye), project No. FSMG-2024-0011.

© 2025 Alexander A. Degtyarev, Sergei A. Bakhurin
This work is licensed under the Creative Commons Attribution-NoDerivs 3.0 Unported License. To view a copy of this
license, visit <http://creativecommons.org/licenses/by-nd/3.0/> or send a letter to Creative Commons, PO Box 1866,
Mountain View, CA 94042, USA.

Введение

Одним из ключевых направлений исследований в области телекоммуникаций является задача многопользовательского обслуживания с высокими скоростями передачи данных [Saad et al., 2020]. Среди множества подходов к решению этой задачи перспективным направлением является технология полнодуплексной связи, позволяющая удвоить эффективность использования спектра. Идея такого подхода основана на совместном использовании одной полосы частот передатчиком и приёмником [Luo et al., 2021; Cirik et al., 2016; Lee, Min, 2015], при этом мощный сигнал передатчика является помехой для приемника. Таким образом, подавление собственных помех передатчика является ключевой задачей при проектировании полнодуплексных систем связи.

Поскольку собственные помехи в FD-системах возникают из-за одновременной работы передатчика и приемника, известный сигнал передатчика необходимо вычесть из сигнала на приемной части устройства. Тем не менее, вычитание сигнала передатчика не достаточно для решения задачи SIC, поскольку сигнал передатчика искажается в канале распространения помех TX-RX (от передатчика к приёмнику того же устройства), в передающем тракте за счёт фазового дисбаланса квадратурного модулятора [Korpi et al., 2014], фазовых шумов [Syrgjala et al., 2014], нелинейных искажений АЦП/ЦАП [Balatsoukas-Stimming et al., 2015], а также нелинейностей усилителя мощности и диплексера [Ghannouchi et al., 2015], которые вносят наиболее существенный вклад в нелинейные искажения сигнала передатчика. Таким образом, для компенсации нелинейных помех необходимо построить математическую модель и произвести оптимизацию параметров модели по заданному критерию качества.

Компенсация собственных помех устройства связи традиционно производится в несколько этапов. Во-первых, аналоговая компенсация применяется к входному сигналу приемника, затем он обрабатывается в цифровой области [Tehrani, 2012]. В данной статье мы рассматриваем активное цифровое подавление.

В современных исследованиях на тему SIC выделяются два основных направления. Первое связано с адаптацией классических моделей, таких как модель Винера-Гаммерштейна [Tapiro, Juntti, 2021; Campo et al., 202]. Второе направление представлено построением моделей собственных помех на основе нейронных сетей с действительными и комплексными параметрами [Kurzo et al., 2020; Kristensen et al., 2020; Elsayed et al., 2021].

Все вышеупомянутые модели, как правило, обучаются градиентными методами, поскольку модели являются двух- и трехслойными, и целевая функция имеет седловые точки, к которым притягивается метод Ньютона в случае знаконеопределенной матрицы Гессе. В данной статье представлены результаты применения смешанного метода Ньютона [Bakhurin et al., 2022] для адаптации классической модели Гаммерштейна. Использование только смешанных производных при построении матрицы Гессе позволяет обеспечить свойство отталкивания метода второго порядка от седловой точки, тем самым решить проблему останова алгоритма адаптации. Представлено сравнение результирующих уровней компенсации помех, а также скоростей сходимости, метода MNM с классическими методами первого порядка.

Модель системы

В данном разделе представлена упрощенная схема FD-приемопередатчика с модулем активной цифровой компенсации помех. Также вводится классическая модель нелинейный помех, используемая в текущих экспериментах.

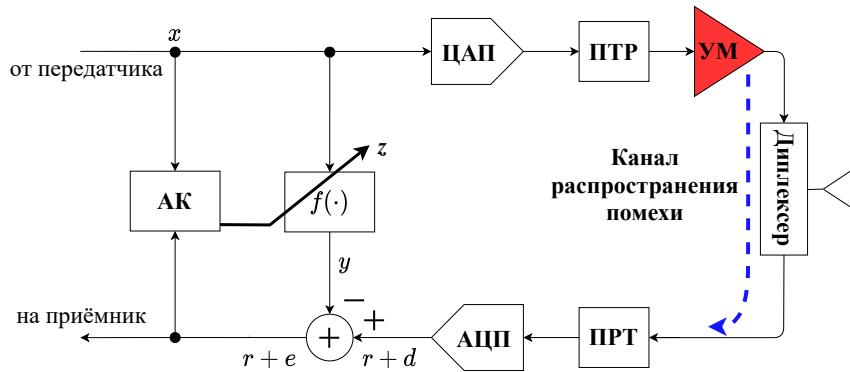


Рис. 1. Схема идентификации собственных помех. x – цифровой сигнал передатчика, r – цифровой полезный сигнал приёмника, d – цифровой сигнал помехи, e – цифровой сигнал обработанной помехи. АК – модуль адаптивной компенсации, УМ – усилитель мощности, ЦАП –цифро-аналоговый преобразователь, АЦП – аналогово-цифровой преобразователь, ПТР – передающий тракт, ПРТ – приёмный тракт.

Задача компенсации собственных помех

Блок-схема идентификации помех показана на рис. 1. Отсчеты цифрового сигнала передатчика x_n преобразуются в аналоговый сигнал, после чего сигнал проходит через нелинейные аналоговые компоненты, такие как аналоговый УМ (усилитель мощности) и диплексер. Результирующий сигнал представляет собой нелинейные искажения, распространяющиеся далее в приёмный тракт по каналу TX-RX. В результате, полезный сигнал приемника смешивается с паразитными помехами, после чего преобразуется в цифровые отсчеты $r_n + d_n$, где r_n – отсчеты полезного сигнала, d_n – отсчеты нелинейных помех.

Для того, чтобы избавиться от паразитных помех d_n , модуль адаптивной компенсации АК на рис. 1 настраивает параметры нелинейного блока $f = f(x_n, z)$, где z – параметры нелинейного блока. Поскольку полезный сигнал r_n не коррелирует с сигналом помех, блок адаптивной компенсации минимизирует SI d_n . В результате, отсчеты сигнала приёмника могут быть представлены как $r_n + e_n$, где $e_n = d_n - y_n = d_n - f(x_n, z)$.

Критерий качества идентификации помех представлен средним квадратом отклонения выхода модели от сигнала помехи (MSE, Mean Square Error) [Tehrani, 2012]:

$$\text{MSE} = \sum_{n=0}^{N-1} e_n^* e_n = \mathbf{e}^H \mathbf{e}, \quad (1)$$

где $\mathbf{e} = \mathbf{d} - \mathbf{y}$, $\mathbf{e} \in \mathbb{C}^{N \times 1}$, $\mathbf{d} \in \mathbb{C}^{N \times 1}$, $\mathbf{y} \in \mathbb{C}^{N \times 1}$, N – длина последовательности сигнала, $(\cdot)^H$ – оператор эрмитового сопряжения.

Модель паразитных помех

Модели помех [Tehrani, 2012] должны отражать физические процессы возникновения данного рода помех. Например, согласно рис. 1 в задаче компенсации собственных помех, сигнал передатчика искажается вследствие нелинейности характеристик усилителя передатчика и диплексера. Затем сигнал искажается вследствие распространения по каналу TX-RX. Таким образом, модель Гаммерштейна [Tehrani, 2012], представленная на рис. 2 отражает процесс возникновения собственной помехи в примённом тракте и может быть использована для решения задачи идентификации. Заметим, что модель Гаммерштейна является упрощением модели Винера-Гаммерштейна [Tehrani, 2012] для случая нелинейных

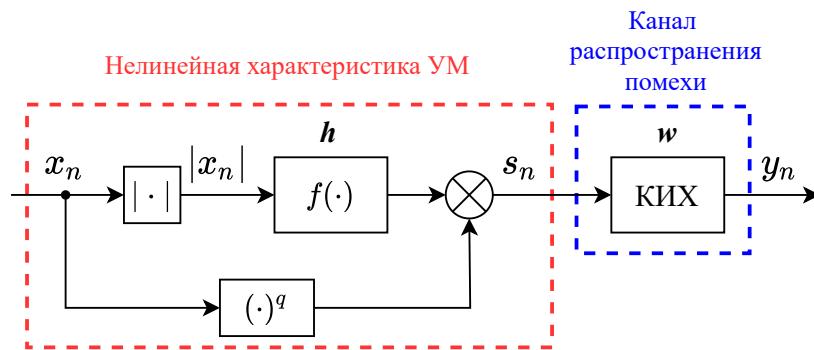


Рис. 2. Модель нелинейных искажений. x_n – отсчёт сигнала передатчика, s_n – выходной отсчёт нелинейности, y_n – выходной отсчёт модели помех, $f(\cdot)$ – нелинейная функция, q – порядок интермодуляционных искажений, \mathbf{h} – обучаемые параметры нелинейной функции, \mathbf{w} – обучаемые параметры КИХ-фильтра. КИХ – фильтр с конечной импульсной характеристикой, УМ – усилитель мощности.

эффектов без памяти. Блок нелинейности $f(\cdot)$ на рис. 2 может быть представлен степенными, тригонометрическими, сплайновыми полиномами и др. В данной работе одномерный сплайновый полином [Molina et al., 2017] был выбран в качестве нелинейности ввиду его низкой вычислительной сложности и простоты аппаратной реализации. Адаптивные параметры блока нелинейности представлены вектором $\mathbf{h} \in \mathbb{C}^{P \times 1}$, P – порядок полинома. Параметры адаптивного КИХ-фильтра [Haykin, 2014] $\mathbf{w} \in \mathbb{C}^{L \times 1}$, $L = 1$ – порядок фильтра. Степень q на практике выбирается в соответствии с порядком интермодуляции. Так как передатчик и приемник работают в одной полосе частот, в данной работе выбран $q = 1$. В результате, выходной отсчёт Гаммерштейна может быть выражен в виде:

$$y_n = \sum_{m=-D}^D w_m \sum_{k=0}^{P-1} h_k x_{n-m} f(|x_{n-m}|), \quad (2)$$

где $L = 2D + 1$, $f(\cdot)$ – базисная функция на основе сплайнов [Molina et al., 2017].

Алгоритм компенсации собственных помех

Текущий раздел разделен на следующие части. Во-первых, приводится общий подход и свойства смешанного метода Ньютона. Затем показано применение MNM к модели Гаммерштейна.

Смешанный метод Ньютона. Общий подход

Смешанный метод Ньютона – метод второго порядка, предназначенный для обучения моделей с комплексными параметрами путём минимизации (1). Для корректной работы алгоритма вектор ошибки \mathbf{e} (1) должен быть голоморфным относительно параметров модели \mathbf{z} , то есть должен не зависеть от сопряжённых параметров \mathbf{z}^* : $\mathbf{e} = \mathbf{e}(\mathbf{x}, \mathbf{z})$. В этом случае MNM обладает свойством отталкивания от седловых точек [Bakhurin et al., 2022].

Пусть $D_{\mathbf{z}}(\cdot)$ – оператор дифференцирования по параметрам модели $\mathbf{z} \in \mathbb{C}^{K \times 1}$, где K – количество параметров всей модели:

$$D_{\mathbf{z}}(\cdot) = \begin{pmatrix} \frac{\partial}{\partial z_0} & \frac{\partial}{\partial z_1} & \cdots & \frac{\partial}{\partial z_{K-1}} \end{pmatrix} \quad (3)$$

Тогда $D_{\mathbf{z}} J \in \mathbb{C}^{1 \times K}$ – производная скалярной вещественнозначной функции $J \in \mathbb{R}$ по вектору параметров всей модели. Таким образом, дифференциал первого порядка функции $J = J(\mathbf{z}, \mathbf{z}^*)$ может быть представлен как [Hjorungnes, 2011]:

$$dJ = D_{\mathbf{z}} J d\mathbf{z} + D_{\mathbf{z}^*} J d\mathbf{z}^*, \quad (4)$$

где $D_{\mathbf{z}} J \in \mathbb{C}^{1 \times K}$, $D_{\mathbf{z}^*} J \in \mathbb{C}^{1 \times K}$, $d\mathbf{z} \in \mathbb{C}^{K \times 1}$, $d\mathbf{z}^* \in \mathbb{C}^{K \times 1}$.

Производная второго порядка скалярной вещественной функции J по вектору параметров описывается следующим образом:

$$D_{\mathbf{z}}(D_{\mathbf{z}} J)^T = H_{\mathbf{z}, \mathbf{z}} J \in \mathbb{C}^{K \times K} – матрица Гессе, \quad (5)$$

где $(\cdot)^T$ – оператор транспонирования.

Используя введённые выше обозначения, дифференциал второго порядка функции J может быть описан следующим образом [Hjorungnes, 2011]:

$$d^2 J = (d\mathbf{z}^T \ d(\mathbf{z}^*)^T) \begin{pmatrix} H_{\mathbf{z}, \mathbf{z}} J & H_{\mathbf{z}^*, \mathbf{z}} J \\ H_{\mathbf{z}, \mathbf{z}^*} J & H_{\mathbf{z}^*, \mathbf{z}^*} J \end{pmatrix} \begin{pmatrix} d\mathbf{z} \\ d\mathbf{z}^* \end{pmatrix} \quad (6)$$

Поскольку алгоритм MNM использует только смешанный гессиан $H_{\mathbf{z}^*, \mathbf{z}} J$ для обучения параметров модели [Bakhurin et al., 2022], то разностное уравнения смешанного метода Ньютона может быть представлено следующим выражением:

$$\mathbf{z}_{k+1} = \mathbf{z}_k - \mu(H_{\mathbf{z}^*, \mathbf{z}} J)^{-1}(D_{\mathbf{z}^*} J)^T, \quad (7)$$

где J – MSE, определенная выражением (1).

Пусть вектор отклонения выхода модели от значений помехи $\mathbf{e}(\mathbf{x}, \mathbf{z}) \in \mathbb{C}^{N \times 1}$ – голоморфная функция относительно \mathbf{z} . Вычислим градиент и смешанный гессиан функции среднеквадратичного отклонения (7). Градиент MSE имеет вид:

$$\begin{aligned} (D_{\mathbf{z}^*} J)^T &= (D_{\mathbf{z}^*}(\mathbf{e}^H \mathbf{e}))^T = (D_{\mathbf{z}^*}(\mathbf{e}^T \mathbf{e}^*))^T = \\ &= (\mathbf{e}^T D_{\mathbf{z}^*} \mathbf{e}^*)^T = (\mathbf{e}^T (D_{\mathbf{z}} \mathbf{e})^*)^T = (D_{\mathbf{z}} \mathbf{e})^H \mathbf{e}. \end{aligned} \quad (8)$$

Поскольку $\mathbf{e} = \mathbf{d} - \mathbf{y}(\mathbf{x}, \mathbf{z})$, где \mathbf{d} не зависит от параметров модели \mathbf{z} , то получим градиент функции потерь вида:

$$(D_{\mathbf{z}^*} J)^T = -(D_{\mathbf{z}} \mathbf{y})^H \mathbf{e}, \quad (9)$$

где $D_{\mathbf{z}} \mathbf{y} \in \mathbb{C}^{N \times K}$ – матрица Якоби выхода модели относительно вектора параметров модели.

В то же время, согласно обозначению (5) и выражению для градиента (9), смешанный гессиан можно рассчитать следующим способом:

$$H_{\mathbf{z}^*, \mathbf{z}} J = D_{\mathbf{z}}(D_{\mathbf{z}^*} J)^T = -D_{\mathbf{z}}((D_{\mathbf{z}} \mathbf{y})^H \mathbf{e}). \quad (10)$$

Поскольку вектор выхода модели \mathbf{y} является голоморфным относительно параметров модели, то смешанный гессиан может быть представлен как:

$$H_{\mathbf{z}^*, \mathbf{z}} J = (D_{\mathbf{z}} \mathbf{y})^H D_{\mathbf{z}} \mathbf{y}. \quad (11)$$

Таким образом, учитывая выражения (9), (11) и (7), алгоритм смешанного метода Ньютона можно представить следующим разностным уравнением:

$$\mathbf{z}_{k+1} = \mathbf{z}_k - \mu((D_{\mathbf{z}} \mathbf{y})^H D_{\mathbf{z}} \mathbf{y})^{-1}(D_{\mathbf{z}} \mathbf{y})^H \mathbf{e}. \quad (12)$$

Смешанный метод Ньютона для модели Гаммерштейна

Согласно разностному уравнению (12), смешанный метод Ньютона требует вычисления матрицы Якоби выходного вектора модели относительно параметров модели \mathbf{z} . Модель Гаммерштейна — двухслойная модель, включающая в себя нелинейность без памяти с адаптивными параметрами $\mathbf{h} \in \mathbb{C}^{P \times 1}$ и адаптивный КИХ-фильтр с параметрами $\mathbf{w} \in \mathbb{C}^{L \times 1}$. Тогда, вектор параметров модели Гаммерштейна и вектор сопряженных параметров представляются следующим образом:

$$\mathbf{z} = \begin{pmatrix} \mathbf{h} \\ \mathbf{w} \end{pmatrix}, \mathbf{z}^* = \begin{pmatrix} \mathbf{h}^* \\ \mathbf{w}^* \end{pmatrix} \in \mathbb{C}^{(P+L) \times 1}. \quad (13)$$

Якобиан выходного вектора модели Гаммерштейна по параметрам представлен выражением:

$$D_{\mathbf{z}} \mathbf{y} = (D_{\mathbf{h}} \mathbf{y} \quad D_{\mathbf{w}} \mathbf{y}) \in \mathbb{C}^{N \times (P+L)}. \quad (14)$$

Таким образом, для подсчета матрицы Гёссе и градиента целевой функции необходимо вычислить производную выходного вектора модели по параметрам слоя нелинейности и КИХ-фильтра.

Согласно теории адаптивной фильтрации [Haykin, 2014] вектор выходных отсчетов КИХ-фильтра может быть выражен через матрицу состояния \mathbf{U} :

$$\mathbf{y} = \mathbf{U} \mathbf{w}, \mathbf{U} \in \mathbb{C}^{N \times L}. \quad (15)$$

Матрица состояния имеет следующую структуру [Haykin, 2014]:

$$\mathbf{U} = \begin{pmatrix} s_D & \cdots & s_1 & s_0 & 0 & \cdots & 0 \\ s_{D+1} & \cdots & s_2 & s_1 & s_0 & \cdots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \cdots & s_{N-1} & s_{N-2} & s_{N-3} & \cdots & s_{N-D-2} \\ 0 & \cdots & 0 & s_{N-1} & s_{N-2} & \cdots & s_{N-D-1} \end{pmatrix},$$

здесь s_n — входной отсчёт фильтра (рис. 2). В соответствии с выражением (2) входной отсчёт КИХ-фильтра в модели Гаммерштейна может быть выражен как:

$$s_n = \sum_{k=0}^{P-1} h_k x_n f(|x_n|). \quad (16)$$

Поскольку матрица состояния адаптивного КИХ-фильтра не зависит от его параметров (16), то, согласно (15), производная выходного вектора модели по параметрам фильтра равна:

$$D_{\mathbf{w}} \mathbf{y} = D_{\mathbf{w}}(\mathbf{U} \mathbf{w}) = \mathbf{U}. \quad (17)$$

Полином на основе сплайнов первого порядка реализует кусочно-линейную интерполяцию при построении нелинейной модели. Данный вид базисных функций эффективен с точки зрения числа операций, осуществляемых в единицу времени при вычислении значений полинома, поэтому он был выбран в качестве модели нелинейности аналоговых компонент системы.

Представим выходной вектор модели Гаммерштейна \mathbf{y} в виде перемножения матрицы входного сигнала модели и вектора параметров нелинейности. Для этого рассмотрим вектор входных отсчетов КИХ-фильтра:

$$\mathbf{s} = \mathbf{V} \mathbf{h}, \mathbf{V} \in \mathbb{C}^{N \times P}. \quad (18)$$

Матрица состояния нелинейности \mathbf{V} для сплайнового полинома показана ниже [Molina et al., 2017]:

$$\mathbf{V} = \begin{pmatrix} 0 & 0 & \cdots & x_0(1 - \Delta_0) & x_0\Delta_0 & \cdots & 0 \\ 0 & \cdots & x_1(1 - \Delta_1) & x_1\Delta_1 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & \cdots & x_k(1 - \Delta_k) & x_k\Delta_k & \cdots & 0 \end{pmatrix}.$$

Каждая строка матрицы состояния \mathbf{V} включает только 2 ненулевых элемента: $x_k(1 - \Delta_k)$ и $x_k\Delta_k$, где x_k — входной отсчёт модели Гаммерштейна, $\Delta_k = ||x_k| - \lfloor |x_k| \rfloor|$, $\lfloor \cdot \rfloor$ — оператор округления числа к ближайшему меньшему.

Текущая матрица состояния может быть записана через её столбцы:

$$\mathbf{V} = (\mathbf{v}_0 \quad \mathbf{v}_1 \quad \cdots \quad \mathbf{v}_{P-1}), \quad (19)$$

тогда входной вектор фильтра (18) может быть выражен как:

$$\mathbf{s} = h_0 \mathbf{v}_0 + h_1 \mathbf{v}_1 + \cdots + h_{P-1} \mathbf{v}_{P-1}. \quad (20)$$

Выходной вектор модели Гаммерштейна может быть представлен в виде:

$$\mathbf{y} = \text{conv}_{\mathbf{w}}(\mathbf{s}), \quad (21)$$

где $\text{conv}_{\mathbf{w}}(\cdot)$ — одномерный линейный оператор свёртки. Используя свойство линейности свёртки из уравнений (20) и (21) получаем выход модели Гаммерштейна через матрично-векторное умножение:

$$\begin{aligned} \mathbf{y} = \text{conv}_{\mathbf{w}}(h_0 \mathbf{v}_0 + \cdots + h_{P-1} \mathbf{v}_{P-1}) &= h_0 \text{conv}_{\mathbf{w}}(\mathbf{v}_0) + \cdots + h_{P-1} \text{conv}_{\mathbf{w}}(\mathbf{v}_{P-1}) = \\ &= (\text{conv}_{\mathbf{w}}(\mathbf{v}_0) \quad \cdots \quad \text{conv}_{\mathbf{w}}(\mathbf{v}_{P-1})) \mathbf{h} \equiv \mathbf{V}_f \mathbf{h}. \end{aligned} \quad (22)$$

Таким образом, выходной вектор модели Гаммерштейна выводится через произведение матрицы состояния нелинейности, фильтрованной по столбцам, и вектора параметров нелинейности. Отметим, что матрица \mathbf{V}_f не зависит от параметров полинома \mathbf{h} .

В результате, используя уравнение (22), получаем производную выхода модели по параметрам нелинейности:

$$D_{\mathbf{h}} \mathbf{y} = D_{\mathbf{h}}(\mathbf{V}_f \mathbf{h}) = \mathbf{V}_f. \quad (23)$$

Наконец, используя выражения для матрицы Якоби выхода модели (14), (17), (23), разностное уравнение смешанного метода Ньютона (12) можно переписать как:

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \mu \begin{pmatrix} \mathbf{V}_f^H \mathbf{V}_f & \mathbf{V}_f^H \mathbf{U} \\ \mathbf{U}^H \mathbf{V}_f & \mathbf{U}^H \mathbf{U} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{V}_f^H \\ \mathbf{U}^H \end{pmatrix} \mathbf{e}. \quad (24)$$

Следует отметить, что алгоритм (24) может быть модифицирован путём поиска оптимального значения шага адаптации μ . В текущих вычислениях параметр μ настраивается в соответствии с решением одномерной задачи оптимизации на каждом шаге алгоритма.

Также следует заметить, что метод градиентного спуска для модели Гаммерштейна представлен следующим выражением:

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \mu (D_{\mathbf{z}^*} J)^T = \mathbf{z}_k + \mu \begin{pmatrix} \mathbf{V}_f^H \\ \mathbf{U}^H \end{pmatrix} \mathbf{e}. \quad (25)$$

Вычислительная сложность шага MNM (24) определяется вычислением матриц \mathbf{U} (17), \mathbf{V}_f (23), градиента $D_{\mathbf{z}^*}\mathbf{J}$ (25), матрично-векторного произведения матрицы Гёссе и градиента, а также вычислением и обращением гессиана $H_{\mathbf{z}^*, \mathbf{z}}\mathbf{J}$ (11).

Вычислительная сложность выходного отсчёта нелинейности на основе сплайна первого порядка (16) оценивается как $\mathcal{O}(1)$ операций с плавающей точкой (FLOPs). Сложность вычисления выхода нелинейности для всей последовательности входных отсчётов $\mathbf{x} \in \mathbb{C}^{N \times 1}$ и, как следствие, сложность вычисления матриц \mathbf{U} и \mathbf{V} может быть оценена как $\mathcal{O}(N)$ FLOPs. Для фильтрации матрицы \mathbf{V} (19) — $\mathcal{O}(PLN)$ FLOPs.

Вычисление градиента и матрицы Гёссе требуют $\mathcal{O}((P+L)N)$ и $\mathcal{O}((P+L)^2N)$ операций соответственно. Сложность вычисления произведения обратного гессиана и градиента равна $\mathcal{O}((P+L)^2)$. Инверсия матрицы Гёссе является одной из наиболее вычислительно трудных операций. Поскольку смешанный гессиан для голоморфных функций является эрмитовой матрицей, то инверсия матрицы может быть реализована с помощью спектрального разложения, сложность которого, оценивается как $\mathcal{O}((P+L)^3)$ [Golub, Loan, 2013] FLOPs.

В соответствии с вышесказанным, сложность шага смешанного алгоритма Ньютона оценивается как:

$$\begin{aligned} \chi_{MNM} = & \mathcal{O}(N) + \mathcal{O}(PLN) + \mathcal{O}((P+L)N) + \mathcal{O}((P+L)^2N) + \mathcal{O}((P+L)^2) + \\ & + \mathcal{O}((P+L)^3) = \mathcal{O}((P+L)^3 + (P+L)^2N). \end{aligned} \quad (26)$$

Таким образом, вычисление и инверсия гессиана являются наиболее сложными с точки зрения числа операций с плавающей точкой.

Заметим, что сложность вычисления шага градиентного спуска существенно ниже, так как для этого требуются только вычисление матриц \mathbf{U} , \mathbf{V}_f , а также градиента:

$$\chi_{grad} = \mathcal{O}(N) + \mathcal{O}(PLN) + \mathcal{O}((P+L)N) = \mathcal{O}(PLN + (P+L)N). \quad (27)$$

Тем не менее, число необходимых шагов метода градиентного спуска для достижения оптимума в реальном приложении может быть много выше по сравнению с числом шагов метода Ньютона. Таким образом, использование вычислительно сложного смешанного метода Ньютона может обеспечить более быстрое достижение оптимума за счет меньшего числа шагов.

Экспериментальная установка

Сигнал передатчика представляет собой QAM-модулированный OFDM сигнал с шириной полосы 60 МГц. Частота дискретизации составляет 484 МГц. Полный набор данных состоит из 80000 отсчётов, 50% из которых используется как тренировочная выборка, остальные отсчёты используются как тестовая выборка.

Схема экспериментальной установки показана на рис. 3. Отсчёты цифрового передатчика отправляются на генератор сигналов, где преобразуются в аналоговый сигнал и перемещаются на несущую частоту 1,7 ГГц. На радиочастоте сигнал проходит через УМ со средней выходной мощностью 20 дБм и захватывается с помощью спектроанализатора на частоте 1,7 ГГц. Захваченный сигнал далее отправляется на ПК. Распространение сигнала по каналу TX-RX моделируется с помощью цифрового КИХ-фильтра 50 порядка. Оцифрованный сигнал с выхода усилителя мощности пропускается через генерированный КИХ-фильтр. Выход фильтра далее рассматривается как последовательность отсчетов собственной помехи.

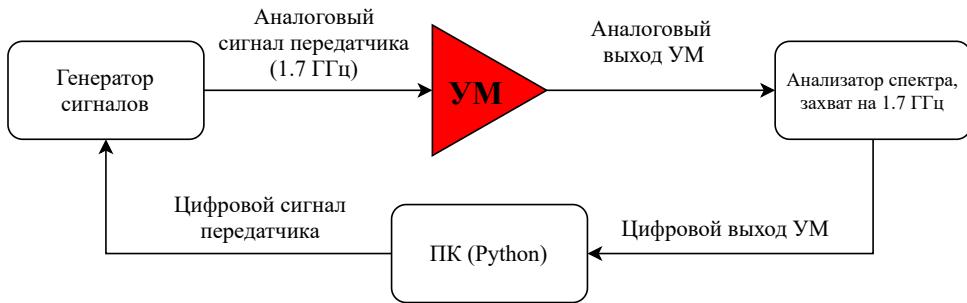


Рис. 3. Схема установки для генерации нелинейных искажений сигнала. УМ – усилитель мощности.

Заметим, что в реальном приложении помехи суммируются с полезным сигналом приемника, как на рис. 1. Тем не менее, модуль аддитивной коррекции подавляет только сигнал собственных помех, поскольку полезный сигнал, приходящий на приёмник с других устройств связи, не коррелирует с сигналом передатчика [Haykin, 2014]. Поэтому в наших симуляциях мы не учитываем полезный RX-сигнал и рассматриваем только собственные помехи в качестве сигнала в приёмном тракте.

Для экспериментов была выбрана модель Гаммерштейна (рис. 2) с аддитивным фильтром с $L = 45$ числом параметров, а также кусочно-линейной функцией, содержащей $P = 8$ сплайнов первого порядка. Степень $q = 1$.

Все предоставленные результаты были получены на графическом процессоре NVIDIA Tesla V100 с 16 ГБ ОЗУ.

Численные результаты

В данном разделе приводится краткое описание классических градиентных методов, таких как блочный градиентный спуск (BGD, Block Gradient Descent) и стохастический градиентный спуск (SGD, Stochastic Gradient Descent), с которыми производится сравнение предложенного метода второго порядка. Кроме того, проводится сравнение результирующего уровня компенсации помехи и скорости сходимости.

Шаг адаптации смешанного метода Ньютона и BGD пересчитывается на всей тренировочной последовательности. Другими словами, матрицы V_f и U в уравнениях (24), (25) вычисляются для 40000 отсчётов.

С другой стороны, в случае SGD вся тренировочная последовательность делится на неперекрывающиеся блоки, которые упаковываются в 50 блоков по 200 отсчётов каждый. Таким образом, за одну эпоху (одно прохождение всей тренировочной последовательности) модель тренируется на 4 блоках. Градиенты вычисляются для каждого блока (по 200 отсчётов) и усредняются (по 50 блоков). Размер блока был выбран экспериментально с точки зрения максимизации результирующего уровня качества модели.

Также следует обратить внимание, что в случае BGD и MNM параметры обновляются один раз за эпоху, то есть после подсчёта шага на полной последовательности тренировочных данных. В то время как в SGD параметры пересчитываются после вычисления стохастического градиента на каждом блоке (50×200 отсчётов).

В следующих симуляциях нормированная среднеквадратичная ошибка NMSE отслеживается на каждой эпохе на тренировочной и тестовой последовательности, что отражено на рис. 4, 5.

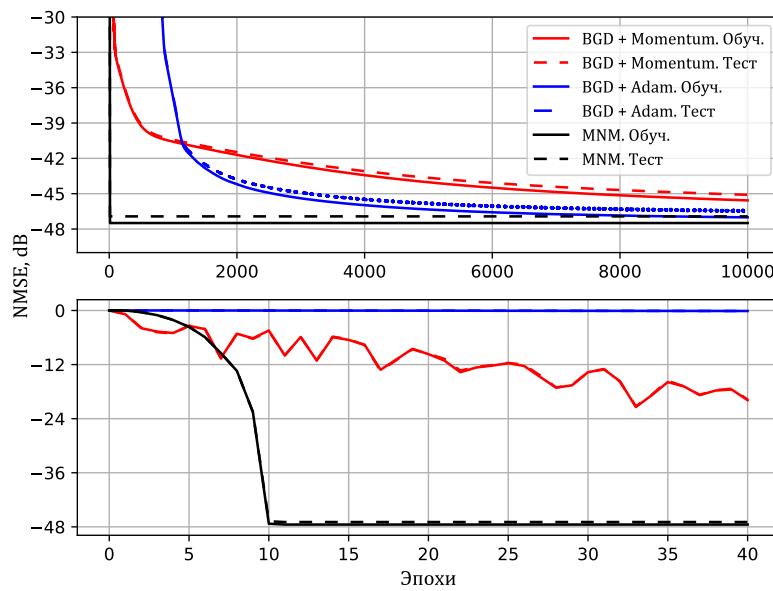


Рис. 4. Кривые сходимости на тренировочной и тестовой последовательностях. Сравнение BGD с оптимизаторами Momentum, Adam и смешанного метода Ньютона. BGD – блочный градиентный спуск, MNM – смешанный метод Ньютона, NMSE – нормированная среднеквадратичная ошибка, Adam, Momentum – оптимизаторы градиентного спуска.

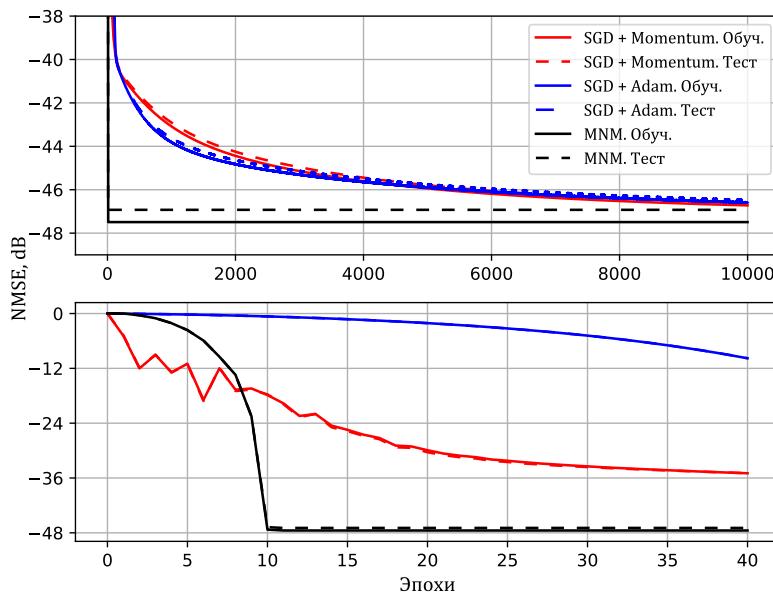


Рис. 5. Кривые сходимости на тренировочной и тестовой последовательностях. Сравнение SGD с оптимизаторами Momentum, Adam и смешанного метода Ньютона. SGD – стохастический градиентный спуск, MNM – смешанный метод Ньютона, NMSE – нормированная среднеквадратичная ошибка, Adam, Momentum – оптимизаторы градиентного спуска.

Сравнение скорости сходимости блочного градиентного спуска с оптимизаторами Momentum и Adam, а также смешанного метода Ньютона показано на рис. 4. MNM требует приблизительно 30 эпох для достижения результирующего значения уровня компенсации, в то время как BGD с обоими оптимизаторами требуют ~ 10000 эпох для достижения такого же значения NMSE.

На рис. 5 представлены кривые обучения, полученные для SGD с оптимизаторами Momentum и Adam и смешанного метода Ньютона. Графики сходимости отражают тот факт, что SGD также требует приблизительно 10000 эпох для достижения уровня компенсации паразитной помехи, полученного при помощи MNM.

Отметим, что параметры оптимизаторов и темп обучения градиентного спуска, были выбраны с точки зрения наилучшего результирующего качества модели и высокой скорости сходимости.

Спектральные плотности мощности (СПМ) исходной и компенсированной различными методами помехи изображены на рис. 6. Графики СПМ построены на основе тестовой последовательности.

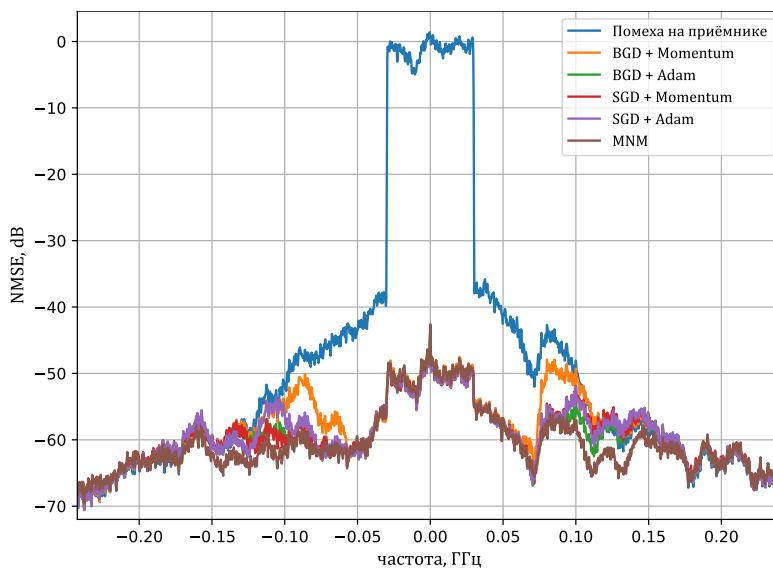


Рис. 6. Спектральные плотности мощности исходной и компенсированных помех. BGD – блочный градиентный спуск, SGD – стохастический градиентный спуск, Adam, Momentum – оптимизаторы градиентного спуска, MNM – смешанный метод Ньютона, NMSE – нормированная среднеквадратичная ошибка.

Таблица 1 отражает заметное увеличение скорости сходимости смешанного метода Ньютона по сравнению с классическими градиентными методами. Для достижения сопоставимого уровня компенсации для MNM требуется всего 30 эпох. Несмотря на то, что при использовании метода второго порядка требуется существенно меньшее число шагов оптимизации, каждый шаг адаптации вычисляется ~ 5 в раз дольше. Тем не менее, общее время, затрачиваемое на обучение модели, существенно меньше по сравнению с методами первого порядка.

Согласно расчетам, приведенным в теоретической части, вычислительная сложность шага MNM (26) высока по сравнению с алгоритмами на основе градиентного спуска (27). Тем не менее, число шагов, необходимых для достижения такого же уровня компенсации, значительно меньше для смешанного Ньютона (таблица 1). Таким образом, MNM отлично подходит для исследования возможного уровня качества моделей. Кроме того, смешанный метод Ньютона можно использовать для онлайн обучения моделей, так как в реальных приложениях требуется накопление матрицы Гессе и, как следствие, частота шагов оптимизации существенно меньше по сравнению с методами первого порядка.

Таблица 1. Сравнение уровня компенсации помехи и скорости сходимости

Алгоритм	BGD Moment.	BGD Adam	SGD Moment.	SGD Adam	MNM
Номер эпохи	10000	10000	10000	10000	30
Время за эпоху, 10^{-2} с	3.8	4.0	3.7	4.1	21
Общее время работы, с	380	403	386	412	6.2
NMSE, dB	-45.1	-46.7	-46.6	-46.5	-46.9

Заключение

В данной работе предложено использование метода второго порядка для компенсации собственных помех в приёмнике полнодуплексного устройства связи. С использованием инструмента матричного дифференцирования, метод строго выведен для случая использования модели Гаммерштейна в качестве модели паразитных помех.

Несмотря на то, что вычисление шага оптимизации смешанного метода Ньютона требует в ~ 5 раз больше времени, общая длительность обучения уменьшена до 30 эпох, по сравнению 10000 эпох необходимых для классических методов оптимизации первого порядка, для достижения того же уровня компенсации 46.9 dB.

В последующих работах планируется рассмотрение применимости предложенного алгоритма второго порядка для обучения нелинейных моделей, представленных нейронными сетями с комплексными параметрами. Смешанный метод Ньютона требует большого числа операций для расчета матрицы Гессе, и ее обращения. Таким образом, важнейшим направлением исследований является модификация смешанного метода Ньютона для минимизации накладных расходов вычисления и инверсии гессиана, что может быть реализовано, например, при помощи различных оценок обратной матрицы.

Список литературы (References)

- Bakhurin S., Hildebrand R., Alkousa M., Titov A. Optimization in complex spaces with the Mixed Newton Method. — 2022. — Preprint at <https://arxiv.org/pdf/2207.04265.pdf>.
- Balatsoukas-Stimming A., Austin A. C. M., Belanovic P., Burg A. Baseband and RF Hardware Impairments in Full-Duplex Wireless Systems: Experimental Characterisation and Suppression // EURASIP Journal on Wireless Communications and Networking. — 2015. — no. 142.
- Campo P. P., Anttila L., Korpi D., Valkama M. Cascaded Spline-Based Models for Complex Nonlinear Systems: Methods and Applications // IEEE Transactions on Signal Processing. — 202. — Vol. 69. — P. 370–384.
- Cirik A. C., Biswas S., Vuppala S., Ratnarajah T. Robust Transceiver Design for Full Duplex Multiuser MIMO Systems // IEEE Wireless Communications Letters. — 2016. — Vol. 5, no. 3. — P. 260–263.
- Elsayed M., El-Banna A. A. A., Dobre O. A. et al. Full-Duplex Self-Interference Cancellation Using Dual-Neurons Neural Networks // IEEE Communications Letters. — 2021. — Vol. 26, no. 3. — P. 557–561.

- Ghannouchi F. M., Hamm O., Helaoui M.* Behavioral modelling and predistortion of wideband wireless transmitters. — John Wiley Sons Ltd, 2015.
- Golub G. H., Loan C. F. V.* Matrix Computations. Fourth Edition. — Baltimore: The Johns Hopkins University Press, 2013.
- Haykin S.* Adaptive Filter Theory, 5th edition. — Communications Research Laboratory, McMaster University, Hamilton, Ontario, Canada: Pearson Education Limited, 2014.
- Hjorungnes A.* Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications. — University of Oslo, Norway: Cambridge University Press, 2011.
- Korpi D., Anttila L., Syrjälä V., Valkama M.* Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver // IEEE Journal on Selected Areas in Communications. — 2014. — Vol. 32, no. 9. — P. 1674–1687.
- Kristensen A. T., Burg A., Balatsoukas-Stimming A.* Advanced Machine Learning Techniques for Self-Interference Cancellation in Full-Duplex Radios. — 2020. — Paper presented at the 2019 53rd Asilomar Conference on Signals, Systems, and Computers, 3–6 November 2019.
- Kurzo Y., Kristensen A. T., Burg A., Balatsoukas-Stimming A.* Hardware Implementation of Neural Self-Interference Cancellation // IEEE Journal on Emerging and Selected Topics in Circuits and Systems. — 2020. — Vol. 10, no. 2. — P. 204–216.
- Lee D., Min B.-W.* Results and trade-off of self-interference cancellation in a full-duplex radio front-end. — 2015. — Paper presented at the 2015 International Workshop on Antenna Technology (iWAT), 4–6 March 2015.
- Luo H., Holm M., Ratnarajah T.* On the performance of active analog self-interference cancellation techniques for beyond 5G systems // China Communications. — 2021. — Vol. 18, no. 10. — P. 158–168.
- Molina A., Rajamania K., Azadet K.* Digital Predistortion Using Lookup Tables With Linear Interpolation and Extrapolation: Direct Least Squares Coefficient Adaptation // IEEE Transactions on Microwave Theory and Techniques. — 2017. — Vol. 65. — P. 980–987.
- Saad W., Bennis M., Chen M.* A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems // IEEE Network. — 2020. — P. 134–142.
- Syrjala V., Valkama M., Anttila L. et al.* Analysis of Oscillator Phase-Noise Effects on Self-Interference Cancellation in Full-Duplex OFDM Radio Transceivers // IEEE Transactions on Wireless Communications. — 2014. — Vol. 13, no. 6. — P. 2977–2990.
- Tapio V., Juntti M.* Non-Linear Self-Interference Cancelation for Full-Duplex Transceivers Based on Hammerstein-Wiener Model // IEEE Communications Letters. — 2021. — Vol. 25, no. 11. — P. 3684–3688.
- Tehrani A. S.* Behavioral modeling of wireless transmitters for distortion mitigation. — Chalmers University of Technology, Gothenburg, Sweden: Chalmers Reproservice, 2012.