

УДК: 621.394.441

Исследование традиционных и ИИ-моделей в задаче подавления интермодуляционных продуктов второго порядка

А. А. Дегтярев^a, Н. В. Бахолдин^b, А. Ю. Масловский^c, С. А. Бахурин^d

Московский физико-технический институт,
Россия, 141701, Московская обл., г. Долгопрудный, Институтский пер., д. 9

E-mail: ^a degtyarev.aa@phystech.edu, ^b bakholdin.nikita@gmail.com, ^c aleksandr.maslovskiy@phystech.edu,
^d bakhurin.sa@mipt.ru

Получено 28.10.2024, после доработки — 13.11.2024
Принято к публикации 25.11.2024

В данной работе рассматриваются нейросетевые модели и полиномиальные модели на основе полинома Чебышёва для компенсации помех. Показано, что нейросетевая модель обеспечивает компенсацию паразитных помех без необходимости настройки параметров, в отличие от полиномиальной модели, где требуется подбор оптимальных задержек. Для обеих архитектур использован метод L-BFGS, который достигает уровня компенсации, сопоставимого с решением LS для полиномиальной модели, с результатом NMSE = -23,59 дБ и требует менее 2000 итераций, что подтверждает его высокую эффективность. Также благодаря высокой обобщающей способности нейросетевых моделей метод первого порядка для нейросетевых архитектур демонстрирует более быструю сходимость по сравнению с полиномиальной моделью. За 20 000 итераций нейросетевая модель достигает прироста уровня компенсации на 0,44 дБ по сравнению с полиномом. В отличие от этого полиномиальная модель может достичь высокого уровня компенсации только при оптимальной настройке параметров методов первого порядка, что подчеркивает одно из ключевых преимуществ нейросетевых моделей.

Ключевые слова: интермодуляционные помехи второго порядка, адаптивный фильтр, нейросетевые модели, полиномы Чебышёва

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

© 2024 Александр Андреевич Дегтярев, Никита Владимирович Бахолдин, Александр Юрьевич Масловский, Сергей Алексеевич Бахурин

Статья доступна по лицензии Creative Commons Attribution-NoDerivs 3.0 Unported License.
Чтобы получить текст лицензии, посетите веб-сайт <http://creativecommons.org/licenses/by-nd/3.0/>
или отправьте письмо в Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

UDC: 621.394.441

A study of traditional and AI-based models for second-order intermodulation product suppression

A. A. Degtyarev^a, N. V. Bakholdin^b, A. Yu. Maslovskiy^c, S. A. Bakhurin^d

Moscow Institute of Physics and Technology,
9 Institutskiy per., Dolgoprudny, Moscow region, 141701, Russia

E-mail: ^a degtyarev.aa@phystech.edu, ^b bakholdin.nikita@gmail.com, ^c aleksandr.maslovskiy@phystech.edu,
^d bakhurin.sa@mipt.ru

*Received 28.10.2024, after completion – 13.11.2024
Accepted for publication 25.11.2024*

This paper investigates neural network models and polynomial models based on Chebyshev polynomials for interference compensation. It is shown that the neural network model provides compensation for parasitic interference without the need for parameter tuning, unlike the polynomial model, which requires the selection of optimal delays. The L-BFGS method is applied to both architectures, achieving a compensation level comparable to the LS solution for the polynomial model, with an NMSE result of -23.59 dB and requiring fewer than 2000 iterations, confirming its high efficiency. Additionally, due to the strong generalization ability of neural network architectures, the first-order method for neural networks demonstrates faster convergence compared to the polynomial model. In 20 000 iterations, the neural network model achieves a 0.44 dB improvement in compensation level compared to the polynomial model. In contrast, the polynomial model can only achieve high compensation levels with optimal first-order method parameter tuning, highlighting one of the key advantages of neural network models.

Keywords: IMD2, adaptive filter, artificial intelligence, Chebyshev polynomials

Citation: *Computer Research and Modeling*, 2024, vol. 16, no. 7, pp. 1569–1578 (Russian).

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

Введение

В последние десятилетия приемник прямого преобразования широко используется в современных смартфонах, так как почти все компоненты передатчика и приемника могут быть размещены на одном чипе [Kim, Silva-Martinez, 2013; Fabiano et al., 2013; Ramella et al., 2017]. Простота такого подхода обеспечивается одноступенчатым понижением частоты, что позволяет обеспечить поддержку нескольких частотных диапазонов.

Ввиду сложности обеспечения изоляции дуплексного фильтра приемопередатчики, работающие в режиме частотного разделения каналов, подвержены утечке передаваемого сигнала в приемный тракт. Утечка сигнала в сочетании с нелинейностью малошумящего усилителя (МШУ) и смесителя понижения частоты приемника может привести к генерации интермодуляционных искажений второго порядка, что значительно снижает чувствительность приемника.

Общий подход к борьбе с помехами заключается в их компенсации в аналоговом радиочастотном [Chang, Shin, 2020] и/или цифровых [Gebhard et al., 2019] областях до уровня шума приемника. В данной статье мы сосредотачиваемся на цифровой компенсации интермодуляционных продуктов второго порядка. Подавление помех в традиционном виде осуществляется с помощью адаптации классических структур, таких как полиномиальные модели с памятью и без памяти, сплайн-полиномы и т. д. [Tehrani, 2012; Campo et al., 2021; Tapio, Juntti, 2021].

Другой многообещающий подход представлен цифровым компенсатором помех с использованием нейронной сети (НС) [Ploder et al., 2019; Kurzo et al., 2020; Kristensen, Burg, Balatsoukas-Stimming, 2020; Elsayed et al., 2021]. Современные архитектуры искусственного интеллекта способны конкурировать с классическими методами по характеристикам и вычислительной сложности в задачах цифрового предсказания, компенсации помех для систем с частотным разделением каналов и т. д. [Elsayed et al., 2022; Hu et al., 2022]. Основным свойством моделей, предлагаемых для задач адаптивной обработки сигналов, является способность к извлечению временных признаков. Для такой задачи в качестве входных данных модель получает последовательность отсчетов сигнала передатчика. Более того, входная последовательность делится на несколько ветвей с различными задержками для учета инерционных свойств нелинейности.

В данной статье мы исследуем интермодуляционные продукты второго порядка, генерируемые в результате прохождения сигнала через нелинейный частотный смеситель в приемном тракте. Таким образом, данные передатчика представлены в виде последовательности комплексных отсчетов сигнала с квадратичной модуляцией, в то время как данные приемника являются последовательностью вещественных отсчетов.

Классические и нейросетевые модели

Классические полиномиальные модели

Классическим методом компенсации нелинейных интермодуляционных продуктов является адаптация обобщенной полиномиальной модели с памятью по методу наименьших квадратов (МНК). Полиномиальная модель известна как структура, описывающая физические свойства усилителя мощности (УМ) для различных типов и режимов их работы [Sytjala et al., 2014].

Для задачи компенсации интермодуляционных искажений второго порядка в данной статье используется модель с памятью на основе полиномов Чебышёва, выход которой представлен следующим образом:

$$y_n = \sum_{k=0}^{K-1} \sum_{p=0}^{P-1} \theta_{k,p} T_p(|x_{n-d_k}|), \quad (1)$$

$$T_p(|x_{n-d_k}|) = \cos(n \cdot \arccos(|x_{n-d_k}|)), \quad (2)$$

где $\theta_{k,p} \in \mathbb{R}$ — обучаемые параметры полинома Чебышёва, x_n — отсчеты сигнала, d_k — задержки отсчетов сигнала, $T_p(\cdot)$ — порядок p полинома Чебышёва первого рода.

Коэффициенты полинома находятся с помощью метода наименьших квадратов (МНК), так как модель является однослойной и линейной по параметрам. Поэтому МНК гарантирует глобальную сходимость за один шаг оптимизации [Tehrani, 2012].

Тем не менее в реальных приложениях используются алгоритмы оптимизации первого порядка. В этом случае модель требует тщательной настройки параметров оптимизации гиперпараметров для обеспечения высокой скорости сходимости, а также стабильности работы алгоритма оптимизации. Ввиду сложности данной задачи ниже рассматриваются структуры нейронных сетей. Архитектуры НС обладают высокой обобщающей способностью и, как следствие, менее чувствительны к настройке гиперпараметров алгоритма.

Нейросетевая модель

Архитектура нейронной сети для компенсации интермодуляционных помех второго порядка представлена на рис. 1. Поскольку нелинейность обладает инерционностью, структура НС учитывает эффект памяти.

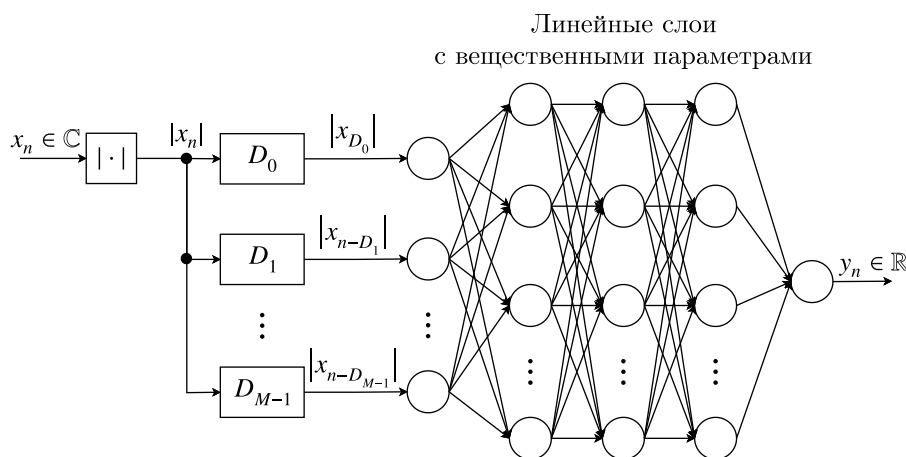


Рис. 1. Архитектура модели на основе ИИ

В данной работе реализована компактная нелинейная модель, учитывающая физический процесс формирования помехи. В качестве основы выбранной архитектуры использована модель Винера–Гаммерштейна. Эффекты памяти учитываются путем разделения входного сигнала на M ветвей с разными задержками, как показано на рис. 1. Вектор отсчетов задержанного сигнала подается на последовательность линейных слоев, разделенных функциями активации:

$$y_n = \mathbf{W}_{\text{out}} \sigma_{L-1}(\mathbf{W}_{L-1} \cdots \sigma_1(\mathbf{W}_1 \sigma_0(\mathbf{W}_0 \mathbf{f}_n))), \quad (3)$$

$$\mathbf{f}_n = (|x_{n-d_0}| \quad |x_{n-d_1}| \quad \cdots \quad |x_{n-d_{M-1}}|), \quad (4)$$

где M — количество задержек сигнала, $\sigma_i(\cdot)$ — функция активации i -го слоя, $\mathbf{W}_0 \in \mathbb{R}^{K \times M}$, $\mathbf{W}_j \in \mathbb{R}^{K \times K}$, $j \in \overline{0, L-1}$, — матрицы весов скрытых линейных слоев, K — количество выходных каналов скрытого слоя, L — количество скрытых слоев, $\mathbf{W}_{\text{out}} \in \mathbb{R}^{1 \times K}$ — матрица весов выходного линейного слоя, \mathbf{f}_n — вектор, состоящий из абсолютных значений входного задержанного сигнала.

Отметим, что согласно уравнению (3) смещение в линейных слоях отсутствует ввиду экономии вычислительных ресурсов.

как задачу идентификации путем минимизации среднеквадратичного отклонения выхода нелинейной модели от сигнала помехи на приемнике:

$$f(\theta) := \frac{1}{m} \sum_{k=1}^m ([\mathcal{M}_\theta(x)]_k - \bar{y}_k)^2 \rightarrow \min_{\theta}. \quad (5)$$

Для оценки качества решения, полученного в результате оптимизации функционала, будем использовать нормированную среднеквадратичную ошибку:

$$\text{NMSE}(y, \bar{y}) := 10 \log_{10} \left\{ \frac{\sum_{k=1}^m (y_k - \bar{y}_k)^2}{\sum_{k=1}^m x_k^2} \right\} \text{ дБ.}$$

Метод наименьших квадратов

Ввиду того что полиномиальная модель 1 является линейной по коэффициентам, глобальный оптимум функционала 5 может быть найден за один шаг по методу наименьших квадратов:

$$\nabla_{\theta} (\|A\theta - b\|^2 + \lambda \|\theta\|^2) = 0, \quad (6)$$

$$\theta = (A^H \cdot A + \lambda)^{-1} \cdot A^H \cdot b. \quad (7)$$

Тем не менее метод наименьших квадратов требует накопления гессиана, а также подсчета обратной матрицы, вычислительная сложность которого, вообще говоря, является кубической. В результате данный метод редко используется при аппаратной реализации системы компенсации помех.

Adam

При аппаратной реализации используются методы первого порядка, поскольку вычислительная сложность градиентных методов является линейной. Классический метод стохастического градиентного спуска сходится к ближайшему локальному оптимуму. Кроме того, в окрестности оптимума норма градиента становится меньше, в результате чего локальный оптимум не может быть достигнут.

Ввиду вышеназванных причин в данной работе применяется оптимизатор градиентного спуска Adam, который позволяет решить проблему «застывания» в локальных оптимумах, а также ускоряет сходимость за счет механизма автоматической регулировки шага [Kingma, Ba, 2014; Maslovskiy et al., 2021]:

$$\theta_k = \theta_{k-1} - \frac{\alpha}{\sqrt{v_k + \epsilon}} \widehat{m}_k, \quad (8)$$

где

$$m_k = \beta_1 m_{k-1} + (1 - \beta_1) \nabla f(x), \quad v_k = \beta_2 v_{k-1} + (1 - \beta_2) \nabla f(x)^2.$$

L-BFGS

Как известно, глобальный оптимум невыпуклой задачи оптимизации, вообще говоря, не может быть найден по методу наименьших квадратов за один шаг. Помимо методов первого порядка, для поиска решения используют методы второго порядка, такие как итеративный метод Ньютона, обладающий квадратичной скоростью сходимости:

$$\theta_k = \theta_{k-1} + \alpha H \nabla_{\theta} f(x), \quad (9)$$

где $H = [\nabla_{\theta}^2 f]^{-1}$.

Однако данный алгоритм подразумевает обращение матрицы Гессе. Для аппроксимации матрицы вторых производных H применяются квазиньютоновские методы. Квазиньютоновский алгоритм BFGS обладает сверхлинейной скоростью сходимости, а также квадратичной вычислительной сложностью.

$$\theta_{k+1} = \theta_k - h_k \cdot H_k \nabla f(\theta_k), \quad (10)$$

где $h_k = \arg \min_{h>0} f(\theta_k - h \cdot H_k \nabla f(\theta_k))$,

$$H_{k+1} = H_k + \frac{H_k \gamma_k \delta_k^\top + \delta_k \gamma_k^\top H_k}{\langle H_k \gamma_k, \gamma_k \rangle} - \beta_k \frac{H_k \gamma_k \gamma_k^\top H_k}{\langle H_k \gamma_k, \gamma_k \rangle}, \quad (11)$$

где

$$\beta_k = 1 + \frac{\langle \gamma_k, \delta_k \rangle}{\langle H_k \gamma_k, \gamma_k \rangle}, \quad \gamma_k = \nabla f(\theta_{k+1}) - \nabla f(\theta_k), \quad \delta_k = \theta_{k+1} - \theta_k, \quad H_0 = I. \quad (12)$$

Тем не менее этот метод не подходит для оптимизации функционала с большим числом параметров ввиду квадратичной сложности по памяти, необходимой для вычисления и хранения матрицы H_k . Поэтому на практике используется метод L-BFGS, который подразумевает пересчет матрицы H_k с использованием лишь l векторов γ_k и δ_k [Zhu et al., 1997; Nocedal, 1980]. Данный метод также используется для вычисления в данной работе для компенсации интермодуляционных помех.

Результаты экспериментов

Для экспериментов использовался программный пакет PyTorch. В текущих экспериментах мы используем значения критериев качества и скорость сходимости в результате адаптации полинома Чебышёва (1), (2), где $K = 3$, $P = 8$, и нейросетевой модели (3), где $M = 3$, $L = 2$, $W_0 \in \mathbb{R}^{3 \times 3}$, $W_1 \in \mathbb{R}^{2 \times 3}$, $W_{\text{out}} \in \mathbb{R}^{1 \times 2}$. Таким образом, полиномиальная модель имеет 24 вещественных параметра, в то время как модель на основе нейронной сети имеет 17 вещественных параметров.

Качество полиномиальной модели оценивалось при помощи МНК, Adam и L-BFGS. Уровень компенсации при помощи нейросетевой модели оценивался по итогам работы Adam и L-BFGS.

Как правило, в реальных аппаратных приложениях для задач подавления помех используются алгоритмы первого порядка из-за высокой вычислительной стоимости методов второго порядка и квазиньютоновских методов. Поэтому в текущих экспериментах была сравнена скорость сходимости моделей на основе нейронной сети и полинома, обученных с помощью градиентного спуска с оптимизатором Adam (8) (рис. 3).

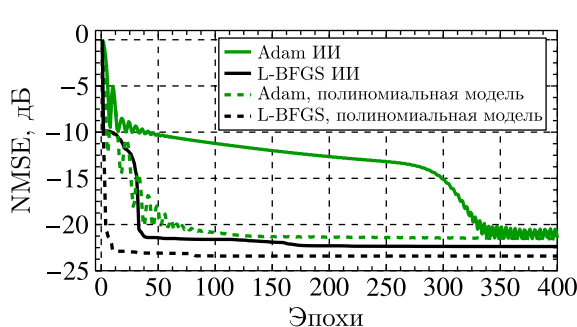
Графики спектральной плотности мощности (СПМ) интермодуляционных помех второго порядка и остаточных ошибок после работы алгоритма компенсации помехи показаны на рис. 4.

Согласно результатам, представленным в таблице 1 и на рис. 3, метод L-BFGS обеспечивает уровень компенсации для обеих архитектур, близкий к решению МНК для полиномиальной модели с NMSE = -23,59 дБ. Более того, скорость сходимости L-BFGS для обеих структур занимает менее 2000 итераций (таблица 1). Текущие результаты показывают его практическую значимость данного метода.

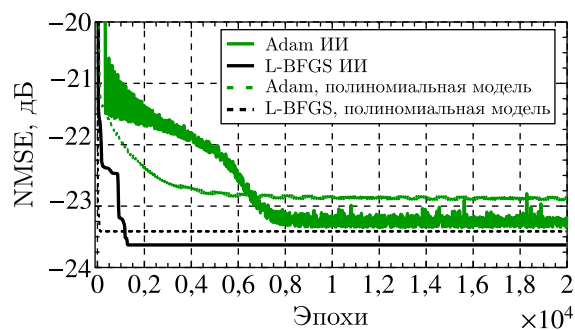
В то же время метод первого порядка для модели на основе нейронной сети демонстрирует более высокую скорость сходимости по сравнению с полиномиальной моделью благодаря обобщающей способности нейросетевых моделей. Архитектура НС обеспечивает улучшение производительности на 0,44 дБ по сравнению с полиномиальной моделью за 20 000 итераций.

Таблица 1. Сравнение результатов моделирования для полиномиальной и нейросетевой модели при использовании различных оптимизаторов: уровень компенсации помехи в зависимости от числа итераций

Модель	Алгоритм	Количество итераций				
		1000	2000	5000	10 000	20 000
Полиномиальная	МНК	-23,59	-23,59	-23,59	-23,59	-23,59
	Adam	-21,96	-22,40	-22,76	-22,83	22,91
	L-BFGS(100)	-23,41	-23,41	-23,41	-23,41	-23,41
ИИ	МНК	N/A	N/A	N/A	N/A	N/A
	Adam	-21,00	-21,69	-22,03	-23,28	-23,35
	L-BFGS(100)	-23,20	-23,63	-23,63	-23,63	-23,63

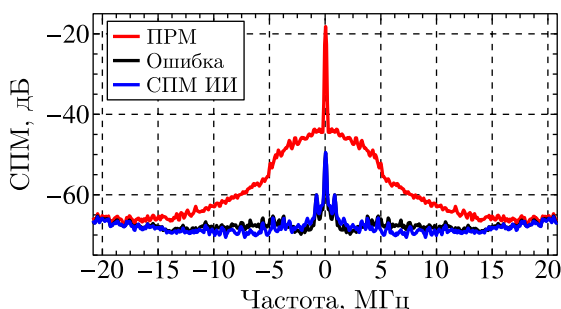


(а) Adam и L-BFGS, 400 итераций

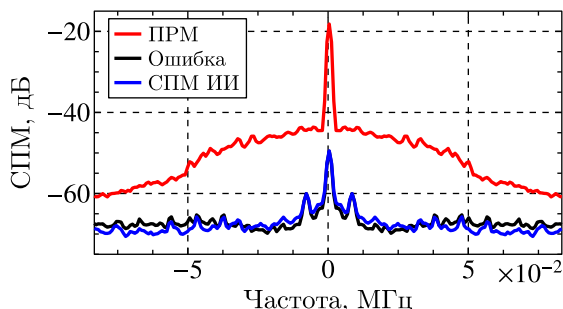


(б) Adam и L-BFGS, 20 000 итераций

Рис. 3. Кривые обучения при различном количестве эпох



(а) Помеха до и после подавления, полоса 40 МГц



(б) Помеха до и после подавления, полоса 20 МГц

Рис. 4. Графики СПМ сигнала приемника, ошибки после компенсации помехи в результате адаптации полинома Чебышёва и нейросетевых моделей

Тем не менее полином может достичь высокого значения уровня подавления путем тонкой настройки параметров оптимизатора первого порядка, что демонстрирует одно из значимых преимуществ нейросетевых структур.

Заключение

В данной работе показано, что предложенная нейросетевая модель и полином Чебышёва способны достигать высокого уровня компенсации помех; при этом нейросетевая модель обеспечивает компенсацию паразитных помех без настройки параметров, тогда как для полиномиальной модели требуется подбор оптимального набора задержек.

Результаты показывают, что метод L-BFGS обеспечивает уровень компенсации для обеих архитектур, сопоставимый с решением LS для полиномиальной модели, достигая значения $NMSE = -23,59$ дБ. Кроме того, метод L-BFGS требует менее 2000 итераций для обеих структур, что подтверждает его эффективность при оценке возможного уровня компенсации для используемой архитектуры.

Благодаря высокой обобщающей способности нейросетевых архитектур метод первого порядка для нейросетевых моделей также демонстрирует более высокую скорость сходимости по сравнению с полиномиальной моделью. Например, за 20 000 итераций нейросетевая архитектура достигает прироста уровня компенсации на 0,44 дБ по сравнению с полиномом. Кроме того, полиномиальная модель может достичь высокого уровня компенсации при только при оптимальной настройке параметров методов первого порядка, что подчеркивает одно из ключевых преимуществ нейросетевых архитектур.

Список литературы (References)

- 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*. — 2021. — Vol. 69. — P. 370–384.
- Chang S., Shin H.* 2.4-GHz CMOS Bluetooth RF receiver with improved IM2 distortion tolerance // *IEEE Transactions on Microwave Theory and Techniques*. — 2020. — Vol. 68, No. 11. — P. 4589–4598. — DOI: 10.1109/TMTT.2020.3009945
- Elsayed M., El-Banna A. A. A., Dobre O. A., Shiu W., Wang P.* Full-duplex self-interference cancellation using dual-neurons neural networks // *IEEE Communications Letters*. — 2021. — Vol. 26, No. 3. — P. 557–561.
- Elsayed M., El-Banna A. A. A., Dobre O. A., Shiu W., Wang P.* Hybrid-layers neural network architectures for modeling the self-interference in full-duplex systems // *IEEE Transactions on Vehicular Technology*. — 2022. — Vol. 71, No. 6. — P. 6291–6307.
- Fabiano I., Sosio M., Liscidini A., Castello R.* SAW-less analog front-end receivers for TDD and FDD // *IEEE Journal of Solid-State Circuits*. — 2013. — Vol. 48, No. 12. — P. 3067–3079.
- Gebhard A., Lang O., Lunglmayr M., Motz C., Kanumalli R. S., Auer C., Paireder T., Wagner M., Pretl H., Huemer M.* A robust nonlinear RLS type adaptive filter for second-order-intermodulation distortion cancellation in FDD LTE and 5G direct conversion transceivers // *IEEE Transactions on Microwave Theory and Techniques*. — 2019. — Vol. 67, No. 5. — P. 1946–1961. — DOI: 10.1109/TMTT.2019.2896513
- Hu X., Liu Z., Yu X., Zhao Y., Chen W., Hu B., Du X., Li X., Helaoui M., Wang W., Ghannouchi F. M.* Convolutional neural network for behavioral modeling and predistortion of wideband power amplifiers // *IEEE Transactions on Neural Networks and Learning Systems*. — 2022. — Vol. 33, No. 8. — P. 3923–3937. — DOI: 10.1109/TNNLS.2021.3054867
- Kim J., Silva-Martinez J.* Low-power, low-cost CMOS direct-conversion receiver front-end for multistandard applications // *IEEE Journal of Solid-State Circuits*. — 2013. — Vol. 48, No. 9. — P. 2090–2103.
- Kingma D. P., Ba J.* Adam: A method for stochastic optimization // *arXiv preprint*. — 2014. — arXiv:1412.6980
- Kristensen A. T., Burg A., Balatsoukas-Stimming A.* Advanced machine learning techniques for self-interference cancellation in full-duplex radios // Paper presented at the 2019 53rd Asilomar Conference on Signals, Systems, and Computers, 3–6 November 2019. — 2020.
- 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.

- Maslovskiy A., Pasechnyuk D., Gasnikov A., Anikin A., Rogozin A., Gornov A., Antonov L., Vlasov R., Nikolaeva A., Begicheva M.* Non-convex optimization in digital pre-distortion of the signal // International Conference on Mathematical Optimization Theory and Operations Research. — 2021. — P. 54–70.
- Nocedal J.* Updating quasi-Newton matrices with limited storage // Mathematics of computation. — 1980. — Vol. 35, No. 151. — P. 773–782.
- Ploder O., Motz C., Paireder T., Huemer M.* A neural network approach for the cancellation of the second-order-intermodulation distortion in future cellular RF transceivers // 2019 53rd Asilomar Conference on Signals, Systems, and Computers. — 2019. — DOI: 10.1109/IEEECONF44664.2019.9048754
- Ramella M., Fabiano I., Manstretta D., Castello R.* A SAW-less 2.4-GHz receiver front-end with 2.4-mA battery current for SoC coexistences // IEEE Journal of Solid-State Circuits. — 2017. — Vol. 52, No. 9. — P. 2292–2305.
- Syrjala V., Valkama M., Anttila L., Riihonen T., Korpi D.* 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. — Gothenburg, Sweden: Chalmers Reproservice, 2012.
- Zhu C., Byrd R. H., Lu P., Nocedal J.* Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization // ACM Transactions on mathematical software (TOMS). — 1997. — Vol. 23, No. 4. — P. 550–560.