Technical Program

Time	Room 1	Room 2
Monday, April 12		
08:45-09:00	Conference Opening	
09:00-10:00	Plenary - Tree archeology: Root finding and broadcasting	
10:10-10:50	Channel Capacity	Quantum Information
10:50-11:00		
11:10-12:00	Coding I: Caching and Distributed storage	Information Theory and Applications
12:00-12:10		
18:00-19:50	Tutorial - An Introduction to Quantum Computation	
20:10-22:00	Tutorial - Explicit and Implicit Inductive Bias in Deep Learning	

2020 IEEE Information Theory Workshop (ITW) | 978-1-7281-5962-1/20/$31.00 ©2021 IEEE | DOI: 10.1109/ITW46852.2021.9457668

Tuesday, April 13

09:00-10:00	Plenary - Coded Caching: Past, Present, Future	
10:10-11:10	Wireless I	
16:00-17:40	Themed Session - Topics in Wireless Communications	
18:00-18:50	Cryptography, Privacy and Security I	Machine Learning I
19:00-20:00	Plenary - Data Driven Algorithm Design	
20:20-22:00	Themed Session - Statistical Physics and Machine Learning	

Wednesday, April 14

09:00-09:50	Coding II: Codes on Graphs and Polar Codes	Multi-User Information Theory
10:00-10:40	Coding III: Coded Caching	Source Coding
11:00-11:50	Coding IV: Coding applications	Statistics and Information Theory I
12:00-12:50	Coding V: Coding at Large	
18:00-18:50	Cryptography, Privacy and Security II	Statistics and Information Theory II
19:00-20:00	Plenary - The generalization error of overparametrized models: Insights from exact asymptotics	
20:10-21:50	Themed Session - Blockchain	

Thursday, April 15

09:00-10:00	Plenary - Information-directed Exploration in Bandits and Reinforcement Learning	
10:00-10:10	Announcements	
10:20-11:10	Machine Learning II	Cryptography, Privacy and Security III
11:20-12:10	Coding VI: Coding theory and practice	
16:00-17:40	Themed Session - Learning Theory	
18:00-18:50	Wireless II	Coding VII: Coding at Large
19:00-20:00	Plenary - Diversity vs. Parallelism in Distributed Computing with Redundancy	
20:10-21:50	Themed Session - Coding Theory and Applications	

Monday, April 12

Monday, April 12 8:45 - 9:00 (Europe/Rome)

Conference Opening

Chairs: Marco Dalai (University of Brescia, Italy), Enrico Paolini (University of Bologna, Italy)

Monday, April 12 9:00 - 10:00 (Europe/Rome)

Plenary - Tree archeology: Root finding and broadcasting

Gábor Lugosi

Chair: Nicolò Cesa-Bianchi (Università degli Studi di Milano, Italy)

Networks are often modeled by random processes in which nodes are added one-by-one, according to some simple random rule. Uniform and preferential attachment trees are among the simplest examples of such dynamically growing networks. The statistical problems we address in this talk regard discovering the past of the tree when a present-day snapshot is observed. We present results that show that, even in gigantic networks, a lot of information is preserved from the early days. In particular, we discuss the problem of finding the root and the broadcasting problem.

Monday, April 12 10:10 - 11:00 (Europe/Rome)

Channel Capacity

Room 1

Chair: Yuval Kochman (The Hebrew University of Jerusalem, Israel)

10:10 *A Single-Letter Upper Bound on the Mismatch Capacity via a Multicasting Approach*

Anelia Somekh-Baruch (Bar-Ilan University, Israel)

We introduce a new analysis technique that we use to derive a single-letter upper bound on the mismatch capacity of a stationary point-to-point memoryless channel with decoding metric \(q\). Our bound is obtained by considering a multicast transmission over a two-user broadcast channel with decoding metrics \(q\) and \(\rho\) at the receivers, referred to as \((q, \rho)\)-surely degraded. This channel has the property that the intersection event of the correct \(q\)-decoding of receiver \(1\) and the erroneous \(\rho\)-decoding of receiver \(2\) has zero probability for any codebook of a certain composition \(P\). Our bound holds in the strong converse sense of exponential decay of the probability of correct decoding at rates above the bound. Several examples that demonstrate the strict improvement of our bound compared to

previous results are analyzed. Further, we detect equivalence classes of isomorphic channel-metric pairs (W,q) that share the same mismatch capacity. We prove that if the class contains a matched pair, then our bound is tight and the mismatch capacity of the entire class is fully characterized and is equal to the LM rate, which is the highest rate achievable by random coding, and may be strictly lower than the Shannon (matched) capacity.

10:20 *Encoder-Assistance for Additive Noise Channels*

Amos Lapidoth (ETHZ, Switzerland); Gian Marti (ETH Zurich, Switzerland)

Flash helping has recently been shown to be an effective technique for describing additive noise to a decoder. It is shown here to be effective also in assisting the encoder: it achieves the helper capacity on the single-user Gaussian channel, on the multiple-access Gaussian channel, on the Exponential channel, and on the discrete modulo-additive noise channel. Most of the results hold irrespective of whether the helper observes the noise causally or noncausally.

10:30 *Feedback Capacity of Gaussian Multiple-Access Wiretap Channel with Degraded Message Sets*

Bin Dai (Southwest Jiaotong University, China); Chong Li (Nakamoto & Turing Labs, USA); Yingbin Liang (The Ohio State University, USA); Zheng Ma (Southwest Jiaotong University, China); Shlomo (Shitz) Shamai (The Technion, Israel)

The Schalkwijk-Kailath (SK) feedback scheme is a capacity-achieving coding scheme for the point-to-point white Gaussian channel with feedback. Recently, it has been shown that the SK scheme, which is not designed with consideration of secrecy, already achieves perfect weak secrecy by itself, i.e., the secrecy capacity of the Gaussian wiretap channel with feedback equals the capacity of the same model without secrecy constraint. In this paper, we propose a capacity-achieving SK type feedback scheme for the two-user Gaussian multiple-access channel with degraded message sets (GMAC-DMS). Similarly to the inherent secrecy nature of the classical SK scheme, we show that the proposed scheme is also secure by itself, which indicates that the feedback secrecy capacity of the two-user Gaussian multiple-access wiretap channel with degraded message sets (GMAC-WT-DMS) equals the capacity of the same model without secrecy constraint.

10:40 *On the Capacity of the Continuous-Space SSFM Model of Optical Fiber*

Milad Sefidgaran (Télécom Paris, France); Mansoor Yousefi (Télécom ParisTech, France)

The limit of a discrete-time model of the optical fiber described by the split-step Fourier method (SSFM) when the number of segments in distance K tends to infinity is considered. It is shown that if $K \geq \mathcal{P}^{2/3}$ and $\mathcal{P} \rightarrow \infty$, where $\mathcal{P}$ is the average input power, the capacity of the resulting continuous-space lossless model is lower bounded by $\frac{1}{2}\log_2(1+\text{SNR}) - \frac{1}{2}+ o(1)$, where $o(1)$ tends to zero with the signal-to-noise ratio SNR. This implies that at least half of the signal degrees-of-freedom remain asymptotically in this model.

10:50 *On the Optimality of Dolinar's Receiver*

Itamar Katz and Yuval Kochman (The Hebrew University of Jerusalem, Israel)

Dolinar's receiver is an architecture for distinguishing between two possible coherent states, using a

photon detector and a local signal which may depend on past detector measurements. The optimal local signal satisfies very favorable properties: it is independent of the time horizon, and the resulting error probability is independent of the measurements. It was also shown that the same signal is optimal in the sense of maximizing the mutual information between the identity of the state and the measurements. In this work we show that the same signal is optimal for the optimization of the expected value of a wide class of objective functions. Our proof is based entirely on convex optimization and functional analysis, without resorting to any "quantum" arguments.

Monday, April 12 10:10 - 10:50 (Europe/Rome)

Quantum Information

Room 2

Chair: Christian Deppe (Technical University of Munich, Germany)

10:10 ***Correcting Erasures with Topological Subsystem Color Codes***

Hiteshvi Manish Solanki and Pradeep K Sarvepalli (Indian Institute of Technology Madras, India)

Qubit loss is one of the forms of noise encountered in some quantum technologies. Such noise is modeled using the quantum erasure channel. Unlike the depolarizing noise, it is much more tractable, yet the performance of many quantum codes over the erasure channel has not been studied as extensively. In this paper, we study the performance of topological subsystem color codes (TSCCs) over the quantum erasure channel. It is the first such study of TSCCs over the erasure channel. We propose multiple decoding algorithms for TSCC and obtain the highest threshold of about 9.7% for the subsystem color code derived from the square octagon lattice.

10:20 ***Linear programming decoder for hypergraph product quantum codes***

Omar Fawzi (ENS de Lyon, France); Lucien Grouès (Sorbonne Université & Inria Paris, France); Anthony Leverrier (INRIA, France)

We introduce a decoder for quantum CSS codes that is based on linear programming. Our definition is a priori slightly different from the one proposed by Li and Vontobel as we have a syndrome oriented approach instead of an error oriented one, but we show that the success condition is equivalent. Although we prove that this decoder fails for quantum codes that do not have good soundness property (i.e., having large errors with syndrome of small weight) such as the toric code, we obtain good results from simulations. We run our decoder for hypergraph products of two random LDPC codes, showing that it performs better than belief propagation, even combined with the small-set-flip decoder that can provably correct a constant fraction of random errors.

10:30 ***Universal Communication Efficient Quantum Threshold Secret Sharing Schemes***

Kaushik Senthoor and Pradeep K Sarvepalli (Indian Institute of Technology Madras, India)

Quantum secret sharing (QSS) is a cryptographic protocol in which a quantum secret is distributed among

a number of parties where some subsets of the parties are able to recover the secret while some subsets are unable to recover the secret. In the standard \(((k,n))\) quantum threshold secret sharing scheme, any subset of \(k\) or more parties out of the total \(n\) parties can recover the secret while other subsets have no information about the secret. But recovery of the secret incurs a communication cost of at least \(k\) qudits for every qudit in the secret. Recently, a class of communication efficient QSS schemes were proposed which can improve this communication cost to \(\frac{d}{d-k+1}\) by contacting \(d\geq k\) parties where \(d\) is fixed prior to the distribution of shares. In this paper, we propose a more general class of \(((k,n))\) quantum secret sharing schemes with low communication complexity. In these schemes the combiner can contact any \(d\) parties at the time of recovery where \(k\leq d\leq n\). This is the first such class of universal communication efficient quantum threshold schemes.

10:40 ***Quantum Channel State Masking***

Uzi Pereg and Christian Deppe (Technical University of Munich, Germany); Holger Boche (Technical University Munich, Germany)

Communication over a quantum channel that depends on a quantum state is considered, when the encoder has channel side information (CSI) and is required to mask information on the quantum channel state from the decoder. A full characterization is established for the entanglement-assisted masking equivocation region, and a regularized formula is given for the quantum capacity-leakage function without assistance. For Hadamard channels without assistance, we derive single-letter inner and outer bounds, which coincide in the standard case of a channel that does not depend on a state.

Monday, April 12 11:10 - 12:00 (Europe/Rome)

Coding I: Caching and Distributed storage

Room 1

Chair: Lakshmi Prasad Natarajan (Indian Institute of Technology Hyderabad, India)

11:10 ***Explicit Construction of Minimum Storage Rack-Aware Regenerating Codes for All Parameters***

Liyang Zhou and Zhifang Zhang (Academy of Mathematics and Systems Science, Chinese Academy of Sciences, China)

We consider the rack-aware storage system where \(n=\bar{n}u\) nodes are organized in \(\bar{n}\) racks each containing \(u\) nodes, and any \(k=\bar{k}u+u_0~(0\leq u_0<u)\) nodes can retrieve the original data file. More importantly, the cross-rack communication cost is much more expensive than the intra-rack communication cost, so that the latter is usually neglected in the system bandwidth. The MSRR (minimum storage rack-aware regenerating) code is an important variation of regenerating codes that achieves the optimal repair bandwidth for single node failures in the rack-aware model. However, explicit construction of MSRR codes for all parameters were not developed until Chen&Barg's work. In this paper we present another explicit construction of MSRR codes for all parameters that improve Chen&Barg's construction in two aspects: (1) The sub-packetization is reduced from \((\bar{d}-\bar{k}+1)^{\bar{n}}\) to \((\bar{d}-\bar{k}+1)^{\lceil\frac{\bar{n}}{u-u_{0}}\rceil}\) where \(\bar{d}\) is the number of helper racks that participate in the repair process; (2) The field size is reduced to \(|F|>n\) which is almost half of the field used in Chen&Barg's construction. Besides, our code keeps the same access level as Chen&Barg's

low-access construction.

11:20 *Coded Data Rebalancing for Decentralized Distributed Databases*

Sushena Sree (International Institute of Information Technology, India); Prasad Krishnan (IIIT Hyderabad, India)

The performance of replication-based distributed databases is affected due to non-uniform storage across storage nodes (also called data skew) and reduction in the replication factor during operation, particularly due to node additions or removals. Data rebalancing refers to the communication involved between the nodes in correcting this data skew, while maintaining the replication factor. For carefully designed distributed databases, transmitting coded symbols during the rebalancing phase has been recently shown to reduce the communication load of rebalancing. In this work, we look at balanced distributed databases with random placement, in which each data segment is stored in a random subset of r nodes in the system, where r refers to the replication factor of the distributed database. We call these as decentralized databases. For a natural class of such decentralized databases, we propose rebalancing schemes for correcting data skew and the reduction in the replication factor arising due to a single node addition or removal. We give converse arguments which show that our proposed rebalancing schemes are optimal asymptotically in the size of the file.

11:30 *Coded Computing for Master-Aided Distributed Computing Systems*

Haoning Chen and Youlong Wu (ShanghaiTech University, China)

We consider a MapReduce-type task running in a distributed computing model which consists of K edge computing nodes distributed across the edge of the network and a Master node that assists the edge nodes to compute output functions. The Master node and the edge nodes, both equipped with some storage memories and computing capabilities, are connected through a multicast network. We define the communication time spent during the transmission for the sequential implementation (all nodes send symbols sequentially) and parallel implementation (the Master node can send symbols during the edge nodes' transmission), respectively. We propose a mixed coded distributed computing scheme that divides the system into two subsystems where the coded distributed computing (CDC) strategy proposed by Songze Li et al. is applied into the first subsystem and a novel master-aided CDC strategy is applied into the second subsystem. We prove that this scheme is optimal, i.e., achieves the minimum communication time for both the sequential and parallel implementation, and establish an optimal information-theoretic tradeoff between the overall communication time, computation load, and the Master node's storage capacity. It demonstrates that incorporating a Master node with storage and computing capabilities can further reduce the communication time. For the sequential implementation, we deduce the approximately optimal file allocation between the two subsystems, which shows that the Master node should map as many files as possible in order to achieve smaller communication time. For the parallel implementation, if the Master node's storage and computing capabilities are sufficiently large (not necessary to store and map all files), then the proposed scheme requires at most 1/2 of the minimum communication time of system without the help of the Master node.

11:40 *An Umbrella Converse for Data Exchange: Applied to Caching, Computing, Shuffling & Rebalancing*

Prasad Krishnan (IIIT Hyderabad, India); Lakshmi Prasad Natarajan (Indian Institute of Technology Hyderabad, India); V. Lalitha (IIIT Hyderabad, India)

The problem of data exchange between multiple nodes with (not necessarily uniform) storage and

communication capabilities models several current multi-user communication problems like Coded Caching, Data shuffling, Coded Computing, etc. The goal in such problems is to design communication schemes which accomplish the desired data exchange between the nodes with the optimal (minimum) amount of communication load. In this work, we present a converse to such a general data exchange problem between multiple nodes. The expression of the converse depends only on the number of bits to be moved between different subsets of nodes, and does not assume anything further specific about the parameters in the problem. Specific problem formulations, such as those in Coded Caching, Coded Data Shuffling, Coded Distributed Computing, and some of their variants, naturally can be seen as instances of this generic data exchange problem. Applying our generic converse to such problems, we recover known important converses for these settings and some of their variants in a simpler way. Further, for a generic coded caching problem with multiple transmitters, receivers and cache sizes, we show a new general converse which subsumes many existing results. We also employ our bound to obtain a new tight converse bound for the multi-node removal case in the Coded Data Rebalancing problem, in which nodes must exchange information to `rebalance' a storage cluster after some node failures occur.

11:50 ***Blind Updates in Coded Caching***

Suman Ghosh (IIT Hyderabad, India); Prasad Krishnan (IIIT Hyderabad, India); Lakshmi Prasad Natarajan (Indian Institute of Technology Hyderabad, India)

We consider the centralized coded caching system where a library of files is available at the server and their subfiles are cached at the clients as prescribed by a placement delivery array (PDA). We are interested in the problem where a specific file in the library is replaced with a new file at the server, the contents of which are correlated with the file being replaced, and this replacement needs to be communicated to the caches. The server loses the original file when the replacement is done and is unaware of the differences between the two files, whereas each cache has access to specific subfiles of the original file as dictated by the PDA. We model the correlation between the two files by assuming that they differ in at the most \(\epsilon\) subfiles, and aim to reduce the number of bits broadcast by the server to update the caches. We design a new elegant coded transmission strategy for the server to update the caches blindly, and also identify another simple scheme that is based on MDS codes. We then derive converse bounds on the minimum cost \(\ell^*\) among all linear strategies. For two well-known families of PDAs -- the Maddah-Ali-Niesen scheme and a scheme by Tang & Ramamoorthy and Yan et al. -- we show that our new scheme has cost \(\ell^*(1 + o(1))\) when the updates are sufficiently sparse, while the scheme using MDS codes has order-optimal cost when the updates are dense.

Monday, April 12 11:10 - 12:10 (Europe/Rome)

Information Theory and Applications

Room 2

Chair: Stefano Rini (National Yangming Jiaotong University, Taiwan)

11:10 ***On the Optimality of Treating Interference as Noise: General Message Sets Revisited***

Hamdi Joudeh (Eindhoven University of Technology, The Netherlands); Giuseppe Caire (Technische Universität Berlin, Germany)

We study the optimality of power control and treating interference as noise (TIN) in the MxN X channel, from the generalized degrees-of-freedom (GDoF) and constant-gap capacity perspectives. A result by Geng, Sun and Jafar shows that if there exist K=min(M,N) transmitter-receiver pairs such that each direct link strength is no less than the sum of the strongest incoming and strongest outgoing cross link strengths (all in dB), then it is optimal to reduce the MxN X channel to a K-user interference channel and use TIN. The proof of this result relies on a deterministic approximation of the original Gaussian network, specifically for the case M<N. Here we present a simpler proof by working directly with the original Gaussian network. Our proof relies on a new "less noisy under interference" order exhibited by TIN-optimal MxN X channels, akin to the "less noisy" order in broadcast channels.

11:20 *A Coded Caching Scheme with Linear Sub-packetization and its Application to Multi-Access Coded Caching*

Anjana Ambika Mahesh (Indian Institute of Science, Bangalore, India); B. Sundar Rajan (Indian Institute of Science, India)

This paper addresses the problem of exponentially increasing sub-packetization with the number of users in a centralized coded caching system by introducing a new coded caching scheme inspired by the symmetric neighboring consecutive side information index coding problem. The scheme has a placement policy where the number of sub-packets required grows only linearly with the number of users, with no restriction on the number of users or file size, and a delivery policy which is instantaneously decodable. Further, an application of the new delivery scheme in a multi-access coded caching set-up is studied and a few results in that direction are presented. In particular, in the multi-access set-up, for cases where optimality rate-memory trade-off characterizations are available, it is shown that the new delivery scheme achieves optimal or near-optimal rates.

11:30 *An Embedded Index Code Construction Using Sub-packetization*

Shanuja Sasi (Indian Institute of Science, Bangalore, India); Vaneet Aggarwal (Purdue University, USA); B. Sundar Rajan (Indian Institute of Science, India)

A variant of the index coding problem (ICP), the embedded index coding problem (EICP) was introduced in [A. Porter and M. Wootters, ``Embedded Index Coding,'' ITW, Sweden, 2019] which was motivated by its application in distributed computing where every user can act as sender for other users and an algorithm for code construction was reported. The construction depends on the computation of min-rank of a matrix, which is computationally intensive. In [A. A. Mahesh, N. S. Karat and B. S. Rajan, ``Min-rank of Embedded Index Coding Problems,'' ISIT, 2020], the authors have provided an explicit code construction for a class of EICP - \(\textit{Consecutive and Symmetric Embedded Index Coding Problem (CS-EICP)}\). We introduce the idea of sub-packetization of the messages in index coding problems to provide a novel code construction for CS-EICP in contrast to the scalar linear solutions provided in the prior works. For CS-EICP, the normalized rate, which is defined as the number of bits transmitted by all the users together normalized by the total number of bits of all the messages, for our construction is lesser than the normalized rate achieved by Mahesh \(\textit{et al.}\), for scalar linear codes.

11:40 *On the continuous time additive Gaussian noise channel in the presence of perfect feedback*

Aman Chawla (Alliance University, India); Salvatore Domenic Morgera (University of South Florida, USA)

A lower bound to the expected decoding time of a power constrained continuous time additive white

Gaussian noise channel with perfect feedback is studied. The main assumption is of right continuity of sample paths for all processes and a Nack-continuity assumption on the Encoder. We also employ a conjecture for the continuous time case.

11:50 ***On the Effectiveness of Fekete's Lemma in Information Theory***

Holger Boche and Yannik Böck (Technical University Munich, Germany); Christian Deppe (Technical University of Munich, Germany)

Fekete's lemma is a well known assertion that states the existence of limit values of superadditive sequences. In information theory, superadditivity of rate functions occurs in a variety of channel models, making Fekete's lemma essential to the corresponding capacity problems. We analyze Fekete's lemma with respect to effective convergence and computability and show that Fekete's lemma exhibits no constructive derivation. In particular, we devise a superadditive, computable sequence of rational numbers so that the associated limit value in the sense of Fekete's lemma is not a computable number. We further characterize the requirements for effective convergence and investigate the speed of convergence, as proposed by Rudolf Ahlswede in his 2006 Shannon lecture.

12:00 ***Combinatorial Quantitative Group Testing with Adversarially Perturbed Measurements***

Yun-Han Li and I-Hsiang Wang (National Taiwan University, Taiwan)

In this paper, combinatorial quantitative group testing (QGT) with noisy measurements is studied. The goal of QGT is to detect defective items from a data set of size n with counting measurements, each of which counts the number of defects in a selected pool of items. While most literatures consider either probabilistic QGT with random noise or combinatorial QGT with noiseless measurements, our focus is on the combinatorial QGT with noisy measurements that might be adversarially perturbed by additive bounded noises. Since perfect detection is impossible, a partial detection criterion is adopted. With the adversarial noise being bounded by $d_n = \Theta(n^\delta)$ and the detection criterion being to ensure no more than $k_n = \Theta(n^\kappa)$ errors can be made, our goal is to characterize the fundamental limit on the number of measurement, termed pooling complexity, as well as provide explicit construction of measurement plans with optimal pooling complexity and efficient decoding algorithms. We first show that the fundamental limit is $\frac{1}{1-2\delta}\frac{n}{\log n}$ to within a constant factor not depending on (n,κ,δ) for the non-adaptive setting when $0<2\delta\leq\kappa <1$, sharpening the previous result by Chen and Wang [1]. We also provide deterministic constructions of an adaptive method with $\frac{1}{1-2\delta}\frac{n}{\log_{2} n}$ pooling complexity up to a constant factor and $O(n)$ decoding complexity.

Monday, April 12 18:00 - 19:50 (Europe/Rome)

Tutorial - An Introduction to Quantum Computation

Eric Chitambar

Chair: Olgica Milenkovic (University of Illinois at Urbana-Champaign (UIUC), USA)

In this tutorial I will provide a brief introduction to quantum computation. The first half of the tutorial will focus

on the basic principles of quantum circuits and quantum information processing. Topics here include qubits, quantum gates, quantum measurement, and entanglement. In the second half we will apply these principles and study some preliminary quantum algorithms such as the Deutsch-Jozsa algorithm and the Bernstein-Vazirani algorithm.

Monday, April 12 20:10 - 22:00 (Europe/Rome)

Tutorial - Explicit and Implicit Inductive Bias in Deep Learning

Nathan Srebro

Chair: Nicolò Cesa-Bianchi (Università degli Studi di Milano, Italy)

Inductive bias (reflecting prior knowledge or assumptions) lies at the core of every learning system and is essential for allowing learning and generalization, both from a statistical perspective, and from a computational perspective. What is the inductive bias that drives deep learning? A simplistic answer to this question is that we learn functions representable by a given architecture. But this is not sufficient neither computational (as learning even modestly sized neural networks is intractable) nor statistically (since modern architectures are too large to ensure generalization). In this tutorial we will explore these considerations, how training humongous, even infinite, deep networks can ensure generalization, what function spaces such infinite networks might correspond to, and how the inductive bias is tightly tied to the local search procedures used to train deep networks.

Tuesday, April 13

Tuesday, April 13 9:00 - 10:00 (Europe/Rome)

Plenary - Coded Caching: Past, Present, Future

Giuseppe Caire

Chair: Alon Orlitsky (University of California, San Diego, USA)

Coded caching has emerged as a powerful and elegant idea for content distribution over communication networks. Since the initial work of Maddah-Ali and Niesen, a vast set of theoretical results have been developed in the network coding and information theory community. These results range from solving more and more complicated theoretical "puzzles" (i.e., highly involved, but somehow practically irrelevant problems) to addressing more concrete problems of practical relevance for applications. Yet, questions still remain about whether such schemes will ever be used in the real world on a vast scale. This talk provides an account of some recent exciting results including the real-world implementation of coded caching on actual wireless networks, addressing some of the residual skepticism about the feasibility and actual gains achievable by these schemes.

Tuesday, April 13 10:10 - 11:10 (Europe/Rome)

Wireless I

Chair: Mael Le Treust (ETIS UMR 8051, Université Cergy-Pontoise, ENSEA, CNRS, France)

10:10 ***A Sphere Packing Bound for AWGN MIMO Fading Channels under Peak Amplitude Constraints***

Antonino Favano (Politecnico di Milano & CNR-IEIIT, Italy); Marco Ferrari (CNR-IEIIT, Italy); Maurizio Magarini and Luca Barletta (Politecnico di Milano, Italy)

An upper bound on the capacity of multiple-input multiple-output (MIMO) additive white Gaussian noise fading channels is derived under peak amplitude constraints. The tightness of the bound is investigated at high signal-to-noise ratio (SNR), for any arbitrary convex amplitude constraint region. Moreover, a numerical simulation of the bound for fading MIMO channels is analyzed, at any SNR level, for a practical transmitter configuration employing a single power amplifier for all transmitting antennas.

10:20 ***Noncoherent MIMO Multiple-Access Channels: A Joint Constellation Design***

Khac-Hoang Ngo (Chalmers University of Technology, Sweden); Sheng Yang (CentraleSupélec, France); Maxime Guillaud and Alexis Decurninge (Huawei Technologies, France)

We consider the joint constellation design problem for noncoherent multiple-input multiple-output multiple-access channels. By analyzing the noncoherent maximum-likelihood detection error, we propose novel design criteria so as to minimize the error probability. For any given set of constellation sizes, the proposed metrics can be optimized over the set of signal matrices. Based on these criteria, we propose a simple and efficient construction consisting in partitioning a single-user constellation. Numerical results show that our proposed metrics are meaningful, and can be used as objectives to generate constellations through numerical optimization that perform better, for the same transmission rate and power constraint, than a common pilot-based scheme and the constellations optimized with existing metrics.

10:30 ***The Optimal DoF for the Noncoherent MIMO Channel with Generic Block Fading***

Khac-Hoang Ngo (Chalmers University of Technology, Sweden); Sheng Yang (CentraleSupélec, France); Maxime Guillaud (Huawei Technologies, France)

The high-SNR capacity of the noncoherent MIMO channel has been derived for the case of independent and identically distributed (IID) Rayleigh block fading by exploiting the Gaussianity of the channel matrix. This implies the optimal degrees of freedom (DoF), i.e., the capacity pre-log factor. Nevertheless, as far as the optimal DoF is concerned, IID Rayleigh fading is apparently a sufficient but not necessary condition. In this paper, we show that the optimal DoF for the IID Rayleigh block fading channel is also the optimal DoF for a more general class of generic block fading channels, in which the random channel matrix has finite power and finite differential entropy. Our main contribution is a novel converse proof based on the duality approach.

10:40 ***On Secure Degrees of Freedom for K-User MISO Broadcast Channel With***

Alternating CSIT

Leyla Sadighi and Sadaf Salehkalaibar (University of Tehran, Iran); Stefano Rini (National Yangming Jiaotong University, Taiwan)

In this paper, the sum secure degrees of freedom (SDoF) of the K-user Multiple Input/Single Output (MISO) Broadcast Channel with Confidential Messages (BCCM) and alternating Channel State Information at the Transmitter (CSIT) is investigated. In the MISO BCCM, a K-antenna transmitter (TX) communicates toward K single-antenna receivers (RXs), so that message for RX k is kept secret from RX j with j<k. For this model, we consider the scenario in which the CSI of the RXs from 2 to K is instantaneously known at the transmitter while CSI of RX 1 is known at the transmitter (i) instantaneously for half of the time and (ii) with a unit delay for the remainder of the time. We refer to this CSIT availability as \emph{alternating} CSIT. Alternating CSIT has been shown to provide synergistic gains in terms of SDoF and is thus of a viable strategy to ensure secure communication by simply relying on the CSI feedback strategy. Our main contribution is the characterization of sum SDoF for this model as SDoF_{\rm sum}= (2K-1)/2. Interestingly, this SDoF_{\rm sum} is attained by a rather simple achievability in which the TX uses artificial noise to prevent the decoding of the message of the unintended receivers at RX 1. The proof for the case K=3 is discussed in detail.

10:50 *Rate-Memory Trade-Off for the Cache-Aided MISO Broadcast Channel with Hybrid CSIT*

Antonio Bazco-Nogueras and Petros Elia (EURECOM, France)

One of the famous problems in communications was the so-called "PN" problem in the Broadcast Channel, which refers to the setting where a fixed set of users provide perfect Channel State Information (CSI) to a multi-antenna transmitter, whereas the remaining users only provide finite precision CSI or no CSI. The Degrees-of-Freedom (DoF) of that setting were recently derived by means of the Aligned Image Set approach. In this work, we resolve the cache-aided variant of this problem (i.e., the "PN" setting with side information) in the regime where the number of users providing perfect CSI is smaller than or equal to the number of transmit antennas. In particular, we derive the optimal rate-memory trade-off under the assumption of uncoded placement, and characterize the same trade-off within a factor of 2.01 for general placement. The result proves that the "PN" impact remains similar even in the presence of side information, but also that the optimal trade-off is not achievable through serving independently the two sets of users.

11:00 *Point-to-Point Strategic Communication*

Mael Le Treust (ETIS UMR 8051, Université Cergy-Pontoise, ENSEA, CNRS, France); Tristan Tomala (HEC Paris, GREGHEC UMR, France)

We investigate a strategic formulation of the joint source-channel coding problem in which the encoder and the decoder are endowed with distinct distortion functions. We provide the solutions in four different scenarios. First, we assume that the encoder and the decoder cooperate in order to achieve a certain pair of distortion values. Second, we suppose that the encoder commits to a strategy whereas the decoder implements a best response, as in the persuasion game where the encoder is the Stackelberg leader. Third, we consider that the decoder commits to a strategy, as in the mismatched rate-distortion problem or as in the mechanism design framework. Fourth, we study the cheap talk game in which the encoding and the decoding strategies form a Nash equilibrium.

Tuesday, April 13 16:00 - 17:40 (Europe/Rome)

Themed Session - Topics in Wireless Communications

Chairs: Giuseppe Durisi (Chalmers University of Technology, Sweden), Yury Polyanskiy (MIT, USA)

16:00 *Massive Random Access for the Quasi-Static Fading MAC*

Alexey A. Frolov (Skolkovo Institute of Science and Technology & IITP RAS, Russia)

The problem of massive machine-type communications is of critical importance for future 5G/6G wireless networks. Indeed, the number of devices (sensors) connected to the network grows exponentially. At the same time, the traffic of the devices is significantly different from the traffic generated by human-users and consists of short packets that are sent sporadically. The main goal is not to increase the spectral efficiency but to provide connectivity and energy efficiency. Current transmission schemes are highly inefficient in this regime. The most promising way to deal with the problem is to use the random access schemes or, equivalently, a grant-free transmission, i.e. the device transmits the packet without any prior communication to the base station. As the number of devices is extremely large and it is difficult to create different encoders for the users, the promising strategy is to employ the same codebook schemes. Such schemes are called unsourced random access schemes. Currently, this topic is hot and the case of Gaussian MAC is well investigated. In the talk, we focus on the quasi-static Rayleigh fading MAC - a more realistic channel model. We present fundamental limits and compare the unsourced random access schemes for this channel. Our main focus is a coded compressed sensing approach.

16:20 *An information theoretic perspective on low resolution quantization in mmWave communications*

Elza Erkip (New York University, USA)

Millimeter wave (mmWave) networks offer the possibility of orders of magnitude greater capacity than their 4G counterparts. However, there are various challenges in realizing these gains: mmWave signals are highly directional and are prone to blockage, and the wide bandwidth as well as large number of transceiver antennas result in significant energy consumption. This talk focuses on using low resolution digital-to-analog and analog-to-digital converters (DACs and ADCs) to reduce the energy consumption of mmWave terminals. We study communication strategies and capacity bounds under quantization and spectral mask constraints. Our results provide design guidelines for transceiver architectures and modulation schemes under low resolution quantization and suggest that satisfying out-of-band emissions could be a major challenge.

16:40 *A Group Testing Approach to Collision Resolution in Coded Slotted Aloha Protocols*

Gianluigi Liva (German Aerospace Center (DLR), Germany)

The demand of efficient medium access control strategies for emerging massive wireless IoT systems has originated a revived interest in the design of powerful random access protocols. Among various modern RA approaches, coded slotted Aloha (CSA) gained some popularity thanks to its remarkable performance, that is achievable with a limited complexity from a signal processing viewpoint. With a few notable exceptions, CSA-like protocols have been analyzed assuming the lack of feedback, i.e., no retransmission is attempted for packets that are lost due to unresolvable collisions. In this work, we explore the use of group testing

techniques to steer a collision resolution phase in the context of framed CSA protocols. We show how the application of the simple combinatorial orthogonal matching pursuit algorithm is sufficient to resolve most of the collisions that hinder the success of the iterative interference cancellation process. We highlight the design choices that have to be addressed, in order to benefit from the proposed approach. Finally, we outline a few interesting directions for future developments.

17:00 ***Creating Nonlinear Wireless Channels Using Intelligent Reflecting Surfaces***

Emil Björnson (Linköping University, Sweden); Jide Yuan, Elisabeth de Carvalho, Robin Williams and Petar Popovski (Aalborg University, Denmark)

Wireless propagation channels are supposed to be linear systems and, if there is no motion, also time-invariant. But what if we could change those foundational principles; what benefits could that bring to wireless communications? In this invited talk, we will describe the novel concept of a frequency-modulating intelligent reflecting surface (FM-IRS). This is a concept that makes use of real-time reconfigurable metasurfaces that can shape the propagation environments. In previous works, this feature has been used to create new reflection behaviors (diffuse and specular), for example, to increase the received signal power at desired locations or reduce interference. However, there are other potential use cases around the corner. We will describe how an IRS can operate as a frequency mixer, which moves the reflected signal to other bands. This feature will effectively create a non-linear propagation channel where the received signal contains other frequencies than the transmitted signal. The talk will cover the potential benefits this new feature can have when it comes to channel estimation, channel capacity, and diversity.

17:20 ***Adaptive Two-Sided Beam Alignment in mmWave via Posterior Matching***

Fernando Pedraza and Giuseppe Caire (Technische Universität Berlin, Germany)

Millimeter wave band communication is an enabler technology for multi Gbps data rates, though with associated technical challenges. Among them are the very high pathloss and the crucial importance of smart signal processing to overcome it. This paper addresses the problem of identifying the location of the strongest scatterers of the channel in the angle domain. The base station in our system operates agnostically with respect to the users, who estimate the strongest path locally and communicate it back just at the end of the process, resulting in a scalable system. The users decide adaptively the directions to test based on a posterior matching approach. Numerical results confirm the validity of our scheme, especially when the time available for beam alignment is limited.

Tuesday, April 13 18:00 - 18:50 (Europe/Rome)

Cryptography, Privacy and Security I

Room 1

Chair: Marco Baldi (Università Politecnica delle Marche, Italy)

18:00 ***Concentrated Stopping Set Design for Coded Merkle Tree: Improving Security Against Data Availability Attacks in Blockchain Systems***

Debarnab Mitra and Lev Tauz (University of California, Los Angeles, USA); Lara Dolecek (UCLA, USA)

In certain blockchain systems, light nodes are clients that download only a small portion of the block. Light nodes are vulnerable to data availability (DA) attacks where a malicious node hides an invalid portion of the block from the light nodes. Recently, a technique based on erasure codes called Coded Merkle Tree (CMT) was proposed by Yu et al. that enables light nodes to detect a DA attack with high probability. The CMT is constructed using LDPC codes for fast decoding but can fail to detect a DA attack if a malicious node hides a small stopping set of the code. To combat this, Yu et al. used well-studied techniques to design random LDPC codes with high minimum stopping set size. Although effective, these codes are not necessarily optimal for this application. In this paper, we demonstrate a more specialized LDPC code design to improve the security against DA attacks. We achieve this goal by providing a deterministic LDPC code construction that focuses on concentrating stopping sets to a small group of variable nodes rather than only eliminating stopping sets. We design these codes by modifying the Progressive Edge Growth algorithm into a technique called the entropy-constrained PEG (EC-PEG) algorithm. This new method demonstrates a higher probability of detecting DA attacks and allows for good codes at short lengths.

18:10 *Multilevel Secrecy over 1-2-1 Networks*

Yahya H. Ezzeldin (University of Southern California, USA & Alexandria University, Egypt); Martina Cardone (University of Minnesota, USA); Christina Fragouli (UCLA, USA)

This paper studies the problem of secure communication over noiseless 1-2-1 networks, an abstract model for networks with directional communication capabilities such as mmWave networks. A secure transmission scheme is designed and shown to achieve a secure rate that is larger than state-of-the-art lower bounds for a class of 1-2-1 network topologies. The proposed scheme leverages the scheduling nature of 1-2-1 networks, the network topology, as well as storage at intermediate nodes to create shared randomness with the source to improve the secure rate. Finally, a novel outer bound is derived and shown to match the achievability bound under certain network conditions, hence characterizing the secure capacity in such regimes.

18:20 *Pairwise Oblivious Transfer*

Remi A Chou (Wichita State University, USA)

We consider oblivious transfer between one client and two non-colluding servers that store independent contents. The client requests one file from each server such that (i) the servers do not learn the file selection of the client, and (ii) the client does not learn information about the non-selected files. We show under the honest-but-curious assumption that oblivious transfer with non-zero rates can be achieved with a multiuser protocol between the client and the servers. This contrasts with the case where the client engages in independent protocols with each of the two servers, for which it is known that oblivious transfer with information-theoretic security is impossible in the absence of additional resources. Furthermore, we derive a capacity result for the proposed setting.

18:30 *Higher Rates and Information-Theoretic Analysis for the RLWE Channel*

Georg Maringer (Technische Universität München, Germany); Sven Puchinger (Technical University of Munich, Germany); Antonia Wachter-Zeh (Technical University of Munich (TUM), Germany)

The \emph{Learning with Errors} (LWE) problem is considered to be a hard problem and lies the foundation of various cryptographic algorithms. Several cryptosystems based on the closely related \emph{Ring Learning with Errors} (RLWE) problem have been proposed within the NIST PQC standardization process,

e.g., the systems LAC and NewHope. The combination of encryption and decryption for these kinds of algorithms can be interpreted as data transmission over noisy channels. To the best of our knowledge this paper is the first work that analyzes the capacity of this channel. We extend this channel from binary to q-ary alphabets and show that this does not compromise the security of the related RLWE-based schemes if appropriate error correcting codes are used to prevent the \textit{decryption failure rate} (DFR) from increasing. We give a lower bound on the capacity of this channel showing that the achievable asymptotic rates are substantially (5.7 times for LAC and 10.7 times for NewHope) higher than the currently deployed ones for the finite length regime. Furthermore, under the assumption of stochastically independent coefficient failures, we show that substantially higher rates can also be achieved in the finite length setting by using the Gilbert-Varshamov bound. Moreover, we give explicit code constructions increasing the achievable rate by a factor of 2 for LAC and a factor of 7 for NewHope without increasing the DFR for the respective parameter sets achieving a security level equivalent to AES256.

18:40 *Optimal Linear Coding Schemes for the Secure Decentralized Pliable Index Coding Problem*

Tang Liu (Princeton University, USA); Daniela Tuninetti (University of Illinois Chicago, USA)

This paper studies the secure decentralized Pliable Index CODing (PICOD) problem, where the security constraint forbids users to decode more than one message while the decentralized setting imposes that there is no central transmitter in the system, and thus transmissions occur only among users. A converse bound from the Authors' previous work showed a factor of three difference in optimal code-length between the centralized and the decentralized versions of the problem, under the constraint of linear encoding. This paper first lists all linearly infeasible cases, that is, problems where no linear code can simultaneously achieve both correctness/decodability and security. Then, it proposes linear coding schemes for the remaining cases and shows that their code-length is to within an additive constant gap from the converse bound.

Machine Learning I

Room 2

Chair: Stefan M. Moser (ETH Zurich, Switzerland & National Chiao Tung University (NCTU), Taiwan)

18:00 *Jensen-Shannon Information Based Characterization of the Generalization Error of Learning Algorithms*

Gholamali Aminian (University College London (UCL), United Kingdom (Great Britain)); Laura Toni and Miguel Rodrigues (University College London, United Kingdom (Great Britain))

Generalization error bounds are critical to understanding the performance of machine learning models. In this work, we propose a new information-theoretic based generalization error upper bound applicable to supervised learning scenarios. We show that our general bound can specialize in various previous bounds. We also show that our general bound can be specialized under some conditions to a new bound involving the Jensen-Shannon information between a random variable modelling the set of training samples and another random variable modelling the hypothesis. We also prove that our bound can be tighter than

mutual information-based bounds under some conditions.

18:10 *Learning, compression, and leakage: Minimizing classification error via meta-universal compression principles*

Fernando Rosas (Imperial College London & Centre of Complexity Science, United Kingdom (Great Britain)); Pedro Mediano (University of Cambdrige, United Kingdom (Great Britain)); Michael Gastpar (EPFL, Switzerland)

Learning and compression are driven by the common aim of identifying and exploiting statistical regularities in data, which opens the door for fertile collaboration between these areas. A promising group of compression techniques for learning scenarios is normalised maximum likelihood (NML) coding, which provides strong guarantees for compression of small datasets - in contrast with more popular estimators whose guarantees hold only in the asymptotic limit. Here we consider a NML-based decision strategy for supervised classification problems, and show that it attains heuristic PAC learning when applied to a wide variety of models. Furthermore, we show that the misclassification rate of our method is upper bounded by the maximal leakage, a recently proposed metric to quantify the potential of data leakage in privacy-sensitive scenarios.

18:20 *RNN-based Detection for Coded Partial-Response Channels*

Simeng Zheng, Yi Liu and Paul H. Siegel (University of California, San Diego, USA)

In this paper, we investigate the use of recurrent neural network (RNN)-based detection of magnetic recording channels with inter-symbol interference (ISI). We refer to the proposed detection method, which is intended for recording channels with partial-response equalization, as Partial-Response Neural Network (PR-NN). We train bi-directional gated recurrent units (bi-GRUs) to recover the ISI channel inputs from noisy channel output sequences and evaluate the network performance when applied to continuous, streaming data. The recording system on which the experiments were conducted uses a rate-2/3, (1,7) runlength-limited (RLL) code with an E2PR4 partial-response channel target. Experimental results with ideal PR signals show that the performance of PR-NN detection approaches that of Viterbi detection in additive white gaussian noise (AWGN). Moreover, the PR-NN detector outperforms Viterbi detection and achieves the performance of Noise-Predictive Maximum Likelihood (NPML) detection in additive colored noise (ACN).

18:30 *Measuring Dependencies of Order Statistics: An Information Theoretic Perspective*

Alex Dytso (New Jersey Institute of Technology, USA); Martina Cardone (University of Minnesota, USA); Cynthia Rush (Columbia University, USA)

This work considers a random sample \(X_1, X_2, ..., X_n\) drawn independently and identically distributed from some known parent distribution \(P_X\) with \(X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}\) being the order statistics of the sample. Under the assumption of an invertible cumulative distribution function associated with the parent distribution \(P_X\), a distribution-free property is established showing that the \(f\)-divergence between the joint distribution of order statistics and the product distribution of order statistics does not depend on \(P_X\). Moreover, it is shown that the mutual information between two subsets of order statistics also satisfies a distribution-free property; that is, it does not depend on \(P_X\). Furthermore, the decoupling rates between \(X_{(r)}\) and \(X_{(m)}\) (i.e., rates at which the mutual information approaches zero) are characterized for various choices of \((r, m)\). The work also considers discrete distributions, which do not satisfy the previously-stated invertibility assumption, and it is shown

that no such distribution-free property holds: the mutual information between order statistics does depend on the parent distribution \(P_X\). Upper bounds on the decoupling rates in the discrete setting are also established.

18:40 *Sphere Covering for Poisson Processes*

Hui-An Shen (University of Bern, Switzerland); Stefan M. Moser (ETH Zurich, Switzerland & National Chiao Tung University (NCTU), Taiwan); Jean-Pascal Pfister (University of Bern, Switzerland)

The geometric interpretation of sphere covering describing the rate distortion problem of a Gaussian source with the squared-error distortion measure is generalized to a Laplacian source and the \(\ell_1\)-distortion measure. Using additional constraints on the distortion measure, sphere covering is further generalized to exponential sources and to Poisson point processes.

Tuesday, April 13 19:00 - 20:00 (Europe/Rome)

Plenary - Data Driven Algorithm Design

Maria Florina Balcan

Chair: Olgica Milenkovic (University of Illinois at Urbana-Champaign (UIUC), USA)

Data driven algorithm design for combinatorial problems is an important aspect of modern data science. Rather than using off the shelf algorithms that only have worst case performance guarantees, practitioners typically optimize over large families of parametrized algorithms and tune the parameters of these algorithms using a training set of problem instances from their domain to determine a configuration with high expected performance over future instances. However, most of this work comes with no performance guarantees. The challenge is that for many combinatorial problems, including partitioning and subset selection problems, a small tweak to the parameters can cause a cascade of changes in the algorithm's behavior, so the algorithm's performance is a discontinuous function of its parameters.

In this talk, I will present new work that helps put data driven combinatorial algorithm selection on firm foundations. This includes strong computational and statistical performance guarantees, both for the batch and online scenarios where a collection of typical problem instances from the given application are presented either all at once or in an online fashion, respectively. I will describe both specific examples (for clustering, partitioning, and subset selection problems) and general principles that emerge in this context (including general techniques for sample complexity guarantees in the batch setting and no-regret guarantees in the online settings).

Tuesday, April 13 20:20 - 22:00 (Europe/Rome)

Themed Session - Statistical Physics and Machine Learning

Chairs: Sebastian Goldt (International School of Advanced Studies (SISSA), Italy), Florent Krzakala (Ecole Normale Superieure, France)

20:20 *Exactly solvable models for high-dimensional machine learning problems*

Bruno Loureiro (EPFL, Switzerland)

In this talk I will introduce an exactly-solvable model describing learning with correlated features. I will motivate how this model encompasses many learning problems of interest, e.g. ridge and logistic regression, learning with random features and scattering transforms and transfer learning. Finally, I will show how in the specific case of ridge regression this model can surprisingly reproduce learning curves from these learning problems on real data.

20:40 *Information-Theoretic Limits for the Matrix Tensor Product*

Galen Reeves (Duke University, USA)

This talk considers a high-dimensional inference problem involving the matrix tensor product of random matrices. This problem generalizes a number of contemporary data science problems including the spiked matrix models used in sparse principal component analysis and covariance estimation. We show that the information-theoretic limits can be described succinctly by formulas involving low-dimensional quantities. Our approach introduces some new techniques for the analysis of high-dimensional matrix-valued signals including a novel extension of the adaptive interpolation method that uses order-preserving positive semidefinite interpolation paths and a variance inequality based on continuous-time I-MMSE relations.

21:00 *Learning with Random Features: Sharp Asymptotics and Universality Laws*

Yue M. Lu (Harvard University, USA)

We study the problem of learning an unknown function using random feature models. Under mild regularity conditions for the feature matrix, we provide an exact characterization of the asymptotic training and generalization errors, valid in both the under-parameterized and over-parameterized regimes. Our results reveal the important roles played by the regularization, the loss function and the activation function in the mitigation of the "double descent phenomenon" in learning. The asymptotic analysis is made possible by a universality theorem, which states that, in terms of training and generalization errors, the random feature model with a nonlinear activation function is asymptotically equivalent to a surrogate Gaussian model with a matching covariance matrix. Our method for proving the universality builds on the classical Lindeberg approach. Major ingredients of the proof include a leave-one-out analysis for the optimization problem associated with the training process and a central limit theorem, obtained via Stein's method, for weakly correlated random variables.

21:20 *Strong concentration of measure for log-concave measures and in optimal Bayesian inference*

Jean Barbier (The Abdus Salam International Center for Theoretical Physics, Italy)

The phenomenon of concentration of measure is a pillar of our understanding of high-dimensional statistical models. There exists a large body of work on concentration in the context of log-concave measures -that appear, e.g., in convex optimisation in machine learning or robust statistics like M-estimation etc-, and optimal Bayesian inference. Such results are crucial in order to analyse fundamental information-theoretic limitations to inference and learning. I will discuss recent results that prove that this phenomenon takes place in a sense stronger than previously studied (we coin this « strong replica symmetry « , borrowing terminology from statistical mechanics), and under very generic conditions.

21:40 *Algorithmic advances on the negative spherical perceptron problem*

Ahmed El Alaoui (UC Berkeley, USA)

We consider the spherical perceptron problem: this is the problem of finding a point on the sphere subject to a number of random linear inequalities. Variants of this problem are also known as discrepancy minimization in combinatorics and theoretical computer science. I will discuss an algorithmic approach attempting to compute a valid solution up to the satisfiability threshold predicted by statistical physics, conditional on the validity of a conjecture concerning a certain variational formula, known as 'no overlap gap' or 'full replica-symmetry breaking'. I will discuss a possible extension to the binary perceptron, where the desired solution must have binary coordinates. This is joint work with Mark Sellke.

Wednesday, April 14

Wednesday, April 14 9:00 - 9:50 (Europe/Rome)

Coding II: Codes on Graphs and Polar Codes weconf

Room 1

Chair: Luca Barletta (Politecnico di Milano, Italy)

9:00 *Construction of Systematic Polar Codes: BER Optimization Perspective*

Bolin Wu, Kai Niu and Jincheng Dai (Beijing University of Posts and Telecommunications, China)

Code construction is a critical issue for polar coding. The expected construction method is of an accurate estimate of the reliability of bit-channels, but the current methods usually require high computational complexity. In this paper, we concern the input-output weight distribution of each bit-channel and derive its recursive calculation algorithm for systematic coding. The union bound and union-Bhattacharyya bound on the bit error probability are also derived to evaluate the reliability of bit-channels. Furthermore, by calculating the logarithmic form of the union-Bhattacharyya bound, we also propose two novel construction methods named the union-Bhattacharyya bound weight of the bit error probability (UBWB) and the simplified UBWB (SUBWB). Numerical results show that the proposed UBWB/SUBWB construction methods can achieve comparable performance to current methods under successive cancellation (SC) decoding and obtain obvious performance gain under SC list (SCL) decoding.

9:10 *Polar Coded Repetition for Low-Capacity Channels*

Fariba Abbasi (Monash University, Australia); Hessam Mahdavifar (University of Michigan, USA); Emanuele Viterbo (Monash University, Australia)

Constructing efficient low-rate error-correcting codes with low-complexity encoding and decoding have become increasingly important for applications involving ultra-low-power devices such as Internet-of-Things (IoT) networks. To this end, schemes based on concatenating the state-of-the-art codes at moderate rates with repetition codes have emerged as practical solutions deployed in various standards. In this paper, we propose a novel mechanism for concatenating outer polar codes with inner repetition codes

which we refer to as polar coded repetition. More specifically, we propose to transmit a slightly modified polar codeword by deviating from Arıkan's standard 2 × 2 Kernel in a certain number of polarization recursions at each repetition block. We show how this modification can improve the asymptotic achievable rate of the polar-repetition scheme, while ensuring that the overall encoding and decoding complexity is kept almost the same. The achievable rate is analyzed for the binary erasure channels (BEC).

9:20 ***Matched Quantized Min-Sum Decoding of Low-Density Parity-Check Codes***

Emna Ben Yacoub (Technical University of Munich, Germany)

A quantized message passing decoding algorithm for low-density parity-check codes is presented. The algorithm relies on the min approximation at the check nodes, and on modelling the variable node inbound messages as observations of an extrinsic discrete memoryless channel. The performance of the algorithm is analyzed and compared to quantized min-sum decoding by means of density evolution, and almost closes the gap with the performance of the sum-product algorithm. A stability analysis is derived, which highlights the role played by degree-3 variable nodes in the stability condition. Finite-length simulation results confirm large gains predicted by the asymptotic analysis.

9:30 ***On Decoding Fountain Codes with Erroneous Received Symbols***

Xuan He (Southwest Jiaotong University, China); Kui Cai (Singapore University of Technology and Design, Singapore)

Motivated by the application of fountain codes in the DNA-based data storage systems, in this paper, we propose the basis-finding algorithm (BFA) for decoding fountain codes for an erasure-error channel where each received symbol has a fixed probability to be correct. The key idea of the BFA is to find a basis of the received symbols, and then use the most reliable basis elements to recover the source symbols with the inactivation decoding. Gaussian elimination can be used to find the basis and to identify the most reliable basis elements. For random fountain codes, we are able to derive some theoretical bounds for the frame error rate (FER) of the BFA, which reveal that the BFA can perform very well for decoding fountain codes for the considered channel.

9:40 ***Two is Better than One: Reducing the Loss of the Window Decoder for SC-LDPC Codes***

Alberto Tarable and Marco Ferrari (CNR-IEIIT, Italy); Luca Barletta (Politecnico di Milano, Italy)

In this paper, we consider spatially coupled LDPC codes derived from protographs. In particular, we analyze the performance of the window decoder (WD), which allows reducing the complexity, the memory requirements, and the latency of the flood belief-propagation decoder. We show that the performance degradation of WD is due to the fact that it exploits a single decoding wave instead of two. This has effect both in the ideal case of infinite code length, where it may imply a threshold loss, and in the case of finite length, where it affects the slope of the BER curve in the waterfall region. We show how a forward-backward decoder can reduce such problems at the price of a limited increase of average complexity.

Multi-User Information Theory

Room 2

Chair: Franco Chiaraluce (Università Politecnica delle Marche, Italy)

9:00 ***Information Freshness and Packet Drop Rate Interplay in a Two-User Multi-Access Channel***

Emmanouil Fountoulakis (Linköping University, Sweden); Themistoklis Charalambous (Aalto University, Finland); Nikolaos Nomikos (University of Cyprus, Cyprus); Anthony Ephremides (University of Maryland, USA); Nikolaos Pappas (Linköping University, Sweden)

In this work, we combine the two notions of timely delivery of information to study their interplay; namely, deadline-constrained packet delivery due to latency constraints and freshness of information. More specifically, we consider a two-user multiple access setup with random-access, in which user 1 is a wireless device with a queue and has external bursty traffic which is deadline-constrained, while user 2 monitors a sensor and transmits status updates to the destination. We provide analytical expressions for the throughput and drop probability of user 1, and an analytical expression for the average Age of Information (AoI) of user 2 monitoring the sensor. The relations reveal that there is a trade-off between the average AoI of user 2 and the drop rate of user 1: the lower the average AoI, the higher the drop rate, and vice versa. Simulations corroborate the validity of our theoretical results.

9:10 ***Zero-Error Capacity Region of a Class of Multiple Access Channels with Inter-User Correlation***

Ghassen Zafzouf and Girish N. Nair (University of Melbourne, Australia)

In this paper, we investigate zero-error communication over M-user multiple access channels with correlated transmitters. The correlation among the users is modeled by means of both a common message seen by all encoders as well as pairwise shared messages. Motivated by networked control systems, we use tools from nonstochastic information theory to characterize the zero-error capacity region of the investigated channel. To this end, both converse and achievability proofs are provided.

9:20 ***Sliding-Window Gelfand-Pinsker Coding: General K-User Broadcast Channels***

Shouvik Ganguly (University of California, San Diego, USA); Lele Wang (University of British Columbia, Canada)

A low-complexity coding scheme, termed as sliding-window Gelfand-Pinsker coding, is proposed. It is shown that in a general K-user broadcast channel, every rate point in the Marton's inner bound can be achieved using single-user encoders and decoders. The scheme provides us with a low-complexity alternative to implement the conceptual K dimensional multi-coding, which is an irreplaceable component in many important network communication schemes, such as Marton coding in Gaussian MIMO broadcast channels and distributed decode-forward in cloud radio access networks, but has not been adopted in practical systems due to high computational complexity. Key features in the proposed scheme include staggered message scheduling, successive Gelfand-Pinsker coding, and sliding-window decoding.

9:30 ***Joint Sensing and Communication over Memoryless Broadcast Channels***

Mehrasa Ahmadipour and Michele A Wigger (Telecom Paris, France); Mari Kobayashi (CentraleSupelec, France)

A memoryless state-dependent broadcast channel (BC) is considered, where the transmitter wishes to convey two private messages to two receivers while simultaneously estimating the respective states via

generalized feedback. The model at hand is motivated by a joint radar and communication system where radar and data applications share the same frequency band. For physically degraded BCs with i.i.d. state sequences, we characterize the capacity-distortion region tradeoff. For general BCs, we provide inner and outer bounds on the capacity-distortion region, as well as a sufficient condition when it is equal to the product of the capacity region and the set of achievable distortion. Interestingly, the proposed synergetic design significantly outperforms a conventional approach that splits the resource either for sensing or communication.

9:40 *Random User Activity with Mixed Delay Traffic*

Homa Nikbakht (Inria, France); Michele A Wigger (Telecom Paris, France); Shlomo (Shitz) Shamai (The Technion, Israel)

This paper analyses the multiplexing gain (MG) achievable over a general interference network with random user activity and random arrival of mixed-delay traffic. The mixed-delay traffic is composed of delay-tolerant traffic and delay-sensitive traffic where only the former can benefit from receiver cooperation since the latter is subject to stringent decoding delays. Two setups are considered. In the first setup, each active transmitter always has delay-tolerant data to send and delay-sensitive data arrival is random. In the second setup, both delay-tolerant and delay-sensitive data arrivals are random, and only one of them is present at any given transmitter. The MG regions of both setups are completely characterized for Wyner's soft-handoff network. For Wyner's symmetric linear and hexagonal networks inner bounds on the MG region are presented.

Wednesday, April 14 10:00 - 10:40 (Europe/Rome)

Coding III: Coded Caching

Room 1

Chair: Philippe Ciblat (Telecom Paris & Institut Polytechnique de Paris, France)

10:00 *Resolving the Worst-User Bottleneck of Coded Caching: Exploiting Finite File Sizes*

Hui Zhao, Antonio Bazco-Nogueras and Petros Elia (EURECOM, France)

In this work, we address the worst-user bottleneck of coded caching, which is known to diminish any caching gains due to the fundamental requirement that the multicast transmission rate should be limited by that of the worst channel among the served users. We consider the quasi-static Rayleigh fading Broadcast Channel, for which we first show that the coded caching gain of the XOR-based standard coded-caching scheme completely vanishes in the low-SNR regime. Yet, we show that this collapse is not intrinsic to coded caching by presenting a novel scheme that can completely recover the caching gains. The scheme exploits an aspect that has remained unexploited: the shared side information brought about by the file size constraint. The worst-user effect is dramatically ameliorated because it is replaced by the worst-group-of-users effect, where the users within a group have the same side information and the grouping is decided before the channel or the demands are known.

10:10 *Multi-access Coded Caching Schemes From Cross Resolvable Designs*

Digvijay Katyal and Pooja Nayak Muralidhar (Indian Institute of Science Bangalore, India); B. Sundar Rajan (Indian Institute of Science, India)

We present a novel caching and coded delivery scheme for a multi-access network where multiple users can have access to the same cache (shared cache) and any cache can assist multiple users. This scheme is obtained from resolvable designs satisfying certain conditions which we call {\it cross resolvable designs}. To be able to compare different multi-access coded schemes with different number of users we normalize the rate of the schemes by the number of users served. Based on this per-user-rate we show that our scheme performs better than the recently proposed ("Multi-access coded caching: gains beyond cache-redundancy" by Serbetci, Parrinello and Elia) SPE scheme. It is shown that the resolvable designs from affine planes are cross resolvable designs and our scheme can be used based on these. The SPE scheme considers only the cases where the product of the number of users and the normalized cache size is 2, whereas the proposed scheme allows different choices depending on the choice of the cross resolvable design.

10:20 *Coded Caching For Relay Networks: The Impact of Caching Memories*

Shu-Jie Cao, Jiahui Chen, Youlong Wu and Ke Wang (ShanghaiTech University, China)

Relay is a traditional key technology to improve the communication reliability and enlarge the covering range of service. Recently, coded caching schemes that reduce traffic congestion through coding and injecting duplicate data among users have attracted wide interests. This paper studies a relay network where all nodes including the central server, relay nodes and users are equipped with cache memories. Each user demands a file from the server's library, and is connected to the server through a specific relay node. We define the communication delay for this model and propose new coded caching schemes for the deterministic and random caching setups, respectively. The proposed schemes exploit the spared transmission time resource and can greatly reduce the transmission delay compared to the previously known caching schemes. Surprisingly, we show that even when relay nodes do not cooperate with each other, using a small amount of caching memories at each relay node is sufficient to achieve the same communication delay as if each relay had access to the full library. To our best knowledge, this is the first result showing that even the caching size is strictly smaller than the library's size, increasing the caching size is wasteful in reducing the transmission latency.

10:30 *Cache Updating Strategy Minimizing the Age of Information with Time-Varying Files' Popularities*

Haoyue Tang (Tsinghua University, China); Philippe Ciblat (Telecom Paris & Institut Polytechnique de Paris, France); Jintao Wang (Tsinghua University, China); Michele A Wigger (Telecom Paris, France); Roy Yates (Rutgers University, USA)

We consider updating strategies for a local cache which downloads time-sensitive files from a remote server through a bandwidth-constrained link. The files are requested randomly from the cache by local users according to a popularity distribution which varies over time according to a Markov chain structure. We measure the freshness of the requested time-sensitive files through their Age of Information (AoI). The goal is then to minimize the average AoI of all requested files by appropriately designing the local cache's downloading strategy. To achieve this goal, the original problem is relaxed and cast into a Constrained Markov Decision Problem (CMDP), which we solve using a Lagrangian approach and Linear Programming. Inspired by this solution for the relaxed problem, we propose a practical cache updating strategy that

meets all the constraints of the original problem. Under certain assumptions, the practical updating strategy is shown to be optimal for the original problem in the asymptotic regime of a large number of files. For a finite number of files, we show the gain of our practical updating strategy over the traditional square-root-law strategy (which is optimal for fixed non time-varying file popularities) through numerical simulations.

Source Coding

Room 2

Chair: Ferdinando Cicalese (Universita' di Verona, Italy)

10:00 ***Zero-Error Sum Modulo Two with a Common Observation***

Milad Sefidgaran (Télécom Paris, France); Aslan Tchamkerten (Telecom ParisTech, France)

This paper investigates the classical modulo two sum problem in source coding, but with a common observation: a transmitter observes \((X, Z)\), the other transmitter observes \((Y, Z)\), and the receiver wants to compute \(X \oplus Y \) without error. Through a coupling argument, this paper establishes a new lower bound on the sum-rate when \(X-Z-Y\) forms a Markov chain.

10:10 ***An Algorithm for Constructing the Optimal Code Trees for Binary Alphabetic AIFV-m Codes***

Ken-ichi Iwata (University of Fukui, Japan); Hirosuke Yamamoto (The University of Tokyo, Japan)

We call the alphabetic version of the AIFV-m code the alphabetic AIFV-m codes. This paper defines binary alphabetic AIFV-m codes and proposes an algorithm to design the optimal binary alphabetic AIFV-m codes in terms of the minimum average codeword length for stationary memoryless sources. The proposed method is based on an iterative optimization algorithm and a dynamic programming algorithm.

10:20 ***Cooperative Multi-Sensor Detection under Variable-Length Coding***

Mustapha Hamad (Télécom Paris, France); Michele A Wigger (Telecom Paris, France); Mireille Sarkiss (Telecom SudParis, France)

We investigate the testing-against-independence problem over a cooperative MAC with two sensors and a single detector under an average rate constraint on the sensors-detector links. For this setup, we design a variable-length coding scheme that maximizes the achievable type-II error exponent when the type-I error probability is limited to \(\epsilon\). Similarly to the single-link result, we show here that the optimal error exponent depends on \(\epsilon\) and that variable-length coding allows to increase the rates over the optimal fixed-length coding scheme by the factor \((1-\epsilon)^{-1}\).

10:30 ***Universal Decoding for Asynchronous Slepian-Wolf Encoding***

Neri Merhav (Technion, Israel)

We consider the problem of (almost) lossless source coding of two correlated memoryless sources using separate encoders and a joint decoder, that is, Slepian-Wolf (S-W) coding. In our setting, the encoding and decoding are asynchronous, i.e., there is a certain relative delay between the two sources. Neither the

source parameters nor the relative delay are known to the encoders and the decoder. Since we assume that both encoders implement standard random binning, which does not require such knowledge anyway, the focus of this work is on the decoder. Our main contribution is in proposing a universal decoder, that independent of the unknown source parameters and the relative delay, and at the same time, is asymptotically as good as the optimal maximum a posteriori probability (MAP) decoder in the sense of the random coding error exponent achieved. Consequently, the achievable rate region is also the same as if the source parameters and the delay were known to the decoder.

Wednesday, April 14 11:00 - 11:50 (Europe/Rome)

Coding IV: Coding applications

Room 1

Chair: Luca Barletta (Politecnico di Milano, Italy)

11:00 ***Decoding of Lifted Affine-Invariant Codes***

Lukas Holzbaur and Nikita Polyanskii (Technical University of Munich, Germany)

Lifted Reed-Solomon codes, a subclass of lifted affine-invariant codes, have been shown to be of high rate while preserving locality properties similar to generalized Reed-Muller codes, which they contain as subcodes. This work introduces a simple bounded distance decoder for (subcodes of) lifted affine-invariant codes that is guaranteed to decode up to half of an asymptotically tight bound on their minimum distance. Further, long q-ary lifted affine-invariant codes are shown to correct almost all error patterns of relative weight $\frac{q-1}{q}-\epsilon$ for $\epsilon>0$.

11:10 ***Feedback insertion-deletion codes***

Georg Maringer (Technische Universität München, Germany); Nikita Polyanskii (Technical University of Munich, Germany); Ilya Vorobyev (Skolkovo Institute of Science and Technology, Russia); Lorenz Welter (Technical University of Munich(TUM), Germany)

In this paper, a new problem of transmitting information over the adversarial insertion-deletion channel with feedback is introduced. Suppose that we can transmit n binary symbols one-by-one over the channel, in which some symbols can be deleted and some additional symbols can be inserted. After each transmission, we see whether insertions or deletions have occurred so that our further encoding strategy can be adjusted. The goal is to transmit as much information as possible under the assumption that the total number of deletions and insertions is limited by τn, $0<\tau<1$. We show how this problem can be reduced to a well-researched problem, in which the only type of errors is substitution. Thereby, the maximal asymptotic rate of insertion-deletion codes is completely established.

11:20 ***On Lifted Multiplicity Codes***

Lukas Holzbaur (Technical University of Munich, Germany); Rina Polyanskaya (Institute for Information Transmission Problems, Russia); Nikita Polyanskii (Technical University of Munich, Germany); Ilya Vorobyev (Skolkovo Institute of Science and

Technology, Russia); Eitan Yaakobi (Technion, Israel)

Lifted Reed-Solomon codes and multiplicity codes are two classes of evaluation codes that allow for the design of high-rate codes that can recover every codeword or information symbol from many disjoint sets. Recently, the underlying approaches have been combined to construct lifted bi-variate multiplicity codes, that can further improve on the rate. We continue the study of these codes by providing lower bounds on the rate and distance for lifted multiplicity codes obtained from polynomials in an arbitrary number of variables. Specifically, we investigate a subcode of a lifted multiplicity code formed by the linear span of m-variate monomials whose restriction to an arbitrary line in $\mathbb{F}_q^m$ is equivalent to a low-degree uni-variate polynomial. We find the tight asymptotic behavior of the fraction of such monomials when the number of variables m is fixed and the alphabet size $q=2^\ell$ is large. For some parameter regimes, lifted multiplicity codes are then shown to have a better trade-off between redundancy and the number of disjoint recovering sets for every codeword or information symbol than previously known constructions.

11:30 ***Reconstructing Mixtures of Coded Strings from Prefix and Suffix Compositions***

Ryan Gabrys (SPAWAR Pacific, USA); Srilakshmi Pattabiraman (University of Illinois at Urbana-Champaign, USA); Olgica Milenkovic (University of Illinois at Urbana-Champaign (UIUC), USA)

The problem of string reconstruction from substring information has found many applications due to its relevance in DNA- and polymer-based data storage. One practically important and challenging paradigm requires reconstructing mixtures of strings based on the union of compositions of their prefixes and suffixes, generated by mass spectrometry readouts. We describe new coding methods that allow for unique joint reconstruction of subsets of strings selected from a code and provide matching upper and lower bounds on the asymptotic rate of the underlying codebooks. Under certain mild constraints on the problem parameters, one can show that the largest possible rate of a codebook that allows for all subcollections of less than or equal to h codestrings to be uniquely reconstructable from the prefix-suffix information equals $1/h$.

11:40 ***Success Probability of Decoding Interleaved Alternant Codes***

Lukas Holzbaur (Technical University of Munich, Germany); Hedongliang Liu (Institute for Communications Engineering & Technical University of Munich, Germany); Alessandro Neri (University of Zurich, Switzerland); Sven Puchinger (Technical University of Munich, Germany); Johan S. H. Rosenkilde (Technical University of Denmark, Denmark); Vladimir Sidorenko (Technical University of Munich, Germany); Antonia Wachter-Zeh (Technical University of Munich (TUM), Germany)

Interleaved Reed-Solomon codes admit efficient decoding algorithms which correct burst errors far beyond half the minimum distance in the random errors regime, e.g., by computing a common solution to the Key Equation for each Reed-Solomon code, as described by Schmidt et al. If this decoder does not succeed, it may either fail to return a codeword or miscorrect to an incorrect codeword, and good upper bounds on the fraction of error matrices for which these events occur are known. The decoding algorithm immediately applies to interleaved alternant codes as well, i.e., the subfield subcodes of interleaved Reed-Solomon codes, but the fraction of decodable error matrices differs, since the error is now restricted to a subfield. In this paper, we present new general lower and upper bounds on the fraction of decodable error matrices

by Schmidt et al.'s decoding algorithm, thereby making it the only decoding algorithm for interleaved alternant codes for which such bounds are known.

Statistics and Information Theory I

Room 2

Chair: Massimo Battaglioni (Università Politecnica delle Marche, Italy)

11:00 ***A Generalization of the DMC***

Sergey Tridenski and Anelia Somekh-Baruch (Bar-Ilan University, Israel)

We consider a generalization of the discrete memoryless channel, in which the channel probability distribution is replaced by a uniform distribution over clouds of channel output sequences. For a random ensemble of such channels, we derive an achievable error exponent and a converse bound on the correct-decoding exponent. As a corollary of these results, we obtain the channel ensemble capacity.

11:10 ***On the Communication Exponent of Distributed Testing for Gaussian Correlations***

Yuval Kochman (The Hebrew University of Jerusalem, Israel); Ligong Wang (ETIS & CNRS, France)

This work addresses distributed binary hypothesis testing, where observations at two terminals are jointly Gaussian, each one standard, with two possible correlation coefficients. We assume that one of the terminals is colocated with the decision center, and focus on a single (Stein) error exponent. Rather than the traditional exponent that is defined with respect to the source blocklength, we assume the source data to be unlimited, and consider the error exponent as a function of the communication message length. We examine two different approaches, one by quantization and the other by sending the index of the maximum, and find them to yield the same exponent. We further find that binning improves upon both approaches in the same way. Finally we compare the obtained exponents to two upper bounds and determine the optimal exponent in some very special cases.

11:20 ***Unbiased Estimation Equation under \(f\)-Separable Bregman Distortion Measures***

Masahiro Kobayashi and Kazuho Watanabe (Toyohashi University of Technology, Japan)

We discuss unbiased estimation equations in a class of objective function using a monotonically increasing function \(f\) and Bregman divergence. The choice of the function \(f\) gives desirable properties such as robustness against outliers. In order to obtain unbiased estimation equations, analytically intractable integrals are generally required as bias correction terms. In this study, we clarify the combination of Bregman divergence, statistical model, and function \(f\) in which the bias correction term vanishes. Focusing on Mahalanobis and Itakura-Saito distances, we provide a generalization of fundamental existing results and characterize a class of distributions of positive reals with a scale parameter, which includes the gamma distribution as a special case. We discuss the possibility of latent bias minimization when the proportion of outliers is large, which is induced by the extinction of the bias correction term.

11:30 ***On Functions of Markov Random Fields***

Bernhard C. Geiger (Know-Center GmbH, Austria); Ali Al-Bashabsheh (Beijing Advanced Innovation Center for Big Data and Brain Computing (BDBC), Beihang University, China)

We derive two sufficient conditions for a function of a Markov random field (MRF) on a given graph to be a MRF on the same graph. The first condition is information-theoretic and parallels a recent information-theoretic characterization of lumpability of Markov chains. The second condition, which is easier to check, is based on the potential functions of the corresponding Gibbs field. We illustrate our sufficient conditions at the hand of several examples and discuss implications for practical applications of MRFs. As a side result, we give a partial characterization of functions of MRFs that are information preserving.

11:40 ***Two-layer Coded Gradient Aggregation with Straggling Communication Links***

L Kai (Shanghaitech University, China); Youlong Wu (ShanghaiTech University, China)

In many distributed learning setups such as federated learning, client nodes at the edge use individually collected data to compute the local gradients and send them to a central master server, and the master aggregates the received gradients and broadcasts the aggregation to all clients with which the clients can update the global model. As straggling communication links could severely affect the performance of distributed learning system, Prakash et al. proposed to utilize helper nodes and coding strategy to achieve resiliency against straggling client-to-helpers links. In this paper, we propose two coding schemes: repetition coding (RC) and MDS coding both of which enable the clients to update the global model in the presence of only helpers but without the master. Moreover, we characterize the uplink and downlink communication loads, and prove the tightness of uplink communication load. Theoretical tradeoff between uplink and downlink communication loads is established indicating that larger uplink communication load could reduce downlink communication load. Compared to Prakash's schemes which require a master to connect with helpers though noiseless links, our scheme can even reduce the communication load in the absence of master when the number of clients and helpers is relatively large compared to the number of straggling links.

Wednesday, April 14 12:00 - 12:50 (Europe/Rome)

Coding V: Coding at Large

Chair: Vitaly Skachek (University of Tartu, Estonia)

12:00 ***An Efficient Block Error Probability Estimation of Reed-Muller Codes under Permutation Decoding***

Mikhail Kamenev (Huawei Technologies Co., Ltd., Russia)

A method for estimating the block error probability of Reed-Muller codes on a binary erasure channel and a binary symmetric channel under a permutation decoding algorithm is proposed. This method is based on an analysis of the block error probability under a recursive decoding algorithm for a fixed number of input erasures or errors. Simulation results demonstrate that the proposed method allows to effectively predict the block error rate performance of Reed-Muller codes under permutation decoding with a different

number of permutations used.

12:10 *On Decoding of Reed-Muller Codes Using a Local Graph Search*

Mikhail Kamenev (Huawei Technologies Co., Ltd., Russia)

We present a novel iterative decoding algorithm for Reed-Muller (RM) codes, which takes advantage of a graph representation of the code. Vertices of the considered graph correspond to codewords, with two vertices being connected by an edge if and only if the Hamming distance between the corresponding codewords equals the minimum distance of the code. The algorithm starts from a random node and uses a greedy local search to find a node with the best heuristic, e.g. Euclidean distance between the received vector and the corresponding codeword. In addition, the cyclic redundancy check can be used to terminate the search as soon as the correct codeword is found, leading to a decrease in the average computational complexity of the algorithm. Simulation results for an additive white Gaussian noise channel show that the presented decoder achieves the performance of maximum likelihood decoding for RM codes of length 512 and for the second order RM codes of length up to 4096. Moreover, it is demonstrated that the considered decoding approach significantly outperforms state-of-the-art decoding algorithms of RM codes with similar computational complexity for a large number of cases.

12:20 *Using List Decoding to Improve the Finite-Length Performance of Sparse Regression Codes*

Haiwen Cao and Pascal Vontobel (The Chinese University of Hong Kong, Hong Kong)

We consider sparse superposition codes (SPARCs) over complex AWGN channels. Such codes can be efficiently decoded by an approximate message passing (AMP) decoder, whose performance can be predicted via so-called state evolution in the large-system limit. In this paper, we mainly focus on how to use concatenation of SPARCs and cyclic redundancy check (CRC) codes on the encoding side and use list decoding on the decoding side to improve the finite-length performance of the AMP decoder for SPARCs over complex AWGN channels. Simulation results show that such a concatenated coding scheme works much better than SPARCs with the original AMP decoder and results in a steep waterfall-like behavior in the bit-error rate performance curves.

12:30 *On the Finite Length Performance of Sparse Regression Codes with Peak-Power Limitation*

Shansuo Liang and Bo Bai (Huawei Technologies Co., Ltd., Hong Kong); Gong Zhang (Huawei Technologies Co., Ltd., China)

This paper concerns practical issues of sparse regression codes (SRCs) with approximate message passing (AMP) decoding. First, Gaussian signaling of SRC incurs a high peak-to-average-power ratio (PAPR) problem. Second, the finite length performance of SRCs is poor at low-to-medium rates and cannot be improved by spatial coupling or power allocation. We confront the two challenges by introducing clipping to SRC. For the encoder, clipping is applied to the Gaussian codeword of SRC for reducing the high PAPR. For the decoder, generalized approximate message passing (GAMP) is used to handle the nonlinear clipping distortion. Interestingly, we observe that clipping with proper thresholds can improve the performance of SRC and the performance gain is large at low rates. Based on the state evolution analysis of GAMP decoding, we provide an explanation for such observation from the curve matching perspective. In the end, some guidelines are provided for choosing proper clipping thresholds empirically.

12:40 *Optimization of Irregular NB QC-LDPC Block Codes over Small Alphabets*

Irina Bocharova and Boris D. Kudryashov (St. Petersburg University of Information Technologies, Mechanics and Optics, Russia); Evgenii P. Ovsyannikov (State University of Aerospace Instrumentation, Russia); Vitaly Skachek and Tähvend Uustalu (University of Tartu, Estonia)

We propose a novel approach for optimization of nonbinary (NB) quasi-cyclic (QC)-LDPC codes. In this approach, the base parity-check matrices are constructed by the simulated annealing method, and then labeled while maximizing the so-called generalized girth of the NB LDPC code Tanner graph. Random coding bounds on the ML decoding error probability for ensembles of ``almost regular'' NB LDPC codes of finite lengths over extensions of the binary Galois field are derived. These bounds are based on the average bit weight spectra for the ensembles of NB LDPC codes. The observed FER performance of the sum-product BP decoding of ``almost regular'' NB QC-LDPC block codes is presented and compared to the finite-length random coding bounds, as well as to the performance of the optimized binary QC-LDPC block code in the 5G standard. In the waterfall region, the gap between the finite-length bounds on the error probability of the ML decoding and the simulation performance of the BP decoding is about 0.1~--~0.2 dB.

Wednesday, April 14 18:00 - 18:50 (Europe/Rome)

Cryptography, Privacy and Security II

Room 1

Chair: Ilan Shomorony (University of Illinois at Urbana-Champaign, USA)

18:00 *Capacity Theorems for Covert Bosonic Channels*

Michael S Bullock, Christos Gagatsos and Boulat A. Bash (University of Arizona, USA)

We study quantum-secure covert-communication over lossy thermal-noise bosonic channels, the quantum-mechanical model for many practical channels. We derive the expressions for the covert capacity of these channels: \(L_{\text{no-EA}}\), when Alice and Bob share only a classical secret, and \(L_{\text{EA}}\), when they benefit from entanglement assistance. Entanglement assistance alters the fundamental scaling law for covert communication. Instead of \(L_{\text{no-EA}}\sqrt{n}-r_{\text{no-EA}}(n)\), \(r_{\text{no-EA}}(n)=o(\sqrt{n})\), entanglement assistance allows \(L_{\text{EA}}\sqrt{n}\log n-r_{\text{EA}}(n)\), \(r_{\text{EA}}(n)=o(\sqrt{n}\log n)\), covert bits to be transmitted reliably over \(n\) channel uses. However, noise in entanglement storage erases the \(\log n\) from our achievability; work on the matching converse is ongoing.

18:10 *Key Superposition Simultaneously Achieves Security and Privacy in Cache-Aided Linear Function Retrieval*

Qifa Yan (Southwest Jiaotong University, China); Daniela Tuninetti (University of Illinois Chicago, USA)

A coded caching scheme, referred to as key superposition, is proposed in the cache-aided content Secure and demand Private Linear Function Retrieval (SP-LFR) setup, where the following conditions are imposed: (a) each user is interested in retrieving an arbitrary linear combination of the files in the server's library;

(b) the content of the library must be kept secure from a wiretapper who obtains the signal sent by the server; and (c) any subset of users together can not obtain any information about the demands of the remaining users. The scheme uses the superposition of security keys and privacy keys in both the placement and delivery phases to guarantee content security and demand privacy, respectively. The achieved load-memory tradeoff is optimal to within a constant multiplicative gap, except for the small memory regime when there are less file than users. The memory-load tradeoff does not increase compared to the best known schemes that only guarantee content security in all regimes or only demand privacy in some regime.

18:20 *Secret Sharing from Correlated Gaussian Random Variables and Public Communication*

Vidhi Rana, Remi A Chou and Hyuck Kwon (Wichita State University, USA)

We study a secret sharing problem, where a dealer distributes shares of a secret among a set of participants under the constraints that (i) authorized sets of users can recover the secret by pooling their shares, (ii) non-authorized sets of colluding users cannot learn any information about the secret. We assume that the dealer and the participants observe the realizations of correlated Gaussian random variables and that the dealer can communicate with the participants through a one-way, authenticated, rate-limited, and public channel. Our main result is a closed-form characterization of the trade-off between secret rate and public communication rate. Unlike traditional secret sharing protocols, in our setting, no perfectly secure channel is needed between the dealer and the participants, and the size of the shares does not depend exponentially but rather linearly on the number of participants and the size of the secret for arbitrary access structures.

18:30 *Data Disclosure Mechanism Design with Non-zero Leakage*

Amirreza Zamani, Tobias J. Oechtering and Mikael Skoglund (KTH Royal Institute of Technology, Sweden)

We study an information-theoretic privacy problem, where an agent observes useful data \(Y\) and wants to reveal the information to a user. Since the useful data is correlated with sensitive data \(X\), the agent employs a privacy mechanism to produce data \(U\) that can be disclosed. Thus, we study the privacy mechanism design that maximizes the revealed information about \(Y\) while satisfying an \(\ell_1\)-privacy criterion under the Markov chain \(X-Y-U\). When a sufficiently small leakage is allowed, we show that the optimizer of the design problem has a specific structure which allows us to use a local approximation of mutual information. More specifically, we show that the optimizer vectors are perturbations of fixed distributions. By using this approximation the original optimization problem can be reduced to a linear programming problem and an approximate solution for privacy mechanism design can be obtained.

18:40 *Private DNA Sequencing: Hiding Information in Discrete Noise*

Kayvon Mazooji, Roy Dong and Ilan Shomorony (University of Illinois at Urbana-Champaign, USA)

When an individual's DNA is sequenced, sensitive medical information becomes available to the sequencing laboratory. A recently proposed way to hide an individual's genetic information is to mix in DNA samples of other individuals. We assume these samples are known to the individual but unknown to the sequencing laboratory. Thus, these DNA samples act as "noise" to the sequencing laboratory, but still allow the individual to recover their own DNA samples afterward. Motivated by this idea, we study the problem of hiding a binary random variable X (a genetic marker) with the additive noise provided by

mixing DNA samples, using mutual information as a privacy metric. This is equivalent to the problem of finding a worst-case noise distribution for recovering X from the noisy observation among a set of feasible discrete distributions. We characterize upper and lower bounds to the solution of this problem, which are empirically shown to be very close. The lower bound is obtained through a convex relaxation of the original discrete optimization problem, and yields a closed-form expression. The upper bound is computed via a greedy algorithm for selecting the mixing proportions.

Statistics and Information Theory II

Room 2

Chair: Ugo Vaccaro (University of Salerno, USA)

18:00 ***Achievable error exponents for the two-way parallel DMC***

Kenneth Palacio-Baus (University of Illinois at Chicago, USA & University of Cuenca, Ecuador); Natasha Devroye (University of Illinois at Chicago, USA)

We investigate error exponent regions for the parallel two-way DMC in which each terminal sends its own message and provides feedback to the other terminal. Various error exponents are presented in different rate-region regimes based on the relative rates and zero-error capacities of both directions. The schemes employed are extensions of error exponents for one-way DMCs with noiseless, rate-limited and noisy feedback.

18:10 ***Real-Time Variable-to-Fixed Lossless Source Coding of Randomly Arriving Symbols***

Uri Abend (Tel-Aviv University, Israel); Anatoly Khina (Tel Aviv University, Israel)

We address the recently suggested problem of causal lossless coding of randomly arriving source samples. We construct variable-to-fixed coding schemes and show that they outperform the previously considered fixed-to-variable schemes when traffic is high, in terms of both delay and Age of Information by appealing to tools from queueing theory. We supplement our theoretical bounds with numerical simulations.

18:20 ***On the Distribution of the Conditional Mean Estimator in Gaussian Noise***

Alex Dytso (New Jersey Institute of Technology, USA); H. Vincent Poor (Princeton University, USA); Shlomo (Shitz) Shamai (The Technion, Israel)

Consider the conditional mean estimator of the random variable \(X\) from the noisy observation \(Y=X+N\) where \(N\) is zero mean Gaussian with variance \(\sigma^2\) (i.e., \(E[X|Y]\)). This work characterizes the probability distribution of \(E[X|Y]\). As part of the proof, several new identities and results are shown. For example, it is shown that the \(k\)-th derivative of the conditional expectation is proportional to the \((k+1) \)-th conditional cumulant. It is also shown that the compositional inverse of the conditional expectation is well-defined and is characterized in terms of a power series.

18:30 ***Relationship Between Source Resolvability with Normalized f-Divergence and Fixed-Length Coding***

Ryo Nomura (Waseda University, Japan)

This paper deals with the relationship between the source resolvability problem (or resolvability problem

for short) and the fixed-length source coding problem. In the literature, optimum achievable rates in the resolvability problem (optimum resolvability rate) with respect to the variational distance as well as the Kullback-Leibler (KL) divergence, have already been analyzed. The relationship between the optimum resolvability rate and the optimum rate of the fixed-length source coding has also been clarified in each cases. In particular, it has been reported that the optimum source resolvability rate with respect to the normalized KL divergence has a close relationship with the optimum fixed-length source coding rate with the correct decoding exponent. Recently, the optimum resolvability rate with respect to a class of f-divergences has been analyzed. This result can be considered as a generalization of the optimum resolvability rate with respect to the unnormalized KL divergence. However, unnormalized f-divergences has not been considered yet in the resolvability problem. Hence, in this paper, we consider the resolvability problem with respect to a class of unnormalized f-divergences. In particular, we derive the relationship between the optimum resolvability rate with a class of normalized f-divergences and the optimum rate of the fixed-length source coding.

18:40 ***On-The-Fly Stochastic Codebook Re-generation for Sources with Memory***

Ahmed Elshafiy and Mahmoud Namazi (University of California, Santa Barbara, USA); Ram Zamir (Tel Aviv University, Israel); Kenneth Rose (University of California, Santa Barbara, USA)

This paper proposes a generalized stochastic mechanism for codebook generation in lossy coding settings for sources with memory. Earlier work has shown that the rate-distortion bound can be asymptotically achieved for discrete memoryless sources by a ``natural type selection'' (NTS) algorithm. In iteration n, the distribution that is most likely to produce the types of a sequence of K codewords of finite length l that ``d-match'' a respective sequence of K source words of length l, (i.e., which satisfy the distortion constraint), is used to regenerate the codebook for iteration n+1. The resulting sequence of codebook generating distributions converges to the optimal distribution Q* that achieves the rate-distortion bound for the memoryless source, asymptotically in l, K, and n. This work generalizes the NTS algorithm to account for sources with memory. The algorithm encodes ml-length source words consisting of l vectors (or super-symbols) of length m. We show that for finite m and l, the sequence of codebook reproduction distributions Q_{0, m, l}, Q_{1, m, l}, ... (each computed after observing a sequence of K d-match events) converges to the optimal achievable distribution Q*_{m, l} (within a set of achievable distributions determined by m and l), asymptotically in K and n. It is further shown that Q*_{m, l} converges to the optimal reproduction distribution Q* that achieves the rate-distortion bound for sources with memory, asymptotically in m and l.

Wednesday, April 14 19:00 - 20:00 (Europe/Rome)

Plenary - The generalization error of overparametrized models: Insights from exact asymptotics weconf

Andrea Montanari

Chair: Alon Orlitsky (University of California, San Diego, USA)

Deep learning models are often so complex that they achieve vanishing classification error on the training set. Despite their huge complexity, the same architectures achieve small generalization error. This phenomenon

has been rationalized in terms of a so-called double descent curve. As the model complexity increases, the generalization error follows the usual U-shaped curve at the beginning, first decreasing and then peaking around the interpolation threshold (when the model achieves vanishing training error). However, it descends again as model complexity exceeds this threshold.

I will focus on the case of a fully-connected two-layers neural network, and consider its linearization around a random initial condition. I will show that many intersting phenomena can be demonstrated and mathematically understood in this simple setting. I will then describe a few open problems and directions for future research.

Wednesday, April 14 20:10 - 21:50 (Europe/Rome)

Themed Session - Blockchain weconf

Chair: David Tse (Stanford University, USA)

20:10 *Byzantine Consensus Through the Lens of Information Theory*

David Tse (Stanford University, USA)

The study of the fundamental limits of communication in the presence of errors was initiated by Shannon in 1948. The study of the fundamental limits of distributed consensus in the presence of faults was initiated by Lamport, Pease, and Shostak in the early 1980s. Both fields study the problem of designing optimal reliable systems but there has been surprisingly limited cross-fertilization of ideas in the past 40 years. In this talk, we give an example of the utility of such cross-fertilization by using an analogy from network information theory (degraded message set) to formulate and solve a central problem in blockchains: the availability-finality dilemma. This is joint work with Joachim Neu and Ertem Nusret Tas.

20:30 *Scaling On-Chain Asset Exchanges via Arrow-Debreu Exchange Markets*

Ashish Goel (Stanford University, USA)

Digital currencies present an opportunity to equalize access to financial systems and reduce the cost and latency of trading assets or transferring money. An ideal exchange for these assets would itself be a blockchain; such a system could be decentralized and transparent, among other desirable properties. The primary difficulty in implementing a multi-asset exchange on a blockchain, as opposed to in a traditional centralized setting, is that blockchain nodes must operate deterministically and replicably. In this work, we introduce a new approach to matching requests to trade assets such that all transactions is a block commute with each other. Our approach is to view an asset exchange as an instance of the Arrow-Debreu exchange market, where each asset is valued in an abstract "phantom" currency, removing the need for a reserve currency. Transaction execution can now be parallelized over an arbitrary number of off-the-shelf CPU cores with minimal synchronization overhead. The central algorithmic difficulty of this approach is in computing equilibrium prices. For the case where the exchange only supports sell orders, we show how existing convex programming techniques can be adapted to compute equilibrium prices efficiently, in both theory and empirically. We also show how even approximate (or "noisy") prices can be used to find a market-clearing solution with many desirable properties. Finally, we mention that supporting both sell and buy orders in these markets makes the price computation a PPAD-hard problem. We believe that our work points to a viable implementation path for sophisticated and robust multi-asset exchanges for digital currencies. This represents joint work with Geoffrey Ramseyer and David Mazi`eres.

20:50 *Stochastic Analysis of Blockchain Protocols*

Aggelos Kiayias (University of Edinburgh, United Kingdom (Great Britain))

Analyzing blockchain protocols in the Byzantine setting, i.e. in the presence of an adversary that follows an arbitrary strategy unbound by rationality, is an important direction towards understanding their security. We overview a number of recent results that abstract blockchain protocol execution as a stochastic process and discuss analytical tools and resulting bounds capturing essential properties - consistency and liveness - both for proof of work and proof of stake protocols.

21:10 *On blockchain models, analysis, and counting processes*

Dongning Guo (Northwestern University, USA)

The last few years have witnessed a number of breakthroughs in the security analysis of blockchain systems employing some longest-chain protocols. The underlying models are often complicated and their analyses are often difficult to parse. In this talk I introduce a minimalistic continuous-time model for block mining times and blockchains. This model is sufficient for establishing many consistency and liveness properties of longest-chain protocols. We discuss a concrete, practical security-latency trade-off obtained using the simple model. We also relate our model and results to prior arts in the literature.

21:30 *Fair-Ordered Protocols for Permissionless Blockchains*

Sreeram Kannan (University of Washington Seattle, USA)

Over the past five years, a significant line of research has investigated the blockchain con-sensus problem in the general permissionless setting, where protocol nodes can leave and join dynamically. The work of Garay et al. (Eurocrypt 2015) and Pass et al. (Eurocrypt 2017) showed the security properties of consistency and liveness for Nakamoto's seminal proof-of-work protocol. However, consistency and liveness do not provide any guarantees on the relationship between the order in which transactions arrive into the network and the finalized order in theledger, making protocols prone to transaction order-manipulation attacks. As a solution, a recent paper by Kelkar et al. (Crypto 2020) introduced a third useful property for consensus protocols: transaction-order-fairness. Their model was limited to the classical (permissioned)setting, where the set of protocol nodes is fixed a priori, and does not fit well for permissionless environments where order-manipulation attacks have been most prominent. In this work, we initiate the investigation of order-fairness in the permissionless setting and provide two protocols that realize it. Our protocols work in a synchronous network and use an underlying longest-chain blockchain. As an added contribution, we show that any fair ordering protocol achieves a powerful zero-block confirmation property, through which honest transactions can be securely confirmed even before they are included in any block. This is joint work with Mahimna Kelkar and Soubhik Deb.

Thursday, April 15

Thursday, April 15 9:00 - 10:00 (Europe/Rome)

Plenary - Information-directed Exploration in Bandits and Reinforcement Learning

Andreas Krause

Chair: Nicolò Cesa-Bianchi (Università degli Studi di Milano, Italy)

The exploration-exploitation dilemma is a central challenge when making decisions under uncertainty. Most common approaches explore by favouring actions with uncertain outcomes. However, aleatoric uncertainty in the outcomes is different from epistemic uncertainty in the estimation task, thus the resulting observations may not necessarily be informative. In this talk, I will present approaches towards efficient information-directed exploration in stochastic multi-armed bandits, Bayesian optimization, reinforcement learning and a rich family of sequential decision problems called partial monitoring. These approaches use information measures for guiding exploration, and their submodularity allows to establish sublinear regret even in non-parametric settings. I will present the theoretical background, as well as empirical demonstrations on deep reinforcement learning tasks.

Thursday, April 15 10:00 - 10:10 (Europe/Rome)

Announcements

Chair: Brian Michael Kurkoski (Japan Advanced Institute of Science and Technology (JAIST), Japan)

Thursday, April 15 10:20 - 11:10 (Europe/Rome)

Machine Learning II

Room 1

Chair: Nicolò Cesa-Bianchi (Università degli Studi di Milano, Italy)

10:20 *On Random Subset Generalization Error Bounds and the Stochastic Gradient Langevin Dynamics Algorithm*

Borja Rodríguez-Gálvez, Germán Bassi, Ragnar Thobaben and Mikael Skoglund (KTH Royal Institute of Technology, Sweden)

In this work, we unify several expected generalization error bounds based on random subsets using the framework developed by Hellström and Durisi [1]. First, we recover the bounds based on the individual sample mutual information from Bu et al. [2] and on a random subset of the dataset from Negrea et al. [3].

Then, we introduce their new, analogous bounds in the randomized subsample setting from Steinke and Zakynthinou [4], and we identify some limitations of the framework. Finally, we extend the bounds from Haghifam et al. [5] for Langevin dynamics to stochastic gradient Langevin dynamics and we refine them for loss functions with potentially large gradient norms.

10:30 *Query Complexity of k-NN based Mode Estimation*

Anirudh Singhal, Subham Pirojiwala and Nikhil Karamchandani (Indian Institute of Technology Bombay, India)

Motivated by the mode estimation problem of an unknown multivariate probability density function, we study the problem of identifying the point with the minimum k-th nearest neighbor distance for a given dataset of n points. We study the case where the pairwise distances are apriori unknown, but we have access to an oracle which we can query to get noisy information about the distance between any pair of points. For two natural oracle models, we design a sequential learning algorithm, based on the idea of confidence intervals, which adaptively decides which queries to send to the oracle and is able to correctly solve the problem with high probability. We derive instance-dependent upper bounds on the query complexity of our proposed scheme and also demonstrate significant improvement over the performance of other baselines via extensive numerical evaluations.

10:40 *Approximating Probability Distributions by ReLU Networks*

Manuj Mukherjee (Bar Ilan University, Israel); Aslan Tchamkerten (Telecom ParisTech, France); Mansoor Yousefi (Télécom ParisTech, France)

How many neurons are needed to approximate a target probability distribution using a neural network with a given input distribution and approximation error? This paper examines this question for the case when the input distribution is uniform, and the target distribution belongs to the class of histogram distributions. We obtain a new upper bound on the number of required neurons, which is strictly better than previously existing upper bounds. The key ingredient in this improvement is an efficient construction of the neural nets representing piecewise linear functions. We also obtain a lower bound on the minimum number of neurons needed to approximate the histogram distributions.

10:50 *On Compressed Sensing Matrices Breaking the Square-Root Bottleneck*

Shohei Satake (Kumamoto University, Japan); Yujie Gu (Kyushu University, Japan)

Compressed sensing is a celebrated framework in signal processing and has many practical applications. One of the challenging problems in compressed sensing is to construct deterministic matrices having the restricted isometry property (RIP). So far, there are only a few publications providing deterministic RIP matrices beating the square-root bottleneck on the sparsity level. In this paper, we investigate RIP of certain matrices defined by higher power residues modulo primes. Moreover, we prove that the widely-believed generalized Paley graph conjecture implies that these matrices have RIP breaking the square-root bottleneck. Also the compression ratio realized by these RIP matrices is significantly larger than 2.

11:00 *Social Learning is Almost as Good as Centralized Detection with Slight Global Knowledge*

Yu-Chieh Huang and I-Hsiang Wang (National Taiwan University, Taiwan)

Fundamental limits on the error probability of the social learning rule proposed by Lalitha et al. [1] and its variants for decentralized detection over a directed graph is investigated. We show that while social learning algorithms enjoy the benefits that each node in the graph weights the messages received

from its neighbors locally to form its private belief and only requires knowledge of the data generating distributions of its own observation, it suffers a gap in the achievable error exponent compared to the centralized case. We propose a generalization of the social learning rule achieving the error exponent in centralized detection with the aid of slight global knowledge. A further analysis reveals that the price of decentralization is at most a constant term in the higher-order asymptotics. To obtain the slight global knowledge needed at each node for achieving the centralized error exponent, we develop a decentralized estimation method for each node to come up with a local estimate of that piece of global knowledge.

Cryptography, Privacy and Security III

Room 2

Chair: Parastoo Sadeghi (University of New South Wales, Australia)

10:20 *Biometric Identification Systems With Noisy Enrollment for Gaussian Source*

Vamoua Yachongka (The University of Electro-Communications, Japan); Hideki Yagi and Yasutada Oohama (University of Electro-Communications, Japan)

In the present paper, we investigate the fundamental trade-off of identification, secrecy, storage, and privacy-leakage rates in biometric identification systems for hidden or remote Gaussian sources. We introduce a technique for deriving the capacity region of these rates by converting the system to one where the data flow is in one-way direction. Also, we provide numerical calculations of three different examples for the generated-secret model. The numerical results imply that it seems hard to achieve both high secrecy and small privacy-leakage rates simultaneously. In addition, as special cases, the characterization coincides with several known results in previous studies.

10:30 *Non-Stochastic Private Function Evaluation*

Farhad Farokhi and Girish N. Nair (University of Melbourne, Australia)

We consider private function evaluation to provide query responses based on private data of multiple untrusted entities in such a way that each cannot learn something substantially new about the data of others. First, we introduce perfect non-stochastic privacy in a two-party scenario. Perfect privacy amounts to conditional unrelatedness of the query response and the private uncertain variable of other individuals conditioned on the uncertain variable of a given entity. We show that perfect privacy can be achieved for queries that are functions of the common uncertain variable, a generalization of the common random variable. We compute the closest approximation of the queries that do not take this form. To provide a trade-off between privacy and utility, we relax the notion of perfect privacy. We define almost perfect privacy and show that this new definition equates to using conditional disassociation instead of conditional unrelatedness in the definition of perfect privacy. Then, we generalize the definitions to multi-party function evaluation (more than two data entities). We prove that uniform quantization of query responses, where the quantization resolution is a function of privacy budget and sensitivity of the query (cf., differential privacy), achieves function evaluation privacy.

10:40 *Privacy-Utility Tradeoff with Nonspecific Tasks: Robust Privatization and Minimum Leakage*

Ta-Yuan Liu and I-Hsiang Wang (National Taiwan University, Taiwan)

Privacy-preserving data release mechanisms aiming to minimize the privacy leakage under utility

constraints of nonspecific tasks are studied through the lens of information theory. While the private feature to be protected is typically determined and known by the users who release their data, the specific task where the release data is utilized is usually unknown. To address the lack of information of the specific task, utility constraints laid on a set of multiple possible tasks are considered. The mechanism protects the privacy of a given feature of the to-be-released data while satisfying utility constraints of all possible tasks in the set. First, the single-letter characterization of the privacy-utility tradeoff region is derived. Characterization of the minimum privacy under log-loss utility constraints turns out to be a non-convex optimization problem involving mutual information in the objective function and the constraints. Second, focusing on the case where the raw data consists of multiple independent components, we show that the above optimization problem can be decomposed into multiple parallel privacy funnel (PF) problems [1] with different weightings. We explicitly derive the optimal solution to each PF problem when the private feature is a deterministic function of a data component. The solution is characterized by the leakage-free threshold, and the minimum leakage is zero while the utility constraint is below the threshold. Once the utility requirement is above the threshold, the privacy leakage increases linearly. Finally, we show that the optimal weighting of each privacy funnel problem can be found by solving a linear program (LP). Numerical results are shown to illustrate the robustness of our approach.

10:50 *Measuring Information Leakage in Non-stochastic Brute-Force Guessing*

Farhad Farokhi (University of Melbourne, Australia); Ni Ding (The University of Melbourne, Australia)

We propose an operational measure of information leakage in a non-stochastic setting to formalize privacy against a brute-force guessing adversary. We use uncertain variables, non-probabilistic counterparts of random variables, to construct a guessing framework in which an adversary is interested in determining private information based on uncertain reports. We consider brute-force trial-and-error guessing in which an adversary can potentially check all the possibilities of the private information that are compatible with the available outputs to find the actual private realization. The ratio of the worst-case number of guesses for the adversary in the presence of the output and in the absence of it captures the reduction in the adversary's guessing complexity and is thus used as a measure of private information leakage. We investigate the relationship between the newly-developed measure of information leakage with maximin information and stochastic maximal leakage that are shown to arise in one-shot guessing.

11:00 *On Properties and Optimization of Information-theoretic Privacy Watchdog*

Parastoo Sadeghi (University of New South Wales, Australia); Ni Ding (The University of Melbourne, Australia); Thierry Rakotoarivelo (Data61, CSIRO, Australia)

We study the problem of privacy preservation in data sharing, where \(S\) is a sensitive variable to be protected and \(X\) is a non-sensitive useful variable correlated with \(S\). Variable \(X\) is randomized into variable \(Y\), which will be shared or released according to \(p_{Y|X}(y|x)\). We measure privacy leakage by \emph{information privacy} (also known as \emph{log-lift} in the literature), which guarantees mutual information privacy and differential privacy (DP). Let \(\mathcal{X}_{\epsilon}^c \subseteq \mathcal{X}\) contain elements in the alphabet of \(X\) for which the absolute value of log-lift (abs-log-lift for short) is greater than a desired threshold \(\epsilon\). When elements \(x\in \mathcal{X}_{\epsilon}^c\) are randomized into \(y\in \mathcal{Y}\), we derive the best upper bound on the abs-log-lift across the resultant pairs \((s,y)\). We then prove that this bound is achievable via an \emph{\(X\)-invariant} randomization \(p(y|x) = R(y)\) for \(x,y\in\mathcal{X}_{\epsilon}^c\). However, the utility measured by the mutual information \(I(X;Y)\) is severely damaged in imposing a strict upper bound \(\epsilon\) on the abs-log-lift. To remedy this and inspired by the probabilistic \((\epsilon, \delta)\)-DP, we propose a relaxed

$((\epsilon, \delta))$-log-lift framework. To achieve this relaxation, we introduce a greedy algorithm which exempts some elements in $\mathcal{X}_{\epsilon}^c$ from randomization, as long as their abs-log-lift is bounded by ϵ with probability $1-\delta$. Numerical results demonstrate efficacy of this algorithm in achieving a better privacy-utility tradeoff.

Thursday, April 15 11:20 - 12:10 (Europe/Rome)

Coding VI: Coding theory and practice

Chair: Ferdinando Cicalese (Universita' di Verona, Italy)

11:20 *Dimension of a Subset of Residue Classes*

Vladislav Shchukin (Huawei Technologies Co. Ltd., Moscow, Russia)

This paper introduces a problem in additive number theory which is motivated by optimization of hardware implementation of QC-LDPC codes. For a fixed subset S, $S \subset \mathbb{Z}_q$, of residue classes modulo q the object of interest is a basis set G, $G \subset \mathbb{Z}_q$, of minimal size, such that every element of S is representable as a sum of several elements from G. For a fixed number k, $k \le \lceil \log_2 q \rceil$, the object of interest is a function $\zeta(q, k)$ defined as a maximum number such that, for every set of cardinality $< \zeta(q, k)$, there exists a basis set of cardinality $< k$.

11:30 *Transmission of a Bit over a Discrete Poisson Channel with Memory*

Niloufar Ahmadypour (Sharif University of Technology, Iran); Amin Gohari (Tehran Institute for Advanced Studies, Iran)

A coding scheme for transmission of a bit maps a given bit to a sequence of channel inputs (called the codeword associated to the transmitted bit). In this paper, we study the problem of designing the best code for a discrete Poisson channel with memory (under peak-power and total-power constraints). The outputs of a discrete Poisson channel with memory are Poisson distributed random variables with a mean comprising a fixed additive noise and a linear combination of past input symbols. Assuming a maximum-likelihood (ML) decoder, we find the best codebook design by minimizing the error probability of the decoder over all codebooks. For the case of having only a total-power constraint, the optimal code structure is obtained provided that the blocklength is greater than the memory length of the channel. For the case of having only a peak-power constraint, the optimal code is derived for arbitrary memory and blocklength in the high-power regime. For the case of having both the peak-power and total-power constraints, the optimal code is derived for memoryless Poisson channels when both the total-power and the peak-power bounds are large.

11:40 *Universal interactive Gaussian quantization with side information*

Shubham K Jha (Indian Institute of Science Bangalore, India); Himanshu Tyagi (Indian Institute of Science, India)

We consider universal quantization with side information for Gaussian observations, where the side information is a noisy version of the sender's observation with an unknown noise variance. We propose a universally rate optimal and practical quantization scheme for all values of unknown noise variance. Our scheme is interactive, uses Polar lattices from prior work, and proceeds by checking in each round if a

reliable estimate has been formed. In particular, our scheme is based on a structural decomposition of the underlying auxiliaries so that even when recovery fails in a round, the parties agree on a common "reference point" that is closer than the previous one.

11:50 ***Improved Memory-Rate Trade-off for Caching with Demand Privacy***

Chinmay Shekhar Gurjarpadhye (Indian Institute of Technology Bombay, India); Jithin Ravi (Universidad Carlos III de Madrid, Spain); Bikash K Dey and Nikhil Karamchandani (Indian Institute of Technology Bombay, India)

We consider the demand-private coded caching problem in a noiseless broadcast network. It is known from past works that a demand-private scheme for \(N\) files and \(K\) users can be obtained from a non-private scheme for \(N\) files and \(NK\) users. We first propose a scheme that improves on this idea by removing some redundant transmissions. The memory-rate trade-off achieved using this scheme is shown to be within a multiplicative factor of 3 from the optimal for all the memory regimes when \(K < N\). We further show that a demand-private scheme for \(N\) files and \(K\) users can be obtained from a particular known non-private scheme for \(N\) files and \(NK-K+1\) users. Finally, we give the exact memory-rate trade-off for demand-private coded caching problems with \(N>K=2\).

12:00 ***Over-The-Air Computation in Correlated Channels***

Matthias Frey (Technische Universität Berlin, Germany); Igor Bjelakovic (Fraunhofer Heinrich Hertz Institute, Germany); Slawomir Stanczak (Technische Universität Berlin & Fraunhofer Heinrich Hertz Institute, Germany)

This paper addresses the problem of Over-The-Air (OTA) computation in wireless networks which has the potential to realize huge efficiency gains for instance in training of distributed ML models. We provide non-asymptotic, theoretical guarantees for OTA computation in fast-fading wireless channels where the fading and noise may be correlated. The distributions of fading and noise are not restricted to Gaussian distributions, but instead are assumed to follow a distribution in the more general sub-gaussian class. Furthermore, our result does not make any assumptions on the distribution of the sources and therefore, it can, e.g., be applied to arbitrarily correlated sources. We illustrate our analysis with numerical evaluations for OTA computation of two example functions in large wireless networks: the arithmetic mean and the Euclidean norm.

Thursday, April 15 16:00 - 17:40 (Europe/Rome)

Themed Session - Learning Theory weconf

Chair: Gergely Neu (Universitat Pompeu Fabra, Spain)

16:00 ***Suboptimality of Constrained Linear Least Squares and Improvements via Improper Learners***

Nikita Zhivotovskiy (ETH Zurich, Switzerland)

We study the problem of predicting as well as the best linear predictor in a bounded Euclidean ball with respect to the squared loss. When only boundedness of the data generating distribution is assumed, we establish that the least squares estimator constrained to a bounded Euclidean ball does not attain

the classical O(d/n) excess risk rate, where d is the dimension of the covariates and n is the number of samples. In particular, we construct a bounded distribution such that the constrained least squares estimator incurs an excess risk of order \Omega(d^{3/2}/n) hence refuting a recent conjecture of Ohad Shamir [JMLR 2015]. In contrast, we observe that non-linear predictors can achieve the optimal rate O(d/n) with no assumptions on the distribution of the covariates. Our non-linear predictor is a modification of the renowned Vovk-Azoury-Warmuth forecaster. We discuss additional distributional assumptions sufficient to guarantee an O(d/n) excess risk rate for the least squares estimator. Among them are certain moment equivalence assumptions often used in the robust statistics literature. While such assumptions are central in the analysis of unbounded and heavy-tailed settings, our work indicates that in some cases, they also rule out unfavorable bounded distributions. The talk is based on a joint work with Tomas Vaškevičius.

16:20 *A Non-Asymptotic Analysis for Stein Variational Gradient Descent*

Anna Korba (ENSAE/CREST, France)

We study the Stein Variational Gradient Descent (SVGD) algorithm, which optimises a set of particles to approximate a target probability distribution \(\pi\propto e^{-V}\) on \(\mathbb{R}^d\). In the population limit, SVGD performs gradient descent in the space of probability distributions on the KL divergence with respect to \(\pi\), where the gradient is smoothed through a kernel integral operator. In this paper, we provide a novel finite time analysis for the SVGD algorithm. We provide a descent lemma establishing that the algorithm decreases the objective at each iteration, and rates of convergence. We also provide a convergence result of the finite particle system corresponding to the practical implementation of SVGD to its population version. Joint work with Adil Salim (Visual Computing Center, KAUST), Michael Arbel (Gatsby Unit, University College London) Giulia Luise (Computer Science Department, University College London), Arthur Gretton (Gatsby Unit, University College London). To appear at Neurips 2020.

16:40 *An introduction to Gated Linear Networks*

Joel Veness (DeepMind, United Kingdom (Great Britain))

Gated Linear Networks (GLNs) are a recently introduced family of deep neural network architectures. Their distinguishing feature is the use of hard gating and local learning to add representational power, as opposed to the more widely used combination of non-linear transfer functions and backpropagation. The simultaneous interaction of local learning and gating gives rise to very different learning dynamics whose range of applications are only just starting to be explored. They can be viewed as a machine learning specific generalization of the PAQ family of context mixing networks, which are a key component of state of the art online language/data compression models. In particular a new form of gating, half-space gating, was proposed to deal with real-valued vector feature spaces. This formulation enjoys universality guarantees, empirical capacity that compares favourably with Deep ReLu networks and has recently been shown to give state of the art results in contextual bandits and regression applications. The key observation was to interpret context mixing networks as data dependant linear networks, explain the local learning procedure in terms of online convex programming to model a feature dependent target density, to abstract and understand the role of various gating functions and their effect on representation power and finally show that certain forms of gating allow for universal learning. Here will we present a unified view of these architectures, discuss their strengths and current limitations, and highlight promising directions for future investigation. Joint work with Avishkar Bhoopchand, David Budden, Agnieszka Grabska-Barwinska, Marcus Hutter, Tor Lattimore, Adam Marblestone, Christopher Mattern, Eren Sezener, Peter Toth, Jianan Wang, Simon Schmitt, Greg Wayne.

17:00 ***Active Regret Minimization with Expert Advice***

Jacob Abernethy and Bhuvesh Kumar (Georgia Institute of Technology, USA); Venkatesh Saligrama (Boston University, USA)

We consider the classical problem of prediction with expert advice, but with an active learning twist. In this new setting, the algorithm may reorder the sequence of examples on which a prediction is made, it aims to minimize regret as usual, but it can observe only a small fraction of the true labels along the way. We consider a variant of the Hedge algorithm for this setting, and we show that under a very particular combinatorial constraint on the matrix of expert predictions we can obtain a very strong regret guarantee while querying very few labels. This constraint, which we refer to as theta-compactness, can be viewed as a non-stochastic variant of the disagreement coefficient, another popular parameter used to reason about the sample complexity of active learning in the IID setting. We also give a polynomial time algorithm to calculate the theta-compactness of a matrix up to an approximation factor of 3.

17:20 ***Smoothed Analysis of Online and Differentially Private Learning***

Nika Haghtalab (University of California, Berkeley, USA)

Practical and pervasive needs for robustness and privacy in algorithms have inspired the design of online adversarial and differentially private learning algorithms. The primary quantity that characterizes learnability in these settings is the Littlestone dimension of the class of hypotheses [Ben-David et al., 2009, Alon et al., 2019]. This characterization is often interpreted as an impossibility result because classes such as linear thresholds and neural networks have infinite Littlestone dimension. In this paper, we apply the framework of smoothed analysis [Spielman and Teng, 2004], in which adversarially chosen inputs are perturbed slightly by nature. We show that fundamentally stronger regret and error guarantees are possible with smoothed adversaries than with worst-case adversaries. In particular, we obtain regret and privacy error bounds that depend only on the VC dimension and the bracketing number of a hypothesis class, and on the magnitudes of the perturbations. Joint work with Abhishek Shetty and Tim Roughgarden.

Thursday, April 15 18:00 - 18:50 (Europe/Rome)

Wireless II

Room 1

Chair: Richard Wesel (University of California, Los Angeles, USA)

18:00 ***Structured Index Coding Problems and Multi-access Coded Caching***

Srinivas Reddy Kota (Indian Institute of Technology, Bombay, India); Nikhil Karamchandani (Indian Institute of Technology Bombay, India)

Index coding and coded caching are two active research topics in information theory with strong ties to each other. Motivated by the multi-access coded caching problem, we study a new class of structured index coding problems (ICPs) which are formed by the union of several symmetric ICPs. We derive upper and lower bounds on the optimal server transmission rate for this class of ICPs and demonstrate that they differ by at most a factor of two. Finally, we apply these results to the multi-access coded caching problem to derive better bounds than the state of the art.

18:10 ***Feedback capacity of ISI MIMO channel with colored Noise***

Abhishek Rawat (University of Minnesota Twin Cities, USA); Nicola Elia (Iowa State University, USA)

We consider the problem of power constrained noiseless feedback capacity of single user MIMO ISI channel with colored noise. We generalize the approach in [1] by allowing non-minimum phase zeros and delays in the channel. When we consider the noise process which is non-minimum phase, we run into internal stability issues when we use the state-space approach while it is implicit in the frequency-domain characterization. Considering this issue, we focus on the state-space characterization and formulate the problem as a convex optimization problem. Through a numerical example, we demonstrate that delays in the channel hampers the gain in feedback capacity and consequently for large delays it approaches feedforward capacity.

18:20 ***Deterministic Identification Over Fading Channels***

Mohammad Javad Salariseddigh and Uzi Pereg (Technical University of Munich, Germany); Holger Boche (Technical University Munich, Germany); Christian Deppe (Technical University of Munich, Germany)

Deterministic identification (DI) is addressed for Gaussian channels with fast and slow fading, where channel side information is available at the decoder. In particular, it is established that the number of messages scales as \(2^{n\log(n)R}\), where \(n\) is the block length and \(R\) is the coding rate. Lower and upper bounds on the DI capacity are developed in this scale for fast and slow fading. Consequently, the DI capacity is infinite in the exponential scale and zero in the double-exponential scale, regardless of the channel noise.

18:30 ***Operating Half-Duplex Diamond Networks with Two Interfering Relays***

Sarthak Jain (University of Minnesota, Twin Cities, USA); Martina Cardone and Soheil Mohajer (University of Minnesota, USA)

This paper considers a diamond network with two interfering relays, where the source communicates with the destination via a layer of 2 half-duplex relays that can communicate with each other. The main focus is on characterizing the 3 relay receive/transmit configuration states (out of the 4 possible ones) that suffice to achieve the approximate capacity of the network. Towards this end, the binary linear deterministic approximation of the Gaussian noise channel is analyzed, and explicit scheduling and relaying schemes are presented. These schemes quantify the amount of information that each relay is responsible for sending to the destination, as well as the fraction of time each relay should receive and transmit.

18:40 ***Finite-Support Capacity-Approaching Distributions for AWGN Channels***

Derek Xiao, Linfang Wang, Dan Song and Richard Wesel (University of California, Los Angeles, USA)

Previously, dynamic-assignment Blahut-Arimoto (DAB) was used to find capacity-achieving probability mass functions (PMFs) for binomial channels and molecular channels. As it turns out, DAB can efficiently identify capacity-achieving PMFs for a wide variety of channels. This paper applies DAB to power-constrained (PC) additive white Gaussian Noise (AWGN) Channels and amplitude-constrained (AC) AWGN Channels. This paper modifies DAB to include a power constraint and finds low-cardinality PMFs that approach capacity on PC-AWGN Channels. While a continuous Gaussian PDF is well-known to be capacity-

achieving on the PC-AWGN channel, DAB identifies low-cardinality PMFs within 0.01 bits of the mutual information provided by a Gaussian PDF. Recall the results of Ozarow and Wyner requiring a constellation cardinality of 2^(C+1) to approach capacity C to within the asymptotic shaping loss of 1.53 dB at high SNR. PMF's found by DAB approach capacity with essentially no shaping loss with cardinality less than 2^(C+1.2). DAB identifies PMFs with better mutual information vs. SNR performance than the analytical approaches to finite-support constellations examined by Wu and Verdu. This paper also uses DAB to find capacity-achieving PMFs with small cardinality support sets for AC-AWGN Channels. The resulting evolution of capacity-achieving PMFs as a function of SNR is consistent with the approximate cardinality transition points of Sharma and Shamai.

Coding VII: Coding at Large

Room 2

Chair: Paul H. Siegel (University of California, San Diego, USA)

18:00 ***Efficient Storage Schemes for Desired Service Rate Regions***

Fatemeh Kazemi (Texas A&M University, USA); Sascha Kurz (University of Bayreuth, Germany); Emina Soljanin (Rutgers University, USA); Alex Sprintson (Texas A&M University, USA)

A major concern in cloud/edge storage systems is serving a large number of users simultaneously. The service rate region is introduced recently as an important performance metric for coded distributed systems, which is defined as the set of all data access requests that can be simultaneously handled by the system. This paper studies the problem of designing a coded distributed storage system storing k files where a desired service rate region $\mathcal{R}$ of the system is given and the goal is 1) to determine the minimum number of storage nodes $n(\mathcal{R})$ for serving all demand vectors inside the set $\mathcal{R}$ and 2) to design the most storage-efficient redundancy scheme with the service rate region covering the set $\mathcal{R}$. Towards this goal, we propose three general lower bounds for $n(\mathcal{R})$. Also, for $k=2$, we characterize $n(\mathcal{R})$, i.e., we show that the proposed lower bounds are tight, via designing a novel storage-efficient redundancy scheme with $n(\mathcal{R})$ storage nodes and the service rate region covering $\mathcal{R}$.

18:10 ***Adaptive Doping of Spatially Coupled LDPC Codes***

Min Zhu (Xidian University, China); David G. M. Mitchell (New Mexico State University, USA); Michael Lentmaier (Lund University, Sweden); Daniel J. Costello, Jr. (University of Notre Dame, USA)

In this paper, we study the problem of error propagation in sliding window decoding (SWD) of spatially coupled LDPC (SC-LDPC) codes. A general decoder model that accounts for error propagation is proposed and analyzed, and the decoded block error rate (BLER) is calculated using the model. In order to improve the BLER performance under decoder error propagation conditions, adaptive variable node (VN) doping is proposed, assuming a noiseless binary feedback channel is available. Example calculations using the proposed model, as well as numerical simulation results, are used to show that adaptive VN doping improves the BLER performance compared to the periodic VN doping and to the undoped case.

18:20 ***Deep Ensemble of Weighted Viterbi Decoders for Tail-Biting Convolutional***

Codes

Tomer Raviv, Asaf Schwartz and Yair Be'ery (Tel-Aviv University, Israel)

Tail-biting convolutional codes extend the classical zero-termination convolutional codes: Both encoding schemes force the equality of start and end states, but under the tail-biting each state is a valid termination. This paper proposes a machine-learning approach to improve the state-of-the-art decoding of tail-biting codes, focusing on the widely employed short length regime as in the LTE standard. This standard also includes a CRC code. First, we parameterize the circular Viterbi algorithm, a baseline decoder that exploits the circular nature of the underlying trellis. An ensemble combines multiple such weighted decoders, each decoder specializes in decoding words from a specific region of the channel words' distribution. A region corresponds to a subset of termination states; the ensemble covers the entire states space. A non-learnable gating satisfies two goals: it filters easily decoded words and mitigates the overhead of executing multiple weighted decoders. The CRC criterion is employed to choose only a subset of experts for decoding purpose. Our method achieves FER improvement of up to 0.75dB over the CVA in the waterfall region for multiple code lengths, adding negligible computational complexity compared to the circular Viterbi algorithm in high SNRs.

18:30 *On Skew Convolutional and Trellis Codes*

Vladimir Sidorenko (Technical University of Munich, Germany); Wenhui Li (Skolkovo Institute of Science and Technology (Skoltech), Russia); Onur Günlü (University of Siegen, Germany); Gerhard Kramer (Technical University of Munich, Germany)

Two new classes of skew codes over a finite field F are proposed, called skew convolutional codes and skew trellis codes. These two classes are defined by, respectively, left or right sub-modules over the skew fields of fractions of skew polynomials over F. The skew convolutional codes can be represented as periodic time-varying ordinary convolutional codes. The skew trellis codes are in general nonlinear over F. Every code from both classes has a code trellis and can be decoded by Viterbi or BCJR algorithms.

18:40 *Polar Coding for Multi-level 3-Receiver Broadcast Channels*

Karthik Nagarjuna Tunuguntla (UC San Diego, USA); Paul H. Siegel (University of California, San Diego, USA)

We consider achieving the rates in the capacity region of a multi-level 3-receiver broadcast channel, in which the second receiver is degraded with respect to the first receiver, with degraded message sets. We propose a two-level chaining strategy based on polar codes that achieves the capacity region of the considered setting without time-sharing. We also look at a slight variation of this problem, where the first receiver only requires to decode its own private message and the other two receivers require to decode another private message common to them. We observe that the capacity region does not enlarge and so the proposed polar coding strategy achieves the capacity region for this problem as well.

Thursday, April 15 19:00 - 20:00 (Europe/Rome)

Plenary - Diversity vs. Parallelism in Distributed Computing with Redundancy weconf

Emina Soljanin

Chair: Olgica Milenkovic (University of Illinois at Urbana-Champaign (UIUC), USA)

Distributed computing enables parallel execution of tasks that make up a large computing job. In large scale systems, even small random fluctuations in service times (inherent to computing environments) often cause a non-negligible number of straggling tasks with long completion times. Redundancy, in the form of simple task replication and erasure coding, has emerged as a potentially powerful way to curtail the variability in service time, as it provides diversity that allows a job to be completed when only a subset of redundant tasks gets executed. Thus both redundancy and parallelism reduce the execution time, but compete for resources of the system. In situations of constrained resources (e.g., fixed number of parallel servers), increasing redundancy reduces the available level of parallelism. This talk will present the diversity vs. parallelism trade off for some common models of task size dependent execution times, and show that different models operate optimally at different levels of redundancy, and thus require very different code rates.

Thursday, April 15 20:10 - 21:50 (Europe/Rome)

Themed Session - Coding Theory and Applications weconf

Chairs: Ryan Gabrys (SPAWAR Pacific, USA), Eitan Yaakobi (Technion, Israel)

20:10 *Capacity and Construction of Recoverable Systems*

Ohad Elishco (Ben-Gurion University of the Negev, Israel); Alexander Barg (University of Maryland, USA)

Motivated by the established notion of storage codes, we consider sets of infinite sequences over a finite alphabet such that every \(k\)-tuple of consecutive entries is uniquely recoverable from its \(l\)-neighborhood in the sequence. We address the problem of finding the maximum growth rate of the set, which we term capacity, as well as constructions of explicit families that approach the optimal rate. We will present two construction of recoverable systems. One is of recursive nature, and the other relies on a modified of de-Bruijn graphs. The techniques that we employ rely on the connection of this problem with constrained systems. In addition, we consider a modification of the problem wherein the entries in the sequence are viewed as random variables over a finite alphabet that follow some joint distribution, and the recovery condition requires that the Shannon entropy of the \(k\)-tuple conditioned on its \(l\)-neighborhood be bounded above by some \(\epsilon>0\). We present some properties of measures that maximize the entropy and show a connection between recoverable systems and \(\epsilon\)-recoverable systems.

20:30 *On tilings of asymmetric limited-magnitude balls*

Hengjia Wei and Moshe Schwartz (Ben-Gurion University of the Negev, Israel)

We study whether an asymmetric limited-magnitude ball may tile \(Z^n\). This ball generalizes previously studied shapes: crosses, semi-crosses, and quasi-crosses. Such tilings act as perfect error-correcting codes in a channel which changes a transmitted integer vector in a bounded number of entries by limited-magnitude errors. A construction of lattice tilings based on perfect codes in the Hamming metric is given. Several non-existence results are proved, both for general tilings, and lattice tilings. A complete classification of lattice tilings for two certain cases is proved.

20:50 ***Repairing Reed-Solomon Codes in Practice***

Lakshmi J Mohan (RMIT University, Australia); Xinh Dinh (Tay Nguyen University, Vietnam); Luu Y Nhi Nguyen (Deakin University, Australia); Son Hoang Dau (RMIT University, Australia)

Recently, there has been considerable interest in finding the optimal repair bandwidths for Reed-Solomon codes. In this talk, we will discuss several aspects of this problem for short-length codes that are currently or will be potentially used in major distributed storage systems. For instance, we will discuss the impact of the set of evaluation points on the optimal repair bandwidth as well as different techniques used in the heuristic search for low-bandwidth repair schemes for such Reed-Solomon codes.

21:10 ***Robust Neural Computation From Error Correcting Codes***

Netanel Raviv (Washington University in Saint Louis, USA)

Neural networks (NNs) are a driving force behind the ongoing information revolution, with a broad spectrum of applications affecting most aspects of science and technology. The interest in robust neural computation under adversarial noise has increased lately, due applications in sensitive tasks ranging from healthcare to finance and autonomous vehicles. This has ignited an influx of research on the topic, which for the most part focuses on obtaining robustness by altering the training process. In contrast, this paper surveys and develops a recently proposed novel approach to obtain robustness \textit{after} training, by adding redundancy to the network and to the data in the form of error correcting codes. Since neural networks are essentially a concatenation of linear classifiers, we focus on obtaining robustness for a single linear classifier by coding the input and the classifier, and then apply the results on the network. We address two different types of adversaries, a worst-case one and an average-case one. For a worst-case adversary, that can choose the input to be attacked, we focus on binarized classifiers and show that the problem is related to construction of certain linear codes with restricted weight patterns. As a result, it is shown that the parity code can obtain robustness against any 1-erasure in any binarized NN, and no decoding is required. For an average-case adversary, that is given a uniformly random input to be attacked, it is shown that the optimal weights for any classifier and any code are given by the Fourier coefficients of that classifier. We demonstrate the latter experimentally, exposing improved accuracy-robustness tradeoff in neural classification of several popular datasets under state-of-the-art attacks.

21:30 ***Concatenated Codes for Recovery From Multiple Reads of DNA Sequences***

Andreas Lenz (Technische Universität München, Germany); Issam Maarouf (University of Bergen, Norway); Lorenz Welter (Technical University of Munich(TUM), Germany); Antonia Wachter-Zeh (Technical University of Munich (TUM), Germany); Eirik Rosnes (Simula UiB, Norway); Alexandre Graell i Amat (Chalmers University of Technology, Sweden)

Decoding sequences that stem from multiple transmissions of a codeword over an insertion, deletion, and substitution channel is a critical component of efficient deoxyribonucleic acid (DNA) data storage systems. In this paper, we consider a concatenated coding scheme with an outer low-density parity-check code and either an inner convolutional code or a block code. We propose two new decoding algorithms for inference from multiple received sequences, both combining the inner code and channel to a joint hidden Markov model to infer symbolwise a posteriori probabilities (APPs). The first decoder computes the exact APPs by jointly decoding the received sequences, whereas the second decoder approximates the APPs by combining the results of separately decoded received sequences. Using the proposed algorithms, we

evaluate the performance of decoding multiple received sequences by means of achievable information rates and Monte-Carlo simulations. We show significant performance gains compared to a single received sequence.