Lecture 12: Constant Degree Lossless Expanders

Topics in Complexity Theory and Pseudorandomness (Spring 2013)
Rutgers University
Swastik Kopparty
Scribes: Meng Li, Yun Kuen Cheung

1 Overview

In this lecture, we will construct constant degree D vertex-expanders with expansion of $(1-\varepsilon)D$ (a.k.a. constant degree lossless expanders). This beautiful construction is due to Capalbo-Reingold-Vadhan-Wigderson. Concretely, for every constant $\varepsilon > 0$ and every N, we will construct a bipartite graph (L, R, E), $|L| = N$, $|R| = M = \mathtt{poly}(\varepsilon)N)$, with left degree $D = \mathtt{poly}(\frac{1}{\varepsilon})$, such that every subset S of L of size at most $\frac{\mathtt{poly}(\varepsilon)M}{D}$, the size of its neighborhood $\Gamma(S)$ is at least $(1-\varepsilon) \cdot D \cdot |S|$. In contrast, our earlier lossless condenser construction (which actually is a lossless expander) had polylogarithmic degree.

Remark. Such extreme vertex expansion is not implied by any kind of eigenvalue expansion. There exist D-regular λ absolute eigenvalue expanders with $\lambda = 2\sqrt{D-1}$ (this is the smallest possible; such graphs are called Ramanujan graphs), but which have only $D/2$ vertex expansion.

Our main tool will be the notion of a randomness "conductor". Conductors provide a unifying framework for thinking about and working with extractors, condensers and expanders.

2 Randomness Conductors

Definition 1. *A $(k_{max}, a, \varepsilon)$ conductor is a function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ such that $\forall k \leq k_{max}$, for all distributions X on $\{0,1\}^n$ with $H_\infty(X) \geq k$, $C(X, U_d)$ is $\varepsilon-$closed to a distribution Y with $H_\infty(Y) \geq k + a$.*

Note that the definitions of conductors and condensers are similar. The difference is that condensers are required to work for only one value of k while conductors work for any value of k smaller than k_{max}.

Observe that in the definition, we cannot have $k_{max} > m - a$. If we have $k_{max} = m - a$, then the conductor is actually an $(m-a, \varepsilon)$ extractor (since the output entropy, when the input entropy is $m - a$, is m). This motivates the following definition.

Definition 2. *An (a, ε) extracting conductor is an $(m - a, a, \varepsilon)$ conductor.*

As noted above, an (a, ε) extracting conductor is also an $(m - a, \varepsilon)$ extractor (but not the other way around; an extractor gives no promises about what it will do on sources of small entropy).

A lossless conductor is one which does not "lose" any entropy: i.e. if $a = d$.

Definition 3. *A (k_{max}, ε) lossless conductor is a $(k_{max}, d, \varepsilon)$ conductor.*

Note that a lossless conductor is also a lossless condenser (but not the other way around).

Via the usual way of constructing a bipartite graph from an extractor/condenser, the following theorem is straight-forward.

Theorem 4. *A (k_{max}, ε) lossless conductor gives a (N, M, D) expander with $N = 2^n, M = 2^m, D = 2^d$ such that for any set of size at most $K_{max} = 2^{k_{max}}$, its neighbor set expands by a factor of $(1 - \varepsilon)D$.*

Just to get warmed up, let us show that absolute eigenvalue expanders give good conductors.

Theorem 5. *Let $L = \{0,1\}^n, R = \{0,1\}^m$, $D = 2^d$, $m = n$ and λ is the second largest absolute eigenvalue. $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is defined as follows: $C(v, i)$ is the i^{th} neighbor of v. Then C is an $\left(n - d, 2\log\frac{2^d}{\lambda} - \log\frac{1}{\varepsilon} - 1, \varepsilon\right)$ conductor.*

We being with a useful sufficient criterion for a distribution to be close to high min-entropy: it should have low ℓ_2 norm.

Lemma 6. *Suppose $\overrightarrow{q}$ is a distribution such that $\sum q_i^2 \leq 2^{-b}$. Then $\overrightarrow{q}$ is ε -close to a distribution with min entropy at least $b - \log\frac{1}{\varepsilon}$.*

Proof. Let J be the set of indices such that $q_i \geq \dfrac{2^{-b}}{\varepsilon}$. Therefore we have

$$2^{-b} \geq \sum q_i^2 \geq \sum_{i \in J} q_i^2 \geq \frac{2^{-b}}{\varepsilon} \sum_{i \in J} q_i$$

$$\sum_{i \in J} q_i \leq \varepsilon$$

Now consider the vector $\overrightarrow{x}$ with the same support as q given by:

$$x_i = \begin{cases} 0 & i \in J \\ q_i & \text{others} \end{cases}$$

Note that $\overrightarrow{x}$ is not a probability distribution. But we do have that following two nice properties: (1) $\overrightarrow{x}$ is ε-close to $\overrightarrow{q}$, and (2) that it is pointwise at most $\frac{2^{-b}}{\varepsilon}$. With some patience, one can perturb $\overrightarrow{x}$ into a probability distribution, while preserving these properties (but we will not do it here).

Then $H_\infty(x) \geq b - \log\dfrac{1}{\varepsilon}$, and $\overrightarrow{q}$ is ε-close to distribution $\overrightarrow{x}$. This is what we wanted. □

Proof. (of Theorem 5) Take $S \subseteq \{0,1\}^n, |S| = 2^k, k \leq k_{max} = n - d$. Denote $D = 2^d, N = 2^n, A$ as the adjacency matrix of the bipartite λ absolute eigenvalue expander G. Let $\overrightarrow{p}$ be the uniform distribution on S. Hence

$$\overrightarrow{p} = \overrightarrow{1_S} \cdot \frac{1}{|S|} = \sum_{i=1}^{N} \alpha_i v_i$$

where $\{v_i\}$ is a set of orthonormal basis eigenvectors of A for $\mathbb{R}^N$, $\alpha_i = \langle \overrightarrow{p}, v_i \rangle$. In particular, $\alpha_1 = \frac{1}{\sqrt{N}}$.

Recall that the distribution of $C(\overrightarrow{p}, U_D)$ is $\frac{1}{D}Ap$. To argue that this distribution has sufficient entropy, we can show that the l_2 norm of it is small and apply Lemma 6 to complete the proof.

$$\left\| \frac{1}{D} A(\sum_{i=1}^{N} \alpha_i v_i) \right\|_2^2 = \left\| \frac{1}{D} (\sum_{i=1}^{N} \alpha_i \lambda_i v_i) \right\|_2^2 = \frac{1}{D^2} \sum_{i=1}^{N} \alpha_i^2 \lambda_i^2$$

Notice that $\overrightarrow{p}$ is represented by an orthonormal basis $\{v_i\}$, v_i are just the coordinate of $\overrightarrow{p}$. Hence, we have

$$\|\overrightarrow{p}\|_2^2 = \sum_{i=1}^{N} \alpha_i^2 = \frac{1}{|S|}$$

Substitute it into above formula, we have

$$\begin{aligned} \frac{1}{D^2} \left(\sum_{i=1}^{N} \alpha_i^2 \lambda_i^2 \right) &\leq \alpha_1^2 + \frac{\lambda^2}{D^2} \sum_{i=2}^{N} \alpha_i^2 = \frac{1}{N} + \frac{\lambda^2}{D^2} \left(\frac{1}{|S|} - \frac{1}{N} \right) \\ &\leq \frac{1}{N} + \frac{\lambda^2}{D^2 |S|} \\ &\leq \frac{2\lambda^2}{D^2 |S|} \end{aligned}$$

(where the last inequality used the fact that $\lambda \geq \sqrt{D}$ and that $|S|D \leq N$).

Now by applying Lemma 6, we get:

$$H_\infty(p) \geq \log \frac{|S|D^2}{2\lambda^2} - \log \frac{1}{\varepsilon} = k + 2 \log \frac{D}{\lambda} - \log \frac{1}{\varepsilon} - 1$$

Therefore, it follows that the λ-eigenvalue expander gives an $\left(n - d, 2 \log \frac{D}{\lambda} - \log \frac{1}{\varepsilon} - 1, \varepsilon \right)$ conductor. □

2.1 The lossiness of the zig-zag product

Given a large expander G with degree $d(G)$, and a smaller expander H with $|V(H)| = d(G)$ and degree $d(H)$, the zig-zag product of G and H gives a new expander with $|V(G)| \cdot |V(H)|$ vertices and degree $d(H)^2$. The zig-zag product produces another constant degree expander. Is this a lossless expander?

Let us start by translating the zig-zag product in the language of conductors, and let us observe the flow of randomness in this language.

Let $v = (g, h)$ denote a vertex in the new expander, which corresponds to $g \in V(G)$ and $h \in V(H)$. Given two random strings with range $[1, d(H)]$, we start from v and then reach a new vertex by

1. use the first random string to go from $v = (g, h)$ to (g, h');
2. (g, h') is equivalent to an edge in G which is incident to g; across this edge, go from (g, h') to (g', h'');

3. use the second random string to go from (g', h'') to (g', h''').

Suppose $v = (g, h)$ is chosen from a random source . Which of the steps below are responsible for (g', h'') having more entropy than (g, h)?

$$(g, h) \longrightarrow (g, h') \longrightarrow (g', h'') \longrightarrow (g', h''').$$

The second step is simply a permutation, and so it can only be the first or third step which increases entropy.

If h is far from uniform (low entropy), the first step will add entropy into h'. However, if h is almost uniform, the first step becomes useless and only the third step adds entropy; it is not lossless. As we will see in the next section, in order to get a construction of expanders with better expansion, we replace the first random step with a "buffer conductor", which retains the entropy which would have lost originally. The third step will be replaced by a lossless conductor that will carry the buffer entropy and the entropy from the second random string to output.

A buffer conductor is an extracting conductor equipped with a buffer that collects the entropy that the extracting conductor loses.

Definition 7. *An (a, ε) buffer conductor is a function $(C, B) : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m \times \{0,1\}^b$ such that*

1. *C is an (a, ε) extracting conductor;*
2. *(C, B) is an (n, ε) lossless conductor.*

3 Constant Degree Lossless Expanders

We now construct a constant degree D expander with $(1 - \varepsilon)D$ vertex expansion. Figure 1 gives an overview of the construction.

There are two powerful and memorable ideas that show up in the construction and analysis of these objects.

- **Key Idea 1:** If some amazing constant-sized object is known to exist, we are free to use it without having to find a nice description for it. The reason is that constant sized objects with desired properties can be found in constant time by brute force search, and so we can explicitly construct this object in constant time.

  In our construction, we will freely use buffer conductors and lossless conductors of constant size. Their role in the construction will be as follows: we will first take a conductor given by a good absolute eigenvalue expander; this will not be lossless, but it will lose some constant number of bits of entropy. We will then use constant sized buffer conductors and lossless conductors to "mop up" this leftover entropy.

The second idea will show up in the analysis.

The next lemma shows that various conductors exist (we will use constant-sized versions of these in our construction). The proof is by probabilistic method, showing that random functions are such conductors with high probability.

Lemma 8. *1. A (k, ε) lossless conductor $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ so long as $m > k + d + \log \frac{1}{\varepsilon}$.*

2. An (c, ε) extracting conductor $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ exists so long as $d > c + 2\log \frac{1}{\varepsilon}$.

3. An (c, ε) buffer conductor $\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m \times \{0,1\}^b$ exists so long as $d > c + 2\log \frac{1}{\varepsilon}$ and $m + b > n + d + \log \frac{1}{\varepsilon}$.

We now come to our main theorem.

Theorem 9. *Let $a = 3\log \frac{1}{\varepsilon}$. We can explicitly construct a bipartite graph (L, R) with $|L| = 2^n$, $|R| = 2^{n-a}$ and the degree of each vertex in L is 2^{2a}, such that it is an $(1-\varepsilon)D$ vertex expander for any subset of L with size at most 2^{n-10a}.*

In the language of conductors: there is an explicit $(n - 10a, \varepsilon)$ lossless conductor $F : \{0,1\}^n \times \{0,1\}^{2a} \to \{0,1\}^{n-a}$

The construction is given in Figure 1. Here BC is a $(0, \varepsilon)$ buffer conductor. LC is a $(17a, \varepsilon)$ lossless conductor. These are both of constant size. Since $a = 3\lg \frac{1}{\varepsilon}$, Lemma 8 implies that these objects exist.

PC is a *permutation conductor* created out of the rotation map of a D-regular 2^{n-20a}-vertex absolute eigenvalue expander (where $D = 2^{14a}$): it takes in a vertex name v and an integer e in $[D]$, and it outputs the rotation map $(v', e') = \mathsf{Rot}(v, e)$. Note that this map is a permutation (hence the name). By the analysis in the earlier section, the map sending $(v, e) \mapsto v'$ is a conductor. By using a suitable explicit Ramanujan graph here, we can get it to be a $(n - 30a, 8a, \varepsilon)$ conductor (This is a 2^{14a}-regular graph, so we may take its absolute eigenvalue expansion to be about 2^{7a}, and so we can get it to be a $(n - 14a, 14a - \log \frac{1}{\varepsilon}, \varepsilon)$-conductor, which comfortably accommodates the parameters we want).

3.1 Proof of Theorem 9

Before analyzing this construction, let us make some remarks.

- Constructing a lossless conductor is totally trivial if the length of the output (Y_1, Y_2) is $n + 2a$. The nontriviality of the theorem comes from the fact that the length of (Y_1, Y_2) is $\leq n$.
- Just by entropy-chasing, if we were allowed to make LC output a larger number of bits, it would be obvious that the overall construction gives a lossless conductor. The crux is to show that the net entropy going into LC (conditioned on Y_1) is can fit inside the $19a$ output bits of LC.
- Consider the special case: For every x_1 in the support of X_1, the conditional distribution $X_2|(X_1 = x_1)$ has high entropy. Then the buffer conductor BC will be in extracting mode, and so $R_3|(X_1 = x_1)$ will be a uniformly distributed string in $\{0,1\}^{14a}$. This will make the permutation conduction PC go into conducting mode, and so Y_1 will have noticeably more entropy than X_1. Meanwhile Z_1 and Z_2 have collected all the leftover entropy, and LC gathers it all, along with the entropy of R_2, into Y_2. It is crucial here that Y_2 is long enough to contain all the leftover entropy from Z_1, Z_2, R_2, and that turns out to be true by keeping track of the flow of entropy (we will do this in detail later).

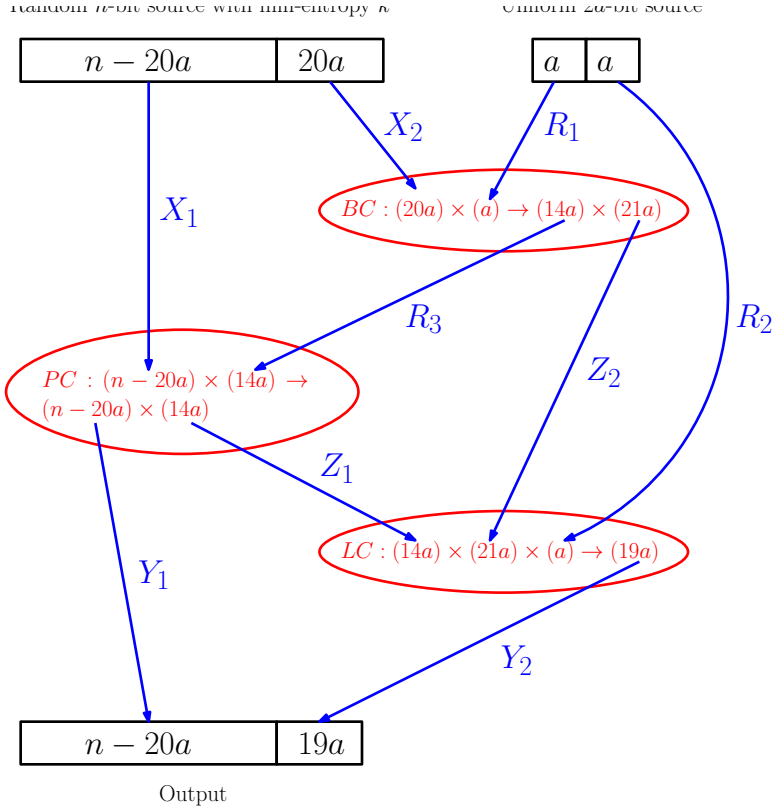

Figure 1: Construction of constant-degree $(1-\varepsilon)d$ vertex expander. $X_1, X_2, \cdots$ beside the arrows denotes the distribution of the input/output. BC is a buffer conductor, PC is a permutation conductor and LC is a lossless conductor.

- Consider another special case: For every x_1 in the support of X_1, the conditional distribution $X_2|(X_1 = x_1)$ has low entropy. Thus X_1 will have high entropy. Now the buffer conductor will not be in extracting mode, and all we can say is that (X_1, R_3) has high entropy, and so Y_1 will have a good amount of entropy. This will imply that there is not too much leftover entropy in Z_1, Z_2, R_2, and so it can all be pushed into Y_2 by LC.
- **Key Idea 2: The Amazing Aspect of this Proof.** Our overall argument will show that we can always express the distribution of (X_1, X_2) as a convex combination of the above two cases, *except for a small loss in entropy*. This seems terrible, because we are trying to show that this is a LOSSLESS conductor, which means there can be absolutely no losses in entropy. But here is the magic: will only use the lossy convex combination decomposition to show that Y_1 has high entropy. Once we know that Y_1 has high entropy, we get losslessness by a completely different argument. Since Y_1 has high entropy, $(Z_1, Z_2, R_2)|(Y_1 = y_1)$ has low

entropy[1]. Thus all the entropy of $(Z_1, Z_2, R_2)|(Y_1 = y_1)$ can be packed into Y_2, and we get losslessness.

We start by giving two technical lemmas.

Lemma 10. *If $H_\infty(W_1, W_2) \geq q$ and $H_\infty(W_2|W_1 = w_1) \leq q_2$ for all $w_1 \in$ Support(W_1), then $H_\infty(W_1) \geq q - q_2$.*

Proof. For any $w_1 \in \text{Support}(W_1)$, $H_\infty(W_2|W_1 = w_1) \leq q_2$ implies that for some w_2, $\Pr[W_2 = w_2|W_1 = w_1] \geq 2^{-q_2}$. Hence $\Pr[W_2 = w_2, W_1 = w_1] \geq 2^{-q_2}\Pr[W_1 = w_1]$.

However, $H_\infty(W_1, W_2) \geq q$ implies that $\Pr[W_2 = w_2, W_1 = w_1] \leq 2^{-q}$. Hence $2^{-q} \geq 2^{-q_2}\Pr[W_1 = w_1]$, i.e. $\Pr[W_1 = w_1] \leq 2^{-(q-q_2)}$. Then $H_\infty(W_1) \geq q - q_2$. □

Lemma 11. *Suppose $X = (X_1, X_2)$ is a distribution with $H_\infty(X) = k$. For every b, ε, X is ε-close to a convex combination of distributions $X' = (X'_1, X'_2)$ and $X'' = (X''_1, X''_2)$ such that:*

1. *$H_\infty(X'), H_\infty(X'') \geq k - \log\frac{1}{\varepsilon}$.*
2. *For each x_1 in the support of X'_1, $H_\infty(X'_2|X'_1 = x_1) \geq b$.*
3. *For each x_1 in the support of X''_1, $H_\infty(X''_2|X''_1 = x_1) < b$.*

Proof. Let $S = \{x_1 \mid H_\infty(X_2|X_1 = x_1) \geq b\}$. Let $X' = (X \mid X_1 \in S)$ and $X'' = (X \mid X_1 \notin S)$, and note that X is a convex combination $pX' + (1-p)X''$ for some p, and X', X'' have disjoint supports.

If $p, (1-p)$ are both $\geq \varepsilon$, then X' and X'' both have the desired min-entropy, and we are done. Otherwise, if $p < \varepsilon$, we note that X is ε-close to X'' (Which has the desired min-entropy), and if $1 - p < \varepsilon$, then X is ε-close to X' (which has the desired min-entropy). □

Let $k \leq n - 10a$. Start with a distribution $X = (X_1, X_2)$ with min-entropy at least k. We want to show that the output $F(X, U_{2a})$ is $O(\varepsilon)$-close to having min-entropy at least $k + 2a$. Throughout we will ignore ε statistical distance and pretend that the output random variables themselves have high entropy (when we actually mean that some random variable statistically close to them has high entropy).

The key step in proof of Theorem 9 is to show that for every input distribution X, the entropy of Y_1 is high.

Lemma 12. *If $H_\infty(X_1, X_2) \geq k$, then $H_\infty(Y_1) \geq k - 15a$.*

Proof. By the previous lemma, we may assume that we are in one of the following two cases.

Case 1. (X_1, X_2) is such that (1) $H_\infty(X_1, X_2) \geq k - \log\frac{1}{\varepsilon}$ (2) for all $x_1 \in \text{Support}(X_1)$, $H_\infty(X_2|X_1 = x_1) \geq 14a$.

By the extracting property of BC, $(R_3|X_1 = x_1)$ is ε-close to U_{14a}. Hence, the joint distribution (X_1, R_3) is ε-close to (X_1, U_{14a}), where X_1 and U_{14a} are independent. Thus PC is in conducting mode, and we will be able to conclude that Y_1 has high min-entropy.

[1]This is not quite formal, the actual proof will do it correctly.

By Lemma 10, $H_\infty(X_1) \geq H_\infty(X_1, X_2) - 20a \geq k - 21a$. Since $k - 21a \leq n - 30a$, we have:

$$H_\infty(Y_1) \geq k - 21a + 8a = k - 13a.$$

Case 2. (X_1, X_2) is such that (1) $H_\infty(X_1, X_2) \geq k - \log\frac{1}{\varepsilon}$ (2) for all $x_1 \in \text{Support}(X_1)$, $H_\infty(X_2|X_1 = x_1) < 14a$.

Since $H_\infty(X_1, X_2) \geq k - a$, we have that for all $x_1 \in \text{Support}(X_1)$,

$$H_\infty(X_2|X_1 = x_1)) + \log\frac{1}{\Pr[X_1 = x_1]} \geq k - a.$$

By the conductor property of BC, along with our hypothesis on $H_\infty((\,)X_2|X_1 = x_1)$, we have that $H_\infty(R_3|X_1 = x_1) \geq H_\infty(X_2|X_1 = x_1)$. Hence,

$$H_\infty(X_1, R_3) = \min_{x_1}\left(H_\infty(R_3|X_1 = x_1) + \lg\frac{1}{\Pr[X_1 = x_1]}\right) \geq k - a.$$

As PC is a permutation, $H_\infty(Y_1, Z_1) \geq k - a$. Trivially $H_\infty(Z_1|Y_1 = y_1) \leq 14a$ for every y_1. Thus Lemma 10 tells us that:

$$H_\infty(Y_1) \geq k - 15a.$$

□

We now argue losslessness of F.

Lemma 13. *$H_\infty(Y_1, Y_2) \geq k + 2a$.*

Proof. Note that

$$k + a \leq H_\infty(X_1, X_2, R_1) = H_\infty(X_1, R_3, Z_2) = H_\infty(Y_1, Z_1, Z_2), \tag{1}$$

where the first equality is by the buffer property of BC, and the second equality is by the permuation property of PC.

Now by the losslessness property of LC, $H_\infty(Y_2|Y_1 = y_1) \geq \min\{H_\infty(Z_1, Z_2|Y_1 = y_1) + a, 17a\} =: h$, i.e.
$\Pr[Y_2 = y_2|Y_1 = y_1] \leq 2^{-h}$. Hence $\Pr[Y_2 = y_2, Y_1 = y_1] \leq 2^{-h}\Pr[Y_1 = y_1]$.

If $h = 17a$, note that by Lemma 12, $\Pr[Y_1 = y_1] \leq 2^{-(k-15a)}$, hence $\Pr[Y_2 = y_2, Y_1 = y_1] \leq 2^{-(k+2a)}$.

If $h = H_\infty(Z_1, Z_2|Y_1 = y_1) + a$, then

$$\begin{aligned}
2^{-h} = 2^{-a}2^{-H_\infty(Z_1,Z_2|Y_1=y_1)} &= 2^{-a}\max_{z_1,z_2}\Pr[Z_1 = z_1, Z_2 = z_2|Y_1 = y_1] \\
&= 2^{-a}\max_{z_1,z_2}\frac{\Pr[Z_1 = z_1, Z_2 = z_2, Y_1 = y_1]}{\Pr[Y_1 = y_1]} \\
&\leq 2^{-a}\frac{2^{-(k+a)}}{\Pr[Y_1 = y_1]}. \qquad \text{by (1)}
\end{aligned}$$

Hence $\Pr[Y_2 = y_2, Y_1 = y_1] \leq 2^{-h}\Pr[Y_1 = y_1] \leq 2^{-(k+2a)}$. □