Universal Serially Concatenated Trellis Coded Modulation for Space-Time Channels

Wen-Yen Weng, Cenk Köse, Bike Xie, and Richard D. Wesel

Electrical Engineering Department, UCLA,

Los Angeles, CA, 90095-1594.

Abstract

In this paper, we propose serially concatenated trellis coded modulations (SCTCMs) that perform consistently close to the available mutual information for periodic erasure channel (PEC), periodic fading channel (PFC) and the 2×2 compound matrix channel. We use both the maximum-likelihood decoding criteria and iterative decoding criteria to design universal SCTCMs for the PEC and the PFC. For the space-time channel, by de-multiplexing the symbols across the antennas, the proposed universal SCTCMs for the period-2 PFC deliver consistent performance over the eigenvalue skew of the matrix channel. Within the family of channels having the same eigenvalue skew, a time-varying linear transformation (TVLT) is used to mitigate the performance variation over different eigenvectors. The proposed space-time SCTCMs of 1.0, 2.0 and 3.0 bits per transmission require excess mutual information in the ranges 0.11-0.15, 0.23-0.26 and 0.35-0.53 bits per antenna, respectively. Because of their consistent performance over all channels, the proposed codes will have good frame-error-rate (FER) performance over any quasi-static fading distribution. In particular, the codes provide competitive FER performance in quasi-static Rayleigh fading.

I. Introduction

Often, design of channel codes focuses on the optimization of performance on a specific channel such as the additive white Gaussian noise (AWGN) channel, or on average performance under a specific channel probability distribution, such as the Rayleigh fading channel. Powerful error-correcting codes such as turbo codes [1] and Low-Density Parity-Check (LDPC) codes [2] have been shown to operate within a dB of the Shannon limit on the AWGN channel. Numerous turbo TCMs [3][4][5][6] have been designed to optimize average performance for space-time Rayleigh fading channels. Typically, recursive space-time trellis codes that satisfy the slow-fading criteria are used as constituent codes. These codes perform well on the intended channel or distribution. However, the performance can degrade significantly over some specific channel realizations.

Root and Varaiya's compound channel coding theory for linear Gaussian channels indicates that a single code can reliably transmit information at R bits/symbol on each channel in the ensemble of linear Gaussian channels with mutual information (MI) larger than the attempted rate [7]. In related work, Sutskover and Shamai [8] recently proposed a setting for decoding of LDPC codes jointly with channel estimation for transmission over memoryless compound channels.

In [9], [10] the term *universal* was used to describe a channel code that has a good bit-error-rate (BER) or frame-error-rate (FER) for every channel in a family, $\mathcal{H}$, of channels. The code is said to be universal over $\mathcal{H}$. These papers presented trellis codes that approach universal behavior on space-time channels with a proximity to the Shannon limit similar to that with which trellis codes approach capacity on the AWGN channel. This paper presents serially concatenated trellis coded modulations (SCTCMs) with universal behavior on three channel families – the period-2 periodic erasure channel (PEC), the period-2 periodic fading channel (PFC) and the 2×2 space-time channel.

The period-2 PFC is specified by

$$y_t = a_{(t \bmod 2)} x_t + n_t, \tag{1}$$

where x_t is the transmitted symbol with average energy E_s, n_t is the additive white Gaussian noise (AWGN) with variance $N_0/2$ per dimension, and y_t is the received symbol. The vector $\underline{a} = [a_0 \;\, a_1]$ describes the nature of the time-varying attenuation behavior of the channel. Assuming $a_0 \geq a_1$, the period-2 PFC, $\underline{a} = [a_0 \;\, a_1]$, can be normalized as the $[1 \;\, \mathbf{q}]$ channel with $0 \leq \mathbf{q} \leq 1$ where $\mathbf{q} = a_1/a_0$. Such channels appear in frequency-hopped or multicarrier transmissions as well as diagonally-layered space-time architectures [11]. Wesel et al. [12], [13] identified universal trellis codes for the PFC.

As a subset of the PFC, the period-2 PEC consist of two instances of the attenuation vector. When $\underline{a} = [1 \;\, 0]$, every other symbol is erased, and when $\underline{a} = [1 \;\, 1]$, there are no erasures, the channel is the standard AWGN channel. We have found (see [11] Section IV and [9] Theorem 3) that the performance on the normalized [1 $\mathbf{q}$] channel is approximately bounded by the [1 1] and [1 0] channels. Hence, the problem of designing universal SCTCMs for the PFC can be simplified to designing SCTCMs for the PEC.

For the linear Gaussian vector channels (space-time channels)

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2}$$

where $\mathbf{H}$ denotes the $N_r \times N_t$ channel matrix, $\mathbf{x}$ is the $N_t \times 1$ vector of transmitted symbols, one for each transmit antenna, $\mathbf{n} \sim \mathcal{N}(0, N_o I_{N_r})$ denotes the additive white Gaussian noise vector, and $\mathbf{y}$ is the $N_r \times 1$ vector of received symbols. N_t and N_r are the number of transmit antennas and receive antennas respectively in the Multiple-Input Multiple-Output (MIMO) system.

In [14], [15], Zheng et al. examine the trade-off between diversity and multiplexing in MIMO systems. This trade-off turns out to be related to universal behavior. In [16], Tivildar and Viswanath give a precise characterization of approximately universal codes. These universal codes are also universally optimal in their trade-off between diversity and multiplexing. Tse and Viswanath provide an excellent overview of universal codes and their role in the diversity-multiplexing tradeoff in [17].

In this paper, we present SCTCMs of 0.5, 1.0 and 1.5 bits per symbol that that provide approximately universal performance for the PEC and the PFC. For the 2×2 space-time channel, our technique de-multiplexes universal SCTCMs for the PFC (as designed in [10], [18]) across the antennas to transmit at 1, 2, and 3 bits per channel use. This structure can deliver consistent performance over a wide range of possible eigenvalue skews. However, nearly singular channels still display a variation due to the particular eigenvectors. The performance difference over eigenvectors can be largely mitigated, but not completely eliminated, by a time-varying linear transformation (TVLT) technique introduced in [19]. The current paper employs a generalized form of TVLT, which uniformly sweeps three phase parameters over $[0, 2\pi)$.

The rest of the paper is organized as follows. Section II defines the channel model and the parameters used to sample the channel space. In Section III, we present the SCTCM design criteria on PEC for maximum-likelihood (ML) decoding and iterative decoding. Section IV shows the simulation results of the proposed SCTCMs on PEC and PFC. Section V analyzes the TVLT technique. Section VI presents the consistent excess mutual information (EMI) requirement of universal SCTCMs and their performance on the quasi-static Rayleigh fading channel. Section VII concludes the paper.

II. Channel Models and Performance Measure

As in [9], [11], and [12], we use EMI to measure the proximity with which the code approaches the theoretical limit of the channel. For the period-2 PFC and PEC in Eq. (1), the mutual information (MI) between the input symbol x and output symbol y is

$$\mathrm{MI}\left(\underline{a}, E_s\right) = \frac{1}{2}\sum_{i=0}^{1}\log_2\left(1+\frac{|a_i|^2 E_s}{N_0}\right). \tag{3}$$

According to the compound channel coding theorem [7], if MI $\geq R$, the error probability of a universal code can decrease to zero in the limit as its blocklength goes to infinity.

In practice, we design a finite-blocklength code transmitting at R-bits per symbol that achieves its target BER, say 10^{-5}, at $SNR^* = \frac{E_s^*}{N_o}$. The asterisks (*) indicate that this is the SNR at which the target BER is achieved. Then, we define the EMI of this code for the PFC as:

$$\mathrm{EMI}(\underline{a}) = \mathrm{MI}(\underline{a}, E_s^*) - R. \tag{4}$$

For the space-time channel in Eq. (2), the MI between the input vector $\mathbf{x}$ and output vector $\mathbf{y}$ is

$$\mathrm{MI}\left(\mathbf{H}, E_s\right) = \log_2\ \det\left(I_{N_r} + \frac{E_s}{N_o}\mathbf{H}\mathbf{H}^{\dagger}\right) = \sum_{i=1}^{\min(N_t,N_r)}\log_2\left(1+\frac{E_s}{N_o}\lambda_i\right). \tag{5}$$

where E_s is the energy per symbol per transmit antenna, $\mathbf{H}^{\dagger}$ is the Hermitian matrix of $\mathbf{H}$ and $\lambda_1, \lambda_2, \cdots, \lambda_{\min(N_t,N_r)}$ are the eigenvalues of $\mathbf{H}\mathbf{H}^{\dagger}$. For the 2×2 space-time channel, assume that $\lambda_1 \geq \lambda_2$ and define the eigenvalue skew $\kappa = \frac{\lambda_2}{\lambda_1}$, then the MI is given by

$$\mathrm{MI}\left(\mathbf{H}, E_s\right) = \log_2\left(1+\frac{E_s}{N_o}\lambda_1\right)\left(1+\frac{E_s}{N_o}\kappa\lambda_1\right), \tag{6}$$

and again using E_s^* to denote the symbol energy that achieves the target BER. The EMI of this code for the channel $\mathbf{H}$ is

$$\mathrm{EMI}(\mathbf{H}) = \mathrm{MI}(\mathbf{H}, E_s^*) - R. \tag{7}$$

At a fixed BER and with a fixed input information blocklength, the universal channel coding problem is a multicriterion optimization problem. Our goal is to find an SCTCM that is a Pareto optimal over all channels in the family. One possible objective function is the maximum EMI requirement over all the channels in $\mathcal{H}$, yielding a minimax problem. Although Root and Varaiya [7] showed that EMI goes to zero over each of the channels in the ensemble when the blocklength goes to infinity, it is still an open problem how fast the EMI decreases as a function of the blocklength and whether the EMIs are the same for all channels. Empirical evidence thus far [9], [11], and [12] indicates that universal codes tend to have relatively constant EMIs. As a result, it is natural to use the minimax criteria:

$$\min \max_{H_i \in \mathcal{H}} \mathrm{EMI}(H_i). \tag{8}$$

Another possible objective function is the average EMI per channel,

$$\overline{\mathrm{EMI}} = \frac{1}{|\mathcal{H}|} \sum_i \mathrm{EMI}(H_i), \tag{9}$$

which is equivalent to the objective function, J_{MI}, in [12]. Empirical evidence also shows that minimax and average EMI criteria usually yield the same universal SCTCM.

To verify the universality of the proposed codes, codes are simulated over all possible realizations of the channel ensemble. For the period-2 PEC, this is not a problem since there are only two instances. As for the period-2 PFC and the 2×2 space-time channel which have infinite channel realizations, we use a fine sampling of the channel space under the assumption of continuity.

For the $[1\ \mathbf{q}]$ PFC, the sampling is done over $\mathbf{q}$ between 0 and 1. For the 2×2 space-time channel, the channel is parameterized as follows. First, by singular value decomposition (SVD), $\mathbf{H}$ can be written as

$$\mathbf{H} = U \Lambda V^{\dagger}, \tag{10}$$

where U and V are unitary matrices. Using the following notations:

$$\Lambda \triangleq \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix}, \; M(\psi) \triangleq \begin{bmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{bmatrix}, \; D(\psi) \triangleq \begin{bmatrix} 1 & 0 \\ 0 & e^{j\psi} \end{bmatrix}, \tag{11}$$

$V^{\dagger}$ can be written as

$$V^{\dagger} = e^{j\mu} D(\omega) M(\phi) D(\theta). \tag{12}$$

At the receiver, we multiply on the left by a unitary matrix $P = e^{-j\mu} D(-\omega) U^{\dagger}$ so that the equivalent channel becomes

$$\mathbf{H}_{eq} = \Lambda M(\phi) D(\theta) = \begin{bmatrix} \sqrt{\frac{1}{1+\kappa}} & 0 \\ 0 & \sqrt{\frac{\kappa}{1+\kappa}} \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi \cdot e^{j\theta} \\ -\sin\phi & \cos\phi \cdot e^{j\theta} \end{bmatrix}. \tag{13}$$

Note that $\kappa = \frac{\lambda_2}{\lambda_1}$ and $\lambda_1 + \lambda_2$ is normalized to 1. We will use Eq. (13) as the channel model to evaluate the performance of SCTCMs on the 2×2 channel space by sampling over κ, ϕ and θ.

III. SCTCM Design Criteria for the PEC

A. Maximum-Likelihood Decoding Criteria

An SCTCM consists of a rate-R_o outer convolutional encoder, C_o, an interleaver, a rate-R_i inner convolutional encoder, C_i, and a two-dimensional 2^n-point constellation mapper. The throughput of the overall scheme is nR_oR_i bits per symbol. The uniform interleaver analysis of [20] has shown that the BER under ML decoding decreases exponentially with the effective free Euclidean distance of the inner code. For a large interleaver size of N_{IL} bits

$$-\log \text{BER} \propto \lfloor (d_o+1)/2 \rfloor \log N_{IL} + \delta^2 \frac{E_s}{4N_0} \tag{14}$$

where d_o is the free Hamming distance of the outer code, and

$$\delta^2 = \begin{cases} \frac{1}{2} d_o d_i^2 & \text{for } d_o \text{ even} \\ \frac{1}{2}(d_o - 3) d_i^2 + (h_m^3)^2 & \text{for } d_o \text{ odd} \end{cases} \tag{15}$$

where d_i is the effective free Euclidean distance of the inner code, defined to be the minimum Euclidean distance of error events caused by two information bit errors, and $h_m^{(3)}$ is the minimum Euclidean distance of error events caused by three information bit errors.

The BER approximation in Eq. (14) extends to the period-2 PEC when the input to the channel is considered to be a vector of two consecutive symbols. Two necessary conditions for universal performance over the compound period-2 PEC are:

1. The concatenated code should have positive redundancy under periodic erasures, i.e., $2R_oR_i < 1$.
2. The inner TCM should have nonzero effective free Euclidean distance under periodic erasures. In particular, for a given number of states, the rate of the inner TCM should be small enough to avoid parallel transitions. This condition results from an extension of the uniform interleaver analysis of [20].

The specific requirement on rate, expressed as a bound on k, to avoid parallel transitions is described by the following lemma:

Lemma: For $k > 1$, an S-state rate-k/n inner encoder mapped to a 2^n-point constellation has *zero* minimum effective free Euclidean distance under erasure if

$$\binom{k}{2} > S. \tag{16}$$

Proof: For $k > 1$, the encoder trellis has $\binom{k}{2}$ transitions per state to describe the inputs with Hamming weight 2. If $\binom{k}{2} > S$, there are parallel transitions with input Hamming weight 2 between states and shortest error events are 1-symbol long, and that symbol is subject to erasure. ■

For a 4-state inner code, the maximum k for nonzero d_i on the [1 0] channel is less than or equal to 3 and for an 8-state inner code, the maximum k for nonzero d_i is 4.

B. Iterative Decoding Criteria

Unlike the ML-decoding design criteria which focus on maximizing the minimum distance, the iterative decoding design criteria focus on finding codes with low pinch-off thresholds. In general, iterative decoding design criteria employ Extrinsic Information Transfer (EXIT) charts [21] or density evolution [22] to predict the pinch-off thresholds of the SCTCMs. The height of the error floor depends on the minimum distance of the code and cannot be predicted from these techniques.

Although not rigorously proved, it is generally observed by [2], [20], [23] that for turbo codes and LDPC codes, there exists a tradeoff between low pinch-off threshold and low error floor. Therefore, we considered both ML-decoding criteria and iterative decoding criteria in our search such that the target BER occurs at the smallest possible SNR.

EXIT chart analysis predicts the waterfall region in the BER curve as the blocklength goes to infinity. The technique analyzes iterative decoding by tracking the density (probability distribution) of the extrinsic information of soft-input soft-output (SISO) *a posteriori* probability (APP) modules as this density evolves from iteration to iteration. In [24], El Gamal and Hammons showed that when the noise is Gaussian, the extrinsic information in the process of iterative decoding can be closely approximated by symmetric Gaussian distribution, where the variance is twice the mean. Using Gaussian approximation, the threshold of convergence of turbo codes can be predicted efficiently.

The threshold can indeed be predicted accurately in this way on the [1 1] channel. However, when the SCTCM is working on the [1 0] channel, the density of the extrinsic information at the output of the inner SISO doesn't follow a symmetric Gaussian distribution. This is because the erased bits tend to have lower log-likelihood ratios (LLRs) than that of the unerased bits. So, the overall pdf resembles a bi-modal, not necessarily symmetric, Gaussian mixture distribution. This effect makes it difficult to analyze the outer and inner SISO separately. The simple extrinsic information calculation resulting from the symmetric Gaussian approximation becomes invalid. The exact densities must be manipulated to analyze the SISOs. As a result, there is no computational advantage for this exact density evolution over explicit simulation. Hence, our approach of finding the threshold is to simulate the candidate SCTCMs and take the SNR with BER = 10^{-1} as its pinch-off threshold.

1) Design of the outer encoder: We use convolutional codes with large free Hamming distance as the outer code. For design simplicity and because it is the inner TCM that directly interfaces the channel, we fix the outer code and find the inner TCM such that the SCTCM has the lowest pinch-off threshold. Table I lists the rate-1/2, 2-, 4- and 8-state maximal free Hamming distance outer codes

$C_{o,1}$, $C_{o,2}$, and $C_{o,3}$ respectively.

2) Design of the inner TCM: Although we cannot use EXIT chart analysis with Gaussian approximation to design SCTCMs with low thresholds on the [1 0] channel, some design guidelines developed from EXIT chart analysis still help. They are described as follows:

- Complexity of constituent codes: It has been found that SCTCMs with lower complexity constituent codes tend to have lower pinch-off thresholds. However, the SCTCMs may suffer a high error floor because of the tradeoff. So, low complexity convolutional codes are preferred if the corresponding error floor is low enough to meet the BER requirement.
- Trellis-collapse check: An SCTCM that works on the [1 0] channel is equivalent to another SCTCM with twice the code rate working on the AWGN channel. In other words, because the erased symbols can't transmit any information at all, pairs of consecutive trellis stages collapse into a single stage. By checking the equivalent trellis-collapsed inner codes, we found the following properties useful to reduce the population of candidate codes.
  1. State reduction: The number of states of the collapsed inner code may become fewer than that of the original encoder because some registers simply never affect the transmitted symbol and are thus useless. For example, a 4-state, rate-2/3 inner code, $C_{i,5}$ (see Table I), on the [0 1] channel, becomes equivalent to a trellis-collapsed 2-state, rate-4/3 code, $C_{i,5}^{\text{tc}}$(see Table I), operating on the AWGN channel. Here equivalent means having exactly the same input-output behavior on the unerased symbols. By the iterative decoding design criteria, constituent encoders with fewer states tend to yield lower pinch-off thresholds. This is especially important for the [0 1], [1 0] channels.
  2. Non-transmitted input bits: After collapsing the inner code, some portion of the input bits to the inner code may not be transmitted at all. For example, $C_{i,12}$ (see Table I) has a [0 1] channel trellis-collapsed equivalent encoder, $C_{i,12}^{\text{tc}}$ (see Table I), which has one non-transmitted bit per symbol (the last row of its generator matrix consists of all zeros). With this encoder, no extrinsic information for this non-transmitted bit can be extracted from the inner SISO. This structural defect is not desired and it results in high error floors

In essence, we search for inner codes whose trellis-collapsed equivalent codes have state reduction and without non-transmitted bits. The qualified codes are then simulated to find the best code that has the lowest pinch-off threshold.

- Constellation labeling

  Constellation labeling affects the performance of SCTCM on both [1 1] and [1 0] channels. As ten Brink [21] shows, the systematic bits play an important role in SCTCM design. The EXIT chart analysis indicates that systematic inner codes have higher initial extrinsic information in their transfer curves than non-systematic codes and thus are more likely to have low pinch-off thresholds. Using a similar concept, our constellation labeling is selected such that the systematic bits of the inner code are best protected.

  Here we take two extreme cases of 8-PSK labelings as examples to illustrate this idea. Assume that we have a systematic rate-$2/3$ inner code with the two MSB as systematic bits. A Gray-labeled 8-PSK ([0, 1, 3, 2, 6, 7, 5, 4]) provides the best protection on the two systematic bits while natural labeling ([0, 1, 2, 3, 4, 5, 6, 7]) offers very poor protection. Figure 1 is the EXIT chart of an SCTCM with the above two 8-PSK labelings working on the AWGN with E_s/N_o=1.1 dB. The Gray-labeled inner TCM has already opened up the decoding tunnel, while the tunnel of the Natural-labeled inner TCM is still closed at this E_s/N_o. The latter will require the E_s/N_o=3.2 dB to open the tunnel.

  When the inner TCM consists of a rate-3/4 code mapped to 16-QAM, it is not possible as in 8-PSK to have maximal protection on each of the systematic bits because there are three input bits but the constellation is 2-dimensional. In this case, a Gray-labeled 16-QAM in raster order using hexadecimal [3, 1, 5, 7; 2, 0, 4, 6; a, 8, c, e; b, 9, d, f] provides the best possible protection of the systematic bits.

Table I lists the proposed inner convolutional codes with code rate-1/3, -2/3, and -3/4, and the corresponding design criteria used. The mapping is Gray-labeled 8PSK or 16QAM. The ML-optimal inner TCM has the largest d_i. As for the iterative-decoding optimal inner TCM, we perform an exhaustive search for inner TCMs whose trellis-collapsed equivalent codes have state reduction and without non-transmitted bits. The qualified candidate codes are then simulated to determine which one has the lowest threshold. These codes are used to construct the SCTCMs in Section IV.

3) Interleaver design: The interleaver proposed is the extended spread interleaver [25] with a further constraint depending on the outer code and the periodic erasure pattern. Intuitively, if the intermediate bits, i.e. the output bits of C_o, that contain the information of a single input information bit are all interleaved to erased symbols, it becomes very difficult for the decoder to decode this input bit since only a small amount of information is preserved through the parity bits. So, the interleaver design aims to prevent the *all-erased* cases from happening for any of the input bits.

For example, let's design an interleaver for an SCTCM consisting of the rate-1/2 ($k_o = 1, n_o = 2$) $C_{o,2}$ from Table I and a rate-3/4 systematic C_i ($k_i = 3, n_i = 4$) with a 16-QAM constellation. The impulse response of $C_{o,2}$ is $h = [11\ 01\ 11\ 00 \cdots 00]$ indicating that five of the intermediate bits are associated with a single input bit. A bit-interleaver of length N_{IL}, is defined as $\{\pi(i) = j,\ i = 0, \cdots, N_{IL} - 1\}$ which interleaves the i^{th} interleaver input to the j^{th} interleaver output. Then the further constraint of the interleaver considering the periodic-2 erasure patterns, [1 0] and [0 1], is:

$$1 \leq \sum_{v \in S_h} \{\lfloor \pi(n_o \cdot l + v)/k_i \rfloor (\text{modulo } 2)\} \leq |S_h| - 1 = 4,\ \forall\ l = 0, \cdots, N_i - 1. \quad (17)$$

where N_i is the input blocklength, $S_h = \{\, w : h(w) = 1\} = \{\, 0, 1, 3, 4, 5\}$, $n_o = 2$, and $k_i = 3$ in this example. Essentially, Eq. (17) guarantees that at least one of these five bits will not have its inner-code output-symbol erased.

Compared with spread interleavers without this further constraint, the new interleaver can improve the threshold of this SCTCM by about 0.05 bits of EMI (0.4 dB of E_s/N_o) on the [1 0] and [0 1] channels while maintaining almost the same performance on the [1 1] channel.

IV. Example SCTCM Designs for the PEC

Table II lists the proposed SCTCMs transmitting at 0.5, 1.0 and 1.5 bits per symbol using the constituent codes in Table I. The EMI requirement, corresponding E_s/N_o and error floor of each SCTCM on the [1 1] and [1 0] channels are presented. In this section, we will discuss and compare the performance of these SCTCMs. All SCTCMs in this section have blocklength of 10,000 bits.

A. SCTCM Design of 0.5 bits per symbol

We used a rate-1/2 C_o and a rate-1/3 C_i mapped to an 8-PSK constellation to provide 0.5 bits/symbol overall throughput. The inner TCM still has redundancy under period-2 erasures becoming effectively a rate-2/3 TCM. An exhaustive search over the 8-state, rate-1/3 systematic feedback encoders using the trellis-collapse check yielded several inner encoders with reduced complexity and among these encoders, $C_{i,1}$ has the lowest threshold.

SC-1 (Table II) concatenates $C_{i,1}$ with a 4-state maximum free-distance outer code, $C_{o,2}$. This scheme has an error floor around BER = 10^{-5} on the [1 0] channel. Two different methods are used to lower the error floor. SC-2 increases the inner effective free distance by replacing $C_{i,1}$ with an ML-optimal $C_{i,2}$, which has the same complexity as C_{i_1}. On the other hand, SC-3 replaced the 4-state $C_{o,2}$ with the 8-state $C_{o,3}$ to increase the outer distance. SC-2 and SC-3 performs similarly and both trade a little pinch-off threshold for lower error floors.

B. SCTCM Design of 1.0 bit per symbol

Our 1.0 bit per symbol SCTCM uses a rate-1/2 C_o and a rate-2/3 C_i driving an 8-PSK constellation. The rate-2/3 inner TCM has *negative redundancy* becoming effectively a rate-4/3 TCM under period-2 erasures and therefore has zero minimum free distance under period-2 erasures. However, it is still possible to find an inner code with a nonzero minimum *effective* free distance during the ML-optimal code search. Iterative decoding criteria, on the other hand, concentrates on low pinch-off threshold and may allow the zero effective free distance.

SC-4 and SC-5 use the same inner TCM ($C_{i,5}$ with a Gray-labeled 8-PSK) found by iterative decoding criteria. The difference here is the complexity of the outer code, 2-state versus 4-state. It shows that we need at least a 4-state C_o to deliver the desired error floor. SC-5, SC-6 and SC-7 provides another set of comparison with a fixed outer code, $C_{o,2}$, and three different 4-state inner codes. $C_{i,5}$ (for SC-5) and $C_{i,4}$ (for SC-6) are good candidate codes according to iterative decoding criteria. They both have reduced complexity without non-transmitted bits after the [0 1] channel trellis collapse. On the other hand, $C_{i,6}$ (for SC-7) has maximal-d_i for both [1 1] and [1 0] channels. The [1 1] channel performances are all similar. On the [1 0] channel, iterative-optimal SC-5 and SC-6 converge significantly earlier than the ML-optimal SC-7. If the objective BER is 10^{-5}, SC-6 will be the best choice with an average EMI of 0.270 bits at BER=10^{-5}. SC-5 requires slightly larger average EMI of 0.280 bits but it has an error floor lower than 10^{-7}.

SC-9 and SC-10 use 8-state inner TCMs designed to be iterative decoding optimal and ML-optimal respectively. Compared to the 4-state inner TCM schemes, SC-5, SC-6 and SC-7, the increased complexity results in higher pinch-off thresholds. Again, the iterative decoding optimal code outperforms the ML-optimal code on the [1 0] channel (0.155 bits less EMI) without sacrificing on the [1 1] channel (0.003 bits more EMI).

C. SCTCM Design of 1.5 bits per symbol

Using the same criteria, we designed 1.5 bits/symbol SCTCMs by concatenating a rate-1/2 C_o with a rate-3/4 C_i mapped to a 16-QAM constellation (SC-11, SC-12 and SC-13).

D. Periodic fading channel performance

Figure 2 shows the EMI requirements of SC-2, SC-5, and SC-11 (the best code at each throughput with an error floor $< 10^{-7}$) on the [1 $\mathbf{q}$] channels for $\mathbf{q}$=0.2, 0.4, 0.6 and 0.8. For all three codes, the EMI requirements are approximately bounded by that of the [1 1] and [1 0] channels. This result is consistent with [11] Section IV and [9] Theorem 3. Also, the EMI is relatively flat over $\mathbf{q}$ for all three cases indicating consistently good performance for all of the [1 $\mathbf{q}$] channels. Therefore, these universal SCTCMs designed for the PEC are also universal for the PFC.

V. Time-Varying Linear Transformation (TVLT)

The proposed space-time SCTCM scheme uses an SCTCM designed for the PFC and de-multiplexes the output symbols into N_t symbol streams. The N_t-symbol vectors are then multiplied by a time-varying $N_t \times N_t$ unitary matrix before transmitted. Therefore, the throughput of the overall scheme is $nR_oR_iN_t$ (bits/s/Hz). On the decoder side, it is assumed that the decoder has perfect channel state information and the decoder also knows exactly the TVLT matrices. Therefore, the LLR of the symbols can be calculated and passed to the turbo decoder. Note that the received symbols form a $N_r \times 1$ vector and each received symbol is a superposition of N_t transmitted symbols. So, the inner SISO is based on the *collapsed trellis* which combines N_t trellis stages together into a super-trellis.

The 2×2 space-time channel is parameterized using κ and the channel angles ϕ and θ as shown in Eq. (13). The angle ϕ determines the amount of interference between the two antennas. When $|\cos(\phi)|$=1 or 0, the channel is diagonal with no interference. On the other hand, the interference is the largest when $|\cos\phi| = |\sin\phi| = \frac{1}{\sqrt{2}}$. The other parameter, θ, represents the phase difference between the constellations of the two antennas. Our experiments showed that although the MI is only a function of κ, different ϕ and θ can result in very different performances for SCTCM especially when κ is close to zero (the singular channel case).

We need to characterize EMI as a function of ϕ and θ but this is computationally difficult. Similar to the constellation labeling experiments in Sec. III-B, we propose an approximate method to predict the EMI performance by observing the correlation between the EMI and the initial extrinsic information (IEI). The higher the IEI of the inner SISO is, the more likely the whole inner-SISO EXIT curve stays above the outer-SISO EXIT curve and thus lower convergence threshold. Note that the IEI can be easily calculated by numerically computing the parallel independent decoding capacity of the systematic bits of the inner code.

Figure 3 shows the EMI requirement and the corresponding inner-SISO IEI of SC-11 at $E_b/N_o = 6.35$ dB as a function of ϕ where $\theta = \kappa = 0$. It is clear that the EMI requirement and the IEI are closely related. Based on the above observations, achieving a uniform IEI over ϕ and θ is way to enhance consistent EMI behavior and thereby approach universal performance. We tested three very different approaches to achieve a uniform IEI over ϕ and θ.

Our first idea was to design a combination of constellation and labeling that delivers a more uniform IEI over ϕ and θ than the conventional schemes using PSK/QAM and Gray labeling. Different constellation shapes were tested. However, their performances were still as dependent on ϕ and θ as those of the PSK/QAM constellations. As for the labeling, it is possible to design a labeling which yields higher IEI for some (ϕ,θ) angles, but at the same time, it also decreases the IEI at other angles. As a result, labelings other than the Gray labeling yield a less uniform IEI and an inferior worst-case performance.

Motivated by [26], our second idea was to make the two constellations of the two antennas correlated. Instead of mapping the n coded bits to a 2^n-point 2D constellation, the general form of a correlated constellation considers the two transmitted antennas together and uses $2n$ consecutive bits to map a 2^{2n}-point 4D constellation (2D per antenna). However, the performance becomes worse because increasing the number of constellation points actually decreases the IEI.

The third idea, which fortunately does work, uses a time-varying linear transformation (TVLT) to rotate the channel to different angles, hoping that the dependence on ϕ and θ can be "averaged out". This idea was first proposed in [19]. In the current paper, this concept is generalized so that a time-varying unitary matrix,

$$Q_t = D(\beta_t)M(\alpha_t)D(\gamma_t), \tag{18}$$

is multiplied to the signal vector, $\mathbf{x}$, before it is transmitted (see Eq. (11) for definitions of $M(\cdot)$ and $D(\cdot)$). By controlling the three parameters, this technique creates a time-varying equivalent channel $\tilde{\mathbf{H}} = \mathbf{H}Q_t$. Using (13) and suppressing the subscript t,

$$\begin{aligned} \tilde{\mathbf{H}} &= \mathbf{H}Q = [\Lambda M(\phi)D(\theta)]\,[D(\beta)M(\alpha)D(\gamma)] \\ &= \Lambda\,[M(\phi)D(\theta+\beta)M(\alpha)]\,D(\gamma). \end{aligned} \tag{19}$$

Because $M(\phi)D(\theta+\beta)M(\alpha)$ is also unitary, it can be decomposed into $e^{j\zeta}D(\tilde{\theta}_L)M(\tilde{\phi})D(\tilde{\theta}_R)$, where

$$\tilde{\phi} = \frac{1}{2}\cos^{-1}\left(\cos(2\phi)\cos(2\alpha) - \sin(2\phi)\sin(2\alpha)\cos(\theta+\beta)\right), \tag{20}$$

$$\zeta = \measuredangle\left\{\cos\phi\cos\alpha - \sin\phi\sin\alpha\cos(\theta+\beta) - j\sin\phi\sin\alpha\sin(\theta+\beta)\right\}, \tag{21}$$

$$\tilde{\theta}_L = \measuredangle\left\{\sin(2\phi)\cos(2\alpha) + \cos(2\phi)\sin(2\alpha)\cos(\theta+\beta) + j\sin(2\alpha)\sin(\theta+\beta)\right\}, \tag{22}$$

$$\tilde{\theta}_R = \measuredangle \left\{ \cos(2\phi)\sin(2\alpha) + \sin(2\phi)\cos(2\alpha)\cos(\theta+\beta) + j\sin(2\phi)\sin(\theta+\beta) \right\}. \tag{23}$$

Note that $e^{j\zeta}$ and $D(\tilde{\theta}_L)$ can be canceled out in the receiver. Hence the equivalent channel, represented in the same form as that in Eq. (13), becomes $\tilde{\mathbf{H}}_{eq} = \Lambda M(\tilde{\phi})D(\tilde{\theta})$, where $\tilde{\phi}$ is given by Eq. (20) and $\tilde{\theta} = \tilde{\theta}_R + \gamma$.

The ideal TVLT would vary α, β, and γ such that $\tilde{\phi}$ and $\tilde{\theta}$ both sweep uniformly over $[0, 2\pi)$. In that case, the SCTCM sees the same set of channels regardless of the actual channel angles ϕ and θ. Without any information of ϕ and θ in the transmitter, the best strategy is to sample uniformly over the 3-dimensional (α, β, γ) space. For fixed α and β values, if γ sweeps uniformly over $[0, 2\pi)$, $\tilde{\theta}$ also varies uniformly. However, due to the nonlinearity in Eq. (20), $\tilde{\phi}$ is not swept uniformly and still depends on ϕ and θ. Therefore, the performance of the SCTCM cannot be completely universal using TVLT.

Here we provide an analysis of the approximate performance of an SCTCM with TVLT. From the simulation results, we propose a model for EMI as a function of ϕ and θ as

$$\text{EMI}(\phi, \theta) \approx \text{EMI}_{\text{min}} + K_1[1 - \cos(4\phi)][K_2 + \cos(2\theta)] \tag{24}$$

where EMI_{min}, K_1 and K_2 depend only on κ and are found to minimize the mean square error. Figure 4 shows that the model corresponds well to the simulated EMI for SC-11 on singular channels ($\kappa = 0$) with $\text{EMI}_{\text{min}} = 0.40$, $K_1 = 0.0289$ and $K_2 = 3.74$. To obtain the approximation of EMI with TVLT, we assume that the EMI is the average of the instantaneous EMIs. Further, we assume a TVLT in which α, β, and γ are uniformly distributed, the average EMI with such a TVLT is given by

$$\begin{aligned} \text{EMI}^*(\phi, \theta) &\approx \frac{1}{N_\alpha N_\beta N_\gamma} \sum_\alpha \sum_\beta \sum_\gamma \text{EMI}(\tilde{\phi}, \tilde{\theta}) \\ &\approx \text{EMI}_{\text{min}} + \frac{K_1 K_2}{4} \cdot [5 - \cos(4\phi)]. \end{aligned} \tag{25}$$

where N_α, N_β and N_γ are the number of α, β and γ swept over $[0, 2\pi)$ respectively. The full derivation of Eq. (25) is in the Appendix. Note that the dependence on θ is eliminated but the EMI still depends on ϕ, although with a smaller variance.

A more relaxed sufficient condition than uniformly distributed α, β, and γ for Eq. (25) to hold is

$$\sum_{\alpha} e^{j4\alpha} = 0, \ \ \sum_{\beta} \cos(2\theta + 2\beta) = 0, \ \ \sum_{\gamma} \cos(2\gamma) = 0. \tag{26}$$

Therefore, We propose two TVLT schemes:

- Fine-sampled TVLT (F-TVLT): Sweep α, β and γ over $[0, 2\pi)$ uniformly with a step size of $\frac{2\pi}{\sqrt[3]{N_s}}$, where N_s denotes the number of symbols per block.
- Simplified TVLT (S-TVLT): By Eq. (26), use $\alpha = 0$ or $\pi/4$, $\beta = 0$ or $\pi/2$, and $\gamma = 0$ or $\pi/2$ which yields a total of 8 different unitary matrices compared with N_s matrices in F-TVLT.

As shown in Figure 5, TVLT makes the EMI requirements more concentrated and the overall performance is close to universal. The S-TVLT performance, which is not shown in the plot, is almost identical to that of the F-TVLT.

VI. Simulation Results of space-time SCTCM

Figure 5 shows the EMI for SC-2, SC-5 and SC-11 with and without F-TVLT over eigenvalue skew. For the same eigenvalue skew, the maximum and minimum EMIs are marked and the shaded region represents the region of operation. All the simulations are done with an input blocklength of 10,000 bits, 12 iterations at the decoder, and BER of 10^{-5}. The TVLT technique improves the universality of the SCTCMs. As a result, SC-2 and SC-5 provide uniform EMI requirements of no more than 0.15 bits and 0.26 bits per antenna respectively. SC-11 provides a uniform EMI of no more than 0.41 bits except for the sudden increase to 0.53 bits for the worst-case singular channel.

Next, we will compare the proposed universal SCTCMs with other coding schemes designed specifically for Rayleigh fading environment in terms of both average performance and universality (channel to channel performance). Among them, the best Rayleigh fading performance comes from Stefanov's [27] parallel concatenated convolutional code with bit-interleaved coded modulation (PCCC-BICM) which is 1.8 dB from the outage probability at FER=10^{-2}. Turbo-TCMs [5], [19], [28] are reported to operate at 2.0 to 2.2 dB from the outage probability. The universal SC-5 with F-TVLT is only 1.5 dB from the outage probability. SC-8, which maps a binary SCCC designed for

AWGN channel by Benedetto [20] directly to 8PSK, is 1.6 dB from the outage probability. For the Rayleigh fading performance, SC-5 and SC-8 were simulated with input blocklength of 260 bits, which is the same as that in [27] and [5]. [19] and [28], however, have blocklengths of 8,196 and 1,024 bits respectively. Moreover, the number of iterations in these papers are not exactly the same. So a fair conclusion would be that these codes all perform similarly on the Rayleigh fading channel.

On the other hand, Figure 7 and Figure 8 shows the EMI requirement region for SC-5, SC-8 and two PCCC-BICMs, all with blocklength of 10,000 bits. SC-5 clearly has a much narrower EMI region and hence is more universal than others, especially after applying TVLT. Note that the PCCC-BICM in [27] was originally rate-1/2 and mapped to a QPSK constellation to transmit 2 bits/s/Hz over two antennas. The system, in the worst condition of $\kappa = 0, \phi = 0$, becomes uncoded and thus requires a very large EMI. Therefore, we also designed a rate-1/3 PCCC-BICM mapped to 8-PSK. The worst case EMI is reduced to 0.8 bits per antenna but it is still not as universal as SC-5.

In Sec. IV-C in [19], Shi found that by multiplying $U^\dagger$ (see Eq. (10)), i.e. the R matrix in Shi's notation, at the receiver, the FER improves significantly but the effect of TVLT is invisible with or without $U^\dagger$. We also found that $U^\dagger$ is crucial for the SCTCM performance on the Rayleigh fading channel. However, unlike Shi, who only uses *real* unitary matrices in their experiments, our generalized TVLT matrices scan the whole *complex* unitary matrix space and provide a better averaging effect. As a result, for SC-5, TVLT improves the FER by 0.9 dB at FER=10^{-2} in the case of no $U^\dagger$ and merely 0.1 dB when $U^\dagger$ is applied. In general, the averaging over channels in a Rayleigh simulation hides the channel-by-channel consistency provided by the TVLT.

Since SC-5 requires a uniform EMI no more than 0.26 bits per antenna (i.e. 0.52 bits per transmission), SC-5 should be able to transmit 2 bits successfully over any channel which provides an MI of more than 2.52 bits. Figure 6 shows that the FER of SC-5 with F-TVLT is close to the outage probability of MI=2.52 bits. The beauty of the universal SCTCM is that SC-5 can perform close to the outage probability of a MI=2.52 bits not only for Rayleigh fading but *any* quasi-static fading channels.

VII. Conclusions

Designs of universal SCTCM schemes are proposed to deliver close-to-capacity performance for three families of channels – the period-2 PEC, the period-2 PFC, and the 2×2 space-time channel. First we design universal SCTCMs for the PEC considering both the ML-decoding criteria and the iterative decoding criteria. Every component of the SCTCM including constituent outer and inner codes, interleaver, constellation, and labeling are all considered and carefully chosen. At BER = 10^{-5}, the proposed SCTCMs transmitting at 0.5 bits/symbol, 1 bit/symbol and 1.5 bits/symbol require an average EMI of 0.116, 0.270, 0.370 bits respectively on the PEC. These codes are then shown to be also universal on the PFC through simulations.

By de-multiplexing the output symbols of an SCTCM to the transmit antennas sequentially, the PFC becomes equivalent to diagonal MIMO channels and PFC-universal SCTCMs show their universality over eigenvalue skew. Then for non-diagonal MIMO channels, it is found that the performance of SCTCM may degrade significantly due to the convergence behavior of the iterative decoder. We observed that the relation between EMI and IEI can be used to facilitate the comparison of different constellations and labelings. Experiments with fixed constellations and labelings could not achieve universal IEI. Hence, TVLT is used to rotate the channel matrix over time and is shown to effectively mitigate the dependence of EMI on the eigenvectors. As a result, the proposed SCTCMs of 1.0, 2.0 and 3.0 bits per channel use require a consistent EMI of 0.11-0.15, 0.23-0.26 and 0.35-0.53 bits per antenna and are 1.1, 1.5, and 2.1 dB respectively from the outage probability at FER=10^{-2} on the quasi-static Rayleigh fading channel. However, the main point of designing such codes is that they are not designed specifically for Rayleigh fading, and will work close to the theoretical limit on any quasi-static fading channel.

References

[1] C. Berrou and A. Glavieux. Near Optimum Error Correcting Coding and Decoding: Turbo-codes. *IEEE Trans. on Comm.*, 44(10):1261–1271, Oct. 1996.

[2] D. J. C. MacKay, S. T. Wilson, and M. C. Davey. Comparison of Constructions of Irregular Gallager Codes. *IEEE Trans. on Comm.*, 47(10):1449–1454, Oct. 1999.

[3] X. Lin and R. S. Blum. Guidelines for Serially Concatenated Space-Time Code Design in Flat Rayleigh Fading Channels. In *IEEE Third Workshop on Signal Processing Advances in Wireless Communications*, Mar. 2001.

[4] D. Tujkovic. High bandwidth efficiency space-time turbo coded modulation. In *ICC*, June 2001.

[5] Y. Hong, J. Yuan, Z. Chen, and B. Vucetic. Space-time turbo trellis codes for two, three, and four transmit antennas. *IEEE Trans. on Veh. Tech.*, 53(2):318–328, Mar. 2004.

[6] Y. Liu, M. P. Fitz, and O. Y. Takeshita. Full rate space-time turbo codes. *IEEE Journal on Sel. Areas in Comm.*, 19(5):969–980, May 2001.

[7] W. L. Root and P. P. Varaiya. Capacity of Classes of Gaussian Channels. *SIAM J. of Applied Math*, 16(6):1350–1393, 1968.

[8] I. Sutskover and S. Shamai. Iterative Decoding of Low-Density Parity-Check Codes Over Compound Channelsl. *IEEE Trans. on Comm.*, 54(2):308–318, Feb. 2006.

[9] C. Köse and R. D. Wesel. Universal Space-time Trellis Codes. *IEEE Trans. on Inform. Theory*, 49(10):2717–2727, Oct. 2003.

[10] C. Köse and R. D. Wesel. Universal Space-Time Codes from Demultiplexed Trellis Codes. *IEEE Trans. on Comm.*, 54(7):1243–1250, July 2006.

[11] A. Matache and R. D. Wesel. Universal Trellis Codes for Diagonally Layered Space-Time Systems. *IEEE Trans. on Signal Processing*, 51(11):2773 –2783, Nov. 2003.

[12] R. D. Wesel, X. Liu, and W. Shi. Trellis Codes for Periodic Erasures. *IEEE Trans. on Comm.*, 48(6):938–947, June 2000.

[13] R. D. Wesel and J. M. Cioffi. Trellis Codes for Periodic Interleavers. *IEEE Comm. Letters*, 3(4):103–105, Apr. 1999.

[14] L. Zheng and D. Tse. Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels. *IEEE Trans. on Inform. Theory*, 49(5):1073–1096, May 2003.

[15] L. Zheng and D. Tse. Diversity-Multiplexing Tradeoff in Multiple-Access Channels. *IEEE Trans. on Inform. Theory*, 50(9):1859–1874, Sept. 2004.

[16] S. Tavildar and P. Viswanath. Approximately Universal Codes over Slow Fading Channels. *To appear in IEEE Trans. on Inform. Theory*.

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

[18] C. Köse, W. Weng, and R. D. Wesel. Serially Concatenated Trellis Coded Modulation for the Compound Periodic Erasures Channel. In *ICC*, Anchorage, AK, May 2003.

[19] W. Shi, C. Komninakis, B. Daneshrad, and R. D. Wesel. Robustness of Space-Time Turbo Codes. In *ICC*, pages 1700–1704, Helsinki, Finland, June 2001.

[20] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. Serial Concatenation of Interleaved Codes: Performace Analysis, Design and Iterative Decoding. *IEEE Trans. on Inform. Theory*, 44(3):909–926, May 1998.

[21] S. ten Brink. Design of Serially Concatenated Codes Based on Iterative Decoding Convergence. In *Second International Symposium on Turbo Codes and Related Topics, Brest, France*, 2000.

[22] D. Divsalar, S. Dolinar, and F. Pollara. Iterative Turbo Decoder Based on Density Evolution. *IEEE Journal on Sel. Areas in Comm.*, 19(5):891–907, May 2001.

[23] T. Tian, C. Jones, J. Villasenor, and R. D. Wesel. Constructions of Irregular LDPC Codes with Low Error Floors. In *IEEE Int. Conf. on Comm.*, May 2003.

[24] H. El Gamal and A. R. Hammons. Analyzing the Turbo Decoder Using the Gaussian Approximation. *IEEE Journal on Sel. Areas in Comm.*, 47:672–686, Feb. 2001.

[25] C. Fragouli and R. D. Wesel. Semi-Random Interleaver Design Criteria. In *Communication Theory Symposium at Globecom 99, Rio de Juanero, Brazil*, Dec. 1999.

[26] B. Clerckx, D. Vanhoenacker-Janvier, C. Oestges, and L. Vandendorpe. Robust signal constellations for spatial multiplexing in the presence of real fading propagation channels. In *ICC*, Anchorage, AK, May 2003.

[27] A. Stefanov and T. M. Duman. Turbo-coded modulation for systems with transmit and receive antenna diversity over block fading channels: system model, decoding approaches, and practical considerations. *IEEE Journal on Sel. Areas in Comm.*, 19(5):958–968, May 2001.

[28] W. Firmanto, Z. Chen, B. Vucetic, and J. Yuan. Design of space-time turbo trellis coded modulation for fading channels. In *IEEE Globecom*, Nov. 2001.

APPENDIX

The derivation of the approximate average EMI with TVLT in Eq. (25) is as follows. Assume that α, β, and γ are uniformly distributed over $[0, 2\pi)$ and use the EMI model in Eq. (24),

$$\begin{aligned}
\mathrm{EMI}^*(\phi,\theta) &\approx \frac{1}{N_\alpha N_\beta N_\gamma}\sum_\alpha\sum_\beta\sum_\gamma \mathrm{EMI}\left(\tilde{\phi},\tilde{\theta}\right) && (27)\\
&\approx \mathrm{EMI}_{\min} + \frac{K_1}{N_\alpha N_\beta N_\gamma}\sum_\alpha\sum_\beta\left[1-\cos\left(4\tilde{\phi}\right)\right]\sum_\gamma\left[K_2+\cos\left(2\tilde{\theta}\right)\right] && (28)\\
&= \mathrm{EMI}_{\min} + \frac{K_1}{N_\alpha N_\beta N_\gamma}\sum_\alpha\sum_\beta\left[1-\cos\left(4\tilde{\phi}\right)\right]\sum_\gamma\left[K_2+\cos\left(2\tilde{\theta}_R+2\gamma\right)\right] && (29)\\
&= \mathrm{EMI}_{\min} + \frac{K_1K_2}{N_\alpha N_\beta}\sum_\alpha\sum_\beta\left[1-\cos\left(4\tilde{\phi}\right)\right]. && (30)
\end{aligned}$$

By Eq. (20), $\cos\left(2\tilde{\phi}\right) = \cos(2\phi)\cos(2\alpha) - \sin(2\phi)\sin(2\alpha)\cos(\theta+\beta)$. Then,

$$\begin{aligned}
1-\cos\left(4\tilde{\phi}\right) &= 2-2\cos^2\left(2\tilde{\phi}\right) && (31)\\
&= 2-2\cos^2(2\phi)\cos^2(2\alpha) - 2\sin^2(2\phi)\sin^2(2\alpha)\cos^2(\theta+\beta)\\
&\quad + 4\cos(2\phi)\sin(2\phi)\cos(2\alpha)\sin(2\alpha)\cos(\theta+\beta) && (32)\\
&= \frac{5}{4}-\frac{1}{4}\cos(4\phi)-\frac{1}{4}\cos(4\alpha)-\frac{1}{4}\cos(2\theta+2\beta)-\frac{3}{4}\cos(4\phi)\cos(4\alpha)\\
&\quad +\frac{1}{4}\cos(4\alpha)\cos(2\theta+2\beta)+\frac{1}{4}\cos(4\phi)\cos(2\theta+2\beta)\\
&\quad -\frac{1}{4}\cos(4\phi)\cos(4\alpha)\cos(2\theta+2\beta)+\sin(4\phi)\sin(4\alpha)\cos(\theta+\beta). && (33)
\end{aligned}$$

The first two terms are constant over α and β while the averages of all the other terms are zero for uniform α and uniform β. Therefore,

$$\mathrm{EMI}^*(\phi,\theta) \approx \mathrm{EMI}_{\min} + \frac{K_1K_2}{4}\cdot\left[5-\cos(4\phi)\right]. \qquad (34)$$

Note that a relaxed sufficient condition for Eq. (34) and (30) to hold is

$$\sum_\alpha e^{j4\alpha}=0,\quad \sum_\beta\cos(2\theta+2\beta)=0,\quad \sum_\gamma\cos(2\gamma)=0, \qquad (35)$$

which motivates the S-TVLT.

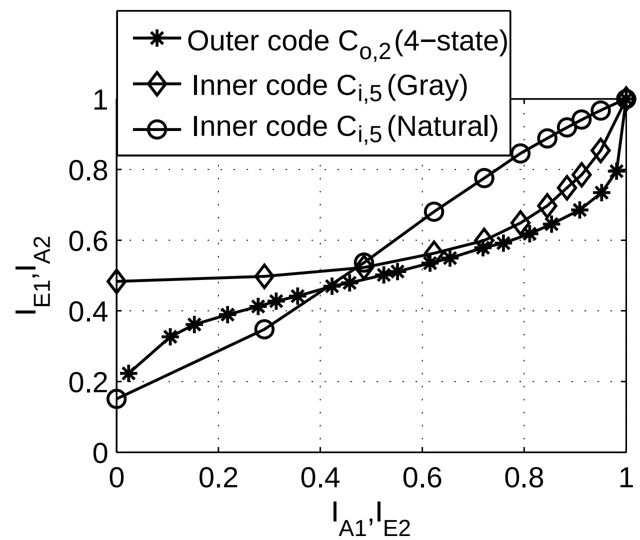

Fig. 1. EXIT chart showing the effect of constellation labeling. The inner TCM with Gray-labeled 8-PSK has a significantly higher initial extrinsic information compared to the same inner TCM with Natural-labeled 8-PSK at the same $E_s/N_o = 1.1$ dB because of better protection of systematic bits. The Gray-labeled inner TCM has already opened up the decoding tunnel, while the tunnel of the Natural-labeled inner TCM is still closed at this E_s/N_o. The latter will require the $E_s/N_o = 3.2$ dB to open the tunnel.

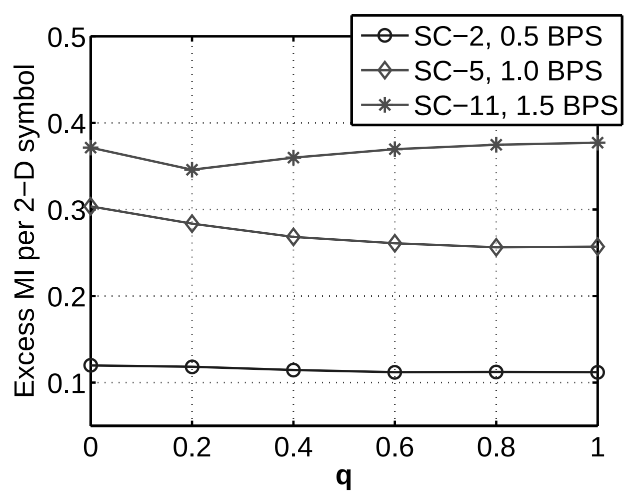

Fig. 2. Periodic fading channel performances of SC-2, SC-5, and SC-11. $\mathbf{q}$ is the attenuation factor in the normalized period-2 fading channel denoted as $[1\ \mathbf{q}]$ channel. EMI is computed for target BER of 10^{-5}.

TABLE I

CONSTITUENT CONVOLUTIONAL ENCODERS OF RATE-$\frac{k}{n}$ WITH MEMORY ν

Encoder	Crit.	k	n	ν	$G(D)$
$C_{o,1}$	ML	1	2	1	$[1+D \quad D]$
$C_{o,2}$	ML	1	2	2	$[1+D^2 \quad 1+D+D^2]$
$C_{o,3}$	ML	1	2	3	$[1+D^2+D^3 \quad 1+D+D^2+D^3]$
$C_{i,1}$	I	1	3	3	$[1 \quad \frac{D}{1+D^2} \quad 1+D]$
$C_{i,2}$	ML	1	3	3	$[1 \quad \frac{D^2}{1+D+D^2} \quad \frac{1}{1+D}]$
$C_{i,3}$	ML	2	3	2	$\begin{bmatrix} 1 & 0 & \frac{1+D^2}{1+D+D^2} \\ 0 & 1 & \frac{1+D}{1+D+D^2} \end{bmatrix}$
$C_{i,4}$	I	2	3	2	$\begin{bmatrix} 1 & 0 & \frac{D}{1+D^2} \\ 0 & 1 & \frac{D}{1+D^2} \end{bmatrix}$
$C_{i,5}$	I	2	3	2	$\begin{bmatrix} 1 & 0 & \frac{1+D}{1+D^2} \\ 0 & 1 & \frac{D}{1+D^2} \end{bmatrix}$
$C_{i,6}$	ML	2	3	2	$\begin{bmatrix} 1 & 0 & \frac{D}{1+D+D^2} \\ 0 & 1 & \frac{1+D}{1+D+D^2} \end{bmatrix}$
$C_{i,7}$	I	2	3	3	$\begin{bmatrix} 1 & 0 & \frac{1}{1+D^3} \\ 0 & 1 & \frac{D}{1+D^3} \end{bmatrix}$
$C_{i,8}$	ML	2	3	3	$\begin{bmatrix} 1 & 0 & \frac{D}{1+D^2+D^3} \\ 0 & 1 & \frac{1+D+D^2}{1+D^2+D^3} \end{bmatrix}$
$C_{i,9}$	I	3	4	2	$\begin{bmatrix} 1 & 0 & 0 & \frac{D}{1+D^2} \\ 0 & 1 & 0 & \frac{D}{1+D^2} \\ 0 & 0 & 1 & \frac{1+D}{1+D^2} \end{bmatrix}$
$C_{i,10}$	I	3	4	2	$\begin{bmatrix} 1 & 0 & 0 & \frac{D}{1+D^2} \\ 0 & 1 & 0 & \frac{D}{1+D^2} \\ 0 & 0 & 1 & \frac{D}{1+D^2} \end{bmatrix}$
$C_{i,11}$	ML	3	4	3	$\begin{bmatrix} 1 & 0 & 0 & \frac{1}{1+D+D^3} \\ 0 & 1 & 0 & \frac{D^2}{1+D+D^3} \\ 0 & 0 & 1 & \frac{1+D+D^2}{1+D+D^3} \end{bmatrix}$
$C_{i,5}^{\text{tc}}$	n/a	4	3	1	$\begin{bmatrix} 1 & 0 & \frac{1}{1+D} \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{1+D} \\ 0 & 0 & \frac{1}{1+D} \end{bmatrix}$
$C_{i,12}$	n/a	2	3	2	$\begin{bmatrix} 1 & 0 & \frac{1}{1+D^2} \\ 0 & 1 & \frac{D}{1+D^2} \end{bmatrix}$
$C_{i,12}^{\text{tc}}$	n/a	4	3	1	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{1+D} \\ 0 & 0 & \frac{1}{1+D} \\ 0 & 0 & 0 \end{bmatrix}$

TABLE II

SCTCM SCHEMES SC-1 TO SC-13 TRANSMITTING AT 0.5, 1.0 AND 1.5 BITS/SYMBOL RESULTED FROM THE SEARCHES BY EITHER ML DECODING CRITERIA OR ITERATIVE DECODING CRITERIA. THE OPERATING E_s/N_o ARE MEASURED AT BER = 10^{-5} ON THE [1 1] AND [1 0] CHANNELS TO COMPUTE THE CORRESPONDING EXCESS MI. SIMULATIONS OF BERS ARE RUN DOWN TO 10^{-7}. NONE OF THESE SCTCMS HAS ERROR FLOOR OBSERVED ON THE [1 1] CHANNEL, SO ONLY ERROR FLOORS ON THE [1 0] CHANNEL ARE REPORTED. NOTE THAT SC-8 IS ORIGINALLY A BPSK, AWGN CODE DESIGNED ACCORDING TO ML CRITERIA [20] BUT HERE IT IS MAPPED TO AN 8-PSK AS AN EXAMPLE OF NON-UNIVERSAL SCTCMS. SC-1 TO SC-13 HAVE INPUT BLOCKLENGTH EQUALS 10,000.

SC	C_o	C_i	ν^o	ν^i	BPS	Inner code Crit.	EMI(bits) at BER = 10^{-5}			E_s/N_o(dB)		Err. floor ([1 0])
							[1 1]	[1 0]	Avg.	[1 1]	[1 0]	
1	2	1	2	3	0.5	I	0.106	0.120	0.113	-2.82	1.35	10^{-5}
2	2	2	2	3	0.5	ML	0.112	0.120	0.116	-2.76	1.35	$< 10^{-7}$
3	3	1	3	3	0.5	I	0.107	0.126	0.117	-2.80	1.42	10^{-6}
4	1	5	1	2	1.0	I	0.274	N/A	N/A	1.52	N/A	10^{-3}
5	2	5	2	2	1.0	I	0.256	0.304	0.280	1.43	7.07	$< 10^{-7}$
6	2	4	2	2	1.0	I	0.256	0.284	0.270	1.42	6.92	10^{-6}
7	2	6	2	2	1.0	ML	0.252	0.442	0.347	1.41	8.05	$< 10^{-7}$
8	2	3	2	2	1.0	ML*	0.244	0.486	0.365	1.36	8.35	$< 10^{-7}$
9	2	7	2	3	1.0	I	0.267	0.330	0.299	1.48	7.26	$< 10^{-7}$
10	2	8	2	3	1.0	ML	0.264	0.485	0.375	1.47	8.35	$< 10^{-7}$
11	2	9	2	2	1.5	I	0.377	0.372	0.375	4.27	10.93	$< 10^{-7}$
12	2	10	2	2	1.5	I	0.382	0358	0.370	4.29	10.84	10^{-7}
13	2	11	2	3	1.5	ML	0.396	0.603	0.500	4.35	12.42	$< 10^{-7}$

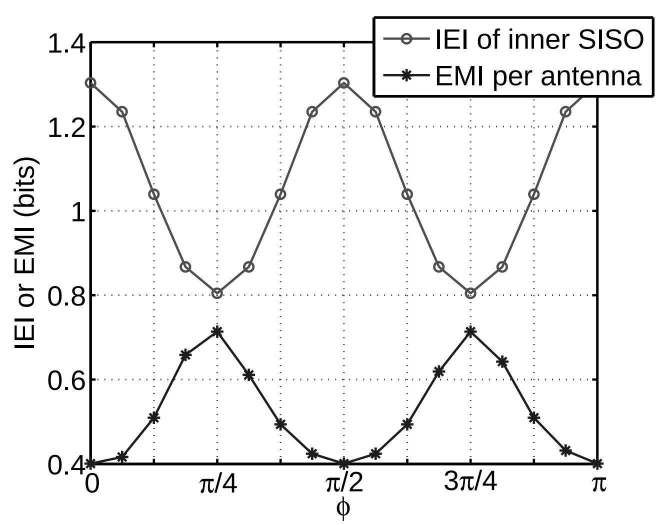

Fig. 3. The inner-SISO IEI at $E_b/N_o = 6.35$ dB and corresponding EMI per antenna for SC-11 as a function of ϕ when $\kappa = \theta = 0$. It is observed that IEI is highly correlated with EMI and thus can be used as a tool to facilitate the code search and to understand and approximate the TVLT performance.

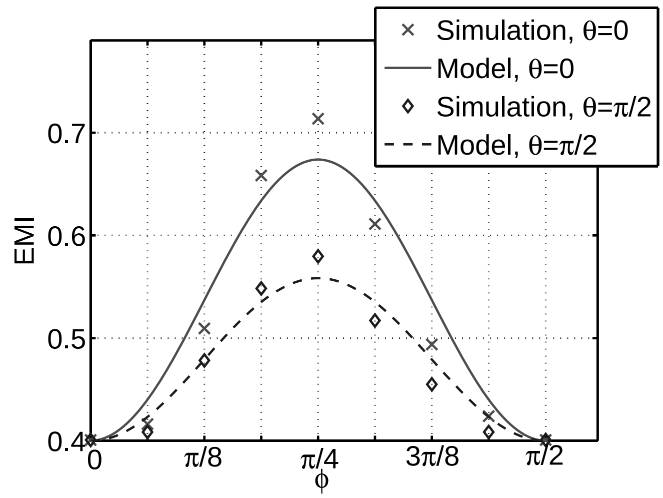

Fig. 4. Simulated EMI requirement for SC-11 on singular channels ($\kappa = 0$) as a function of ϕ when $\theta = 0$ and $\theta = \pi/2$. The proposed model for EMI in Eq. (24) with $\text{EMI}_{\text{min}} = 0.40$, $K_1 = 0.0289$ and $K_2 = 3.74$ which minimizes the mean square error corresponds well to the simulation results.

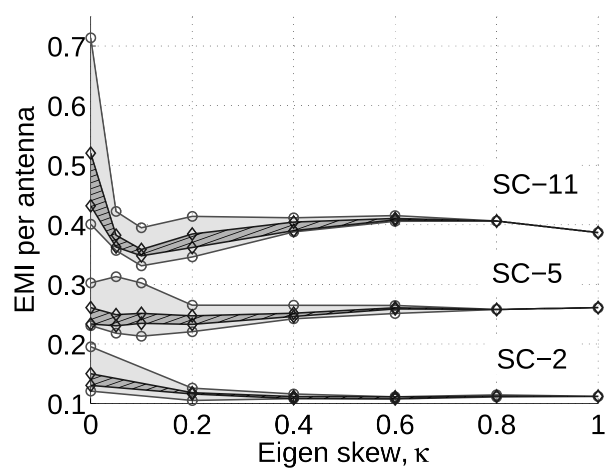

Fig. 5. EMI requirements of SC-2, SC-5 and SC-11 over eigenvalue skew and eigenvectors. The gray area is the EMI region without TVLT while the shaded area is the EMI region with F-TVLT. The proposed SCTCMs perform consistently close to channel capacity for any of the 2×2 channels.

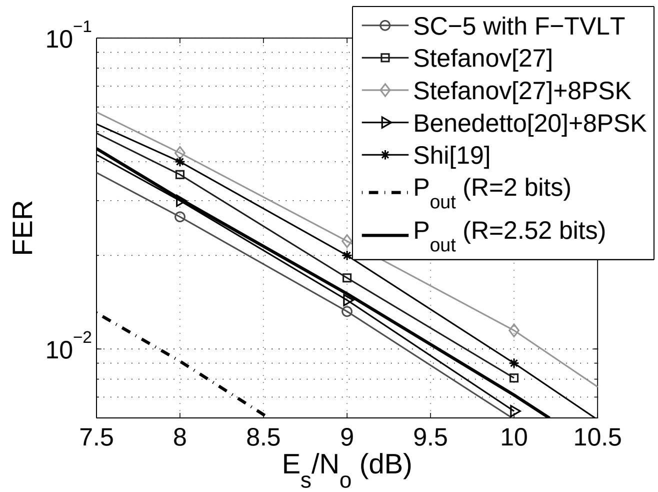

Fig. 6. FER comparison of 2 bits/s/Hz codes including the SC-5 with TVLT, two PCCC-BICMs (by Stefanov [27]), a binary SCCC (by Benedetto [20]) mapped to 8PSK (SC-8), and a PCTCM (by Shi [19]) over quasi-static Rayleigh fading channel.

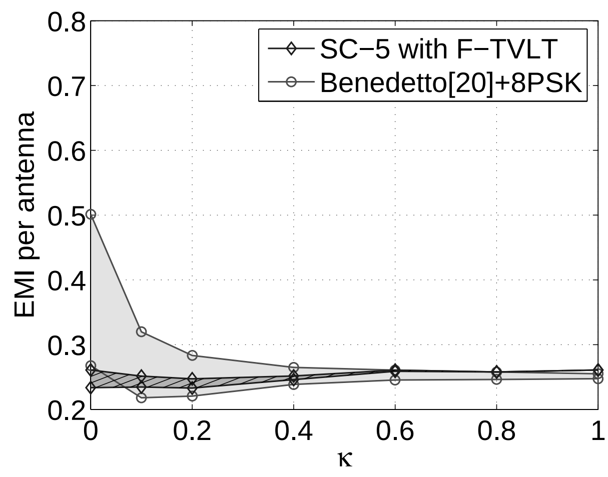

Fig. 7. EMI requirement regions for SC-5 with F-TVLT and a binary SCCC (by Benedetto [20]) mapped to 8PSK (SC-8) transmitting at 2 bits/s/Hz. SC-5 with F-TVLT provides a more consistent channel-by-channel performance than SC-8, and a slightly better (0.1 dB) Rayleigh fading performance as shown in Figure 6.

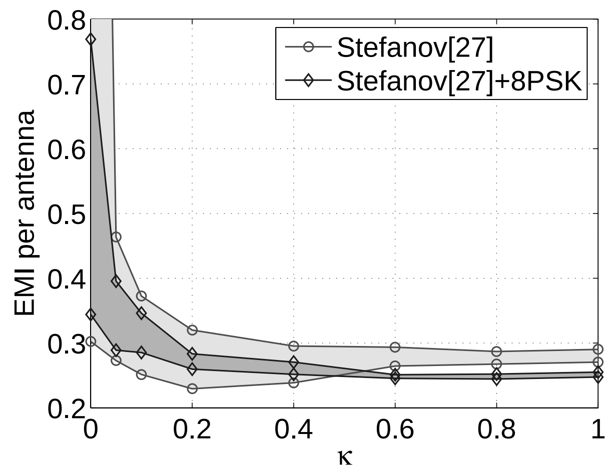

Fig. 8. EMI requirement regions for two PCCC-BICMs (by Stefanov [27]) transmitting at 2 bits/s/Hz over the 2×2 matrix channel in Eq. (13). Using a larger constellation (8PSK) improves the universality, but both codes are still not as universal as SC-5.