

## **Оценка точности и синтез нелинейных импульсных математических операций**

**Романов А.М.**

*"МИРЭА - Российский технологический университет", проспект Вернадского, 78,*

*Москва, 119454, Россия*

*e-mail: [romanov@mirea.ru](mailto:romanov@mirea.ru)*

**Статья поступила 26.05.2020**

### **Аннотация**

Программируемые логические интегральные схемы (ПЛИС) широко применяются в аэрокосмической отрасли как для реализации бортовых устройств, так и наземного оборудования. Данная работа посвящена оценке точности и синтезу нелинейных импульсных математических операций (ИМО). ИМО или прямая обработка сигма-дельта модулированных импульсных потоков является перспективным методом сокращения ресурсоёмкости ПЛИС, который позволяет реализовывать сложные алгоритмы цифровой обработки сигналов, в том числе на базе радиационно-стойких отечественных ПЛИС, чьи ресурсы существенно меньше, чем у зарубежных аналогов в коммерческом исполнении. Ранее полученные оценки точности нелинейных ИМО являлись эмпирическими и охватывали лишь малый набор математических операций в ограниченном диапазоне возможных параметров реализации. В данной работе был предложен новый подход к анализу точности и синтезу нелинейных ИМО, который позволяет проводить произвольную нелинейную цифровую обработку сигналов в форме сигма-дельта модулированных

импульсных потоков, обеспечивая заданную точность. Данный подход позволяет гибко определять соотношение между ресурсоёмкостью блоков ИМО и точностью вычислений. На примере реализации умножения импульсных потоков было показано превосходство предложенного решения над ранее известным широко применяемым аналогом.

**Ключевые слова:** нелинейные импульсные математические операции, сигма-дельта модуляция, прямая обработка импульсных потоков, программируемые логические интегральные схемы.

## Введение

Программируемые логические интегральные схемы (ПЛИС) широко применяются в аэрокосмической отрасли как для реализации бортовых устройств, так и наземного оборудования [1-<http://trudymai.ru/published.php?ID=100780>

3]. При реализации цифровой обработки сигналов на базе ПЛИС неизбежно встает вопрос сокращения ресурсоёмкости аппаратно-программного обеспечения [<http://trudymai.ru/published.php?ID=69720>

4]: чем меньше аппаратных ресурсов требует каждая операция, тем меньше энергии требуется для её выполнения, и тем более дешевые ПЛИС можно использовать для её реализации. Ещё более остро данная проблема стоит при использовании ПЛИС отечественного производства [<http://trudymai.ru/published.php?ID=91978>

2, <http://trudymai.ru/published.php?ID=69720>

4] и радиационно стойких ПЛИС [<http://trudymai.ru/published.php?ID=100780>

3], которые традиционно имеют меньше доступных для разработчика ресурсов, чем их коммерческие аналоги зарубежного производства.

Перспективным подходом для сокращения ресурсоёмкости ПЛИС является прямая обработка сигма-дельта модулированных импульсных потоков (ИП). Суть данного подхода заключается в использовании для кодирования сигналов не импульсно-кодовой модуляции (ИКМ), а высокочастотной сигма-дельта модуляции, и проведении математических операций непосредственно над сигма-дельта модулированными импульсными сигналами.

Исследователи неоднократно обращались к подобным подходам, начиная с 70-х годов 20 века [<http://trudymai.ru/published.php?ID=105741>

5-17]. В начале 21 века появляются работы, посвященные применению этих подходов при создании аппаратно-программного обеспечения ПЛИС [14-17]. В данных работах неизменно демонстрируется существенная экономия логических ресурсов, требуемых для аппаратной реализации математических операций, однако, несмотря на общий объем исследований, они были посвящены решению конкретных прикладных проблем и синтезу отдельных вычислительных блоков, не давая общих подходов к синтезу произвольных математических операций над сигма-дельта модулированными импульсными потоками, обладающих заданной точностью.

В работе [18] мной впервые в мире была сформулирована законченная теория, включающая в себя обобщённую структуру импульсной математической операции (ИМО), необходимую терминологию для описания ИМО и математический аппарат

оценки точности ИМО, в совокупности позволяющие синтезировать произвольные математические преобразования с требуемой точностью. Тем не менее, оценка точности нелинейных вычислительных блоков, проведенная в [18], по сути, являлась эмпирической: она охватывала только малый набор математических операций в ограниченном диапазоне возможных параметров реализации. Устранению этого пробела и посвящена данная работа.

### **Оценка точности и синтез нелинейных импульсных математических операций**

Важной особенностью реализации нелинейных ИМО — это то, что для их выполнения обязательно требуется демодуляция ИП. Дело в том, что ни один единичный импульс ИП в каждый момент времени не отражает мгновенное значение соответствующего ему ИКМ или аналогового сигнала. В [19] было экспериментально показано, что в случае демодуляции входного сигнала усредняющим фильтром, выход которого поступает на обычную нелинейную математическую операцию, построенная таким образом нелинейная ИМО будет осуществлять линейную интерполяцию заложенной в нее математической операции. В этой же работе отмечалось, что обязательным условием для работы нелинейной ИМО является наличие передискретизации (oversampling), то есть существенного превышения частоты ИП над частотой входного сигнала.

Для аналитической оценки точности нелинейных ИМО сделаем предположение, что выполнение условия передискретизации соответствует постоянному значению  $x$ , на основе которого формируется входной ИП  $X$ . В

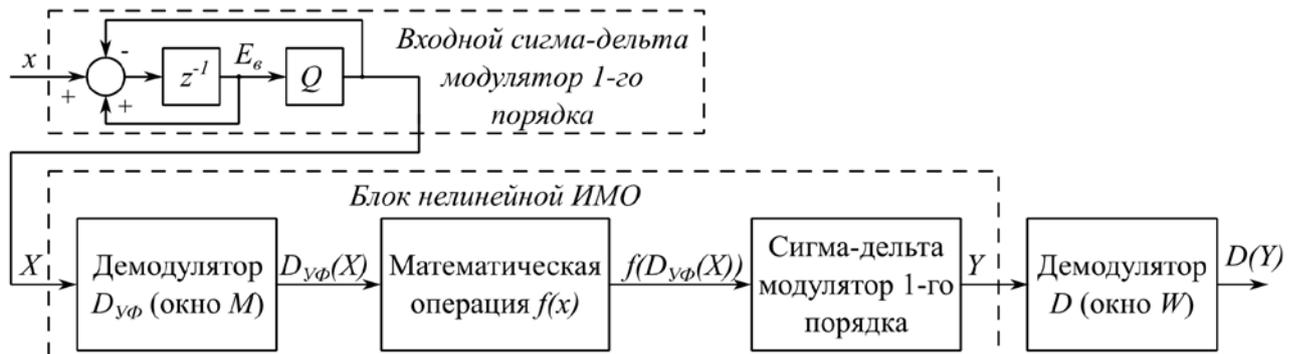
действительности данное допущение не является крайне жестким, так как реальные сигналы в контурах систем управления исполнительного уровня изменяются с частотой много меньшей, чем возможная частота ИП. В то же время, если при реализации линейных операций частоту изменения сигналов соответствующих ИП можно было не учитывать, то при реализации нелинейных ИМО может потребоваться установка дополнительных входных фильтров. С учетом этого допущения докажем следующее утверждение.

**Утверждение 1:** Пусть задан блок, реализующий нелинейную ИМО  $Y(X)$ , состоящий из цифрового усредняющего фильтра  $D_{УФ}$  с окном  $M$  импульсов, выход которого поступает на математическую операцию  $f(x)$ , выход которой подключен на сигма-дельта модулятор первого порядка, который формирует выходной ИП  $Y$ , и на вход этого блока нелинейной ИМО поступает трехуровневый ИП  $X$ , который формируется при помощи сигма-дельта модулятора первого порядка с трехуровневым квантователем, функция квантования которого соответствует реле с зоной нечувствительности, и описывается выражением (1), на вход которого поступает ИКМ константа  $x$ , такая что  $-1 \leq x \leq 1$ . Тогда нелинейная ИМО  $Y(X)$  соответствует линейной интерполяции  $f(x)$  по  $2 \cdot M + 1$  опорным точкам, равномерно распределенным на диапазоне  $[-1; 1]$  при демодуляции в цифровом усредняющем

фильтре  $D$  с окном  $W > M$  с точностью  $\frac{2 \cdot M + 1}{W}$ .

$$Q(x) = \begin{cases} R\left((x+1)\frac{(N-1)}{2}\right) \frac{2}{(N-1)} - 1, & \text{при } -1 < x < 1 \\ 1, & \text{при } x \geq 1 \\ -1, & \text{при } x \leq -1 \end{cases}, \quad (1)$$

**Доказательство:** Для упрощения понимания построим структурную схему



вычислительного тракта, описанного в условии утверждения **1** (рис. 1).

Рис. 1. Структурная схема вычислительного тракта, описанного в условии утверждения **1**.

Условие постоянства сигнала  $x$  позволяет дать более точную оценку диапазона значений, которые может принимать интегратор сигма-дельта модулятора, формирующего  $X$ , по сравнению с той, что даётся в работе [18].

Для функции округления до ближайшего целого числа  $R$  верно неравенство

$$-\frac{1}{2} < R(p) - p \leq \frac{1}{2} \quad (2)$$

Подставим вместо  $p$  значение  $(E_e + 1)\frac{N-1}{2}$ , так чтобы аргумент  $R$  в выражении (2) стал таким же, как в функции квантования (1), и получим неравенство

$$-\frac{1}{2} < R\left((E_e + 1)\frac{N-1}{2}\right) - (E_e + 1)\frac{N-1}{2} \leq \frac{1}{2} \quad (3)$$

Домножим все части неравенства (3) на  $\frac{2}{N-1}$ , получим

$$-\frac{1}{2} \frac{2}{N-1} < R((E_\epsilon + 1) \frac{N-1}{2}) \frac{2}{N-1} - (E_\epsilon + 1) \leq \frac{1}{2} \frac{2}{N-1} \quad (4)$$

Упростив выражение (4), получим

$$-\frac{1}{N-1} < R((E_\epsilon + 1) \frac{N-1}{2}) \frac{2}{N-1} - 1 - E_\epsilon \leq \frac{1}{N-1} \quad (5)$$

Подставим значение  $Q(E_\epsilon)$  для  $E_\epsilon \in [-1; 1]$  из уравнения (1) и домножим все его части на -1:

$$-\frac{1}{N-1} \leq E_\epsilon - Q(E_\epsilon) < \frac{1}{N-1} \quad (6)$$

Обозначим  $E_{\epsilon|n}$  значение  $E_\epsilon$  на  $n$ -ом такте работы сигма-дельта модулятора, формирующего  $X$ , на котором  $E_{\epsilon|h} \in [-1; 1]$ .

Прибавим ко всем частям неравенства (6)  $x$ . Поскольку по условиям утверждения  $x$  — константа, результат сложения примет следующий вид:

$$-\frac{1}{N-1} + x \leq Q(E_{\epsilon|h}) - E_{\epsilon|h} + x < x + \frac{1}{N-1} \quad (7)$$

Теперь рассмотрим дискретную модель цифрового сигма-дельта модулятора первого порядка, которая описывается системой уравнений (8)

$$\begin{cases} E_{n+1} = x_n - Q(E_m) + E_n \\ X_n = Q(E_m) \end{cases}, \quad (8)$$

где  $m$  — шаг последнего квантования,  $n$  — текущий шаг расчета,  $Q$  — функция квантования,  $x$  — вход сигма-дельта модулятора,  $X$  — выход сигма-дельта модулятора,  $E$  — состояние интегратора сигма-дельта модулятора.

Подставив в (7) выражение для  $E_{\varepsilon|n+1}$  из (8), получим окончательное

неравенство

$$-\frac{1}{N-1} + x \leq E_{\varepsilon|n+1} < x + \frac{1}{N-1} \quad (9)$$

В случае, если  $E_{\varepsilon|n+1} \in [-1; 1]$ , то  $E_{\varepsilon|n+2}$  оценивается аналогично  $E_{\varepsilon|n+2}$  при

помощи выражений (6) - (9). Если  $E_{\varepsilon|n+1} \in [1; 1 + \frac{1}{N-1})$ , то  $Q(E_{\varepsilon|n+2})$  согласно (1)

примет значение 1, оценка разницы между значением интегратора и выходом квантователя модулятора примет следующий вид:

$$0 \leq E_{\varepsilon|n+1} - Q(E_{\varepsilon|n+1}) < \frac{1}{N-1} \quad (10)$$

Прибавив к обоим частям выражения (10) постоянную константу получим (11)

и (12) соответственно.

$$x \leq E_{\varepsilon|n+1} - Q(E_{\varepsilon|n+1}) + x < x + \frac{1}{N-1} \quad (11)$$

$$x \leq E_{\varepsilon|n+2} < x + \frac{1}{N-1} \quad (12)$$

Аналогичным образом можно доказать, что в случае, если

$E_{\varepsilon|n+1} \in [-1 - \frac{1}{N-1}; -1]$ , значение  $E_{\varepsilon|n+2}$  будет оцениваться неравенством (13).

$$\frac{-1}{N-1} + x \leq E_{\varepsilon|n+2} \leq x \quad (13)$$

Нетрудно заметить, что с одной стороны, выражения (9), (12), (13) покрывают все возможные частные случаи изменения значения интегратора сигма-дельта модулятора при постоянном входном сигнале  $x \in [-1; 1]$ , а с другой стороны, если значение  $E_{\sigma}$  на текущем периоде удовлетворяет условиям (12) и (13), то оно также будет удовлетворять условию (9). Таким образом, (9) дает оценку значениям интегратора сигма-дельта модулятора, формирующего ИП Х.

На основе полученной оценки уточним оценку выходного сигнала усредняющего фильтра, который входит в состав определенного в утверждении блока нелинейной ИМО.

На основе выражений (8), описывающих работу сигма-дельта модулятора первого порядка, запишем уравнения для значения сигнала на выходе интегратора в течение М последовательных периодах импульсного элемента модулятора, приняв за нулевой период произвольный период после того, как впервые выполнилось условие  $-1 \leq E_{\sigma} \leq 1$ , что неминуемо произойдет согласно теореме 1 из работы [18].

$$E_{\sigma|1} = x - X_0 + E_{\sigma|0}, \quad E_{\sigma|2} = x - X_1 + E_{\sigma|1}, \quad \dots, \quad E_{\sigma|M+1} = x - X_M + E_{\sigma|M} \quad (14)$$

Теперь выразим  $E_{\sigma|n+1}$  через значение  $E_{\sigma|1}$ ,  $E_{\sigma|2}$ ,  $E_{\sigma|3}$  и так далее. В итоге после всех подстановок получится уравнение

$$E_{\sigma|M+1} = x - X_M + x - X_{M-1} \dots x - X_0 + E_{\sigma|0} \quad (15)$$

Сгруппировав почленно и перенеся  $E_{\sigma|0}$  в левую часть, получим

$$E_{\varepsilon_{M+1}} - E_{\varepsilon_0} = M x - \sum_{i=0}^M X_i \quad (16)$$

Поделим обе части выражения (16) на  $M$  и получим:

$$\frac{E_{\varepsilon_{M+1}} - E_{\varepsilon_0}}{M} = x - \frac{\sum_{i=0}^M X_i}{M} \quad (17)$$

Не трудно заметить, что  $\frac{\sum_{i=0}^M X_i}{M}$  — это значение на выходе усредняющего фильтра, заданного в условии доказываемого утверждения. Для упрощения обозначим его далее  $D_{\text{yф}}(X_M)$ . Используя оценку полученную в выражении (9), оценим рассогласование между ИКМ константой  $x$  и выходом усредняющего фильтра:

$$-\frac{2}{M(N-1)} < x - D_{\text{yф}}(X) < \frac{2}{M(N-1)} \quad (18)$$

Вследствие структуры и принципа работы цифрового усредняющего фильтра, на вход которого поступает двух- или трехуровневый ИП, его выход может принимать только дискретные значения с минимальным шагом  $\frac{1}{M}$ . При  $N = 3$  (трехуровневый поток, заданный в условиях утверждения) выражение (18) примет вид:

$$-\frac{1}{M} < x - D_{\text{yф}}(X) < \frac{1}{M} \quad (19)$$

Из (19) следует, что при постоянном значении  $x$  выход  $D_{y\phi}(X_M)$  будет принимать не более 2 разных значений в окрестности  $x$  (рис. **Ошибка! Источник ссылки не найден.**). Обозначим эти два значения как  $d_1$  и  $d_2$ .



Рис. 2. Иллюстрация диапазона возможных значений выходного сигнала усредняющего фильтра  $D_{y\phi}(X_M)$

В случае, если  $D_{y\phi}(X_M)$  продемодулировать выходным демодулятором  $D$ , то верно будет равенство (20).

$$D(D_{y\phi}(X_n)) = \frac{d_1 \cdot n_1 + d_2 \cdot n_2}{n_1 + n_2}, \quad (20)$$

где  $n_1$  и  $n_2$  — это количество тактов, в течение которых выход демодулятора  $D_{y\phi}$  принимал значения  $d_1$  и  $d_2$ , соответственно, за время, начиная с такта  $n - W + 1$  и кончая  $n$ -ым. Также можно отметить, что верно равенство  $n_1 + n_2 = W$ .

В свою очередь, выход  $F$  математической операции  $f$  (рис. 1), продемодулированный демодулятором  $D$ , будет описываться выражением (21).

$$D(f(D_{y\phi}(X_n))) = \frac{f(d_1) \cdot n_1 + f(d_2) \cdot n_2}{n_1 + n_2}, \quad (21)$$

Выделим из правой части выражений (20) и (21) множитель  $\frac{n_1}{n_1 + n_2}$  и получим выражения (22) и (23) соответственно.

$$D(D_{y\phi}(X_n)) - d_2 = \frac{n_1}{n_1 + n_2} (d_1 - d_2) \quad (22)$$

$$D(f(D_{y\phi}(X_n))) - f(d_2) = \frac{n_1}{n_1 + n_2} (f(d_1) - f(d_2)) \quad (23)$$

Поделив (22) на (23), получим выражение (24).

$$\frac{D(f(D_{y\phi}(X_n))) - f(d_2)}{D(D_{y\phi}(X_n)) - d_2} = \frac{f(d_1) - f(d_2)}{d_1 - d_2} \quad (24)$$

Легко заметить, что (24) — это уравнение прямой, проходящей через точки  $(d_1, f(d_1))$  и  $(d_2, f(d_2))$ . Исходя из расположения  $d_1$  и  $d_2$  для каждого значения  $x$  (рис. **Ошибка! Источник ссылки не найден.**), можно сделать вывод, что  $D(f(D_{y\phi}(X_n)))$  — это линейная интерполяция  $f(x)$  по  $2 \cdot M + 1$  опорным точкам, равномерно распределенным на диапазоне  $[-1; 1]$  в точке  $D_{y\phi}(X_n)$ .

Выразим  $D(f(D_{y\phi}(X_n)))$  через  $D_{y\phi}(X_n)$ :

$$D(f(D_{y\phi}(X_n))) = \frac{f(d_1) - f(d_2)}{d_1 - d_2} D(D_{y\phi}(X_n)) + \frac{d_1 f(d_2) - d_2 f(d_1)}{d_1 - d_2} \quad (25)$$

Перенесем все элементы равенства (25) в левую часть.

$$D(f(D_{y\phi}(X_n))) - \frac{f(d_1) - f(d_2)}{d_1 - d_2} D(D_{y\phi}(X_n)) - \frac{d_1 f(d_2) - d_2 f(d_1)}{d_1 - d_2} = 0 \quad (26)$$

$Y_n$  получается путем преобразования  $f(D_{y\phi}(X_n))$  в трехуровневый импульсный поток при помощи сигма-дельта модулятора 1-го порядка. Согласно следствию 4, приведенному в работе [18], ошибка преобразования, вносимая таким модулятором при демодуляции выходного сигнала в усредняющем фильтре  $D$ , не превышает  $\frac{3}{W}$ .

Таким образом, на основе выражения (26) можно записать неравенство (27).

$$\left| D(Y_n) - \frac{f(d_1) - f(d_2)}{d_1 - d_2} D(D_{y\phi}(X_n)) - \frac{d_1 f(d_2) - d_2 f(d_1)}{d_1 - d_2} \right| \leq \frac{3}{W} \quad (27)$$

Если предположить, что значения  $X_{n-W+1-(M-1)} \dots X_{n-W}$  равны соответственно значениям  $X_{n-M} \dots X_n$ , то было бы верно выражение (28), т. к. каждый выход фильтра  $D_{y\phi}$  войдет в итоговую сумму в выходном фильтре  $D$   $M$ -раз.

$$D(D_{y\phi}(X_n)) = M \frac{1}{M} \frac{1}{W} \sum_{i=n-W+1}^n X_i = D(X_n) \quad (28)$$

В действительности же первые  $M-1$  значение на выходе  $D_{y\phi}$  будут формироваться на основе значений  $X_n$  за пределами окна фильтра  $D$  (рис.2).

Однако, как было показано ранее, при постоянном значении  $x$  выход  $D_{y\phi}$  принимает только два возможных значения, которые отличаются друг от друга на

$\frac{1}{M}$ . Соответственно, максимальное отклонение  $D(D_{y\phi}(X_n))$  от  $D(X_n)$  в условиях

постоянного значения  $x$  не будет превышать  $\frac{M-1}{M} \frac{1}{W}$ , что описывается неравенством (29).

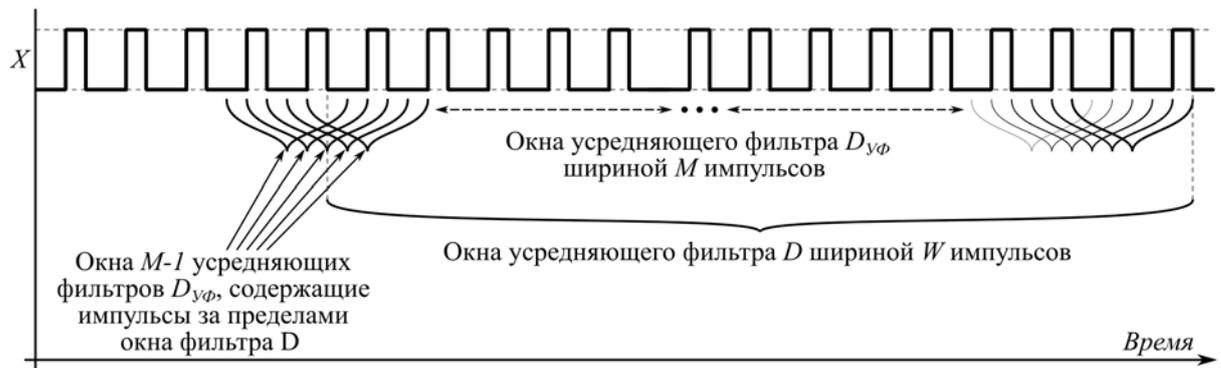


Рис. 2. Значения  $D_{уф}$ , сформированные на основе импульсов ИП  $X$  за пределами окна фильтра  $D$ .

$$\left| D(D_{уф}(X_n)) - D(X_n) \right| \leq \frac{M-1}{M W} \quad (29)$$

На основе выражений (27) и (29) запишем оценку точности соответствия нелинейной ИМО линейной интерполяции  $f(x)$  (30).

$$\left| D(Y_n) - \left( \frac{f(d_1) - f(d_2)}{d_1 - d_2} D(X_n) + \frac{d_1 f(d_2) - d_2 f(d_1)}{d_1 - d_2} \right) \right| \leq \frac{3}{W} + \frac{M-1}{M W} \max \left( \left| \frac{f(d_1) - f(d_2)}{d_1 - d_2} \right| \right) \quad (31)$$

подставив  $|d_1 - d_2| = \frac{1}{M}$  в выражение (31), получим

$$\left| D(Y_n) - \left( \frac{f(d_1) - f(d_2)}{d_1 - d_2} D(X_n) + \frac{d_1 f(d_2) - d_2 f(d_1)}{d_1 - d_2} \right) \right| \leq \frac{3 + (M-1) \cdot \max \left( \left| \frac{f(d_1) - f(d_2)}{d_1 - d_2} \right| \right)}{W} \quad (32)$$

исходя из того, что для сохранения возможности модуляции выхода математической операции  $f(x) \in [-1; 1]$  максимальное значение  $|f(d_1) - f(d_2)|$  не превышает 2 вне зависимости от конкретного типа функции  $f(x)$ . Это позволяет упростить (32) до (33).

$$\left| D(Y_n) - \left( \frac{f(d_1) - f(d_2)}{d_1 - d_2} D(X_n) + \frac{d_1 f(d_2) - d_2 f(d_1)}{d_1 - d_2} \right) \right| \leq \frac{2 \cdot M + 1}{W} \quad (33)$$

Из выражения (33) следует, что по определению соответствия между математической операцией над ИКМ сигналами и ИМО [18] нелинейная ИМО  $Y(X)$  соответствует линейной интерполяции  $f(x)$  по  $2 \cdot M + 1$  опорным точкам, равномерно распределенным на диапазоне  $[-1; 1]$  при демодуляции в цифровом усредняющем фильтре  $D$  с окном  $W > M$  с точностью  $\frac{2 \cdot M + 1}{W}$ , что и требовалось доказать.

Проанализируем утверждение **1** и его доказательство, так это позволит понять основные компромиссы, которые необходимо найти при создании блоков нелинейных ИМО.

Наиболее строгим допущением в формулировке утверждения **1** является постоянство сигнала  $x$ , на основе которого формируется входной ИП  $X$ , поскольку сигналы в контуре системы управления непрерывно меняются, и расчет нелинейной операции над постоянным сигналом вроде бы не имеет практического смысла. В то же время, анализируя доказательство утверждения **1**, можно заметить, что достаточным является постоянство значения  $x$  в течение последних  $W + M - 1$  ИП  $X$ . Данное условие является вполне выполнимым для многих систем управления роботов. Так, например, при частоте ИП в 50 МГц,  $W = 2^{16} - 1$  и  $M = 2^8 - 1$ ,  $x$  должен не менять свое значение на протяжении не менее 1.32 мс. Если говорить о контурах регулирования положения, то их сигналы в большинстве случаев изменяются в куда меньшем темпе.

Также стоит отметить, что оценка точности соответствия, которая дается в утверждении **1**, является универсальной и независимой от типа аппроксимируемой нелинейности, что делает её довольно грубой. Для уточнения этой оценки обратимся к выражению (32). Видно, что в его правой части  $(M-1)$  умножается максимальное абсолютное изменение нелинейной функции  $f$  между соседними опорными точками линейной аппроксимации. Оценка в формулировке утверждения **1** получается принятием этой разницы за 2 (максимальное возможное значение) вне зависимости от  $M$ . В реальности же, для большинства нелинейных функций, применяемых на практике, модуль разности  $\max(|f(d_1) - f(d_2)|)$  при увеличении количества опорных точек будет уменьшаться. На рис. **3** представлены оценки точности соответствия математической функции  $f(x) = \sin(4\pi \cdot x)$  и реализующей её нелинейной ИМО, полученные: экспериментально, путем моделирования нелинейной ИМО (красная линия, а); при помощи оценки на основе выражения (32) (черная линия, б) и при помощи оценки на основе утверждения **1**. Для удобства интерпретации оценки представлены в логарифмическом масштабе. Все оценки получены при размере окна усредняющего фильтра  $D$  в 2047 импульсов. Также важно отметить, что оценки на рис. **3** включают не только ошибки, вносимые ИМО, но и разницу между  $f(x)$  и её линейной аппроксимацией.

Из рис. **3** видно, что обе оценки достаточно консервативны, и реальная точность нелинейной ИМО оказывается выше, чем любая из них. В то же время, нельзя забывать, что проведенные эксперименты оценивают точность только в ограниченном наборе частных случаев, поэтому их прямое сравнение с

аналитическими оценками некорректно. Тем не менее, важным является то, что все три оценки однозначно показывают нелинейную зависимость точности ИМО от ширины окна демодулятора  $D_{y\phi}$ .

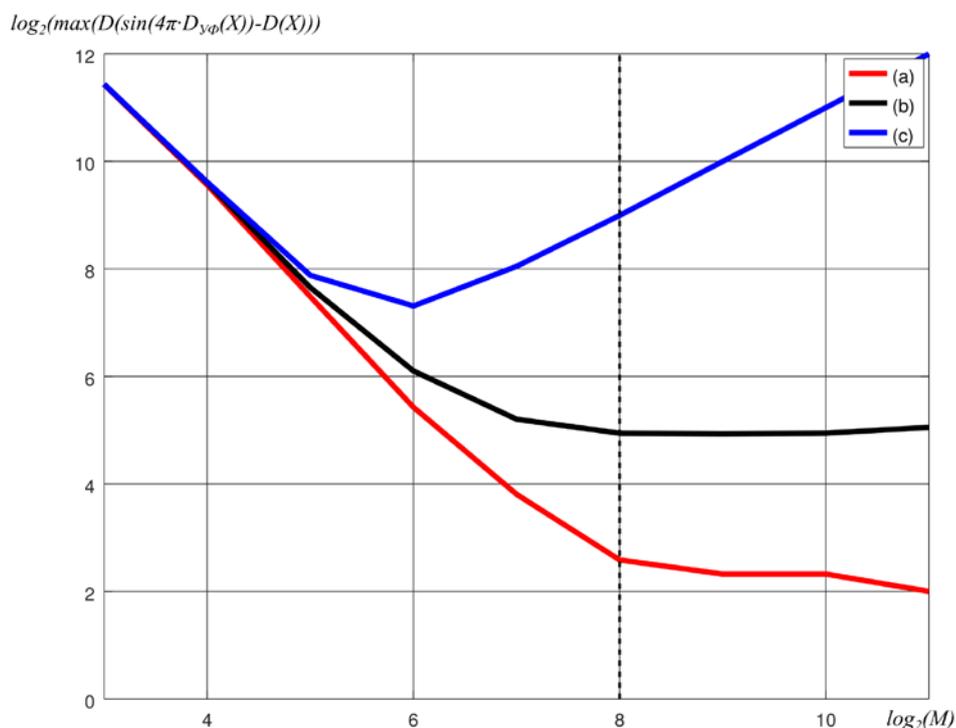


Рис. 3. Логарифмы по основанию 2 оценки точности соответствия математической функции  $f(x)=\sin(4\pi\cdot x)$  и реализующей её нелинейной ИМО, полученные (а) экспериментально, путем моделирования нелинейной ИМО; (б) при помощи оценки на основе выражения (32); (с) при помощи оценки на основе утверждения 1

Из этой зависимости видно, что с определенного момента увеличение ширины окна фильтра более не дает существенного увеличения точности, тем самым позволяя инженеру избежать нецелесообразного использования лишних ресурсов ПЛИС. При этом точка перегиба графика экспериментальной зависимости (рис. 3, а) и точка перегиба графика оценки при помощи выражения (32) расположены на линии, соответствующей  $\log_2(M)=8$ , что говорит о возможности использования выражения (32) как инструмента нахождения субоптимального размера ширины окна демодулятора  $D_{y\phi}$ .

Также надо отметить, что допущение постоянства  $x$  в условии утверждения **1** приводит к тому, что в полученных оценках никак не учитываются фазовые сдвиги, которые вносит усредняющий фильтр  $D_{y\phi}$ . При этом если рассматривать в качестве выходного демодулятора объект управления, то для минимизации влияния вычислительных процессов на качество управления задержки, вносимые фильтром  $D_{y\phi}$ , должны быть на порядок ниже минимальной постоянной времени объекта управления. Таким образом, в качестве эмпирической рекомендации предлагается использовать значения  $M$  не более чем  $0.1 \cdot W$ .

Следующее, на что необходимо обратить внимание — это тип модулятора, формирующего входной ИП. Нетрудно заметить, что при  $N=2$  согласно (18)  $D_{y\phi}(X)$  может отличаться от  $x$  более чем на  $\frac{1}{M}$ , что приведет к наличию более 2 возможных значений на выходе демодулятора  $D_{y\phi}$ . Таким образом, как и указано в условии утверждения **1**, линейная интерполяция будет выполняться только при условии, если входной ИП  $X$  сформирован трехуровневым сигма-дельта модулятором. Данный факт подтверждается в том числе экспериментально (рис.4).

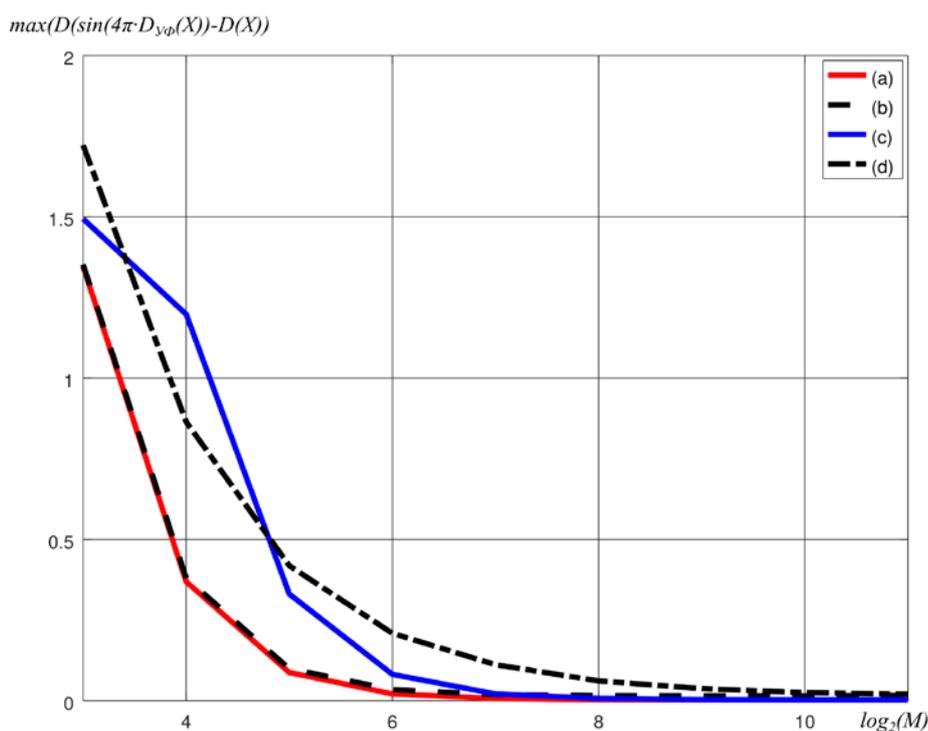


Рис. 4. Оценки точности соответствия математической функции  $f(x)=\sin(4\pi \cdot x)$  и реализующей её нелинейной ИМО, полученные (a) экспериментально, путем моделирования нелинейной ИМО с трехуровневыми ИП  $X$ ; (b) при помощи оценки на основе выражения (32); (c) экспериментально, путем моделирования нелинейной ИМО с двухуровневыми ИП  $X$ ; (d) оценка точности аппроксимации  $f(x)$  методом округления до ближайшей опорной точки

Таким образом, для реализации блока нелинейной ИМО со структурой, представленной рис. 1, обладающего заданным уровнем точности и обеспечивающего линейную интерполяцию заложенной в него нелинейности, необходимо, чтобы входной ИП этого блока был сформирован при помощи сигма-дельта модулятора первого порядка с трехуровневым квантователем на основе ИКМ сигнала, меняющего свои значения не чаще, чем  $W+M-1$  импульсов сигма-дельта модуляции. Но как выполнить это требование для сигналов внутри системы управления, формируемых на основе выходов других ИМО? Здесь возможно несколько путей. Наиболее удачной является такая организация вычислений, при которой ИП на вход нелинейных ИМО поступают с ИМО, осуществляющих

низкочастотную фильтрацию сигналов, например, интеграторов. В этом случае для такой операции будет достаточно использовать выходной модулятор с трехуровневым квантователем и оценить частоту изменения ИКМ сигнала на входе этого модулятора, исходя из фильтрующих свойств этой ИМО. Если же организовать такой порядок вычислений невозможно, то вторым решением является использование дополнительного демодулятора между предыдущей операцией и нелинейной ИМО, выход которого будет промодулирован в трехуровневый ИП. Идеальным в этом случае может стать демодулятор на основе усредняющего фильтра с фиксатором нулевого порядка, описанный в [18], так как он позволяет точно гарантировать выполнение условия постоянства сигнала  $x$  в течение  $W+M-1$  импульсов. Недостатками использования промежуточных демодуляторов являются дополнительный расход ресурсов ПЛИС на их реализацию, и что более критично, дополнительное влияние на динамику системы управления. Однако, с точки зрения теории автоматического управления, добавление усредняющего фильтра с фиксатором нулевого порядка в вычислительный тракт ИМО можно рассматривать как добавление звена задержки с фиксатором нулевого порядка и учитывать при синтезе закона управления, что позволит получить прогнозируемый результат. В случае же, если демодулятор не будет добавлен, точность нелинейной ИМО нельзя будет гарантировать, что потенциально создаст куда больше проблем для обеспечения требуемого качества управления во всех возможных режимах.

Здесь стоит отметить, что выходной модулятор самого блока нелинейной ИМО не обязан формировать трехуровневый ИП. В случае его замены на сигма-дельта

модулятор с двухуровневым квантователем оценка точности соответствия между нелинейной ИМО и соответствующей ей математической операции согласно выражениям (27) и (32) ухудшится на  $\frac{1}{W}$ , в то же время это позволит несколько сократить ресурсоёмкость реализации такой нелинейной ИМО на ПЛИС.

В данной работе не рассматривается аналитическая оценка точности нелинейных ИМО при демодуляции их в других типах демодуляторов, отличных от усредняющего фильтра, однако результаты экспериментальных исследований позволяют утверждать, что при использовании апериодического демодулятора точность соответствия нелинейных ИМО будет сопоставима с точностью, получаемой при использовании в качестве демодулятора усредняющего фильтра с

окном  $W = \frac{T + \Delta t}{\Delta t}$ , где  $T$  — постоянная времени апериодического демодулятора, а  $\Delta t$  — длительность импульсов ИП.

### **Синтез многомерных нелинейных импульсных математических операций**

Доказательство утверждения **1** для многомерного случая является крайне трудоемкой задачей. В то же время результаты экспериментальных исследований говорят о том, что описанная в этом утверждении закономерность сохраняется и при количестве входов нелинейной ИМО более 1. Таким образом, с некоторой осторожностью утверждение **1** можно распространить на многомерный случай и предположить, что для многомерных нелинейных ИМО, построенных на предварительной демодуляции входных ИП при помощи усредняющих фильтров с

окном  $M$  (рис. 5), выходной сигнал будет многомерной линейной аппроксимацией математической функции, опорные точки которой равномерно распределены по каждой из координат в диапазоне  $[-1;1]$  с шагом  $\frac{1}{M}$ .

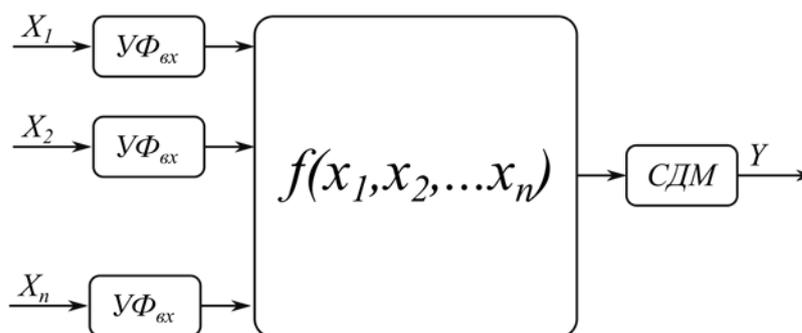


Рис. 5. Обобщенная структура блока нелинейной ИМО

Для оценки эффективности данной структуры сравним построенный на её основе блок умножения ИП с широко известным умножителем Фуджисака, реализованным в версии для трехуровневой сигма-дельта модуляции [13, 15-17]. Результаты сравнения представлены в таблице 1. Параметры ресурсоёмкости обоих умножителей получены в результате их синтеза в ПЛИС Xilinx Spartan-3.

Как видно при  $M=3$ , предложенный умножитель имеет такую же максимальную ошибку и в два раза большее среднее квадратическое отклонение результата от эталонного умножения по сравнению с умножителем Фуджисака. При этом предложенный умножитель эффективнее более чем в 3 раза с точки зрения экономии ресурсов в ПЛИС. При  $M=7$  ресурсоёмкость предложенного умножителя

на базе усредняющих фильтров становится сравнима с умножителем Фуджисака, а точность вычислений по сравнению с ним улучшается приблизительно в 10 раз.

Таблица 1. Результаты сравнения умножителя Фуджисака и полученного на основе результатов данной работы умножителя на базе усредняющих фильтров

Тип умножителя	Кол-во использованных SLICE	Кол-во использованных LUT	Кол-во использованных LUTRAM	Максимальная ошибка	Средне-квадратическое отклонение
Умножитель Фуджисака	138	209	0	0.25	0.01
Умножитель на базе усредняющих фильтров с окном ( $M=3$ )	31	43	22	0.25	0.02
Умножитель на базе усредняющих фильтров с окном ( $M=7$ )	76	141	88	0.0283	0.0023

### Заключение

В данной работе был предложен новый подход к анализу точности и синтезу нелинейных ИМО, который позволяет проводить произвольную нелинейную цифровую обработку сигналов в форме ИП, обеспечивая заданную точность. Данный подход позволяет гибко разрешать компромисс между ресурсоёмкостью блоков ИМО и точностью вычислений. На примере реализации умножения ИП было показано превосходство предложенного решения над ранее известным и широко применяемым умножителем Фуджисака. Несмотря на то, что предложение использовать частичную демодуляцию входных ИП в усредняющих фильтрах для выполнения нелинейных операций предлагалось и ранее (например, [20]), в данной работе впервые исследовано и предложено теоретическое обоснование зависимости

точности операций от параметров входных демодуляторов, а также описаны ограничения на их применяемость.

### Библиографический список

1. Кузьмин Р.Э. Фазовая автоподстройка приемных каналов систем спутниковой связи // Труды МАИ. 2018. № 99. URL: <http://trudymai.ru/published.php?ID=91978>
2. Матафонов Д.Е. Создание и отработка маршрутизатора в стандарте SpaceWire на отечественной программируемой логической интегральной схеме // Труды МАИ. 2018. № 103. URL: <http://trudymai.ru/published.php?ID=100780>
3. Муллов К.Д. Воздействие космической радиации на цифровые устройства на базе ПЛИС и методы повышения радиационной стойкости данных систем // Труды МАИ. 2016. № 87. URL: <http://trudymai.ru/published.php?ID=69720>
4. Романов А.М. Новая архитектура подчиненных устройств Ethernet POWERLINK на базе программируемых логических интегральных схем // Труды МАИ. 2019. № 106. URL: <http://trudymai.ru/published.php?ID=105741>
5. Lockhart G.B. Digital encoding and filtering using delta modulation // Radio and Electronic Engineer, 1972, vol. 42, no. 12, pp. 547 - 551. DOI: [10.1049/ree.1972.0093](https://doi.org/10.1049/ree.1972.0093)
6. Kouvaras N. Operations on delta-modulated signals and their application in the realization of digital filters // Radio and Electronic Engineer, 1978, vol. 48, no. 9, pp. 431 - 438. DOI:[10.1049/ree.1978.0062](https://doi.org/10.1049/ree.1978.0062)

7. Zrilic D. et al. Implementing signal processing functions on ternary encoded delta-modulated pulse streams // IEEE International Symposium on Circuits and Systems, IEEE, 1988, pp. 1553 - 1556. DOI: [10.1109/ISCAS.1988.15227](https://doi.org/10.1109/ISCAS.1988.15227)
8. Wong P.W., Gray R.M. FIR filters with sigma-delta modulation encoding // IEEE Transactions on Acoustics, Speech, and Signal Processing, 1990, vol. 38, no. 6, pp. 979 - 990. DOI: [10.1109/29.56058](https://doi.org/10.1109/29.56058)
9. Murray A.F., Del Corso D., Tarassenko L. Pulse-stream VLSI neural networks mixing analog and digital techniques // IEEE Transactions on Neural Networks, 1991, vol. 2, no. 2, pp. 193 - 204. DOI: [10.1109/72.80329](https://doi.org/10.1109/72.80329)
10. Kershaw S.M., Sandler M.B. Digital signal processing on a sigma-delta bitstream // IEE Colloquium on Oversampling techniques and sigma-delta modulation, IET, 1994, pp. 9/1 - 9/8.
11. Pneumatikakls A., Anastassopoulos V., Deliyannis T. Realization of a high-order IIR delta sigma filter // International journal of electronics, 1995, vol. 78, no. 6, pp. 1071 - 1089.
12. Angus J. Direct DSP on sigma-delta encoded audio signals // Audio Engineering Society Conference: UK 14th Conference: Audio-The Second Century, Audio Engineering Society, 1999.
13. Fujisaka H. et al. Bit-stream signal processing and its application to communication systems // IEE Proceedings-Circuits, Devices and Systems, 2002, vol. 149, no. 3, pp. 159 - 166. DOI: [10.1049/ip-cds:20020396](https://doi.org/10.1049/ip-cds:20020396)

14. Wu X. et al. A novel processor architecture for real-time control //Asia-Pacific Conference on Advances in Computer Systems Architecture, Springer, Berlin, Heidelberg, 2006, pp. 270 - 280. DOI: [10.1007/11859802\\_22](https://doi.org/10.1007/11859802_22)
15. Ng C.W., Wong N., Ng T.S. Tri-level bit-stream signal processing circuits and applications // International Conference on Signal Processing and Communication Systems, 2007.
16. Ng C.W., Wong N., Ng T.S. Bit-stream adders and multipliers for tri-level sigma–delta modulators // IEEE Transactions on Circuits and Systems II: Express Briefs, 2007, vol. 54, no. 12, pp. 1082 - 1086. DOI: [10.1109/TCSII.2007.906173](https://doi.org/10.1109/TCSII.2007.906173)
17. Sadik A.Z., O'Shea P.J. Realization of ternary sigma-delta modulated arithmetic processing modules // EURASIP Journal on Advances in Signal Processing, 2009, vol. 2009, no. 1, pp. 574627.
18. Романов А.М. Развитие технологии сигма-дельта модуляции для создания в архитектуре ПЛИС ресурсоёмких устройств управления мехатронно-модульными системами: дисс. канд. техн. наук: 05.13.05. – М.: МИРЭА, 2013. – 234 с.
19. Romanow A., Romanow M. FPGA based implementation of content-addressed memory based on using direct sigma-delta bitstream processing // 2016 IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference (EIconRusNW), IEEE, 2016, pp. 320 - 324. DOI: [10.1109/EIconRusNW.2016.7448184](https://doi.org/10.1109/EIconRusNW.2016.7448184)
20. Maloberti F. Non conventional signal processing by the use of sigma delta technique: a tutorial introduction // IEEE International Symposium on Circuits and Systems, 1992, vol. 6, no. 6, pp. 2645 - 2648. DOI: [10.1109/ISCAS.1992.230677](https://doi.org/10.1109/ISCAS.1992.230677)

# Accuracy Estimation and Synthesis of Nonlinear Direct Mathematical Operations on Bitstreams

**Romanov A.M.**

*MIREA - Russian Technological University,  
78, Vernadsky prospect, Moscow, 119454, Russia  
e-mail: [romanov@mirea.ru](mailto:romanov@mirea.ru)*

## **Abstract**

Field programmable gate arrays (FPGA) are widely used in the aerospace industry, both for on-board devices and ground equipment. This paper is devoted to accuracy estimation and synthesis of nonlinear direct mathematical operations on bitstreams (DMOB). DMOB or direct processing of sigma-delta modulated bitstreams is a promising method to reduce the resource area consumption of the FPGA designs which allows to implement complex digital signal processing algorithms on the basis of radiation-resistant FPGAs as well as FPGAs produced by Russian vendors, the logical resources of which are significantly smaller than those of foreign counterparts in commercial versions. The estimates of nonlinear DMOB accuracy which were previously obtained were empirical and covered only a small set of mathematical operations in a limited range of possible implementation parameters. This paper proposes a new approach to accuracy analysis and synthesis of nonlinear DMOBs which allows to implement a wide range of nonlinear digital signal processing algorithms directly on sigma-delta modulated bitstreams providing a given accuracy. The key point of this paper is that nonlinear DMOBs implemented by combining input averaging filters and nonlinear mathematical operation

perform linear interpolation of this mathematical operation, and the number of interpolation reference points depends on the filter window. This statement supported by a number of assumptions is analytically proved for a one-dimensional case, and it is experimentally confirmed with regard to a multidimensional case. All the assumptions used in the proof are analyzed in detail, and the conditions of the proposed method of applicability for real projects are formulated on the basis of the assumptions. Solutions described in this paper allow to achieve a flexible compromise on the nonlinear DMOB FPGA core area usage and the computation accuracy. The example of the implementation of bitstream multiplier demonstrates that the proposed solution is superior to the previously known and widely used analogue.

**Keywords:** nonlinear bitstream mathematical operations, sigma-delta modulation, direct bitstream processing, field programmable gate arrays.

### References

1. Kuz'min R.E. *Trudy MAI*, 2018, no. 99, available at: <http://trudymai.ru/eng/published.php?ID=91978>
2. Matafonov D.E. *Trudy MAI*, 2018, no. 103, available at: <http://trudymai.ru/eng/published.php?ID=100780>
3. Mullov K.D. *Trudy MAI*, 2016, no. 87, available at: <http://trudymai.ru/eng/published.php?ID=69720>

4. Romanov A.M. *Trudy MAI*, 2019, no. 106, available at: <http://trudymai.ru/eng/published.php?ID=105741>
5. Lockhart G.B. Digital encoding and filtering using delta modulation, *Radio and Electronic Engineer*, 1972, vol. 42, no. 12, pp. 547 - 551. DOI: [10.1049/ree.1972.0093](https://doi.org/10.1049/ree.1972.0093)
6. Kouvaras N. Operations on delta-modulated signals and their application in the realization of digital filters, *Radio and Electronic Engineer*, 1978, vol. 48, no. 9, pp. 431 - 438. DOI:[10.1049/ree.1978.0062](https://doi.org/10.1049/ree.1978.0062)
7. Zrilic D. et al. Implementing signal processing functions on ternary encoded delta-modulated pulse streams, *IEEE International Symposium on Circuits and Systems*, IEEE, 1988, pp. 1553 - 1556. DOI:[10.1109/ISCAS.1988.15227](https://doi.org/10.1109/ISCAS.1988.15227)
8. Wong P.W., Gray R.M. FIR filters with sigma-delta modulation encoding, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1990, vol. 38, no. 6, pp. 979 - 990. DOI:[10.1109/29.56058](https://doi.org/10.1109/29.56058)
9. Murray A.F., Del Corso D., Tarassenko L. Pulse-stream VLSI neural networks mixing analog and digital techniques, *IEEE Transactions on Neural Networks*, 1991, vol. 2, no. 2, pp. 193 - 204. DOI: [10.1109/72.80329](https://doi.org/10.1109/72.80329)
10. Kershaw S.M., Sandler M.B. Digital signal processing on a sigma-delta bitstream, *IEE Colloquium on Oversampling techniques and sigma-delta modulation*, IET, 1994, pp. 9/1 - 9/8.
11. Pneumatikakls A., Anastassopoulos V., Deliyannis T. Realization of a high-order IIR delta sigma filter, *International journal of electronics*, 1995, vol. 78, no. 6, pp. 1071 - 1089.

12. Angus J. Direct DSP on sigma-delta encoded audio signals, *Audio Engineering Society Conference: UK 14th Conference: Audio-The Second Century*, Audio Engineering Society, 1999.
13. Fujisaka H. et al. Bit-stream signal processing and its application to communication systems, *IEE Proceedings-Circuits, Devices and Systems*, 2002, vol. 149, no. 3, pp. 159 - 166. DOI: [10.1049/ip-cds:20020396](https://doi.org/10.1049/ip-cds:20020396)
14. Wu X. et al. A novel processor architecture for real-time control, *Asia-Pacific Conference on Advances in Computer Systems Architecture*, Springer, Berlin, Heidelberg, 2006, pp. 270 - 280. DOI: [10.1007/11859802\\_22](https://doi.org/10.1007/11859802_22)
15. Ng C.W., Wong N., Ng T.S. Tri-level bit-stream signal processing circuits and applications, *International Conference on Signal Processing and Communication Systems*, 2007.
16. Ng C.W., Wong N., Ng T.S. Bit-stream adders and multipliers for tri-level sigma-delta modulators, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2007, vol. 54, no. 12, pp. 1082 - 1086. DOI: [10.1109/TCSII.2007.906173](https://doi.org/10.1109/TCSII.2007.906173)
17. Sadik A.Z., O'Shea P.J. Realization of ternary sigma-delta modulated arithmetic processing modules, *EURASIP Journal on Advances in Signal Processing*, 2009, vol. 2009, no. 1, pp. 574627.
18. Romanov A.M. *Razvitie tekhnologii sigma-del'ta modulyatsii dlya sozdaniya v arkhitekture PLIS resursoemkikh ustroystv upravleniya mekhatronno-modul'nyimi sistemami* (Development of sigma-delta modulation technology for the creation of

resource-intensive mechatronic-modular systems control devices in the FPGA architecture), Doctoral thesis, Moscow, MIREA, 2013, 234 p.

19. Romanow A., Romanow M. FPGA based implementation of content-addressed memory based on using direct sigma-delta bitstream processing, *2016 IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference (EIconRusNW)*, IEEE, 2016, pp. 320 - 324. DOI: [10.1109/EIconRusNW.2016.7448184](https://doi.org/10.1109/EIconRusNW.2016.7448184)

20. Maloberti F. Non conventional signal processing by the use of sigma delta technique: a tutorial introduction, *IEEE International Symposium on Circuits and Systems*, 1992, vol. 6, no. 6, pp. 2645 - 2648. DOI: [10.1109/ISCAS.1992.230677](https://doi.org/10.1109/ISCAS.1992.230677)