SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2018, Vol. 61 100307:1–100307:14
https://doi.org/10.1007/s11432-018-9517-9

Special Focus on Distributed Storage Coding

New results on multilevel diversity coding with secure regeneration†

Shuo SHAO[1], Tie LIU[2], Chao TIAN[2] & Cong SHEN[3]

[1]*Department of Electrical Engineering, Shanghai Jiaotong University, Shanghai* 200240*, China;*
[2]*Department of Electrical Engineering, Texas A&M University, College Station* 77841*, USA;*
[3]*School of Information Science and Technology, University of Science & Technology of China, Hefei* 230022*, China*

Received 9 March 2018/Revised 27 June 2018/Accepted 13 July 2018/Published online 5 September 2018

Abstract The problem of multilevel diversity coding with secure regeneration is revisited. Under the assumption that the eavesdropper can access the repair data for all compromised storage nodes, Shao et al. provided a precise characterization of the minimum-bandwidth-regeneration (MBR) point of the achievable normalized storage-capacity repair-bandwidth tradeoff region. In this paper, it is shown that the MBR point of the achievable normalized storage-capacity repair-bandwidth tradeoff region remains the same even if we assume that the eavesdropper can access the repair data for some compromised storage nodes (defined as the type II compromised nodes) but only the data contents of the remaining compromised nodes (defined as the type I compromised nodes), as long as the number of type I compromised nodes is no greater than that of type II compromised nodes.

Keywords distributed storage, regenerating code, exact repair, secrecy constraint, MBR point

Citation Shao S, Liu T, Tian C, et al. New results on multilevel diversity coding with secure regeneration. Sci China Inf Sci, 2018, 61(10): 100307, https://doi.org/10.1007/s11432-018-9517-9

1 Introduction

Diversity coding, node repair, and security are three basic ingredients of modern distributed storage systems. The interplay of all three ingredients is captured by a fairly general mathematical model known multilevel diversity coding with secure regeneration (MDC-SR) [1].

More specifically, in an (n, d, ℓ) MDC-SR problem, a total of $d - \ell$ independent files $M_{\ell+1}, \ldots, M_d$ of size $B_{\ell+1}, \ldots, B_d$, respectively, are to be encoded and stored in n distributed storage nodes, each of capacity α. The encoding needs to ensure that:

• (Diversity coding) the file M_j can be perfectly recovered by having full access to any j out of the total n storage nodes for any $j \in \{\ell + 1, \ldots, d\}$;

• (Node repair) when node failures occur and there are d remaining nodes in the system, any failed node can be recovered by downloading data of size β from each one of the remaining nodes;

• (Security) the files $M_{\ell+1}, \ldots, M_d$ needs to be kept information-theoretically secure against an eavesdropper, which can access the repair data for ℓ compromised storage nodes.

* Corresponding author (email: shuoshao@sjtu.edu.cn)
† Invited article

Setting $\ell = 0$, the above problem reduces to the problem of multilevel diversity coding with regeneration (MDC-R) considered in [2, 3]. Setting $B_j = 0$ for all $j \neq k$, the above problem reduces to the (n, k, d, ℓ) secure regenerating code (SRC) problem considered in [4–11]. The goal is to understand the optimal tradeoffs between the storage capacity and repair bandwidth in satisfying all three aforementioned requirements.

From the code construction perspective, it is natural to consider the so-called separate coding scheme, i.e., to construct a code for the (n, d, ℓ) MDC-SR problem, we can simply use an (n, j, d, ℓ) SRC to encode the file M_j for each $j \in \{\ell + 1, \ldots, d\}$, and hence, the coded messages for each file will remain separate when stored in the storage nodes and during the repair processes. However, despite being a natural scheme, it was shown in [2] that separate coding is in general suboptimal in achieving the optimal tradeoffs between the normalized storage-capacity and repair-bandwidth for the MDC-R problem (which is a special case of the MDC-SR problem as mentioned previously). On the other hand, it has been shown [1] that separate coding can, in fact, achieve the minimum-bandwidth-regenerating (MBR) point of the achievable normalized storage-capacity and repair-bandwidth tradeoff region for the general MDC-SR problem. Nevertheless, the optimal tradeoffs between the storage capacity α and download bandwidth β, and, the performance of the minimum-storage-regenerating (MSR) point are still not fully understood. Especially for the MSR point, a code was given in [6] for SRC problem by extending the known MSR code without any security constraint. This coding scheme can achieve the MSR point when $d \geqslant 2k - 2$ and the eavesdropper can only observe type I compromised nodes (the definition of type I compromised node will be defined in the following part). However, it is still unknown as to whether this code is optimal for the more general eavesdropper model in our paper.

In this paper, we shall revisit the MDC-SR problem with a more general eavesdropping model. More specifically, instead of assuming that the eavesdropper can access the repair data for all compromised storage nodes, we shall assume that the compromised storage nodes can be divided into two different categories: type I compromised nodes and type II compromised nodes. While for the type II compromised nodes, we assume that the eavesdropper can access the repair data as before, for the type I compromised nodes we assume that the eavesdropper can only access the stored data contents.

Let ℓ_1 and ℓ_2 be the number of type I compromised nodes and type II compromised nodes respectively, and $\ell := \ell_1 + \ell_2$ be the total number of compromised nodes. By the node repair requirement, the data contents stored at any node can be fully recovered from its repair data. Therefore, for any fixed ℓ, the eavesdropper becomes weaker as ℓ_1 increases, which leads to a potentially larger achievable normalized storage-capacity and repair-bandwidth tradeoff region. A question of fundamental interest is to understand whether increasing ℓ_1 can lead to a strictly larger achievable normalized storage-capacity and repair-bandwidth tradeoff region. Our main result of the paper is to show that the MBR point of the achievable normalized storage-capacity and repair-bandwidth tradeoff region remains the same, as long as $\ell_1 \leqslant \ell/2$ (or equivalently, $\ell_1 \leqslant \ell_2$ by the fact that $\ell_2 = \ell - \ell_1$). From the technical viewpoint, this is mainly accomplished by establishing two outer bounds (one of them must be "horizontal", i.e., on the normalized repair-bandwidth only) on the achievable normalized storage-capacity and repair-bandwidth tradeoff region, which intersect precisely at the MBR point.

The rest of the paper is organized as follows. In Section 2 we formally introduce the problem of MDC-SR with the generalized eavesdropping model. The main results of the paper are then presented in Section 3. In Section 4, we introduce two "exchange" lemmas and use them to establish the main results of the paper. Finally, we conclude the paper in Section 5.

Notation. Sets and random variables will be written in calligraphic and sans-serif fonts respectively, to differentiate from the real numbers written in normal math fonts. For any two integers $t \leqslant t'$, we shall denote the set of consecutive integers $\{t, t + 1, \ldots, t'\}$ by $[t : t']$. The use of the brackets will be suppressed otherwise.

2 The generalized MDC-SR problem

In this paper, we study a distributed storage system that share the same file recovery and node repair function with [1]. Let $(n, d, N_1, \ldots, N_d, K, T, R)$ be a tuple of positive integers such that $d < n$. Formally, an $(n, d, N_1, \ldots, N_d, K, T, R)$ code consists of:

- for each $i \in [1 : n]$, a message-encoding function $f_i : (\prod_{j=1}^{d}[1 : N_j]) \times [1 : K] \to [1 : T]$;
- for each $\mathcal{A} \subseteq [1 : n] : |\mathcal{A}| \in [1 : d]$, a message-decoding function $g_{\mathcal{A}} : [1 : T]^{|\mathcal{A}|} \to [1 : N_{|\mathcal{A}|}]$;
- for each $\mathcal{B} \subseteq [1 : n] : |\mathcal{B}| = d$, $i' \in \mathcal{B}$, and $i \in [1 : n] \setminus \mathcal{B}$, a repair-encoding function $f^{\mathcal{B}}_{i' \to i} : [1 : T] \to [1 : R]$;
- for each $\mathcal{B} \subseteq [1 : n] : |\mathcal{B}| = d$ and $i \in [1 : n] \setminus \mathcal{B}$, a repair-decoding function $g^{\mathcal{B}}_i : [1 : R]^d \to [1 : T]$.

For each $j \in [1 : d]$, let M_j be a message that is uniformly distributed over $[1 : N_j]$. The messages $M_1, \ldots, M_d$ are assumed to be mutually independent. Let $\tilde{K}$ be a random key that is uniformly distributed over $[1 : K]$ and independent of the messages $(M_1, \ldots, M_d)$. For each $i \in [1 : n]$, let $W_i = f_i(M_1, \ldots, M_d, \tilde{K})$ be the data stored at the i-th storage node, and for each $\mathcal{B} \subseteq [1 : n] : |\mathcal{B}| = d$, $i' \in \mathcal{B}$, and $i \in [1 : n] \setminus \mathcal{B}$, let $S^{\mathcal{B}}_{i' \to i} = f^{\mathcal{B}}_{i' \to i}(W_{i'})$ be the data downloaded from the i'-th storage node in order to regenerate the data originally stored at the i-th storage node under the context of repair group $\mathcal{B}$. Obviously,

$$(B_j = \log N_j : j \in [1 : d]), \quad \alpha = \log T, \quad \text{and } \beta = \log S$$

represent the message sizes, storage capacity, and repair bandwidth, respectively.

The main deference between our definition in this paper and that in [1] is the model of eavesdropper. The eavesdropper now can observer a more complicated data combination consisted of both stored content and repair content. Let ℓ_1 and ℓ_2 be two nonnegative integers such that $\ell := \ell_1 + \ell_2 < d$. A normalized message-rate storage-capacity repair-bandwidth tuple $(\bar{B}_{\ell+1}, \ldots, \bar{B}_d, \bar{\alpha}, \bar{\beta})$ is said to be achievable for the (n, d, ℓ_1, ℓ_2) generalized MDC-SR problem if an $(n, d, 1, \ldots, 1, N_{\ell+1}, \ldots, N_d, K, T, R)$ code (i.e., $N_j = 1$ for all $j \in [1 : \ell]$) can be found such that:

- (rate normalization)

$$\frac{\alpha}{\sum_{t=\ell+1}^{d} B_t} = \bar{\alpha}, \quad \frac{\beta}{\sum_{t=\ell+1}^{d} B_t} = \bar{\beta}, \quad \frac{B_j}{\sum_{t=\ell+1}^{d} B_t} = \bar{B}_j \tag{1}$$

for any $j \in [\ell + 1 : d]$;

- (message recovery)

$$M_{|\mathcal{A}|} = g_{\mathcal{A}}(W_i : i \in \mathcal{A}) \tag{2}$$

for any $\mathcal{A} \subseteq [1 : n] : |\mathcal{A}| \in [\ell + 1 : d]$;

- (node regeneration)

$$W_i = g^{\mathcal{B}}_i(S^{\mathcal{B}}_{i' \to i} : i' \in \mathcal{B}) \tag{3}$$

for any $\mathcal{B} \subseteq [1 : n] : |\mathcal{B}| = d$ and $i \in [1 : n] \setminus \mathcal{B}$;

- (repair secrecy)

$$I((M_{\ell+1}, \ldots, M_d); (W_i : i \in \mathcal{E}_1), (S_{\to j} : j \in \mathcal{E}_2)) = 0 \tag{4}$$

for any $\mathcal{E}_1, \mathcal{E}_2 \subseteq [1 : n]$ such that $|\mathcal{E}_1| = \ell_1$, $|\mathcal{E}_2| = \ell_2$ and $\mathcal{E}_1 \cap \mathcal{E}_2 = \emptyset$ (so $\mathcal{E}_1$ and $\mathcal{E}_2$ represent the sets of types I and II compromised storage nodes, respectively), where $S_{\to i} := (S^{\mathcal{B}}_{i' \to i} : \mathcal{B} \subseteq [1 : n],\ |\mathcal{B}| = d,\ \mathcal{B} \not\ni i,\ i' \in \mathcal{B})$ is the collection of data that can be downloaded from the other nodes to regenerate node i.

The closure of all achievable $(\bar{B}_{\ell+1}, \ldots, \bar{B}_d, \bar{\alpha}, \bar{\beta})$ tuples is the achievable normalized message-rate storage-capacity repair-bandwidth tradeoff region $\mathcal{R}_{n,d,\ell_1,\ell_2}$ for the (n, d, ℓ_1, ℓ_2) generalized MDC-SR problem. For a fixed normalized message-rate tuple $(\bar{B}_{\ell+1}, \ldots, \bar{B}_d)$, the achievable normalized storage-capacity repair-bandwidth tradeoff region is the collection of all normalized storage-capacity repair-bandwidth pairs $(\bar{\alpha}, \bar{\beta})$ such that $(\bar{B}_{\ell+1}, \ldots, \bar{B}_d, \bar{\alpha}, \bar{\beta}) \in \mathcal{R}_{n,d,\ell_1\ell_2}$ and is denoted by $\mathcal{R}_{n,d,\ell_1,\ell_2}(\bar{B}_{\ell+1}, \ldots, \bar{B}_d)$.

Fixing ℓ and setting $\ell_1 = 0$, the (n, d, ℓ_1, ℓ_2) generalized MDC-SR problem reduces to the (n, d, ℓ) MDC-SR problem considered previously in [1], where it was shown that any achievable normalized message-rate storage-capacity repair-bandwidth tuple $(\bar{B}_{\ell+1}, \ldots, \bar{B}_d, \bar{\alpha}, \bar{\beta}) \in \mathcal{R}_{n,d,\ell}$ must satisfy

$$\bar{\beta} \geqslant \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} \bar{B}_j, \tag{5}$$

$$\bar{\alpha} + (d(d-\ell) - \ell)\bar{\beta} \geqslant (d-\ell)(d+1) \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} \bar{B}_j, \tag{6}$$

where $T_{d,k,\ell} := \sum_{t=\ell+1}^{k} (d+1-t)$. When set as equalities, the intersection of (5) and (6) is given by

$$\left(\bar{\alpha}, \bar{\beta}\right) = \left(d \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} \bar{B}_j, \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} \bar{B}_j \right), \tag{7}$$

which can be achieved by separate encoding with a previous scheme proposed by Shah et al. [6]. This provides a precise characterization of the MBR point for the (n, d, ℓ) MDC-SR problem.

3 Main results

The MBR point is the rate tuple $(\alpha_{\mathrm{MBR}}, \beta_{\mathrm{MBR}})$ such that

$$\beta_{\mathrm{MBR}} = \min_{(\alpha,\beta) \in \mathcal{R}_{n,d,\ell_1,\ell_2}} \beta, \quad \alpha_{\mathrm{MBR}} = \min_{(\alpha,\beta_{\mathrm{MBR}}) \in \mathcal{R}_{n,d,\ell_1,\ell_2}} \alpha.$$

Our main result of the paper is to show that the tradeoff point (7) remains to be the MBR point of $\mathcal{R}_{n,d,\ell_1,\ell_2}$ for the generalized MDC-SR problem as long as $\ell_1 \leqslant \ell_2$. The results are summarized in Theorem 1.

Theorem 1. For the generalized MDC-SR problem, any achievable normalized message-rate storage-capacity repair-bandwidth tuple $(\bar{B}_{\ell+1}, \ldots, \bar{B}_d, \bar{\alpha}, \bar{\beta}) \in \mathcal{R}_{n,d,\ell_1,\ell_2}$ must satisfy

$$\bar{\beta} \geqslant \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} \bar{B}_j, \tag{8}$$

and in addition, when $\ell_1 \leqslant \ell_2 = \ell - \ell_1$, we also have

$$\bar{\alpha} + T_{d,d,\ell_1+1}\bar{\beta} \geqslant (T_{d,d,\ell_1} + \ell_1) \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} \bar{B}_j. \tag{9}$$

When set as equalities, the intersection of (8) and (9) is precisely given by (7). We may thus conclude immediately that (7) is the MBR point of $\mathcal{R}_{n,d,\ell_1,\ell_2}$ for the generalized MDC-SR problem as long as $\ell_1 \leqslant \ell_2$.

Note that setting $\ell_1 = 0$, the outer bound (9) reduces to

$$\bar{\alpha} + \frac{d(d-1)}{2}\bar{\beta} \geqslant \frac{d(d+1)}{2} \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} \bar{B}_j. \tag{10}$$

So while the outer bound (8) coincides with (5), the outer bound (9) does not reduce to (6) when setting $\ell_1 = 0$. Simple calculations yield that the outer bound (10) is stronger than (6) if and only if $\ell \leqslant d/2$. In particular, when $\ell = 0$, the outer bound (10) reduces to that for the (n, d) MDC-R problem [3], while the outer bound (6) is strictly weaker. Figure 1 shows the comparison of (10) and (6) when $(\bar{B}_1, \bar{B}_2, \bar{B}_3) = (0, 1/3, 2/3)$ in $(4, 3, 0, 0)$ MDC-SR problem. In this figure, the outer bound (6) is below outer bound (10), though both of them intersect with (8) at the MBR point.

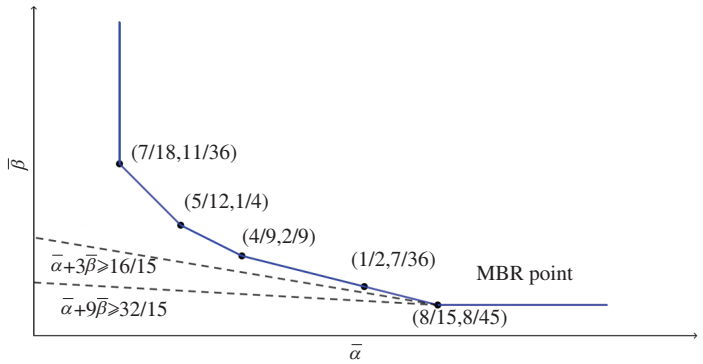

Figure 1 (Color online) The optimal tradeoff region for $(4,3,0,0)$ MDC-SR problem when $(\bar{B}_1,\bar{B}_2,\bar{B}_3)=(0,1/3,2/3)$. $(\bar{\alpha},\bar{\beta})$ are defined in (1). The outer bounds (8), (10) and (6) are evaluated as $\bar{\beta}\geqslant 8/45$, $\bar{\alpha}+3\bar{\beta}\geqslant 16/15$, and $\bar{\alpha}+9\bar{\beta}\geqslant 32/15$, respectively. When set as equalities, they intersect precisely at the MBR point $(8/15,8/45)$.

4 Proof of the main results

Let us first outline the main ingredients for proving the outer bounds (8) and (9).

(1) Total number of nodes. To prove the outer bounds (8) and (9), let us first note that these bounds are independent of the total number of storage nodes n in the system. Therefore, in our proof, we only need to consider the cases where $n=d+1$. For the cases where $n>d+1$, since any subsystem consisting of $d+1$ out of the total n storage nodes must give rise to a $(d+1,d,\ell)$ MDC-SR problem. Therefore, these outer bounds must apply as well. When $n=d+1$, any repair group $\mathcal{B}$ of size d is uniquely determined by the node j to be repaired, i.e., $\mathcal{B}=[1:n]\setminus\{j\}$, and hence can be dropped from the notation $S^{\mathcal{B}}_{i\to j}$ without causing any confusion.

(2) Code symmetry. Due to the built-in symmetry of the problem, to prove the outer bounds (8) and (9), we only need to consider the so-called symmetrical codes [12] for which the joint entropy of any subset of random variables from

$$((M_1,\ldots,M_d),\tilde{K},(W_i:i\in[1:n]),(S_{i\to j}:i,j\in[1:n],i\neq j))$$

remains unchanged under any permutation over the storage-node indices. For example, for a four nodes distributed storage system, we have

$$H(W_1,S_{3\to 2})=H(W_2,S_{4\to 3})$$

according to a permutation that every index shifts to the next index.

(3) Key collections of random variables. Focusing on the symmetrical $(n=d+1,d,N_1,\ldots,N_d,K,T,R)$ codes, the following collections of random variables play a key role in our proof:

$$\begin{aligned}
&M_{\mathcal{A}}:=(M_i:i\in\mathcal{A}),\quad \mathcal{A}\subseteq[1:d],\\
&M^{(m)}:=M_{[1:m]},\quad m\in[1:d],\\
&W_{\mathcal{A}}:=(W_i:i\in\mathcal{A}),\quad \mathcal{A}\subseteq[1:n],\\
&S_{i\to\mathcal{B}}:=(S_{i\to j}:j\in\mathcal{B}),\quad i\in[1:n],\ \mathcal{B}\subseteq[1:n]\setminus\{i\},\\
&S_{\mathcal{B}\to j}:=(S_{i\to j}:i\in\mathcal{B}),\quad j\in[1:n],\ \mathcal{B}\subseteq[1:n]\setminus\{j\},\\
&S_{\to j}:=S_{[1:j-1]\cup[j+1:n]\to j},\quad j\in[1:n],\\
&S_{\to\mathcal{B}}:=(S_{\to j}:j\in\mathcal{B}),\quad \mathcal{B}\subseteq[1:n],\\
&\underline{S}_{\to j}:=S_{[1:j-1]\to j},\quad j\in[1:n],\\
&\underline{S}_{\to\mathcal{B}}:=(\underline{S}_{\to j}:j\in\mathcal{B}),\quad \mathcal{B}\subseteq[1:n],\\
&\overline{S}_{\to j}:=S_{[j+1:n]\to j},\quad j\in[1:n],\\
&\overline{S}_{\to\mathcal{B}}:=(\overline{S}_{\to j}:j\in\mathcal{B}),\quad \mathcal{B}\subseteq[1:n],
\end{aligned}$$

$$U^{(t,s)} := (W_{[1:t]}, \overline{S}_{\to[t+1:s]}), \quad s \in [1:n],\ t \in [0:s],$$
$$U^{(s)} := U^{(0,s)}.$$

These collections of random variables have also been used in [3, 11].

An important part of the proof is to understand the relations between the collections of random variables defined above, and to use them to derive the desired converse results. We shall discuss this next.

4.1 Technical lemmas

Lemma 1. For any $(n = d+1, d, N_1, \ldots, N_d, K, T, R)$ code that satisfies the node regeneration requirement (3), $(\underline{S}_{\to[t+1:s]}, W_{[t+1:s]})$ is a function of $U^{(t,s)}$ for any $s \in [1:n]$ and $t \in [0:s-1]$.

The above lemma, which was first introduced in [1, 11], demonstrates the "compactness" of $U^{(t,s)}$ and has a number of direct consequences. For example, for any fixed $s \in [1:n]$, it is clear from Lemma 1 that $U^{(t_2,s)}$ is a function of $U^{(t_1,s)}$ and hence $H(U^{(t_2,s)}) \leqslant H(U^{(t_1,s)})$ for any $0 \leqslant t_1 \leqslant t_2 \leqslant s-1$.

Lemma 2 (Exchange Lemma 1 [1])**.** For any symmetrical $(n = d+1, d, N_1, \ldots, N_d, K, T, R)$ code that satisfies the node regeneration requirement (3), we have

$$\frac{d+1-j}{d-m} H(U^{(i,m)}|M^{(m)}) + H(U^{(i',j)}|M^{(m)}) \geqslant \frac{d+1-j}{d-m} H(U^{(i,m+1)}|M^{(m)}) + H(U^{(i',j-1)}|M^{(m)}) \quad (11)$$

for any $m \in [1:d-1]$, $i \in [0:m-1]$, $i' \in [0:i]$, and $j \in [i'+1 : m-i+i'+1]$.

Corollary 1. For any symmetrical $(n = d+1, d, N_1, \ldots, N_d, K, T, R)$ code that satisfies the node regeneration requirement (3), we have

$$\frac{T_{d,j_1+1,j_2}}{d-j_1} H(U^{(i,j_1)}|M^{(j_1)}) \geqslant \frac{T_{d,j_1,j_2}}{d-j_1} H(U^{(i,j_1+1)}|M^{(j_1)}) + H(U^{(i,j_2)}|M^{(j_1)}) \quad (12)$$

for any $j_1 \in [1:d-1]$, $i = [0:j_1]$ and $j_2 \in [i:j_1-1]$.

Proof. Set $m = j_1$ and $i' = i$ in (11). We have

$$\frac{d+1-j}{d-j_1} H(U^{(i,j_1)}|M^{(j_1)}) + H(U^{(i,j)}|M^{(j_1)}) \geqslant \frac{d+1-j}{d-j_1} H(U^{(i,j_1+1)}|M^{(j_1)}) + H(U^{(i,j-1)}|M^{(j_1)}) \quad (13)$$

for any $j \in [j_2+1 : j_1]$. Add the inequalities (13) for $j \in [j_2+1 : j_1]$ and cancel the common term $\sum_{j=j_2+1}^{j_1-1} H(U^{(i,j)}|M^{(j_1)})$ from both sides. We have

$$\begin{aligned}\frac{T_{d,j_1+1,j_2}}{d-j_1} H(U^{(i,j_1)}|M^{(j_1)}) &= \frac{T_{d,j_1,j_2}}{d-j_1} H(U^{(i,j_1)}|M^{(j_1)}) + H(U^{(i,j_1)}|M^{(j_1)}) \\ &\geqslant \frac{T_{d,j_1,j_2}}{d-j_1} H(U^{(i,j_1+1)}|M^{(j_1)}) + H(U^{(i,j_2)}|M^{(j_1)}).\end{aligned}$$

Corollary 2. For any symmetrical $(n = d+1, d, N_1, \ldots, N_d, K, T, R)$ code that satisfies the node regeneration requirement (3), we have

$$T_{d,m,\ell}^{-1} H(U^{(\ell_1,m)}|M^{(m)}) \geqslant T_{d,m+1,\ell}^{-1} H(U^{(\ell_1,m+1)}|M^{(m)}) + (T_{d,m,\ell}^{-1} - T_{d,m+1,\ell}^{-1}) H(U^{(\ell_1,\ell)}|M^{(m)}) \quad (14)$$

for any $\ell \in [0:d-1]$, $\ell_1 \in [0:\ell]$ and $m \in [\ell+1 : d-1]$.

Proof. Set $i = i' = \ell_1$, $j_1 = m$ and $j_2 = \ell$ in (12). We have

$$\frac{T_{d,m,\ell}}{d-m} H(U^{(\ell_1,m)}|M^{(m)}) + H(U^{(\ell_1,m)}|M^{(m)}) \geqslant \frac{T_{d,m,\ell}}{d-m} H(U^{(\ell_1,m+1)}|M^{(m)}) + H(U^{(\ell_1,\ell)}|M^{(m)}), \quad (15)$$

which can be equivalently written as

$$\frac{T_{d,m+1,\ell}}{d-m} H(U^{(\ell_1,m)}|M^{(m)}) \geqslant \frac{T_{d,m,\ell}}{d-m} H(U^{(\ell_1,m+1)}|M^{(m)}) + H(U^{(\ell_1,\ell)}|M^{(m)}) \quad (16)$$

by the fact that $T_{d,m,\ell}+(d-m)=T_{d,m+1,\ell}$. Multiplying both sides of (16) by

$$\frac{d-m}{T_{d,m+1,\ell}T_{d,m,\ell}}=T_{d,m,\ell}^{-1}-T_{d,m+1,\ell}^{-1}$$

completes the proof of (12).

Lemma 3 (Exchange Lemma 2)**.** For any symmetrical $(n=d+1,d,N_1,\ldots,N_d,K,T,R)$ code that satisfies the node regeneration requirement (3), we have

$$\frac{d-\ell_1}{d-\ell}H(U^{(\ell_1,\ell)})+H(U^{(\ell_1,\ell_1+1)},S_{\ell_1+1\to[1:\ell_1]})\geqslant\frac{d-\ell_1}{d-\ell}H(U^{(\ell_1,\ell+1)})+H(U^{(\ell_1,\ell_1)},S_{\ell_1+1\to[1:\ell_1]}) \quad (17)$$

for any $\ell\in[1:d-1]$ and $\ell_1\in[0:\lfloor\ell/2\rfloor]$.

Proof. See the Appendix A.

We note here that when setting $\ell_1=0$, Lemma 3 coincides with Lemma 2 with $i=i'=0$ and $j=1$.

4.2 The proof

Consider a symmetrical $(n=d+1,d,1,\ldots,1,N_{\ell+1},\ldots,N_d,K,T,R)$ regenerating code that satisfies the rate normalization requirement (1), the message recovery requirement (2), the node regeneration requirement (3), and the repair secrecy requirement (4). Let us first prove a few intermediate results. The outer bounds (8) and (9) will then follow immediately.

Proposition 1.

$$\frac{1}{d-\ell}H(U^{(\ell_1,\ell+1)})\geqslant\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+T_{d,m,\ell}^{-1}H(U^{(\ell_1,m)}|M_{[\ell+1:m]})+\left(\frac{1}{d-\ell}-T_{d,m,\ell}^{-1}\right)H(U^{(\ell_1,\ell)}) \quad (18)$$

for any $m\in[\ell+1:d]$. Consequently,

$$\frac{1}{d-\ell}H(U^{(\ell_1,\ell+1)})\geqslant\sum_{j=\ell+1}^{d}T_{d,j,\ell}^{-1}B_j+\frac{1}{d-\ell}H(U^{(\ell_1,\ell)}). \quad (19)$$

Proof. To see (18), consider proof by induction. For the base case with $m=\ell+1$, we have

$$\begin{aligned}\frac{1}{d-\ell}H(U^{(\ell_1,\ell+1)})&\overset{(a)}{=}\frac{1}{d-\ell}H(U^{(\ell_1,\ell+1)},M_{\ell+1})\\&\overset{(b)}{=}\frac{1}{d-\ell}(H(M_{\ell+1})+H(U^{(\ell_1,\ell+1)}|M_{\ell+1}))\\&\overset{(c)}{=}\frac{1}{d-\ell}(B_{\ell+1}+H(U^{(\ell_1,\ell+1)}|M_{\ell+1}))\\&\overset{(d)}{=}T_{d,\ell+1,\ell}^{-1}B_{\ell+1}+T_{d,\ell+1,\ell}^{-1}H(U^{(\ell_1,\ell+1)}|M_{\ell+1}),\end{aligned}$$

where (a) follows from the fact that $M_{\ell+1}$ is a function of $W_{[1:\ell+1]}$, which is a function of $U^{(\ell_1,\ell+1)}$ by Lemma 1; (b) follows from the chain rule for entropy; (c) follows from the fact that $H(M_{\ell+1})=B_{\ell+1}$; and (d) follows from the fact that $T_{d,\ell+1,\ell}=d-\ell$. Assuming that (18) holds for some $m\in[\ell+1:d-1]$, we have

$$\begin{aligned}&\frac{1}{d-\ell}H(U^{(\ell_1,\ell+1)})\\&\overset{(a)}{\geqslant}\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+T_{d,m,\ell}^{-1}H(U^{(\ell_1,m)}|M_{[\ell+1:m]})+\left(\frac{1}{d-\ell}-T_{d,m,\ell}^{-1}\right)H(U^{(\ell_1,\ell)})\\&\overset{(b)}{\geqslant}\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+T_{d,m+1,\ell}^{-1}H(U^{(\ell_1,m+1)}|M_{[\ell+1:m]})+\left(\frac{1}{d-\ell}-T_{d,m+1,\ell}^{-1}\right)H(U^{(\ell_1,\ell)})\end{aligned}$$

$$
\begin{aligned}
&\overset{(c)}{\geqslant} \sum_{j=\ell+1}^{m} T_{d,j,\ell}^{-1} B_j + T_{d,m+1,\ell}^{-1} H(U^{(\ell_1,m+1)}, M_{m+1}|M_{[\ell+1:m]}) + \left(\frac{1}{d-\ell} - T_{d,m+1,\ell}^{-1}\right) H(U^{(\ell_1,\ell)})\\
&\overset{(d)}{=} \sum_{j=\ell+1}^{m} T_{d,j,\ell}^{-1} B_j + T_{d,m+1,\ell}^{-1} H(M_{m+1}|M_{[\ell+1:m]}) + T_{d,m+1,\ell}^{-1} H(U^{(\ell_1,m+1)}|M_{[\ell+1:m+1]})\\
&\quad + \left(\frac{1}{d-\ell} - T_{d,m+1,\ell}^{-1}\right) H(U^{(\ell_1,\ell)})\\
&\overset{(e)}{=} \sum_{j=\ell+1}^{m} T_{d,j,\ell}^{-1} B_j + T_{d,m+1,\ell}^{-1} B_{m+1} + T_{d,m+1,\ell}^{-1} H(U^{(\ell_1,m+1)}|M_{[\ell+1:m+1]}) + \left(\frac{1}{d-\ell} - T_{d,m+1,\ell}^{-1}\right) H(U^{(\ell_1,\ell)})\\
&= \sum_{j=\ell+1}^{m+1} T_{d,j,\ell}^{-1} + T_{d,m+1,\ell}^{-1} H(U^{(\ell_1,m+1)}|M_{[\ell+1:m+1]}) + \left(\frac{1}{d-\ell} - T_{d,m+1,\ell}^{-1}\right) H(U^{(\ell_1,\ell)}),
\end{aligned}
$$

where (a) follows from the induction assumption; (b) follows from Corollary 2; (c) follows from the fact that M_{m+1} is a function of $W_{[1:m+1]}$, which is is a function of $U^{(\ell_1,m+1)}$ by Lemma 1; (d) follows from the chain rule for entropy; and (e) follows from the facts that M_{m+1} is independent of $M_{[\ell+1:m]}$ and that $H(M_{m+1}) = B_{m+1}$. This completes the induction step and hence the proof of (18).

To see (19), simply set $m = d$ in (18). We have

$$
\frac{1}{d-\ell} H(U^{(\ell_1,\ell+1)}) \geqslant \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} B_j + T_{d,d,\ell}^{-1} H(U^{(\ell_1,d)}|M_{[\ell+1:d]}) + \left(\frac{1}{d-\ell} - T_{d,d,\ell}^{-1}\right) H(U^{(\ell_1,\ell)}). \tag{20}
$$

Note that

$$
H(U^{(\ell_1,d)}|M_{[\ell+1:d]}) \geqslant H(U^{(\ell_1,\ell)}|M_{[\ell+1:d]}) = H(U^{(\ell_1,\ell)}), \tag{21}
$$

where the last equality follows from the fact that $I(U^{(\ell_1,\ell)}; M_{[\ell+1:d]}) = 0$ by the repair secrecy requirement (4). Substituting (21) into (20) completes the proof of (19).

Proposition 2.

$$
H(S_{\ell_1+1\to[1:\ell_1]}) + \frac{\ell_1}{d-\ell} H(U^{(\ell_1,\ell)}) \geqslant \frac{\ell_1}{d-\ell} H(U^{(\ell_1,\ell+1)}). \tag{22}
$$

Proof. First note that for any $m \in [1 : \ell_2 + 1]$ and $k \in [\ell + 1 : d + 1]$, we have

$$
\begin{aligned}
&H(S_{\ell_1+1\to[1:m]}) + H(U^{(\ell_1,\ell)}, S_{[\ell+2:k]\to\ell+1})\\
&\overset{(a)}{=} H(S_{k+1\to[\ell_1+1:\ell_1+m-1]\cup\{\ell+1\}}) + H(U^{(\ell_1,\ell)}, S_{\to[\ell_1+1:\ell]}, S_{[\ell+2:k]\to\ell+1})\\
&\overset{(b)}{\geqslant} H(S_{k+1\to[\ell_1+1:\ell_1+m-1]}) + H(U^{(\ell_1,\ell)}, S_{[\ell+2:k+1]\to\ell+1})\\
&\overset{(c)}{=} H(S_{\ell_1+1\to[1:m-1]}) + H(U^{(\ell_1,\ell)}, S_{[\ell+2:k+1]\to\ell+1}),
\end{aligned} \tag{23}
$$

where (a) and (c) follow from the fact that $H(S_{\ell_1+1\to[1:m]}) = H(S_{k+1\to[1:m-1]\cup\{\ell+1\}})$ and $H(S_{k+1\to[1:m-1]}) = H(S_{\ell_1+1\to[1:m-1]})$ due to the symmetrical code that we consider, and (b) follows from the submodularity of the entropy function. Add (23) over $m \in [1 : \ell_1]$ (since $\ell_1 \leqslant \ell_2$ as our setup) and cancel $\sum_{m=1}^{\ell_1-1} H(S_{d+1\to[1:m]})$ from both sides. We have

$$
H(S_{d+1\to[1:\ell]}) + \ell_1 H(U^{(\ell_1,\ell)}, S_{[\ell+2:k+1]\to\ell+1}) \geqslant \ell_1 H(U^{(\ell_1,\ell)}, S_{[\ell+2:k+1]\to\ell+1}). \tag{24}
$$

Add (25) over $k \in [\ell + 1 : d]$ and cancel $\sum_{k=\ell+1}^{d-1} H(U^{(\ell_1,\ell)}, S_{[\ell+2:k+1]\to\ell+1})$ from both sides. We have

$$
(d-\ell) H(S_{d+1\to[1:\ell]}) + \ell_1 H(U^{(\ell_1,\ell)}) \geqslant \ell_1 H(U^{(\ell_1,\ell)}, S_{[\ell+2:d+1]\to\ell+1}) = \ell_1 H(U^{(\ell_1,\ell+1)}). \tag{25}
$$

Multiplying both sides by $(d-\ell)^{-1}$ completes the proof of (22).

Proposition 3.

$$H(U^{(\ell_1+1,m)})+\frac{d-m}{d-\ell}H(U^{(\ell_1,\ell+1)})\geqslant(d-m)\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+H(U^{(\ell_1+1,m+1)})+\frac{d-m}{d-\ell}H(U^{(\ell_1,\ell)}) \quad (26)$$

for any $m\in[\ell+1:d-1]$. Consequently,

$$H(U^{(\ell_1+1,\ell+1)})+\frac{T_{d,d,\ell+1}}{d-\ell}H(U^{(\ell_1,\ell+1)})\geqslant T_{d,d,\ell}\sum_{j=\ell+1}^{d}T_{d,j,\ell}^{-1}B_j+\frac{T_{d,d,\ell}}{d-\ell}H(U^{(\ell_1,\ell)}). \quad (27)$$

Proof. To see (26), note that for any $m\in[\ell+1:d-1]$, we have

$$\begin{aligned}
&H(U^{(\ell_1+1,m)}|M_{[\ell+1:m]})+\frac{d-m}{d-\ell}H(U^{(\ell_1,\ell+1)})\\
&\overset{(a)}{\geqslant} H(U^{(\ell_1+1,m)}|M_{[\ell+1:m]})\\
&\quad+(d-m)\left(\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+T_{d,m,\ell}^{-1}H(U^{(\ell_1,m)}|M_{[\ell+1:m]})+\left(\frac{1}{d-\ell}-T_{d,m,\ell}^{-1}\right)H(U^{(\ell_1,\ell)})\right)\\
&=H(U^{(\ell_1+1,m)}|M_{[\ell+1:m]})+(d-m)T_{d,m,\ell}^{-1}H(U^{(\ell_1,m)}|M_{[\ell+1:m]})\\
&\quad+(d-m)\left(\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+\left(\frac{1}{d-\ell}-T_{d,m,\ell}^{-1}\right)H(U^{(\ell_1,\ell)})\right)\\
&\overset{(b)}{\geqslant} H(U^{(\ell_1+1,m+1)}|M_{[\ell+1:m]})+(d-m)T_{d,m,\ell}^{-1}H(U^{(\ell_1,\ell)}|M_{[\ell+1:m]})\\
&\quad+(d-m)\left(\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+\left(\frac{1}{d-\ell}-T_{d,m,\ell}^{-1}\right)H(U^{(\ell_1,\ell)})\right)\\
&\overset{(c)}{=} H(U^{(\ell_1+1,m+1)}|M_{[\ell+1:m]})+(d-m)T_{d,m,\ell}^{-1}H(U^{(\ell_1,\ell)})\\
&\quad+(d-m)\left(\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+\left(\frac{1}{d-\ell}-T_{d,m,\ell}^{-1}\right)H(U^{(\ell_1\ell)})\right)\\
&=H(U^{(\ell_1+1,m+1)}|M_{[\ell+1:m]})+(d-m)\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j+\frac{d-m}{d-\ell}H(U^{(\ell_1,\ell)}),
\end{aligned}$$

where (a) follows from (18) of Proposition 1; (b) follows from Corollary 2; and (c) follows from the fact that $I(U^{(\ell_1,\ell)};M_{[\ell+1:m]})=0$ due to the repair secrecy requirement (4). Adding $H(M_{[\ell+1:m]})$ to both sides and using the facts that

$$H(U^{(\ell_1+1,m)}|M_{[\ell+1:m]})+H(M_{[\ell+1:m]})=H(U^{(\ell_1+1,m)},M_{[\ell+1:m]})\overset{(a)}{=}H(U^{(\ell_1+1,m)}),$$

and that

$$H(U^{(\ell_1+1,m+1)}|M_{[\ell+1:m]})+H(M_{[\ell+1:m]})=H(U^{(\ell_1+1,m+1)},M_{[\ell+1:m]})\overset{(b)}{=}H(U^{(\ell_1+1,m+1)})$$

complete the proof of (26). Here, (a) and (b) are due to the facts that $M_{[\ell+1:m]}$ is a function of $W_{[1:m]}$, which is a function of both $U^{(\ell_1+1,m)}$ and $U^{(\ell_1+1,m+1)}$ by Lemma 1.

To see (27), add (26) over $m\in[\ell+1:d-1]$ and cancel $\sum_{m=\ell+2}^{d-1}H(U^{(\ell_1+1,m)})$ from both sides of the inequality. We have

$$H(U^{(\ell_1+1,\ell+1)})+\frac{T_{d,d,\ell+1}}{d-\ell}H(U^{(\ell_1,\ell+1)})$$

$$\geqslant \sum_{m=\ell+1}^{d-1}\left((d-m)\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j\right)+H(U^{(\ell_1+1,d)})+\frac{T_{d,d,\ell+1}}{d-\ell}H(U^{(\ell_1,\ell)}). \tag{28}$$

Note that

$$\sum_{m=\ell+1}^{d-1}\left((d-m)\sum_{j=\ell+1}^{m}T_{d,j,\ell}^{-1}B_j\right)=\sum_{j=\ell+1}^{d-1}T_{d,j,\ell}^{-1}B_j\left(\sum_{m=j}^{d-1}(d-m)\right)=\sum_{j=\ell+1}^{d-1}T_{d,j,\ell}^{-1}T_{d,d,j}B_j. \tag{29}$$

Furthermore,

$$\begin{aligned}
H(U^{(\ell_1+1,d)}) &\overset{\text{(a)}}{=} H(U^{(\ell_1+1,d)},M_{[\ell+1:d]})\\
&\overset{\text{(b)}}{=} H(U^{(\ell_1+1,d)}|M_{[\ell+1:d]})+H(M_{[\ell+1:d]})\\
&\overset{\text{(c)}}{=} H(U^{(\ell_1+1,d)}|M_{[\ell+1:d]})+\sum_{j=\ell+1}^{d}B_j\\
&\geqslant H(U^{(\ell_1,\ell)}|M_{[\ell+1:d]})+\sum_{j=\ell+1}^{d}B_j\\
&\overset{\text{(d)}}{=} H(U^{(\ell_1,\ell)})+\sum_{j=\ell+1}^{d}B_j,
\end{aligned} \tag{30}$$

where (a) follows from the fact that $M_{[\ell+1:d]}$ is a function of $W_{[1:d]}$, which is a function of $U^{(\ell_1+1,d)}$ by Lemma 1; (b) follows from the chain rule for entropy; (c) follows from the fact that $H(M_{[\ell+1:d]})=\sum_{j=\ell+1}^{d}B_j$; and (d) follows from the fact that $I(U^{(\ell_1,\ell)};M_{[\ell+1:d]})=0$ due to the repair secrecy requirement (4).

Substituting (29) and (30) into (28) gives

$$\begin{aligned}
&H(U^{(\ell_1+1,\ell+1)})+\frac{T_{d,d,\ell+1}}{d-\ell}H(U^{(\ell_1,\ell+1)})\\
&\geqslant \sum_{j=\ell+1}^{d-1}T_{d,j,\ell}^{-1}T_{d,d,j}B_j+\sum_{j=\ell+1}^{d}B_j+\left(1+\frac{T_{d,d,\ell+1}}{d-\ell}\right)H(U^{(\ell_1,\ell)})\\
&=\sum_{j=\ell+1}^{d-1}T_{d,j,\ell}^{-1}(T_{d,d,j}+T_{d,j,\ell})B_j+B_d+\frac{T_{d,d,\ell}}{d-\ell}H(U^{(\ell_1,\ell)})\\
&\overset{\text{(a)}}{=}T_{d,d,\ell}\sum_{j=\ell+1}^{d-1}T_{d,j,\ell}^{-1}B_j+B_d+\frac{T_{d,d,\ell}}{d-\ell}H(U^{(\ell_1,\ell)})\\
&=T_{d,d,\ell}\sum_{j=\ell+1}^{d}T_{d,j,\ell}^{-1}B_j+\frac{T_{d,d,\ell}}{d-\ell}H(U^{(\ell_1,\ell)}),
\end{aligned}$$

where (a) follows from the fact that $T_{d,d,j}+T_{d,j,\ell}=T_{d,d,\ell}$. This completes the proof of the proposition.

We are now ready to prove the outer bounds (8) and (9). To prove (8), note that

$$\begin{aligned}
\beta+\frac{1}{d-\ell}H(U^{(\ell_1,\ell)}) &\overset{\text{(a)}}{\geqslant} \frac{1}{d-\ell}\left(H(\overline{S}_{\to\ell+1})+H(U^{(\ell_1,\ell)})\right)\\
&\overset{\text{(b)}}{\geqslant} \frac{1}{d-\ell}H(U^{(\ell_1,\ell+1)})\\
&\overset{\text{(c)}}{\geqslant} \sum_{j=\ell+1}^{d}T_{d,j,\ell}^{-1}B_j+\frac{1}{d-\ell}H(U^{(\ell_1,\ell)}),
\end{aligned}$$

where (a) follows from the fact that $H(\overline{S}_{\to \ell+1}) \leqslant (d-\ell)\beta$; (b) follows from the union bound on entropy; and (c) follows from (19) of Proposition 1. Cancelling $\frac{1}{d-\ell}H(U^{(\ell_1,\ell)})$ from both sides of the inequality and normalizing both sides by $\sum_{t=\ell+1}^{d} B_t$ complete the proof of (8).

To prove (9), note that

$$
\begin{aligned}
&\alpha + T_{d,d,\ell_1+1}\beta + \frac{\ell_1 + T_{d,d,\ell_1}}{d-\ell} H(U^{(\ell_1,\ell)}) \\
&\overset{(a)}{=} \alpha + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,\ell+1,\ell_1+1}}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell} + \frac{d-\ell_1}{d-\ell}\right) H(U^{(\ell_1,\ell)}) \\
&= \frac{T_{d,\ell+1,\ell_1+1}}{d-\ell} H(U^{(\ell_1,\ell)}) + \alpha + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell} + \frac{d-\ell_1}{d-\ell}\right) H(U^{(\ell_1,\ell)}) \\
&\overset{(b)}{\geqslant} \frac{T_{d,\ell,\ell_1+1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) + H(U^{(\ell_1,\ell_1+1)}) + \alpha + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell} + \frac{d-\ell_1}{d-\ell}\right) H(U^{(\ell_1,\ell)}) \\
&= \frac{d-\ell_1}{d-\ell} H(U^{(\ell_1,\ell)}) + H(U^{(\ell_1,\ell_1+1)}, S_{\ell_1+1\to[1:\ell_1]}) + \alpha + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell}\right) H(U^{(\ell_1,\ell)}) \\
&\quad + \frac{T_{d,\ell,\ell_1+1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(c)}{\geqslant} \frac{d-\ell_1}{d-\ell} H(U^{(\ell_1,\ell+1)}) + H(U^{(\ell_1,\ell_1)}, S_{\ell_1+1\to[1:\ell_1]}) + \alpha + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell}\right) H(U^{(\ell_1,\ell)}) \\
&\quad + \frac{T_{d,\ell,\ell_1+1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(d)}{=} \alpha + H(U^{(\ell_1,\ell_1)}, S_{\ell_1+1\to[1:\ell_1]}) + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell}\right) H(U^{(\ell_1,\ell)}) + \frac{T_{d,\ell,\ell_1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(e)}{\geqslant} H(W_{\ell+1}) + H(U^{(\ell_1,\ell_1)}, S_{\ell_1+1\to[1:\ell_1]}) + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell}\right) H(U^{(\ell_1,\ell)}) + \frac{T_{d,\ell,\ell_1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(f)}{=} H(W_{\ell+1}, S_{\ell+1\to[1:\ell_1]}) + H(U^{(\ell_1,\ell_1)}, S_{\ell_1+1\to[1:\ell_1]}) + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell}\right) H(U^{(\ell_1,\ell)}) \\
&\quad + \frac{T_{d,\ell,\ell_1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(g)}{\geqslant} H(S_{\ell_1+1\to[1:\ell_1]}) + H(U^{(\ell_1+1,\ell_1+1)}, S_{\ell_1+1\to[1:\ell_1]}) + T_{d,d,\ell_1+1}\beta + \left(\frac{\ell_1}{d-\ell} + \frac{T_{d,d,\ell+1}}{d-\ell}\right) H(U^{(\ell_1,\ell)}) \\
&\quad + \frac{T_{d,\ell,\ell_1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(h)}{=} H(S_{\ell_1+1\to[1:\ell_1]}) + \frac{\ell_1}{d-\ell} H(U^{(\ell_1,\ell)}) + H(U^{(\ell_1+1,\ell_1+1)}) + T_{d,d,\ell_1+1}\beta + \frac{T_{d,d,\ell+1}}{d-\ell} H(U^{(\ell_1,\ell)}) \\
&\quad + \frac{T_{d,\ell,\ell_1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(i)}{\geqslant} \frac{\ell_1}{d-\ell} H(U^{(\ell_1,\ell+1)}) + H(U^{(\ell_1+1,\ell_1+1)}) + T_{d,d,\ell_1+1}\beta + \frac{T_{d,d,\ell+1}}{d-\ell} H(U^{(\ell_1,\ell)}) + \frac{T_{d,\ell,\ell_1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(j)}{=} H(U^{(\ell_1+1,\ell_1+1)}) + T_{d,\ell+1,\ell_1+1}\beta + \frac{T_{d,d,\ell+1}}{d-\ell} H(U^{(\ell_1,\ell)}) + T_{d,d,\ell+1}\beta + \frac{T_{d,\ell,\ell_1} + \ell_1}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(k)}{\geqslant} H(U^{(\ell_1+1,\ell+1)}) + \frac{T_{d,d,\ell+1}}{d-\ell} H(U^{(\ell_1,\ell+1)}) + \frac{T_{d,\ell,\ell_1} + \ell_1}{d-\ell} H(U^{(\ell_1,\ell+1)}) \\
&\overset{(l)}{\geqslant} (T_{d,d,\ell} + T_{d,\ell,\ell_1} + \ell_1) \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} B_j + \frac{T_{d,d,\ell} + T_{d,\ell,\ell_1} + \ell_1}{d-\ell} H(U^{(\ell_1,\ell)}) \\
&\overset{(m)}{=} (T_{d,d,\ell_1} + \ell_1) \sum_{j=\ell+1}^{d} T_{d,j,\ell}^{-1} B_j + \frac{T_{d,d,\ell_1} + \ell_1}{d-\ell} H(U^{(\ell_1,\ell)}),
\end{aligned}
$$

where (a) follows from the fact that $T_{d,\ell+1,\ell_1+1} + T_{d,d,\ell+1} + d - \ell_1 = T_{d,d,\ell_1}$; (b) follows from Corollary 1

that

$$\frac{T_{d,\ell+1,\ell_1+1}}{d-\ell}H(U^{(\ell_1,\ell)}) \geqslant \frac{T_{d,\ell,\ell_1+1}}{d-\ell}H(U^{(\ell_1,\ell+1)}) + H(U^{(\ell_1,\ell_1+1)})$$

by setting $j_1 = \ell$, $j_2 = \ell_1 + 1$ and $i = \ell_1$ in (12), while $M^{(\ell)} = \emptyset$ as our problem setup; (c) follows from (17) in Lemma 3; (d) follows from the fact that $T_{d,\ell,\ell_1+1} + d - \ell_1 = T_{d,\ell,\ell_1}$; (e) follows from the fact that $H(W_{\ell+1}) \leqslant \alpha$; (f) and (h) follows from the fact that $S_{\ell_1+1\to[1:\ell_1]}$ is a function of W_{ℓ_1+1}; (g) follows from the fact that $H(W_{\ell+1}, S_{\ell_1+1\to[1:\ell_1]}) + H(U^{(\ell_1,\ell_1)}, S_{\ell_1+1\to[1:\ell_1]}) \geqslant H(S_{\ell_1+1\to[1:\ell_1]}) + H(U^{(\ell_1+1,\ell_1+1)}, S_{\ell_1+1\to[1:\ell_1]})$ due to submodularity; (i) follows from (22) in Proposition 2; (j) follows from the fact that $T_{d,d,\ell_1+1} = T_{d,d,\ell+1} + T_{d,\ell+1,\ell_1+1}$; (k) follows from the facts that $T_{d,\ell+1,\ell_1+1}\beta \geqslant \overline{S}_{\to[\ell_1+2:\ell+1]}$ and $(d-\ell)\beta \geqslant \overline{S}_{\to\ell+1}$; (l) follows from (19) and (27) of Propositions 1 and 3, respectively; (m) follows from the fact that $T_{d,d,\ell} + T_{d,\ell,\ell_1} = T_{d,d,\ell_1}$. Cancelling $\frac{T_{d,d,\ell_1+\ell_1}}{d-\ell}H(U^{(\ell_1,\ell)})$ from both sides of the inequality and normalizing both sides by $\sum_{t=\ell+1}^{d} B_t$ complete the proof of (9).

5 Concluding remarks

This paper considered the problem of MDC-SR with a generalized eavesdropping model. It was demonstrated that the MBR point of the achievable normalized storage-capacity repair-bandwidth tradeoff region does not depend on the total numbers of Types I and II compromised storage nodes, if the number of Type I compromised nodes is less than or equal to the number of Type II compromised nodes. As a future study, it would be interesting to see whether this result extends to the entire achievable normalized storage-capacity repair-bandwidth tradeoff region.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61631017) and National Science Foundation of USA (Grant Nos. CCF-15-24839, CCF-15-26095).

References

1 Shao S, Liu T, Tian C, et al. Multilevel diversity coding with secure regeneration: separate coding achieves the MBR point. 2017. ArXiv:1712.03326
2 Tian C, Liu T. Multilevel diversity coding with regeneration. IEEE Trans Inf Theory, 2016, 62: 4833–4847
3 Shao S, Liu T, Tian C. Multilevel diversity coding with regeneration: separate coding achieves the MBR point. In: Proceedings of Annual Conference on Information Science and Systems, Princeton, 2016. 602–607
4 Pawar S, Rouayheb S E, Ramchandran K. On secure distributed data storage under repair dynamics. In: Proceedings of IEEE International Symposium on Information Theory, Austin, 2010. 2543–2547
5 Pawar S, El Rouayheb S, Ramchandran K. Securing dynamic distributed storage systems against eavesdropping and adversarial attacks. IEEE Trans Inf Theory, 2011, 57: 6734–6753
6 Shah N B, Rashmi K V, Kumar P V. Information-theoretically secure regenerating codes for distributed storage. In: Proceedings of IEEE Global Telecommunications Conference, Kathmandu, 2011
7 Goparaju S, Rouayheb S E, Calderbank R, et al. Data secrecy in distributed storage systems under exact repair. In: Proceedings of IEEE International Symposium on Network Coding (NetCod), Calgary, 2013
8 Rawat A S, Koyluoglu O O, Silberstein N, et al. Optimal locally repairable and secure codes for distributed storage systems. IEEE Trans Inf Theory, 2014, 60: 212–236
9 Tandon R, Amuru S D, Clancy T C, et al. Toward optimal secure distributed storage systems with exact repair. IEEE Trans Inf Theory, 2016, 62: 3477–3492
10 Ye F, Shum K W, Yeung R W. The rate region for secure distributed storage systems. IEEE Trans Inf Theory, 2017, 63: 7038–7051
11 Shao S, Liu T, Tian C, et al. On the tradeoff region of secure exact-repair regenerating codes. IEEE Trans Inf Theory, 2017, 63: 7253–7266
12 Tian C. Characterizing the rate region of the (4,3,3) exact-repair regenerating codes. IEEE J Sel Areas Commun, 2014, 32: 967–975

Appendix A Proof of the exchange Lemma 2

First note that $d-\ell_1 > d-\ell$, so we may write $d-\ell_1 = s(d-\ell)+r$ for some integer $s \geqslant 1$ and $r \in [1:d-\ell]$. Next, let

$$a_t := \begin{cases} t, & t \in [1:\ell_1], \\ t+\ell_1, & t \in [\ell_1+1:\ell-\ell_1], \\ t+\ell+1, & t \in [\ell-\ell_1+1:d-\ell]. \end{cases}$$

Finally, let $\tau_0 := \{a_t : t \in [1:r]\}$ and

$$\tau_q := \{a_t : t \in [r+1+(q-1)(d-\ell) : r+q(d-\ell)]\}$$

for any $q \in [1:s]$. It is straightforward to verify that:

- $\tau_q \cap \tau_{q'} = \emptyset$ for any $q \neq q'$;
- $\bigcup_{q=0}^{s-1} \tau_q = [1:\ell_1] \cup [2\ell_1+1:\ell]$;
- $\tau_s = [\ell+2:d+1]$.

Consider a symmetrical $(n=d+1, d, N_1, \ldots, N_d, T, R)$ code that satisfies the node regeneration requirement (3). Let us show by induction that for any $p \in [1:s]$, we have

$$\begin{aligned} &pH(U^{(\ell_1,\ell)}|M^{(\ell)}) + H(U^{(\ell_1,\ell_1+1)}|M^{(\ell)}) \\ &\geqslant pH(U^{(\ell_1,\ell+1)}|M^{(\ell)}) + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell_1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-p}\tau_q\to\ell+1}|M^{(\ell)}\Big). \end{aligned} \tag{A1}$$

To prove the base case of $p=1$, first note that

$$\begin{aligned} H(U^{(\ell_1,\ell)}|M^{(\ell)}) &\overset{\text{(a)}}{=} H(U^{(\ell_1,\ell)}, W_{[1:\ell_1]}, \underline{S}_{\to[\ell_1+1:\ell]}|M^{(\ell)}) \\ &= H(W_{[1:\ell_1]}, S_{\to[\ell_1+1:\ell]}|M^{(\ell)}) \\ &\overset{\text{(b)}}{=} H(W_{[1:\ell]}, S_{\to[\ell_1+1:\ell]}, S_{[1:\ell]\to\ell+1}|M^{(\ell)}), \end{aligned}$$

where (a) follows from the fact that $(W_{[\ell_1:\ell]}, \underline{S}_{\to[\ell_1+1:\ell]})$ is a function of $U^{(\ell_1,\ell)}$ by Lemma 1, and (b) follows from the fact that $S_{[1:\ell]\to\ell+1}$ is a function of $W_{[1:\ell]}$. Furthermore,

$$\begin{aligned} H(U^{(\ell_1,\ell_1+1)}|M^{(\ell)}) &\overset{\text{(a)}}{=} H(U^{(\ell_1,\ell_1+1)}, \underline{S}_{\to\ell_1+1}|M^{(\ell)}) \\ &= H(W_{[1:\ell_1]}, S_{\to\ell_1+1}|M^{(\ell)}) \\ &\overset{\text{(b)}}{=} H(W_{[\ell_1+1:2\ell_1]}, S_{\to 2\ell_1+1}|M^{(\ell)}) \\ &\overset{\text{(c)}}{=} H(W_{[\ell_1+1:2\ell_1]}, S_{\to\ell+1}|M^{(\ell)}) \\ &\overset{\text{(d)}}{=} H(W_{[\ell_1+1:2\ell_1]}, S_{[1:\ell_1]\to\ell+1}, S_{[2\ell_1+1:\ell]\to\ell+1}, S_{[\ell+2:d+1]\to\ell+1}, S_{\ell+1\to[\ell_1+1:2\ell_1]}|M^{(\ell)}), \end{aligned}$$

where (a) follows from the fact that $\underline{S}_{\to\ell_1+1}$ is a function of $U^{(\ell_1,\ell_1+1)}$ by Lemma 1, and (b) and (c) follow from the symmetrical code that we consider; (d) follows that $S_{\ell+1\to[\ell_1+1:2\ell_1]}$ is a function of $S_{\to\ell+1}$. It follows that

$$\begin{aligned} &H(U^{(\ell_1,\ell)}|M^{(\ell)}) + H(U^{\ell_1,\ell_1+1}|M^{(\ell)}) \\ &\geqslant H(W_{[1:\ell]}, S_{\to[\ell_1+1:\ell]}, S_{[1:\ell]\to\ell+1}|M^{(\ell)}) + H(W_{[\ell_1+1:2\ell_1]}, S_{[1:\ell_1]\to\ell+1}, S_{[2\ell_1:\ell]\to\ell+1}, S_{[\ell+2:d+1]\to\ell+1}, S_{\ell+1\to[\ell_1+1:2\ell_1]}|M^{(\ell)}) \\ &\overset{\text{(a)}}{\geqslant} H(W_{[1:\ell]}, S_{\to[\ell_1+1:\ell]}, S_{[1:\ell]\to\ell+1}, S_{[\ell+2:d+1]\to\ell+1}|M^{(\ell)}) + H(W_{[\ell_1+1:2\ell_1]}, S_{[1:\ell_1]\to\ell+1}, S_{[2\ell_1:\ell]\to\ell+1}, S_{\ell+1\to[\ell_1+1:2\ell_1]}|M^{(\ell)}) \\ &= H(U^{(\ell_1,\ell)}, \underline{S}_{\to\ell+1}|M^{(\ell)}) + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-1}\tau_q\to\ell+1}|M^{(\ell)}\Big) \\ &= H(U^{(\ell_1,\ell+1)}|M^{(\ell)}) + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-1}\tau_q\to\ell+1}|M^{(\ell)}\Big), \end{aligned}$$

where (a) follows from the submodularity of the entropy function. This completes the proof of the base case of $p=1$.

Assume that (A1) holds for some $p \in [1:s-1]$. We have

$$\begin{aligned} &(p+1)H(U^{(\ell_1,\ell)}|M^{(\ell)}) + H(U^{(\ell_1,\ell_1+1)}|M^{(\ell)}) \\ &= H(U^{(\ell_1,\ell)}|M^{(\ell)}) + (pH(U^{(\ell_1,\ell)}|M^{(\ell)}) + H(U^{(\ell_1,\ell_1+1)}|M^{(\ell)})) \\ &\geqslant H(U^{(\ell_1,\ell)}|M^{(\ell)}) + pH(U^{(\ell_1,\ell+1)}|M^{(\ell)}) + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-p}\tau_q\to\ell+1}|M^{(\ell)}\Big). \end{aligned} \tag{A2}$$

Note that both $S_{\ell+1\to[\ell_1+1:2\ell_1]}$ and $S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}$ are functions of $W_{[1:\ell]}$, which is in turn a function of $U^{(\ell_1,\ell)}$ by Lemma 1. We thus have

$$H(U^{(\ell_1,\ell)}|M^{(\ell)}) = H\Big(U^{(\ell_1,\ell)}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}|M^{(\ell)}\Big).$$

Furthermore, by the symmetrical code that we consider we have

$$H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-p}\tau_q\to\ell+1}|M^{(\ell)}\Big)$$

$$= H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}, S_{[\ell+2:d+1]\to\ell+1}|M^{(\ell)}\Big).$$

It follows that

$$\begin{aligned}
&H(U^{(\ell_1,\ell)}|M^{(\ell)}) + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-p}\tau_q\to\ell+1}|M^{(\ell)}\Big)\\
&= H\Big(U^{(\ell_1,\ell)}, S_{\ell+1\to[\ell_1+1,2\ell_1]}, S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}|M^{(\ell)}\Big) + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]},\\
&\quad S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}, S_{[\ell+2:d+1]\to\ell+1}|M^{(\ell)}\Big)\\
&\overset{(a)}{\geqslant} H\Big(U^{(\ell_1,\ell)}, S_{\ell+1\to[\ell_1+1,2\ell_1]}, S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}, S_{[\ell+2:d+1]\to\ell+1}|M^{(\ell)}\Big)\\
&\quad + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}|M^{(\ell)}\Big)\\
&= H(U^{(\ell_1,\ell+1)}|M^{(\ell)}) + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}|M^{(m)}\Big),
\end{aligned} \tag{A3}$$

where (a) follows from the submodularity of the entropy function. Substituting (A3) into (A2) gives

$$\begin{aligned}
&(p+1)H(U^{(\ell_1,\ell)}|M^{(\ell)}) + H(U^{(\ell_1,\ell_1+1)}|M^{(\ell)})\\
&\geqslant (p+1)H(U^{(\ell_1,\ell+1)}|M^{(\ell)}) + H\Big(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\bigcup_{q=0}^{s-(p+1)}\tau_q\to\ell+1}|M^{(\ell)}\Big),
\end{aligned}$$

which completes the induction step and hence the proof of (A1).

Setting $p = s$ in (A1), we have

$$\begin{aligned}
&sH(U^{(\ell_1,\ell)}|M^{(\ell)}) + H(U^{(\ell_1,\ell_1+1)}|M^{(\ell)})\\
&\geqslant sH(U^{(\ell_1,\ell+1)}|M^{(\ell)}) + H(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, S_{\tau_0\to\ell+1}|M^{(\ell)})\\
&= sH(U^{(\ell_1,\ell+1)}|M^{(\ell)}) + H(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}|M^{(\ell)}) + H(S_{\tau_0\to\ell+1}|W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, M^{(\ell)}).
\end{aligned} \tag{A4}$$

By the symmetrical codes that we consider, we have

$$H(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}|M^{(\ell)}) = H(W_{[1:\ell_1]}, S_{\ell+1\to[1:\ell_1]}|M^{(\ell)}) = H(W_{[1:\ell_1]}, S_{\ell_1+1\to[1:\ell_1]}|M^{(\ell)}) \tag{A5}$$

and

$$H(S_{\tau_0\to\ell+1}|W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, M^{(\ell)}) = H(S_{\tau\to\ell+1}|W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, M^{(\ell)})$$

for any subset $\tau \subseteq [\ell+2:d+1]$ such that $|\tau| = r$. By Han's subset inequality[1)], we have

$$\begin{aligned}
&H(S_{\tau_0\to\ell+1}|W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, M^{(\ell)})\\
&\geqslant \frac{r}{d-\ell}H(S_{[\ell+2:d+1]\to\ell+1}|W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, M^{(\ell)})\\
&\geqslant \frac{r}{d-\ell}H(S_{[\ell+2:d+1]\to\ell+1}|W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]}, M^{(\ell)}, U^{(\ell_1,\ell)})\\
&\overset{(a)}{=} \frac{r}{d-\ell}H(S_{[\ell+2:d+1]\to\ell+1}|U^{(\ell_1,\ell)}, M^{(\ell)})\\
&= \frac{r}{d-\ell}(H(S_{[\ell+2:d+1]\to\ell+1}, U^{(\ell_1,\ell)}|M^{(\ell)}) - H(U^{(\ell_1,\ell)}|M^{(\ell)}))\\
&= \frac{r}{d-\ell}(H(U^{(\ell_1,\ell+1)}|M^{(\ell)}) - H(U^{(\ell_1,\ell)}|M^{(\ell)})),
\end{aligned} \tag{A6}$$

where (a) follows from the fact that $(W_{[\ell_1+1:2\ell_1]}, S_{\ell+1\to[\ell_1+1:2\ell_1]})$ is a function of $U^{(\ell_1,\ell)}$ by Lemma 1. Substituting (A5) and (A6) into (A4) gives

$$\left(s + \frac{r}{d-\ell}\right)H(U^{(\ell_1,\ell)}|M^{(\ell)}) + H(U^{(\ell_1,\ell_1+1)}|M^{(\ell)}) \geqslant \left(s + \frac{r}{d-\ell}\right)H(U^{(\ell_1,\ell+1)}|M^{(\ell)}) + H(U^{(\ell_1,\ell_1)}|M^{(\ell)}),$$

which is equivalent to (11) by noting that

$$s + \frac{r}{d-\ell} = \frac{s(d-\ell)+r}{d-\ell} = \frac{d-\ell_1}{d-\ell}.$$

This completes the proof of the exchange lemma.

1) Han T S. Nonnegative entropy measures of multivariate symmetric correlations. Inf Control, 1978, 36: 133–156