Multilevel 2-Cell t-Write Codes

Aman Bhatia, Aravind R. Iyengar, and Paul H. Siegel
University of California, San Diego, La Jolla, CA 92093 – 0401, USA
Email: {amanbh, aravind, psiegel}@ucsd.edu

***Abstract*—We consider t-write codes for write-once memories with cells that can store multiple levels. Using worst-case sum-rate optimal 2-cell t-write code constructions for the asymptotic case of continuous levels, we derive 2-cell t-write code constructions that give good sum-rates for cells that support q discrete levels. A general encoding scheme for q-level 2-cell t-write codes is provided.**

I. Introduction

Write-once memory (WOM) codes have been of interest in the field of data storage [1], [2] with applications to punch cards, optical storage and, more recently, flash memory. Flash memory stores information in the form of charge in a floating gate, referred to as a *cell*. While increasing the charge level of an individual cell is a simple operation with low latency, decreasing the level of a cell requires a complex erase operation on a large group of cells that also decreases the lifetime of the device. WOM codes have been proposed for flash memory as a way to write information multiple times before an erase operation is required, thereby enhancing the lifetime of the device [3], [4], [5].

While most WOM constructions proposed previously consider codes for n cells in a flash memory device where cells support $q = 2$ levels [3], [5], [6], recent works consider the construction of rewrite codes for multilevel cells that support $q \geq 3$ levels [7], [8], [9], [10]. The maximum sum-rate of rewrite codes for t writes on q-ary cells was shown to be $\log_2 \binom{q+t-1}{t}$ [11]. In [10], the author derived achievable rates for lattice-based, 2-write WOM codes over n cells in the asymptotic setting of continuous cell levels, where the cardinality of the message set is the same on each write. In this paper, we consider bounds and constructions for lattice-based t-write codes for two q-level cells. In Section II we formulate the code design problem and provide a general encoding scheme for use with our codes. In Section III we extend ideas presented by Kurkoski [10] and invoke the continuous approximation to obtain worst-case sum-rate optimal t-write regions for two cells supporting continuous levels. In Section IV we then derive a code construction for two cells with $q \geq 3$ levels that gives good sum-rates. Finally, we present our conclusions in Section V.

II. Partition-based Rewrite Codes

A. Problem Setup

Assume a memory device where each cell supports q levels such that the level of the cell can only be increased during the write operation. We are interested in storing information on these cells t times before they are erased. We propose to do this by using a *t-write code* that encodes messages for each write in 2 cells jointly. Naturally, the message stored in a certain write affects the number of messages that can be stored on subsequent writes. Let $M_{i,t}$ be the number of messages that can be stored in 2 cells in the worst case at the i^{th} write. The *instantaneous rate* for the i^{th} write of the t-write code is defined as

$$R_{i,t} = \frac{1}{2}\log_2 M_{i,t} \text{ bits per cell per write.} \tag{1}$$

The total number of messages stored in t writes is $S_t = \prod_{i=1}^{t} M_{i,t}$. The *worst-case sum-rate* for the t-write code is then defined as

$$R_t = \frac{1}{2}\log_2 S_t \text{ bits per cell per erase.} \tag{2}$$

Note that $R_t = \sum_{i=1}^{t} R_{i,t}$.

Without coding, we can store q^n messages in n cells in one write which must be followed by an erase operation. Thus, the instantaneous rate and the worst-case sum-rate without encoding is $\log_2 q$. Though the instantaneous rate for t-write codes is less than $\log_2 q$, the sum-rate can be made larger. Since the lifetime of the memory device depends on the number of erase operations, we will consider the sum-rate R_t (or equivalently, S_t) as the figure of merit for t-write codes.

Assume two cells that store levels $(x, y) \in \mathbb{L} = [q]^2$ where $[q] = \{1, 2, \cdots, q\}$. We will refer to a pair of cell levels as a *point*. To define the t-write code, we partition the set $\mathbb{L}$ into $\{\mathbb{L}_{i,t}\}_{i=1}^{t}$, such that $\mathbb{L}_{i,t}$ is the set of points that may be stored at the i^{th} write. Let $\mathbb{L}_{0,t} = \{(0,0)\}$ for ease of notation. Suppose the point stored at the $(i-1)^{\text{th}}$ write is (x, y); then the set of points that may be stored at the i^{th} write is

$$\begin{aligned}&\mathbb{L}_{i,t}(x,y)\\ &\quad\triangleq \mathbb{L}_{i,t} \cap \{x, x+1, \cdots, q\} \times \{y, y+1, \cdots, q\}\end{aligned}$$

for $1 \leq i \leq t$. Thus

$$M_{i,t} = \min_{(x,y)\in\mathbb{L}_{i-1,t}} |\mathbb{L}_{i,t}(x,y)|,\ 1 \leq i \leq t.$$

We will denote the set of messages that can be stored on the i^{th} write as $\mathbb{M}_{i,t} \triangleq [M_{i,t}]$. The t-write code is defined by the encoder-decoder pair (ϕ, ψ) where

$$\phi : \bigcup_{i=1}^{t} \left(\mathbb{M}_{i,t} \times \mathbb{L}_{i-1,t}\right) \mapsto \mathbb{L}, \quad \psi : \mathbb{L} \mapsto [t] \times \bigcup_{i=1}^{t} \mathbb{M}_{i,t}$$

such that

$$\phi(m, x, y) \in \mathbb{L}_{i,t}(x, y) \quad \forall\, m \in \mathbb{M}_{i,t}, (x, y) \in \mathbb{L}_{i-1,t}, \tag{3}$$
$$\psi(x, y) \in \{i\} \times \mathbb{M}_{i,t} \ \ \forall\ (x, y) \in \mathbb{L}_{i,t} \tag{4}$$

for $1 \leq i \leq t$. The condition in (3) implies that at the i^{th} write, the t-write code encodes a message by only increasing the cell levels and the condition in (4) implies that the decoder

978-1-4673-0223-4/12/$31.00 ©2012 IEEE

maps every point in the set $\mathbb{L}_{i,t}$ to a message in set $\mathbb{M}_{i,t}$. For a consistent encoder-decoder pair, we will further require

$$\psi\left(\phi(m,x,y)\right)=(i,m)\ \ \forall\ m\in\mathbb{M}_{i,t},(x,y)\in\mathbb{L}_{i-1,t} \quad (5)$$

for $1 \leq i \leq t$. Thus, a 2-cell t-write code $(\{\mathbb{L}_{i,t}\},\phi,\psi)$ that satisfies conditions (3), (4) and (5) achieves worst-case sum-rate $R_t = \frac{1}{2}\sum_{i=1}^{t}\log M_{i,t}$. In the next subsection, we show that given $\{\mathbb{L}_{i,t}\}_{i=1}^{t}$, there exists an encoder-decoder pair (ϕ,ψ) that satisfy the required conditions.

B. Encoding Scheme

We associate with the encoder-decoder pair, a $q\times q$ matrix Φ, such that $\Phi(x,y)$ denotes the message associated with any point $(x,y)\in[q]^2$. Then, the encoder and decoder are defined as in Algorithm 1. We propose Algorithm 2 to construct matrix Φ given a partition $\{\mathbb{L}_{i,t}\}$. We will present the proof of correctness of the algorithm in the longer version of the paper [12].

Algorithm 1 Encoder-decoder pair (ϕ,ψ) based on matrix Φ

```
  // Input is message to be written and current cell levels
  function φ(m, x, y)
Require: (m, x, y) ∈ ∪_{j=1}^t (M_{j,t} × L_{j−1,t})
      i ← Σ_{j=1}^t j · 1_{(x,y)∈L_{j−1,t}}
      Find (x′, y′) ∈ L_{i,t}(x, y) : Φ(x′, y′) = m
      return (x′, y′)
  end function

  // Input is current cell levels
  function ψ(x, y)
      i ← Σ_{j=1}^t j · 1_{(x,y)∈L_{j,t}}
      m ← Φ(x, y)
      return (i, m)
  end function
```

Now that we have shown the construction of an encoding-decoding scheme for any partition of $\mathbb{L}$, we turn to the question of finding partitions that maximize the sum-rate. We first answer this question in Section III for the case where the cells are allowed to store any level x from a continuum of levels, i.e. $x\in[0,\ell]\subset\mathbb{R}$. We will derive the optimal $\{\mathbb{L}_{i,t}\}_{i=1}^{t}$ that maximizes the sum-rate for this continuous cell levels case. Based on these results, we will then construct codes for the case where cells can store discrete levels $[q]$ in Section IV.

III. Optimization of Partitions - Continuous Approximation

A. Optimal 2-Writes

We first consider storing information using two cells that can together store all levels $(x,y)\in[0,\ell_1]\times[0,\ell_2]\triangleq\mathbb{L}$ where $x,y\in\mathbb{R}$. In this subsection, we consider the case where these cells are used to store information twice before being erased. We will assume that the levels of the two cells can only be increased from previously written levels.

Let the set of points that can be written on the two cells in the first write be $\mathbb{L}_1\subset\mathbb{L}$. For technical convenience, we will let $\mathbb{L}_1$ be a closed set. We will assume that if $(x,y)\in\mathbb{L}_1$, then $(\underline{x},y)$ and $(x,\underline{y})\in\mathbb{L}_1$ for every $\underline{x}\in[0,x],\underline{y}\in[0,y]$.

Algorithm 2 Algorithm to construct Φ given $\{\mathbb{L}_{i,t}\}_{i=1}^{t}$

```
  procedure INITIALIZEPHI ({L_{i,t}}_{i=1}^t)
    for all i ∈ {1, ··· , t} do
      for all (x, y) ∈ L_{i,t} do // Initialize Φ to 0
        Φ(x, y) ← 0
      end for
      // Start with point that achieves the min. volume
      (x̂, ŷ) ← arg min_{(x,y)∈L_{i−1,t}} |L_{i,t}(x, y)|
      // Assign messages to points accessible from (x̂, ŷ)
      ASSIGNPHI (L_{i,t}(x̂, ŷ), M_{i,t})
      (x′, y′) ← (x̂, ŷ)
      // Assign messages from other points in L_{i−1,t}
      while x′ > 1 ∧ y′ < q do
        (x″, y″) ← (x′ − 1, y) : y ≥ y′
        M_lost ← Φ(L_{i,t}(x′, y′)) \ Φ(L_{i,t}(x″, y″))
        ASSIGNPHI (L_{i,t}(x″, y″), M_lost)
        (x′, y′) ← (x″, y″)
      end while
      (x′, y′) ← (x̂, ŷ)
      while x′ < q ∧ y′ ≥ 1 do
        (x″, y″) ← (x′ + 1, y) : y ≤ y′
        M_lost ← Φ(L_{i,t}(x′, y′)) \ Φ(L_{i,t}(x″, y″))
        ASSIGNPHI (L_{i,t}(x″, y″), M_lost)
        (x′, y′) ← (x″, y″)
      end while
    end for
  end procedure

  // Procedure to assign messages M′ to points in set L′
  procedure ASSIGNPHI (L′, M′)
Require: |M′| ≤ |{(x, y) ∈ L′ : Φ(x, y) = 0}|
    for all m ∈ M′ do
      (x_unset, y_unset) ← (x, y) : (x, y) ∈ L′ ∧ Φ(x, y) = 0
      Φ(x_unset, y_unset) ← m
    end for
  end procedure
```

This ensures that all available points are utilized. Let $x_{\sup}\triangleq\sup\{x:\exists\, y\in[0,\ell_2]:(x,y)\in\mathbb{L}_1\}$. With this, we can let $y=\beta_1(x)$ denote the *boundary* of $\mathbb{L}_1$, i.e.,

$$\beta_1(x)=\max\{y\in[0,\ell_2]:(x,y)\in\mathbb{L}_1\}\quad\forall\, x\in[0,x_{\sup}]\,.$$

Equivalently, we can write $\mathbb{L}_1$ given the boundary β_1 as $\mathbb{L}_1(\beta_1)=\{(x,y):y\in[0,\beta_1(x)]\ \forall\, x\in[0,x_{\sup}]\}$. If a certain point $(x_1,y_1)\in\mathbb{L}_1$ is written in the first write, the set of possible levels that can be written in the second write, $\mathbb{L}_2(x_1,y_1)$, is

$$\mathbb{L}_2(x_1,y_1)=\{(x,y)\in\mathbb{L}:x\in[x_1,\ell_1],y\in[y_1,\ell_2]\}.$$

According to the *continuous approximation* principle for dense lattices [13], the volume of a set $\mathbb{S}$, denoted by $|\mathbb{S}|$, is a measure of the number of messages that can be stored in that set. With this notion, the number of messages that can be written in the first write is

$$M_1(\beta_1)\triangleq|\mathbb{L}_1(\beta_1)|=\int_0^{x_{\sup}}\beta_1(x)\,\mathrm{d}x,$$

and the number of messages that can be stored in the second write after writing (x_1, y_1) on the first write is

$$|\mathbb{L}_2(x_1, y_1)| = (\ell_1 - x_1)(\ell_2 - y_1).$$

In the worst case, we can therefore guarantee a total number of messages

$$S_2(\beta_1) \triangleq |\mathbb{L}_1(\beta_1)| \cdot \inf_{(x,y)\in\mathbb{L}_1(\beta_1)} |\mathbb{L}_2(x, y)| \triangleq M_1(\beta_1) \cdot M_2(\beta_1)$$

with two writes. Note that this quantity is a function of the boundary function β_1 alone. We refer to

$$\beta_1^* = \arg\max_{\beta_1\in\mathbb{W}(\mathbb{L})} S_2(\beta_1),$$

where $\mathbb{W}(\mathbb{L}) = \cup_{x_{\sup}\in[0,\ell_1]}\{f : [0, x_{\sup}] \mapsto [0, \ell_2]\}$, as the *optimal* boundary for the first write assuming that the cells are to be used twice before each erase. Then, we have the following.

Lemma 1 (Optimal boundary for 2 writes): The optimal boundary for the first write for a 2-write WOM code is

$$\beta_1^*(x) = \ell_2 - \frac{c}{\ell_1 - x} \quad \forall x \in [0, x_{\sup}]$$

where c is a constant independent of x.

Proof: Let β_1^* be the optimal boundary for the first write. Then

$$M_2(\beta_1^*) = \inf_{x\in[0,x_{\sup}]} (\ell_1 - x)(\ell_2 - \beta_1^*(x))$$

is the minimum volume achievable on the second write. Let

$$x'_{\sup} = \ell_1 - \frac{M_2(\beta_1^*)}{\ell_2}.$$

It can be shown that $x'_{\sup} \geq x_{\sup}$. Consider the set

$$\mathcal{E} \triangleq \{x \in [0, x_{\sup}] : (\ell_1 - x)(\ell_2 - \beta_1^*(x)) > M_2(\beta_1^*)\} \cup \left[x_{\sup}, x'_{\sup}\right].$$

Then, define β'_1 such that

$$\beta'_1(x) = \begin{cases} \ell_2 - M_2(\beta_1^*)(\ell_1 - x)^{-1}, & \forall x \in \mathcal{E}, \\ \beta_1^*(x), & \forall x \in [0, x_{\sup}] \setminus \mathcal{E}. \end{cases}$$

Clearly,

$$\begin{aligned} \beta'_1(x) &> \beta_1^*(x) \quad \forall x \in \mathcal{E} \\ \beta'_1(x) &\geq \beta_1^*(x) \quad \forall x \in [0, x_{\sup}] \\ \beta'_1(x) &\geq 0 \quad\quad \forall x \in \left[x_{\sup}, x'_{\sup}\right]. \end{aligned}$$

Then $M_1(\beta'_1) > M_1(\beta_1^*)$ and $M_2(\beta'_1) = M_2(\beta_1^*)$. Thus, using β'_1 as the boundary increases the volume achievable on the first write while the volume achievable on the second write in the worst-case is still $M_2(\beta_1^*)$. This contradicts the assumption that β_1^* is optimal, unless $\beta_1^* = \beta'_1$. ■

Remark 1: Note that the optimal boundary is actually a rectangular hyperbola with center (ℓ_1, ℓ_2), and $x = \ell_1$ and $y = \ell_2$ as the asymptotes. In [10], the author hypothesized that the optimal boundary for 2 writes is a hyperbola. We will shortly generalize the above result and show that the optimal boundaries for t writes are rectangular hyperbolas as well.

Theorem 2 (Optimal sum-rates for 2 writes): The optimal boundary for the first write for a 2-write WOM code is given by

$$\beta_1^*(x) = \ell_2 - \frac{\omega_2 \ell_1 \ell_2}{\ell_1 - x}$$

for every $x \in [0, \ell_1(1 - \omega_2)]$, where $\omega_2 = -\frac{1}{2}\left[W_{-1}\left(\frac{-1}{2\sqrt{e}}\right)\right]^{-1} \approx 0.284668$ and W_{-1} is the real branch of the Lambert W function satisfying $W(x) < -1$ [14]. The optimal number of messages on each write is

$$\begin{aligned} M_1^* &\triangleq M_1(\beta_1^*) = \frac{1}{2}(1 - \omega_2) \cdot \ell_1 \ell_2 \\ M_2^* &\triangleq M_2(\beta_1^*) = \omega_2 \cdot \ell_1 \ell_2, \end{aligned}$$

so that the total number of messages for two writes is

$$S_2^*(\ell_1, \ell_2) \triangleq S_2(\beta_1^*) = \frac{1}{2}\omega_2(1 - \omega_2) \cdot (\ell_1 \ell_2)^2.$$

Proof: From Lemma 1,

$$\beta_1^*(x) = \ell_2 - \frac{c}{\ell_1 - x} \quad \forall x \in [0, x_{\sup}]$$

where $x_{\sup} = \ell_1 - (c/\ell_2)$. The volumes achievable on the first and the second writes with boundary β_1^* are

$$M_1(\beta_1^*) = \int_0^{x_{\sup}} \beta_1^*(x)\,\mathrm{d}x = \ell_1\ell_2 - c - c\log\frac{\ell_1\ell_2}{c}$$

and $M_2(\beta_1^*) = c$. Therefore,

$$\begin{aligned} S_2(\beta_1^*) &= M_1(\beta_1^*) \cdot M_2(\beta_1^*) \\ &= \left(\ell_1\ell_2 - c - c\log\frac{\ell_1\ell_2}{c}\right) c = (\ell_1\ell_2)^2 f\left(\frac{c}{\ell_1\ell_2}\right), \end{aligned}$$

where $f(x) \triangleq x(1 - x + x\log x)$. Now, we choose $c = \hat{c}$ such that $S_2(\beta_1^*)$ is maximized; that is,

$$\hat{c} \triangleq \arg\max_{c\in[0,\ell_1\ell_2]} S_2(\beta_1^*) = \ell_1\ell_2 \cdot \arg\max_{x\in[0,1]} f(x) = \ell_1\ell_2 \cdot \omega_2.$$

■

B. Optimal t-Writes

For t writes using 2 cells with maximum levels ℓ_1 and ℓ_2, let the sets $\mathbb{L}_{i,t} \subset \mathbb{L}$ denote the sets of points that may be used on the i^{th} write, for $1 \leq i \leq t$. If a certain point $(x_i, y_i) \in \mathbb{L}_{i,t}$ is written in the i^{th} write, the set of points that can be written in the $(i+1)^{\text{th}}$ write, $\mathbb{L}_{i+1,t}(x_i, y_i)$, is

$$\mathbb{L}_{i+1,t}(x_i, y_i) = \{(x, y) \in \mathbb{L}_{i+1,t} : x \in [x_i, \ell_1], y \in [y_i, \ell_2]\}.$$

We will use the notation $\beta_{i,t}$ to denote the boundary for the i^{th} write, $M_{i,t}$ to denote the number of messages for the i^{th} write, and $S_t(\ell_1, \ell_2)$ to denote the total number of messages.

We now state and prove without details the main theorem for t-writes on 2 cells.

Theorem 3 (Optimal t-writes): The optimal boundary for the i^{th} write to store information t times in 2 cells such that the total number of messages is maximized is given by

$$\beta_{i,t}^*(x) = \ell_2 - \frac{\left(\prod_{j=t-i+1}^{t} \omega_j\right)\ell_1\ell_2}{\ell_1 - x}$$

for all $0 \leq x \leq \ell_1 \left(1 - \prod_{j=t-i+1}^{t} \omega_j\right)$, $1 \leq i \leq t-1$. The number of messages that can be stored on the i^{th} write, $1 \leq i \leq t$, is

$$M_{i,t}^* = \frac{1}{t-i+1} \cdot \left(\prod_{j=t-i+2}^{t} \omega_j \right) (1 - \omega_{t-i+1}) \cdot \ell_1 \ell_2,$$

and the total number of messages that can be stored in t writes is

$$S_t^*(\ell_1, \ell_2) = \frac{1}{t!} \left(\prod_{j=2}^{t} (\omega_j)^{j-1} (1 - \omega_j) \right) \cdot (\ell_1 \ell_2)^t ,$$

where $\omega_j = \frac{\tau_j}{W_{-1}(\tau_j e^{\tau_j})}$ and $\tau_j \triangleq -\frac{j-1}{j}$ for all $1 \leq j \leq t$.

Sketch of Proof: The proof proceeds by induction. From Theorem 2, the claim is true for $t = 2$ writes. Suppose the claim is true for k writes. Consider the case for $(k+1)$ writes. The total number of messages stored in the last k writes is upper bounded by $\inf_x S_k^*(\ell_1 - x, \ell_2 - \beta_{1,k+1}(x))$ which, by the induction hypothesis, is a function of $\inf_x (\ell_1 - x)(\ell_2 - \beta_{1,k+1}(x))$. If $\beta_{1,k+1}$ is a rectangular hyperbola, then this upper bound is achieved from any point $(x_1, \beta_{1,k+1}(x_1))$ stored on the first write because the boundaries for the subsequent k writes are hyperbolas independent of x_1 as shown in Fig. 1. Thus, $\beta_{1,k+1}^*$ is a hyperbola. To complete the proof, choose the hyperbola that maximizes the total number of messages in $k+1$ writes as $\beta_{1,k+1}^*$. ■

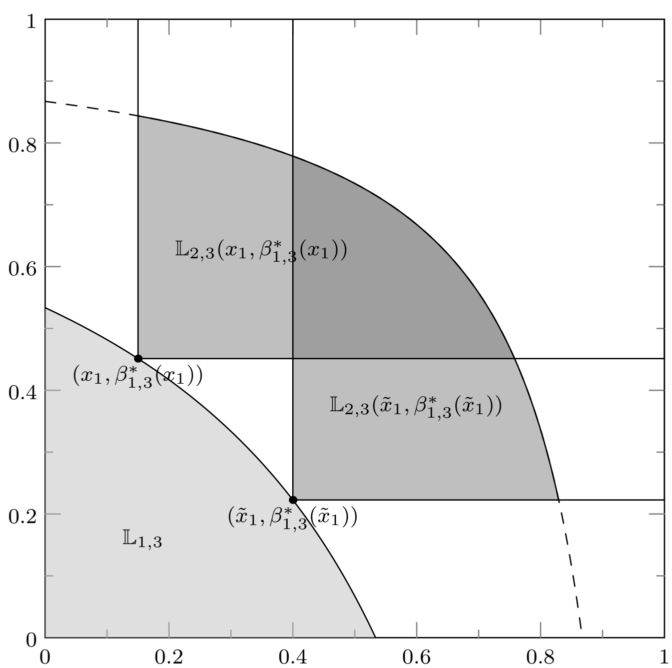

Fig. 1. Optimal boundaries for 3 writes over 2 cells with $\ell_1 = \ell_2 = 1$. We can see that the boundaries for the second write are independent of the points $(x_1, \beta^*_{1,3}(x_1))$ and $(\tilde{x}_1, \beta^*_{1,3}(\tilde{x}_1))$ written on the first write.

IV. Codes For Cells With Discrete Support

In this section we will discretize the results of Section III to construct t-write codes where the cells support q discrete

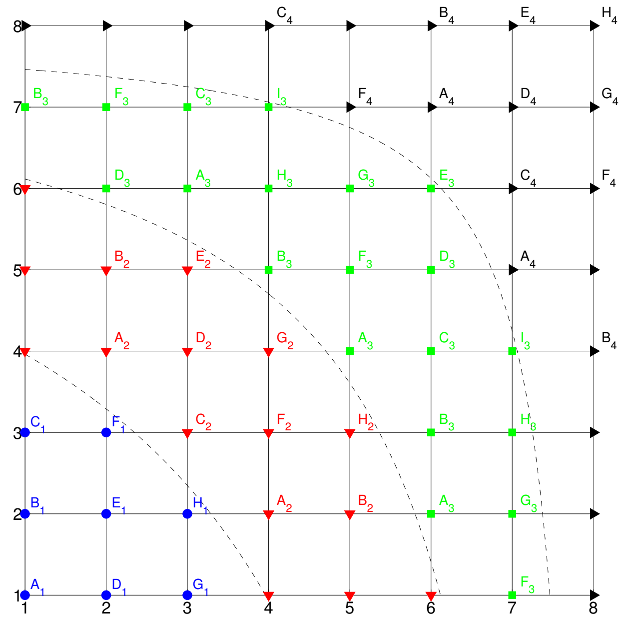

Fig. 2. 4-write code for 2 cells with 8 levels as described in Section IV. Points in the i^{th} partition are assigned messages $\{\mathrm{A}_i, \mathrm{B}_i, \cdots\}$ according to the encoder defined in Section II-B such that after any $i-1$ writes, $M_{i,t}$ messages may be stored on the next write.

levels. The discretization parameter will be denoted by $\Delta \triangleq \frac{\ell}{q-1}$ where continuous levels $[0, \ell]$ are discretized to q discrete levels. Without loss of generality, we will assume that $\ell = 1$ throughout this section. For 2 cells, we discretize $[0,1]^2$ into points indexed by ordered pairs in $[q]^2$. Then, the partition $\{\mathbb{L}_{i,t}\}_{i=1}^t$ that defines the 2-cell t-write code is given by

$$\begin{aligned} \mathbb{L}_{i,t} = \{(x,y) \in [q]^2 :& \quad\quad (6)\\ &\beta_{i-1,t}((x-1)\Delta) \leq (y-1)\Delta < \beta_{i,t}((x-1)\Delta) \\ &\forall x \in [q]\}, \qquad 1 \leq i \leq t-1 \end{aligned}$$

where $\beta_{i,t}$ is defined as in Theorem 3, $\beta_{0,t} \triangleq 0$, and $\mathbb{L}_{t,t} = \mathbb{L} \setminus \cup_{i=1}^{t-1} \mathbb{L}_{i,t}$. Given $\{\mathbb{L}_{i,t}\}_{i=1}^t$, the encoder-decoder for the t-write code are defined as in Section II-B. Then

$$M_{i,t} = \min_{(x,y) \in \mathbb{L}_{i-1,t}} |\mathbb{L}_{i,t}(x,y)|$$

and the instantaneous rate for the i^{th} write and sum-rate for the t-write code are given by (1) and (2), respectively. We shall use R_t^{D} to denote the sum-rate for the proposed t-write codes for cells supporting discrete levels.

As an example, we demonstrate in Fig. 2 how a 4-write code is constructed for 2 cells with 8 levels each. First, the optimal write boundaries in $[0,1]^2$ for continuous level cells, indicated in the figure with dashed lines, are obtained as described in Section III. The partition $\{\mathbb{L}_{i,4}\}_{i=1}^4$, as defined in (6) for the cells discretized with $\Delta = \frac{1}{q-1}$ for $q = 8$, is shown along with the messages assigned to each point by the encoding scheme defined in Section II-B. In the worst case, the 4-write code allows 8, 8, 9 and 8 messages to be written on the sequence of four writes. Thus,

$$R_t^{\mathsf{D}} = \frac{1}{2} \log_2(8 \cdot 8 \cdot 9 \cdot 8) = 6.085 \text{ bits per cell per erase.}$$

In Table I, we list the worst-case sum rate R_t^{D} achieved by the proposed codes for various numbers of writes, t, and cell levels, q. For each value of q, there is a particular number of writes t^* that achieves the best sum-rate. Note that $\log_2 q$ is the sum-rate achieved with no coding. The advantage of the proposed codes is apparent from the values in Table I.

TABLE I
WORST-CASE SUM-RATES R_t^{D} IN BITS PER CELL PER ERASE ACHIEVED BY t-WRITE CODES ON 2 CELLS WITH q LEVELS.

q	4	8	16	32
$\downarrow t \backslash \log_2 q \rightarrow$	2	3	4	5
2	2.70	4.55	6.44	8.40
3	2.95	5.48	8.25	11.13
4	2.59	6.09	9.71	13.46
5	2.09	6.55	10.90	15.40
6	1.79	6.61	11.80	17.21
7	–	6.70	12.54	18.72
8	–	6.42	13.15	20.22
9	–	6.38	13.73	21.43
10	–	5.88	14.19	22.57
t^*	3	7	14	29
$R_{t^*}^{\mathsf{D}}$	2.95	6.70	14.78	30.42

We can show that the rates R_t^{D} achieved by the proposed t-write code for cells with q levels satisfy

$$\frac{1}{2}\sum_{i=1}^{t}\log_2\left(1-\frac{\Delta^2}{M_{i,t}^*}\right) \leq R_t^{\mathsf{D}}-R_t^{\mathsf{C}} \leq \frac{1}{2}\sum_{i=1}^{t}\log_2\left(1+\frac{4\Delta}{M_{i,t}^*}\right) \tag{7}$$

$$\text{where } R_t^{\mathsf{C}} \triangleq \frac{1}{2}\cdot\log\left(\prod_{i=1}^{t}\frac{M_{i,t}^*}{\Delta^2}\right) = \frac{1}{2}\cdot\log S_t^*(1,1)+t\log\frac{1}{\Delta} \text{ bits per cell per erase}$$

is the sum-rate estimate obtained by the continuous approximation in the previous section. Thus, for small discretization parameters (or equivalently, large number of discrete levels), the sum-rate achieved by the proposed t-write code is well approximated by R_t^{C}. Table II lists upper and lower bounds on R_t^{D} obtained from (7), for various q and $t = 2, 3$. It is clear from the table that the approximation gets better as q increases.

TABLE II
R_t^{C}, BOUNDS ON R_t^{D} FROM (7), AND R_t^{D} ACHIEVED BY t-WRITE CODES

q	R_t^{C}	Lower Bound	Upper Bound	R_t^{D}
		$t = 2$		
4	1.52	0.90	3.90	2.70
8	3.97	3.87	5.45	4.55
16	6.17	6.15	7.04	6.44
32	8.26	8.26	8.75	8.40
64	10.31	10.31	10.57	10.37
128	12.32	12.32	12.47	12.36
256	14.34	14.34	14.41	14.36
		$t = 3$		
4	0.76	2.05	5.62	2.95
8	4.43	4.13	7.74	5.48
16	7.73	7.66	9.87	8.25
32	10.87	10.85	12.17	11.13
64	13.94	13.93	14.67	14.07
128	16.97	16.97	17.37	17.04
256	19.99	19.99	20.20	20.02

V. CONCLUSIONS

In this paper, we derived worst-case sum-rate optimal write regions for t writes on two cells for memories with cells that support continuous levels. These results were used to construct t-write codes for memories with multilevel cells that achieve high sum-rates over multiple writes per erase. We showed that the rates achieved by the proposed t-write codes for cells with a large number of discrete levels are well approximated by the rates estimated through the continuous approximation. We also gave a general encoding scheme for a 2-cell t-write code.

A number of generalizations of the work presented in this paper are possible. Often, in practice, symmetric-writes are more desirable, even though they achieve lower sum rates. In [12], we generalize the ideas in this paper to the case of symmetric writes, as well as to the case of $n > 2$ cells. However, finding an encoding scheme for codes over multiple cells is an open problem. It might also be useful to compute the average number of writes and rates achievable, instead of considering worst-case rates.

VI. ACKNOWLEDGMENTS

The authors would like to thank Brian Kurkoski for helpful discussions. This work was supported in part by NSF Grant CCF-1116739 and by the Center for Magnetic Recording Research at UCSD.

REFERENCES

[1] R. L. Rivest and A. Shamir, "How to reuse a write-once memory," *Information and Control*, vol. 55, nos. 1 –3, pp. 1 –19, Dec. 1982.

[2] J. K. Wolf, A. D. Wyner, J. Ziv, and J. Körner, "Coding for a write-once memory," *AT&T Bell Labs. Tech. J.*, vol. 63, no. 6, pp. 1089 –1112, 1984.

[3] A. Jiang, V. Bohossian, and J. Bruck, "Rewriting Codes for joint information storage in flash memories," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5300 –5313, Oct. 2010.

[4] S. Kayser, E. Yaakobi, P. H. Siegel, A. Vardy, and J. K. Wolf, "Multiple-write WOM-codes," in *Proc. 48th Ann. Allerton Conf. on Comm., Control and Computing*, Monticello, IL, Sep. 29 - Oct. 1, 2010.

[5] Y. Wu and A. Jiang, "Position modulation code for rewriting write-once memories," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3692 –3697, Jun. 2011.

[6] E. Yaakobi, S. Kayser, P. H. Siegel, A. Vardy, and J. K. Wolf, "Efficient two-write WOM-codes," in *IEEE Information Theory Workshop*, Dublin, Ireland, August 2010.

[7] Q. Huang, S. Lin, and K. A. S. Abdel-Ghaffar, "Error-correcting codes for flash coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6097 – 6108, Sep. 2011.

[8] R. Gabrys, E. Yaakobi, L. Dolecek, P. H. Siegel, A. Vardy, and J. K. Wolf, "Non-binary WOM-codes for multilevel flash memories," in *IEEE Information Theory Workshop*, Paraty, Brazil, pp. 40 – 44, Oct. 16 - 20, 2011

[9] Y. Cassuto and E. Yaakobi, "Practical re-write codes with access considerations," in *Non-Volatile Memories Workshop*, San Diego, CA, March 2012.

[10] B. M. Kurkoski, "Notes on a lattice-based WOM construction that guarantees two writes," in *Proc. 34th Symp. Information Theory and Applications*, Ousyuku, Iwate, Japan, pp. 520-524, Nov-Dec 2011.

[11] F. Fu, and A. J. Han Vinck, "On the capacity of generalized write-once memory with state transitions described by an arbitrary directed acyclic graph," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp 308 – 313, Jan. 1999.

[12] A. Bhatia, A. R. Iyengar, M. Qin, and P. H. Siegel, "Multilevel t-write codes," *In preparation*.

[13] G. D. Forney, R. G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. H. Qureshi, "Efficient modulation for band-limited channels," *IEEE J. on Selected Areas in Comm.*, JSAC-2, pp. 632 – 647, Sep. 1984.

[14] R. L. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. Knuth, "On the Lambert W function," *Adv. in Comp. Math.* 5(4):329–359, 1996.