Perturbation Codes

Adi Akavia and Ramarathnam Venkatesan

***Abstract*— We present a new family of codes with good asymptotic properties. These codes are constructed from simple old codes using a new *perturbation operator* that we introduce. We provide an error reduction algorithm for these codes that uses only elementary operations with small precision. We also present a soft error reduction algorithm for the expander based codes of Alon-Bruck-Naor-Naor-Roth when concatenated with any binary code.**

I. Introduction

We present a new family of codes achieving good asymptotic properties while admitting an error reduction algorithm that uses only elementary operations with small precision. This is done by introducing a new *perturbation operator* with which we construct new codes from old codes. In particular we apply the perturbation operator to combine random linear codes together with the expander based codes of Alon-Bruck-Naor-Naor-Roth (ABNNR) [3] achieving a new code that maintains both the good distance of the random linear codes as well as the error-reduction capabilities of the ABNNR codes.

ABNNR codes [3] are expander based codes enabling *distance amplification* from any constant normalized distance to normalized distance approaching 1. This is achieved by composing an appropriate ABNNR code to a code C_0 of constant normalized distance resulting in a new code C' of normalized distance $1 - \Theta(1)$, the same encoding length as C_0, and alphabet size $\Theta(q)$ for q the alphabet size of C_0. The resulting code C' can be then transformed into a *binary* code achieving distance approaching $\frac{1}{2}$ and constant rate via the standard concatenation technique.

Guruswami and Indyk [12] presented an *error reduction algorithm* for ABNNR codes concatenated with binary codes. Their algorithm follows Forney's Generalized Minimum Distance (GMD) methodology [8] for decoding concatenated codes requiring an efficient decoding algorithm for the inner code, and an algorithm for decoding from erasures and errors for the outer code. (Here, following standard terminology, we refer to the non-binary code as the *outer code* and to the binary code concatenated with it as the *inner code*.)

A. Results

1) New codes: We present a new family of binary codes achieving normalized distance approaching $\frac{1}{2}$, and constant rate. Furthermore, we present an error reduction algorithm for these codes. That is, an algorithm that, given a codeword corrupted by adversarial noise flipping a constant fraction of its bits, outputs a message that agrees with the original message on at least $(1 - \varepsilon)$-fraction of its bits, for ε an arbitrarily small constant. (Specifically, the normalized distance may be $\frac{1}{2} - c$ for arbitrarily small constant c, and the rate is a constant depending on both c and ε.)

2) New code operator: We present a new code operator – the *perturbation operator*– that given any binary linear code R, outputs a new binary linear code $Pert(R)$ such that $Pert(R)$ preserves the distance of R, preserves the rate of R up to a constant factor, and furthermore, $Pert(R)$ admits an error reduction algorithm even if R does not. The codes above are obtained by taking R to be random linear code. The perturbation operator can be viewed as combining the given code R with a variant of the ABNNR code.

3) Soft error reduction algorithm for ABNNR codes when concatenated with any binary code: We present soft error reduction algorithm for the ABNNR codes [3] when concatenated with any binary code. Our algorithm achieves similar performance to the previously known algorithm of Guruswami-Indyk [12], albeit, in a different way. One particular difference is that we make no use of Forney's GMD decoding approach, instead, we take a soft decoding approach that does not require an erasures-and-errors decoding algorithm for the outer code. Our algorithm is easy to implement both in software as well as in hardware: it is *local*, *i.e.*, each message bit is found with $O(1)$ queries to the codewords, it can be implemented in a constant size (*i.e.*, NC^0) circuit, and it executes only bit-wise operations. (This is in contrast to arithmetic in finite fields of high characteristic that is needed in decoding algebraic codes such as the Reed-Solomon codes).

We remark that improvements of our algorithm to handle expander graphs with poly-logarithmic degrees were recently developed and employed by Gopalan-Guruswami [11] in the context of proving hardness amplification results within the class NP. The two properties of our algorithm —(1) avoiding the use of GMD and (2) being local— are crucial for the applications of [11].

B. Paper Organization

In section II we give some preliminary definitions and notations. In section III we formally define the perturbation operator. In section IV we analyze the rate and distance of perturbation codes. In section V we present the error reduction algorithm for the perturbation codes. In section VI

This work was not supported in part by Microsoft Research fellowship, NSF Grant CCF0514167, IAS/DIMACS post-doctoral fellowship.

A. Akavia is with the School of Mathematics, Institute of Advanced Study, Princeton NJ, USA akavia@ias.edu

R. Venkatesan is with Microsoft Research, Redmond WA, USA and Bangalore India. venkie@microsoft.com

we present the soft error reduction algorithm for ABNNR codes concatenated with binary codes.

II. Preliminaries

We present some preliminary definitions and notations.

We use the notation $[n] \stackrel{def}{=} \{1, \ldots, n\}$ for every positive integer n.

A. Error correcting codes.

A binary *error correcting code* is an infinite family $\mathcal{C} = \{\mathcal{C}_k\}_{k=1}^{\infty}$ of sets $\mathcal{C}_k$ of codewords $C(m) \in \{0,1\}^n$ encoding messages $m \in \{0,1\}^k$ by length n binary vectors; k is the *information rate*, and n is the *block length*. The *rate* of the code is the ratio k/n. The *normalized distance* of two codewords $C(m), C(m') \in \{0,1\}^n$ is the fraction of bits where they differ $\delta(C(m), C(m')) = \frac{1}{m} |\{ i \,|\, C(m)_i \neq C(m')_i \}|$; the *distance* of $\mathcal{C}_k$ is the minimum normalized distance over all pairs of codewords $\delta(\mathcal{C}_k) = \min_{C(m),C(m') \in \mathcal{C}_k} \delta(C(m), C(m'))$; the distance of $\mathcal{C}$ is $\delta(\mathcal{C}) = \inf_{k \to \infty} \delta(\mathcal{C}_k)$. *Encoding* is the process of mapping messages into codewords. For *linear codes* $\mathcal{C}_k$, codewords $C(m) = mG$ are obtained by multiplying the message $m \in \{0,1\}^k$ by a *generating matrix* $G \in \{0,1\}^{k \times n}$. In the *adversarial noise model of parameter* $p \in [0,1]$ codewords are corrupted by noise flipping up to p-fraction of their bits. *Decoding* is the process of mapping corrupted codewords into the messages corresponding to the original uncorrupted codeword.

B. Small biased sets.

We say that $S_m \subseteq \{0,1\}^m$ is a *γ-biased set* if $\forall x \in \{0,1\}^m$, $\frac{1}{|S_m|} \left| \sum_{\alpha \in S_m} \chi_\alpha(x) \right| \leq \gamma$ where $\chi_\alpha(x) = (-1)^{\langle \alpha, x \rangle}$. A random set S_m of size $|S_m| = O(m/\gamma^2)$ is γ-biased with very high probability. Explicit constructions of γ-biased sets of size $O(m^{\Theta(1)}/\gamma^{\Theta(1)})$ are achieved in [17], [4], [15], [1], [18], [9], [5].

C. Chernoff/Hoeffding bound [14].

Let $X_1, \ldots, X_t$ be independent random variables of expectations $\mu_1, \ldots, \mu_y$ and bounded values $|X_i| \leq M$. Then, for any η, $\Pr[\frac{1}{t} \sum_{i=1}^{t} X_i - \frac{1}{t} \sum_{i=1}^{t} \mu_i \geq \eta] \leq 2 \cdot exp\left(-\frac{2t\eta^2}{M^2}\right)$

D. Expander graphs.

Let $H = ([k], [b], E)$ be a bipartite graph. We say that H is (d_L, d_R)-regular if its left degrees are d_L and its right degrees are d_R. We say that H is λ-expander if it satisfies *expander mixing lemma*, that is, if for every $A \subseteq [k]$ and $B \subseteq [b]$,

$$\left| |E(A,B)| - \frac{d_L |A| |B|}{b} \right| < \lambda d_L \sqrt{|A| |B|}$$

We remark that the expander mixing lemma is usually derived from more standard definitions of expander graph, rather than being given as a definition.

E. ABNNR codes concatenated with binary codes.

We ABNNR codes and their concatenation with binary codes. Let $\mathcal{H} = \{([n], [n], E_n)\}_{n \in \mathbb{N}}$ be a family of d-regular λ-expander graphs, and let $\mathcal{C}_0$ be a family of binary codes encoding d bits into $O(d)$ bits. Let $m = m_1 \ldots m_n$ be a message in $\{0,1\}^n$.

- The *ABNNR code* $\mathcal{C}_{\mathcal{H}}$ encodes m by the codeword $C_{H_n}(m) \in (\{0,1\}^d)^n$ whose i-th symbol is $(m_{i_1}, \ldots, m_{i_d})$ for $i_1, \ldots, i_d$ the left neighbors of right node i in the graph H_n.[1]
- The concatenation $\mathcal{C}(\mathcal{H}, \mathcal{C}_0)$ of an ABNNR code $\mathcal{C}_{\mathcal{H}}$ with binary code $\mathcal{C}_0$ ("binary ABNNR code", in short) is a code in which the codeword encoding message m is the vector $C(m) = (\mathcal{C}_0(C_{H_n}(m)_1), \ldots, \mathcal{C}_0(C_{H_n}(m)_n) \in (\{0,1\}^{O(d)})^n$.

We use the terminology "*the i-th block of* $C(m)$" to refer to bits $\mathcal{C}_0(C_{H_n}(m)_i)$ of $C(m)$. We denote the restriction of $C(m)$ to its i-th block by $C(m)^i$. Likewise, for a corrupted codeword w, we denote by w^i the restriction of w to the bits corresponding to the i-th codeword block.

III. The Perturbation Operator

In this section we present the perturbation operator.

The perturbation operator, given a generating matrix $R \in \{0,1\}^{k \times n}$ of any binary linear code R, outputs a generating matrix $Pert(R)$ of a new binary linear code $Pert(R)$. The columns of the generating matrix $Pert(R)$ are sums (modulo 2) $r \oplus \ell$ of columns r of R and vectors ℓ from low dimensional subspaces of $\{0,1\}^k$. We think of the sums $r \oplus \ell$ as a perturbation of the column r by ℓ. Specifically, the columns of $Pert(R)$ are all sums $r \oplus \ell$ for $r \in R_i$, $\ell \in L_i$, $i = 1, \ldots, \Theta(n)$ for R_i and L_i chosen as follows.

The R_i's are random subsets of the columns of R of size $s = \Theta(1)$ that are chosen uniformly at random (with repetitions).

The L_i's consist of all columns of a generating matrix of (unbalanced variant of) ABNNR codes concatenated with a binary code, where each L_i consist of the columns corresponding to the i-th symbol of the ABNNR code.

Definition 1 (Perturbation Operator): For any information rate k and any perturbation operator parameters $(b, d, s, \gamma, \lambda)$, the perturbation operator is defined by a bipartite $(bd/k, d)$-regular λ-expander graph $G = ([k], [b], E)$, and γ-biased sets $L_i \subseteq span\{ e_j \,|\, (j,i) \in E \}$ (for e_j the standard basis element with 1 at coordinate j and zero elsewhere) as follows.

For any binary linear code with generating matrix $R \in \{0,1\}^{k \times n}$, the generating matrix of $Pert(R)$ is composed of b sub-matrices $P_i \in \{0,1\}^{k \times |R_i||L_i|}$, for $i = 1, \ldots, b$, such that the columns of each P_i are

$$P_i\text{'s columns} = \{ r \oplus \ell \,|\, r \text{ a column of } R_i \text{ and } \ell \in L_i \}$$

for $R_1, \ldots, R_b$ subsets of the columns of R, each of size s, chosen uniformly at random (with repetitions).

[1] The ABNNR code applies to any alphabet Σ. We focus on binary alphabet for simplicity.

We remark that the random choices in the definition of the perturbation operator can be derandomized.

Notations. In the following we fix notations $k, b, d, s, \gamma, \lambda, \{L_i\}_{i=1}^b$, $R, Pert(R)$ and $\{P_i\}_{i=1}^b$ to be as in the above definition. For all $i \in [b]$, we let

$$V_i = span\{ e_j \mid (j,i) \in E\}$$

be the subspaces of $\{0,1\}^k$ defined by the structure of the graph G. We denote by P the generating generating matrix of $Pert(R)$; and by $P(m)$ (or mP) the encoding of $m \in \{0,1\}^k$ by $Pert(R)$. We index coordinates of codewords $P(m)$ and corrupted codewords w by pairs r, ℓ for $r \in R_i$, $\ell \in L_i$, $i \in [b]$ where

$$P(m)_{r,\ell} = \langle m, r \oplus \ell \rangle$$

IV. Perturbation Codes: Rate & Distance

In this section we analyze the rate and distance of perturbation codes $Pert(R)$ showing how they relate to the rate ρ and normalized distance δ of the given code R. In particular, we show that $Pert(R)$ has rate $\Theta(\rho)$, and with probability at least half its normalized distance is at least $\delta - \Theta(1)$ for any $\zeta > 0$ provided that the perturbation parameters $d, s, \gamma = \Theta(1)$ and $b = \Theta(k)$ are sufficiently large. We remark that the success probability half can be amplified by increasing s.

Theorem 2: Let R be a binary linear code of rate ρ and normalized distance δ, the following holds:

1) $Pert(R)$ has rate $\rho \cdot \Theta(\gamma^2 / s \log d)$
2) For any $\zeta \in (0, \delta)$, with probability at least half, $Pert(R)$ has normalized distance at least $\delta - \zeta$, as long as the perturbation parameters $s = \Theta(1/\zeta^2)$ and $b = \Theta(k/\zeta^2)$ are sufficiently large (where the probability is taken over the choice of R_i's).

Proof: Denote by $t = \Theta(\log d / \gamma^2)$ the size of each of the γ-biased sets L_i. For a message $m \in \{0,1\}^k$, the length of the codeword $P(m)$ is $\sum_{i=1}^b |R_i|\,|L_i| = bst$. So the rate of the perturbation code is $k/bst = \rho \cdot \Theta(\gamma^2/s \log d)$.

We next show that with probability at least $\frac{1}{2}$, the normalized distance of $Pert(R)$ is at least $\delta - \zeta$. The normalized distance of $Pert(R)$ (as well as any linear code) is equal to the minimum normalized weight of its non-zero codeword, where the *normalized weight* of a codeword xP is the fraction of non-zero entries in the vector xP. Thus, it suffices to prove that the minimum normalized weight of the Perturbation code is at least $\delta - \zeta$ with probability of at least $\frac{1}{2}$ (where the probability is taken over the random choices of the perturbation operator).

Fix some non-zero $x \in \{0,1\}^k$. Observe that the normalized weight of xP is equal to $\frac{1}{stb}\|xP\|_2^2 = \frac{1}{stb}\sum_{i=1}^b \|xP_i\|_2^2$. In Claim 2.1 below we show that, with probability at least $1 - \frac{1}{2^{k+1}}$, $\frac{1}{st}\|xP_i\|_2^2 \in (\delta \pm \frac{2}{3}\zeta)$ for at least $(1 - \frac{2}{3}\zeta)b$ of the P_i's where $i = 1, \ldots, b$. This implies that the normalized weight of xP is at least $(\delta - \frac{2}{3}\zeta)\cdot(1 - \frac{2}{3}\zeta) + 0 \cdot \frac{2}{3}\zeta \geq \delta - \zeta$. By union bound this holds for all non-zero $x \in \{0,1\}^k$ with probability at least $1 - 2^k/2^{k+1} = \frac{1}{2}$. Thus, we conclude that the minimum distance of the code P is at least $\delta - \zeta$ with probability at least half.

Claim 2.1: For $b = \Theta(k/\zeta^2)$ sufficiently large, for each non-zero $x \in \{0,1\}^k$, with probability at least $1 - \frac{1}{2^{k+1}}$ over the choice of R_i's, $\frac{1}{st}\|xP_i\|_2^2 \in (\delta \pm \frac{2}{3}\zeta)$ for at least $(1 - \frac{2}{3}\zeta)b$ of the P_i's where $i = 1, \ldots, b$.

Proof: Fix a non-zero $x \in \{0,1\}^k$. For each $i = 1, \ldots, b$, we say that x is *good with respect to* R_i if the fraction of 1's in the set $\{ \langle x, r \rangle \mid r \in R_i \}$ is within $(\delta \pm \frac{2}{3}\zeta)$.

We first fix i and show that if x is good with respect to R_i, then $\frac{1}{st}\|xP_i\|_2^2 \in (\delta \pm \frac{2}{3}\zeta)$. Observe that $\frac{1}{st}\|xP_i\|_2^2$ is equal to the fraction of 1's in the set $\{ \langle x, r \oplus \ell \rangle \mid r \in R_i, \ell \in L_i \}$. For each *fixed* $\ell \in \{0,1\}^k$ the fraction of 1's in the set $\{ \langle x, r \oplus \ell \rangle \mid r \in R^i \}$ is within $(\delta \pm \frac{2}{3}\zeta)$ (because $\langle x, r \oplus \ell \rangle = \langle x, r \rangle \oplus \langle x, \ell \rangle$, where $\langle x, \ell \rangle \in \{0,1\}$ is constant when varying only over r, and because x is good w.r. to R_i). Since this holds for every ℓ, then it holds also when varying over all $\ell \in L_i$.

We next show that x is good with respect to R_i with probability at least $1 - \zeta/3$. Denote by α_i the fraction of 1's in the set $\{ \langle x, r \rangle \mid r \in R_i \}$. Since R has normalized distance δ and the vectors in R_i are chosen uniformly and independently at random, then $E[\alpha_i] = \delta$, and by Hoeffding bound, $\Pr[|\alpha_i - \delta| \geq \frac{2}{3}\zeta] \leq 2\exp(-2s(\frac{2}{3}\zeta)^2) \leq \zeta/3$ where the last equality hold for our choice of $\zeta = \Theta(\sqrt{s})$.

Finally, we apply Hoeffding bound to conclude that, with probability at least $1 - \frac{1}{2^{k+1}}$, x is good with respect to R_i for at least $1 - \frac{2}{3}\zeta$ of the $i = 1, \ldots, b$. Denote by $I_1, \ldots, I_b$ indicator random variables such that $I_i = 1$ if x is *not* good with respect to R_i, and $I_i = 0$ otherwise (where $i = 1, \ldots, b$). Observe that $I_1, \ldots, I_b$ are independent random variables, and by the above, $I_i = 1$ with probability at most $\zeta/3$. By Hoeffding bound, $\Pr[\left|\frac{1}{b}\sum_{i=1}^b I_i - \frac{\zeta}{3}\right| \geq \frac{\zeta}{3}] \leq 2\exp(-2b(\zeta/3)^2) \leq \frac{1}{2^{k+1}}$ for $b = \Theta(sk)$ sufficiently large. ∎ ∎

V. Perturbation Codes: Error Reduction

In this section we present our error reduction algorithm for perturbations codes $Pert(R)$. The error reduction algorithm, given a binary vector w which is at normalized distance at most $1/64 - \Theta(1)$ from $P(m)$ (for $P(m)$ the codeword of $Pert(R)$ encoding the message $m \in \{0,1\}^k$), outputs $m' \in \{0,1\}^k$ which is at normalized Hamming distance to $(1-c)$ from m for $c = \Theta(1)$ an arbitrarily small constant depending on the parameters of the perturbation operator.

A. The Algorithm

The algorithm is composed of two parts. First we reduce the problem of error reduction perturbation codes to the problem of error reduction in another code, specifically, the ABNNR code used in the perturbation operator when concatenated with some binary code. Then, we use known techniques [12] to reduce errors in the latter code.

1) Computing new inputs: The idea in this part of the algorithm is to map the input w to an input w' which is a corrupted codeword of another code: the ABNNR code corresponding to the graph G from the definition of the

perturbation operator when concatenated with some good binary code.

The vector w' is obtained by xor-ing appropriate entries of w. Specifically, we choose for each L_i a uniformly random matching L_i^{pairs} of the $\ell \in L_i$ into disjoint pairs $\{\ell_1, \ell_2\}$, and a uniformly random $r \in R_i$, and define for each pair $\{\ell_1, \ell_2\} \in L_i^{pairs}$,

$$w'_{\ell_1,\ell_2} = w_{r,\ell_1} \oplus w_{r,\ell_2}$$

We then for each $i \in [b]$ find $\alpha^i(w') \in V_i$ with highest agreement with w', where the agreement of α^i and w' is

$$agree_{i,\alpha^i} = \left\{ \{\ell_1, \ell_2\} \in L_i^{pairs} \;\middle|\; w'_{\ell_1,\ell_2} = \langle \alpha^i, \ell_1 \oplus \ell_2 \rangle \right\}.$$

To increase accuracy we repeat the above procedure $T = O(\log b)$ times, and set for each $i \in [b]$, $\alpha^i \in V_i$ to be the most frequent value out of all repetitions. That is, denoting by $w^1, \ldots, w^T$ the w''s from the various repetitions, we have

$$\alpha^i = \text{most frequent value in } \left\{\alpha^i(w^1), \ldots, \alpha^i(w^T)\right\}.$$

2) Error reduction: We apply an error reduction algorithm on the computed values α^i. For each left node $\ell \in [k]$ in the expander graph G from the definition of perturbation codes, we set the ℓ-th bit m'_ℓ by majority vote over all neighbors i of ℓ: For each neighbor i of ℓ and bit value $v \in \{0,1\}$, we say that α^i *assigns value* v *to the* ℓ*-th bit* if $\langle \alpha^i, e_\ell \rangle = v$. We set m'_ℓ to be v if in at least half of the neighbors i of ℓ, α^i assigns v to the ℓ-th bit (breaking ties arbitrarily).

The outputs of the algorithm is $m' = m'_1 ... m'_k$.

B. Analysis

We give a sketch of the analysis of the algorithm showing that for $(1 - \Theta(1))$ fraction of the $\ell \in [k]$, m' has the correct value $m'_\ell = m_\ell$. We remark that the parameters can be optimized using a tighter proof.

Denote by $\rho_{i,\alpha} = \Delta_{|R_i \times L_i}(w, P(\alpha))$ the distance of w from $P(\alpha)$ when restricting the view to entries $(r, \ell) \in R_i \times L_i$. Denote $\rho_i = \rho_{i,m}$. Denote by $disagree_{r,L_i^{pairs},i,m}$ the number of entries $\{\ell, \ell'\} \in L_i^{pairs}$ s.t. $w'_{\ell,\ell'}$ disagrees with $\langle m, \ell + \ell' \rangle$, *i.e.*, $disagree_{r,L_i^{pairs},i,m} = \left| \left\{ \{\ell, \ell'\} \in L_i^{pairs} \;\middle|\; w_{r,\ell} \oplus w_{4,\ell'} \neq \langle m, \ell + \ell' \rangle \right\} \right|$.

We first show that if $\Delta(w, P(m)) < n'/64$, then for at least $(\frac{1}{2} + \Theta(1))$ of the $i \in [b]$, for at least $(\frac{1}{2} + \Theta(1))$ of the $r \in R_i$, it holds that for all L_i^{pairs}, $disagree_{r,L_i^{pairs},i,m} < 1/4 - \Theta(1)$.

Claim 3: If $\Delta(w, P(m)) < n'/64$, then the following holds. There exists a subset $I \subseteq [b]$ of size $|I| \geq (\frac{1}{2} + \Theta(1))b$ s.t. $\forall i \in I$ there exists a subset $R_i^{good} \subseteq R_i$ of size $\left|R_i^{good}\right| \geq (\frac{1}{2} + \Theta(1))|R_i|$ s.t. for all $r \in R_i^{good}$ and for all partitions L_i^{pairs} of L_i into disjoint pairs,

$$disagree_{r,L_i^{pairs},i,m} < 1/4 - \Theta(1).$$

Proof: Let w be close to $P(m)$ in Hamming distance: $\Delta(w, P(m)) < \varepsilon n'$ for $n' = O(bsd)$ the codeword length.

By an averaging argument, this implies that for at least $(\frac{1}{2} + \Theta(1))$ of the $i \in b$,

$$\rho_i < 2\varepsilon |L_i \times R_i|.$$

By definition of ρ_i, for a random $r \in R_i$, the number of entries $(r, \ell) \in \{r\} \times L_i$ that disagree with $\langle m, r + \ell \rangle$ is expected to be ρ_i. By averaging argument, for at least $(\frac{1}{2} + \Theta(1))$ of the r's, the number of disagreements is at most

$$\rho'_i = 2 \cdot 2\varepsilon |L_i| = 4\varepsilon |L_i|.$$

For all such r's, for any L_i^{pairs}, the number of entries $w'_{\ell,\ell'}$ which disagree with $\langle m, \ell + \ell' \rangle$ is at most

$$disagree_{r,L_i^{pairs},i,m} \leq 2\rho'_i \leq 16\varepsilon \left|L_i^{pairs}\right|$$

(because $w'_{\ell,\ell'} \neq \langle m, \ell + \ell' \rangle$ iff exactly one of the inequalities holds: $w_{r,\ell} \neq \langle m, r + \ell \rangle$ or $w_{r,\ell'} \neq \langle m, r + \ell' \rangle$).

Setting $\varepsilon \leq \frac{1}{4}\frac{1}{16} - \Theta(1))$, we get that for at least $(\frac{1}{2} + \Theta(1))$ of the $i \in [b]$, for at least $(\frac{1}{2} + \Theta(1))$ of the $r \in R_i$, it holds that

$$disagree_{r,L_i^{pairs},i,m} < 1/4 - \Theta(1).$$

■

Next, we show that with high probability, α^i is equal to the restriction $m_{|V_i}$ of m to V_i for all $i \in I$; where we say that α^i is equal to the restriction of m to V_i if $\langle \alpha^i, e_\ell \rangle = \langle m, e_\ell \rangle$ for every ℓ left neighbor of right node i in the expander graph G from the definition of the perturbation operator.

Claim 4: For $(\frac{1}{2} + \Theta(1))$ of the $i \in [b]$, $\alpha^i = m_{|V_i}$.

Proof: Fix some $i \in I$. We first show that $\alpha^i(w') = m_{|V_i}$ with probability $(\frac{1}{2} + \Theta(1))$ over the choice of r and L_i^{pairs}: By the above claim $disagree_{r,L_i^{pairs},i,m} < 1/4 - \Theta(1)$ for $(\frac{1}{2} + \Theta(1))$ of the r's. Moreover, for such r's, m is unique in satisfying the above, as long as L_i^{pairs} is a small biased set; where the latter occurs with very high probability over the choice of the random matching L_i^{pairs}. So, $\alpha^i(w') = m_{|V_i}$ with probability $(\frac{1}{2} + \Theta(1))$ over the choice of r and L_i^{pairs}.

Now, since α^i was chosen to be the *most frequent* value over all $\alpha^i(w^t)$'s, then by Chernoff bound, $\alpha^i = m_{|V_i}$ with probability $\exp(O(T)) = 1/b^c$.

Last, by union bound, $\alpha^i = m_{|V_i}$ holds *for all* $i \in I$ with probability $1/b^{c-1}$ for C a constant determined by the number T of repetitions. ■

Finally, by expansion arguments, for $(1 - \Theta(1))$ fraction of the left nodes ℓ, the fraction of their neighbors in I and the fraction of their neighbors in $[b] \setminus I$ are roughly $|I|/b$ and $1 - |I|/b$ respectively.[2] Since $|I| = (\frac{1}{2} + \Theta(1))b$, this implies that for $(1 - \Theta(1))$ fraction of the ℓ's, *most* of their neighbors are in I. By the above claim this implies that for $(1 - \Theta(1))$ fraction of the ℓ's, for most of their neighbors i, $\alpha^i = m_{|V^i}$. This in turn imply that the value assigned m'_ℓ is correct.

[2]See details of similar expansion arguments in section VI.

VI. Soft Error Reduction for Concatenated ABNNR Codes

We present a soft error reduction algorithm that, given a corrupted codeword w of a binary ABNNR code, recovers $(1 - O(1))$-fraction of the bits of the encoded message x. The soft error reduction algorithm works by finding independently for each left node ℓ the best assignment to its neighbors $\Gamma(\ell)$, as follows. For each left node ℓ, we restrict our attention to the corrupted codeword block w^i corresponding to right neighbors $i \in \Gamma(\ell)$ of ℓ. We then find the codeword $C(m')$ closest to w^i on those blocks (by exhaustive search). If the closest codeword is sufficiently close, we set the ℓ-th output bit z_ℓ to be the value of the ℓ-th bit of the encoded message m'. We repeat this procedure for each left node $\ell \in [n]$, outputting $z_1, \ldots, z_n$. A pseudo code follows.

Algorithm 5: **Soft Error Reduction Algorithm.**

Input: A description of the code $\mathcal{C}(\mathcal{H}, \mathcal{C}_0)$, a noise parameter ε, and a corrupted codeword $w \in \{0,1\}^{n/r}$ s.t. $\Delta(C(x), w) \leq \frac{1}{4} - \varepsilon$.

Output: $z \in \{0,1\}^n$ s.t. $\Delta(z, x) \leq (\sqrt{2}\lambda)^{1/3}$.

Steps:

1) For each left node $\ell \in [n]$
   a) By exhaustive search, find assignment $y \in \{0,1\}^{d^2}$ to all message bits $i \in \Gamma_L(\Gamma_R(\ell))$ such that y minimizes
   $$dist_\ell(y) = \frac{1}{d} \sum_{i \in \Gamma(\ell)} \Delta(C_0(y)^i, w^i)$$
   b) If $dist_\ell(y) < \frac{1}{4} - \frac{\varepsilon}{4}$, set $z_\ell = y_\ell$
   c) Else set $z_\ell = \perp$
2) Output $z_1 \ldots z_k$

To show that this is an efficient algorithm, we argue that the exhaustive search step can be done in constant running time. This is because there is a $d = O(1)$ number of blocks i neighboring each left node ℓ, and each of these blocks is determined by a $d = O(1)$ number of message bits (specifically, the bits corresponding to left neighbors of right node i). Thus, the number of assignment to be examined in the exhaustive search is $2^{d^2} = O(1)$.

To prove the correctness of this algorithm, we first apply the Expander Mixing Lemma to show that for the correct message x, $dist_\ell(x) < \frac{1}{4} - \frac{\varepsilon}{4}$ for at least $(1 - O(\lambda^{1/3}))$-fraction of the bits $\ell \in [n]$. Second, we show that for any y that disagrees with x on the ℓ-th bit (*i.e.*, $y_\ell \neq x_\ell$), $dist_\ell(y) > \frac{1}{4} - \frac{\varepsilon}{4}$. Combined together this proves that the output z of the algorithm agrees with the encoded message x on at least $(1 - O(\lambda^{1/3}))n$ of its bits.

We analyze the soft error reduction algorithm a general case of variants of ABNNR codes constructed from *possibly unbalanced* expander graphs. Namely, graphs $H = ([k], [b], E)$, satisfying *expander mixing lemma*: For every $A \subseteq [k]$ and $B \subseteq [b]$,

$$\left| |E(A,B)| - \frac{d_L |A| |B|}{b} \right| < \lambda d_L \sqrt{|A| |B|}$$

Let $\mathcal{C}(\mathcal{H}, \mathcal{C}_0)$ be a binary ABNNR code with $\mathcal{H}$ a family of (d_L, d_R)-regular λ-expander graphs $H = ([k], [b], E)$, and $\mathcal{C}_0$ a family of binary error correcting codes of rate r and normalized Hamming distance $\Delta(\mathcal{C}_0) = 1/2$.

Theorem 6 (Soft Error Reduction): For any $k \in \mathbb{N}$, a message $x \in \{0,1\}^k$, a agreement parameter $\varepsilon > 8(\sqrt{2}\lambda)^{1/3}$, and a corrupted codeword $w \in \{0,1\}^{k/r}$ s.t. $\Delta(w, C(x)) \leq \frac{1}{4} - \varepsilon$, the Soft Error Reduction Algorithm 5, given ε and w, outputs a message $z \in \{0,1\}^k$ s.t.

$$\Delta(x, z) < (\sqrt{2}\lambda)^{1/3}$$

The running time of the algorithm in $O\left(k \cdot 2^{d_L d_R}\right)$.

Proof. Fix an input corrupted codeword w and let x be the encoded message s.t. $\Delta(C(x), w) < \frac{1}{4} - \varepsilon$. Recall that for each left node ℓ, we denote by $dist_\ell(y) = \frac{1}{d} \sum_{i \in \Gamma(\ell)} \Delta(C_0(y)^i, w^i)$ the average distance from w of codewords encoding (messages agreeing with) y on the block corresponding to right neighbors i of left node ℓ.

By Lemma 7 below when assigning $\gamma = \varepsilon/8$, for at least $(1 - \frac{\varepsilon}{8})k$ of the bits $\ell \in [k]$ of the encoded message x,

$$dist_\ell(x) < \frac{1}{4} - \frac{\varepsilon}{4}$$

Conversely, by Lemma 8 below, for each left node $\ell \in [k]$ and any $y \in \{0,1\}^k$ such that $y_\ell \neq x_\ell$,

$$dist_\ell(y) > \frac{1}{4} - \frac{\varepsilon}{4}$$

Combining both lemmas together, we conclude that for at least $(1 - \frac{\varepsilon}{8})k$ of the bits $\ell \in [k]$, $z_\ell = x_\ell$ (for z_ℓ from Algorithm 5 above). That is, $\Delta(x, z) < (\sqrt{2}\lambda)^{1/3}$.

The running time of this algorithm is $k \cdot 2^{d^2}$. □

We now prove the lemmas used in the proof of Theorem 6 above. We use the following notation: For each right node $i \in [b]$ and any assignment $y \in \{0,1\}^{d_R}$ to the left neighbors of i, we denote the distance of the i-th block of w from $C_0(y)$ (*i.e.*, the i-th block of the encoding of any message agreeing with y) by

$$\Delta_i(y) \stackrel{def}{=} \Delta(C_0(y)^i, w^i)$$

We keep the same notation as in Theorem 6.

Lemma 7 (Main Lemma): Let $\gamma \in ((\sqrt{2}\lambda)^{1/3}, 1)$. Then for any $x \in \{0,1\}^k$, for at least $(1 - \gamma)k$ of the bits $\ell \in 1, \ldots, k$,

$$\left| \mathop{\mathbb{E}}_{i \in \Gamma(\ell)} [\Delta_i(x)] - \mathop{\mathbb{E}}_{i \in 1, \ldots, b} [\Delta_i(x)] \right| \leq 2\gamma$$

Proof. We first fix some notation. Denote $\delta = \gamma^2 d_L$. Fix some $x \in \{0,1\}^k$. For $j = 1, \ldots, 1/\gamma$, denote

$$T_j = \{\, i \in 1, \ldots, b \mid \Delta_i(x) \in [(j-1)\gamma, j\gamma) \}$$

Claim 7.1: For any $\ell \in 1, \ldots, k$, if $\forall j \in 1, \ldots, 1/\gamma$, $\left| E(\{\ell\}, T_j) - d_L \frac{|T_j|}{b} \right| \leq \delta$, then

$$\left| \mathop{\mathbb{E}}_{i \in \Gamma(\ell)} [\Delta_i(x)] - \mathop{\mathbb{E}}_{i \in 1, \ldots, b} [\Delta_i(x)] \right| \leq 2\gamma$$

Proof: First we bound $\mathbb{E}_{i\in\Gamma(\ell)}[\Delta_i(x)]$. Rewriting $\mathbb{E}_{i\in\Gamma(\ell)}[\Delta_i(x)]$ in terms of the T_j's we get

$$\mathop{\mathbb{E}}_{i\in\Gamma(\ell)}[\Delta_i(x)] = \frac{1}{|\Gamma(\ell)|}\sum_{i\in\Gamma(\ell)}\Delta_i = \frac{1}{d_L}\sum_{j=1}^{1/\gamma}\sum_{i\in T_j\cap\Gamma(\ell)}\Delta_i$$

(where in the last step we use the fact that $|\Gamma(\ell)| = d_L$, because H is (d_L, d_R)-regular). By the definition of T_j, for all $i \in T_j$, $\Delta_i \in [(j-1)\gamma, j)$ for all $i \in T_j$. Therefore,

$$\frac{1}{d_L}\sum_{j=1}^{1/\gamma}|T_j\cap\Gamma(\ell)|\,(j-1)\gamma \leq \mathop{\mathbb{E}}_{i\in\Gamma(\ell)}[\Delta_i(x)] \text{ and}$$

$$\mathop{\mathbb{E}}_{i\in\Gamma(\ell)}[\Delta_i(x)] < \frac{1}{d_L}\sum_{j=1}^{1/\gamma}|T_j\cap\Gamma(\ell)|\,j\gamma$$

Now, $|T_j\cap\Gamma(\ell)| = |E(\{\ell\},T_j)| \in d_L\frac{|T_j|}{b} \pm \delta$, therefore,

$$\begin{aligned}&\frac{1}{d_L}\sum_{j=1}^{1/\gamma}\left(d_L\frac{|T_j|}{b}+\delta\right)(j-1)\gamma\\ \leq\ &\mathop{\mathbb{E}}_{i\in\Gamma(\ell)}[\Delta_i(x)]\\ <\ &\frac{1}{d_L}\sum_{j=1}^{1/\gamma}\left(d_L\frac{|T_j|}{b}-\delta\right)j\gamma\end{aligned}$$

Canceling the d_L in the denominator and the nominator we get

$$\sum_{j=1}^{1/\gamma}\left(\frac{|T_j|}{b}-\frac{\delta}{d_L}\right)(j-1)\gamma \leq \mathop{\mathbb{E}}_{i\in\Gamma(\ell)}[\Delta_i(x)] \quad (1)$$

$$\mathop{\mathbb{E}}_{i\in\Gamma(\ell)}[\Delta_i(x)] < \sum_{j=1}^{1/\gamma}\left(\frac{|T_j|}{b}+\frac{\delta}{d_L}\right)j\gamma \quad (2)$$

Second, we bound $\mathbb{E}_{i=1,\ldots,b}[\Delta_i(x)]$. Rewriting $\mathbb{E}_{i=1,\ldots,b}[\Delta_i(x)]$ in terms of the T_j's we get

$$\mathop{\mathbb{E}}_{i=1,\ldots,b}[\Delta_i(x)] = \frac{1}{b}\sum_{i=1}^{b}\Delta_i = \frac{1}{b}\sum_{j=1}^{1/\gamma}\sum_{i\in T_j}\Delta_i$$

By the definition of T_j, for all $i \in T_j$, $\Delta_i \in [(j-1)\gamma, j)$ for all $i \in T_j$. Therefore,

$$\sum_{j=1}^{1/\gamma}\frac{|T_j|}{b}(j-1)\gamma \leq \mathop{\mathbb{E}}_{i=1,\ldots,b}[\Delta_i(x)] < \sum_{j=1}^{1/\gamma}\frac{|T_j|}{b}j\gamma \quad (3)$$

Combining the bounds from Equations 1 and 3 we conclude that

$$\begin{aligned}&\left|\mathop{\mathbb{E}}_{i\in\Gamma(\ell)}[\Delta_i(x)] - \mathop{\mathbb{E}}_{i=1,\ldots,b}[\Delta_i(x)]\right|\\ \leq\ &\sum_{j=1}^{1/\gamma}\frac{|T_j|}{b}\gamma + \sum_{j=1}^{1/\gamma}\frac{\delta}{d_L}j\gamma \leq 2\gamma\end{aligned}$$

(where the last inequality is true since $\sum_{j=1}^{1/\gamma}\frac{|T_j|}{b}\gamma = \gamma$, and $\sum_{j=1}^{1/\gamma}\frac{\delta}{d_L}j\gamma = \frac{\delta}{d_L}\gamma\frac{\frac{1}{\gamma}(\frac{1}{\gamma}+1)}{2} < \frac{\delta}{\gamma d_L} = \gamma$ for $\delta = \gamma^2 d_L$). ■

Claim 7.2: Denote $Bad = \left\{\ell \mid \exists j \in 1,\ldots,1/\gamma \text{ s.t. } \left|E(\{\ell\},T_j) - d_L\frac{|T_j|}{b}\right| > \delta\right\}$. If $\lambda < \gamma^3/\sqrt{2}$, then $|Bad| < \gamma b$

Proof: For each $j \in 1,\ldots,\gamma$, denote

$$\begin{aligned}Bad_j^+ &= \left\{\ell \mid |E(\{\ell\},T_j)| > d_L\frac{|T_j|}{b}+\delta\right\}\\ Bad_j^- &= \left\{\ell \mid |E(\{\ell\},T_j)| < d_L\frac{|T_j|}{b}-\delta\right\}\end{aligned}$$

By a counting argument, there exists $j \in 1,\ldots,1/\gamma$ such that either $\left|Bad_j^+\right| \geq \frac{\gamma}{2}|Bad|$ or $\left|Bad_j^-\right| \geq \frac{\gamma}{2}|Bad|$. Without loss of generality assume

$$\left|Bad_j^+\right| \geq \frac{\gamma}{2}|Bad| \quad (4)$$

By definition of Bad_j^+, for each $\ell \in Bad_j^+$, $|E(\{\ell\},T_j)| > d_L\frac{|T_j|}{b}+\delta$. Therefore,

$$\left|E(Bad_j^+,T_j)\right| > \left|Bad_j^+\right|\left(d_L\frac{|T_j|}{b}+\delta\right) \quad (5)$$

On the other hand, by Expander Mixing Lemma ,

$$\begin{aligned}&\left|E(Bad_j^+,T_j)\right| \quad (6)\\ \leq\ &d_L\left|Bad_j^+\right|\frac{|T_j|}{b} + \lambda d_L\sqrt{\left|Bad_j^+\right||T_j|} \quad (7)\end{aligned}$$

Combining Equations 5 and 6, we get that

$$\begin{aligned}&\left|Bad_j^+\right|\left(d_L\frac{|T_j|}{b}+\delta\right)\\ <\ &d_L\left|Bad_j^+\right|\frac{|T_j|}{b} + \lambda d_L\sqrt{\left|Bad_j^+\right||T_j|}\end{aligned}$$

Reorganizing the above expression and assigning $\delta = \gamma^2 d_L$ and $|T_j| \leq b$, we get

$$\sqrt{\left|Bad_j^+\right|} < \frac{\lambda d_L}{\delta}\sqrt{|T_j|} \leq \frac{\lambda}{\gamma^2}\sqrt{b}$$

Combining the above with Equation 4, we conclude that

$$|Bad| < \frac{2}{\gamma}\left(\frac{\lambda}{\gamma^2}\right)^2 b < \gamma b$$

(where the last inequality is true by the condition of λ). ■ □

Lemma 8: For any $x, y \in \{0,1\}^k$ s.t. $x_\ell \neq y_\ell$, if $\mathbb{E}_{i\in\Gamma(\ell)}[\Delta_i(x)] < \frac{1}{4} - \frac{\varepsilon}{4}$ then

$$\mathop{\mathbb{E}}_{i\in\Gamma(\ell)}[\Delta_i(y)] > \frac{1}{4} - \frac{\varepsilon}{4}$$

Proof. Recall that x^i, y^i denote the restrictions of x and x, y, respectively, to the bits $i_1,\ldots,i_{d_R} \in 1,\ldots,k$ neighboring the i-th right node of H. For all $i \in \Gamma(\ell)$, $x^i \neq y^i$, because ℓ is a neighbor of each $i \in \Gamma(\ell)$, *i.e.*, $\ell \in i_1,\ldots,i_{d_R}$. Therefore,

$$\forall i \in \Gamma(\ell),\ \Delta(x^i,y^i) \geq dist(C_0) \geq \frac{1}{2} - \frac{\varepsilon}{2} \quad (8)$$

By the triangle inequality,

$$\Delta(x^i, y^i) \leq \Delta(x^i, w^i) + \Delta(w^i, y^i) \tag{9}$$

Combining Equations 8 and 9 above, we get that

$$\Delta(w^i, y^i) \geq \Delta(x^i, y^i) - \Delta(x^i, w^i) \geq \frac{1}{2} - \frac{\varepsilon}{2} - \Delta(x^i, w^i)$$

Taking the expectation over all $i \in \Gamma(\ell)$, we conclude that

$$\begin{aligned} \mathop{\mathbb{E}}_{i \in \Gamma(\ell)} \left[\Delta(w^i, y^i)\right] &\geq \frac{1}{2} - \frac{\varepsilon}{2} - \mathop{\mathbb{E}}_{i \in \Gamma(\ell)} \left[\Delta(x^i, w^i)\right] \\ &\geq \frac{1}{2} - \frac{\varepsilon}{2} - \left(\frac{1}{4} - \frac{\varepsilon}{4}\right) = \frac{1}{4} - \frac{\varepsilon}{4} \end{aligned}$$

□

VII. ACKNOWLEDGMENTS

We are grateful to Shafi Goldwasser, Abishek Kumarasubramanian, Avinash Vaidyanathan, Irit Dinur, Yuval Isahi, Venkat Guruswami, Alex Samorodnitsky, Adi Shamir, Amir Shpilka and Madhu Sudan for helpful discussions.

REFERENCES

[1] M. Ajtai, H. Iwaniec, J. Komlos, J. Pintz and E. Szemeredi, "Constructions of a this set with small Fourier coefficients," *B*ull. London Math. Soc. 22, pp. 583–590, 1990.

[2] A. Akavia, S. Goldwasser, and S. Safra. "Proving Hardcore Predicates Using List Decoding," *IEEE Symposium on Foundations of Computer Science*, 11-14 October 2003, Cambridge, MA, USA, Proceedings, pp. 146-158.

[3] N. Alon, J. Bruck, J. Naor, M. Naor, R.M. Roth, "Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs", *IEEE Trans. Inform. Theory*, 38 (1992), 509-516.

[4] N. Alon, O. Goldreich, J. Hastad and R. Peralta, "Simple constructions of almost k-wise independent random variables," *J*ournal of random structures and algorithms, 3:3 (1992), pp. 289–304.

[5] N. Alon and Y. Mansour, "ε-discrepancy sets and their applications for interpolation of sparse polynomial," *I*nformation Processing Letters, 54:337–342 (1995).

[6] Eli Ben-Sasson, Oded Goldreich, Prahladh Harsha, Madhu Sudan and Salil Vadhan, "Robust PCPs of Proximity, Shorter PCPs and Applications to Coding", *ACM Symposium on the Theory of Computing*, 2004.

[7] R. Blahut. "Algebraic Methods for Signal Processing and Communications Coding," *Springer Verlag*, 1991.

[8] G. D. Forney, "Generalized Minimum Distance decoding", *IEEE Transactions on Information Theory*, 12, 125-131, 1966

[9] G. Even, O. Goldreich, M. Luby, N. Nisan and B. Velickovic, "Approximations of general independent distributions," *ACM Symposium on the Theory of Computing*'92, pp. 10–16, 1992.

[10] O. Goldreich and L. Levin "A Hard-Core Predicate for All One-Way Functions," *ACM Symposium on the Theory of Computing*, May 1989, New York, NY, USA, Proceedings, pp. 25–32.

[11] V. Guruswami and P. Gopalan, "Hardness Amplification within NP against Deterministic Algorithms," *IEEE Conference on Computational Complexity*, pp.19–30, 2008.

[12] V. Guruswami and P. Indyk, "Expander-Based Constructions of Efficiently Decodable Codes", *Proc. IEEE Annual Symposium on Foundations of Computer Science (FOCS'01)*, 658-667, 2001.

[13] , V. Guruswami and P. Indyk, "Near-optimal linear-time codes for unique decoding and new list-decodable codes over smaller alphabets", *Proc. 34th ACM Annual Symposium on Theory of Computing (STOC'02)*, 812–821, 2002.

[14] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *American Statistical Association Journal*, March 1962, pp. 13-30.

[15] M. Katz, "An estimate for characters sum," *J*. AMS 2, (1963) pp. 197–200.

[16] V.Pless, W.C Huffman R. Brualdi "Handbook of Coding Theory" Vol I and II, North Holland. 1999

[17] J. Naor and M. Naor, "Small biased probability spaces: efficient constructions and applications," 22nd ACM Symposium on the Theory of Computing, pp. 213–223, 1990.

[18] A. A. Razborov, A. Wigderson and E. Szemeredi, "Constructing small sets that are uniform in arithmetic progressions," *C*ombinatorics, Probability and Computing 2:513–518, 1993.

[19] D. Spielman, "Linear-time encodable and decodable error-correcting codes", *IEEE Transactions on Information Theory*, Vol 42, No 6, pp. 1723-1732. 1996.