

## ТЕОРЕТИКО-ИНФОРМАЦИОННЫЙ АНАЛИЗ МНОГОУРОВНЕВОЙ FLASH-ПАМЯТИ

### Часть 2: Оценка параметров кодов, обеспечивающих заданную надежность хранения данных

А. Н. Трофимов<sup>а</sup>, канд. техн. наук, доцент  
Ф. А. Таубин<sup>а</sup>, доктор техн. наук, профессор

<sup>а</sup>Санкт-Петербургский государственный университет аэрокосмического приборостроения,  
Санкт-Петербург, РФ

**Введение:** многоуровневая flash-память благодаря высокой плотности записи занимает доминирующее место на рынке энергонезависимых устройств хранения данных и интенсивно используется для хранения данных в весьма обширной сфере приложений. Повышение плотности записи, достигаемое за счет уменьшающегося физического размера ячейки наряду с возрастающим количеством используемых состояний, приводит к снижению надежности хранения данных, что требует использования помехоустойчивого кодирования. **Цель:** исследование потенциальных возможностей помехоустойчивого кодирования для одной из основных моделей многоуровневой NAND flash-памяти в зависимости как от условий записи и хранения данных, так и заданного уровня надежности хранения данных. **Результаты:** в явном виде получены соотношения, связывающие между собой основные параметры гипотетических кодов, обеспечивающих заданный уровень надежности хранения данных для одной из возможных моделей flash-памяти, которая описывает страницу чипа памяти как систему с независимыми многоуровневыми ячейками. С использованием полученных соотношений устанавливаются обменные соотношения между скоростью гипотетического кода, его длиной и минимальным расстоянием, который соответствует параметрам (гипотетического) кода с наименьшим расстоянием и, следовательно, с потенциально невысокой сложностью декодирования. Скорости таких кодов для типичных случаев оказываются весьма близкими к пропускной способности канала (разница составляет примерно 2–3 %). **Практическая значимость:** полученные результаты позволяют оценить степень близости параметров конкретной схемы кодирования для flash-памяти к их предельным значениям. Показано, что для типичных значений параметров канала хранения данных flash-памяти минимальное расстояние гипотетического кода с потенциально невысокой сложностью декодирования составляет 4 или 5, а длины соответствующих кодов варьируются в пределах примерно от 600 до 900. При этом оказывается, что увеличение числа циклов перезаписи и (или) длительности хранения данных при сохранении требуемой надежности хранения данных приводит к уменьшению длины гипотетического кода с потенциально невысокой сложностью декодирования.

**Ключевые слова** — многоуровневая NAND flash-память, канал записи, граница случайного кодирования, граница Варшавова — Гилберта.

#### Введение

Увеличение плотности записи в многоуровневой flash-памяти достигается за счет уменьшения разницы между соседними пороговыми уровнями напряжения в ячейке. Это приводит к возрастанию влияния шума записи/считывания и других факторов, снижающих надежность хранения данных. Как отмечалось в работе [1], вероятность ошибки (raw bit error rate) в многоуровневой NAND flash-памяти оказывается неприемлемо высокой — порядка  $10^{-4}$  и более, тогда как требуемая вероятность ошибки лежит в диапазоне  $10^{-12}$ – $10^{-16}$ . В результате обеспечение требуемой надежности хранения данных может быть достигнуто лишь при использовании помехоустойчивого кодирования.

К настоящему времени предложено довольно много вариантов помехоустойчивого кодирования, предназначенного для многоуровневой NAND flash-памяти: коды БЧХ, коды Рида — Соломона, низкоплотностные коды, решетчатые

коды (см., например, публикации [2–8]). Отметим, что на практике в большинстве чипов с многоуровневой NAND flash-памятью используются коды БЧХ, что объясняется приемлемой (для индустрии) сложностью реализации декодеров БЧХ-кодов.

Как представляется, объективный сравнительный анализ эффективности предложенных вариантов помехоустойчивого кодирования для flash-памяти должен включать, наряду с традиционными характеристиками, и сопоставление с характеристиками, полученными на основе теоретико-информационного анализа. В данной работе в качестве таких характеристик рассматриваются параметры гипотетических кодов, обеспечивающих заданный уровень надежности хранения данных. Под *гипотетическим* кодом далее понимается код, для которого вероятность ошибки декодирования  $P_e$  кодового блока (слова) по максимуму правдоподобия удовлетворяет неравенству  $P_e < 2^{-nE(R)}$ , где  $E(R)$  — экспонента случайного кодирования. Как правило, гипотетиче-

ский код является случайным кодом и поэтому не имеет регулярной структуры. Параметры гипотетических кодов находятся с использованием а) границ случайного кодирования для многоуровневой NAND flash-памяти, полученными в работе [1], и б) границы Варшавова — Гилберта для  $q$ -ичных блоковых кодов. Полученные таким образом параметры представляют собой своего рода опорные отметки, т. е. они позволяют оценить, в какой степени конкретная схема кодирования для flash-памяти близка к потенциально предельным возможностям.

### Модель канала

Модель, описывающая канал записи flash-памяти и используемая в настоящей работе, с необходимыми деталями, обоснованиями и ссылками подробно рассмотрена в работе [1]. Здесь для связности изложения приводится краткое формальное описание этой модели.

В изучаемой модели блок flash-памяти рассматривается как множество независимых ячеек. В работе [1] и здесь полагается, что физический носитель может рассматриваться как стационарный канал без заметной тенденции к пакетированию ошибок, поэтому математическая модель представляет собой модель канала без памяти, которая полностью определяется моделью одиночной ячейки flash-памяти. Входные уровни каждой ячейки принимают некоторые фиксированные значения  $x_0, x_1, \dots, x_{q-1}$ , а выходные значения представляют собой случайные величины. Распределения этих случайных величин параметризованы значениями  $x_0, x_1, \dots, x_{q-1}$  и описываются условными функциями плотности вероятности  $p_{y|x}(y|x)$ , где  $-\infty < y < \infty$ ,  $x = x_0, x_1, \dots, x_{q-1}$ . Условные функции плотности вероятности  $p_{y|x}(y|x)$  задаются следующим образом:

$$p_{y|x}(y|x) = \begin{cases} p_0(y|x), & \text{если } x = x_0; \\ p_0(y|x) * p_r(y) * p_c(y) * p_d(y), & \text{если } x = x_1, x_2, \dots, x_{q-1}. \end{cases} \quad (1)$$

В равенстве (1) использованы следующие обозначения:  $p_0(\cdot|\cdot)$  — функции плотности вероятности, задающие распределения номинальных выходных значений;  $p_r(\cdot)$ ,  $p_c(\cdot)$ ,  $p_d(\cdot)$  — функции плотности вероятности, определяющие соответственно влияние записи/стирания, взаимную интерференцию ячеек и влияние времени хранения (детали и пояснения — см. [1]). Знаком \* в правой части равенства (1) обозначена операция свертки.

Как и в работе [1], будем далее считать, что  $q = 4$  и в рассматриваемом примере значения но-

минальных, или начальных, уровней  $x_0, x_1, \dots, x_3$  равны:  $x_0 = 1,4$ ,  $x_1 = 2,6$ ,  $x_2 = 3,2$ ,  $x_3 = 3,93$ .

Функции плотности вероятности  $p_0(\cdot|\cdot)$ , задающие распределения начальных уровней, определены как гауссова плотность для  $x = x_0$

$$p_0(y|x_0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-x_0)^2}{2\sigma^2}}, \quad \sigma = 0,35$$

и как равномерные плотности для  $x = x_1, \dots, x_{q-1}$ , т. е.

$$p_0(y|x_i) = \begin{cases} \frac{1}{\Delta_i}, & \text{если } |y - x_i| \leq \frac{\Delta_i}{2}; \\ 0 & \text{в противном случае,} \end{cases}$$

$i = 1, 2, \dots, q - 1$ ;  $\Delta_1 = \Delta_2 = \Delta_3 = 0,2$ .

Функция плотности вероятности  $p_r(\cdot)$  задается как двусторонняя симметричная экспоненциальная плотность

$$p_r(y) = \frac{1}{2\lambda_r} e^{-\frac{|y|}{\lambda_r}},$$

где параметр  $\lambda_r$  связан с числом циклов записи/стирания  $N$  зависимостью  $\lambda_r = K_\lambda \sqrt{N}$ ,  $K_\lambda = 0,00025$ .

Функция плотности вероятности  $p_c(\cdot)$  в рассматриваемой модели задана как усеченная гауссова плотность, т. е. она имеет вид

$$p_c(y) = \begin{cases} \frac{c}{\sqrt{2\pi}\sigma_c} e^{-\frac{(x-\mu_c)^2}{2\sigma_c^2}}, & \text{если } |y - \mu_c| \leq w_c; \\ 0 & \text{в противном случае,} \end{cases}$$

где  $c$  — нормирующий коэффициент, обеспечивающий выполнения равенства  $\int_{-\infty}^{\infty} p_c(y) dy = 1$ , а остальные параметры имеют значения  $w_c = 0,1\mu_c$  и  $\sigma_c = 0,4\mu_c$ ,  $\mu_c = 0,2$ . Функция плотности вероятности  $p_d(\cdot)$  задается как нормальная плотность вероятности

$$p_d(y) = \frac{1}{\sqrt{2\pi}\sigma_d} e^{-\frac{(y-\mu_d)^2}{2\sigma_d^2}},$$

где величины  $\mu_d$  и  $\sigma_d$  определяются как

$$\mu_d = -K_s(x - x_0)K_d N^{0,5} \ln(1 + T / T_0);$$

$$\sigma_d = K_s(x - x_0)K_m N^{0,6} \ln(1 + T / T_0).$$

В этих выражениях, как и ранее,  $N$  обозначает число циклов перезаписи;  $T$  — время хранения;  $T_0$  — интервал времени, соответствующий одному часу;  $K_s = 0,38$ ,  $K_d = 10^{-4}$  и  $K_m = 4 \cdot 10^{-6}$ .

Функции плотности вероятности  $p_{y|x}(y|x)$  могут быть найдены численно в соответствии с равенствами (1). Для некоторых значений величин  $N$  и  $T$  функции  $p_{y|x}(y|x)$  могут быть с приемлемой точностью аппроксимированы гауссовыми

плотностями. Детали аппроксимации приведены в работе [1]. Рассмотренная модель, определяемая функциями плотности вероятности  $p_{y|x}(y|x)$  или их гауссовыми аппроксимациями, называемая также моделью с гауссовым шумом, стандартное отклонение которого зависит от входного значения (input-dependent additive Gaussian noise) [9], допускает обобщение на большее число входных уровней, в частности, на шесть, восемь и двенадцать уровней [3]. Такое обобщение не представляется сложным и может быть при необходимости выполнено.

### Параметры кодов, обеспечивающих заданный уровень надежности

С использованием построенных границ случайного кодирования для вероятности ошибки на блок можно оценить параметры гипотетического блочного кода, позволяющего достичь требуемого уровня надежности хранения данных. Оценивание основных параметров корректирующего гипотетического кода для канала flash-памяти далее выполняется на основе границы случайного кодирования, рассмотренной в работе [1], для декодера максимального правдоподобия, использующего непрерывный выход канала (мягкие решения). Практическая реализация такого, строго оптимального, алгоритма декодирования может оказаться затруднительной и даже неприемлемой для систем flash-памяти из-за высокой сложности. Возможным выходом может быть применение субоптимальных алгоритмов [10, 11]. Ключевая особенность этих и подобных им алгоритмов состоит в том, что сравнительно простой алгоритм декодирования выполняется несколько раз для различных вариантов формирования жесткого выхода канала (квантования), а окончательное решение принимается с использованием результатов этих попыток декодирования и непрерывного выхода канала. Такие алгоритмы были разработаны несколько десятилетий назад, но интерес к ним вновь возник, и они оказались востребованными при проектировании высокоскоростных систем передачи и хранения информации. Алгоритмы декодирования из этого класса очень незначительно уступают по вероятности ошибки декодирования оптимальным алгоритмам, использующим непрерывный выход канала, но в ряде случаев значительно превосходят их по сложности реализации. Заметную долю сложности таких алгоритмов составляет сложность алгебраического или комбинаторного декодирования жесткого выхода канала, поскольку на этой стадии выполняется несколько попыток декодирования. Сложность алгебраического декодирования зависит от многих параметров, и одним из важнейших факторов, влияющих на слож-

ность, является кратность исправляемых ошибок, определяемая минимальным расстоянием кода. Поэтому минимальное расстояние кода и граница вероятности ошибки для декодирования в канале с непрерывным выходом могут рассматриваться как основы для оценки параметров кода для субоптимального алгоритма декодирования.

Примем для определенности, что надежность хранения данных оценивается требуемым уровнем вероятности ошибки на бит  $P_b$ . Альтернативным выбором критерия для оценки надежности хранения данных может быть вероятность ошибки на блок. Выбор в качестве критерия вероятности ошибки на бит представляется более приемлемым, поскольку в этом случае не возникает вопрос о длине блока, разрешение которого хотя и является важным для практических приложений, может создать ненужные трудности при теоретическом рассмотрении.

Предположим, что канал описывается рассмотренной моделью многоуровневой ячейки flash-памяти. Теорема кодирования для канала без памяти [12] утверждает, что существует блочный код длины  $n$ , объема  $M$  слов, со скоростью  $R$ ,  $R = n^{-1} \log_2 M$  бит/символ канала, для которого вероятность ошибки на блок  $P_e$  ограничена сверху как

$$P_e < 2^{-nE(R)}, \quad (2)$$

где  $E(R)$  — показатель экспоненты случайного кодирования, вычисляемый по формуле

$$E(R) = \max_{p_x} \max_{0 < \rho < 1} (E_0(\rho, p_x) - \rho R), \quad (3)$$

где

$$E_0(\rho, p_x) = -\log_2 \int_{-\infty}^{\infty} \left( \sum_x p_x(x) p_{y|x}(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} dy. \quad (4)$$

Важное следствие из теоремы кодирования состоит в том, что пропускная способность канала  $C$  и функция Галлагера (4) связаны соотношением

$$C = \max_{p_x} \left. \frac{\partial E_0(p_x, \rho)}{\partial \rho} \right|_{\rho=0}.$$

Важной характеристикой канала является вычислительная скорость канала  $R_0$ , определяемая как

$$\begin{aligned} R_0 &= \max_{p_x} E_0(p_x, 1) = \\ &= -\log_2 \min_{p_x} \int_{-\infty}^{\infty} \left( \sum_x p_x(x) \sqrt{p_{y|x}(y|x)} \right)^2 dy. \end{aligned}$$

Эта величина может рассматриваться как предельное значение *практически достижимой* скорости кода.

Для решения задачи оценивания параметров кода требуется связать вероятность ошибки на бит  $P_b$  с вероятностью ошибки на блок  $P_e$ . Для этого обозначим через  $P_s$  вероятность ошибки в одном  $q$ -ичном кодовом символе, возникающей в результате неверного декодирования кодового блока, и найдем сначала связь между вероятностью ошибки на блок  $P_e$  и величиной  $P_s$ , а потом выразим  $P_b$  через  $P_s$ .

При ошибочном декодировании декодер отдает получателю кодовое слово, отличающееся от истинного слова в некотором числе символов. Наиболее вероятным будет ошибочное декодирование в пользу ближайшего слова кода, т. е. отличающегося от истинного в  $d$  символах, где  $d$  — минимальное расстояние кода. Поэтому в первом приближении можно считать, что с приемлемой точностью выполняется соотношение (см. также [13])

$$P_s \approx \frac{d}{n} P_e = \delta P_e, \quad (5)$$

где  $\delta = d/n$ . Вероятность ошибки на символ и вероятность ошибки на бит связаны неравенствами

$$(1 / \log_2 q) P_s \leq P_b \leq P_s. \quad (6)$$

Тогда с использованием выражений (3)–(6) можно записать, что существует код, для которого

$$P_b \approx \delta P_e < \delta 2^{-nE(R)}. \quad (7)$$

Величина  $\delta$  и скорость кода  $R$  могут быть связаны между собой границей Варшавова — Гилберта. Для случая  $q$ -ичных кодов эта граница может быть сформулирована как [14]

$$\begin{aligned} R &= \frac{\log_2 M}{n} \geq \frac{1}{n} \log_2 \frac{q^n}{\sum_{i=0}^d C_n^i (q-1)^i} = \\ &= \log_2 q - \frac{1}{n} \log_2 \sum_{i=0}^d C_n^i (q-1)^i \end{aligned}$$

и при  $n \rightarrow \infty$  представлена в виде

$$R \geq \log_2 q - h(\delta), \quad (8)$$

где  $h(x) = x \log_2 q - x \log_2 x - (1-x) \log_2 (1-x)$ . Полагая в выражениях (7) и (8) выполнение равенств, можно записать, что параметры кода  $n$  и  $\delta$  для требуемого значения вероятности ошибки на бит  $P_b$  могут быть заданы как функции от скорости кодирования  $R$ ,  $0 < R < C$ , в виде

$$\delta(R) = h^{-1}(\log_2 q - R); \quad (9)$$

$$n(R) = \frac{\log_2(P_b / \delta(R))}{E(R)}. \quad (10)$$

Определим также величину

$$d(R) = \delta(R)n(R). \quad (11)$$

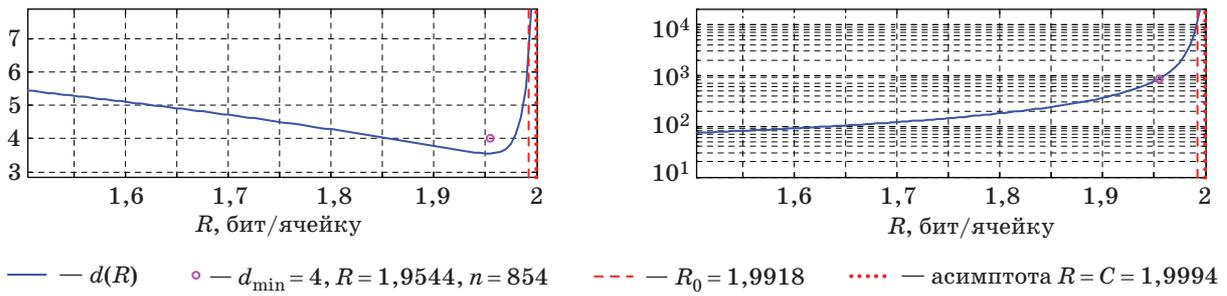
Итак, выражения (9)–(11) в совокупности устанавливают взаимосвязь между длиной и минимальным расстоянием гипотетического кода, который обеспечивает заданную вероятность ошибки на бит  $P_b$ .

### Численные результаты

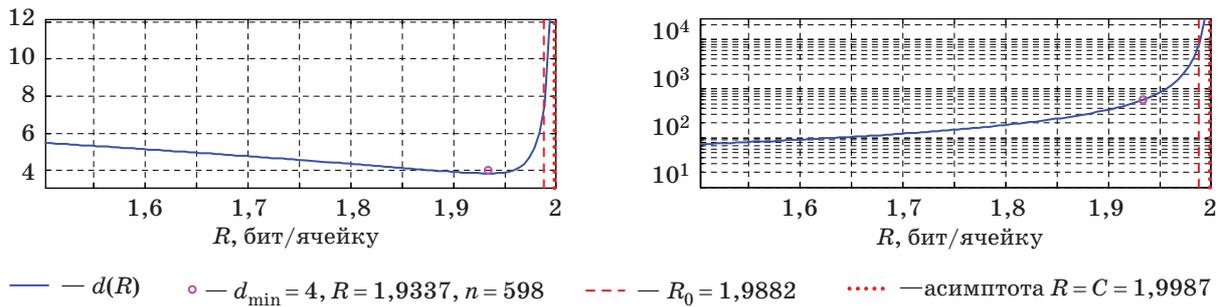
Зависимости  $d(R)$  и  $n(R)$  вычислены согласно выражениям (9)–(11) при заданных значениях вероятности ошибки на бит  $P_b = 10^{-12}$  и  $P_b = 10^{-14}$ . Результаты вычислений представлены на рис. 1–4.

Как видно из приведенных графиков, кривая  $d(R)$  имеет минимум, который соответствует параметрам (гипотетического) кода с наименьшим расстоянием и, следовательно, с малой сложностью декодирования. Этому минимуму соответствует точка на графике  $n(R)$ , определяющая длину этого кода с потенциально невысокой сложностью декодирования. Как следует из рис. 1, для  $P_b = 10^{-12}$ ,  $N = 100$ ,  $T = 1$  мес. этот код имеет минимальное расстояние  $d = 4$ , длину 850 и скорость  $R \approx 1,95$  бит/символ канала (бит/ячейку). Из рис. 2, представляющего канал записи с другими параметрами, следует, что для  $P_b = 10^{-12}$ ,  $N = 1000$ ,  $T = 12$  мес. параметры кода с потенциально небольшой сложностью имеют значения  $d = 4$ ,  $n \approx 600$ ,  $R \approx 1,93$  бит/символ канала. Рис. 3 и 4 иллюстрируют взаимосвязь параметров (гипотетического) кода для случая, когда требуемая вероятность ошибки и  $P_b = 10^{-14}$ . При числе циклов перезаписи  $N = 100$  и времени хранения  $T = 1$  мес. код с потенциально невысокой сложностью декодирования имеет параметры  $d = 4$ ,  $n \approx 900$  и скорость  $R \approx 1,95$  бит/символ канала. При увеличении числа циклов перезаписи  $N$  до 1000 и длительности хранения данных  $T$  до 12 мес. параметры кода с потенциально небольшой сложностью принимают следующие значения:  $d = 5$ ,  $n \approx 700$ ,  $R \approx 1,93$  бит/символ канала.

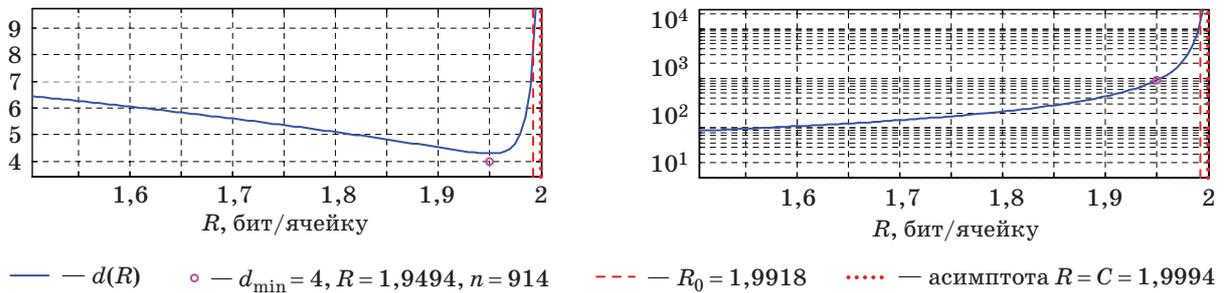
Как видим, снижение требуемой вероятности ошибки  $P_b$  на два порядка — с  $10^{-12}$  до  $10^{-14}$  — сопровождается увеличением длины гипотетического кода примерно в 1,07–1,2 раза. Увеличение числа циклов перезаписи и (или) длительности хранения данных при сохранении требуемой надежности хранения данных, наоборот, приводит к уменьшению длины гипотетического кода с потенциально невысокой сложностью декодирования. Отметим также, что скорость указанных гипотетических кодов (с потенциально невысокой сложностью декодирования) оказывается весьма близка к пропускной способности канала записи — разница составляет



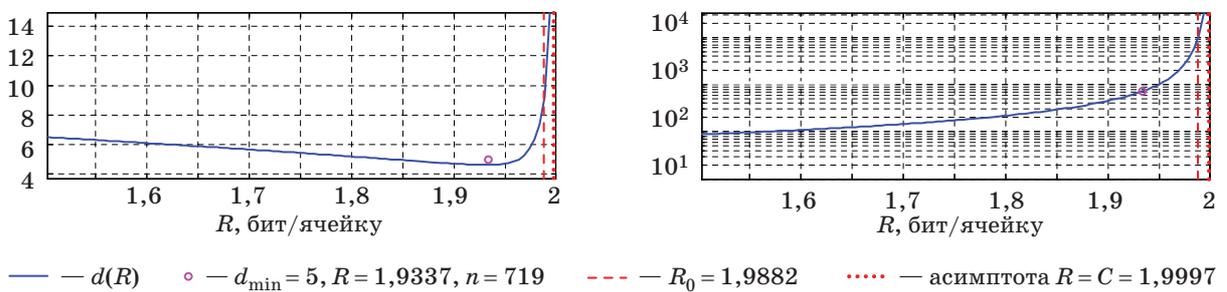
■ Рис. 1. Графики функций  $d(R)$  и  $n(R)$  для  $P_b = 10^{-12}, N = 100, T = 1$  мес.



■ Рис. 2. Графики функций  $d(R)$  и  $n(R)$  для  $P_b = 10^{-12}, N = 1000, T = 12$  мес.



■ Рис. 3. Графики функций  $d(R)$  и  $n(R)$  для  $P_b = 10^{-14}, N = 100, T = 1$  мес.



■ Рис. 4. Графики функций  $d(R)$  и  $n(R)$  для  $P_b = 10^{-14}, N = 1000, T = 12$  мес.

примерно 2,4 % (в случае  $N = 100, T = 1$  мес.) и примерно 3,3 % (в случае  $N = 1000, T = 12$  мес.).

Разумеется, приведенные значения могут служить лишь предварительными и приблизительными ориентирами для поиска параметров ре-

лизуемых схем кодирования для каналов flash-памяти. Очевидно также, что если величина скорости кода, соответствующая наименьшему значению минимального расстояния, оказывается недостаточной (или избыточность кода слишком

велика) для практических приложений, то следует сместиться по графикам  $d(R)$  и  $n(R)$  вправо и выбрать код с большим минимальным расстоянием и большей длиной. Например, для достижения скорости кода  $R = R_0 = 1,9982$  бит/символ канала при  $N = 1000$ ,  $T = 12$  мес.,  $P_b = 10^{-12}$  параметры кода должны быть порядка  $n = 7000 \dots 8000$  и  $d = 8 \dots 9$ , и в этих пределах следует проводить дальнейший поиск приемлемых вариантов. Еще раз стоит обратить внимание, что приведенное рассмотрение и примеры справедливы при использовании непрерывного выхода канала. Указанный подход может быть непосредственно распространен для исследования схем кодирования в канале записи с квантованным выходом. Также полученные оценки могут быть полезны для выбора параметров кодов и для важных для практики субоптимальных алгоритмов декодирования, использующих мягкий выход канала.

### Заключение

В работе был рассмотрен круг вопросов, связанных с оценкой параметров гипотетических кодов, обеспечивающих заданный уровень надежности хранения данных применительно к одной из возможных моделей flash-памяти. Эта модель описывает систему с независимыми многоуровневыми ячейками. Распределение выходных значений представляет собой случайную величину, определенную на множестве одиночных элементов памяти, распределение которой существенно зависит от ряда мешающих факторов, включающих влияние циклов записи/стирания, взаимную интерференцию ячеек и влияние времени хранения данных. В рассматриваемой модели предполагается отсутствие группирования искажений, поэтому канал хранения данных описывается моделью без памяти с непрерывным выходом. Такая модель задает наилучшее использование выхода канала. Параметры гипотетических кодов определялись путем комбинирования границ случайного кодирования для многоуровневой NAND flash-памяти, полученными в работе [1], и границы Варшамова — Гилберта для  $q$ -ичных блоковых кодов.

На основе представленных результатов можно сформулировать следующие выводы:

— функция  $d(R)$ , описывающая зависимость минимального расстояния кода от скорости  $R$ ,

имеет минимум, который соответствует параметрам (гипотетического) кода с наименьшим расстоянием и, следовательно, с потенциально невысокой сложностью декодирования;

— для типичных значений параметров канала хранения данных flash-памяти минимальное расстояние гипотетического кода с потенциально невысокой сложностью декодирования составляет 4 или 5; при этом длины кодов  $n$  варьируются в ориентировочных пределах 600–900;

— снижение требуемой вероятности ошибки  $P_b$  на два порядка (с  $10^{-12}$  до  $10^{-14}$ ) сопровождается увеличением длины гипотетического кода примерно в 1,07–1,2 раза;

— увеличение числа циклов перезаписи и (или) длительности хранения данных при сохранении требуемой надежности хранения данных приводит к уменьшению длины гипотетического кода с потенциально невысокой сложностью декодирования;

— скорость указанных гипотетических кодов (с потенциально невысокой сложностью декодирования) оказывается весьма близкой к пропускной способности канала — разница составляет примерно 2,4 % (в случае  $N = 100$ ,  $T = 1$  мес.) и примерно 3,3 % (в случае  $N = 1000$ ,  $T = 12$  мес.).

Приведенные результаты получены, как отмечалось, для модели канала с непрерывным выходом. На практике сигнал, считываемый из ячейки flash-памяти, подвергается равномерному квантованию (uniform sensing quantization) на 16–32 уровня. Предложенный в данной работе подход может быть непосредственно распространен и на модель с квантованным выходом. Представляется весьма вероятным, что введение 16-уровневого и тем более 32-уровневого квантования не приведет к сколько-нибудь существенному изменению приведенных выше основных выводов. В пользу этого говорит и тот факт, что введение квантования для конкретных кодовых схем (см., например, [6, 7]) не приводит к заметному снижению эффективности кодирования.

Работа выполнена при поддержке Минобрнауки РФ при выполнении научно-исследовательской работы в рамках проектной части государственного задания в сфере научной деятельности по заданию № 2.2716.2014/К от 17.07.2014.

### Литература

1. Трофимов А. Н., Таубин Ф. А. Теоретико-информационный анализ многоуровневой flash-памяти. Ч. 1: Модель канала и границы случайного кодирования // Информационно-управляющие системы. 2016. № 2. С. 56–67. doi:10.15217/issn1684-8853.2016.2.56

2. Sun F., Devarajan S., Rose K., Zhang T. Design of On-Chip Error Correction Systems for Multilevel NOR and NAND Flash Memories // IET Circuits Devices and Systems. 2007. Vol. 1. N 3. P. 241–249.

3. Sun F., Rose K., Zhang T. On the Use of Strong BCH Codes for Improving Multilevel NAND Flash Memo-

- ry Storage Capacity. [http://www.researchgate.net/publication/254376882\\_On\\_the\\_Use\\_of\\_Strong\\_BCH\\_Codes\\_for\\_Improving\\_Multilevel\\_NAND\\_Flash\\_Memory\\_Storage\\_Capacity](http://www.researchgate.net/publication/254376882_On_the_Use_of_Strong_BCH_Codes_for_Improving_Multilevel_NAND_Flash_Memory_Storage_Capacity) (дата обращения: 26.10.2014).
4. Kurkoski B. M. The E8 Lattice and Error Correction in Multi-Level Flash Memory // Proc. of the IEEE Intern. Conf. on Communications, 5 June–9 June 2011, Kyoto, Japan. 2011. P. 1–5.
  5. Dong G., Xie N., Zhang T. On the Use of Soft-Decision Error-Correction Codes in NAND Flash Memory // IEEE Transactions on Circuits and Systems I: Regular Papers. 2011. Vol. 58. N 2. P. 429–439.
  6. Li S., Zhang T. Improving Multi-Level NAND Flash Memory Storage Reliability Using Concatenated BCH-TCM Coding // IEEE Transactions on VLSI Systems. 2010. Vol. 18. N 10. P. 1412–1420.
  7. Xu Q., Gong P., Chen T. M. Concatenated LDPC-TCM Coding for Reliable Storage in Multi-level Flash Memories // Proc. of the 9<sup>th</sup> Intern. Symp. on Communication Systems, Networks & Digital Signal Processing (CSNDSP 2014). July 2014. P. 166–170.
  8. Kurkoski B. M. Coded Modulation Using Lattices and Reed-Solomon Codes, with Applications to Flash Memories // IEEE Transactions on Selected Areas in Communications. 2014. Vol. 32. N 5. P. 900–908.
  9. Huang X., et al. Multilevel Flash Memories: Channel modeling, Capacities and Optimal Coding Rates // X. Huang, A. Kavcic, X. Ma, G. Dong, T. Zhang // Intern. Journal on Advances in Systems and Measurement. 2013. Vol. 6. N 3&4. P. 364–373. [http://www.iariajournals.org/systems\\_and\\_measurements/sysmea\\_v6\\_n34\\_2013\\_paged.pdf](http://www.iariajournals.org/systems_and_measurements/sysmea_v6_n34_2013_paged.pdf) (дата обращения: 26.10.2014).
  10. Chase D. A Class of Algorithms for Decoding Block Codes with Channel Measurement Information // IEEE Transactions on Information Theory. 1972. Vol. 18. N 1. P. 170–182.
  11. Weldon E. Decoding Binary Block Codes on Q-ary Output Channel // IEEE Transactions on Information Theory. 1971. Vol. 17. N 6. P. 713–718.
  12. Gallager R. G. Information Theory and Reliable Communication. — N. Y.: John Wiley & Sons, 1968. — 588 p.
  13. Clark G., Cain J. Error-Correction Coding for Digital Communications. — N. Y.: Plenum Press, 1981. — 422 p.
  14. Van Lint J. H. Introduction to Coding Theory (Graduate Texts in Mathematics, vol. 86). 3rd edition. — Berlin: Springer, 1999. — 227 p.

UDC 621.391

doi:10.15217/issn1684-8853.2016.3.64

### Information Theory Analysis for Multilevel Flash Memory. Part 2: Evaluation of Code Parameters for Given Data Storage Reliability

Trofimov A. N.<sup>a</sup>, PhD, Tech., Associate Professor, andrei.trofimov@vu.spb.ruTaubin F. A.<sup>a</sup>, Dr. Sc., Tech., Professor, ftaubin@yahoo.com<sup>a</sup>Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., Saint-Petersburg, 190000, Russian Federation

**Introduction:** Multilevel flash memory due to its higher recording density has a dominant position at the market of non-volatile storage devices, being used extensively in a very wide range of applications. Increasing the recording density achieved by decreasing the physical size of a cell, and the increasing number of states lead to a decrease in the reliability of the data storage and to the need of error-correcting coding. **Purpose:** The goal is to study the capabilities of error-correcting coding for a model of multilevel NAND flash memory depending on the conditions of the recording and on a given level of the data storage reliability. **Results:** Explicit relations have been obtained connecting the main parameters of the hypothetical codes which provide a given level of data storage reliability for a flash memory model specifying a memory page as a system with independent multilevel cells. Using the obtained dependencies, trade-off relationships are set up for a hypothetical code between its rate, length and minimum distance. It is shown that the function describing the dependence of the minimum distance on the code rate has a minimum which corresponds to a code with the smallest distance and, hence, with a potentially low decoding complexity. Rates of such codes for typical cases are very close to the channel capacity (the difference is about 2-3%). **Practical relevance:** The obtained results allow us to estimate how close the parameters of a coding scheme for flash memory are to their limit values. It is shown that for the typical values of flash memory data-storage channel parameters, the minimum distance of a hypothetical code with a potentially low decoding complexity is 4 or 5, and the length of the respective codes are in the range from 600 to 900. A larger number of the rewriting cycles and/or longer duration of data storage with the same required reliability lead to a shorter hypothetical code with a potentially low decoding complexity.

**Keywords** — Multilevel NAND Flash Memory, Recording Channel, Random Coding Bound, Varshamov–Gilbert Bound.

### References

1. Trofimov A. N., Taubin F. A. Information Theory Analysis of Multilevel Flash Memory. Part 1. Channel Model and Random Coding Bounds. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2016, no. 2, pp. 56–67 (In Russian). doi:10.15217/issn1684-8853.2016.2.56
2. Sun F., Devarajan S., Rose K., Zhang T. Design of On-chip Error Correction Systems for Multilevel NOR and NAND Flash Memories. *IET Circuits Devices and Systems*, 2007, vol. 1, no. 3, pp. 241–249.
3. Sun F., Rose K., Zhang T. On the Use of Strong BCH Codes for Improving Multilevel NAND Flash Memory Storage Capacity. Available at: [http://www.researchgate.net/publication/254376882\\_On\\_the\\_Use\\_of\\_Strong\\_BCH\\_Codes\\_for\\_Improving\\_Multilevel\\_NAND\\_Flash\\_Memory\\_Storage\\_Capacity](http://www.researchgate.net/publication/254376882_On_the_Use_of_Strong_BCH_Codes_for_Improving_Multilevel_NAND_Flash_Memory_Storage_Capacity)

- Memory\_Storage\_Capacity (accessed 26 October 2014).
4. Kurkoski B. M. The E8 Lattice and Error Correction in Multi-Level Flash Memory. *Proc. of the IEEE Intern. Conf. on Communications*, 5 June–9 June 2011, Kyoto, Japan, 2011, pp. 1–5.
  5. Dong G., Xie N., Zhang T. On the Use of Soft-Decision Error-Correction Codes in NAND Flash Memory. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2011, vol. 58, no. 2, pp. 429–439.
  6. Li S., Zhang T. Improving Multi-Level NAND Flash Memory Storage Reliability Using Concatenated BCH-TCM Coding. *IEEE Transactions on VLSI Systems*, 2010, vol. 18, no. 10, pp. 1412–1420.
  7. Xu Q., Gong P., Chen T. M. Concatenated LDPC-TCM Coding for Reliable Storage in Multi-Level Flash Memories. *Proc. of the 9<sup>th</sup> Intern. Symp. on Communication Systems, Networks & Digital Signal Processing (CSNDSP 2014)*, July 2014, pp. 166–170.
  8. Kurkoski B. M. Coded Modulation Using Lattices and Reed-Solomon Codes, with Applications to Flash Memories. *IEEE Transactions on Selected Areas in Communications*, 2014, vol. 32, no. 5, pp. 900–908.
  9. Huang X., Kavcic A., Ma X., Dong G., Zhang T. Multilevel Flash Memories: Channel Modeling, Capacities and Optimal Coding Rates. *International Journal on Advances in Systems and Measurement*, 2013, vol. 6, no. 3&4, pp. 364–373. Available at: [http://www.ariajournals.org/systems\\_and\\_measurements/sysmea\\_v6\\_n34\\_2013\\_paged.pdf](http://www.ariajournals.org/systems_and_measurements/sysmea_v6_n34_2013_paged.pdf) (accessed 26 October 2014).
  10. Chase D. A Class of Algorithms for Decoding Block Codes with Channel Measurement Information. *IEEE Transactions on Information Theory*, 1972, vol. 18, no. 1, pp. 170–182.
  11. Weldon E. Decoding Binary Block Codes on Q-ary Output Channel. *IEEE Transactions on Information Theory*, 1971, vol. 17, no. 6, pp. 713–718.
  12. Gallager R. G. *Information Theory and Reliable Communication*. New York, John Wiley & Sons, 1968. 588 p.
  13. Clark G., Cain J. *Error-Correction Coding for Digital Communications*. New York, Plenum Press, 1981. 422 p.
  14. Lint J. H. van. *Introduction to Coding Theory (Graduate Texts in Mathematics, vol. 86)*. 3rd edition. Berlin, Springer, 1999. 227 p.

**Научный журнал  
«ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ»  
выходит каждые два месяца.**

Стоимость годовой подписки (6 номеров) для подписчиков России — 4800 рублей, для подписчиков стран СНГ — 5400 рублей, включая НДС 18%, таможенные и почтовые расходы.

Подписку на печатную версию журнала можно оформить в любом отделении связи по каталогу:

«Роспечать»: № 48060 — годовой индекс, № 15385 — полугодовой индекс,

а также через посредство подписных агентств:

«Северо-Западное агентство „Прессинформ“»

Санкт-Петербург, тел.: (812) 335-97-51, 337-23-05,

эл. почта: [press@crp.spb.ru](mailto:press@crp.spb.ru), [zajavka@crp.spb.ru](mailto:zajavka@crp.spb.ru),

сайт: <http://www.pinform.spb.ru>

«МК-Периодика» (РФ + 90 стран)

Москва, тел.: (495) 681-91-37, 681-87-47,

эл. почта: [export@periodicals.ru](mailto:export@periodicals.ru), сайт: <http://www.periodicals.ru>

«Информнаука» (РФ + ближнее и дальнее зарубежье)

Москва, тел.: (495) 787-38-73, эл. почта: [informnauka3@yandex.ru](mailto:informnauka3@yandex.ru),

сайт: <http://www.informnauka.com>

«Деловая пресса»

Москва, тел.: (495) 962-11-11, эл. почта: [podpiska@delpress.ru](mailto:podpiska@delpress.ru),

сайт: <http://delpress.ru/contacts.html>

«Коммерсант-Курьер»

Казань, тел.: (843) 291-09-99, 291-09-47, эл. почта: [kazan@komcur.ru](mailto:kazan@komcur.ru),

сайт: <http://www.komcur.ru/contacts/kazan/>

«Урал-Пресс» (филиалы в 40 городах РФ)

Сайт: <http://www.ural-press.ru>

«Идея» (Украина)

Сайт: <http://idea.com.ua>

«ВТЛ» (Узбекистан)

Сайт: <http://btl.sk.uz/ru/cat17.html> и др.

На электронную версию нашего журнала (все выпуски, годовая подписка, один выпуск, одна статья)

вы можете подписаться на сайтах НЭБ: <http://elibrary.ru>;

РУКОНТ: <http://www.rucont.ru>; ИВИС: <http://www.ivis.ru/>

Полнотекстовые версии журнала за 2002–2015 гг.

в свободном доступе на сайте журнала (<http://www.i-us.ru>),

НЭБ (<http://www.elibrary.ru>)

и Киберленинки (<http://cyberleninka.ru/>

[journal/n/informatsionno-upravlyayuschiesistemy](http://journal/n/informatsionno-upravlyayuschiesistemy)).