UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Apple Inc.
Petitioner

v.

California Institute of Technology
Patent Owner

Case IPR2017-00211

UNOPPOSED MOTION TO SUBMIT REPLACEMENT EXHIBITS PURSUANT TO 37 C.F.R. § 42.104(c)

Pursuant to 37 C.F.R. § 42.104(c), Petitioner Apple Inc. respectfully requests that the Patent Trial and Appeal Board ("Board") grant leave to the Petitioner to submit replacement Exhibits that address inadvertent clerical errors made when filing the following three exhibits in the IPR2017-00210, -00211, and -00219 petitions:

- Frey, B. J. and MacKay, D. J. C., "Irregular Turbocodes," *Proc. 37th Allerton Conf. on Comm., Control and Computing*, Monticello, Illinois, 1999 (the "Frey Reference").
- D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for 'turbo-like' codes," *Proc. 36th Allerton Conf. on Comm., Control and Computing*, Allerton, Illinois, 1998 (the "Divsalar Reference").
- Declaration of Paul H. Siegel (the "Siegel Declaration").

Petitioner requested a telephonic hearing on this issue on February 21, 2017, which the Board held on February 24, 2017. At the February 24 hearing, the Board directed Petitioner to file this motion under 37 C.F.R. § 42.104(c) together with the replacement exhibits.

Accordingly, and for the reasons set forth more fully below, Petitioner respectfully requests that the Board: (1) replace the Frey Reference originally filed as Exhibit 1102 with the replacement copy attached as Appendix A to this motion;

(2) replace the Divsalar Reference originally filed as Exhibit 1103 with the replacement copy attached as Appendix B to this motion; and (3) replace the Siegel Declaration originally filed as Exhibit 1120 with the replacement copy attached as Appendix C to this motion. Patent Owner does not oppose this motion.

I. APPLICABLE RULE

Pursuant to 37 C.F.R. § 42.104(c), a party may file a motion "to correct a clerical or typographical mistake in a petition." The Board has explained that this rule is remedial in nature and subject to liberal interpretation. *ABB Inc. v. ROY-G-BIV Corp.*, IPR2013-00063, Paper 21 at 7 (PTAB Jan. 16, 2013) (citing *Tcherepnin v. Knight*, 389 U.S. 332, 336 (1967)). The Board has regularly granted motions to correct inadvertent errors related to the filing of exhibits pursuant to § 42.104(c). *See, e.g.*, *Owens Corning v. Certainteed Corp.*, IPR2014-01397, Paper 10 at 2 (PTAB Dec. 17, 2014); *Syntroleum Corp. v. Neste Oil OYJ*, IPR2013-00178, Paper 21 at 5 (PTAB July 22, 2013).

II. FACTS RELEVANT TO THIS MOTION

On November 15, 2016, Petitioner filed three *inter partes* review petitions in IPR2017-00210, -00211, and -00219 directed to U.S. Patent No. 7,116,710. During the preparation and filing of these petitions, lead counsel Richard Goldenberg directed that the copies of the Frey Reference, Divsalar Reference, and

the Siegel Declaration (attached hereto as Appendices A, B, and C, respectively, to distinguish them from originally-filed Exhibits 1102, 1103, and 1120) were to be filed with the petitions. (Goldenberg Declaration, Ex. 1125, ¶¶6, 9, 10.) In carrying out these instructions, however, the associate assisting in uploading these exhibits, Jonathan E. Barbee, mistakenly directed legal staff to upload incorrect copies of the Frey Reference, the Divsalar Reference, and the Siegel Declaration. (Barbee Declaration, Ex. 1126, ¶¶4-7.) As explained below, this occurred due to clerical errors in the preparation of the exhibits to the petitions.

A. The Frey Reference

Counsel for Petitioner had several additional copies of the Frey Reference in the firm's document management database, including the inadvertently-filed exhibit, which lacks a table of contents and a date stamp. The associate assisting with uploading the exhibits, Mr. Barbee, unintentionally selected the wrong copy of the Frey Reference because the inadvertently-filed document had been circulated for a different purpose. (Barbee Declaration, Ex. 1126, ¶5.) The text of the replacement Frey Reference is identical to the text of the inadvertently-filed exhibit and will not affect the substance of the IPR2017-00210, -00211, and -00219 petitions, but merely corrects a clerical error.

The inadvertently-filed exhibit lacks page numbering corresponding to the Table of Contents of the publication in which the Frey Reference was published. The Table of Contents from that publication was filed as a separate exhibit with the IPR2017-00210, -00211, and -00219 petitions as Ex. 1015, Ex. 1115, and Ex. 1215, respectively. The Table of Contents bears the same date stamp as the replacement Frey Reference (*i.e.*, March 20, 2000 from the Cornell University Library) and indicates that the first page is page 241. The pagination of the inadvertently-filed Frey exhibit does not match the pagination identified in the Table of Contents in Ex. 1015, Ex. 1115, and Ex. 1215 because the inadvertently-filed Frey exhibit begins at page 1. In the replacement Frey Reference, the first page of the exhibit is page 241, which matches the pagination indicated in the Table of Contents of the publication in which the Frey Reference was published, as shown in Ex. 1015, Ex. 1115, and Ex. 1215.

B. The Divsalar Reference

Counsel for Petitioner cited to the Divsalar Reference in the IPR2017-00210, -00211, and -00219 petitions using sequential page numbering that designated the first page as page 1, whereas the inadvertently-filed exhibit bears only sequential page numbering that begins with page 201. The associate assisting in uploading the exhibits, Mr. Barbee, inadvertently directed legal staff to upload the Divsalar

Reference without adding sequential page numbers beginning with page 1 below the original page numbers of the exhibit. (Barbee Declaration, Ex. 1126, ¶6.) The addition of the sequential page numbers was needed for the Divsalar Reference to match the citations in the IPR2017-00210, -00211, and -00219 petitions. The replacement Divsalar Reference is identical to the inadvertently-filed exhibit except for the addition of the sequential page numbers and will not affect the substance of the IPR2017-00210, -00211, and -00219 petitions, but merely corrects a clerical error.

C. The Siegel Declaration

Petitioner obtained from Professor Paul. H. Siegel a declaration that includes an attachment called "Exhibit 1." "Exhibit 1" to the Siegel Declaration was inadvertently omitted, however, when the associate assisting with the exhibits, Mr. Barbee, directed legal staff to upload this declaration. (Barbee Declaration, Ex. 1126, ¶7.).

Specifically, the Siegel Declaration and "Exhibit 1" were sent by Professor Siegel to counsel for Petitioner in separate emails. Mr. Barbee inadvertently overlooked the email forwarding "Exhibit 1" in the process of preparing the Siegel Declaration for filing. A copy of the same presentation in "Exhibit 1" was filed with the IPR2017-00210, -00211, and -00219 petitions as Ex. 1005, Ex. 1105, and

Ex. 1205, respectively. This is because the Siegel Declaration was submitted to demonstrate that Ex. 1005, Ex. 1105, and Ex. 1205 were true and accurate copies of the presentation. The replacement Siegel Declaration is identical to the inadvertently-filed exhibit except for the addition of "Exhibit 1", which is identical to Ex. 1005, Ex. 1105, and Ex. 1205 in IPR2017-00210, -00211, and -00219, respectively (specifically, all are copies of the presentation entitled "The Serial Concatenation of Rate-1 Codes Through Uniform Random Interleavers," which is identified by name on p. 3 of the Siegel Declaration), and will not affect the substance of the IPR2017-00210, -00211, and -00219 petitions, but merely corrects a clerical error.

III. THIS MOTION IS UNOPPOSED

While Patent Owner (California Institute of Technology, or "Caltech") does not acquiesce to any factual assertions made by the Petitioner in its motion or corresponding supporting documents, Caltech does not oppose the relief requested in this motion in that replacement documents are accepted for the Frey Reference, the Divsalar Reference, and the Siegel declaration. During the hearing on February 24, 2017, the Board explained that, due to the filing of the corrected exhibits, and should Caltech not oppose the requested replacement of exhibits, the Patent Owner would receive a one-month extension of the deadline for filing its

Preliminary Patent Owner Response, making that response due on April 3. Petitioner does not oppose the one-month extension for filing of the Preliminary Patent Owner Response.

IV. THE INADVERTENT ERRORS IN THE FREY REFERENCE, THE DIVSALAR REFERENCE, AND THE SIEGEL DECLARATION ARE CLERICAL ERRORS THAT SHOULD BE CORRECTED

The inadvertent errors in the filing of the Frey Reference, the Divsalar Reference, and the Siegel Declaration should be corrected because correcting these clerical errors will not prejudice Patent Owner and will not affect the substance of the petitions in IPR2017-00210, -00211, and -00219. The substance of the replacement copies of the Frey Reference, the Divsalar Reference, and the Siegel Declaration is identical to the inadvertently-filed exhibits. As explained above, the differences between the replacement copies and the inadvertently-filed exhibits are clerical in nature.

V. CONCLUSION

For the foregoing reasons, Petitioner respectfully requests that Ex. 1102 (the Frey Reference), Ex. 1103 (the Divsalar Reference), and Ex. 1120 (the Siegel Declaration) to the petition in the above-captioned *inter partes* review proceeding be replaced with the replacement copies attached to this motion as Appendices A, B, and C, respectively.

Dated: February 28, 2017

Respectfully Submitted,

/Dominic E. Massa/

Dominic E. Massa
Registration No. 44,905
WILMER CUTLER PICKERING
HALE AND DORR LLP
60 State Street
Boston, MA 02109
Tel: (617) 526-6386
Fax: (617) 526-5000

Attorney for Petitioner

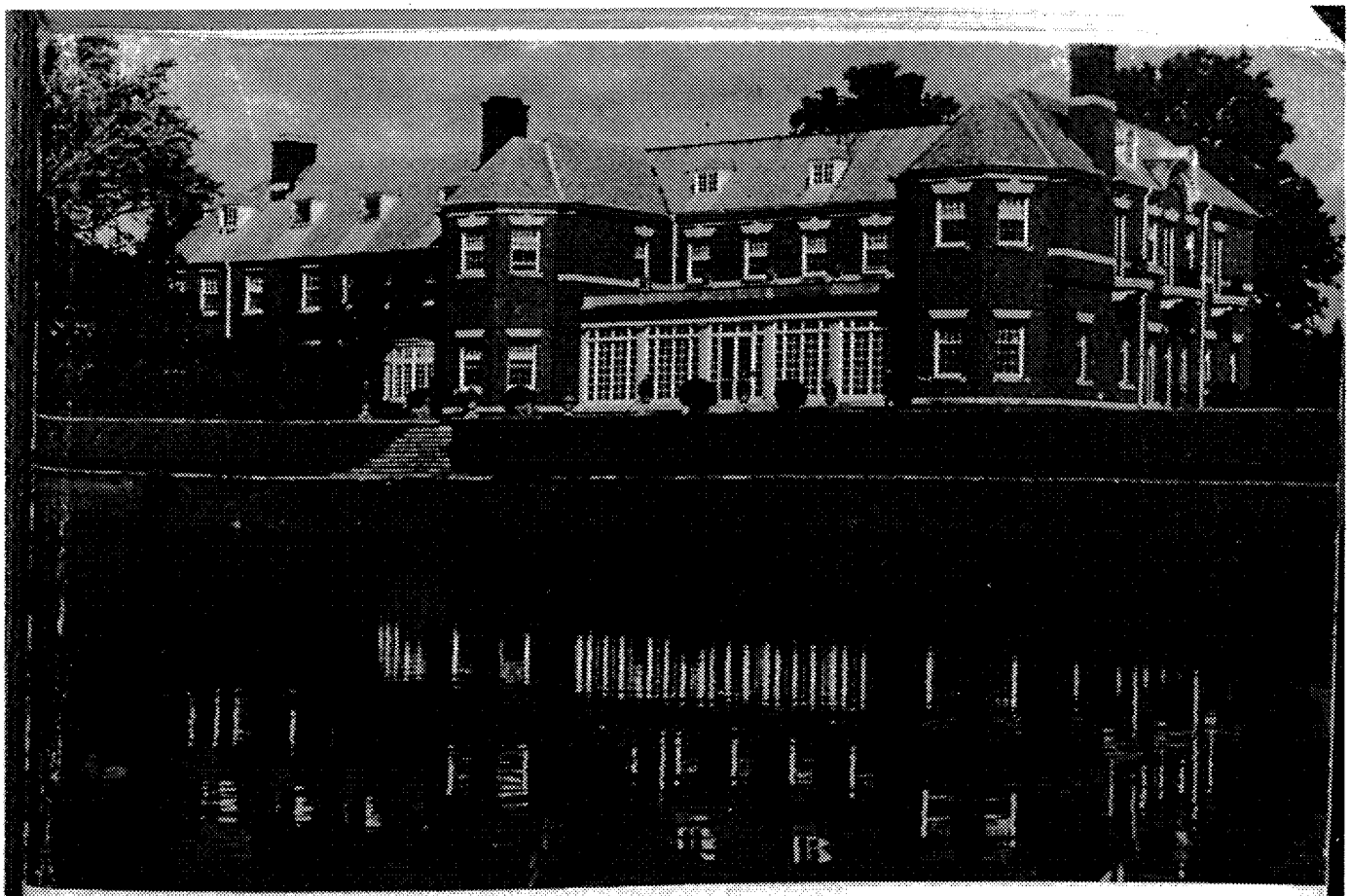
CERTIFICATE OF SERVICE

I hereby certify that on February 28, 2017, a true and correct copy of the foregoing MOTION TO SUBMIT REPLACEMENT EXHIBITS PURSUANT TO 37 C.F.R. § 42.104(c) and Petitioner's Updated Exhibit List were served via electronic mail upon the following attorneys of record:

Michael Rosato (mrosato@wsgr.com)
Matthew Argenti (margenti@wsgr.com)
Richard Torczon (rtorczon@wsgr.com)

/Dominic E. Massa/
Dominic E. Massa
Registration No. 44,905

Appendix A



Proceedings

THIRTY-SEVENTH ANNUAL ALLERTON CONFERENCE ON COMMUNICATION, CONTROL AND COMPUTING

September 22 - 24, 1999

Allerton House, Monticello, Illinois
Sponsored by the
Coordinated Science Laboratory and the
Department of Electrical and Computer Engineering of the
University of Illinois at Urbana-Champaign

PROCEEDINGS

THIRTY-SEVENTH ANNUAL ALLERTON CONFERENCE ON COMMUNICATION, CONTROL, AND COMPUTING

Bruce Hajek
R.S. Sreenivas
Conference Co-Chairs

Conference held
September 22, September 23, and September 24, 1999
Allerton House
Monticello, Illinois

Sponsored by
The Coordinated Science Laboratory
and
The Department of Electrical and Computer Engineering
of the
UNIVERSITY OF ILLINOIS
at
Urbana-Champaign

TABLE OF CONTENTS

II-A: CODING THEORY II: ITERATIVE DECODING AND TURBO CODES

Organizers: R. Koetter and R.E. Blahut
(University of Illinois at Urbana-Champaign)

Chair: R. Koetter
(University of Illinois at Urbana-Champaign)

II-B: STOCHASTIC NETWORKS II

Organizers: S.P. Meyn and R. Srikant
(University of Illinois at Urbana-Champaign)

Chair: R. Srikant
(University of Illinois at Urbana-Champaign)

II-C: LEARNING ALGORITHMS IN SIGNAL PROCESSING

Organizers: A. Singer and M. Feder
(University of Illinois at Urbana-Champaign and Tel Aviv University)
Chair: M. Feder
(Tel Aviv University)

II-D: OPTICAL NETWORKS I

Organizers: M. Médard and E. Modiano
(University of Illinois at Urbana-Champaign
and Massachusetts Institute of Technology)
Chair: E. Modiano
(Massachusetts Institute of Technology)

II-E: COMMUNICATION SYSTEMS AND SERVICES
Chair: S. Lin
(University of Hawaii)

II-F: FADING CHANNELS AND POWER CONTROL
Chair: D. Sarwate
(University of Illinois at Urbana-Champaign)

IV-A: WIRELESS COMMUNICATIONS II: SYSTEM CONSIDERATIONS IN PHYSICAL LAYER DESIGN

Organizers: V.V. Veeravalli and U. Madhow
(Cornell University and University of Illinois at Urbana-Champaign)
Chair: U. Madhow
(University of Illinois at Urbana-Champaign)

IV-B: CODING THEORY IV

Chair: R.E. Blahut
(University of Illinois at Urbana-Champaign)

Irregular Turbocodes

Brendan J. Frey
Computer Science, University of Waterloo
Electrical and Computer Engineering, University of Illinois at Urbana
http://www.cs.uwaterloo.ca/~frey

David J. C. MacKay
Department of Physics, Cavendish Laboratories
Cambridge University
http://wol.ra.phy.cam.ac.uk/mackay

Abstract

Recently, several groups have increased the coding gain of iteratively decoded Gallager codes (low density parity check codes) by varying the number of parity check equations in which each codeword bit participates. In regular turbocodes, each "systematic bit" participates in exactly 2 trellis sections. We construct irregular turbocodes with systematic bits that participate in varying numbers of trellis sections. These codes can be decoded by the iterative application of the sum-product algorithm (a low-complexity, more general form of the turbodecoding algorithm). By making the original rate 1/2 turbocode of Berrou *et al.* slightly irregular, we obtain a coding gain of 0.15 dB at a block length of $N = 131,072$, bringing the irregular turbocode within 0.3 dB of capacity. Just like regular turbocodes, irregular turbocodes are linear-time encodable.

1 Introduction

Recent work on irregular Gallager codes (low density parity check codes) has shown that by making the codeword bits participate in varying numbers of parity check equations, significant coding gains can be achieved [1–3]. Although Gallager codes have been shown to perform better than turbocodes at BERs below 10^{-5} [4][1], until recently Gallager codes performed over 0.5 dB worse than turbocodes for BERs greater than 10^{-5}. However, in [3], Richardson *et al.* found irregular Gallager codes that perform 0.16 dB *better* than the original turbocode at BERs greater than 10^{-5} [5] for a block length of $N \approx 131,072$.

[1]Gallager codes to not exhibit decoding errors, only decoding failures, at long block lengths with $N > 5,000$.

In this paper, we show that by tweaking a turbocode so that it is irregular, we obtain a coding gain of 0.15 dB for a block length of $N = 131,072$. For example, an $N = 131,072$ irregular turbocode achieves $E_b/N_0 = 0.48$ dB at BER $= 10^{-4}$, a performance similar to the best irregular Gallager code published to date [3]. By further optimizing the degree profile, the permuter and the trellis polynomials, we expect to find even better irregular turbocodes. Like their regular cousins, irregular turbocodes exhibit a BER flattening due to low-weight codewords.

2 Irregular turbocodes

In Fig. 1, we show how to view a turbocode so that it can be made irregular. The first picture shows the set of systematic bits (middle row of discs) being fed directly into one convolutional code (the chain at the top) and being permuted before being fed into another convolutional code (the chain at the bottom). For a rate 1/2 turbocode, each constituent convolutional code should be rate 2/3 (which may, for example, be obtained by puncturing a lower-rate convolutional code).

Since the order of the systematic bits is irrelevant, we may also introduce a permuter before the upper convolutional code, as shown in the second picture. In the third picture, we have simply drawn the two permuters and convolutional codes side by side.

For long turbocodes, the values of the initial state and the final state of the convolutional chains do not significantly influence performance (*e.g.*, see [6]). So, as shown in the fourth picture, we can view a turbocode as a code that copies the systematic bits, permutes both sets of these bits, and then feeds them into a convolutional code. We refer to this turbocode as "regular", since each systematic bit is copied exactly once.

The final picture illustrates one way the above turbocode can be made irregular. Some of the systematic bits are "tied" together, in effect causing some systematic bits to be replicated more than once. Notice that too keep the rate of the overall code fixed at 1/2, some extra parity bits must be punctured.

More generally, an *irregular turbocode* has the form shown in Fig. 2, which is a type of "trellis-constrained code" as described in [7]. We specify a *degree profile*, $f_d \in [0,1], d \in \{1,2,\ldots,D\}$. f_d is the fraction of codeword bits that have degree d and D is the maximum degree. Each codeword bit with degree d is repeated d times before being fed into the permuter. Several classes of permuter lead to linear-time encodable codes. In particular, if the bits in the convolutional code are partitioned into "systematic bits" and "parity bits", then by connecting each parity bit to a degree 1 codeword bit, we can encode in linear time.

The average codeword bit degree is

$$\bar{d} = \sum_{d=1}^{D} d \cdot f_d \tag{1}$$

The overall rate R of an irregular turbocode is related to the rate R' of the convolutional code and the average degree $\bar{d}$ by

$$\bar{d}(1 - R') = 1 - R. \tag{2}$$

So, if the average degree is increased, the rate of the convolutional code must also be increased to keep the overall rate constant. This can be done by puncturing the convolutional code or by designing a new, higher rate convolutional code.

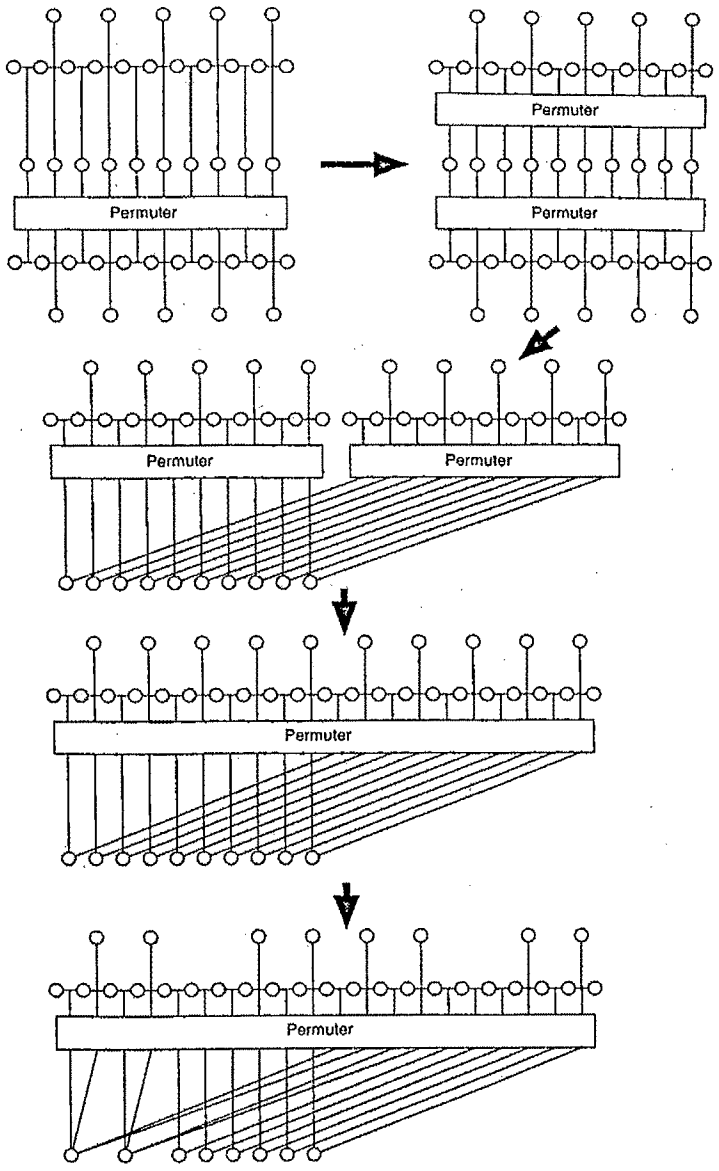


Figure 1: The first 4 pictures show that a turbocode can be viewed as a code that copies the systematic bits, permutes both sets of these bits and then feeds them into a convolutional code. The 5th picture shows how a turbocode can be made irregular by "tying" some of the systematic bits together, *i.e.*, by having some systematic bits replicated more than once. Too keep the rate fixed, some extra parity bits must be punctured. Too keep the block length fixed, we must start with a longer turbocode.

3 Decoding irregular turbocodes

Fig. 2 can be interpreted as the graphical model [6, 8–10] for the irregular turbocode. Decoding consists of the iterative application of the sum-product algorithm (a low-complexity, more general form of turbodecoding) in this graph.

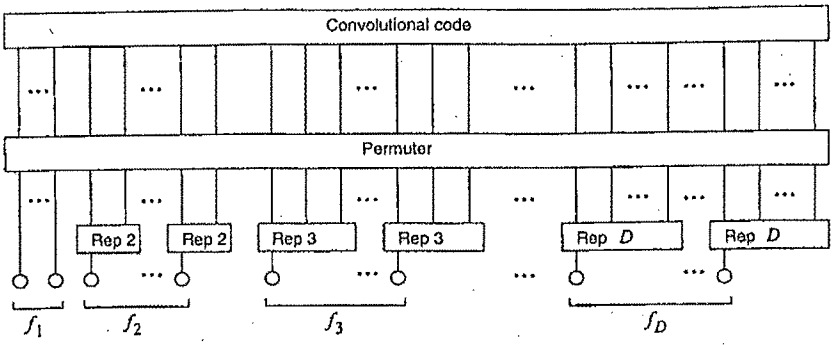


Figure 2: A general *irregular turbocode*. For $d = 1, \ldots, D$, fraction f_d of the codeword bits are repeated d times, permuted and connected to a convolutional code.

After receiving the channel output, the decoder computes the channel output log-likelihood ratios for the N codeword bits,

$$L_1^0, L_2^0, \ldots, L_N^0, \tag{3}$$

and then repeats each log-likelihood ratio appropriately. If codeword bit i has degree d, we set

$$L_{i,1} \leftarrow L_i^0,\ L_{i,2} \leftarrow L_i^0,\ \ldots,\ L_{i,d} \leftarrow L_i^0. \tag{4}$$

Next, the log-likelihood ratios are permuted and fed into the BCJR algorithm [11] for the convolutional code. The BCJR algorithm assumes the inputs are *a priori* log-probability ratios and uses the forward-backward algorithm [12] to compute a set of *a posteriori* log-probability ratios. If codeword bit i has degree d, the algorithm produces d *a posteriori* log-probability ratios,

$$L'_{i,1}, L'_{i,2}, \ldots, L'_{i,d}. \tag{5}$$

For a regular turbocode, there are just two *a posteriori* log-probability ratios, $L'_{i,1}$ and $L'_{i,2}$, for each degree 2 bit and they correspond to the "extrinsic information" produced by each constituent convolutional code.

The current estimate of the log-probability ratio for bit i given the channel output is

$$\hat{L}_i \leftarrow L_i^0 + \sum_{k=1}^{d} (L'_{i,k} - L_{i,k}). \tag{6}$$

To compute the inputs to the BCJR algorithm needed for the next iteration, we subtract off the corresponding outputs from the BCJR algorithm produced by the previous iteration:

$$L_{i,k} \leftarrow \hat{L}_i - L'_{i,k}. \tag{7}$$

So, each iteration consists of computing the inputs to the BCJR algorithm, permuting the inputs, applying the BCJR algorithm, permuting the outputs of the BCJR algorithm, and taking the repetitions into account to combine the outputs to form estimates of the log-probability ratios of the codeword bits given the channel output.

In our simulations, after each iteration, we check to see if the current decision gives a codeword. If it does, the iterations terminate and otherwise, the decoder iterates further until some maximum number of iterations is reached and a decoding failure is declared.

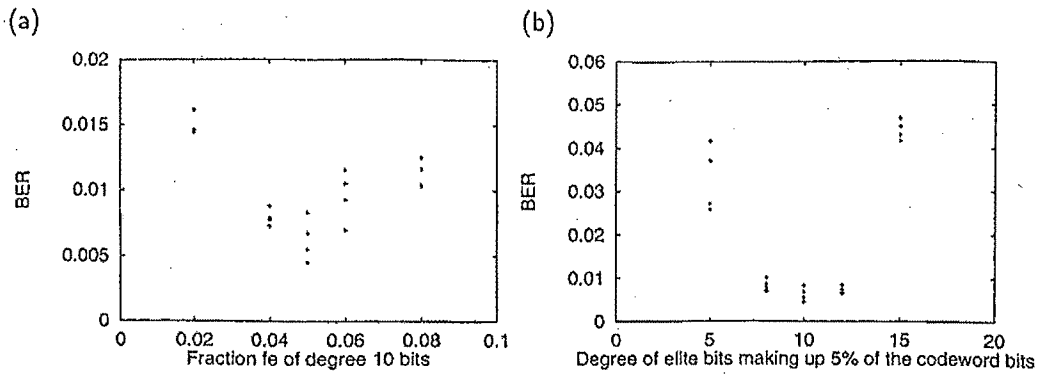


Figure 3: (a) shows the effect of changing the fraction of elite bits on the BER, while keeping the degree of the elite bits fixed at 10. (b) shows the effect of changing the degree of the elite bits on the BER, while keeping the fraction of elite bits fixed at 0.05. For each fraction and degree, the performance of 4 randomly drawn permuters is shown.

4 Selecting the profile

Finding a good profile is not trivial, since the best profile will depend on the parameters of the convolutional code, the permuter and the distortion measure (bit error rate, block error rate, decoding failure rate, high-weight decoding failure rate, *etc.*)

The results we report in this paper were obtained by making small changes to a block length $N = 10,000$ version of the original rate $R = 1/2$ turbocode proposed by Berrou *et al.*. In this turbocode, $f_1 = f_2 = 1/2$ (see Fig. 2) and the convolutional code polynomials are 37 and 21 (octal). The taps associated with polynomial 37 are connected to the degree 2 codeword bits, 1/2 of the taps associated with polynomial 21 are connected to the degree 1 bits, and the remaining 1/2 of the taps associated with polynomial 21 are punctured, giving the required convolutional code rate of $R' = 2/3$.

To simplify our search, we considered profiles where besides degrees 1 and 2, only one other degree, e for "elite", had a nonzero fraction. So, for a code with overall rate R and fraction f_e of degree e elite bits, we have

$$\begin{aligned} f_1 &= 1 - R = 1/2, \\ f_2 &= 1 - f_1 - f_e = 1/2 - f_e. \end{aligned} \tag{8}$$

In this restricted class of codes, irregularity is governed by two parameters, e and f_e.

From (2) it is clear that when the average degree is increased, the rate of the convolutional code must also be increased to keep the overall rate at 1/2. We increased the rate of the punctured convolutional code by further puncturing the taps associated with polynomial 21 to obtain a convolutional code with rate

$$R' = 1 - \frac{1-R}{\bar{d}} = 1 - \frac{1/2}{1/2 + 2(1/2 - f_e) + ef_e}. \tag{9}$$

So, in the codes we explored, the level of puncturing was quite high and some extra low-weight codewords were introduced.

To begin with, we made an irregular turbocode with $e = 10$ (chosen using intuition) and varied f_e from 0.02 to 0.08 while measuring the BER at $E_b/N_0 = 0.6$ dB. In each

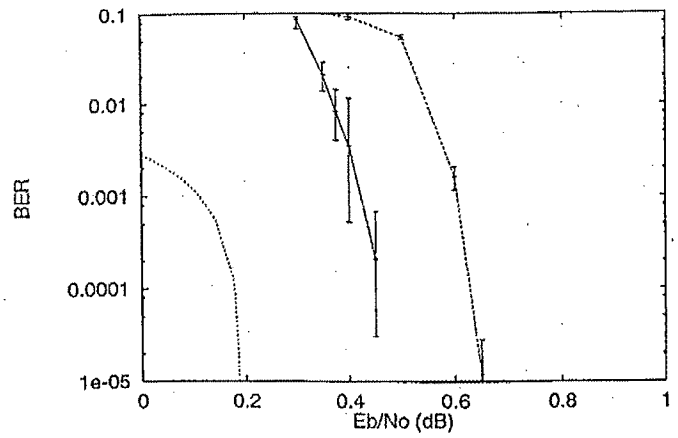


Figure 4: Performance of the original block length $N = 131,072$ turbocode (dashed line) and one of its irregular cousins (solid line). These results are for irregular turbocodes obtained by tweaking the original turbocode – we are currently searching for optimal degree profiles, permuters and trellis polynomials.

experiment, we simulated enough blocks to obtain a relatively small confidence interval. The results are shown in Fig. 3a, which indicates that for degree 10 elite bits, the best fraction is roughly 0.05. Next, we kept the fraction of elite bits fixed at $f_e = 0.05$ and we varied the degree of the elite bits. The results are shown in Fig. 3b, which indicates that for a fraction of 0.05, the best degree is roughly 10.

These results show that for $e = 10$, $f_e = 0.05$ is a good fraction and that for $f_e = 0.05$, $e = 10$ is a good degree. However, values of e and f_e that give good profiles are probably correlated, so we are currently extending our search.

5 Results

Fig. 4 shows the simulated BER-E_b/N_0 curves for the original block length $N = 131,072$ regular turbocode (dashed line) and its irregular cousin (solid line), using profile $e = 10$, $f_e = 0.05$.

The irregular turbocode clearly performs better than the regular turbocode for BER $> 10^{-4}$. At BER $= 10^{-4}$, the $N = 131,072$ irregular turbocode is 0.3 dB from capacity, a 0.15 dB improvement over the regular turbocode.

For high E_b/N_0, most of the errors for the irregular turbocode were due to low-weight codewords. According to preliminary results, the distribution of error weights appears to indicate that the flattening effect for the particular $N = 131,072$ irregular turbocode we constructed occurs at a *higher* BER than it does for the regular turbocode. However, the flattening effect is highly sensitive to the technique used to construct the permuter (we drew it at random) and the design of the convolutional code (we just further punctured the convolutional code used in the original turbocode). We are currently experimenting with techniques for lowering the level of the flattening effect (*e.g.*, see [13]).

6 Conclusions

We have shown that by making the original, regular turbocode irregular, a coding gain of 0.15 dB is obtained, bringing the irregular turbocode within 0.3 dB of capacity at a BER of 10^{-4}. This irregular turbocode performs in the same regime as the best known irregular Gallager code.

We emphasize that we obtained these results by tweaking the regular turbocode originally introduced by Berrou *et al.*. We believe we will be able to improve these performance curves significantly, both by exploring the polynomials used in the convolutional code and by adjusting the degree profile and the permuter structure. (One way to speed up the search is to extend the method of "density evolution" [3] to models with state.) In particular, we are investigating ways to select the permuter and the polynomials to eliminate low-weight codewords, thus reducing the flattening effect [13–15]. We are also studying ways of constraining the degree 1 "parity" bits (*i.e.*, increasing their degree) to eliminate low-weight codewords.

References

[1] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Improved low-density parity-check codes using irregular graphs and belief propagation," in *Proceedings of IEEE International Symposium on Information Theory*, 1998.

[2] D. J. C. MacKay, S. T. Wilson, and M. C. Davey, "Comparison of constructions of irregular Gallager codes," *IEEE Transactions on Communications*, vol. 47, October 1999.

[3] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of provably good low density parity check codes." Submitted to *IEEE Transactions on Information Theory*, July 1999.

[4] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Transactions on Information Theory*, vol. 45, pp. 399–431, March 1999.

[5] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Transactions on Communications*, vol. 44, pp. 1261–1271, October 1996.

[6] B. J. Frey, *Graphical Models for Machine Learning and Digital Communication*. Cambridge MA.: MIT Press, 1998. See `http://www.cs.utoronto.ca/~frey`.

[7] B. J. Frey and D. J. C. MacKay, "Trellis-constrained codes," in *Proceedings of the 35th Allerton Conference on Communication, Control and Computing 1997*, 1998. Available at `http://www.cs.utoronto.ca/~frey`.

[8] N. Wiberg, H.-A. Loeliger, and R. Kötter, "Codes and iterative decoding on general graphs," *European Transactions on Telecommunications*, vol. 6, pp. 513–525, September/October 1995.

[9] F. R. Kschischang and B. J. Frey, "Iterative decoding of compound codes by probability propagation in graphical models," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 219–230, February 1998.

[10] R. J. McEliece, D. J. C. MacKay, and J. F. Cheng, "Turbo-decoding as an instance of Pearl's 'belief propagation' algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 16, February 1998.

[11] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284–287, March 1974.

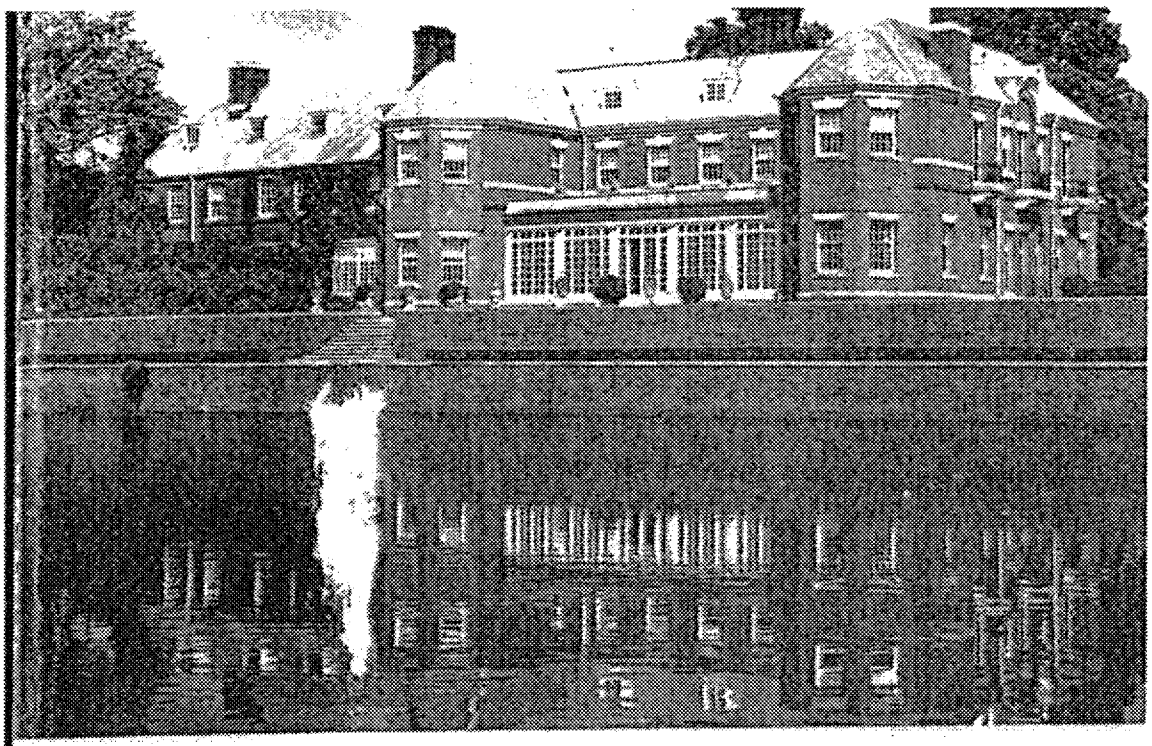
[12] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state markov chains," *Annals of Mathematical Statistics*, vol. 37, pp. 1559–1563, 1966.

[13] M. Öberg and P. H. Siegel, "Application of distance spectrum analysis to turbocode performance improvement," in *Proceedings of the* 35^{th} *Allerton Conference on Communication, Control and Computing 1997*, 1998.

[14] D. Divsalar and F. Pollara, "Turbo-codes for PCS applications," in *Proceedings of the International Conference on Communications*, pp. 54–59, 1995.

[15] D. Divsalar and R. J. McEliece, "Effective free distance of turbocodes," *Electronics Letters*, vol. 32, pp. 445–446, February 1996.

Appendix B



Proceedings

THIRTY-SIXTH ANNUAL ALLERTON CONFERENCE ON COMMUNICATION, CONTROL AND COMPUTING

September 23 - 25, 1998

Allerton House, Monticello, Illinois
Sponsored by the
Coordinated Science Laboratory and the
Department of Electrical and Computer Engineering of the
University of Illinois at Urbana-Champaign

PROCEEDINGS

THIRTY-SIXTH ANNUAL ALLERTON CONFERENCE ON COMMUNICATION, CONTROL, AND COMPUTING

Tamer Başar
Bruce Hajek
Conference Co-Chairs

Conference held
September 23 - 25, 1998
Allerton House
Monticello, Illinois

Sponsored by
The Coordinated Science Laboratory
and
The Department of Electrical and Computer Engineering
of the
UNIVERSITY OF ILLINOIS
at
Urbana-Champaign

Coding Theorems for "Turbo-Like" Codes*

Dariush Divsalar, Hui Jin, and Robert J. McEliece
Jet Propulsion Laboratory and California Institute of Technology
Pasadena, California USA
E-mail: dariush@shannon.jpl.nasa.gov, (hui, rjm)@systems.caltech.edu

Abstract.
In this paper we discuss AWGN coding theorems for ensembles of coding systems which are built from fixed convolutional codes interconnected with random interleavers. We call these systems "turbo-like" codes and they include as special cases both the classical turbo codes [1,2,3] and the serial concatentation of interleaved convolutional codes [4]. We offer a general conjecture about the behavior of the ensemble (maximum-likelihood decoder) word error probability as the word length approches infinity. We prove this conjecture for a simple class of rate $1/q$ serially concatenated codes where the outer code is a q-fold repetition code and the inner code is a rate 1 convolutional code with transfer function $1/(1+D)$. We believe this represents the first rigorous proof of a coding theorem for turbo-like codes.

1. Introduction.

The 1993 discovery of turbo codes by Berrou, Glavieux, and Thitimajshima [1] has revolutionized the field of error-correcting codes. In brief, turbo codes have enough randomness to achieve reliable communication at data rates near capacity, yet enough structure to allow practical encoding and decoding algorithms. This paper is an attempt to illuminate the first of these two attributes, i.e., the "near Shannon limit" capabilities of turbo-like codes on the AWGN channel.

Our specific goal is to prove AWGN coding theorems for a class of generalized concatenated convolutional coding systems with interleavers, which we call "turbo-like" codes. This class includes both parallel concatenated convolutional codes (classical turbo codes) [1, 2, 3] and serial concatenated convolutional codes [4] as special cases. Beginning with a code structure of this type, with fixed component codes and interconnection topology, we attempt to show that as the block length approaches infinity, the ensemble (over all possible interleavers) maximum likelihood error probability approaches zero if E_b/N_0 exceeds some threshold. Our proof technique is to derive an explicit expression for the ensemble input-output weight enumerator (IOWE) and then to use this expression, in combination with either the classical union bound, or the recent "improved" union bound of Viterbi and Viterbi [9], to show that the maximum likelihood word error probability approaches zero as $N \to \infty$. Unfortunately the difficulty of the first step, i.e., the computation of the ensemble IOWE, has kept us from full success, except for some very simple coding systems, which we call *repeat and accumulate* codes. Still, we are optimistic that this technique will yield coding theorems for a much wider class of interleaved concatenated codes. In any case, it is satisfying to have rigorously proved coding theorems for even a restricted class of turbo-like codes.

Here is an outline of the paper. In Section 2 we quickly review the classical union bound on maximum-likelihood word error probability for block codes on the AWGN

* Dariush Divsalar's work, and a portion of Robert McEliece's work, was performed at JPL under contract with NASA. The remainder of McEliece's work, and Hui Jin's work, was performed at Caltech and supported by NSF grant no. NCR-9505975, AFOSR grant no. 5F49620-97-1-0313, and grant from Qualcomm.

channel, which is seen to depend on the code's weight enumerator. In Section 3 we define the class of "turbo-like" codes, and give a formula for the average input-output weight enumerator for such a code. In Section 4 we state a conjecture (the interleaver gain exponent conjecture) about the ML decoder performance of turbo-like codes. In Section 5, we define a special class of turbo-like codes, the repeat-and-accumulate codes, and prove the IGE conjecture for them. Finally, in Section 6 we present performance curves for some RA codes, using an iterative, turbo-like, decoding algorithm. This performance is seen to be remarkably good, despite the simplicity of the codes and the suboptimality of the decoding algorithm.

2. Union Bounds on the Performance of Block Codes.

In this section we will review the classical union bound on the maximum-likelihood word error probability for block codes.

Consider a binary linear (n, k) block code C with code rate $r = k/n$. The *(output) weight enumerator* (WE) for C is the sequence of numbers $A_0, \ldots, A_n$, where A_h denotes the number of codewords in C with (output) weight h. The *input-output weight enumerator* (IOWE) for C is the array of numbers $A_{w,h}$, $w = 0, 1, \ldots, k$, $h = 0, 1, \ldots, n$: $A_{w,h}$ denotes the number of codewords in C with input weight w and output weight h.

The union bound on the word error probability P_W of the code C over a memoryless binary-input channel, using maximum likelihood decoding, has the well-known form

$$P_W \le \sum_{h=1}^{n} A_h z^h \tag{2.1}$$

$$= \sum_{h=1}^{n} \left(\sum_{w=1}^{k} A_{w,h} \right) z^h. \tag{2.2}$$

In (2.1) and (2.2), the function z^h represents an upper bound on the pairwise error probability for two codewords separated by Hamming (output) distance h. For AWGN channels, $z = e^{-rE_b/N_0}$ where E_b/N_0 is the signal-to-noise ratio per bit.

3. The Class of "Turbo-Like" Codes.

In this section, we consider a general class of concatenated coding systems of the type depicted in Figure 1, with q encoders (circles) and $q - 1$ interleavers (boxes). The ith code C_i is an (n_i, N_i) linear block code, and the ith encoder is preceded by an interleaver (permuter) P_i of size N_i, except C_1 which is not preceded by an interleaver, but rather is connected to the input. The overall structure must have no loops, i.e., it must be a graph-theoretic tree. We call a code of this type a "turbo-like" code.

Define $s_q = \{1, 2, \ldots, q\}$ and subsets of s_q by $s_I = \{i \in s_q : C_i \text{ connected to input}\}$, $s_O = \{i \in s_q : C_i \text{ connected to output }\}$, and its complement $\bar{s}_O$. The overall system depicted in Figure 1 is then an encoder for an (n, N) block code with $n = \sum_{i \in s_O} n_i$.

If we know the IOWE $A^{(i)}_{w_i,h_i}$'s for the constituent codes C_i, we can calculate the *average* IOWE $A_{w,h}$ for the overall system (averaged over the set of all possible interleavers), using the *uniform interleaver* technique [2]. (A uniform interleaver is defined as a probabilistic device that maps a given input word of weight w into all distinct $\binom{N_i}{w}$ permutations of it with equal probability $p = 1/\binom{N_i}{w}$.) The result is

$$A_{w,h} = \sum_{\substack{h_i : i \in s_O \\ \sum h_i = h}} \sum_{h_i : i \in \bar{s}_O} A^{(1)}_{w_1,h_1} \prod_{i=2}^{q} \frac{A^{(i)}_{w_i,h_i}}{\binom{N_i}{w_i}} \tag{3.1}$$

In (3.1) we have $w_i = w$ if $i \in s_I$, and $w_i = h_j$ if C_i is preceeded by C_j (see Figure 2.). We do not give a proof of formula (3.1), but it is intuitively plausible if we note that the term $A^{(i)}_{w_i,h_i}/\binom{N_i}{w_i}$ is the probability that a random input word to C_i of weight w_i will produce an output word of weight h_i.

For example, for the $(n_2+n_3+n_4, N)$ encoder of Figure 1 the formula (3.1) becomes

$$A_{w,h} = \sum_{\substack{h_1,h_2,h_3,h_4 \\ (h_2+h_3+h_4=h)}} A^{(1)}_{w_1,h_1} \frac{A^{(2)}_{w_2,h_2}}{\binom{N_2}{w_2}} \frac{A^{(3)}_{w_3,h_3}}{\binom{N_3}{w_3}} \frac{A^{(4)}_{w_4,h_4}}{\binom{N_4}{w_4}}$$

$$= \sum_{\substack{h_1,h_2,h_3,h_4 \\ (h_2+h_3+h_4=h)}} A^{(1)}_{w,h_1} \frac{A^{(2)}_{w,h_2}}{\binom{N}{w}} \frac{A^{(3)}_{h_1,h_3}}{\binom{n_1}{h_1}} \frac{A^{(4)}_{h_1,h_4}}{\binom{n_1}{h_1}}.$$

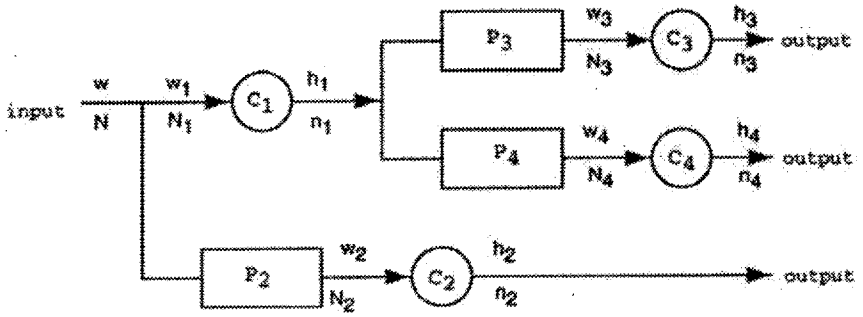


Figure 1. A "turbo-like" code with $s_I = \{1,2\}$, $s_O = \{2,3,4\}$, $3_O = \{1\}$.

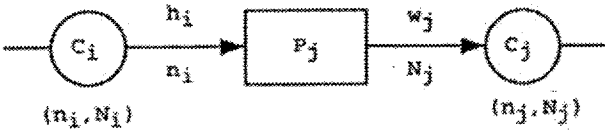


Figure 2. C_i (an (n_i, N_i) encoder) is connected to C_j (an (n_j, N_j) encoder) by an interleaver of size N_j. We have the "boundary conditions" $N_j = n_i$ and $w_j = h_i$.

4. The Interleaving Gain Exponent Conjecture.

In this section we will consider systems of the form depicted in Figure 1, in which the individual encoders are truncated convolutional encoders, and study the behavior of the average ML decoder error probability as the input block length N approaches

infinity. If $A_{w,h}^N$ denotes the IOWE when the input block has length N, we introduce the following notation for the union bound (2.2) for systems of this type:

$$P_W^{\text{UB}} \stackrel{\text{def}}{=} \sum_{h=1}^{n} \left(\sum_{w=1}^{N} A_{w,h}^N \right) z^h. \tag{4.1}$$

Next we define, for each fixed $w \geq 1$ and $h \geq 1$,

$$\alpha(w,h) = \limsup_{N \to \infty} \log_N A_{w,h}^N. \tag{4.2}$$

It follows from this definition that if w and h are fixed,

$$A_{w,h}^N z^h = O(N^{\alpha(w,h)+\epsilon}) \qquad \text{as } N \to \infty,$$

for any $\epsilon > 0$. Thus if we define

$$\beta_M = \max_{h \geq 1} \max_{w \geq 1} \alpha(w,h). \tag{4.3}$$

it follows that for all w and h,

$$A_{w,h}^N z^h = O(N^{\beta_M + \epsilon}) \qquad \text{as } N \to \infty,$$

for any $\epsilon > 0$. The parameter β_M, which we shall call the *interleaving gain exponent* (IGE), was first introduced in [2] and [3] for parallel concatenation and later in [4] for serial concatenation. Extensive numerical simulations, and theoretical considerations that are not fully rigorous lead to the following conjecture about the behavior of the union bound for systems of the type shown in Figure 1.

The IGE Conjecture. *There exists a positive number γ_0, which depends on the q component convolutional codes and the tree structure of the overall system, but not on N, such that for any fixed $E_n/N_0 > \gamma_0$, as the block length N becomes large,*

$$P_W^{\text{UB}} = O(N^{\beta_M}) \tag{4.4}$$

Eq. (4.4) implies that if $\beta_M < 0$, then for a given $E_b/N_0 > \gamma_0$ the *word* error probability of the concatenated code decreases to zero as the input block size is increased. This is summarized by saying that there is *word error probability interleaving gain.*[1]

In [7], we discuss the calculation of $\alpha(w,h)$ and β_M for a concatenated system of the type depicted in Figure 1, using analytical tools introduced in [3] and [4]. For example, for the parallel concatenation of q codes, with $q-1$ interleavers, we have

$$\beta_M \leq -q + 2,$$

with equality if and only if each of the component codes is recursive. For a "classical" turbo code with $q = 2$, we have $\beta_M = 0$, so there is no word error probability interleaving gain. This suggests that the word error probability for classic turbo codes will not improve with input block size, which is in agreement with simulations.

[1] There is a similar conjecture for the bit error probability which we do not discuss in this paper. Suffice it to say that the interleaving gain exponent for bit error probability is $\beta_M - 1$.

As another example, consider the serial concatenation of two convolutional codes. If the inner code is recursive then,

$$\beta_M \leq -\left\lfloor \frac{d^o_{\text{free}} + 1}{2} \right\rfloor + 1,$$

where d^o_{free} is the minimum distance of the outer code. Therefore, for serial concatenated codes, if $d^o_f \geq 3$ there is interleaving gain for word error probability. (If the inner code is nonrecursive $\beta_M \geq 0$ and there is no interleaving gain.)

5. A Class of Simple Turbo-Like Codes.

In this section we will introduce a class of turbo-like codes which are simple enough so that we can prove the IGE conjecture. We call these codes *repeat and accumulate* (RA) codes. The general idea is shown in Figure 3. An information block of length N is repeated q times, scrambled by an interleaver of size qN, and then encoded by a rate 1 *accumulator*. The accumulator can be viewed as a truncated rate-1 recursive convolutional encoder with transfer function $1/(1+D)$, but we prefer to think of it as a block code whose input block $[x_1, \ldots, x_n]$ and output block $[y_1, \ldots, y_n]$ are related by the formula

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= x_1 + x_2 \\ y_3 &= x_1 + x_2 + x_3 \\ &\vdots \\ y_n &= x_1 + x_2 + x_3 + \cdots + x_n. \end{aligned} \tag{5.1}$$

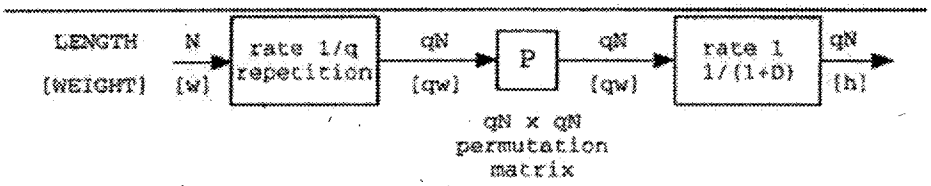


Figure 3. Encoder for a (qN, N) repeat and accumulate code. The numbers above the input-output lines indicate the length of the corresponding block, and those below the lines indicate the weight of the block.

To apply the union bound from Section 2 to the class of RA codes, we need the input-output weight enumerators for both the (qn, n) repetition code, and the (n, n) accumulator code. The outer repetition code is trivial: if the input block has length n, we have

$$A^{(w)}_{w,h} = \begin{cases} 0 & \text{if } h \neq qw \\ \binom{n}{w} & \text{if } h = qw. \end{cases} \tag{5.2}$$

The inner accumulator code is less trivial, but it is possible to show that (again assuming the input block has length n):

$$A_{w,h}^{(i)} = \binom{n-h}{\lfloor w/2 \rfloor} \binom{h-1}{\lceil w/2 \rceil - 1}. \tag{5.3}$$

It follows then from the general formula (3.1), that for the (qN, N) RA code represented by Figure 3, the ensemble IOWE is

$$\begin{aligned} A_{w,h}^{(N)} &= \sum_{h_1=0}^{qN} \frac{A_{w,h_1}^{(o)} A_{h_1,h}^{(i)}}{\binom{qN}{qw}} \\ &= \frac{\binom{N}{w}\binom{qN-h}{\lfloor qw/2 \rfloor}\binom{h-1}{\lceil qw/2 \rceil - 1}}{\binom{qN}{qw}}. \end{aligned} \tag{5.4}$$

From (5.4) it is easy to compute the parameters $\alpha(w,h)$ and β_M in (4.2) and (4.3). The result is

$$\alpha(w,h) = -\left\lceil \frac{(q-2)w}{2} \right\rceil \tag{5.5}$$

$$\beta_M = -\left\lceil \frac{(q-2)}{2} \right\rceil. \tag{5.6}$$

It follows from (5.6) that an RA code can have word error probability interleaving gain only if $q \geq 3$.

We are now prepared to use the union bound to prove the IGE conjecture for RA codes. In order to simplify the exposition as much as possible, we will assume for the rest of this section that $q = 4$, the extension to arbitrary $q \geq 3$ being straightforward but rather lengthy. For $q = 4$, (5.6) becomes $\beta_M = -1$, so the IGE conjecture is $P_W^{UB} = O(N^{-1})$ for $E_b/N_0 > \gamma_0$ in this instance.

The union bound (2.2) for the ensemble of $q = 4$ RA codes is, because of (5.4),

$$P_W^{UB} = \sum_{h=2}^{4N} \sum_{w=1}^{h/2} \frac{\binom{N}{w}\binom{4N-h}{2w}\binom{h-1}{2w-1}}{\binom{4N}{4w}} z^h. \tag{5.7}$$

Denote the (w,h)th term in the sum (5.7) by $T_N(w,h)$:

$$T_N(w,h) \stackrel{\text{def}}{=} A_{w,h} z^h = \frac{\binom{N}{w}\binom{4N-h}{2w}\binom{h-1}{2w-1}}{\binom{4N}{4w}} z^h.$$

Using standard techniques (e.g. [8, Appendix A]), it is possible to show that for all (w,h),

$$T_N(w,h) \leq D 2^{N\lceil F(x,y) + \log_2 z \rceil}, \tag{5.8}$$

where $D = 4/\sqrt{\pi}$ is a constant, $x = w/4N$, $y = h/4N$,

$$F(x,y) = \frac{-\frac{3}{4}H_2(4x) + (1-y)H_2(\frac{2x}{1-y}) + yH_2(\frac{2x}{y})}{y},$$

and $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function. The maximum of the function $F(x,y)$ in the range $0 \le 2x \le y \le 1-2x$ occurs at $(x,y) = (0.100, 0.371)$ and is 0.562281, so that if $\log_2 z < -0.562281$, the exponent in (5.8) will be negative.

Let us therefore assume that $\log_2 z < -0.562281$, which is equivalent to $E_b/N_0 = -(1/r)\ln z = -4\ln z \ge 4 \cdot \ln 2 \cdot 0.562281 = 1.559 = 1.928$ dB. If E is defined to be $E = -\log_2 z + 0.562281$, it follows from (5.8) for all w and h,

$$T_N(w,h) \le D2^{-hE}. \tag{5.9}$$

What (5.9) tells us is that if $E_b/N_0 > 1.928$ dB, most of the terms in the union bound (5.7) will tend to zero rapidly, as $N \to \infty$. The next step in the proof is to break the sum in (5.7) into two parts, corresponding to those terms for which (5.9) is helpful, and those for which it is not. To this end, define

$$h_N \stackrel{\text{def}}{=} \frac{3}{E}\log_2 N,$$

and write

$$\begin{aligned} P_W^{\text{UB}} &= \sum_{h=2}^{4N}\sum_{w=1}^{h/2} T_N(w,h) \\ &= \sum_{h=2}^{h_N}\sum_{w=1}^{h/2} T_N(w,h) + \sum_{h=h_N+1}^{4N}\sum_{w=1}^{h/2} T_N(w,h) \\ &= S_1 + S_2. \end{aligned}$$

It's easy to verify that when N is large enough, $A_{w+1,h}/A_{w,h} < 1$ for $h \le h_N$ and $w \le h/2 \le h_N/2$, which shows $A_{w,h}$ is a decreasing function of w for large N. Thus the sum S_1 can be overbounded as follows (we omit some details):

$$\begin{aligned} S_1 &= \sum_{h=2}^{h_N}\sum_{w=1}^{h/2} T_N(w,h) \\ &= \sum_{h=2}^{h_N} T_N(1,h) + \sum_{h=2}^{h_N}\sum_{w=2}^{h/2} T_N(w,h) \\ &= O(N^{-1}) + \sum_{h=2}^{h_N}\sum_{w=2}^{h/2} T_N(w,h) \\ &\le O(N^{-1}) + \sum_{h=2}^{h_N}\sum_{w=2}^{h/2} A_{2,h}z^h \\ &= O(N^{-1}) + \sum_{h=2}^{h_N}\sum_{w=2}^{h/2} O(h^3/N^2)z^h \\ &= O(N^{-1}) + O(h_N^5/N^2) \\ &= O(N^{-1}). \end{aligned}$$

For the sum S_2, we bound each term $T_N(w,h)$ by (5.9):

$$\begin{aligned} S_2 &= \sum_{h=h_N+1}^{4N} \sum_{w=1}^{h/2} T_N(w,h) \\ &\le \sum_{h_N+1}^{4N} \sum_{w=1}^{h/2} D2^{-hE} \\ &= D/2 \sum_{h_N+1}^{4N} h2^{-hE} \\ &\le D\frac{2^{-Eh_N}(h_N+1)}{(1-2^{-E})^2} \\ &= O(N^{-3}\log_2 N) \\ &= o(N^{-2}). \end{aligned}$$

We have therefore shown that for the ensemble of $q = 4$ RA codes, if $E_b/N_0 > 1.928$ dB,

$$P_W^{UB} = S_1 + S_2 = O(N^{-1}) + o(N^{-1}) = O(N^{-1}), \tag{5.10}$$

which as we saw above, is the IGE conjecture in this case.

Although the union bound gives a proof of the IGE conjecture for RA codes, the resulting value of γ_0 is by no means the best possible. Indeed, if we use the recent Viterbi-Viterbi improved union bound [9] to bound the sum S_2, we can lower the value of γ_0 considerably, e.g. for $q = 4$ from 1.928 dB to 0.313 dB. In Figure 4 and Table 1 we display our numerical results on RA codes. There we compare the "cutoff threshold" γ_0 for RA codes with q in the range $3 \le q \le 8$ using both the classical union bound and the Viterbi-Viterbi improved union bound to the cutoff threshold for the ensemble of all codes (i.e., "random codes") of a fixed rate. We believe that these values of γ_0 can be reduced still further, for example by using the bound of [6] instead of the Viterbi-Viterbi bound.

q	3	4	5	6	7	8
RA Codes (Union Bound)	2.200	1.928	1.798	1.721	1.670	1.631
Random Codes (Union Bound)	2.031	1.853	1.775	1.694	1.651	1.620
RA Codes (Viterbi Bound)	1.112	0.313	−0.125	−0.402	−0.592	−0.731
Random Codes (Viterbi Bound)	0.214	−0.224	−0.486	−0.662	−0.789	−0.885
Binary Shannon Limit	−0.495	−0.794	−0.963	−1.071	−1.150	−1.210

Table 1. Numerical data gleaned from Figure 4.

6. Performance of RA Codes with Iterative Decoding.

The results of this paper show that the performance of RA codes *with maximum-likelihood decoding* is very good. However, the complexity of ML decoding of RA

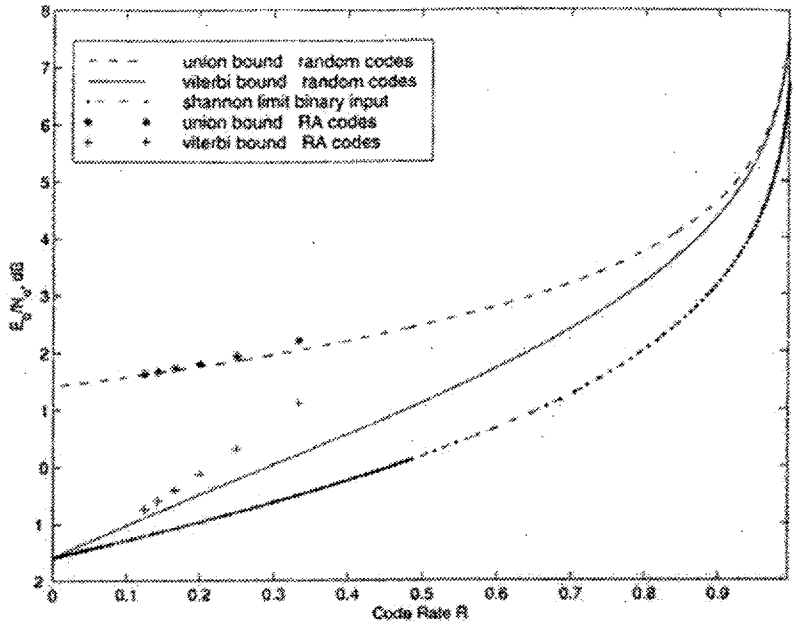


Figure 4. Comparing the RA code "cutoff threshold" to the cutoff rate of random codes using both the classical union bound and the Viterbi-Viterbi improved union bound.

codes, like that of all turbo-like codes, is prohibitively large. But an important feature of turbo-like codes is the availability of a simple iterative, message passing decoding algorithm that approximates ML decoding. We wrote a computer program to implement this "turbo-like" decoding for RA codes with $q = 3$ (rate 1/3) and $q = 4$ (rate 1/4), and the results are shown in Figure 5. We see in Figure 4, for example, that the empirical cutoff threshold for RA codes for $q = 3$ appears to be less than 1 dB, compared to the upper bound of 1.112 dB found in Table 1.

References.

1. C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: turbo codes," *Proc. 1993 IEEE International Conference on Communications*, Geneva, Switzerland (May 1993), pp. 1064–1070.
2. S. Benedetto and G. Montorsi, "Unveiling turbo codes: some results on parallel concatenated coding schemes", *IEEE Trans. on Inf. Theory*, vol. 42, no. 2 (March 1996), pp. 409–428.
3. S. Benedetto and G. Montorsi, "Design of parallel concatenated convolutional codes," *IEEE Transactions on Communications*, vol. 44, no. 5, (May 1996) pp. 591–600.
4. S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding," *IEEE Trans. on Information Theory*, vol. 44, no. 3, (May 1998), pp. 909–926.

Appendix C

UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE PATENT TRIAL AND APPEAL BOARD

In re: Inter Partes Review of:	:	
U.S. Pat. No. 7,116,710	:	
U.S. Pat. No. 7,421,032	:	
U.S. Pat. No. 7,421,781	:	
and U.S. Pat. No. 8,284,833	:	
Inventor: Hui Jin, et al	:	IPR No. Unassigned

Assignee: California Institute of Technology

Common Title: **Serial Concatenation of Interleaved Convolutional Codes Forming Turbo-Like Codes**

Mail Stop PATENT BOARD
Patent Trial and Appeal Board
U.S. Patent and Trademark Office
Box 1450
Alexandria, Virginia 22313-1450

Submitted Electronically via the Patent Review Processing System

DECLARATION OF PAUL H. SIEGEL

Declaration of Paul H. Siegel

I, **Paul H. Siegel**, declare as follows:

1. I am over the age of 18 and am legally competent to make this declaration. I make this declaration based upon my own personal knowledge.

2. I am currently a Professor at the University of California, San Diego, in the Department of Electrical and Computer Engineering, the Jacobs School of Engineering, and the Center for Memory and Recording Research.

3. I have been asked to provide a statement of certain facts related to work I did together with Henry D. Pfister relating to the serial concatenation of rate-1 codes through uniform random interleavers.

4. Beginning in early 1999, I collaborated with Dr. Pfister, then a Ph.D student of mine, to show that improved error-correcting codes can be constructed from simple components, *e.g.*, by serially concatenating an arbitrary outer code of rate $r < 1$ and m identical rate-1 inner codes.

5. The result of that collaboration was a presentation at the 1999 Allerton Conference on Communications, Control and Computing, in Allerton, Illinois ("1999 Allerton Conference") in September 1999. This conference was held September 22-24, 1999, and was open to the public for attendance. Any person who wanted to attend and was able to pay the attendance fee could

attend. The 1999 Allerton Conference was considered one of two primary conferences on the topic of iterative decoding during the time. The 1999 Allerton conference itself, as well as the proceedings that took place there, were publicized and generally known to those who were interested in topics relating to error-correcting codes and iterative decoding.

6. At the 1999 Allerton Conference, I presented a series of slides relating to my work with Dr. Pfister on a class of codes based on serial concatenation of a rate-r code with $m \geq 1$, uniformly interleaved rate-1 codes. In these slides, I discussed an example of these codes, called a "Repeat-Accumulate-Accumulate" (or "RAA") code. An RAA code is is similar to an RA code (Divsalar, et al., Allerton '98) to which an additional accumulator has been added. I compared RAA codes to RA codes and concluded that the RAA codes were able to achieve a lower word-error-probability than the RA codes, demonstrating that the performance of RA codes could be improved by including an additional accumulator. This presentation was entitled "The Serial Concatenation of Rate-1 Codes Through Uniform Random Interleavers," and I presented it in the IIA: Coding Theory: Iterative Decoding and Turbo Codes Session on September 22, 1999, the first day of the conference. A true and accurate copy of the slides that I presented at the 1999 Allerton Conference is attached as Exhibit 1.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information or belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: 11/9/16

Paul H. Siegel

Paul H. Siegel

Exhibit 1

The Serial Concatenation of Rate-1 Codes Through Uniform Random Interleavers

Henry Pfister and Paul H. Siegel
Signal Transmission and Recording (STAR) Lab
University of California, San Diego

{hpfister, psiegel}@ucsd.edu

Allerton Conference
September 22-24 , 1999

Outline

- Union Bounds and Code Performance
- Serial Concatenation and Repeat-Accumulate (RA) Codes
- Serial Concatenation of Rate-1 Codes
- Repeat-Accumulate-Accumulate (RAA) Codes
- Summary

Union Bounds on Performance

- Rate $r = k/n$, linear block code C
- Input Output Weight Enumerator Function (IOWEF) :

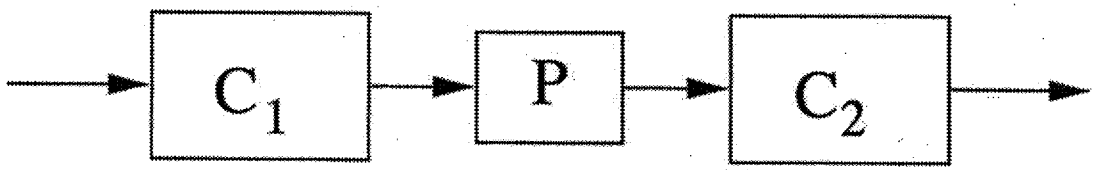
$$A_{w,h} \stackrel{\text{def}}{=} \#\text{ codewords, input weight } w\text{, output weight } h$$

- Union bound on word error probability P_W
  (binary-input, memoryless channel, maximum-likelihood decoding):

$$P_W \leq \sum_{h=1}^{n} \sum_{w=1}^{k} A_{w,h} z^h$$

- z is channel dependent; e.g., for Gaussian channel, $z = e^{-r(E_b/N_0)}$.
- For ensembles, replace $A_{w,h}$ by average IOWEF $\overline{A_{w,h}}$

Serial Concatenation through a Uniform Interleaver

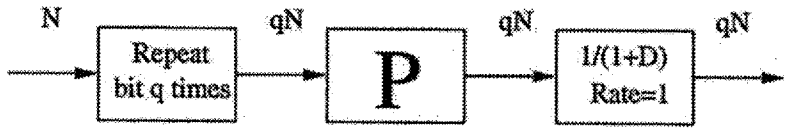


- Let C_1, C_2 be (n_1, k_1), (n_2, k_2) linear block codes with $n_1 = k_2$, and IOWEFs $A^{(1)}_{w,h}$, $A^{(2)}_{w,h}$.

- Let C be the (n_2, k_1) code obtained by serial concatenation of C_1 and C_2 through a uniform interleaver of size n_1, with average IOWEF $A_{w,h}$:

$$A_{w,h} = \sum_{h_1=0}^{n_1} A^{(1)}_{w,h_1} \cdot \frac{A^{(2)}_{h_1,h}}{\binom{n_1}{h_1}}$$

Repeat-Accumulate (RA) Codes
(Divsalar, et al., Allerton'98)



- Repeat input block $x_1 x_2 \cdots x_N$ a total of q times.
- Permute with random interleaver P of size $n = qN$.
- Accumulate over block:

$$u_1 u_2 \cdots u_n \longrightarrow v_1 v_2 \cdots v_n$$

$$v_1 = u_1$$

$$v_2 = u_1 + u_2$$

$$\vdots$$

$$v_n = u_1 + \cdots + u_n$$

A Simple Example

- Rate-1, $(n, k) = (3, 3)$ Accumulate code:

Input Bits	Input Weight w	Output Bits	Output Weight h
000	0	000	0
001	1	001	1
010	1	011	2
100	1	111	3
011	2	010	1
101	2	110	2
110	2	100	1
111	3	101	2

11

RA Weight Enumerator
(Divsalar, et al., Allerton'98)

- (qN, N) Repeat Code

$$A_{w,h}^{(1)} = \begin{cases} 0 & if \quad h \neq qw \\ \begin{pmatrix} qN \\ w \end{pmatrix} & if \quad h = qw \end{cases}$$

(n, n) Accumulate Code (cf. 'Oberg and Siegel, Allerton'98)

$$A_{w,h}^{(2)} = \begin{pmatrix} n-h \\ \lfloor w/2 \rfloor \end{pmatrix} \begin{pmatrix} h-1 \\ \lceil w/2 \rceil - 1 \end{pmatrix}$$

- (qN, N) RA Code

$$\overline{A}_{w,h} = \frac{\binom{N}{w}\binom{qN-h}{\lfloor qw/2 \rfloor}\binom{h-1}{\lceil qw/2 \rceil - 1}}{\binom{qN}{qw}} \, if \; w, h > 0$$

RA Coding Theorem
(Divsalar, et al., Allerton'98)

- **Theorem:** For $q \geq 3$, there exists $\gamma_q > 0$ such that, for any $E_b/N_0 > \gamma_q$, as the block length N becomes large,

$$P_W^{UB} = O(N^{\beta}),$$

where $\beta = -\lceil \frac{q-2}{2} \rceil$. Hence $P_W \to 0$ as $N \to \infty$.

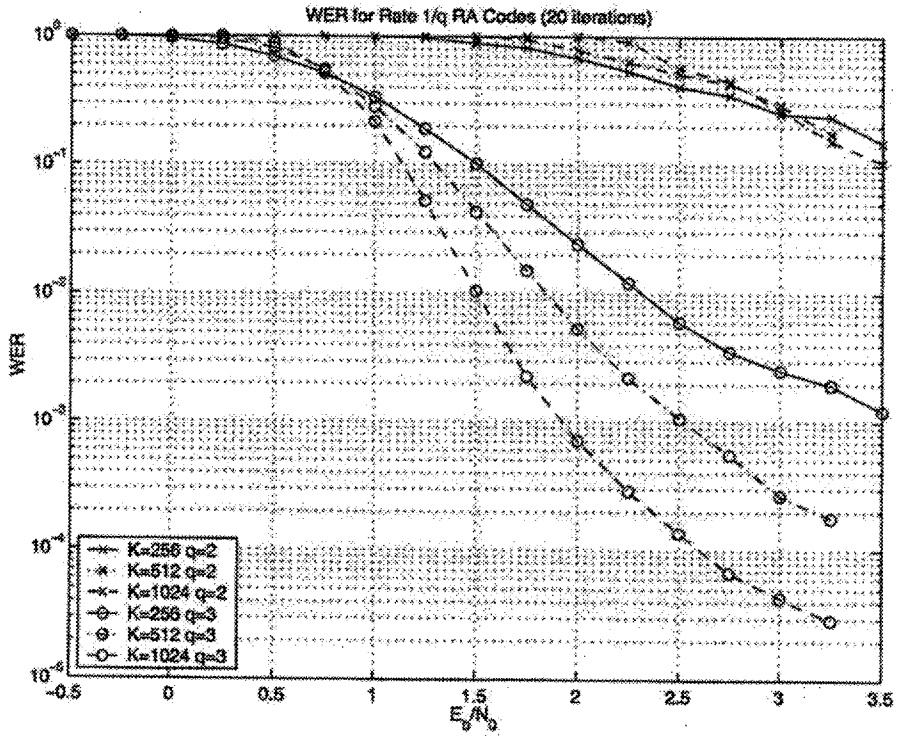
- Example γ_q estimates (from Viterbi-Viterbi Bound):

$$\gamma_3 \approx 1.112 \text{ dB}$$

$$\gamma_4 \approx 0.313 \text{ dB}$$

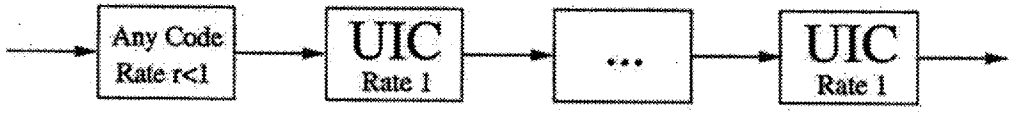
Performance of RA Codes

- RA codes, $r = \frac{1}{3}$ and $r = \frac{1}{2}$, with iterative decoding.



Serial Concatenation

- Consider an encoder architecture of the following form:



- UIC represents a Uniform Interleaver in cascade with a rate-1 Code (e.g., an (n, n) Accumulate Code).

- We want to characterize the average IOWEF when the code is concatenated with m UIC's.

Serial Concatenation with UIC's

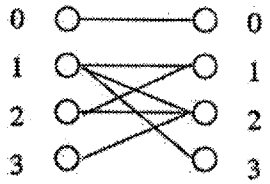
- Let IOWEF for rate-1, (n, n) code be $A = [A_{i,j}], \quad 0 \leq i, j \leq n.$

- Define IOW Transition Probabilities (IOWTP): $P = [P_{i,j}], \quad 0 \leq i, j \leq n,$

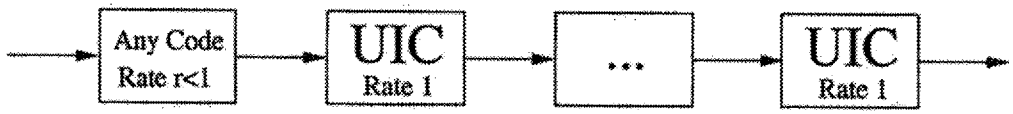
$$P_{i,j} = Pr(h = j | w = i) = \frac{A_{i,j}}{\binom{n}{i}}.$$

- Rate-1, $(n, k) = (3, 3)$ Accumulate code:

$$[P_{i,j}] = [Pr(h = j | w = i)] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.\overline{3} & 0.\overline{3} & 0.\overline{3} \\ 0 & 0.\overline{6} & 0.\overline{3} & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$



Serial Concatenation with UIC's



- Outer (n, k) code C, IOWEF $C_{w,h}$.
- Cascade of m, rate-1 (n, n) UIC's, each with IOWTP $P = [P_{i,j}]$.
- Average IOWEF $\overline{A_{w,h}}$ of serial concatenation:

$$\overline{A_{w,h}} = \sum_{h_1=0}^{n} C_{w,h_1} \left[P^m\right]_{h_1,h}$$

Perron-Frobenius Theory

- **Theorem:** An irreducible stochastic matrix P has a unique stationary distribution π s.t. $\pi P = \pi$ and $\sum \pi_i = 1$.

- **Theorem:** An irreducible aperiodic (primitive) stochastic matrix P with unique stationary distribution π satisfies:

$$\lim_{m \to \infty} P^m = \begin{bmatrix} \pi \\ \vdots \\ \pi \end{bmatrix}$$

- A linear rate-1 code is *primitive* if $P^+ = [P_{i,j}]$, $i,\ j \neq 0,$ is a primitive matrix.

A Simple Example (cont.)

- Rate-1, $(n, k) = (3, 3)$ Accumulate Code:

$$\begin{pmatrix} 0 & \frac{3}{7} & \frac{3}{7} & \frac{1}{7} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.\overline{3} & 0.\overline{3} & 0.\overline{3} \\ 0 & 0.\overline{6} & 0.\overline{3} & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1+2}{7} & \frac{1+1+1}{7} & \frac{1}{7} \end{pmatrix}$$

$$\lim_{m \to \infty} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.\overline{3} & 0.\overline{3} & 0.\overline{3} \\ 0 & 0.\overline{6} & 0.\overline{3} & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}^m = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3/7 & 3/7 & 1/7 \\ 0 & 3/7 & 3/7 & 1/7 \\ 0 & 3/7 & 3/7 & 1/7 \end{pmatrix}$$

Stationary Distributions and Rate-1 Codes

- **Proposition:** Consider the SC of m primitive rate-1 codes through uniform interleavers. For non-zero input weights and as $m \to \infty$, the output weight distribution is independent of the input weight distribution and is

$$\pi_h = \frac{\binom{n}{h}}{2^n - 1}, \quad for \quad h \neq 0.$$

- The ensemble averaged OWEF for any rate $r < 1$ code SC with $m \to \infty$ primitive rate-1 codes is

$$\overline{A_h} = \frac{2^{rn} - 1}{2^n - 1}\binom{n}{h}$$

Iterated Rate-1 Codes Make Good Codes

- **Proposition:** For a code randomly chosen from an ensemble with average OWE $\overline{A_h}$,

$$Pr(d_{min} < d) \leq \sum_{h=1}^{d-1} \overline{A_h}.$$

(cf. Gallager, 1963)

- For length n, rate $r < 1$,and $\epsilon > 0$, let $d^*(n, r, \epsilon)$ be the largest value of d satisfying

$$\sum_{h=0}^{d-1} \binom{n}{h} \leq \frac{2^n}{2^{rn} - 1} \epsilon.$$

Iterated Rate-1 Codes Make Good Codes

- **Proposition:** For the ensemble of length-n codes obtained by serial concatenation of a rate-r linear block code with $m \longrightarrow \infty$ primitive rate-1 linear codes through uniform random interleavers,

$$Pr(d_{min} < d^*(n, r, \epsilon)) \leq \epsilon.$$

- **Corollary:** Let $\epsilon = (2^{(1-r)n} - 1)(2^{rn} - 1)/(2^n - 1)$. Then,

$$Pr(d_{min} < d^*) \leq \epsilon < 1,$$

where $d^*$is the largest value of d satisfying

$$\sum_{h=0}^{d-1} \left(\begin{array}{c} n \\ h \end{array} \right) \leq 2^{n(1-r)} = 2^{n-k}$$

(i.e., at least one code satisfies the Gilbert-Varshamov Bound).

Bounds for CA^m Codes

- Rate r<1, outer code C.
- Serial concatenation with m uniformly interleaved Accumulate codes.
- Compute largest value d^* satisfying

$$\sum_{h=1}^{d-1} \overline{A_h} < \frac{1}{2}.$$

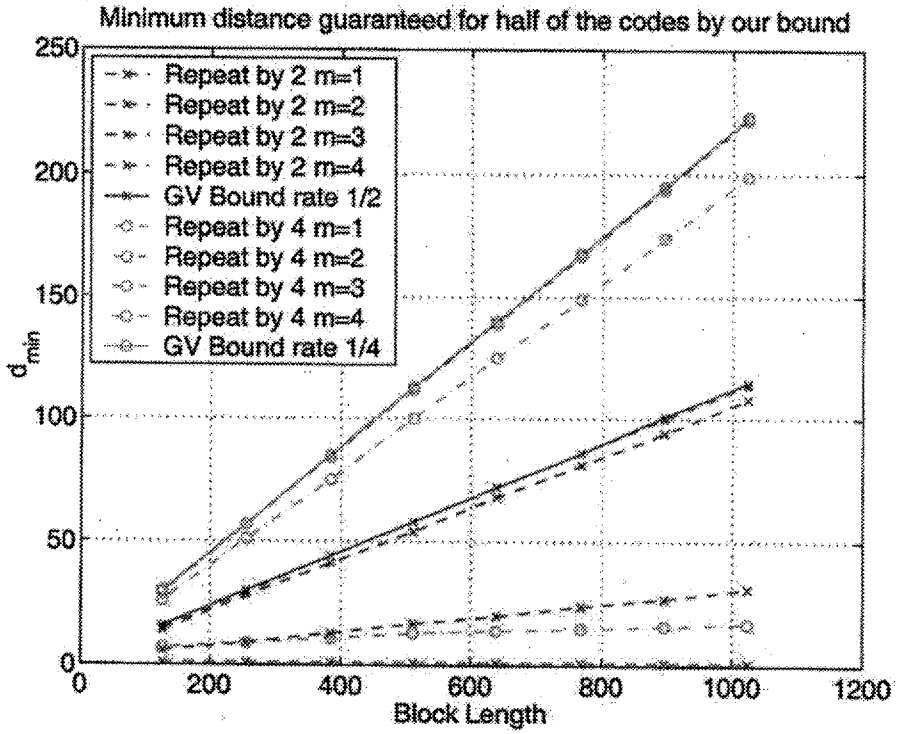
- Then, by the bound,

$$Pr(d_{min} < d^*) < \frac{1}{2},$$

(i.e., at least half the codes satisfy $d_{min} \geq d^*$).

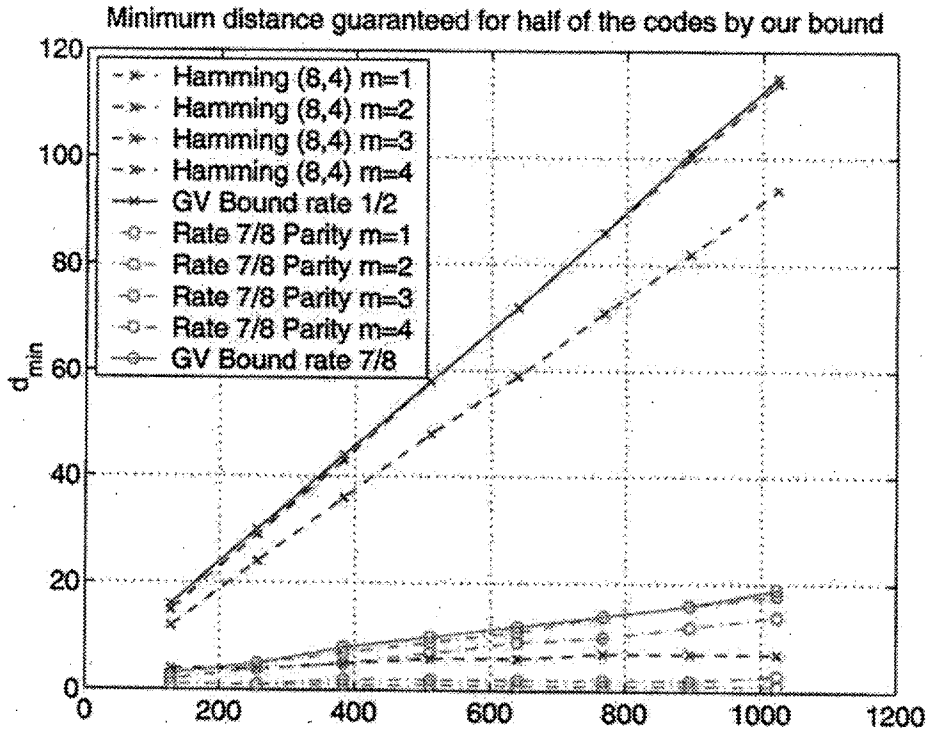
Numerical Results

- CA^m bounds for Repeat-2 and Repeat-4 codes, $1 \leq m \leq 4$.



Numerical Results

- For (8,4) Hamming and rate 7/8 parity-check codes, $1 \leq m \leq 4$.



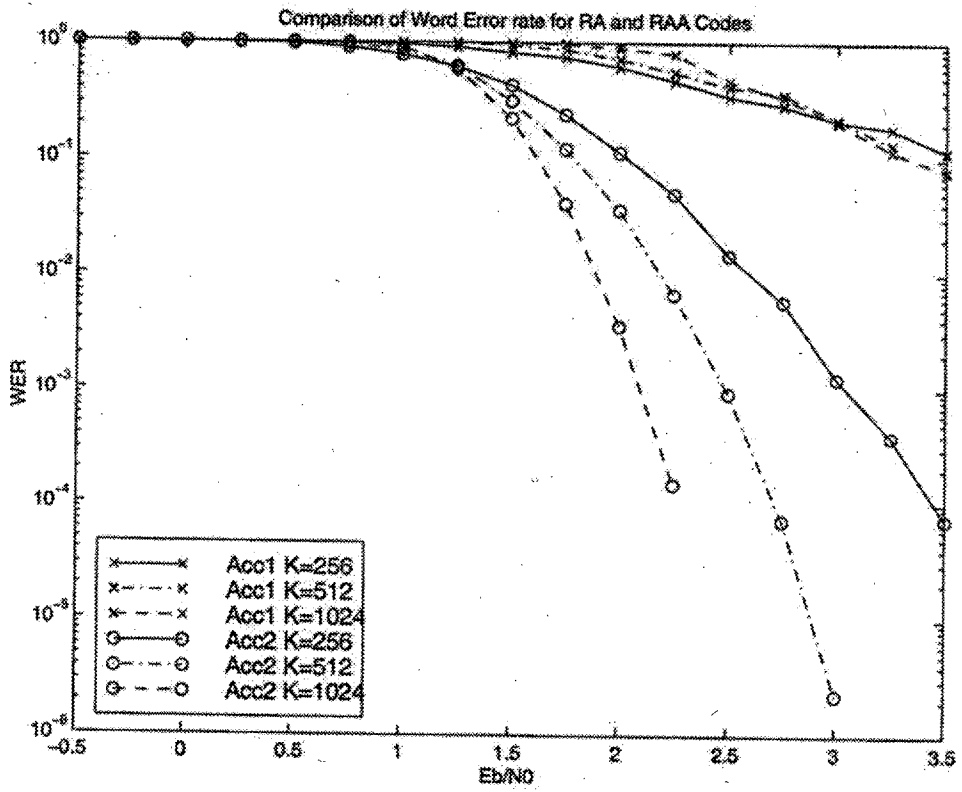
Rate 1/2, RAA Codes

- For rate $r = 1/2$ RA code, analysis and simulations suggest that ***there is no word-error-probability interleaving gain***.

- For rate $r = 1/2$ RAA code, analysis and simulations suggest that ***there is...***

Rate 1/2, RA vs. RAA

- Word-error-probability P_W , from computer simulation.



Summary

- We investigated a new class of codes based upon serial concatenation of a rate-r code with $m \geq 1$, uniformly interleaved rate-1 codes.

- We analyzed the output weight enumerator function $\overline{A_h}$ for finite m, as well as asymptotically for $m \longrightarrow \infty$.

- We evaluated the "goodness" of these ensembles in terms of the Gilbert-Varshamov Bound.

- We compared $r = 1/2$ RAA codes to RA codes.

  Question: For $r = 1/2$, RAA codes, is there a γ such that , for $E_b/N_0 > \gamma$, $P_W \longrightarrow 0$ as $N \longrightarrow \infty$?