Compressed Encoding for Rank Modulation

Eyal En Gad  
Electrical Engineering  
California Institute of Technology  
Pasadena, CA 91125, U.S.A.  
eengad@caltech.edu

Anxiao (Andrew) Jiang  
Computer Science and Engineering  
Texas A&M University  
College Station, TX 77843, U.S.A.  
ajiang@cse.tamu.edu

Jehoshua Bruck  
Electrical Engineering  
California Institute of Technology  
Pasadena, CA 91125, U.S.A.  
bruck@caltech.edu

Abstract—Rank modulation has been recently proposed as a scheme for storing information in flash memories. While rank modulation has advantages in improving write speed and endurance, the current encoding approach is based on the “push to the top” operation that is not efficient in the general case. We propose a new encoding procedure where a cell level is raised to be higher than the minimal necessary subset instead of all of the other cell levels. This new procedure leads to a significantly more compressed (lower charge levels) encoding. We derive an upper bound for a family of codes that utilize the proposed encoding procedure, and consider code constructions that achieve that bound for several special cases.

I. INTRODUCTION

With the recent application to flash memories, the rank-modulation scheme has gained renewed interest as evident in the recent series of papers [1]–[3], [5]–[7]. Basically, instead of a conventional multi-level cell in which a symbol in the input alphabet is represented by the charge level of one cell, in rank modulation, the stored information is represented by the permutation induced by the relative charge levels of several cells. The scheme, first described in [2] in the context of flash memory, works in conjunction with a simple cell-programming operation called “push-to-the-top”, which raises the charge level of a single cell above the rest of the cells. It was suggested in [2] that this scheme speeds up code programming by eliminating the over-programming problem. In addition, it also reduces corruption due to retention.

The cost of changing the state in the scheme – namely, the cost of the rewriting step – is measured by the number of “push-to-top” operations that are used, because it represents by how much the maximum cell level among the n cells has increased [2]. It is important to minimize this cell-level increment because the cells have a physical limit that upper bounds the cell levels. The less the cell levels are increased, the more rewrites can be performed before a block erasure operation becomes necessary, and the longer the lifetime of the memory will be.

We show an example in Fig. 1(a), where the state of n = 4 cells needs to change from \( u = [2, 1, 3, 4] \) to \( v = [2, 1, 4, 3] \). (Here the cells are indexed by \( 1, 2, \cdots, n \). And their state is denoted by the permutation \( [u_1, u_2, \cdots, u_n] \in S_n \), where cell \( u_1 \) has the highest charge level and \( u_n \) has the lowest charge level. For \( i = 1, \cdots, n \), cell \( u_i \) has rank \( i \).) Three “push-to-top” operations are needed, where cell 4, cell 1 and cell 2 are pushed sequentially. They are represented by the three edges in the figure. The cost of this rewriting is 3.

We can see from the above example, however, that the “push-to-top” operation is a conservative approach. To change the state from \( u = [2, 1, 3, 4] \) to \( v = [2, 1, 4, 3] \), when we push cell 4, the level of cell 4 in fact only needs to be greater than cell 3. There is no need to make it greater than the levels of all the other \( n - 1 = 3 \) cells (i.e., cells 1, 2 and 3). Similarly, when cell 1 is pushed, its level only needs to be greater than cell 3 and cell 4, instead of cells 2, 3 and 4. So a more moderate programming approach as shown in Fig. 1(b) can be taken, and the increment of the cell levels (in particular, the increment of the maximum cell level) can be substantially reduced. So the cost of rewriting can be reduced, which is important for the overall rewriting performance and the longevity of the memories.

In this work, we consider a programming approach that minimizes the increase of cell levels. To change the state from \( u = [u_1, u_2, \cdots, u_n] \in S_n \) to \( v = [v_1, v_2, \cdots, v_n] \in S_n \), we program cells based on their order in \( v \), so that every cell’s level increases as little as possible:

- For \( i = n - 1, n - 2, \cdots, 1 \) do:
  - \( \{ \text{Increase the level of cell } v_i, \text{ to make it greater than the level of the cell } v_{i+1} \} \).

Note that in the above programming process, when cell \( v_i \) is programmed, cell \( v_{i+1} \) already has the highest level among the cells \( v_i, v_{i+1}, v_{i+2}, \cdots, v_n \). We call the programming operation here the “minimal-push-up” operation. (In comparison, if we program cell \( v_i \) to make its level greater than the maximum level among the cells \( v_1, v_2, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n \), then it becomes the original “push-to-top” operation.) The “minimal-push-up” approach is robust, too, as it has no risk of overshooting. And it minimizes increment of the maximum level of the \( n \) cells (i.e., the rewrite cost).

In section II of the paper we study a discrete model of the rewrite process that allows us to define the rewrite cost of the “minimal-push-up” operation. We show that the cost equals the maximal rank increase among the elements of the initial permutation. In section III we turn to the problem of designing rewrite codes under a worst-case cost constraint. We...
limit ourselves to codes that assign a symbol to each one of the states, and show that such codes for a worst-case cost of 1 cannot achieve a storage capacity greater than \(1 - \frac{1}{k} \log_2(\frac{k}{2})\) bits per cell. We consider code constructions that achieve that bound for \(n \leq 5\), where the capacity is 54% higher than that of previous rank modulation codes. Finally, we show a way to generalize the construction for higher cost constraints.

II. Rewrite model and the transition graph

To design coding schemes, we need a good robust discrete model for the rewriting. We present here a discrete model for measuring the rewriting cost, which is suitable for both the “push-to-top” approach and the “minimal-push-up” approach. To rigorously define the cost of a rewriting step (i.e., a state transition), we will use the concept of virtual levels. Let \(u = \{u_1, u_2, \ldots, u_n\} \in S_n\) denote the current cell state, and let \(v = \{v_1, v_2, \ldots, v_n\} \in S_n\) denote the new state that the cells need to change into via increasing cell levels. Let \(d(u \rightarrow v)\) denote the number of push-up operations that are applied to the cells in order to change the state from \(u\) into \(v\). For \(i = 1, 2, \cdots, d(u \rightarrow v)\), let \(p_i \in [n] \triangleq \{1, 2, \cdots, n\}\) denote the integer and let \(B_i \subseteq [n] \setminus \{p_i\}\) denote the subset, such that the \(i\)-th push-up operation is to increase the \(p_i\)-th cell’s level to make it greater than the levels of all the cells in \(B_i\). (For example, for the rewriting in Fig. 1 (a), we have \(d(u \rightarrow v) = 3, p_1 = 4, B_1 = \{1, 2, 3\}, p_2 = 1, B_2 = \{2, 3, 4\}, p_3 = 2, B_3 = \{1, 3, 4\}\). For the rewriting in Fig. 1 (b), we have \(d(u \rightarrow v) = 3, p_1 = 4, B_1 = \{3\}, p_2 = 1, B_2 = \{3, 4\}, p_3 = 2, B_3 = \{1, 3, 4\}\).) Clearly, such push-up operations have no risk of overshooting.

For the current state \(u\), we assign the virtual levels \(n, n - 1, \cdot \cdot \cdot , 2, 1\) to the cells \(u_1, u_2, \ldots, u_{n-1}, u_n\), respectively. The greater a cell’s level is, the greater its virtual level is. It should be noted that when the virtual level increases by one, the increase in the actual cell level is not a constant because it depends on the actual programming process, which is noisy. However, it is reasonable to expect that when we program a cell \(a\) to make its level higher than a cell \(b\), the difference between the two cell levels will concentrate around an expected value. (For example, a one-shot programming using hot-electron injection can achieve quite stable programming performance at high writing speed.) Based on this, we formulate a simple yet robust discrete model for rewriting, which will be an important tool for designing coding schemes.

Consider the \(i\)-th push-up operation (for \(i = 1, \ldots, d(u \rightarrow v)\)), where we increase the level of cell \(p_i\) to make it greater than the levels of the cells in \(B_i\). For any \(j \in [n]\), let \(\ell_j\) denote cell \(j\)’s virtual level before this push-up operation. Then after the push-up operation, let the virtual level of cell \(p_i\) be

\[1 + \max_{j \in B_i} \ell_j;\]

namely, it is greater than the maximum virtual level of the cells in \(B_i\) by one. This increase represents the increment of the level of cell \(p_i\). After the \(d(u \rightarrow v)\) push-up operations that change the state from \(u\) to \(v\), for \(i = 1, \ldots, n\), let \(\ell_i’\) denote the virtual level of cell \(i\). We define the cost of the rewriting process as the increase in the maximum virtual level of the \(n\) cells, which is

\[
\max_{i \in [n]} \ell_i’ - n = \ell_{\ell_1}’ - n.
\]

Example 1. For the rewriting process shown in Fig. 1 (a), the virtual levels of cells 1, 2, 3, 4 change as \((3, 4, 2, 1) \rightarrow (3, 4, 2, 5) \rightarrow (6, 4, 2, 5) \rightarrow (6, 7, 2, 5)\). Its cost is 3.

For the rewriting process shown in Fig. 1 (b), the virtual levels of cells 1, 2, 3, 4 change as \((3, 4, 2, 1) \rightarrow (3, 4, 2, 3) \rightarrow (4, 4, 2, 3) \rightarrow (4, 5, 2, 3)\). Its cost is 1.

We would like to comment further on the robustness of the above discrete model. The model captures the typical behavior of cell programming. Yet when the minimal-push-up operations are used, the number of cells to push in practice is not always a constant when the old and new states \(u, v\) are given. An example is shown in Fig. 2 where the state needs to change from \(u = [1, 2, 3, 4]\) to \(v = [2, 1, 4, 3]\). The typical programming process is shown in Fig. 2 (a), where two cells – cell 4 and then cell 2 – are pushed up sequentially. (Note that based on the discrete model, the rewriting cost is 1. This is consistent with the increase of the maximum cell level here.) But as shown in Fig. 2 (b), in the rare case where cell 4’s level is significantly over-raised to the extent that it exceeds
previously, we can compute \( \ell \) of changing the state from \( u \) functions. Let \( i \) denote the virtual level of cell \( u \). Then we have

\[
C(u \rightarrow v) = \max_{i \in [n]} (v^{-1}(i) - u^{-1}(i)).
\]

**Theorem 1.**

It is simple to see that \( 0 \leq C(u \rightarrow v) \leq n - 1 \). An example of the rewriting cost is shown in Fig. 3.

![Figure 3. State diagram for the states of three cells, where every edge represents a state transition of cost 1. The dashed edges are transitions obtained by “minimal-push-up” operations but not “push-to-top” operations.](image)

In the following theorem we present an equivalent definition of the cost. According to the theorem, the cost is equal to the maximal increase in rank among the cells.

**Theorem 1.**

\[
C(u \rightarrow v) = \ell_{v(1)} - n.
\]

**Proof:** Assume by induction on \( k \) that

\[
\ell_{v(k)} = n + 1 - k + \max_{i \in [k,...,n]} (i - u^{-1}(v(i))).
\]

In the base case, \( k = n \), and \( \ell_{v(n)} = n + 1 - n + \max_{i \in [n,...,n]} (i - u^{-1}(v(i))) = 1 + n - u^{-1}(v(n)) \). We can see that this is in fact the result of the programming process. Now we assume that the expression is true for \( k \). For \( k - 1 \), by the programming process,

\[
\ell_{v(k-1)} = \max\{\ell_{v(k)} + 1, n + 1 - u^{-1}(v(k-1))\}
= \max\{n + 1 - k + \max_{i \in [k,...,n]} (i - u^{-1}(v(i))) + 1, n + 1 - u^{-1}(v(k-1))\}
\]

by the induction assumption

\[
= n + 1 - (k-1) + \max_{i \in [k-1,...,n]} (i - u^{-1}(v(i)))
\]

and the induction is proven.

Now we assign \( \ell_{v(1)} \) in the definition of the cost:

\[
C(u \rightarrow v) = \ell_{v(1)} - n
= n + 1 - 1 + \max_{i \in [1,...,n]} (i - u^{-1}(v(i))) - n
= \max_{i \in [n]} (v^{-1}(i) - u^{-1}(i))
\]


Codes for rewriting data based on the “push-to-top” operation were studied in \[2\]. Since the “minimal-push-up” approach has lower rewriting cost than the “push-to-top” operation, we can construct rewrite codes with higher rates.

In order to discuss rewriting, we first need to define a decoding scheme. It is often the case that the alphabet size used by the user to input data and read stored information differs from the alphabet size used as internal representation. In our case, data is stored internally in one of \(n!\) different permutations. Let us assume the user alphabet is \(Q = \{1, 2, \ldots, q\}\). A decoding scheme is a function \(D : S_n \rightarrow Q\) mapping internal states to symbols from the user alphabet. Suppose the current internal state is \(u \in S_n\) and the user inputs a new symbol \(a \in Q\). A rewriting operation given \(a\) is now defined as moving from state \(u \in S_n\) to state \(v \in S_n\) such that \(D(v) = a\). The cost of the rewriting operation is \(C(u \rightarrow v)\).

Next, we define a few terms. Define the transition graph \(G_n = (V_n, A_n)\) as a directed graph with \(V_n = S_n\), i.e., with \(n!\) vertices representing the permutations in \(S_n\). There is a directed edge \(u \rightarrow v\) if and only if \(C(u \rightarrow v) = 1\). Note that \(G_n\) is a regular digraph. Given a vertex \(u \in V_n\) and an integer \(r \in \{0, 1, \ldots, n-1\}\), we define the ball \(B_{n,r}(u)\) as \(B_{n,r}(u) = \{v \in V_n | C(u \rightarrow v) \leq r\}\).

**Theorem 2.**

\[
|B_{n,r}(u)| = r!(r+1)^n-r
\]

**Proof:** We use induction on \(n\). When \(n = 2\) the statement is trivial. (So it is when \(n = r+1\), where \(|B_{r+1,r}(u)| = (r+1)!\). Now we assume that the statement is true for \(n \leq n_0\), and consider \(n = n_0 + 1\) and \(n > r+1\). Let \(u = [u(1), u(2), \ldots, u(n)] \in S_n\), and without loss of generality (w.l.o.g.) let \(u(1) = n\). Let \(v = [v(1), v(2), \ldots, v(n)] \in B_{n,r}(u)\). Let \(u = [u(2), u(3), \ldots, u(n)] \in S_{n-1}\), and let \(v \in S_{n-1}\) be obtained from \(v\) by removing the element \(u(1) = n\). By Theorem 1, the first element in \(u\), namely \(u(1) = n\), can take one of the first \(r+1\) positions in \(v\). Given that position, there is a one-to-one mapping between pushing-up the remaining \(n-1\) elements from \(u\) to \(v\) and pushing-up those \(n-1\) elements from \(u\) to \(\hat{v} \in S_{n-1}\), and we have \(C(\hat{u} \rightarrow \hat{v}) = C(u \rightarrow v) \leq r\).

So we get \(|B_{n,r}(u)| = (r+1)!|B_{n-1,r}(u)|\) for \(r \geq 1\). Thus, \(|B_{n,r}(u)| = r!(r+1)^n-r\).

Note that given \(u\), \(|\{v \in S_n | v^{-1}(i) - u^{-1}(i) \leq r\} |\) for \(i \leq n\) is the size of the ball under infinity norm. When \(r = 1\), that size is known to be a Fibonacci number \(4\).

In addition, we note that \(|B_{n,1}(u)| = 2^{n-1}\). Therefore, the out-degree of each vertex in \(G_n\) is \(2^{n-1} - 1\). In comparison, when we allow only the “push-to-the-top” operation, \(|B_{n,1}(u)| = n\). Hence we get an exponential increase in the degree, which might lead to an exponential increase in the rate of rewrite codes. In the next section we study rewrite codes under a worst-case cost constraint.

### III. Worst-Case Decoding Scheme for Rewrite

In this section, we study codes where the cost of the rewrite operation is limited by \(r\).

**A. The case of \(n \leq 4\)**

We start with the case of \(r = 1\). The first non-trivial case for \(r = 1\) is \(n = 3\). However, for this case the additional “minimal-push-up” transitions do not allow for a better rewrite code. An optimal construction for a graph with only the “push-to-top” transitions was described in \[2\]. That construction assigns a symbol to each state according to the first element in the permutation, for a total of 3 symbols. It is easy to see that this construction is also optimal for a graph with the “minimal-push-up” transitions.

For greater values of \(n\), in order to simplify the construction, we limit ourselves to codes that assign a symbol to each of the \(n!\) states. We call such codes full assignment codes. Note that better codes for which not all the states are assigned to symbols might exist. When all of the states are assigned to symbols, each state must have an edge in \(A_n\) to at least one state labelled by each symbol. We define a set of vertices \(D = D_n\) as a dominating set if any vertex not in \(D\) is the initial vertex of an edge that ends in a vertex in \(D\). Every dominating set is assigned to one symbol. Our goal is to partition the set of \(n!\) vertices into the maximum number of dominating sets. We start by presenting a construction for \(n = 4\).

**Construction 1.** Divide the 24 states of \(S_4\) into 6 sets of 4 states each, where each set is a coset of \((\langle 1, 2, 3, 4 \rangle, \langle 1, 2, 3 \rangle)\), the cyclic group generated by \((1, 2, 3, 4)\). Map each set to a different symbol.

**Theorem 3.** Each set in Construction is a dominating set.

**Proof:** Let \(I_d\) be the identity permutation, \(g = (1, 2, 3, 4)\) and \(G = \langle g \rangle\). For each \(h \in S_4\), \(hG\) is a coset of \(G\). For each \(v = [v(1), \ldots, v(n)] \in hG\) and each \(u = [u(1), \ldots, u(n)] \in S_4\) such that \(u(1) = v(1)\), \(u\) has an edge to either \(v\) or \(v \ast g\). For example, in the coset \(I_d G = G\), for \(v = I_d\) and \(u \in S_4\) such that \(u(1) = v(1) = 1\), if \(u(2) = 2\) or \(3\), \(u\) has an edge to \(I_d = [1, 2, 3, 4]\), and if \(u(2) = 4\), \(u\) has an edge to \(I_d \ast g = [4, 1, 2, 3]\). Since \(G\) is a cyclic group of order 4, for every \(u \in S_4\) there exists \(v \in hG\) such that \(u(1) = v(1)\), and therefore \(hG\) is a dominating set.

For \(k \in [n]\) and \(B \subseteq S_n\), we define

\[
\operatorname{Pref}_k(B) = \{t | s = tu \text{ for } |u| = k \text{ and } s \in B\}
\]

where \(t,u\) are segments of the permutation \(s\). For example, \(\operatorname{Pref}_3([1, 2, 3, 4, 5, 1, 2, 3, 4, 5]) = \{(1, 2, 3), (1, 3, 2)\}\). We provide a lower bound to a dominating set’s size.

**Theorem 4.** If \(D\) is a dominating set of \(G_n\), then

\[
|D| \geq \frac{n!}{\phi (2^{n-1})}
\]

**Proof:** Each \(p_3 \in \operatorname{Pref}_3(S_n)\) is a prefix of 3 different prefixes in \(\operatorname{Pref}_2(S_n)\). For example, for \(n = 5\), \([1, 2]\) is a prefix of \([1, 2, 3], [1, 2, 4], [1, 2, 5]\). Each \(v \in D\) dominates \(2^{n-2}\) prefixes in \(\operatorname{Pref}_2(S_n)\). For example, for \(n = 4\), every
permutation that start with $[1, 2, 3, 4, 5]$, $g_1 = [1, 2, 3, 5, 4]$ dominates the prefixes $\{1, 2, 3, 5, 4\}$ as a representative of the domination over the 4 prefixes in $Pref_3(S_5)$ that are dominated by the coset. According to the construction, a set of representatives in $D$ that share the same parity is a left coset of $\langle g_2 \rangle$. Let one of the cosets of $\langle g_2 \rangle$ in $D$ be called C. For each $v \in C$, the subset $\{v, v * g_2\}$ represents a domination over the prefix $v(2)$. For example, for $v = I_4$, the subset $\{I_d = [1, 2, 3, 4, 5], I_d * g_2 = [2, 4, 5, 3, 1]\}$ represent a domination over the prefix $[2]$. Since $|\langle g_2 \rangle| = 5$, C represents a complete domination over $Pref_4(S_5)$, and therefore $D$ is a dominating set.

The rate of the code is
\[
R = \frac{1}{5} \log_2 12 = 0.717 \text{ bits per cell}
\]

C. The case of $r \geq 2$

When the cost constraint is greater than 1, we can generalize the constructions studied above. We present a construction for the case $r = n - 4$. The construction begins by dividing the $n!$ states $S_n$ into $\frac{n!}{120}$ sets, where two sets are in the same set if and only if their first $n - 5$ elements are the same. The sets are all dominating sets, because we can get to any set by at most $n - 5$ “push-to-top” operations. We further divide each of these sets to 12 sets of 10 members, in the same way as in Construction 2 according to the last 5 elements of the permutations. By the properties of construction 2, each of the smaller sets is still a dominating set. The rate of the code is
\[
R = \frac{1}{n} \log_2 \frac{120}{12} = 0.717 \text{ bits per cell}
\]

IV. Conclusion

We have presented a programming method that minimizes rewriting cost for rank modulation, and studied rewrite codes for a worst-case constraint on the cost. The presented codes are optimal full-assignment codes. It remains our future research to extend the code constructions to general code length, non-full assignment codes and average-case cost constraint.

REFERENCES


---

The parity (oddness or evenness) of a permutation $u$ can be defined as the parity of the number of inversions for $u$.  

2The parity (oddness or evenness) of a permutation $u$ can be defined as the parity of the number of inversions for $u$. 

**Theorem 5.** Each set $D$ in Construction 2 is a dominating set.

**Proof:** Each right coset of $\langle g_1 \rangle$ dominates 4 prefixes in $Pref_3(S_5)$. For example, the coset $\langle g_1 \rangle = \{I_d = [1, 2, 3, 4, 5], g_1 = [1, 2, 3, 5, 4]\}$ dominates the prefixes $\{[1, 2, 3, 5, 4]\}$ as a representative of the domination over the 4 prefixes in $Pref_3(S_5)$ that are dominated by the coset. According to the construction, a set of representatives in $D$ that share the same parity is a left coset of $\langle g_2 \rangle$. Let one of the cosets of $\langle g_2 \rangle$ in $D$ be called C. For each $v \in C$, the subset $\{v, v * g_2\}$ represents a domination over the prefix $v(2)$. For example, for $v = I_4$, the subset $\{I_d = [1, 2, 3, 4, 5], I_d * g_2 = [2, 4, 5, 3, 1]\}$ represent a domination over the prefix $[2]$. Since $|\langle g_2 \rangle| = 5$, C represents a complete domination over $Pref_4(S_5)$, and therefore $D$ is a dominating set.

The rate of the code is
\[
R = \frac{1}{5} \log_2 12 = 0.717 \text{ bits per cell}
\]

C. The case of $r \geq 2$

When the cost constraint is greater than 1, we can generalize the constructions studied above. We present a construction for the case $r = n - 4$. The construction begins by dividing the $n!$ states $S_n$ into $\frac{n!}{120}$ sets, where two sets are in the same set if and only if their first $n - 5$ elements are the same. The sets are all dominating sets, because we can get to any set by at most $n - 5$ “push-to-top” operations. We further divide each of these sets to 12 sets of 10 members, in the same way as in Construction 2 according to the last 5 elements of the permutations. By the properties of construction 2, each of the smaller sets is still a dominating set. The rate of the code is
\[
R = \frac{1}{n} \log_2 \frac{120}{12} = 0.717 \text{ bits per cell}
\]

IV. Conclusion

We have presented a programming method that minimizes rewriting cost for rank modulation, and studied rewrite codes for a worst-case constraint on the cost. The presented codes are optimal full-assignment codes. It remains our future research to extend the code constructions to general code length, non-full assignment codes and average-case cost constraint.

REFERENCES