Information-Theoretic Foundations of Mismatched Decoding

Other titles in Foundations and Trends® in Communications and Information Theory

Group Testing: An Information Theory Perspective
Matthew Aldridge, Oliver Johnson and Jonathan Scarlett
ISBN: 978-1-68083-596-0

Sparse Regression Codes
Ramji Venkataramanan, Sekhar Tatikonda and Andrew Barron
ISBN: 978-1-68083-580-9

Fundamentals of Index Coding
Fatemeh Arbabjolfaei and Young-Han Kim
ISBN: 978-1-68083-492-5

Community Detection and Stochastic Block Models
Emmanuel Abbe
ISBN: 978-1-68083-476-5

Information-Theoretic Foundations of Mismatched Decoding

Jonathan Scarlett
National University of Singapore, Singapore
scarlett@comp.nus.edu.sg

Albert Guillén i Fàbregas
ICREA, Universitat Pompeu Fabra, Spain
University of Cambridge, UK
guillen@ieee.org

Anelia Somekh-Baruch
Bar-Ilan University, Israel
somekha@biu.ac.il

Alfonso Martinez
Universitat Pompeu Fabra, Spain
alfonso.martinez@ieee.org

the essence of knowledge
Boston — Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

J. Scarlett, A. Guillén i Fàbregas, A. Somekh-Baruch and A. Martinez. *Information-Theoretic Foundations of Mismatched Decoding*. Foundations and Trends® in Communications and Information Theory, vol. 17, no. 2–3, pp. 149–401, 2020.

ISBN: 978-1-68083-713-1

Foundations and Trends® in Communications and Information Theory

Volume 17, Issue 2–3, 2020

Editorial Scope

Topics

Foundations and Trends® in Communications and Information Theory publishes survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2020, Volume 17, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328 . Also available as a combined paper and online subscription.

Contents

Notation **3**

1 Introduction **6**
1.1 Problem Setup 7
1.2 Applications of Mismatched Decoding 10
1.3 Other Channel Uncertainty Models 16
1.4 Outline of the Monograph 18

2 Discrete Memoryless Channels **22**
2.1 Introduction 22
2.2 Properties of the Mismatch Capacity 23
2.3 Achievable Rates: GMI and LM Rate 25
2.4 Examples 32
2.5 Properties of the GMI and LM Rate 40
2.6 Proofs of Achievable Rates 52
2.7 Multi-Letter Improvements 60

3 Continuous-Alphabet Memoryless Channels **62**
3.1 Introduction 62
3.2 Problem Setup 63
3.3 i.i.d. Random Coding and the GMI 64
3.4 Cost-Constrained Random Coding and the LM Rate 66

3.5 Cost-Constrained Random Coding with Multiple Auxiliary Costs 68
3.6 Proofs of Achievable Rates 70
3.7 Ensemble Tightness Results 72
3.8 Examples 76

4 Mismatch in Rate Distortion Theory 90
4.1 Introduction 90
4.2 Problem Setup 91
4.3 Achievability Result 94
4.4 Examples 96
4.5 Achievability Proof 99
4.6 Multi-Letter Improvement and Converse 103
4.7 Mismatched Random Codebooks with Optimal Encoding . 105
4.8 Gaussian Compression for Non-Gaussian Sources 114

5 Multiple-Access Channels 118
5.1 Introduction 118
5.2 Problem Setup 119
5.3 Achievable Rate Region 120
5.4 Examples 123
5.5 Proof of Achievable Rate Region 126
5.6 Ensemble Tightness 131
5.7 Dual Expressions and Continuous Alphabets 136

6 Multi-User Coding Techniques for Single-User Channels 140
6.1 Introduction 140
6.2 Expurgated Parallel Coding 141
6.3 Superposition Coding 145
6.4 Refined Superposition Coding 156

7 Error Exponents 169
7.1 Introduction 169
7.2 Random-Coding Exponents for DMCs 170
7.3 Random-Coding Exponents with Continuous Alphabets . . 175
7.4 Expurgated Exponents 178

7.5 Attaining the Best of Both Exponents 180
7.6 Numerical Example 185
7.7 Error Exponents for Multi-User Coding Techniques 186

8 Upper Bounds on the Mismatch Capacity **190**
8.1 Introduction 190
8.2 Initial Results 191
8.3 The Csiszár–Narayan Conjecture 193
8.4 Other Decoding Rules 203
8.5 Further Multi-Letter Upper Bounds 207
8.6 A Single-Letter Upper Bound 210

9 Overview of Other Topics **217**
9.1 Alternative Achievable Rates and Capacity Notions 217
9.2 Continuous and Fading Channels 218
9.3 Refined Asymptotics 219
9.4 Modified Channel Coding Settings 220
9.5 Other Decoders 221
9.6 Mismatch in Practical Codes 223

10 Conclusion **224**

Appendix **229**

A Appendix **230**
A.1 Continuity Arguments 230
A.2 Other Omitted Proofs 238

Acknowledgments **243**

References **244**

Information-Theoretic Foundations of Mismatched Decoding

Jonathan Scarlett[1], Albert Guillén i Fàbregas[2],
Anelia Somekh-Baruch[3] and Alfonso Martinez[4]

[1] *National University of Singapore, Singapore; scarlett@comp.nus.edu.sg*
[2] *ICREA, Universitat Pompeu Fabra, Spain, and University of Cambridge, UK; guillen@ieee.org*
[3] *Bar-Ilan University, Israel; somekha@biu.ac.il*
[4] *Universitat Pompeu Fabra, Spain; alfonso.martinez@ieee.org*

ABSTRACT

Shannon's channel coding theorem characterizes the maximal rate of information that can be reliably transmitted over a communication channel when optimal encoding and decoding strategies are used. In many scenarios, however, practical considerations such as channel uncertainty and implementation constraints rule out the use of an optimal decoder. The mismatched decoding problem addresses such scenarios by considering the case that the decoder cannot be optimized, but is instead fixed as part of the problem statement. This problem is not only of direct interest in its own right, but also has close connections with other long-standing theoretical problems in information theory.

In this monograph, we survey both classical literature and recent developments on the mismatched decoding problem, with an emphasis on achievable random-coding rates for memoryless channels. We present two widely-considered

Jonathan Scarlett, Albert Guillén i Fàbregas, Anelia Somekh-Baruch and Alfonso Martinez (2020), "Information-Theoretic Foundations of Mismatched Decoding", Foundations and Trends® in Communications and Information Theory: Vol. 17, No. 2–3, pp 149–401. DOI: 10.1561/0100000101.

achievable rates known as the generalized mutual information (GMI) and the LM rate, and overview their derivations and properties. In addition, we survey several improved rates via multi-user coding techniques, as well as recent developments and challenges in establishing upper bounds on the mismatch capacity, and an analogous mismatched encoding problem in rate-distortion theory. Throughout the monograph, we highlight a variety of applications and connections with other prominent information theory problems.

Notation

(Introduced in Section 1)

$\mathcal{X}$, $\mathcal{Y}$	Input and output alphabets
m, $\hat{m}$	Message and its estimate
M	Number of codewords
n	Block length
R	Coding rate
W, W^n	Channel transition law and its n-letter extension
q, q^n	Decoding metric and its n-letter extension
$\mathcal{C}$	Codebook
$\boldsymbol{x}^{(m)}$	m-th codeword
p_{e}	Average error probability
$p_{\mathrm{e,max}}$	Maximal error probability
C_{M}	Mismatch capacity
C	Matched capacity

(Introduced in Section 2)

H_2	Binary entropy function
$\boldsymbol{X}$, $\boldsymbol{Y}$	Transmitted codeword and received sequence
$\overline{\boldsymbol{X}}$	Non-transmitted codeword
$P_{\boldsymbol{X}}$	Codeword distribution for random coding
$\widehat{P}_{\boldsymbol{x}}$	Empirical distribution of $\boldsymbol{x}$
I_{GMI}, I_{LM}	Generalized mutual information (GMI) and LM rate
C_{GMI}, C_{LM}	Rates with optimized input distributions
Q_X	Input distribution
$\mathcal{P}$	Set of all distributions on a given alphabet
$\mathcal{P}_n$	Set of all empirical distributions on a given alphabet
$\widetilde{P}_{XY}$	Auxiliary distribution in primal rates
s, a	Auxiliary parameters in dual rates
b	Auxiliary function on the output alphabet
$\mathcal{T}^n$	Type class
$Q_{X,n}$	Type approximating Q_X
$\bar{p}_{\mathrm{e}}$	Random-coding error probability
i_s^n	Information density quantity used in proofs
$C_{\mathrm{GMI}}^{(k)}$, $C_{\mathrm{LM}}^{(k)}$	Multi-letter achievable rates
$C_{\mathrm{GMI}}^{(\infty)}$, $C_{\mathrm{LM}}^{(\infty)}$	Limiting multi-letter achievable rates

(Introduced in Section 3)

c, Γ	System cost function and threshold
a, a_l	Auxiliary cost functions
c^n, a^n, a_l^n	Additive multi-letter extensions of cost functions
ϕ_c, ϕ_a, ϕ_l	Means of cost functions
Ω_n	Normalizing constant for cost-constrained ensemble
$\mathcal{D}_n$	Constraint set for cost-constrained ensemble
I'_{LM}	Fixed-cost LM rate
r_l, $\bar{r}_l$	Auxiliary parameters for cost-constrained ensemble
L	Number of auxiliary costs
μ, σ^2	Noise mean and variance in additive channels

(Introduced in Section 4)

Π_X, Π_X^n	Source distribution and n-letter extension
$\widehat{\boldsymbol{x}}^{(m)}$	Codeword in rate-distortion theory
d_0, d_1	Encoding metric and true distortion measure
d_0^n, d_1^n	Additive n-letter extensions of distortion functions
D_1, D_1^*	Distortion threshold and distortion-rate function
$Q_{\widehat{X}}$	Auxiliary distribution in rate-distortion theory
$\overline{D}_1$, $\overline{D}_1^{(k)}$	Achievable distortion and multi-letter version
$\widetilde{\mathcal{P}}$	Constraint set in achievable distortion expression
d, D	Distortion measure and level when $d_0 = d_1$
R^*_{Matched}, D^*_{Matched}	Matched rate-distortion function and distortion-rate function
D_{min}, D_{prod}	Extreme values of distortion level
$\overline{R}_{\text{iid}}$, $\overline{R}_{\text{cc}}$	Random coding rate-distortion functions
σ^2	Source variance

(Introduced in Section 5)

M_1, M_2	Multiple-access channel codebook sizes
R_1, R_2	Multiple-access channel rates
Q_1, Q_2	Multiple-access channel input distributions
$P_{\boldsymbol{X}_1}$, $P_{\boldsymbol{X}_2}$	Multiple-access channel codeword distributions

(Introduced in Section 6)

M_0, M_1, $\{M_{1u}\}$	Superposition coding codebook sizes
R_0, R_1, $\{R_{1u}\}$	Superposition coding rates
$\{n_u\}$	Refined superposition coding sub-block lengths
Q_{UX}	Superposition coding input distribution
$P_{\boldsymbol{UX}}$	Superposition coding codeword distribution
$\mathcal{U}$	Auxiliary alphabet for superposition coding
$\boldsymbol{U}$	Auxiliary codeword for superposition coding
$p_{\text{e},\nu}$, $\overline{p}_{\text{e},\nu}$	Multi-user coding error probabilities

(Introduced in Section 7)

$E_{\mathrm{r}}^{\mathrm{iid}}$, E_0^{iid}	i.i.d. exponent and Gallager function
$E_{\mathrm{r}}^{\mathrm{cc}}$, E_0^{cc}	Constant-composition exponent and Gallager function
$E_{\mathrm{r}}^{\mathrm{cost}}$, E_0^{cost}	Cost-constrained exponent and Gallager function
$E_{\mathrm{ex}}^{\mathrm{cc}}$, $E_{\mathrm{x}}^{\mathrm{cc}}$	Constant-composition expurgated exponent and Gallager function
ρ	Dual error exponent parameter
E_{CK}	Csiszár-Körner exponent
E_{RGV}	Random Gilbert-Varshamov exponent
d, Δ	RGV distance function and parameter

(Introduced in Section 8)

ϵ_k	Error probability with block length k
η_k	Minimal difference of k-letter log-metric values
$q_{\max}$	Maximal metric value
$q_{\max}^k(y^k)$	Maximal metric value for fixed output sequence
B	Upper bound on $\lvert \log q(x,y) \rvert$
$\mathcal{A}_k$	Output vectors with a unique metric maximizer
Φ_k	Conditional probability of random codeword exceeding metric value
$\mathcal{X}_q^*$	Set of inputs with maximal metric difference
$\mathcal{M}_{\max}$, $\mathcal{M}_{\max,n}$	Set of maximal conditional joint distributions and type-based variant
$\overline{R}$	Single-letter mismatch capacity upper bound
$\mathcal{G}_{\boldsymbol{x}}$	Bipartite graph associated with two channels
$\mathcal{E}_{\boldsymbol{x}}$	Edge set associated with two channels

References

[1] C. E. Shannon, "A mathematical theory of communication," pp. 379–423, *Bell Sys. Tech. Journal*, vol. 27, Jul. 1948.

[2] A. Alvarado, T. Fehenberger, B. Chen, and F. Willems, "Achievable information rates for fiber optics: Applications and computations," pp. 424–439, *J. Lightwave Tech.*, vol. 36, no. 2, Jan. 2018.

[3] W. Huleihel, S. Salamatian, N. Merhav, and M. Médard, "Gaussian intersymbol interference channels with mismatch," pp. 4499–4517, *IEEE Trans. Inf. Theory*, vol. 65, no. 7, Jul. 2019.

[4] A. Lapidoth and S. Shamai, "Fading channels: How perfect need 'perfect side information' be?" Pp. 1118–1134, *IEEE Trans. Inf. Theory*, vol. 48, no. 5, May 2002.

[5] P. Elias, "Coding for two noisy channels," in *London Symp. Inf. Theory*, pp. 61–74, 1955.

[6] R. Gallager, *Information Theory and Reliable Communication.* John Wiley & Sons, 1968.

[7] I. Csiszár and P. Narayan, "Channel capacity for a given decoding metric," pp. 35–43, *IEEE Trans. Inf. Theory*, vol. 45, no. 1, Jan. 1995.

[8] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," pp. 1953–1967, *IEEE Trans. Inf. Theory*, vol. 40, no. 6, Nov. 1994.

[9] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," pp. 1520–1529, *IEEE Trans. Inf. Theory*, vol. 42, no. 5, Sep. 1996.

[10] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[11] H. Weingarten, Y. Steinberg, and S. Shamai, "Gaussian codes and weighted nearest neighbor decoding in fading multiple-antenna channels," pp. 1665–1686, *IEEE Trans. Inf. Theory*, vol. 50, no. 8, Aug. 2004.

[12] A. Guillén i Fàbregas, A. Martinez, and G. Caire, "Bit-interleaved coded modulation," pp. 1–153, *Found. Trends Comms. Inf. Theory*, vol. 5, no. 1–2, Jan. 2008.

[13] A. Martinez, A. Guillén i Fàbregas, G. Caire, and F. Willems, "Bit-interleaved coded modulation revisited: A mismatched decoding perspective," pp. 2756–2765, *IEEE Trans. Inf. Theory*, vol. 55, no. 6, Jun. 2009.

[14] N. Binshtok and S. Shamai, "Integer metrics for binary input symmetric output memoryless channels," pp. 1636–1645, *IEEE Trans. Comms.*, vol. 47, no. 11, Nov. 1999.

[15] J. Salz and E. Zehavi, "Decoding under integer metrics constraints," pp. 307–317, *IEEE Trans. Comms.*, vol. 43, no. 234, Feb. 1995.

[16] M. Secondini and E. Forestieri, "Scope and limitations of the non-linear Shannon limit," pp. 893–902, *J. Lightwave Tech.*, vol. 35, no. 4, Feb. 2017.

[17] H. Ghozlan and G. Kramer, "Models and information rates for Wiener phase noise channels," pp. 2376–2393, *IEEE Trans. Inf. Theory*, vol. 63, no. 4, 2017.

[18] R. Ahlswede, N. Cai, and Z. Zhang, "Erasure, list, and detection zero-error capacities for low noise and a relation to identification," pp. 55–62, *IEEE Trans. Inf. Theory*, vol. 42, no. 1, Jan. 1996.

[19] J. Körner and A. Orlitsky, "Zero-error information theory," pp. 2207–2229, *IEEE Trans. Inf. Theory*, vol. 44, no. 6, Oct. 1998.

[20] C. E. Shannon, "The zero error capacity of a noisy channel," pp. 8–19, *IRE Trans. Inf. Theory*, vol. 2, no. 3, Sep. 1956.

[21] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," pp. 2148–2177, *IEEE Trans. Inf. Theory*, vol. 44, no. 6, Oct. 1998.

[22] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.

[23] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[24] Y. Feldman and A. Somekh-Baruch, "Channels with state information and mismatched decoding," in *IEEE Inf. Theory Workshop*, pp. 419–423, 2016.

[25] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," pp. 1229–1241, *Ann. Math. Stats.*, vol. 30, no. 4, 1959.

[26] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacities of certain channel classes under random coding," pp. 558–567, *Ann. Math. Stats.*, vol. 31, no. 3, 1960.

[27] I. Csiszar and P. Narayan, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," pp. 181–193, *IEEE Trans. Inf. Theory*, vol. 34, no. 2, Feb. 1988.

[28] I. Stiglitz, "Coding for a class of unknown channels," pp. 189–195, *IEEE Trans. Inf. Theory*, vol. 12, no. 2, Apr. 1966.

[29] R. Gallager, "Fixed composition arguments and lower bounds to error probability," 1992, [Online]. Available: http://web.mit.edu/gallager/www/notes/notes5.pdf.

[30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., 2006.

[31] J. Hui, "Fundamental issues of multiple accessing," *PhD Thesis*. MIT, 1983.

[32] I. Csiszár and J. Körner, "Graph decomposition: A new key to coding theorems," pp. 5–12, *IEEE Trans. Inf. Theory*, vol. 27, no. 1, Jan. 1981.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[34] R. Gallager, "A simple derivation of the coding theorem and some applications," pp. 3–18, *IEEE Trans. Inf. Theory*, vol. 11, no. 1, Jan. 1965.

[35] T. R. M. Fischer, "Some remarks on the role of inaccuracy in Shannon's theory of information transmission," in *Trans. 8th Prague Conf. Inf. Theory*, pp. 211–226, 1971.

[36] A. Lapidoth, "Mismatched decoding and the multiple-access channel," pp. 1439–1452, *IEEE Trans. Inf. Theory*, vol. 42, no. 5, Sep. 1996.

[37] J. Scarlett, "Reliable communication under mismatched decoding," [Online: http://itc.upf.edu/biblio/1061]. *PhD Thesis.* University of Cambridge, 2014.

[38] A. Ganti, A. Lapidoth, and E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," pp. 2315–2328, *IEEE Trans. Inf. Theory*, vol. 46, no. 7, Nov. 2000.

[39] V. Balakirsky, "A converse coding theorem for mismatched decoding at the output of binary-input memoryless channels," pp. 1889–1902, *IEEE Trans. Inf. Theory*, vol. 41, no. 6, Nov. 1995.

[40] E. Telatar, "Multi-access communications with decision feedback decoding," *PhD Thesis.* MIT, 1992.

[41] C. Bunte, A. Lapidoth, and A. Samorodnitsky, "The zero-undetected-error capacity approaches the Sperner capacity," pp. 3825–3833, *IEEE Trans. Inf. Theory*, vol. 60, no. 7, 2014.

[42] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," pp. 2307–2359, *IEEE Trans. Inf. Theory*, vol. 56, no. 5, May 2010.

[43] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Mismatched decoding: Error exponents, second-order rates and saddlepoint approximations," pp. 2647–2666, *IEEE Trans. Inf. Theory*, vol. 60, no. 5, May 2014.

[44] G. Kaplan and S. Shamai, "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment," pp. 228–239, *Arch. Elek. Über.*, vol. 47, no. 4, 1993.

[45] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Cost-constrained random coding and applications," in *Inf. Theory and Apps. Workshop*, pp. 1–7, San Diego, CA, Feb. 2013.

[46] A. Dembo and L. Kontoyiannis, "Source coding, large deviations, and approximate pattern matching," pp. 1590–1615, *IEEE Trans. Inf. Theory*, vol. 48, no. 6, 2002.

[47] J. Scarlett, V. Y. F. Tan, and G. Durisi, "The dispersion of nearest-neighbor decoding for additive non-Gaussian channels," pp. 81–92, *IEEE Trans. Inf. Theory*, vol. 63, no. 1, Jan. 2017.

[48] A. Lapidoth, "On the role of mismatch in rate distortion theory," pp. 38–47, *IEEE Trans. Inf. Theory*, vol. 43, no. 1, Jan. 1997.

[49] D. Sakrison, "The rate distortion function for a class of sources," pp. 165–195, *Information and Control*, vol. 15, no. 2, 1969.

[50] D. Sakrison, "The rate of a class of random processes," pp. 10–16, *IEEE Trans. Inf. Theory*, vol. 16, no. 1, 1970.

[51] E. Yang and J. C. Kieffer, "On the performance of data compression algorithms based upon string matching," pp. 47–65, *IEEE Trans. Inf. Theory*, vol. 44, no. 1, 1998.

[52] S. I. Bross and A. Lapidoth, "The additive noise channel with a helper," in *IEEE Inf. Theory Workshop*, pp. 1–5, 2019.

[53] E. Yang and Z. Zhang, "On the redundancy of lossy source coding with abstract alphabets," pp. 1092–1110, *IEEE Trans. Inf. Theory*, vol. 45, no. 4, 1999.

[54] C. E. Shannon, "Probability of error for optimal codes in a gaussian channel," pp. 611–656, *Bell Sys. Tech. Journal*, vol. 38, no. 3, 1959.

[55] L. Zhou, V. Y. F. Tan, and M. Motani, "Refined asymptotics for rate-distortion using Gaussian codebooks for arbitrary sources," pp. 3145–3159, *IEEE Trans. Inf. Theory*, vol. 65, no. 5, May 2019.

[56] V. Kostina and S. Verdú, "Fixed-length lossy compression in the finite blocklength regime," pp. 3309–3338, *IEEE Trans. Inf. Theory*, vol. 58, no. 6, 2012.

[57] I. Kontoyiannis and R. Zamir, "Mismatched codebooks and the role of entropy coding in lossy data compression," pp. 1922–1938, *IEEE Trans. Inf. Theory*, vol. 52, no. 5, 2006.

[58] Y. Steinberg and M. Gutman, "An algorithm for source coding subject to a fidelity criterion, based on string matching," pp. 877–886, *IEEE Trans. Inf. Theory*, vol. 39, no. 3, 1993.

[59] Z. Zhang and V. K. Wei, "An on-line universal lossy data compression algorithm via continuous codebook refinement. I. Basic results," pp. 803–821, *IEEE Trans. Inf. Theory*, vol. 42, no. 3, 1996.

[60] R. M. Gray and T. Linder, "Mismatch in high-rate entropy-constrained vector quantization," pp. 1204–1217, *IEEE Trans. Inf. Theory*, vol. 49, no. 5, May 2003.

[61] R. Gray and L. Davisson, "Quantizer mismatch," pp. 439–443, *IEEE Trans. Comms.*, vol. 23, no. 4, Apr. 1975.

[62] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Multiuser random coding techniques for mismatched decoding," pp. 3950–3970, *IEEE Trans. Inf. Theory*, vol. 62, no. 7, Jul. 2016.

[63] D. de Caen, "A lower bound on the probability of a union," pp. 217–220, *Discrete Math.*, vol. 169, 1997.

[64] A. Somekh-Baruch, "On achievable rates and error exponents for channels with mismatched decoding," pp. 727–740, *IEEE Trans. Inf. Theory*, vol. 61, no. 2, Feb. 2015.

[65] J. Scarlett, A. Somekh-Baruch, A. Martinez, and A. Guillén i Fàbregas, "A counter-example to the mismatched decoding converse for binary-input discrete memoryless channels," pp. 5387–5395, *IEEE Trans. Inf. Theory*, vol. 61, no. 10, Oct. 2015.

[66] N. Shulman, "Communication over an unknown channel via common broadcasting," *PhD Thesis.* Tel Aviv University, 2003.

[67] J. Scarlett, L. Peng, N. Merhav, A. Martinez, and A. Guillén i Fàbregas, "Expurgated random-coding ensembles: Exponents, refinements and connections," pp. 4449–4462, *IEEE Trans. Inf. Theory*, vol. 60, no. 8, Aug. 2014.

[68] A. Somekh-Baruch, J. Scarlett, and A. Guillén i Fàbregas, "Generalized random Gilbert-Varshamov codes," pp. 3452–3469, *IEEE Trans. Inf. Theory*, vol. 65, no. 6, Jun. 2019.

[69] S. Shamai and I. Sason, "Variations on the Gallager bounds, connections, and applications," pp. 3029–3051, *IEEE Trans. Inf. Theory*, vol. 48, no. 12, Dec. 2002.

[70] E. N. Gilbert, "A comparison of signalling alphabets," pp. 504–522, *Bell Labs Tech. J.*, vol. 31, no. 3, 1952.

[71] R. R. Varshamov, "Estimate of the number of signals in error correcting codes," 739–741, *Dokl. Akad. Nauk SSSR*, vol. 117, no. 5, 1957.

[72] A. Somekh-Baruch, "A general formula for the mismatch capacity," pp. 4554–4568, *IEEE Trans. Inf. Theory*, vol. 61, no. 9, Sep. 2015.

[73] A. Somekh-Baruch, "Multi-letter converse bounds for the mismatched discrete memoryless channel with an additive metric," in *IEEE Int. Symp. Inf. Theory*, pp. 531–535, 2015.

[74] E. Asadi Kangarshahi and A. Guillén i Fàbregas, "A single-letter upper bound to the mismatch capacity," 2020, [Online]. Available: https://arxiv.org/abs/2004.01785.

[75] A. Somekh-Baruch, "Converse theorems for the DMC with mismatched decoding," pp. 6196–6207, *IEEE Trans. Inf. Theory*, vol. 64, no. 9, Sep. 2018.

[76] T. S. Han, *Information-Spectrum Methods in Information Theory.* Springer, 2002.

[77] S. Verdú and T. S. Han, "A general formula for channel capacity," pp. 1147–1157, *IEEE Trans. Inf. Theory*, vol. 40, no. 4, Jul. 1994.

[78] A. Somekh-Baruch, "On mismatched list decoding," in *IEEE Int. Symp. Inf. Theory*, pp. 526–530, 2015.

[79] R. Ahlswede, "Multi-way communication channels," in *IEEE Int. Symp. Inf. Theory*, pp. 23–52, 1973.

[80] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," pp. 620–643, *IEEE Trans. Inf. Theory*, vol. 55, no. 2, 2009.

[81] E. Asadi Kangarshahi and A. Guillén i Fàbregas, "An upper bound to the mismatch capacity," in *IEEE Int. Symp. Inf. Theory*, pp. 2873–2877, 2019.

[82] E. Asadi Kangarshahi and A. Guillén i Fàbregas, "Properties of a recent upper bound to the mismatch capacity," in *Int. Zürich Sem. Inf. Comm.*, pp. 115–119, 2019.

[83] D. Divsalar, *Performance of mismatched receivers on bandlimited channels, PhD Thesis.* University of California, LA, 1978.

[84] D. Kazakos, “Upper and lower bounds for noisy channel coding under mismatch,” in *Proc. Conf. Info. Sci. Sys.*, pp. 37–42, 1981.

[85] J. Omura and B. Levitt, “Coded error probability evaluation for antijam communication systems,” pp. 896–903, *IEEE Trans. Comms.*, vol. 30, no. 5, May 1982.

[86] I. Abou-Faycal and A. Lapidoth, “On the capacity of reduced-complexity receivers for intersymbol interference channels,” in *IEEE Conv. Elec. Eng. in Israel*, pp. 263–266, 2000.

[87] F. Rusek and D. Fertonani, “Bounds on the information rate of intersymbol interference channels based on mismatched receivers,” pp. 1470–1482, *IEEE Trans. Inf. Theory*, vol. 58, no. 3, 2012.

[88] P. Sadeghi, P. O. Vontobel, and R. Shams, “Optimization of information rate upper and lower bounds for channels with memory,” pp. 663–688, *IEEE Trans. Inf. Theory*, vol. 55, no. 2, Feb. 2009.

[89] F. Rusek and A. Prlja, “Optimal channel shortening for MIMO and ISI channels,” pp. 810–818, *IEEE Trans. Wireless Comms.*, vol. 11, no. 2, 2012.

[90] A. T. Asyhari and A. Guillén i Fàbregas, “MIMO block-fading channels with mismatched CSI,” pp. 7166–7185, *IEEE Trans. Inf. Theory*, vol. 60, no. 11, Nov. 2014.

[91] A. Asyhari and A. Guillén i Fàbregas, “Nearest neighbor decoding in MIMO block-fading channels with imperfect CSIR,” pp. 1483–1517, *IEEE Trans. Inf. Theory*, vol. 58, no. 3, Mar. 2012.

[92] W. Zhang, “A general framework for transmission with transceiver distortion and some applications,” pp. 384–399, *IEEE Trans. Comms.*, vol. 60, no. 2, Feb. 2012.

[93] D. Fehr, J. Scarlett, and A. Martinez, “Fixed-energy random coding with rescaled codewords at the transmitter,” in *Int. Zürich Sem. Comms.*, pp. 170–174, 2016.

[94] V. Y. F. Tan, “Asymptotic estimates in information theory with non-vanishing error probabilities,” pp. 1–184, *Found. Trends Comms. Inf. Theory*, vol. 11, no. 1–2, Sep. 2014.

[95] A. Martinez and A. Guillén i Fàbregas, "Random-coding bounds for threshold decoders: Error exponent and saddlepoint approximation," in *IEEE Int. Symp. Inf. Theory*, pp. 2899–2903, 2011.

[96] A. Martinez and A. Guillén i Fàbregas, "Saddlepoint approximation of random-coding bounds," in *Inf. Theory Apps. Workshop*, pp. 1–6, 2011.

[97] L. Zhou, V. Y. F. Tan, and M. Motani, "Second-order asymptotics of universal JSCC for arbitrary sources and additive channels," in *IEEE Int. Symp. Inf. Theory*, pp. 1500–1504, 2018.

[98] A. Lapidoth, "A note on feedback communication with mismatched decoding," in *IEEE Int. Conf. Sci. Elec. Eng.*, pp. 1–5, 2016.

[99] C. Bunte and A. Lapidoth, "The zero-undetected-error capacity of discrete memoryless channels with feedback," in *Allerton Conf. Comm., Control, and Comp.*, pp. 1838–1842, 2012.

[100] A. Somekh-Baruch, "Mismatched identification via channels," in *IEEE Int. Symp. Inf. Theory*, pp. 2751–2755, 2017.

[101] R. Ahlswede and G. Dueck, "Identification via channels," pp. 15–29, *IEEE Trans. Inf. Theory*, vol. 35, no. 1, Jan. 1989.

[102] M. H. Yassaee, M. R. Aref, and A. Gohari, "A technique for deriving one-shot achievability results in network information theory," 2013, [Online]. Available: http://arxiv.org/abs/1303.0696.

[103] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "The likelihood decoder: Error exponents and mismatch," in *IEEE Int. Symp. Inf. Theory*, pp. 86–90, 2015.

[104] N. Merhav, "The generalized stochastic likelihood decoder: Random coding and expurgated bounds," pp. 5039–5051, *IEEE Trans. Inf. Theory*, vol. 63, no. 8, Aug. 2017.

[105] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Mismatched multi-letter successive decoding for the multiple-access channel," pp. 2253–2266, *IEEE Trans. Inf. Theory*, vol. 64, no. 4, Apr. 2018.

[106] N. Merhav, "Universal decoding for arbitrary channels relative to a given class of decoding metrics," pp. 5566–5576, *IEEE Trans. Inf. Theory*, vol. 59, no. 9, Sep. 2013.

[107] G. Forney, "The Viterbi algorithm," pp. 268–278, *Proc. IEEE*, vol. 61, no. 3, Mar. 1973.

[108] A. Lapidoth and S. Shamai, "A lower bound on the bit-error-rate resulting from mismatched Viterbi decoding," pp. 473–482, *European Trans. Telecom.*, vol. 9, no. 6, 1998.

[109] M. Alsan, "A lower bound on achievable rates by polar codes with mismatch polar decoding," in *IEEE Inf. Theory Workshop*, pp. 1–5, 2013.

[110] M. Alsan, "Performance of mismatched polar codes over BSCs," in *Int. Symp. Inf. Theory Apps.*, pp. 750–755, 2012.

[111] M. Alsan and E. Telatar, "Polarization as a novel architecture to boost the classical mismatched capacity of B-DMCs," in *IEEE Inf. Theory Workshop*, pp. 366–370, 2014.

[112] G. L. Nemhauser and L. A. Wolsey, *Integer Programming and Combinatorial Optimization.* Wiley New York, 1999.