UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Canais Matriciais Multiplicativos sobre Corpos e Anéis Finitos com Aplicações em Codificação de Rede

Tese submetida à
Universidade Federal de Santa Catarina
como parte dos requisitos para a obtenção
do grau de Doutor em Engenharia Elétrica

Roberto Wanderley da Nóbrega

Orientador: Bartolomeu Ferreira Uchôa Filho
Co-orientador: Danilo Silva

Florianópolis, 15 de outubro de 2013.

ROBERTO WANDERLEY DA NÓBREGA

CANAIS MATRICIAIS MULTIPLICATIVOS SOBRE CORPOS E ANÉIS FINITOS COM APLICAÇÕES EM CODIFICAÇÃO DE REDE

FLORIANÓPOLIS
2013

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Nóbrega, Roberto Wanderley da
Canais matriciais multiplicativos sobre corpos e anéis finitos com aplicações em codificação de rede / Roberto Wanderley da Nóbrega ; orientador, Bartolomeu Ferreira Uchôa Filho ; co-orientador, Danilo Silva. - Florianópolis, SC, 2013.
99 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Elétrica.

Inclui referências

1. Engenharia Elétrica. 2. Anéis de cadeia finitos. 3. Capacidade de canal. 4. Codificação de rede. 5. Teoria da informação. I. Uchôa Filho, Bartolomeu Ferreira. II. Silva, Danilo. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Elétrica. IV. Título.

Roberto Wanderley da Nóbrega

CANAIS MATRICIAIS MULTIPLICATIVOS SOBRE CORPOS E ANÉIS FINITOS COM APLICAÇÕES EM CODIFICAÇÃO DE REDE

Esta Tese foi julgada adequada para a obtenção do título de Doutor em Engenharia Elétrica, área de concentração Comunicações e Processamento de Sinais, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.

Florianópolis, 15 de outubro de 2013.

Patrick Kuo Peng, Dr.
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

Banca examinadora

Prof. Bartolomeu Ferreira Uchôa Filho, Ph.D.
Universidade Federal de Santa Catarina

Prof. Danilo Silva, Ph.D.
Universidade Federal de Santa Catarina

Prof. Leonardo Silva Resende, D.Sc.
Universidade Federal de Santa Catarina

Prof. Weiler Alves Finamore, Ph.D.
Universidade Federal de Juiz de Fora

Prof.ª Sueli Irene Rodrigues Costa, Ph.D.
Universidade Estadual de Campinas

Para Melina

Agradecimentos

Desejo expressar meu reconhecimento a todos que, de uma maneira ou outra, colaboraram na realização deste trabalho, em especial

a Bartolomeu Ferreira Uchôa Filho e Danilo Silva, meus amigos e mestres, pela excelente orientação, pelo constante estímulo durante todo o doutorado, e pelos exemplos de caráter e profissionalismo que me proporcionaram ensinamentos muito mais valiosos que qualquer conhecimento técnico;

a Frank Kschischang, meu orientador durante o sanduíche em Toronto, pelos conhecimentos transmitidos, por estar sempre disponível e pelos preciosos conselhos "*about life, the academia and everything*" [46];

a Chen Feng, meu "quarto orientador", pela incrível prestatividade, pela inesgotável sabedoria compartilhada durante a minha estadia em Toronto e pelo conselho "*start with* $\mathbb{Z}_4$";

a Ricardo Bohaczuk Venturelli, pelas valiosas discussões e pelas inúmeras sugestões que contribuíram para a melhoria desta tese e dos slides da defesa;

a Leonardo Silva Resende, Mário de Noronha Neto, Sueli Irene Rodrigues Costa e Weiler Alves Finamore, pela valiosa participação na banca examinadora, seja do exame de qualificação, da tese, ou de ambos;

a Carlos Aurélio Faria da Rocha, Leonardo Silva Resende e Raimes Moraes, pelo incentivo e orientação acadêmica e pelo contínuo esforço na manutenção e melhoria da qualidade do Grupo de Pesquisa em Comunicações (GPqCom);

a Andrei Piccinini Legg, Bruno Sens Chang, Chen Feng, Diego Morschbacher, Felice Manganiello, João Fernando Refosco Baggio, João Luiz Rebelatto, Julián Jair López Salamanca, Maria Claudia de Almeida Castro, Pedro Giassi Junior, Ricardo Bohaczuk Venturelli e demais amigos e colegas do GPqCom, do GSE e do FRK Group, pela agradável convivência e pelos momentos de confraternização;

a Adaucto Wanderley da Nóbrega e Maria Luiza Amarante da Nóbrega, meus queridos pais, pelo estímulo e apoio incondicional desde a primeira hora;

a Maria Isabel Amarante da Nóbrega Wolff e Adaucto Wanderley da Nóbrega Junior, meus estimados irmãos, pelos exemplos de dedicação e caráter que sempre nortearam meu caminho;

a Alvina da Silva, pela bondade, ternura e zelo com que me tratou durante toda a minha existência;

a Drew Brigham, Fernanda de Pinho, Francisco Antônio Machado da Silva, Lucas Barcelos de Oliveira, Manssur Gustavo Cassias Pereira, Renato Herartt, pela amizade e afeição presentes em todos os momentos;

a Melina de Andrade Silveira, minha amada, pelo apoio, carinho e compreensão e por ser a alegria da minha vida;

ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo auxílio financeiro sem o qual o doutorado certamente não seria realizado;

à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo excelente apoio à pesquisa bibliográfica através do seu portal de periódicos.

Resumo da Tese apresentada à UFSC como parte dos requisitos necessários para obtenção do grau de Doutor em Engenharia Elétrica

CANAIS MATRICIAIS MULTIPLICATIVOS SOBRE CORPOS E ANÉIS FINITOS COM APLICAÇÕES EM CODIFICAÇÃO DE REDE

Roberto Wanderley da Nóbrega

15 de outubro de 2013

Orientador: Bartolomeu Ferreira Uchôa Filho
Co-orientador: Danilo Silva
Área de concentração: Comunicações e Processamento de Sinais
Palavras-chave: Anéis de cadeia finitos, Capacidade de canal, Codificação de canal, Codificação de rede, Teoria da informação
Número de páginas: xvii + 99

Um canal matricial multiplicativo (MMC) é um canal de comunicação em que a entrada $\boldsymbol{X}$ e a saída $\boldsymbol{Y}$ são matrizes relacionadas pela expressão $\boldsymbol{Y} = \boldsymbol{G}\boldsymbol{X}$, em que $\boldsymbol{G}$ é chamada de matriz de transferência. Esta tese considera MMCs sobre corpos e anéis de cadeia finitos, os quais têm aplicações práticas em codificação de rede. É adotado um enfoque probabilístico, sob a ótica da teoria da informação, de modo que o canal resultante pode ser visto como um canal discreto sem memória caracterizado essencialmente pela distribuição de probabilidade da matriz $\boldsymbol{G}$.

São abordados dois problemas na tese. Primeiramente, considera-se MMCs sobre corpos finitos, em que é assumido que as instâncias da matriz de transferência são desconhecidas tanto do transmissor quanto do receptor (isto é, o cenário não-coerente). Também é assumido que a distribuição de probabilidade da matriz $\boldsymbol{G}$ seja tal que matrizes de mesmo posto são equiprováveis. Esse modelo generaliza alguns dos considerados anteriormente na literatura. Por ser mais flexível, o modelo permite sua aplicação (no contexto de codificação de rede linear) em um maior número de situações práticas. Como contribuição, obtém-se

a capacidade do canal como um problema de otimização convexa que pode ser resolvido numericamente de maneira eficiente. Além disso, obtém-se formas fechadas para a capacidade em diversas situações especiais.

O segundo problema considera MMCs sobre anéis de cadeia finitos, dos quais corpos finitos são um caso particular. A motivação para tal estudo vem da área codificação de rede na camada física. Desta vez, é assumido que as instâncias da matriz de transferência estão disponíveis ao receptor, mas não ao transmissor (isto é, o cenário coerente). Fora isso, não é exigida nenhuma restrição sobre as estatísticas da matriz $\boldsymbol{G}$. Nesse caso, é obtida uma forma fechada para a capacidade do canal e é proposto um esquema de codificação capaz de atingir a capacidade com complexidade de tempo polinomial.

Abstract of Thesis presented to UFSC as a partial fulfillment
of the requirements for the degree of Doctor in Electrical Engineering

MULTIPLICATIVE MATRIX CHANNELS OVER FINITE FIELDS AND RINGS WITH APPLICATIONS TO NETWORK CODING

Roberto Wanderley da Nóbrega

October 15th, 2013

Advisor: Bartolomeu Ferreira Uchôa Filho
Co-advisor: Danilo Silva
Area of concentration: Communications and Signal Processing
Keywords: Channel capacity, Channel coding, Finite chain rings, Information theory, Network coding
Number of pages: xvii + 99

A multiplicative matrix channel (MMC) is a communication channel in which the input $\boldsymbol{X}$ and the output $\boldsymbol{Y}$ are matrices related by the law $\boldsymbol{Y} = \boldsymbol{G}\boldsymbol{X}$, where $\boldsymbol{G}$ is called the transfer matrix. This thesis considers MMCs over finite fields and chain rings, which have practical applications in network coding. A probabilistic approach is adopted, in the light of information theory, so that the resulting channel can be seen as a discrete memoryless channel which is essentially determined by the probability distribution of $\boldsymbol{G}$.

Two problems are examined in this thesis. First, MMCs over finite fields are considered, where it is assumed that the instances of the transfer matrix are unknown to both the transmitter and receiver (this is known as the non-coherent scenario). It is also assumed that the probability distribution of $\boldsymbol{G}$ is such that matrices with the same rank are equiprobable. This model generalizes some previously considered in the literature. Since it is more flexible, the model allows its application (in the context of linear network coding) in a larger number of practical situations. As a contribution, the channel capacity is obtained as the

solution of a convex optimization problem which can be efficiently solved by numerical methods. Furthermore, closed-form expressions for the capacity are obtained for several special situations.

The second problem considers MMCs over finite chain rings, of which finite fields are special cases. The motivation for such comes from physical-layer network coding. This time, it is assumed that the instances of the transfer matrix are available to the receiver, but not to the transmitter (this is known as the coherent scenario). Apart from that, no restrictions on the statistics of $\boldsymbol{G}$ are imposed. In this case, a closed-form expression for the channel capacity is obtained, and a polynomial time capacity-achieving coding scheme is proposed.

Sumário

Notação

Símbolo	Descrição
$\mathcal{B} \subseteq \mathcal{A}$	$\mathcal{B}$ é um subconjunto de $\mathcal{A}$.
$\mathcal{B} \subset \mathcal{A}$	$\mathcal{B}$ é um subconjunto de $\mathcal{A}$, mas $\mathcal{B} \neq \mathcal{A}$.
$\mathcal{A} \times \mathcal{B}$	Produto cartesiano do conjunto $\mathcal{A}$ com o conjunto $\mathcal{B}$.
$\mathcal{A}^n$	n-ésima potência cartesiana do conjunto $\mathcal{A}$, isto é, o conjunto de todas as n-tuplas com componentes em $\mathcal{A}$.
$\mathcal{A}^{m \times n}$	Conjunto de todas as matrizes com m linhas, n colunas e com entradas em $\mathcal{A}$.
$\lvert\mathcal{A}\rvert$	Cardinalidade (número de elementos) do conjunto $\mathcal{A}$.
$f : \mathcal{A} \to \mathcal{B}$ $x \mapsto y$	f é um mapeamento (função) com domínio $\mathcal{A}$ e contradomínio $\mathcal{B}$, que associa o elemento $x \in \mathcal{A}$ ao elemento $y \in \mathcal{B}$, isto é, $f(x) = y$.
$\operatorname{img} \phi$	Imagem do mapeamento ϕ.
$\ker \phi$	Núcleo do homomorfismo ϕ.
$\mathbb{N}$	Números naturais, incluindo o zero.
$\mathbb{Z}$	Números inteiros.
$\mathbb{R}$	Números reais.
$\mathbb{C}$	Números complexos.

Símbolo	Descrição
$\mathbb{Z}_m$	Inteiros módulo m.
$\mathbb{Z}[i]$	Inteiros gaussianos, isto é, $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$.
$\mathbb{Z}_m[i]$	Inteiros gaussianos módulo m, isto é, $\mathbb{Z}_m[i] = \{a + bi : a, b \in \mathbb{Z}_m\}$.
$\mathbb{F}_q$	Corpo finito com q elementos. Também chamado de corpo de Galois e denotado por $\mathrm{GF}(q)$.
$R[X]$	Anel polinomial na variável X sobre o anel comutativo R, isto é, o conjunto de todos os polinômios em X com coeficientes em R.
$\mathcal{G}_k(W)$	Grassmanniana k-dimensional do espaço vetorial W, isto é, o conjunto de todos os subespaços k-dimensionais de W. Ver p. 23.
$\mathcal{P}(W, d)$	Conjunto de todos os subespaços com dimensão no máximo d do espaço vetorial W. Ver p. 24.
$\mathcal{T}_r(\mathbb{F}_q^{m\times n})$	Subconjunto de $\mathbb{F}_q^{m\times n}$ consistindo de todas as matrizes de posto r. Ver p. 24.
$\mathcal{T}(\mathbb{F}_q^{m\times n})$	Subconjunto de $\mathbb{F}_q^{m\times n}$ consistindo de todas as matrizes de posto $\min\{n, m\}$ (isto é, posto completo). Ver p. 24.
$\mathrm{GL}_n(R)$	Grupo linear geral de grau n sobre o anel comutativo R, isto é, o conjunto de todas as matrizes inversíveis em $R^{n\times n}$.
$\begin{bmatrix} m \\ k \end{bmatrix}_q$	Coeficiente binomial gaussiano. Também chamado de coeficiente q-binomial. Ver p. 24.
$\operatorname{row} A$	Espaço (ou módulo) gerado pelas linhas da matriz A.
$\operatorname{col} A$	Espaço (ou módulo) gerado pelas colunas da matriz A.
$\dim V$	Dimensão do espaço vetorial V.
$\operatorname{rank} A$	Posto da matriz A.
$\operatorname{shape} M$	*Shape* do R-módulo M. Ver p. 60.
$\operatorname{shape} A$	*Shape* da matriz A. Ver p. 62.
$\langle A \rangle$	O mesmo que $\operatorname{row} A$, se A for uma matriz.
$\langle r \rangle$	Ideal gerado pelo elemento $r \in R$ de um anel comutativo R.

Símbolo	Descrição
$\Pr[A]$	Probabilidade do evento A.
$p_{\boldsymbol{x}}(x)$	Probabilidade de a variável aleatória $\boldsymbol{x}$ assumir o valor x.
$p_{\boldsymbol{x}\|\boldsymbol{y}}(x\|y)$	Probabilidade de a variável aleatória $\boldsymbol{x}$ assumir o valor x dado que $\boldsymbol{y}$ assumiu o valor y.
$\mathrm{E}[\boldsymbol{x}]$	Valor esperado da variável aleatória $\boldsymbol{x}$.
$H(\boldsymbol{x})$	Entropia da variável aleatória $\boldsymbol{x}$.
$H(\boldsymbol{x}\|\boldsymbol{y})$	Entropia condicionada da variável aleatória $\boldsymbol{x}$, dado a variável aleatória $\boldsymbol{y}$.
$I(\boldsymbol{x};\boldsymbol{y})$	Informação mútua entre as variáveis aleatórias $\boldsymbol{x}$ e $\boldsymbol{y}$.
$1[P]$	Função indicadora da proposição P, isto é, $1[P] = 1$ se P é verdadeiro, caso contrário, $1[P] = 0$.
$\lfloor x \rfloor$	Função chão de $x \in \mathbb{R}$ (maior inteiro menor ou igual a x).
$\lceil x \rceil$	Função teto de $x \in \mathbb{R}$ (menor inteiro maior ou igual a x).

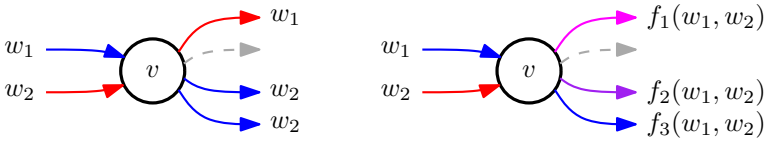
CAPÍTULO 1

Introdução

1.1 Codificação de rede

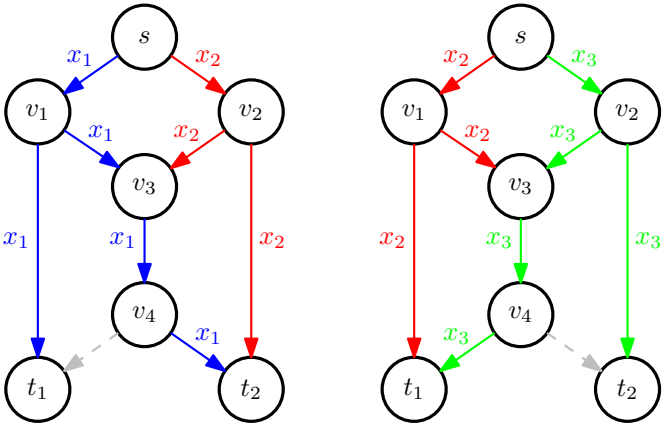
Tradicionalmente, as redes de comunicação adotam a técnica de *roteamento*, que consiste na escolha de caminhos adequados para a informação que trafega pela rede. Nesse esquema, cada nó da rede funciona como um *comutador*: apenas seleciona, replica e repassa o que recebe (Figura 1.1a). Em outras palavras, os dados são considerados unidades atômicas imutáveis.

A área de *codificação de rede* (do inglês *network coding*), introduzida por Ahlswede, Cai, Li e Yeung [2], confronta esse paradigma. Ao invés de limitar a operação dos nós à função de um comutador, nessa



(a) Nó operando como comutador. (b) Nó operando como codificador.

Figura 1.1: Modos de operação de um nó da rede.



(a) Primeiro instante.

(b) Segundo instante.

Figura 1.2: Multidifusão na rede borboleta utilizando roteamento.

nova proposta é permitido que cada nó opere plenamente como um *codificador*: os dados na saída podem ser uma combinação arbitrária dos dados na entrada do nó (Figura 1.1b). Assim, ocorre o processamento da informação que flui pela rede. Uma vez que todo codificador pode operar como um comutador, roteamento é apenas um caso particular de codificação de rede.

Para justificar o uso de codificação de rede, considere o "exemplo canônico" da literatura: multidifusão na *rede borboleta*. A rede borboleta, introduzida em [2], é mostrada nas Figuras 1.2 e 1.3. Nela, cada enlace é capaz de transmitir um único pacote por unidade de tempo, instantaneamente e livre de erros. Um pacote é uma sequência de símbolos (bits, por exemplo). O objetivo é a transmissão de informação do nó fonte s para os nós destino t_1 e t_2.

Se apenas roteamento é permitido, tal tarefa seria executada como na Figura 1.2, em que os recursos da rede são compartilhados no tempo. No primeiro instante, o nó t_1 recebe apenas o pacote x_1, enquanto o nó t_2 recebe os pacotes x_1 e x_2. No segundo instante, o nó t_1 recebe os pacotes x_2 e x_3, enquanto o nó t_2 recebe apenas o pacote x_3. O resultado é a transmissão de três pacotes em dois instantes de tempo, isto é, uma taxa de 1,5 pacotes por instante de tempo.

Em contraste, utilizando codificação de rede, é possível transmitir

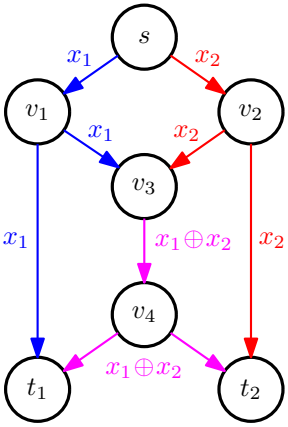


Figura 1.3: Multidifusão na rede borboleta utilizando codificação de rede.

2 pacotes por instante de tempo, de acordo com a Figura 1.3. O nó v_3 calcula e transmite a soma módulo 2 (equivalente à operação de XOR) bit-a-bit dos pacotes que recebe. O nó t_1, que recebe os pacotes x_1 e $x_1 \oplus x_2$, consegue decodificar x_2 calculando $x_1 \oplus (x_1 \oplus x_2) = x_2$. Analogamente, o nó t_2 também é capaz de decodificar x_1 e x_2. Assim, existe um claro benefício, em termos de taxa de transmissão, quando se permite o processamento de informação pelos nós intermediários da rede, justificando o uso de codificação de rede.

Em geral, as operações efetuadas pelos nós da rede podem ser quaisquer. No entanto, como mostrado por Li, Yeung e Cai [29] e posteriormente por Kötter e Médard [27], no caso de um único nó fonte, é suficiente restringir as operações sobre os pacotes a *combinações lineares*, em que cada pacote é um vetor de comprimento ℓ com entradas em um dado *corpo finito* $\mathbb{F}_q$. Nesse caso, chamado de *codificação de rede linear*, é possível mostrar que a *comunicação fim-a-fim* entre um nó fonte e um nó destino da rede pode ser modelada pela expressão

$$Y = GX, \tag{1.1}$$

em que $X \in \mathbb{F}_q^{n \times \ell}$ é a matriz cujas linhas são os n pacotes injetados na rede pelo nó fonte, $Y \in \mathbb{F}_q^{m \times \ell}$ é a matriz cujas linhas são os m pacotes coletados da rede pelo nó destino e $G \in \mathbb{F}_q^{m \times n}$ é a *matriz de transferência* de X para Y, cujas entradas dependem da topologia da

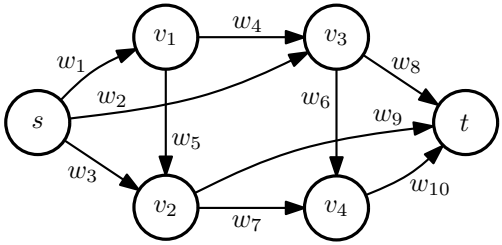


Figura 1.4: Topologia de uma rede simples.

rede e das combinações lineares efetuadas pelos nós intermediários[1].

Exemplo. Considere a rede ilustrada na Figura 1.4. Suponha que a rede opere de acordo com codificação de rede linear sobre um corpo finito $\mathbb{F}_q$. Deseja-se obter a relação entre os n pacotes enviados pelo nó fonte s e os m pacotes recebidos pelo nó destino t, em que, nesse exemplo, $n = m = 3$. Sejam $w_1, w_2, \ldots, w_9, w_{10} \in \mathbb{F}_q^\ell$ os pacotes transportados pelos enlaces da rede, conforme indicado na figura. Então, tem-se

$$\begin{aligned}
w_4 &= a_{14}w_1, \\
w_5 &= a_{15}w_1, \\
w_6 &= a_{26}w_2 + a_{46}w_4, \\
w_7 &= a_{37}w_3 + a_{57}w_5, \\
w_8 &= a_{28}w_2 + a_{48}w_4, \\
w_9 &= a_{39}w_3 + a_{59}w_5, \\
w_{10} &= a_{60}w_6 + a_{70}w_7,
\end{aligned}$$

em que $a_{ij} \in \mathbb{F}_q$ são os coeficientes das combinações lineares efetuadas pelos nós intermediários. Fazendo $x_1 = w_1$, $x_2 = w_2$ e $x_3 = w_3$, além de $y_1 = w_8$, $y_2 = w_9$ e $y_3 = w_{10}$, e resolvendo recursivamente as equações

[1] No caso em que erros de enlace estão presentes, a expressão em (1.1) é substituída por $Y = GX + HZ$, em que Z é a matriz cujas linhas são os pacotes de erro e H é a matriz de transferência de Z para Y. O presente trabalho, no entanto, desconsidera essa situação, assumindo que qualquer pacote errado é descartado pelos nós através do uso de códigos detectores de erro suficientemente poderosos.

acima, obtém-se

$$\begin{aligned} y_1 &= a_{14}a_{48}x_1 + a_{28}x_2, \\ y_2 &= a_{15}a_{59}x_1 + a_{39}x_3, \\ y_3 &= (a_{14}a_{46}a_{60} + a_{15}a_{57}a_{70})x_1 + a_{26}a_{60}x_2 + a_{37}a_{70}x_3, \end{aligned}$$

ou, alternativamente, $Y = GX$, em que

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

e

$$G = \begin{bmatrix} a_{14}a_{48} & a_{28} & 0 \\ a_{15}a_{59} & 0 & a_{39} \\ a_{14}a_{46}a_{60} + a_{15}a_{57}a_{70} & a_{26}a_{60} & a_{37}a_{70} \end{bmatrix},$$

em que G é a matriz de transferência de X para Y. □

A área de codificação de rede recebeu especial atenção após o trabalho de Ho et al. [18, 19] que sugere a ideia de *codificação de rede linear aleatória*, na qual os coeficientes das combinações lineares são escolhidos aleatoriamente. É mostrado que, no caso de um único nó fonte, o método atinge o desempenho ótimo com alta probabilidade, desde que seja empregado um corpo finito de tamanho q suficientemente grande. Com isso, tornou-se possível um funcionamento distribuído do sistema, uma vez que foi eliminada a necessidade do conhecimento prévio da topologia da rede e das combinações lineares efetuadas. Posteriormente, Chou, Wu e Jain [5] propõem um protocolo prático (utilizando o conceito de *gerações*) que implementa codificação de rede linear aleatória. Como consequência, a expressão em (1.1), cuja validade era antes limitada a redes com perfeito sincronismo e topologia fixa, passa agora a ser aplicável em sistemas sujeitos a problemas de sincronismo, perda e atraso de pacotes, congestionamento, entrada e saída de nós, etc.

Note que os pacotes transmitidos (isto é, a matriz X) podem ser recuperados a partir dos pacotes recebidos (isto é, a matriz Y) se a matriz de transferência G tiver posto n. Para tanto, é necessário, em princípio, o conhecimento da matriz G no nó destino. No caso de

codificação de rede linear aleatória, a matriz de transferência G, que depende dos coeficientes das combinações lineares, é desconhecida a priori no destino. Em [5, 18, 19], esse problema é resolvido através do uso de um *cabeçalho* em cada pacote transmitido pelo nó fonte. Mais precisamente, cada pacote enviado pela fonte é prefixado com um vetor da base canônica de $\mathbb{F}_q^n$, de modo que

$$X = \begin{bmatrix} I_{n\times n} & X' \end{bmatrix},$$

em que $I_{n\times n}$ é a matriz identidade $n \times n$ e $X' \in \mathbb{F}_q^{n\times(\ell-n)}$ é a matriz contendo de fato a informação. Assim, tem-se

$$Y = GX = G\begin{bmatrix} I_{n\times n} & X' \end{bmatrix} = \begin{bmatrix} G & GX' \end{bmatrix},$$

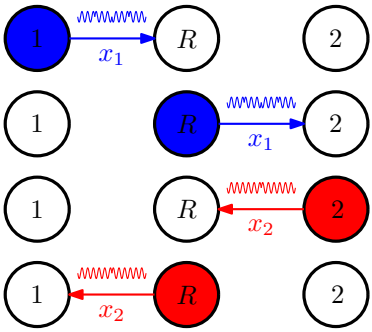
de modo que o nó destino passa a ter conhecimento de G e GX' sendo, portanto, capaz de recuperar X', desde que G tenha posto n. O preço a ser pago é uma sobrecarga (*overhead*) de n^2 símbolos, devido ao cabeçalho; tal sobrecarga, no entanto, pode ser tornada desprezível se for permitido aumentar ℓ (o comprimento do pacote) arbitrariamente.

Para mais detalhes sobre os fundamentos e benefícios de codificação de rede, encaminha-se o leitor para os livros de Yeung et al. [58] e Fragouli e Soljanin [16, 15].

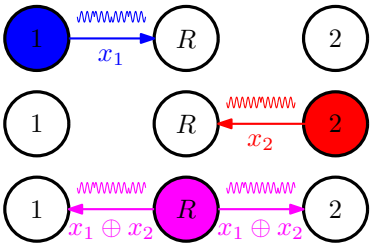
1.2 Codificação de rede na camada física

As técnicas de codificação de rede foram inicialmente concebidas para serem aplicadas em camadas superiores da rede (camada de rede, camada de aplicação). Em particular, no caso de redes sem-fio, a questão do compartilhamento do meio físico seria tratada através de multiplexação (tipicamente no tempo ou na frequência), criando efetivamente enlaces ortogonais e evitando-se, assim, a interferência dos sinais físicos. A área de *codificação de rede na camada física*, proposta independentemente por Popovski e Yomo [44], Nazer e Gastpar [34] e Zhang, Liew e Lam [60] (para *surveys* abrangentes da área, veja [36] e [30]), em contraste, toma vantagem da interferência inerente às redes sem-fio, ao invés de evitá-la.

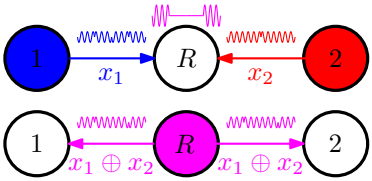
Para efeito de ilustração, considere a rede sem-fio bidirecional com



(a) Roteamento.



(b) Codificação de rede tradicional.

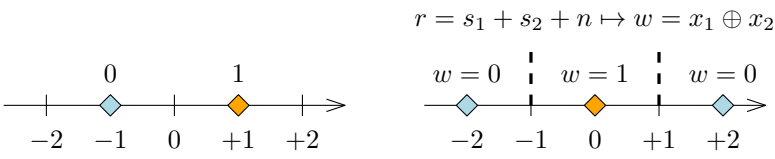


(c) Codificação de rede na camada física.

Figura 1.5: Três soluções para a rede sem-fio bidirecional com nó intermediário.

nó intermediário mostrada na Figura 1.5. Suponha que os nós terminais 1 e 2 desejam trocar pacotes entre si, com auxílio de um nó intermediário R. A solução utilizando roteamento é apresentada na Figura 1.5a, em que os pacotes são trocadas em 4 instantes de tempo. A solução através de codificação de rede tradicional é mostrada na Figura 1.5b. Note a semelhança com a solução apresentada anteriormente na Figura 1.3. Note também que, nesse caso, o sistema *evita* a interferência dos sinais: no primeiro instante de tempo, o nó 2 não transmite sinal algum enquanto o nó 1 envia sua mensagem e analogamente no instante seguinte. Com isso, os pacotes são trocadas em 3 instantes de tempo, uma melhora em relação ao caso anterior (roteamento).

A solução através de codificação de rede na camada física é apresentada na Figura 1.5c. Nela, no primeiro instante de tempo, os nós 1 e 2 transmitem *simultaneamente* suas mensagens. Cabe ao nó R *inferir* uma combinação linear de tais mensagens (no exemplo, o XOR



(a) Constelação BPSK. (b) $(h_1, h_2) = (1, 1)$.

Figura 1.6: Codificação de rede na camada física com modulação BPSK.

das mensagens) com base no sinal físico recebido (a "soma no ar" das mensagens enviadas, sujeita a todas as adversidades do canal de comunicação sem-fio). É importante ressaltar que, nesse caso, o nó intermediário *não* tem conhecimento individual dos pacotes x_1 e x_2 (como é no caso da codificação de rede tradicional), mas apenas da soma $x_1 \oplus x_2$. Com esse esquema, as mensagens são trocadas em apenas 2 instantes de tempo, melhorando ainda mais o desempenho em termos de taxa de transmissão.

Mais detalhadamente (veja a Figura 1.6), assuma que o sistema opera com modulação BPSK e considere um modelo de desvanecimento plano em bloco com perfeito conhecimento dos coeficientes do canal sem-fio na recepção dos sinais [17]. Seja $\phi : \mathbb{Z}_2 \to \mathbb{R}$ um mapeamento de bits para pontos da constelação BPSK (na Figura 1.6a, tem-se $\phi(0) = -1$ e $\phi(1) = +1$). No primeiro instante de tempo, o nó 1 modula o pacote $x_1 \in \mathbb{Z}_2^\ell$ em uma sequência $s_1 \in \mathbb{R}^\ell$ de pontos da constelação (aplicando ϕ entrada-a-entrada). Analogamente, o nó 2 modula $x_2 \in \mathbb{Z}_2^\ell$ em $s_2 \in \mathbb{R}^\ell$. Em seguida, ambos os nós transmitem simultaneamente seus sinais, de modo que o sinal recebido pelo nó R será

$$r = h_1 s_1 + h_2 s_2 + n,$$

em que $h_1, h_2 \in \mathbb{R}$ são os coeficientes de desvanecimento e $n \in \mathbb{R}^\ell$ é o vetor de ruído, aqui assumido gaussiano com entradas i.i.d.

No segundo instante de tempo, o nó intermediário, utilizando apenas r, h_1 e h_2, tenta extrair a soma módulo 2 dos pacotes de informação, isto é, o vetor $w = x_1 \oplus x_2 \in \mathbb{Z}_2^\ell$. Supondo que os coeficientes de desvanecimento sejam dados por $h_1 = h_2 = 1$, o sinal recebido $r = s_1 + s_2 + n$ se situará em torno dos pontos -2 e $+2$, como mostrado na Figura 1.6b,

e o procedimento ótimo de decisão se dá como segue[2] (veja as regiões de decisão indicadas na figura):

$$w = x_1 \oplus x_2 = \begin{cases} 0, & \text{se } |r| > 1, \\ 1, & \text{caso contrário.} \end{cases}$$

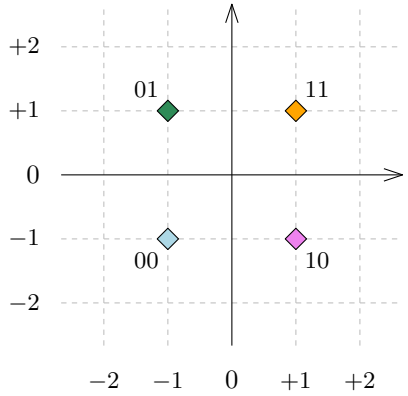
Em seguida, o nó R mapeia $w \in \mathbb{Z}_2^\ell$ em uma sequência de pontos da constelação que, após ser transmitida, é recebida e demodulada pelos nós 1 e 2. Com x_1 e $w = x_1 \oplus x_2$, o nó 1 pode agora obter x_2, como de costume; similarmente, o nó 2 obtém x_1.

Exemplo. Sejam $\ell = 4$, $x_1 = (0, 1, 0, 1)$ e $x_2 = (0, 0, 1, 1)$. Portanto, tem-se $s_1 = (-1, +1, -1, +1)$ e $s_2 = (-1, -1, +1, +1)$, de modo que, se $h_1 = h_2 = 1$, então $r = s_1 + s_2 + n = (-2, 0, 0, +2) + n$. Assim, se o ruído n não for suficientemente severo, o nó R será capaz de extrair $w = x_1 \oplus x_2 = (0, 1, 1, 0)$ a partir de r, de acordo com o procedimento de decisão fornecido acima. □

No entanto, essa técnica esbarra em um obstáculo quando estendida para outros tipos de modulações digitais. O seguinte exemplo, que ilustra o problema, foi introduzido por Feng, Silva e Kschischang [12]. Suponha que seja empregada a modulação complexa QPSK, como ilustrado na Figura 1.7. Uma vez que tal constelação é quaternária, cada pacote deve ser agora uma sequência de ℓ símbolos de um alfabeto contendo 4 elementos. Uma escolha natural para esse alfabeto é $\mathbb{Z}_2 \times \mathbb{Z}_2 = \{00, 01, 10, 11\}$, o conjunto de todos os pares de bits, em que as operações de soma e multiplicação são efetuadas entrada-a-entrada. Um mapeamento $\phi : \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{C}$ de símbolos do alfabeto para pontos da constelação é mostrado na Figura 1.7a.

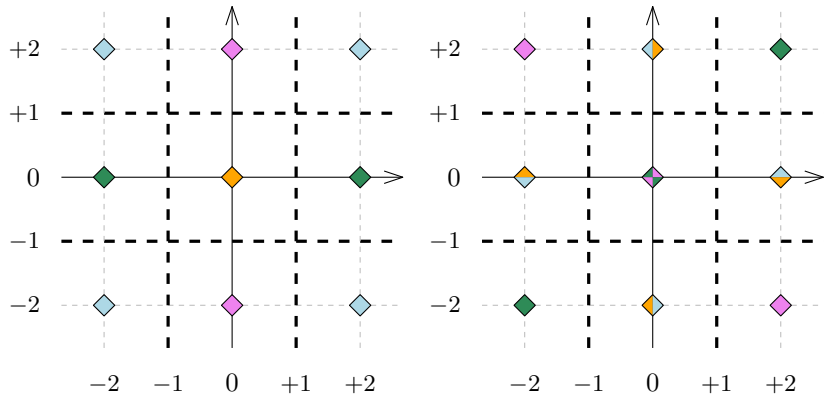
No caso em que os coeficientes de desvanecimento são dados por $h_1 = h_2 = 1$, o sinal recebido $r = s_1 + s_2 + n$ ficará em torno dos pontos mostrados na Figura 1.7b. Nesse situação, o nó intermediário ainda é capaz de extrair a soma $w = x_1 \oplus x_2$ a partir de r. No entanto, no caso em que $h_1 = 1$ e $h_2 = i$, isso se torna impossível. De fato, como mostra a Figura 1.7c, a rotação de fase efetuada pelo canal de comunicação sem-fio gera *ambiguidade* para o nó intermediário.

[2] Para o procedimento ótimo de decisão no caso em que h_1 e h_2 são quaisquer, veja [35] ou [12].



(a) Constelação QPSK rotulada com elementos de $\mathbb{Z}_2 \times \mathbb{Z}_2$.

$r = s_1 + s_2 + n \mapsto w = x_1 \oplus x_2$ $r = s_1 + is_2 + n \mapsto w = x_1 \oplus x_2$



(b) $(h_1, h_2) = (1, 1)$. (c) $(h_1, h_2) = (1, i)$.

Figura 1.7: Codificação de rede na camada física com modulação QPSK considerando o anel $\mathbb{Z}_2 \times \mathbb{Z}_2$.

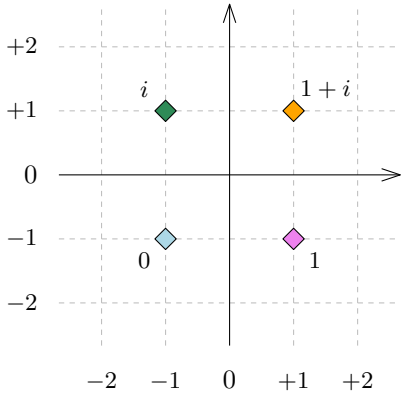
EXEMPLO. Seja $\ell = 1$. Existem precisamente dois pares (x_1, x_2) que fornecem $s_1 + s_2 = 2i$, a saber, $(x_1, x_2) = (01, 11)$ ou $(11, 01)$. Em ambos os casos, tem-se $x_1 \oplus x_2 = 10$, de modo que, para $h_1 = h_2 = 1$, o nó intermediário pode inferir com segurança que $x_1 \oplus x_2 = 10$ sempre que o sinal recebido $r = s_1 + s_2 + n$ se situar nas proximidades de $2i$. Também existem precisamente dois pares (x_1, x_2) que fornecem $s_1 + is_2 = 2i$, a saber, $(x_1, x_2) = (01, 10)$ ou $(11, 11)$. Nesse caso, em contraste, $x_1 \oplus x_2 = 11$ para o primeiro par e $x_1 \oplus x_2 = 00$ para o segundo. Portanto, para $h_1 = 1$ e $h_2 = i$, o nó intermediário ficará em dúvida quanto ao valor de $x_1 \oplus x_2$ caso o sinal recebido $r = s_1 + is_2 + n$ se situe nas proximidades de $2i$. A Figura 1.7 mostra a situação geral. □

Pode-se mostrar que o problema permanece mesmo com outras escolhas para o mapeamento $\phi : \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{C}$. Além disso, a mudança do alfabeto $\mathbb{Z}_2 \times \mathbb{Z}_2$ para o corpo finito $\mathbb{F}_4 = \{0, 1, \alpha, 1 + \alpha\}$, em que $\alpha^2 = 1 + \alpha$, também não resolve o problema, mesmo se forem permitidas combinações lineares mais gerais (por exemplo, $w = x_1 + \alpha x_2$ ao invés de $w = x_1 + x_2$). A solução está no uso de uma terceira opção: $\mathbb{Z}_2[i] = \{0, 1, i, 1 + i\}$, em que $i^2 = 1$ (note que $i \in \mathbb{Z}_2[i]$, apesar de relacionado, é diferente de $i \in \mathbb{C}$). A Figura 1.8 mostra os detalhes da solução. Ressalta-se que, no caso em que $h_1 = 1$ e $h_2 = i$, a combinação linear extraída pelo nó intermediário é $w = x_1 + ix_2$. Isso não apresenta nenhum problema na hora da recuperação dos pacotes por parte dos nós 1 e 2: por exemplo, o nó 1 pode recuperar x_2 a partir de x_1 e w através de $i(x_1 + w) = i(x_1 + x_1 + ix_2) = x_2$.

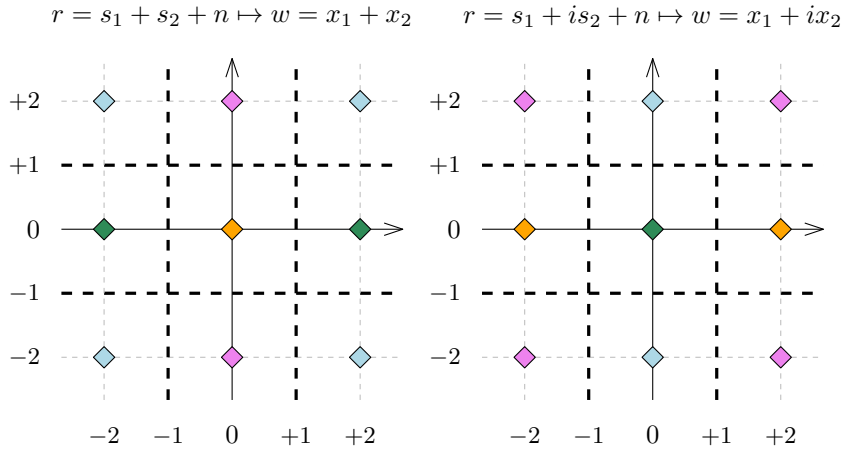
1.3 Codificação de rede sobre anéis finitos

O objeto matemático $\mathbb{Z}_2[i]$ mencionado acima é um exemplo daquilo que a disciplina de álgebra abstrata chama de *anel*. Um anel (comutativo) é uma estrutura algébrica com duas operações, adição e multiplicação, satisfazendo certas propriedades básicas[3] (a saber: associatividade e comutatividade da adição e da multiplicação; distributividade da multiplicação sobre a adição; existência de identidade aditiva, "0", e multiplicativa, "1"; e existência de elementos opostos, isto é, inversos aditivos). Todo corpo é um anel comutativo. No entanto, ao contrário

[3] Veja o Apêndice A.1 para um tratamento mais preciso de anéis.



(a) Constelação QPSK rotulada com elementos de $\mathbb{Z}_2[i]$.



(b) $(h_1, h_2) = (1, 1)$. (c) $(h_1, h_2) = (1, i)$.

Figura 1.8: Codificação de rede na camada física com modulação QPSK considerando o anel $\mathbb{Z}_2[i]$.

do que ocorre em corpos, elementos não-nulos do anel não necessariamente possuem inversos multiplicativos e, além disso, a multiplicação de dois elementos não-nulos não necessariamente fornece um elemento diferente de zero. Exemplos de anéis com um número infinito de elementos incluem os números reais ($\mathbb{R}$), os números complexos ($\mathbb{C}$) e os números inteiros ($\mathbb{Z}$). Exemplos de anéis finitos (isto é, com um número finito de elementos) incluem os corpos finitos ($\mathbb{F}_q$) e os inteiros com aritmética modular ($\mathbb{Z}_m$).

Tradicionalmente, codificação de rede linear tem sido considerada quase sempre sobre corpos finitos. A necessidade de estruturas algébricas mais gerais era vista, até pouco tempo atrás, como uma mera generalização matemática, sendo desinteressante do ponto de vista da engenharia. No entanto, essa situação viria a mudar após a introdução da área de codificação de rede na camada física. De fato, não é difícil se convencer que as técnicas descritas na seção anterior podem ser estendidas para topologias de rede sem-fio mais gerais. Nesse contexto, cada nó da rede infere combinações lineares a partir dos sinais físicos recebidos por sua antena e transmite combinações lineares de combinações lineares previamente obtidas. Assim, pela linearidade, a comunicação fim-a-fim entre dois nós da rede ainda pode ser modelada pela expressão em (1.1). A diferença é que, como discutido acima, o alfabeto envolvido não é necessariamente um corpo finito, mas sim um anel finito (que, em geral, depende da modulação empregada pelo sistema).

Exemplo. Considere a Figura 1.9, que mostra uma rede sem-fio em camadas. Suponha que a rede opere de acordo com codificação de rede na camada física, com os pacotes da camada superior sobre um dado anel finito R. Sejam $w_1, w_2, w_3 \in R^\ell$ os $n = 3$ pacotes enviados pelo nó fonte s e $w_7, w_8, w_9 \in R^\ell$ os $m = 3$ pacotes recebidos pelo nó destino t. Sejam $s_1, s_2, \ldots, s_6 \in \mathbb{C}^\ell$ os sinais físicos (sequências de pontos da constelação) transmitidos pelos nós $v_1, v_2, \ldots, v_6$, respectivamente, e $r_4, r_5, \ldots, r_9 \in \mathbb{C}^\ell$ os sinais físicos recebidos pelos nós $v_4, v_5, \ldots, v_9$, conforme a figura. Note que, nesse exemplo, por simplicidade, os nós v_1, v_2, v_3 não recebem sinais físicos do nó s, mas sim os pacotes w_1, w_2, w_3, respectivamente, vindos diretamente da camada superior; analogamente, os nós v_7, v_8, v_9 não transmitem sinais físicos para o nó t, mas sim os pacotes w_7, w_8, w_9 pela camada superior.

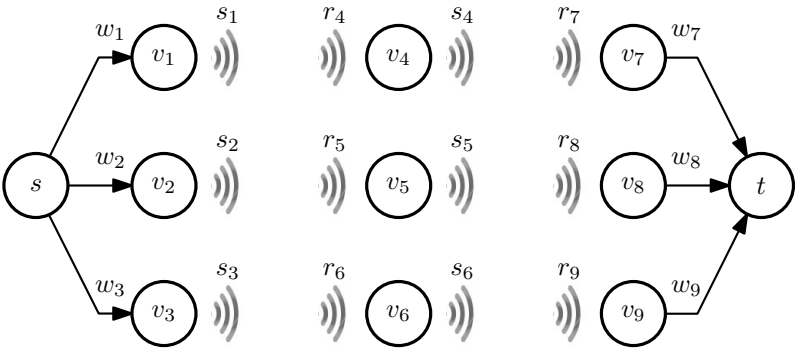


Figura 1.9: Rede sem-fio em camadas.

O funcionamento do sistema é como segue. Os nós v_1, v_2, v_3 iniciam modulando os pacotes $w_1, w_2, w_3 \in R^\ell$ nos sinais $s_1, s_2, s_3 \in \mathbb{C}^\ell$, respectivamente, contendo pontos da constelação. Os sinais s_1, s_2, s_3, são então transmitidos simultaneamente, sendo superpostos no meio físico. Portanto, os sinais recebidos pelos nós v_4, v_5, v_6 são, respectivamente,

$$\begin{aligned} r_4 &= h_{14}s_1 + h_{24}s_2 + h_{34}s_3 + n_4, \\ r_5 &= h_{15}s_1 + h_{25}s_2 + h_{35}s_3 + n_5, \\ r_6 &= h_{16}s_1 + h_{26}s_2 + h_{36}s_3 + n_6, \end{aligned}$$

em que $h_{ij} \in \mathbb{C}$ são coeficientes de desvanecimento e $n_i \in \mathbb{C}^\ell$ são vetores de ruído. A partir de r_i, h_{1i}, h_{2i} e h_{3i}, utilizando os princípios da codificação de rede na camada física, o nó v_i (para $i = 4, 5, 6$) é capaz de inferir uma combinação linear $w_i \in R^\ell$ dos pacotes w_1, w_2, w_3, de modo que

$$\begin{aligned} w_4 &= a_{14}w_1 + a_{24}w_2 + a_{34}w_3, \\ w_5 &= a_{15}w_1 + a_{25}w_2 + a_{35}w_3, \\ w_6 &= a_{16}w_1 + a_{26}w_2 + a_{36}w_3, \end{aligned}$$

em que $a_{ij} \in R$. Os pacotes $w_4, w_5, w_6 \in R^\ell$ são então modulados nos sinais $s_4, s_5, s_6 \in \mathbb{C}^\ell$, que por sua vez são transmitidos pelo meio físico. Os nós v_7, v_8, v_9, a partir dos sinais recebidos $r_7, r_8, r_9 \in \mathbb{C}^\ell$,

respectivamente, são capazes de extrair combinações lineares

$$w_7 = a_{47}w_4 + a_{57}w_5 + a_{67}w_6,$$
$$w_8 = a_{48}w_4 + a_{58}w_5 + a_{68}w_6,$$
$$w_9 = a_{49}w_4 + a_{59}w_5 + a_{69}w_6,$$

que são finalmente entregues ao nó destino t.

Portanto, definindo $x_1 = w_1$, $x_2 = w_2$ e $x_3 = w_3$, além de $y_1 = w_7$, $y_2 = w_8$ e $y_3 = w_9$, tem-se $Y = GX$, em que

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in R^{n\times\ell}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in R^{m\times\ell}$$

e

$$G = \begin{bmatrix} a_{47} & a_{57} & a_{67} \\ a_{48} & a_{58} & a_{68} \\ a_{49} & a_{59} & a_{69} \end{bmatrix} \begin{bmatrix} a_{14} & a_{24} & a_{34} \\ a_{15} & a_{25} & a_{35} \\ a_{16} & a_{26} & a_{36} \end{bmatrix},$$

em que $G \in R^{m\times n}$ é a matriz de transferência. □

1.4 Canais matriciais multiplicativos

Como visto até agora, a expressão em (1.1) é capaz de modelar a comunicação fim-a-fim entre dois nós de um sistema que opera com codificação de rede (possivelmente na camada física), sob hipóteses bastante práticas (por exemplo, a topologia pode ser variante no tempo, o funcionamento da rede pode ser assíncrono e os enlaces da rede podem estar sujeitos a apagamentos). O presente trabalho considera tal expressão como caracterizadora de um legítimo canal de comunicação, o chamado *canal matricial multiplicativo* (MMC, do inglês *multiplicative matrix channel*), também denominado por alguns autores de *canal operador linear* (LOC, do inglês *linear operator channel*).

No estudo de MMCs, costuma-se distinguir dois cenários de interesse, ilustrados pela Figura 1.10. No chamado *cenário não-coerente*, a matriz de transferência é desconhecida tanto do transmissor quanto do receptor. Nesse caso, a entrada do canal é a matriz X e a saída do

(a) Cenário não-coerente. (b) Cenário coerente.

Figura 1.10: Dois cenários para o canal matricial multiplicativo.

Figura 1.11: Transmissão de informação por um DMC.

canal é a matriz Y (Figura 1.10a). Já no chamado *cenário coerente*, a matriz de transferência é desconhecida do transmissor, mas conhecida pelo receptor. Assim, a entrada do canal é a matriz X, ao passo que a saída do canal é o par de matrizes (Y, G) (Figura 1.10b).

Neste trabalho, é adotado um enfoque probabilístico, no qual as matrizes em questão são consideradas *variáveis aleatórias*. Com isso, o MMC pode ser visto como um *canal discreto sem memória* (DMC, do inglês *discrete memoryless channel*). Canais discretos sem memória foram introduzidos por Shannon [45] no trabalho que deu origem à teoria da informação. Um DMC é um sistema utilizado para a transmissão de informação, sendo caracterizado por uma relação probabilística entre sua entrada e sua saída. Como mostrado por Shannon, o emprego de *codificadores* e *decodificadores* apropriados (Figura 1.11) possibilita a transmissão de informação de maneira confiável. A máxima taxa de informação que pode ser transmitida pelo canal com probabilidade de erro tão pequena quanto se queira é chamada de *capacidade* do canal. Shannon determinou uma fórmula para a capacidade de DMCs quaisquer, que depende da relação probabilística entre a entrada e a saída mencionada acima. No caso de MMCs, tal relação é determinada essencialmente pela distribuição de probabilidade da matriz de transferência.

Motivado por aplicações práticas em codificação de rede, o presente trabalho busca fornecer contribuições ao estudo de MMCs sobre corpos e anéis finitos, com ênfase no cálculo da capacidade do canal e na determinação de esquemas de codificação práticos que alcancem ou se aproximem da capacidade. Antes de descrever com mais detalhes as contribuições desta tese, serão mencionados alguns trabalhos anteriores

da literatura que também abordam o canal de comunicação fim-a-fim induzido por codificação de rede linear.

Trabalhos relevantes

Kötter e Kschischang [28] foram pioneiros em considerar a comunicação fim-a-fim em redes operando com codificação de rede linear, abrindo novos horizontes não apenas na área de codificação de rede, mas também nas áreas de teoria da informação e teoria de codificação. No trabalho, é considerado o caso mais geral com erros de enlace, modelado pela expressão $Y = GX + HZ$, em que Z é a matriz cujas linhas são os pacotes de erro e H é a matriz de transferência de Z para Y. É observado que, na ausência de erros de enlace ($Z = 0$) e no caso em que a matriz de transferência G tem posto completo, o *subespaço vetorial* gerado pelas linhas da matriz de entrada é sempre preservado, isto é $\langle Y \rangle = \langle X \rangle$, em que $\langle A \rangle$ denota o subespaço gerado pelas linhas de A. Agora, na presença de erros de enlace ou no caso em que a matriz de transferência apresenta deficiência de posto, pode-se limitar os possíveis subespaços transmitidos a um subconjunto particular de todos os subespaços vetoriais, definindo-se, assim, um *código de subespaço*.

Em [28], a métrica sugerida para o projeto de códigos de subespaço foi a chamada *distância de subespaço*. Também foi obtido um limitante superior (estilo Singleton) sobre o tamanho de códigos de subespaço de dimensão constante e foi proposta uma construção de códigos (estilo Reed-Solomon) que alcança assintoticamente tal limitante.

Posteriormente, os códigos propostos em [28] foram reinterpretados por Silva, Kötter e Kschischang [48] sob a ótica da *métrica de posto*. Em [48], também foi mostrado que a distância de subespaço fornece uma condição suficiente (mas não necessária) para que um código de subespaço seja "bem-sucedido" em sua tarefa de correção de erros de enlace e deficiência de posto no canal matricial não-coerente, sob a hipótese de que tanto a matriz de erros Z quanto as matrizes de transferência G e H sejam escolhidas por um adversário onisciente e com poder computacional ilimitado.

Na busca por uma condição necessária e suficiente, Silva e Kschischang [47] desenvolveram um arcabouço teórico para lidar com canais adversários genéricos; os resultados obtidos são então particularizados para o canal matricial não-coerente sujeito a erros de enlace e defi-

ciência de posto. Em particular, é encontrada a métrica que fornece a condição necessária e suficiente supracitada, sendo tal métrica chamada de *distância de injeção*. Além do caso não-coerente, [47] também considera o cenário coerente. Atualmente, a construção de códigos de subespaço baseada em métricas é uma área bastante ativa de pesquisa (veja, por exemplo, [26] e [9] e suas referências).

Apesar de a ideia de um enfoque probabilístico já ter sido sugerida em [28], o enfoque lá adotado foi essencialmente combinatório. Os primeiros artigos a efetivamente adotar o enfoque probabilístico no estudo da comunicação fim-a-fim em codificação de rede foram os de Montanari e Urbanke [32, 33], Jafari et al. [22, 23, 24], Silva, Kschischang e Kötter [50, 49] e Yang et al. [53, 56, 55]. Os resultados de tais trabalhos relacionados a MMCs serão expostos no Capítulo 2.

1.5 Contribuições

São abordados dois problemas nesta tese, sumarizados a seguir.

MMC não-coerente sobre corpos finitos com matriz de transferência uniforme dado o posto

Primeiramente, são considerados MMCs sobre corpos finitos em um cenário não-coerente. É assumido que a distribuição da matriz de transferência seja tal que matrizes de mesmo posto são equiprováveis. Desse modo, o DMC resultante é essencialmente caracterizado pela distribuição de posto da matriz de transferência. Essa hipótese contrasta com os trabalhos já existentes na literatura, de Silva et al. [49] (no qual a matriz de transferência é uniforme de posto completo) e Jafari et al. [24] (no qual a matriz de transferência possui entradas uniformes i.i.d.). Como será discutido, o modelo proposto é mais flexível e generaliza os modelos considerados em [49] e [24], o que permite sua aplicação em um número maior de cenários realistas (em particular em redes nas quais apagamentos de enlace têm um papel importante), ao mesmo tempo que é suficientemente simples para permitir análise matemática. O modelo também é conservador no sentido de que sua capacidade fornece um limitante inferior para a capacidade de um MMC genérico com a mesma distribuição de posto. Como contribuições, a capacidade do canal é expressa como a solução de um problema de otimização convexa que pode

ser facilmente resolvido com métodos numéricos habituais. Para o caso especial de entrada de posto constante, é obtida uma forma fechada para a capacidade. O comportamento do canal para comprimento do pacote ou tamanho do corpo arbitrariamente grande é estudado, sendo mostrado que entrada de posto constante é suficiente neste caso. Por fim, é provado que comunicação via subespaços ainda é ótima, mesmo nesse modelo mais geral.

MMC coerente sobre anéis de cadeia finitos

O segundo problema abordado nesta tese considera a comunicação em MMCs sobre anéis de cadeia finitos, desta vez considerando, por simplicidade, o cenário coerente. Anéis de cadeia finitos (dos quais corpos finitos são um caso particular) consistem de uma importante classe de anéis que surgem em diversas situações práticas no contexto de codificação de rede na camada física. Inicia-se apresentando conceitos preliminares sobre anéis de cadeia finitos e álgebra linear sobre tais anéis, os quais generalizam vários resultados familiares da álgebra linear sobre corpos. Como contribuições, é obtida uma forma fechada para a capacidade do canal e é proposto um esquema de codificação que alcança a capacidade em complexidade de tempo polinomial, fazendo uso de códigos matriciais sobre corpos finitos previamente existentes. Os resultados apresentados estendem os correspondentes para corpos finitos, obtidos por Yang et al. [55].

• • •

O restante desse documento é organizado como segue. Inicia-se com o Capítulo 2, introduzindo o modelo matemático de MMCs e apresentando uma breve revisão bibliográfica do assunto. O Capítulo 3 trata de MMCs sobre corpos finitos com matriz de transferência uniforme dado o posto. O Capítulo 4 considera MMCs sobre anéis de cadeia finitos. Finalmente, o Capítulo 5 conclui o trabalho, sumarizando as contribuições realizadas e sugerindo futuras investigações na área.

CAPÍTULO 2

Canais matriciais multiplicativos

Seja R um anel finito e sejam n, m e ℓ inteiros não-negativos. Como visto no capítulo introdutório, este trabalho considera o *canal matricial multiplicativo* (MMC), um canal de comunicação induzido pela expressão

$$\boldsymbol{Y} = \boldsymbol{G}\boldsymbol{X},$$

em que $\boldsymbol{X} \in R^{n\times\ell}$ é a *matriz de entrada*, $\boldsymbol{Y} \in R^{m\times\ell}$ é a *matriz de saída*[1] e $\boldsymbol{G} \in R^{m\times n}$ é a *matriz de transferência*, sendo tais matrizes variáveis aleatórias[2].

Neste capítulo, é apresentado o modelo matemático de um MMC, em que são distinguidos os cenários coerente e não-coerente. Por simplicidade, esse capítulo se restringe ao caso em que R é um corpo finito $\mathbb{F}_q$. Também são expostos alguns resultados já existentes na literatura. Antes de prosseguir, inicia-se revisando alguns conceitos preliminares.

[1] Ainda que no cenário coerente a saída do canal seja o par $(\boldsymbol{Y}, \boldsymbol{G})$, continua-se chamando $\boldsymbol{Y}$ de matriz de saída.

[2] De agora em diante, símbolos em negrito serão utilizados para denotar variáveis aleatórias, enquanto que símbolos comuns serão utilizados para suas amostras.

2.1 Preliminares

2.1.1 Canais discretos sem memória

Aqui é apresentada uma revisão das definições e dos resultados básicos acerca de canais discretos sem memória. Para mais detalhes, veja qualquer livro de teoria da informação, como, por exemplo, [6] ou [1]. Um *canal discreto sem memória* (DMC) com entrada $\boldsymbol{x}$ e saída $\boldsymbol{y}$ é definido por uma tripla $(\mathcal{X}, p_{\boldsymbol{y}|\boldsymbol{x}}, \mathcal{Y})$, em que

(*i*) $\mathcal{X}$ e $\mathcal{Y}$, denominados de *alfabeto de entrada* e *saída*, respectivamente, são conjuntos finitos e

(*ii*) $p_{\boldsymbol{y}|\boldsymbol{x}}(y|x)$, denominada de *probabilidade de transição*, fornece a probabilidade de se receber $\boldsymbol{y} = y \in \mathcal{Y}$ dado que $\boldsymbol{x} = x \in \mathcal{X}$ foi enviado.

Observação: A probabilidade de transição $p_{\boldsymbol{y}|\boldsymbol{x}}$ é costumeiramente representada por uma matriz de dimensão $|\mathcal{X}| \times |\mathcal{Y}|$, chamada de *matriz de transição*, cujas linhas e colunas são indexadas, respectivamente, por $\mathcal{X}$ e $\mathcal{Y}$, e cuja entrada x, y é dada por $p_{\boldsymbol{y}|\boldsymbol{x}}(y|x)$.

O canal é *sem memória* no sentido de que o símbolo de saída em um dado instante depende apenas do símbolo de entrada nesse mesmo instante, sendo condicionalmente independente de símbolos de entrada e saída passados. Além disso, é assumida a ausência de realimentação, de modo que o símbolo de entrada em um dado instante é independente dos símbolos de saída passados. Assim, se o canal for utilizado N vezes, com sequência de entrada $\vec{\boldsymbol{x}} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N) \in \mathcal{X}^N$ distribuída de acordo com $p_{\vec{\boldsymbol{x}}}$, então a sequência de saída $\vec{\boldsymbol{y}} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N) \in \mathcal{Y}^N$ será distribuída de acordo com $p_{\vec{\boldsymbol{y}}}$, a qual é induzida por $p_{\vec{\boldsymbol{x}}}$ e $p_{\vec{\boldsymbol{y}}|\vec{\boldsymbol{x}}}$, em que

$$p_{\vec{\boldsymbol{y}}|\vec{\boldsymbol{x}}}(y_1, y_2, \ldots, y_N | x_1, x_2, \ldots, x_n) = \prod_{i=1}^{N} p_{\boldsymbol{y}|\boldsymbol{x}}(y_i|x_i).$$

Considere um DMC com alfabeto de entrada $\mathcal{X}$ e alfabeto de saída $\mathcal{Y}$. Sejam N e M inteiros positivos. Um *código* $\mathcal{C}$ de *comprimento* N e *taxa* $\mathrm{R}(\mathcal{C}) = (\log M)/N$ para esse canal é definido por uma *função de codificação* $\phi : \{1, 2, \ldots, M\} \to \mathcal{X}^N$ e uma *função de decodificação* $\hat{\phi} : \mathcal{Y}^N \to \{1, 2, \ldots, M\}$. Uma sequência $\vec{x} = (x_1, x_2, \ldots, x_N) \in \mathcal{X}^N$

é dita ser uma *palavra-código* de $\mathcal{C}$ se $\vec{x} = \phi(m)$ para algum $m \in \{1, 2, \ldots, M\}$. Em um abuso de terminologia e notação, o conjunto de todas as palavras-código também é denominado de *código* e denotado por $\mathcal{C} \subseteq \mathcal{X}^N$.

Considere um DMC definido por $(\mathcal{X}, p_{\boldsymbol{y}|\boldsymbol{x}}, \mathcal{Y})$. Seja $\boldsymbol{m}$, denominada de *mensagem*, uma variável aleatória uniformemente distribuída sobre $\{1, 2, \ldots, M\}$. Adicionalmente, sejam $\vec{\boldsymbol{x}} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N) = \phi(\boldsymbol{m})$ a sequência na entrada do canal e $\vec{\boldsymbol{y}} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N)$ a sequência correspondente na saída do canal. Finalmente, seja $\hat{\boldsymbol{m}} = \hat{\phi}(\vec{\boldsymbol{y}})$ a estimativa da mensagem $\boldsymbol{m}$ feita pelo decodificador. Então, a *probabilidade de erro* de um código $\mathcal{C}$ é definida por

$$\mathrm{P_e}(\mathcal{C}) = \Pr[\boldsymbol{m} \neq \hat{\boldsymbol{m}}].$$

A *capacidade* de um DMC é definida por

$$C = \max_{p_{\boldsymbol{x}}} I(\boldsymbol{x}; \boldsymbol{y}),$$

em que

$$I(\boldsymbol{x}; \boldsymbol{y}) = \sum_{x \in \mathcal{X}} p_{\boldsymbol{x}}(x) \sum_{y \in \mathcal{Y}} p_{\boldsymbol{y}|\boldsymbol{x}}(y|x) \log \frac{p_{\boldsymbol{y}|\boldsymbol{x}}(y|x)}{p_{\boldsymbol{y}}(y)}$$

é a *informação mútua* entre $\boldsymbol{x}$ e $\boldsymbol{y}$ e a maximização se dá sobre todas as possíveis distribuições de entrada $p_{\boldsymbol{x}}$.

O seguinte resultado memorável foi obtido por Shannon [45]: Se $R < C$, então, para todo $\epsilon > 0$, existe um código $\mathcal{C}$ (de comprimento N suficientemente grande) tal que $\mathrm{R}(\mathcal{C}) \geq R$ e $\mathrm{P_e}(\mathcal{C}) \leq \epsilon$. Por outro lado, se $R > C$, então não existe tal código. Em outras palavras, a capacidade representa a maior taxa na qual informação pode ser transmitida pelo canal com probabilidade de erro arbitrariamente pequena.

2.1.2 Matrizes e subespaços sobre corpos finitos

Seja W um espaço vetorial sobre um corpo finito $\mathbb{F}_q$. O conjunto de todos os subespaços k-dimensionais de W é denominado de *grassmanniana* e é denotado por $\mathcal{G}_k(W)$. O tamanho da grassmanniana é dado por

$$|\mathcal{G}_k(W)| = \begin{bmatrix} m \\ k \end{bmatrix}_q,$$

em que $m = \dim W$ e

$$\begin{bmatrix} m \\ k \end{bmatrix}_q = \begin{cases} \prod_{i=0}^{k-1} \frac{q^m - q^i}{q^k - q^i}, & \text{se } 0 \leq k \leq m, \\ 0, & \text{caso contrário,} \end{cases} \tag{2.1}$$

é o *coeficiente binomial gaussiano*. Além disso, denota-se por $\mathcal{P}(W, d)$ o conjunto de todos os subespaços de W com dimensão no máximo d, isto é,

$$\mathcal{P}(W, d) = \bigcup_{k=0}^{d} \mathcal{G}_k(W).$$

No próximo capítulo, será utilizado o fato de que o coeficiente binomial gaussiano satisfaz

$$q^{k(m-k)} \leq \begin{bmatrix} m \\ k \end{bmatrix}_q \leq \gamma_q \, q^{k(m-k)}, \tag{2.2}$$

em que

$$\gamma_q = \prod_{i=1}^{\infty} \frac{1}{1 - q^{-i}}$$

é tal que $\lim_{q\to\infty} \gamma_q = 1$. Para uma prova de (2.2), veja [28, Lemma 4].

Denota-se por $\mathbb{F}_q^{m\times n}$ o conjunto de todas as matrizes $m \times n$ com entradas em $\mathbb{F}_q$. O subconjunto de $\mathbb{F}_q^{m\times n}$ consistindo de todas as matrizes de posto r é denotado por $\mathcal{T}_r(\mathbb{F}_q^{m\times n})$. Além disso, define-se $\mathcal{T}(\mathbb{F}_q^{m\times n}) = \mathcal{T}_{\min\{n,m\}}(\mathbb{F}_q^{m\times n})$, isto é, $\mathcal{T}(\mathbb{F}_q^{m\times n})$ é o conjunto de todas as matrizes $m \times n$ sobre $\mathbb{F}_q$ de posto completo[3]. Por comodidade, abrevia-se $\mathcal{T}_r = \mathcal{T}_r(\mathbb{F}_q^{m\times n})$ quando a dimensão $m \times n$ e o corpo $\mathbb{F}_q$ são facilmente inferidos pelo contexto. Tem-se [13]

$$|\mathcal{T}(\mathbb{F}_q^{m\times n})| = \begin{cases} \prod_{i=0}^{n-1} (q^m - q^i), & \text{se } 0 \leq n \leq m, \\ 0, & \text{caso contrário,} \end{cases} \tag{2.3}$$

[3]Note que $\mathcal{T}(\mathbb{F}_q^{n\times n})$ nada mais é que o conjunto de todas as matrizes $n \times n$ sobre $\mathbb{F}_q$ inversíveis, o qual é chamado de *grupo linear geral* de grau n sobre $\mathbb{F}_q$ e denotado por $\mathrm{GL}_n(\mathbb{F}_q)$.

e

$$|\mathcal{T}_r(\mathbb{F}_q^{m\times n})| = \frac{|\mathcal{T}(\mathbb{F}_q^{m\times r})||\mathcal{T}(\mathbb{F}_q^{n\times r})|}{|\mathcal{T}(\mathbb{F}_q^{r\times r})|} = |\mathcal{T}(\mathbb{F}_q^{m\times r})| \begin{bmatrix} n \\ r \end{bmatrix}_q . \tag{2.4}$$

2.2 MMC não-coerente

Finalmente, esta seção apresenta a definição de um MMC. Inicia-se pelo caso não-coerente.

No cenário não-coerente, em que a matriz de transferência é desconhecida tanto do transmissor quanto do receptor, o DMC resultante é definido por $(\mathcal{X}, p_{\boldsymbol{Y}|\boldsymbol{X}}, \mathcal{Y})$, em que o alfabeto de entrada é $\mathcal{X} = \mathbb{F}_q^{n\times\ell}$, o alfabeto de saída é $\mathcal{Y} = \mathbb{F}_q^{m\times\ell}$ e a probabilidade de transição, de acordo com a lei da probabilidade total, é dada por

$$p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \sum_G p_{\boldsymbol{G}|\boldsymbol{X}}(G|X)\, p_{\boldsymbol{Y}|\boldsymbol{G},\boldsymbol{X}}(Y|G,X),$$

em que $p_{\boldsymbol{G}|\boldsymbol{X}}(G|X)$ é a probabilidade de $\boldsymbol{G}$ condicionada a $\boldsymbol{X}$ e

$$p_{\boldsymbol{Y}|\boldsymbol{G},\boldsymbol{X}}(Y|G,X) = \begin{cases} 1, & \text{se } Y = GX, \\ 0, & \text{caso contrário.} \end{cases}$$

É natural considerar $\boldsymbol{X}$ e $\boldsymbol{G}$ estatisticamente independentes, de modo que $p_{\boldsymbol{G}|\boldsymbol{X}}(G|X) = p_{\boldsymbol{G}}(G)$ e, portanto,

$$p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \sum_{G:Y=GX} p_{\boldsymbol{G}}(G).$$

Note que, fixado o corpo finito $\mathbb{F}_q$, o canal recém definido é especificado completamente pelos parâmetros n, m, ℓ e $p_{\boldsymbol{G}}$. Neste trabalho, no entanto, denota-se o MMC não-coerente simplesmente por MMC$(\boldsymbol{G}, \ell)$.

Códigos para MMCs são chamados de *códigos matriciais*. Para o caso do MMC não-coerente, um código matricial $\mathcal{C}$ de comprimento N é definido por uma função de codificação $\phi : \{1, 2, \ldots, M\} \to (\mathbb{F}_q^{n\times\ell})^N$ e uma função de decodificação $\hat{\phi} : (\mathbb{F}_q^{m\times\ell})^N \to \{1, 2, \ldots, M\}$. Portanto, cada palavra-código de $\mathcal{C}$ é uma sequência de N matrizes em $\mathbb{F}_q^{n\times\ell}$. Quando $N = 1$, diz-se que $\mathcal{C}$ é um código matricial *one-shot*; caso contrário, $\mathcal{C}$ é dito ser *multi-shot*.

Resultados existentes

Dois casos particulares de importância tanto teórica quanto prática foram estudados na literatura e são descritos a seguir. O primeiro deles, de Silva et al. [49], calcula a capacidade do MMC assumindo que a matriz de transferência é quadrada e uniforme sobre todas as matrizes inversíveis.

Teorema 2.1. *Seja* $\boldsymbol{G} \in \mathbb{F}_q^{n\times n}$ *uniformemente distribuída sobre todas as matrizes inversíveis. Então, a capacidade de* $\mathrm{MMC}(\boldsymbol{G}, \ell)$*, em símbolos q-ários por uso do canal, é dada por*

$$C = \log_q \sum_{k=0}^{n} \begin{bmatrix} \ell \\ k \end{bmatrix}_q .$$

Assintoticamente no tamanho do corpo, q, a capacidade é dada por

$$\lim_{q\to\infty} C = (\ell - n)n.$$

Um código one-shot que alcança a capacidade com probabilidade de erro igual a zero pode ser construído escolhendo uma matrix $X \in \mathbb{F}_q^{n\times \ell}$ *tal que* $\langle X\rangle = U$ *para cada subespaço* $U \in \mathcal{P}(\mathbb{F}_q^{\ell}, n)$

O código proposto pelo Teorema 2.1, apesar de ótimo, possui codificação e decodificação não-triviais. Um esquema sub-ótimo alternativo de fácil codificação e decodificação consiste na utilização de cabeçalhos, como descrito no Capítulo 1. Claramente, o esquema com cabeçalhos possui probabilidade de erro igual a zero e a taxa igual a $(\ell-n)n$ sendo, portanto, assintoticamente ótimo no tamanho do corpo, q. Note também que o Teorema 2.1 justifica a ideia de *comunicação via subespaços* citada no Capítulo 1, na qual a informação é enviada na escolha do subespaço gerado pelas linhas da matriz de entrada.

Jafari et al. [24], em contraste, consideram a matriz de transferência com entradas i.i.d. uniformes sobre $\mathbb{F}_q$, o que equivale a dizer que a matriz de transferência é uniforme sobre todas as matrizes em $\mathbb{F}_q^{m\times n}$. A capacidade é obtida como a solução de um problema de otimização convexa sobre $\min\{n, m\}+1$ variáveis (que não será reproduzido aqui). Uma expressão exata para a capacidade, quando o tamanho q do corpo é superior a um dado limiar, também é fornecida (veja abaixo). Adicionalmente, é mostrado que comunicação via subespaços também é

ótima no modelo particular adotado.

Teorema 2.2. *Seja $\boldsymbol{G} \in \mathbb{F}_q^{m\times n}$ uniformemente distribuída sobre todas as matrizes em $\mathbb{F}_q^{m\times n}$. Então, existe q_0 tal que, se $q > q_0$, a capacidade de* MMC$(\boldsymbol{G}, \ell)$*, em símbolos q-ários por uso do canal, é dada por*

$$C = \sum_v |\mathcal{T}_v(\mathbb{F}_q^{n\times u^*})| q^{-nu^*} \log \frac{\begin{bmatrix}\ell\\ v\end{bmatrix}_q}{\begin{bmatrix}u^*\\ v\end{bmatrix}_q},$$

em que $u^* = \min\{n, m, \lfloor \ell/2 \rfloor\}$. Assintoticamente no tamanho do corpo, q, a capacidade é dada por

$$\lim_{q\to\infty} C = (\ell - u^*)u^*.$$

O Capítulo 3 deste trabalho considera um modelo dos quais ambos os modelos acima são casos particulares. Mais precisamente, é permitido que a matriz de transferência $\boldsymbol{G}$ tenha distribuição de probabilidade do *posto* arbitrária; entretanto, considera-se que todas as matrizes de mesmo posto sejam equiprováveis.

2.3 MMC coerente

Para o cenário coerente, em que a matriz de transferência é desconhecidas do transmissor, mas conhecida pelo receptor, obtém-se um DMC definido por $(\mathcal{X}, p_{\boldsymbol{Y},\boldsymbol{G}|\boldsymbol{X}}, \mathcal{Y})$, em que o alfabeto de entrada é $\mathcal{X} = \mathbb{F}_q^{n\times\ell}$, o alfabeto de saída é $\mathcal{Y} = \mathbb{F}_q^{m\times\ell} \times \mathbb{F}_q^{m\times n}$ e a probabilidade de transição é dada por

$$p_{\boldsymbol{Y},\boldsymbol{G}|\boldsymbol{X}}(Y, G|X) = \begin{cases} p_{\boldsymbol{G}}(G), & \text{se } Y = GX, \\ 0, & \text{caso contrário,} \end{cases}$$

em que, novamente, considerou-se que as matrizes $\boldsymbol{X}$ e $\boldsymbol{G}$ são estatisticamente independentes. Neste trabalho, o canal em questão é denotado por C-MMC$(\boldsymbol{G}, \ell)$.

Note que, no caso de códigos matriciais para o MMC coerente, a função de codificação tem contradomínio $(\mathbb{F}_q^{n\times\ell})^N$, exatamente como no caso não-coerente; no entanto, a função de decodificação tem domínio $(\mathbb{F}_q^{m\times\ell} \times \mathbb{F}_q^{m\times n})^N$, ou seja, o decodificador tem acesso não apenas

às matrizes de saída $\boldsymbol{Y_1}, \boldsymbol{Y_2}, \ldots, \boldsymbol{Y}_N$, mas também às matrizes de transferência $\boldsymbol{G}_1, \boldsymbol{G}_2, \ldots, \boldsymbol{G}_N$.

Resultados existentes

A capacidade do MMC coerente em sua forma mais geral foi encontrada por Yang et al. [55].

Teorema 2.3. *Seja $\boldsymbol{G}$ com distribuição de probabilidade $p_{\boldsymbol{G}}$ qualquer. Então, a capacidade de* C-MMC($\boldsymbol{G}, \ell$) *é dada, em dígitos q-ários por uso do canal, por*

$$C = \mathrm{E}[\boldsymbol{r}]\ell,$$

em que E[·] *denota* valor esperado*, sendo alcançada se a entrada for uniformemente distribuída sobre $\mathbb{F}_q^{n\times\ell}$. Em particular, a capacidade depende de $p_{\boldsymbol{G}}$ apenas através de* E[$\boldsymbol{r}$].

Em [56, 55] também são propostos dois esquemas de codificação *multi-shot* para MMCs sobre corpos finitos, capazes de alcançar a capacidade dada pelo Teorema 2.3. O primeiro faz uso de códigos de métrica de posto [48] e requer $\ell \geq n$; o segundo é baseado em codificação aleatória e não impõe nenhuma restrição sobre ℓ. Ambos têm complexidade de tempo polinomial. Ressalta-se que ambos os esquemas de codificação são "universais" no sentido de que apenas o valor esperado E[$\boldsymbol{r}$] é levado em conta na construção do código (não é necessário o conhecimento completo de $p_{\boldsymbol{G}}$, nem mesmo de $p_{\boldsymbol{r}}$).

O Capítulo 4 deste trabalho considera o MMC coerente sobre anéis de cadeia finitos, dos quais corpos finitos são um caso particular. É apresentada uma generalização do Teorema 2.3. Além disso, é proposto um esquema de codificação que combina códigos existentes sobre corpos finitos (tais como os citados no parágrafo anterior) para se obter um código sobre o anel de cadeia.

CAPÍTULO 3

MMC não-coerente sobre corpos finitos com matriz de transferência uniforme dado o posto

3.1 Introdução

Como já mencionado, MMCs não-coerentes sobre corpos finitos foram estudados anteriormente por Silva et al. [49] e Jafari et al. [24]. Especificamente, em [49], a matriz de transferência $\boldsymbol{G}$ é escolhida uniformemente dentre todas as matrizes de posto completo, enquanto que em [24], é assumido que $\boldsymbol{G}$ possui entradas i.i.d. selecionadas uniformemente (ou, equivalentemente, $\boldsymbol{G}$ é uniforme sobre todas as matrizes). Embora essas distribuições poderiam, em princípio, ser usadas para modelar sistemas operando de acordo com codificação de rede linear aleatória, elas são incapazes de refletir adequadamente diferentes topologias de rede, ou descrever com exatidão sistemas nos quais apagamentos de enlace têm um papel importante. Isto se deve porque nesses modelos, a matriz de transferência é completamente especificada pelo

O conteúdo deste capítulo foi apresentado no *2011 IEEE International Symposium on Information Theory (ISIT'11)* [42] e foi posteriormente publicado nas *IEEE Transactions on Information Theory* [41]. As ideias nas quais este capítulo é baseado são decorrentes de um documento não publicado [51].

tamanho do corpo, q, e pelas dimensões n e m. Por outro lado, uma descrição precisa de uma matriz de transferência com distribuição de probabilidade completamente geral exigiria, adicionalmente, a especificação de q^{nm} parâmetros (a saber, $p_{\boldsymbol{G}}(G)$, para $G \in \mathbb{F}_q^{m\times n}$), sendo portanto impraticável mesmo para valores modestos de q, n e m.

Em face a essa tensão entre tratabilidade e generalidade, o presente capítulo sugere um novo modelo o qual generaliza ambos os modelos de [49] e [24], mas que mantém a um nível realista a quantidade de informação necessária para descrever o canal. Mais especificamente, o modelo aqui adotado permite que a distribuição de probabilidade do *posto* de $\boldsymbol{G}$ seja arbitrária; entretanto, considera-se que todas as matrizes de mesmo posto sejam equiprováveis. Neste trabalho, uma matriz com esse tipo de distribuição de probabilidade é dita ser *uniforme dado o posto* (abreviado como *u.g.r.*, do inglês *uniform given rank*) (§3.2). Sob essa hipótese, a distribuição de probabilidade do posto da matriz de transferência caracteriza completamente a distribuição de probabilidade da própria matriz de transferência e, portanto, também determina completamente o canal. O modelo requer apenas $\min\{n, m\} + 1$ parâmetros para descrever o canal (a saber, $p_{\boldsymbol{r}}(r)$, para $0 \leq r \leq \min\{n, m\}$, em que $\boldsymbol{r} = \operatorname{rank} \boldsymbol{G}$). Embora o cálculo analítico da distribuição do posto em uma topologia de rede genérica seja um problema desafiador (até mesmo no caso mais simples de enlaces livres de apagamento), estimativas razoáveis podem ser obtidas na prática através de simulação de Monte Carlo, dado um modelo de rede. Na verdade, a distribuição (empírica) do posto é uma figura de mérito natural presente em diversas implementações de esquemas de codificação de rede não-coerente (veja, por exemplo, [5]). Assim, é razoável supor que tal informação está, de fato, disponível.

Como justificativa da utilidade prática do modelo proposto, é fornecido um exemplo que ilustra como o modelo u.g.r. é capaz de capturar de forma mais adequada o comportamento de sistemas que operam com codificação de rede linear aleatória, quando comparado com outros modelos existentes. Especificamente, será visto que, para certas topologias de rede, as capacidades fornecidas em [49, 24] desviam cada vez mais da verdadeira capacidade à medida que (i) a distância (no grafo) entre os nós fonte e destino aumenta ou (ii) a probabilidade de apagamento nos enlaces cresce (§3.3). Ademais, como será provado, qualquer MMC

pode ser reduzido para um MMC com matriz de transferência u.g.r. (embora às custas de uma diminuição potencial na capacidade do canal) através de um simples pre-processamento no transmissor e no receptor. Uma vez que tal pre-processamento não altera a distribuição de posto da matriz de transferência, pode-se concluir que, dentre todas as matrizes de transferência que compartilham a mesma distribuição de posto, a u.g.r. é aquela que fornece a menor capacidade. Nesse sentido (isto é, com a distribuição de posto mantida fixa), a distribuição u.g.r. é a distribuição de "pior caso" (§3.4).

O restante do capítulo se concentra na capacidade e informação mútua do MMC não-coerente com matriz de transferência u.g.r. Inicia-se obtendo a probabilidade de transição do canal. Em particular, é mostrado que essa probabilidade de transição depende das matrizes de entrada e saída apenas através de seus postos (§3.5). É mostrado que a capacidade do canal é alcançada quando a matriz de entrada (analogamente à matriz de transferência) é u.g.r., e uma fórmula para a informação mútua do canal é obtida para este tipo de entrada. Como consequência, a complexidade computacional associada ao problema de otimização convexa envolvido na obtenção da capacidade do canal é reduzida consideravelmente quando comparada com o caso de um MMC com matriz de transferência genérica—uma redução de $q^{n\ell}$ para $n+1$ variáveis, como será mostrado (§3.6). Em seguida, é abordado o caso particular de entrada de posto constante. Nesse caso, é possível obter uma forma fechada para a capacidade de posto constante (§3.7). Posteriormente, considera-se o problema no qual é permitido que q ou ℓ cresça arbitrariamente e é mostrado que a capacidade do canal (sem restrição na entrada) é alcançada com entrada de posto constante (§3.8). Como contribuição final, é verificado que comunicação via subespaços permanece ótima mesmo nesse caso mais geral de matriz de transferência u.g.r. (§3.9). Os resultados dessas seções generalizam alguns daqueles obtidos anteriormente em [49] e [24].

O trabalho de Yang et al. [55, 53, 56, 54], realizado simultânea e independentemente deste, considera uma matriz de transferência completamente geral (independente da entrada). Os autores identificam uma classe de entrada (chamada de "α-*type*") que é suficiente para alcançar a capacidade do canal. Como consequência, o número de variáveis envolvidas no problema do cálculo da capacidade de canal é

reduzido, porém para uma quantidade ainda exponencial no tamanho da matriz de transferência. Também é obtido limitantes inferiores e superiores na capacidade, os quais dependem apenas da distribuição de posto da matriz de transferência. Vale mencionar que alguns dos resultados aqui apresentados poderiam ser obtidos como casos particulares dos resultados em [55] (comparações serão feitas ao longo do capítulo quando aplicáveis.) Entretanto, acredita-se que o enfoque adotado aqui é mais simples e mais iluminador para o caso particular de matriz de transferência u.g.r.

Por fim, vale a observação de que alguns dos resultados aqui obtidos foram subsequentemente empregados em [25], no qual o MMC é modelado por um *canal arbitrariamente variável* (AVC, do inglês *arbitrarily varying channel*). Mais precisamente, é assumido em [25] que o posto da matriz de transferência é escolhido aleatoriamente de acordo com uma distribuição de probabilidade conhecida, mas, fora isso, a matriz de transferência pode variar arbitrariamente a cada uso do canal. É mostrado que a capacidade desse canal é a mesma capacidade do MMC com matriz de transferência u.g.r. considerado aqui.

3.2 Modelo do canal

Como dito anteriormente, este capítulo considera canais nos quais a matriz de transferência é "uniforme dado o posto". Tal conceito é definido formalmente a seguir.

Definição. Uma matriz aleatória $\boldsymbol{A} \in \mathbb{F}_q^{m\times n}$, distribuída de acordo com $p_{\boldsymbol{A}}$, é dita ser *uniforme dado o posto* (u.g.r.) se, para todos $A, A' \in \mathbb{F}_q^{m\times n}$ tais que $\operatorname{rank} A = \operatorname{rank} A'$, valer $p_{\boldsymbol{A}}(A) = p_{\boldsymbol{A}}(A')$.

Seja $\boldsymbol{A}$ uma matriz aleatória sobre $\mathbb{F}_q^{m\times n}$ com distribuição de probabilidade $p_{\boldsymbol{A}}$. Adicionalmente, seja $\boldsymbol{k} = \operatorname{rank} \boldsymbol{A}$ a variável aleatória assumindo valores em $\{0, \ldots, \min\{n, m\}\}$ de acordo com a distribuição de probabilidade $p_{\boldsymbol{k}}$ dada por

$$p_{\boldsymbol{k}}(k) = \sum_{A\in\mathcal{T}_k} p_{\boldsymbol{A}}(A).$$

Não é difícil verificar que $\boldsymbol{A}$ é u.g.r. se e somente se

$$p_{\boldsymbol{A}}(A) = \frac{p_{\boldsymbol{k}}(k)}{|\mathcal{T}_k(\mathbb{F}_q^{m\times n})|},$$

em que $k = \operatorname{rank} A$. Portanto, a distribuição de probabilidade do posto, $p_{\boldsymbol{k}}$, determina completamente $p_{\boldsymbol{A}}$ quando $\boldsymbol{A}$ é u.g.r. Adicionalmente, pode-se mostrar que a entropia de $\boldsymbol{A}$ satisfaz

$$H(\boldsymbol{A}) \leq \sum_k p_{\boldsymbol{k}}(k) \log_q \frac{|\mathcal{T}_k(\mathbb{F}_q^{m\times n})|}{p_{\boldsymbol{k}}(k)}, \tag{3.1}$$

com igualdade quando $\boldsymbol{A}$ for u.g.r. Assim, dentre todas as matrizes aleatórias com uma dada distribuição de posto, a u.g.r. é aquela com máxima entropia.

Este capítulo estuda o MMC não-coerente com matriz de entrada $\boldsymbol{X}$, matriz de saída $\boldsymbol{Y}$ e matriz de transferência $\boldsymbol{G}$ u.g.r. Define-se a variável aleatória

$$\boldsymbol{r} = \operatorname{rank} \boldsymbol{G},$$

distribuída de acordo com

$$p_{\boldsymbol{r}}(r) = \sum_{G\in\mathcal{T}_r} p_{\boldsymbol{G}}(G),$$

a qual representa o posto da matriz de transferência. Além disso, por simplicidade, assume-se $\max\{n, m\} \leq \ell$.

Como dito anteriormente, ambos os modelos de Silva et al. [49] e Jafari et al. [24] são casos particulares do modelo u.g.r. aqui considerado. De fato, para o modelo de canal de [49], no qual $\boldsymbol{G}$ é uniformemente distribuida sobre $\mathcal{T}(\mathbb{F}_q^{n\times n})$, tem-se

$$p_{\boldsymbol{r}}(r) = \begin{cases} 1, & \text{se } r = n, \\ 0, & \text{caso contrário,} \end{cases} \tag{3.2}$$

enquanto que para o modelo de canal de [24], no qual $\boldsymbol{G}$ é uniformemente distribuída sobre $\mathbb{F}_q^{m\times n}$, tem-se

$$p_{\boldsymbol{r}}(r) = \frac{|\mathcal{T}_r(\mathbb{F}_q^{m\times n})|}{q^{nm}}. \tag{3.3}$$

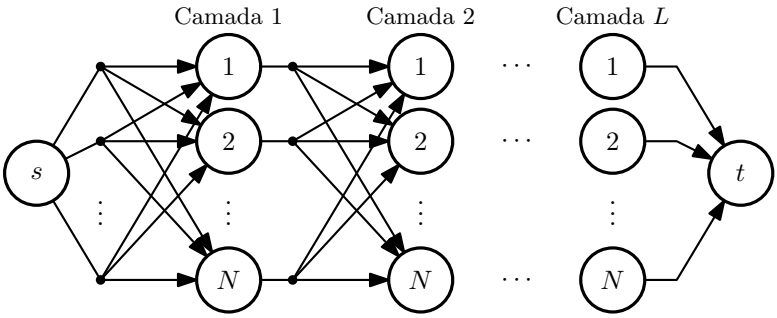


Figura 3.1: Rede sem-fio em camadas. Existem L camadas e cada camada possui N nós retransmissores.

3.3 Exemplo: Rede sem-fio em camadas

Esta seção apresenta um exemplo que mostra como o modelo u.g.r. é capaz de melhor representar um sistema operando de acordo com codificação de rede linear aleatória, quando comparado com modelos existentes. Considere a rede sem-fio ilustrada na Figura 3.1, com L camadas (colunas) e N nós retransmissores por camada. Assuma que o sistema opera com pacotes de comprimento ℓ e que entre duas camadas adjacentes (e também entre o nó fonte e a camada 1 e entre a camada L e o nó destino) existam N enlaces difusão (*broadcast*) ortogonais entre si, os quais estão sujeitos a apagamentos, ocorrendo nos finais do enlace, com probabilidade ϵ. Sempre que um pacote é apagado, considera-se que esse foi recebido como o vetor nulo. Adicionalmente, assuma que não haja comunicação entre nós não-adjacentes, bem como nós situados na mesma camada.

O sistema opera como segue. Primeiramente, o nó fonte s transmite pacotes à primeira camada utilizando todos os N canais de difusão ortogonais. O nó repete esse processo M vezes, de modo que um total de MN pacotes é recebido por cada nó da primeira camada. (É assumido que o nó fonte não efetua qualquer randomização.) Após isso, cada nó da primeira camada calcula M combinações lineares aleatórias (com coeficientes uniformes sobre $\mathbb{F}_q$ e i.i.d.) de todos os seus pacotes recebidos e, em seguida, difunde essas combinações lineares para a segunda camada, novamente em M instantes de tempo, utilizando um dos N canais ortogonais atribuído a si. Dessa forma, um total de MN pacotes

é recebido por cada nó da segunda camada, M de cada nó da primeira camada. O sistema opera analogamente até a camada L. Finalmente, o nó destino t recebe MN pacotes, M de cada nó da camada L.

Será mostrado agora que o sistema em questão pode ser modelado através de um MMC com $n = m = MN$. Seja $\boldsymbol{X} \in \mathbb{F}_q^{MN\times\ell}$ (resp., $\boldsymbol{Y} \in \mathbb{F}_q^{MN\times\ell}$) a matriz cujas linhas são os pacotes transmitidos (resp., recebidos) pelo nó fonte (resp., destino). Seja $\boldsymbol{R}_{i,j} \in \mathbb{F}_q^{MN\times\ell}$ (resp., $\boldsymbol{S}_{i,j} \in \mathbb{F}_q^{M\times\ell}$) a matriz cujas linhas são os pacotes recebidos (resp., transmitidos) pelo j-ésimo nó retransmissor da i-ésima camada, para $1 \le i \le L$ e $1 \le j \le N$. Da operação da rede descrita anteriormente, tem-se

$$\boldsymbol{S}_{i,j} = \boldsymbol{M}_{i,j}\boldsymbol{R}_{i,j},$$

para $1 \le i \le L$ e $1 \le j \le N$, em que $\boldsymbol{M}_{i,j} \in \mathbb{F}_q^{M\times MN}$ são matrizes cujas entradas são i.i.d., uniformes sobre $\mathbb{F}_q$. Também tem-se que

$$\boldsymbol{R}_{1,j} = \boldsymbol{E}_{1,j}\boldsymbol{X},$$

$$\boldsymbol{R}_{i,j} = \boldsymbol{E}_{i,j}\begin{bmatrix}\boldsymbol{S}_{i-1,1}\\ \vdots \\ \boldsymbol{S}_{i-1,N}\end{bmatrix} \quad \text{e} \quad \boldsymbol{Y} = \boldsymbol{E}'\begin{bmatrix}\boldsymbol{S}_{L,1}\\ \vdots \\ \boldsymbol{S}_{L,N}\end{bmatrix},$$

para $2 \le i \le L$ e $1 \le j \le N$, em que $\boldsymbol{E}_{i,j}, \boldsymbol{E}' \in \mathbb{F}_q^{MN\times MN}$ são matrizes diagonais (modelando os apagamentos) cujas entradas na diagonal são i.i.d. com $p(0) = \epsilon$ e $p(1) = 1 - \epsilon$. Disso tudo, deduz-se que

$$\boldsymbol{Y} = \boldsymbol{G}\boldsymbol{X},$$

em que

$$\boldsymbol{G} = \boldsymbol{E}'\boldsymbol{M}_L\boldsymbol{E}_L\cdots\boldsymbol{M}_2\boldsymbol{E}_2\boldsymbol{M}_1\boldsymbol{E}_1, \tag{3.4}$$

e $\boldsymbol{M}_i \in \mathbb{F}_q^{MN\times MN^2}$ (uma matriz bloco-diagonal) e $\boldsymbol{E}_i \in \mathbb{F}_q^{MN^2\times MN}$ são dadas por

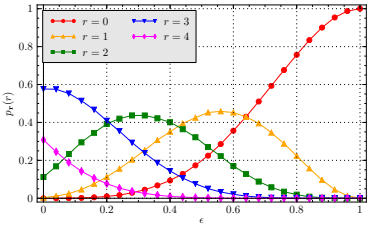
$$\boldsymbol{M}_i = \begin{bmatrix}\boldsymbol{M}_{i,1} & & \\ & \ddots & \\ & & \boldsymbol{M}_{i,N}\end{bmatrix} \quad \text{e} \quad \boldsymbol{E}_i = \begin{bmatrix}\boldsymbol{E}_{i,1}\\ \vdots \\ \boldsymbol{E}_{i,N}\end{bmatrix}.$$

Note que, em geral, a matriz de transferência dada por (3.4) *não é* u.g.r. Assim, como mencionado na introdução (e provado mais adiante na Seção 3.4), os resultados de capacidade das Seções 3.5 a 3.8 servirão apenas como limitantes inferiores para a real capacidade do canal. No entanto, vale observar que o cálculo da real capacidade do canal é uma tarefa computacionalmente intratável, mesmo para parâmetros pequenos. Por exemplo, se $q = 2$ e $n = m = \ell = 8$, seria necessário, a priori, resolver um problema de otimização sobre $q^{n\ell} = 2^{64}$ variáveis, o que é claramente impraticável. De acordo com [55], poderia-se reduzir a quantidade de variáveis para $\sum_{u=0}^{n} \begin{bmatrix} n \\ u \end{bmatrix}_q > 2^{18}$, um número ainda impraticável.

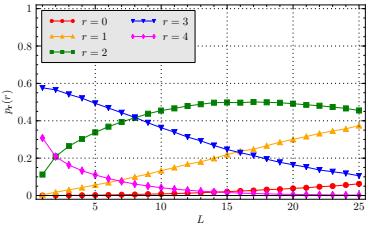
EXEMPLO. As Figuras 3.2a e 3.2b mostram a distribuição de posto, $p_{\boldsymbol{r}}$, induzida pela rede sem-fio em camadas com $q = 2$ e $N = M = 2$ (de modo que $n = m = MN = 4$) em função de ϵ, para $L = 1$, e em função de L, para $\epsilon = 0$, respectivamente. Note que o valor de ℓ não é importante aqui. Ambas as distribuições de posto foram obtidas de (3.4) pelo método de Monte Carlo com 100.000 realizações.

As Figuras. 3.2c e 3.2d mostram a capacidade do MMC correspondente, assumindo matriz de transferência u.g.r., com distribuições de posto das Figuras 3.2a e 3.2b. Considerou-se $\ell = 8$ para o comprimento do pacote. Os resultados foram obtidos do Teorema 3.3 da Seção 3.6. As figuras também mostram a capacidade de um MMC com os mesmos parâmetros q, n, m e ℓ, mas modelado por matriz de transferência uniforme de posto completo (Teorema 2.1) ou uniforme sobre todas as matrizes (Teorema 2.2), bem como o limitante superior considerando o cenário coerente (Teorema 2.3). □

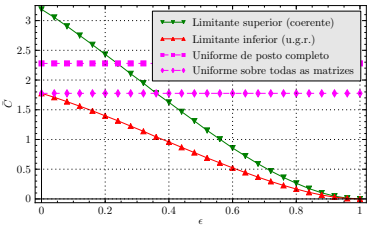
Claramente, os modelos de [49] e [24] são insensíveis aos efeitos de apagamento de enlace e variações na topologia (aqui ilustrado pelo número de camadas). Percebe-se que as capacidades para esses modelos desviam consideravelmente da verdadeira capacidade. Em contraste, a partir das tendências dos limitantes inferior e superior, pode-se inferir que a capacidade do modelo u.g.r. se comporta de forma muito parecida com a capacidade verdadeira (note que o limitante superior tende a zero quando ϵ se aproxima de 1, ou quando L aumenta; portanto, o mesmo acontece com a capacidade verdadeira). De fato, como o próximo exemplo ilustra, o limitante inferior u.g.r. pode se situar bastante próximo da capacidade verdadeira.



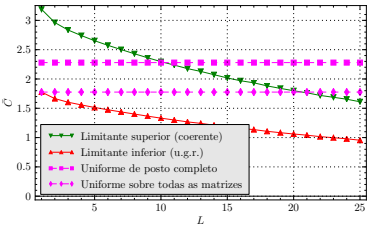
(a) Distribuição de posto em função de ϵ, para $L = 1$.



(b) Distribuição de posto em função de L, para $\epsilon = 0$.

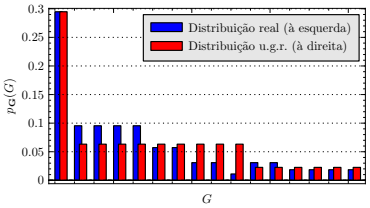


(c) Capacidade em funçao de ϵ, para $L = 1$ e $\ell = 8$.

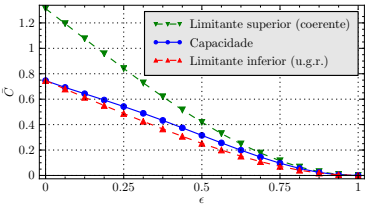


(d) Capacidade em funçao de L, para $\epsilon = 0$ e $\ell = 8$.

Figura 3.2: Distribuição de posto e capacidade de canal para a rede sem-fio em camadas com $N = M = 2$ e $q = 2$.



(a) Distribuição da matriz de transferência, para $\epsilon = 1/4$.



(b) Capacidade real e limitantes, em função de ϵ, para $\ell = 3$.

Figura 3.3: Distribuição da matriz de transferência e capacidade de canal para a rede sem-fio em camadas com $L = 1$, $N = 2$, $M = 1$ e $q = 2$. Na Figura 3.3a, o eixo horizontal consiste de todas as matrizes em $\mathbb{F}_2^{2\times 2}$, ordenadas da esquerda para a direita como segue: $[00;00]$, $[10;00]$, $[01;00]$, $[00;10]$, $[00;01]$, $[11;00]$, $[00;11]$, $[10;10]$, $[01;01]$, $[11;11]$, $[10;01]$, $[01;10]$, $[11;10]$, $[11;01]$, $[10;11]$, $[01;11]$.

EXEMPLO. Este exemplo tem o objetivo de quantificar a perda de taxa ocasionada quando se considera a matriz de transferência u.g.r. quando, de fato, ela não é. Para tanto, seja a rede sem-fio em camadas com tamanho do corpo $q = 2$, uma única camada ($L = 1$) e dois nós retransmissores ($N = 2$). Seja também $M = 1$, de modo que $n = m = 2$. Nesse caso, (3.4) fornece

$$\boldsymbol{G} = \boldsymbol{E}'\boldsymbol{M}_1\boldsymbol{E}_1 = \begin{bmatrix} \boldsymbol{e}_5\boldsymbol{a}_1\boldsymbol{e}_1 & \boldsymbol{e}_5\boldsymbol{a}_2\boldsymbol{e}_2 \\ \boldsymbol{e}_6\boldsymbol{a}_3\boldsymbol{e}_3 & \boldsymbol{e}_6\boldsymbol{a}_4\boldsymbol{e}_4 \end{bmatrix},$$

em que as entradas $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_6 \in \mathbb{F}_2$ (relacionadas aos apagamentos) são i.i.d. com $\Pr[\boldsymbol{e}_i = 0] = \epsilon$ e as entradas $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_4 \in \mathbb{F}_2$ (os coeficientes de codificação de rede) são i.i.d. com $\Pr[\boldsymbol{a}_i = 0] = 1/2$. A distribuição da matriz de transferência, $p_{\boldsymbol{G}}(G)$ com $\epsilon = 1/4$ é mostrada na Figura 3.3a, a qual também mostra a distribuição u.g.r. correspondente.

A Figura 3.3b mostra a verdadeira capacidade do canal (obtida resolvendo o problema de maximização original sobre $q^{n\ell} = 64$ variáveis), juntamente com o limitante inferior u.g.r. (obtido resolvendo o um problema de maximização sobre $n + 1 = 3$ variáveis, de acordo com o Teorema 3.3) e o limitante superior coerente (dado pelo Teorema 2.3), em função de ϵ, para um comprimento do pacote $\ell = 3$. É interessante observar que o limitante inferior u.g.r. é exato para $\epsilon = 0$, uma vez que, nesse caso, $\boldsymbol{G}$ torna-se uniformemente distribuída sobre $\mathbb{F}_q^{m\times n}$ e, portanto, é u.g.r. Além disso, para os demais valores de ϵ, a capacidade verdadeira é bastante próxima do limitantes inferior u.g.r., o que constitui uma evidência de que o modelo u.g.r. é realmente uma boa aproximação para sistemas que operam de acordo com codificação de rede linear aleatória. □

3.4 O modelo u.g.r. como o pior caso

Qualquer MMC pode ser artificialmente transformado em um MMC com matriz de transferência u.g.r. (tendo a mesma distribuição de posto do canal original) por meio de "randomização" tanto no transmissor quanto no receptor. O Teorema 3.1 a seguir torna tal afirmação precisa. A demonstração do teorema foi sugerida por Chen Feng (*University of Toronto*) e se dá através da aplicação de uma versão generalizada

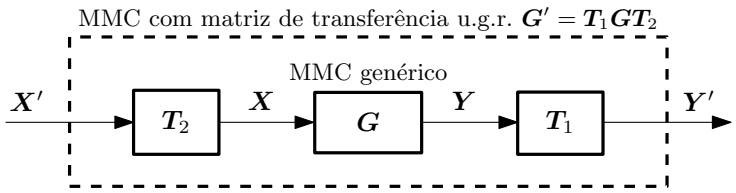


Figura 3.4: Convertendo um MMC genérico em um MMC com matriz de transferência u.g.r. A distribuição de posto do novo canal é a mesma do canal original.

do *cripto-lema* [14], a qual pode ser útil em outras aplicações (veja Apêndice B.1).

Teorema 3.1. *Seja $\boldsymbol{G} \in \mathbb{F}_q^{m\times n}$ uma matriz aleatória com distribuição de probabilidade qualquer e defina $\boldsymbol{G}' = \boldsymbol{T}_1\boldsymbol{G}\boldsymbol{T}_2$, em que $\boldsymbol{T}_1 \in \mathrm{GL}_m(\mathbb{F}_q)$ e $\boldsymbol{T}_2 \in \mathrm{GL}_n(\mathbb{F}_q)$ são matrizes inversíveis uniformemente distribuídas e independentes de $\boldsymbol{G}$ e uma da outra. Então, $\boldsymbol{G}'$ é u.g.r. e possui a mesma distribuição de posto de $\boldsymbol{G}$.*

Demonstração. O resultado segue após a aplicação do Lemma B.2 com $\mathcal{G} = \mathrm{GL}_m(\mathbb{F}_q) \times \mathrm{GL}_n(\mathbb{F}_q)$, em que a operação do grupo é dada por $(T_1', T_2') \cdot (T_1, T_2) = (T_1'T_1, T_2T_2')$, o conjunto da ação é dado por $\mathcal{S} = \mathbb{F}_q^{m\times n}$ e a operação da ação, $\circ : \mathcal{G}\times\mathcal{S} \to \mathcal{S}$, é definida por $(T_1, T_2)\circ M = T_1MT_2$. Os fatos de que $\mathcal{G}$ é um grupo e $\circ$ é uma ação de $\mathcal{G}$ em $\mathcal{S}$ seguem da álgebra linear; as órbitas, nesse caso, são $\{\mathcal{T}_r(\mathbb{F}_q^{m\times n}) : r = 0, 1, \ldots, \min\{n, m\}\}$, as quais são completamente caracterizadas pelo posto de $\boldsymbol{G}$. □

Efetivamente (veja a Figura 3.4), ao invés de transmitir os pacotes originais ($\boldsymbol{X}'$, digamos), o transmissor envia $\boldsymbol{X} = \boldsymbol{T}_2\boldsymbol{X}'$; e ao invés de utilizar a real saída do canal ($\boldsymbol{Y}$, digamos), o receptor considera $\boldsymbol{Y}' = \boldsymbol{T}_1\boldsymbol{Y}$ para a decodificação. (Aqui, $\boldsymbol{T}_1$ e $\boldsymbol{T}_2$ são definidos como no Teorema 3.1.) Consequentemente, se a matriz de transferência do canal original for $\boldsymbol{G}$, então $\boldsymbol{Y}' = \boldsymbol{T}_1\boldsymbol{Y} = \boldsymbol{T}_1\boldsymbol{G}\boldsymbol{X} = \boldsymbol{T}_1\boldsymbol{G}\boldsymbol{T}_2\boldsymbol{X}' = \boldsymbol{G}'\boldsymbol{X}'$, em que $\boldsymbol{G}'$, de acordo com o Teorema 3.1, é u.g.r. e possui a mesma distribuição de posto de $\boldsymbol{G}$. Naturalmente, como consequência da *desigualdade do processamento de dados* [6], tem-se $I(\boldsymbol{X}';\boldsymbol{Y}') \leq I(\boldsymbol{X};\boldsymbol{Y})$, de modo que essa conversão se dá em detrimento de uma potencial redução da capacidade do canal.

Assim, conclui-se que, dentre todas as matrizes de transferência que compartilham a mesma distribuição de posto, a u.g.r. é aquela com a menor capacidade de canal e que qualquer resultado de capacidade obtido para o MMC com matriz de transferência u.g.r. pode ser utilizado como um limitante inferior para MMCs com matrizes de transferências arbitrárias (não u.g.r.).

Alguns comentários adicionais são pertinentes. Primeiro, randomização no transmissor (mas não no receptor) já é uma prática usual em sistemas que empregam codificação de rede linear aleatória [28]. Segundo, uma vez que tanto a multiplicação de matrizes e a geração de matrizes aleatórias inversíveis podem ser realizadas em tempo polinomial, a randomização é também um procedimento com complexidade de tempo polinomial. Terceiro, visto que as matrizes $\boldsymbol{T}_1$ e $\boldsymbol{T}_2$ são independentes de $\boldsymbol{G}$ e uma da outra, não é assumido qualquer conhecimento do canal, e nem mesmo é necessário segredo compartilhado entre transmissor e receptor. Finalmente, para uma quantificação numérica da perda de taxa ocasionada pela randomização, veja o segundo exemplo da Seção 3.3.

3.5 Probabilidade de transição do canal

Nas próximas seções, além de $\boldsymbol{r} = \operatorname{rank} \boldsymbol{G}$, distribuída de acordo com $p_{\boldsymbol{r}}(r) = \sum_{A \in \mathcal{T}_r} p_{\boldsymbol{G}}(A)$, também faz-se uso das variáveis aleatórias $\boldsymbol{u} = \operatorname{rank} \boldsymbol{X}$ e $\boldsymbol{v} = \operatorname{rank} \boldsymbol{Y}$, cujas distribuições de probabilidade são dadas, respectivamente, por $p_{\boldsymbol{u}}(u) = \sum_{X \in \mathcal{T}_u} p_{\boldsymbol{X}}(X)$ e $p_{\boldsymbol{v}}(v) = \sum_{Y \in \mathcal{T}_v} p_{\boldsymbol{Y}}(Y)$.

A *probabilidade de transição de posto* (isto é, a probabilidade de se receber uma matriz de posto v dado que uma matriz de posto u foi transmitida) é de fundamental importância para este canal. Uma vez que $\boldsymbol{u} \to \boldsymbol{X} \to \boldsymbol{Y} \to \boldsymbol{v}$ é uma cadeia de Markov, a probabilidade de transição de posto é dada por

$$\begin{aligned} p_{\boldsymbol{v}|\boldsymbol{u}}(v|u) &= \sum_{X,Y} p_{\boldsymbol{v}|\boldsymbol{Y}}(v|Y) p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) p_{\boldsymbol{X}|\boldsymbol{u}}(X|u) \\ &= \sum_{X \in \mathcal{T}_u} p_{\boldsymbol{X}|\boldsymbol{u}}(X|u) \sum_{Y \in \mathcal{T}_v} p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) \end{aligned}$$

e, portanto, pode depender não apenas de $p_{\boldsymbol{Y}|\boldsymbol{X}}$ (ou seja, de $p_{\boldsymbol{G}}$), mas também de $p_{\boldsymbol{X}|\boldsymbol{u}}$. O teorema a seguir determina a probabilidade de

transição de posto para o caso de matriz de transferência u.g.r. e mostra que tal probabilidade independe de $p_{\boldsymbol{X}|\boldsymbol{u}}$. O teorema também expressa a probabilidade de transição do canal em termos da probabilidade de transição de posto. A demonstração do teorema pode ser encontrada na Seção 3.10, que também apresenta as demais demonstrações omitidas deste capítulo.

Teorema 3.2. *Considere um MMC com matriz de transferência u.g.r.*

(i) Sejam u, v e r inteiros tais que $0 \leq r \leq \min\{n, m\}$. Então,

$$p_{\boldsymbol{v}|\boldsymbol{u},\boldsymbol{r}}(v|u,r) = \frac{\begin{bmatrix} u \\ v \end{bmatrix}_q}{\begin{bmatrix} n \\ r \end{bmatrix}_q} \begin{bmatrix} n-u \\ r-v \end{bmatrix}_q q^{v(n-u-r+v)}. \tag{3.5}$$

Dessa forma, a probabilidade de transição de posto é dada por

$$p_{\boldsymbol{v}|\boldsymbol{u}}(v|u) = \sum_r p_{\boldsymbol{r}}(r) p_{\boldsymbol{v}|\boldsymbol{u},\boldsymbol{r}}(v|u,r)$$

e a distribuição de probabilidade do posto da saída é dada por

$$p_{\boldsymbol{v}}(v) = \sum_u p_{\boldsymbol{u}}(u) p_{\boldsymbol{v}|\boldsymbol{u}}(v|u).$$

(ii) A probabilidade de transição do canal é dada por

$$p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X) = \begin{cases} \dfrac{p_{\boldsymbol{v}|\boldsymbol{u}}(v|u)}{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|}, & \text{se } \langle Y \rangle \subseteq \langle X \rangle, \\ 0, & \text{caso contrário.} \end{cases} \tag{3.6}$$

(iii) Se a entrada $\boldsymbol{X}$ é u.g.r., então a saída $\boldsymbol{Y}$ também é u.g.r.

Observação: Sejam u, v e r inteiros tais que $0 \leq r \leq \min\{n, m\}$. Lembrando que o coeficiente binomial gaussiano $\begin{bmatrix} m \\ k \end{bmatrix}_q$ é não-nulo se e somente se $0 \leq k \leq m$, tem-se, de acordo com (3.5), que $p_{\boldsymbol{v}|\boldsymbol{u},\boldsymbol{r}}(v|u,r) \neq 0$ se e somente se $0 \leq v \leq u$ e $0 \leq r - v \leq n - u$; tais condições, por sua vez, são equivalentes a $u + r - n \leq v \leq \min\{u, r\}$. Isso já era esperado: o limitante superior é equivalente ao fato de que $\operatorname{rank} AX \leq \min\{\operatorname{rank} X, \operatorname{rank} A\}$ e o limitante inferior segue da desigualdade de Sylvester, que afirma que, se A e X são matrizes de dimensões $m \times n$ e $n \times \ell$, respectivamente, então $\operatorname{rank} X + \operatorname{rank} A - n \leq \operatorname{rank} AX$.

Observação: O Teorema 3.2 mostra que a *matriz de transição* $p_{\boldsymbol{Y}|\boldsymbol{X}}$ do canal em questão pode ser particionada em blocos, com a propriedade de que, em cada bloco, as linhas são permutações umas das outras e as colunas são permutações umas das outras. Assim, o MMC com matriz de transferência u.g.r. é um canal que apresenta a chamada "simetria em bloco" introduzida por Pedersen e Topsøe em [43], trabalho cujo autor da presente tese tomou conhecimento apenas após o fechamento deste capítulo. Desse modo, os resultados da Seção 3.6 poderiam, em princípio, ser obtidos a partir dos resultados de [43].

3.6 Capacidade do canal

A capacidade do canal é derivada a seguir. Será visto que entrada u.g.r. é suficiente para alcançar a capacidade, de modo que não há necessidade de se considerar entradas mais gerais. Seja

$$I^*(p_{\boldsymbol{u}}) = \max_{p_{\boldsymbol{X}}:p_{\boldsymbol{u}}} I(\boldsymbol{X};\boldsymbol{Y}), \tag{3.7}$$

em que a maximização é efetuada sobre todas as distribuições de probabilidade $p_{\boldsymbol{X}}$ cuja probabilidade de posto associada é $p_{\boldsymbol{u}}$, isto é, sobre o conjunto

$$\{p_{\boldsymbol{X}} : \textstyle\sum_{X\in\mathcal{T}_u} p_{\boldsymbol{X}}(X) = p_{\boldsymbol{u}},\ \text{para } u = 0, 1, \ldots, n\}.$$

Teorema 3.3. *Seja $\boldsymbol{G} \in \mathbb{F}_q^{m\times n}$ uma matriz u.g.r. Então, a capacidade de* MMC$(\boldsymbol{G}, \ell)$ *é dada, em símbolos q-ários por uso do canal, por*

$$C = \max_{p_{\boldsymbol{u}}} I^*(p_{\boldsymbol{u}}),$$

em que

$$I^*(p_{\boldsymbol{u}}) = \sum_v p_{\boldsymbol{v}}(v) \log_q \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times \ell})|}{p_{\boldsymbol{v}}(v)} - \sum_u h_u p_{\boldsymbol{u}}(u) \tag{3.8}$$

e

$$h_u = \sum_v p_{\boldsymbol{v}|\boldsymbol{u}}(v|u) \log_q \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|}{p_{\boldsymbol{v}|\boldsymbol{u}}(v|u)}, \tag{3.9}$$

sendo alcançada com distribuição de entrada u.g.r.

Do Teorema 3.3, conclui-se que o problema de encontrar a capacidade e a distribuição de entrada ótima correspondente para o MMC com matriz de transferência u.g.r., originalmente um problema de otimização convexa sobre $q^{n\ell}$ variáveis (a saber, $p_{\boldsymbol{X}}(X)$ para $X \in \mathbb{F}_q^{n\times\ell}$) pode ser simplificado para outro problema de otimização convexa, desta vez envolvendo apenas $n+1$ variáveis (a saber, $p_{\boldsymbol{u}}(u)$, para $u = 0, \ldots, n$). A solução de tal problema de otimização pode ser obtida por métodos numéricos eficientes bem estudados (veja, por exemplo, [3]).

3.7 Entrada de posto constante

Foca-se agora no caso especial em que as matrizes de entrada do canal são restritas a terem *posto constante*. Este caso é de interesse por pelo menos dois motivos. Primeiro, entrada de posto constante vem a ser assintoticamente ótima tanto com o comprimento do pacote quanto com o tamanho do corpo finito (como será visto mais adiante). E segundo, a maioria das construções de códigos de subespaço existentes fornecem "códigos na grassmanniana", isto é, códigos de subespaço de dimensão constantes [28].

Denota-se por C_u a máxima informação mútua do canal quando a entrada é restrita a matrizes de posto u. Seja u^* o valor de u que maximiza C_u, de modo que $C_{u^*} = \max_u C_u$. Denomina-se C_u de *capacidade de posto u* e C_{u^*} de *capacidade de posto constante* do canal.

Teorema 3.4. *Seja $\boldsymbol{G} \in \mathbb{F}_q^{m\times n}$ u.g.r. Então, a capacidade de posto u de* MMC($\boldsymbol{G}, \ell$)*, em símbolos q-ários por uso do canal, é dada por*

$$C_u = \sum_v p_{\boldsymbol{v}|\boldsymbol{u}}(v|u) \log_q \frac{\begin{bmatrix}\ell\\ v\end{bmatrix}_q}{\begin{bmatrix}u\\ v\end{bmatrix}_q}, \tag{3.10}$$

sendo alcançada com distribuição de entrada uniforme [sobre $\mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$]. Ademais,

$$C_{u^*} \leq C \leq C_{u^*} + \log_q(\min\{n, m\} + 1). \tag{3.11}$$

Observação: Em particular, para entrada de posto *completo* (isto é, $\boldsymbol{u} = n$), então $\boldsymbol{v} = \boldsymbol{r}$ (pois $\boldsymbol{v} = \operatorname{rank} \boldsymbol{Y} = \operatorname{rank} \boldsymbol{GX} = \operatorname{rank} \boldsymbol{G} = \boldsymbol{r}$). A

capacidade em (3.10) torna-se

$$C_n = \sum_r p_{\boldsymbol{r}}(r) \log_q \frac{\begin{bmatrix}\ell\\ r\end{bmatrix}_q}{\begin{bmatrix}n\\ r\end{bmatrix}_q},$$

um resultado obtido previamente em [51]. Além disso, como $p_{\boldsymbol{v}|\langle \boldsymbol{X}\rangle}(v|U)$ depende de U apenas através de $u = \dim U$ (veja o Teorema 3.2), o resultado concorda com [55, Theorem 7].

3.8 Comportamento assintótico

A seguir, considera-se o comportamento do canal para comprimento do pacote, ℓ, ou tamanho do corpo, q, assintoticamente grande. É mostrado que, em ambos os cenários, entrada de posto constante é suficiente para alcançar a capacidade.

Considera-se primeiramente o comportamento assintótico em ℓ, o comprimento do pacote. Nessa situação é apropriado definir $\bar{C} = C/\ell$, a *capacidade normalizada* do canal matricial, medida em pacotes por uso do canal. Também define-se a *capacidade de posto u normalizada* como $\bar{C}_u = C_u/\ell$ e a *capacidade de posto constante normalizada* como $\bar{C}_{u^*}$, em que u^* é o valor de u que maximiza $\bar{C}_u$.

Teorema 3.5. *Assintoticamente no comprimento do pacote, ℓ, a capacidade normalizada do MMC com matriz de transferência u.g.r. é dada por*

$$\lim_{\ell\to\infty} \bar{C} = \mathrm{E}[\boldsymbol{r}],$$

sendo alcançada com entrada uniforme de posto constante. O posto ótimo de entrada é sempre $u^* = n$.

Observação: Esse resultado também é obtido em [55, Corollary 1] para o caso de um MMC com matriz de transferência com distribuição de probabilidade qualquer.

Volta-se agora para o comportamento assintótico em q, o tamanho do corpo. Em geral, a distribuição do posto pode depender de q [por exemplo, no caso de (3.3)]. Assim, no que segue,

$$p_{\boldsymbol{r}}^{\infty}(r) = \lim_{q\to\infty} p_{\boldsymbol{r}}(r)$$

denota a distribuição limite de $\boldsymbol{r}$, assumindo que tal limite exista. Obviamente, quando a distribuição de posto não depende de q, então $p^{\infty}_{\boldsymbol{r}}(r) = p_{\boldsymbol{r}}(r)$.

Teorema 3.6. *Assintoticamente no tamanho do corpo, q, a capacidade do MMC com matriz de transferência u.g.r. é dada por*

$$\lim_{q\to\infty} C = \max_{u} \left[(\ell - u) \sum_{r} p^{\infty}_{\boldsymbol{r}}(r) \min\{u, r\} \right],$$

sendo alcançada com entrada uniforme de posto constante.

Observação: Considere codificação de rede linear aleatória na ausência de erros e apagamentos de enlace. Quando o tamanho do corpo, q, é assintoticamente grande, sabe-se que a matriz de transferência terá posto h com probabilidade 1, em que h é o corte mínimo (*mincut*) da rede [19]. Nesse caso, $p^{\infty}_{\boldsymbol{r}}(r) = 1[r = h]$, de modo que

$$\lim_{q\to\infty} C = \max_{u} \left[(\ell - u) \min\{u, h\} \right] = (\ell - u^*)u^*,$$

em que $u^* = \min\{h, \lfloor \ell/2 \rfloor\}$. Para o sub-caso no qual $h = \min\{n, m\}$, tem-se $u^* = \min\{n, m, \lfloor \ell/2 \rfloor\}$, o que concorda com os Teoremas 2.1 e 2.2, uma vez que em ambos os casos tem-se $p^{\infty}_{\boldsymbol{r}}(r) = 1[r = \min\{n, m\}]$ [veja as equações em (3.2) e (3.3)].

3.9 Comunicação via subespaços

O último resultado deste capítulo diz respeito à otimalidade da codificação de subespaço [28] para o MMC com matriz de transferência u.g.r. A demonstração do seguinte teorema faz uso do conceito de *agrupamentos sem perda de informação* de letras de entrada (ou saída) em DMCs (veja Apêndice B.2). No que segue, vale lembrar que $\mathcal{P}(\mathbb{F}_q^{\ell}, d)$ denota o conjunto de todos os subespaços de $\mathbb{F}_q^{\ell}$ com dimensão no máximo d.

Teorema 3.7. *Considere o MMC com matriz de transferência u.g.r. Sejam $\boldsymbol{U} = \langle \boldsymbol{X} \rangle$ e $\boldsymbol{V} = \langle \boldsymbol{Y} \rangle$. Então,*

$$I(\boldsymbol{X}; \boldsymbol{Y}) = I(\boldsymbol{U}; \boldsymbol{V}), \tag{3.12}$$

qualquer que seja a distribuição de entrada $p_{\boldsymbol{X}}$*. Adicionalmente, para todo* $U \in \mathcal{P}(\mathbb{F}_q^\ell, n)$ *e* $V \in \mathcal{P}(\mathbb{F}_q^\ell, m)$*, tem-se*

$$p_{\boldsymbol{V}|\boldsymbol{U}}(V|U) = |\mathcal{T}(\mathbb{F}_q^{m\times \dim V})|\ p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X), \tag{3.13}$$

em que $X \in \mathbb{F}_q^{n\times\ell}$ *e* $Y \in \mathbb{F}_q^{m\times\ell}$ *são quaisquer matrizes tais que* $\langle X\rangle = U$ *e* $\langle Y\rangle = V$*.*

Demonstração. A partir do Teorema 3.2, sabe-se que $p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$ depende de X e Y apenas através de $\langle X\rangle$ e $\langle Y\rangle$. Portanto, de acordo com o Lema B.3, os mapeamentos $f(\boldsymbol{X}) = \langle \boldsymbol{X}\rangle$ e $g(\boldsymbol{Y}) = \langle \boldsymbol{Y}\rangle$ preservam informação. Isso prova (3.12). Para provar (3.13), primeiro aplica-se o agrupamento de entrada no canal matricial original $(\mathcal{X}, p_{\boldsymbol{Y}|\boldsymbol{X}}, \mathcal{Y})$, obtendo-se um canal intermediário $(\mathcal{U}, p_{\boldsymbol{Y}|\boldsymbol{U}}, \mathcal{Y})$, com $p_{\boldsymbol{Y}|\boldsymbol{U}}(Y|U) = p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$, em que X é tal que $\langle X\rangle = U$. Em seguida, aplica-se o agrupamento de saída nesse canal intermediário, obtendo-se o canal de subespaço $(\mathcal{U}, p_{\boldsymbol{V}|\boldsymbol{U}}, \mathcal{V})$, com

$$\begin{aligned} p_{\boldsymbol{V}|\boldsymbol{U}}(V|U) &= \sum_{Y':\langle Y'\rangle = V} p_{\boldsymbol{Y}|\boldsymbol{U}}(Y'|U) \\ &= |\mathcal{T}(\mathbb{F}_q^{m\times \dim V})|\ p_{\boldsymbol{Y}|\boldsymbol{U}}(Y|U), \end{aligned}$$

em que Y é tal que $\langle Y\rangle = V$. Note que o último passo da equação acima segue de

$$|\{Y' \in \mathbb{F}_q^{m\times\ell} : \langle Y'\rangle = V\}| = |\mathcal{T}(\mathbb{F}_q^{m\times \dim V})|,$$

o que é válido pois, associado a cada $Y' \in \mathbb{F}_q^{m\times\ell}$ tal que $\langle Y'\rangle = V$, existe uma única matriz de posto completo $T \in \mathcal{T}(\mathbb{F}_q^{m\times \dim V})$ tal que $Y' = T\tilde{Y}$, em que $\tilde{Y} \in \mathcal{T}(\mathbb{F}_q^{\dim V\times\ell})$ é qualquer matriz de posto completo tal que $\langle \tilde{Y}\rangle = V$. □

O Teorema 3.7 mostra que, no que diz respeito ao MMC com matriz de transferência u.g.r., os processamentos de entrada $f(X) = \langle X\rangle$ e saída $g(Y) = \langle Y\rangle$ não causam perda de informação. Como consequência, o canal matricial

$$(\mathcal{X} = \mathbb{F}_q^{n\times\ell},\ p_{\boldsymbol{Y}|\boldsymbol{X}},\ \mathcal{Y} = \mathbb{F}_q^{m\times\ell})$$

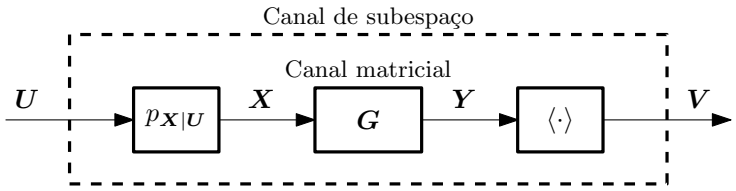


Figura 3.5: Convertendo o canal matricial em um canal de subespaço.

pode ser simplificado para um *canal de subespaço* equivalente

$$(\mathcal{U} = \mathcal{P}(\mathbb{F}_q^\ell, n),\ p_{\boldsymbol{V}|\boldsymbol{U}},\ \mathcal{V} = \mathcal{P}(\mathbb{F}_q^\ell, m))$$

com probabilidade de transição $p_{\boldsymbol{V}|\boldsymbol{U}}$ dada por (3.13). Concretamente, o novo canal é obtido concatenando o canal original (i) na entrada com um dispositivo (possivelmente probabilístico) que transforma um subespaço U em uma matriz X tal que $\langle X\rangle = U$ e (ii) na saída com um dispositivo (determinístico) que calcula $V = \langle Y\rangle$, conforme a Figura 3.5.

Devido a (3.12), qualquer esquema de codificação para o canal matricial possui um correspondente no canal de subespaço, alcançando exatamente a mesma informação mútua, e vice-versa. Em particular, pode-se focar apenas em $(\mathcal{U}, p_{\boldsymbol{V}|\boldsymbol{U}}, \mathcal{V})$ ao projetar e analisar códigos ou esquemas de codificação.

3.10 Demonstrações omitidas

Esta seção apresenta as demonstrações omitidas pelas seções anteriores. De modo a preservar espaço, os subscritos das distribuições de probabilidades poderão ser omitidos [por exemplo, escrevendo-se $p(X)$ no lugar de $p_{\boldsymbol{X}}(X)$]. As demonstrações a seguir farão uso extensivo dos resultados de enumeração da Seção 2.1 do capítulo anterior.

A demonstração do Teorema 3.2 fará uso dos seguintes resultados de contagem. Neste momento, fica registrado o agradecimento a Chen Feng (*University of Toronto*) por sugerir uma prova mais simples do que a originalmente apresentada para o Lema 3.8.

Lema 3.8. *Sejam* $X \in \mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$ *e* $Y \in \mathcal{T}_v(\mathbb{F}_q^{m\times\ell})$. *Defina*

$$\mathcal{H}(r; X, Y) = \{G \in \mathcal{T}_r(\mathbb{F}_q^{m\times n}) : Y = GX\}$$

Então,

$$|\mathcal{H}(r; X, Y)| = \begin{cases} \phi_q(m, n, u, r, v), & \text{se } \langle Y\rangle \subseteq \langle X\rangle, \\ 0, & \text{caso contrário,} \end{cases}$$

em que

$$\phi_q(m, n, u, r, v) = |\mathcal{T}_{r-v}(\mathbb{F}_q^{(m-v)\times(n-u)})|\, q^{v(n-u)}. \tag{3.14}$$

Demonstração. Se $\langle Y\rangle \nsubseteq \langle X\rangle$, então claramente $|\mathcal{H}(r; X, Y)| = 0$. Suponha, então, que $\langle Y\rangle \subseteq \langle X\rangle$. Pelo Lema B.5, existem matrizes inversíveis P, Q, T tais que $X' = PXT$ e $Y' = QYT$. Portanto, $Y = GX$ é equivalente a $QYT = QGP^{-1}PXT$, isto é, $Y' = G'X'$, em que $G' = QGP^{-1}$. Uma vez que o mapeamento $\mathcal{H}(r; X, Y) \to \mathcal{H}(r; X', Y')$ definido por $G \mapsto QGP^{-1}$ é bijetor, tem-se $|\mathcal{H}(r; X, Y)| = |\mathcal{H}(r; X', Y')|$. Em outras palavras, $|\mathcal{H}(r; X, Y)|$ depende de X e Y apenas através de $u = \operatorname{rank} X$ e $v = \operatorname{rank} Y$. Assim, pode-se calcular $|\mathcal{H}(r; X, Y)|$ escolhendo $X = I_u^{n\times\ell}$ e $Y = I_v^{m\times\ell}$. Particionando a matriz G em blocos de acordo com

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \in \mathbb{F}_q^{m\times n},$$

em que os blocos são $G_{11} \in \mathbb{F}_q^{v\times u}$, $G_{12} \in \mathbb{F}_q^{v\times(n-u)}$, $G_{21} \in \mathbb{F}_q^{(m-v)\times u}$ e $G_{22} \in \mathbb{F}_q^{(m-v)\times(n-u)}$, tem-se que as condições $Y = GX$ e $\operatorname{rank} G = r$ são equivalentes a

$$\begin{aligned} &G_{11} = [I_v \;\; 0_{v\times(u-v)}], \\ &G_{12} = \text{qualquer matriz } v \times (n-u), \\ &G_{21} = 0_{(m-v)\times u}, \text{ e} \\ &G_{22} = \text{qualquer matriz } (m-v) \times (n-u) \text{ de posto } r-v, \end{aligned}$$

em que a última exigência segue porque, uma vez que $G_{21} = 0$, tem-se que $\operatorname{rank} G = \operatorname{rank} G_{11} + \operatorname{rank} G_{22}$, ou seja, $\operatorname{rank} G_{22} = r - v$. Das quatro exigências acima, obtém-se o resultado desejado. □

Lema 3.9. *Seja $X \in \mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$. O número de matrizes $Y \in \mathcal{T}_v(\mathbb{F}_q^{m\times\ell})$ tais que $\langle Y\rangle \subseteq \langle X\rangle$ é dado por*

$$|\{Y \in \mathcal{T}_v : \langle Y\rangle \subseteq \langle X\rangle\}| = |\mathcal{T}_v(\mathbb{F}_q^{m\times u})|.$$

Agora, seja $Y \in \mathcal{T}_v(\mathbb{F}_q^{m\times\ell})$. O número de matrizes $X \in \mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$ tais que $\langle Y\rangle \subseteq \langle X\rangle$ é dado por

$$|\{X \in \mathcal{T}_u : \langle Y\rangle \subseteq \langle X\rangle\}| = |\mathcal{T}_v(\mathbb{F}_q^{m\times u})| \frac{|\mathcal{T}_u(\mathbb{F}_q^{n\times\ell})|}{|\mathcal{T}_v(\mathbb{F}_q^{m\times\ell})|}.$$

Demonstração. Para $X \in \mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$, defina o conjunto $\mathcal{J}(X) = \{Y \in \mathcal{T}_v : \langle Y\rangle \subseteq \langle X\rangle\}$. Sejam $X_1, X_2 \in \mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$. Então, existem matrizes $S \in \mathrm{GL}_n(\mathbb{F}_q)$ e $T \in \mathrm{GL}_\ell(\mathbb{F}_q)$ tais que $X_1 = SX_2T$. Uma vez que $Y \mapsto YT^{-1}$ é uma bijeção entre $\mathcal{J}(X_1)$ e $\mathcal{J}(X_2)$, conclui-se que $|\mathcal{J}(X_1)| = |\mathcal{J}(X_2)|$. Assim, para calcular o valor de $|\mathcal{J}(X)|$, pode-se fazer $X = I_u^{n\times\ell}$ sem perda de generalidade. Uma vez que $Y \in \mathcal{J}(I_u^{n\times\ell})$ se e somente se Y é da forma $[Y_0 \;\; 0]$, em que $Y_0 \in \mathcal{T}_v(\mathbb{F}_q^{m\times u})$, conclui-se que $|\mathcal{J}(I_u^{n\times\ell})| = |\mathcal{T}_v(\mathbb{F}_q^{m\times u})|$, como desejado.

Agora, para $Y \in \mathcal{T}_v(\mathbb{F}_q^{m\times\ell})$, defina o conjunto $\mathcal{K}(Y) = \{X \in \mathcal{T}_u : \langle Y\rangle \subseteq \langle X\rangle\}$. Analogamente ao caso anterior, é possível mostrar que $|\mathcal{K}(Y_1)| = |\mathcal{K}(Y_2)|$ para todo $Y_1, Y_2 \in \mathcal{T}_v(\mathbb{F}_q^{m\times\ell})$. Considere portanto um grafo bipartido em que os Xs em $\mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$ são os nós do lado esquerdo e os Ys em $\mathcal{T}_v(\mathbb{F}_q^{m\times\ell})$ são os nós do lado direito; um nó X está conectado com um nó Y se e somente se $\langle Y\rangle \subseteq \langle X\rangle$. O número de ramos conectados ao lado esquerdo, a saber, $|\mathcal{T}_u(\mathbb{F}_q^{n\times\ell})|\,|\mathcal{J}(X)|$, deve ser igual ao número de ramos conectados ao lado direito, a saber, $|\mathcal{T}_v(\mathbb{F}_q^{m\times\ell})|\,|\mathcal{K}(Y)|$. O resultado então segue. □

Demonstração do Teorema 3.2. Sejam $X \in \mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$, $Y \in \mathcal{T}_v(\mathbb{F}_q^{m\times\ell})$ e r tais que $0 \le r \le \min\{n, m\}$. Tem-se

$$\begin{aligned} p(Y|X,r) &= \sum_{G\in\mathcal{T}_r} p(G|r)\, p(Y|X,G) \\ &\overset{(a)}{=} \frac{1}{|\mathcal{T}_r(\mathbb{F}_q^{m\times n})|} \sum_{G\in\mathcal{T}_r} 1[Y = GX] \\ &\overset{(b)}{=} \frac{1}{|\mathcal{T}_r(\mathbb{F}_q^{m\times n})|} \phi_q(m,n,u,r,v)\, 1[\langle Y\rangle \subseteq \langle X\rangle], \end{aligned}$$

em que (a) segue porque $\boldsymbol{G}$ is u.g.r. e (b) segue do Lema 3.8. Portanto, pelo Lema 3.9, pode-se escrever

$$p(v|X,r) = \sum_{Y \in \mathcal{T}_v} p(Y|X,r) = \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|}{|\mathcal{T}_r(\mathbb{F}_q^{m\times n})|}\phi_q(m,n,u,r,v),$$

de modo que

$$\begin{aligned} p(v|u,r) &= \sum_{X\in\mathcal{T}_u} p(X|u)\, p(v|X,r) \\ &= \sum_{X\in\mathcal{T}_u} p(X|u)\frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|}{|\mathcal{T}_r(\mathbb{F}_q^{m\times n})|}\phi_q(m,n,u,r,v) \\ &= \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|}{|\mathcal{T}_r(\mathbb{F}_q^{m\times n})|}\phi_q(m,n,u,r,v), \end{aligned}$$

e (3.6) segue ao se comparar as expressões de $p(Y|X,r)$ e $p(v|u,r)$. A equação em (3.5) segue da expressão acima, substituindo $\phi_q(m,n,u,r,v)$ pela sua definição em (3.14), e por simplificação, fazendo uso das definições e expressões da Seção 2.1.

Para finalizar a prova, suponha que $\boldsymbol{X}$ seja u.g.r. Então, para cada $Y \in \mathcal{T}_v(\mathbb{F}_q^{m\times \ell})$, tem-se

$$\begin{aligned} p(Y) &= \sum_u \sum_{X\in\mathcal{T}_u} p(Y|X)p(X) \\ &\overset{(a)}{=} \sum_u \frac{p(u)}{|\mathcal{T}_u(\mathbb{F}_q^{n\times\ell})|} \sum_{X\in\mathcal{T}_u} p(Y|X) \\ &\overset{(b)}{=} \sum_u \frac{p(u)}{|\mathcal{T}_u(\mathbb{F}_q^{n\times\ell})|}\frac{p(v|u)}{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|} \sum_{X\in\mathcal{T}_u} 1[\langle Y\rangle \subseteq \langle X\rangle] \\ &\overset{(c)}{=} \sum_u \frac{p(u)}{|\mathcal{T}_u(\mathbb{F}_q^{n\times\ell})|}\frac{p(v|u)}{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|}|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|\frac{|\mathcal{T}_u(\mathbb{F}_q^{n\times\ell})|}{|\mathcal{T}_v(\mathbb{F}_q^{m\times\ell})|} \\ &= \frac{p(v)}{|\mathcal{T}_v(\mathbb{F}_q^{m\times\ell})|}, \end{aligned}$$

em que (a) segue porque $\boldsymbol{X}$ é u.g.r., (b) segue de (3.6) e (c) segue do Lema 3.9. Portanto, $\boldsymbol{Y}$ também será u.g.r., como afirmado. □

Demonstração do Teorema 3.3. Para cada $X \in \mathcal{T}_u(\mathbb{F}_q^{n\times\ell})$, tem-se

$$\begin{aligned} H(\boldsymbol{Y}|\boldsymbol{X}=X) &= \sum_v \sum_{Y\in\mathcal{T}_v} p(Y|X)\log_q \frac{1}{p(Y|X)} \\ &= \sum_v p(v|u)\log_q \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|}{p(v|u)} \\ &= h_u, \end{aligned}$$

em que $p(Y|X)$ foi substituído como em (3.6). Tirando a média sobre todo $X \in \mathbb{F}_q^{n\times\ell}$, obtém-se

$$\begin{aligned} H(\boldsymbol{Y}|\boldsymbol{X}) &= \sum_u \sum_{X\in\mathcal{T}_u} H(\boldsymbol{Y}|\boldsymbol{X}=X)p(X) \\ &= \sum_u h_u \sum_{X\in\mathcal{T}_u} p(X) \\ &= \sum_u h_u p(u), \end{aligned}$$

o qual depende de $p_{\boldsymbol{X}}$ apenas através de $p_{\boldsymbol{u}}$. Portanto,

$$\begin{aligned} I^*(p_{\boldsymbol{u}}) &= \max_{p_{\boldsymbol{X}}:p_{\boldsymbol{u}}} I(\boldsymbol{X};\boldsymbol{Y}) \\ &= \max_{p_{\boldsymbol{X}}:p_{\boldsymbol{u}}} [H(\boldsymbol{Y}) - H(\boldsymbol{Y}|\boldsymbol{X})] \\ &= \max_{p_{\boldsymbol{X}}:p_{\boldsymbol{u}}} H(\boldsymbol{Y}) - \sum_u h_u p(u), \end{aligned}$$

e obtém-se o resultado desejado de (3.1). □

Demonstração do Teorema 3.4. Se a entrada é restrita a matrizes de posto u, então $\boldsymbol{u} = u$ é uma constante e, portanto, $p(v) = p(v|u)$. A informação mútua do canal dada pelo Teorema 3.3 simplifica para

$$\sum_v p(v|u)\log_q \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times\ell})|}{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|},$$

e obtém-se (3.10) ao aplicar (2.4). O limitante inferior de (3.11) é imediato. Analogamente a Yang et al. [55, Lemma 4], pode-se reescrever a

informação mútua do canal em (3.8) como

$$\begin{aligned} I^*(p_{\boldsymbol{u}}) &= \sum_v p(v) \log_q \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times \ell})|}{p(v)} - \sum_u p(u) h_u \\ &= \sum_{u,v} p(u)p(v|u) \log_q \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times \ell})|}{p(v)} - \sum_{u,v} p(u)p(v|u) \log_q \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|}{p(v|u)} \\ &= \sum_{u,v} p(u)p(v|u) \log_q \frac{|\mathcal{T}_v(\mathbb{F}_q^{m\times \ell})|}{|\mathcal{T}_v(\mathbb{F}_q^{m\times u})|} + \sum_{u,v} p(u)p(v|u) \log_q \frac{p(v|u)}{p(v)} \\ &= \sum_u p(u) C_u + I(\boldsymbol{u};\boldsymbol{v}), \end{aligned}$$

em que $I(\boldsymbol{u};\boldsymbol{v})$ é a informação mútua entre as variáveis aleatórias $\boldsymbol{u}$ e $\boldsymbol{v}$. O limitante superior de (3.11) então segue, pois $\sum_u p(u)C_u \leq \max_u C_u = C_{u^*}$ e $I(\boldsymbol{u};\boldsymbol{v}) \leq \log_q(\min\{n,m\}+1)$. □

Demonstração do Teorema 3.5. Dividindo (3.11) por ℓ e tomando o limite quando $\ell \to \infty$, tem-se

$$\lim_{\ell\to\infty} \bar{C} = \lim_{\ell\to\infty} \bar{C}_{u^*},$$

de modo que entrada de posto constante é suficiente para alcançar a capacidade com ℓ arbitrariamente grande.

Agora, dividindo (3.10) por ℓ e tomando o limite quando $\ell \to \infty$, obtém-se

$$\begin{aligned} \lim_{\ell\to\infty} \bar{C}_u &= \sum_v p(v|u) \left(\lim_{\ell\to\infty} \frac{1}{\ell} \log_q \frac{\begin{bmatrix}\ell\\ v\end{bmatrix}_q}{\begin{bmatrix}u\\ v\end{bmatrix}_q} \right) \\ &= \sum_v p(v|u) \left(\lim_{\ell\to\infty} \frac{1}{\ell} \log_q \begin{bmatrix}\ell\\ v\end{bmatrix}_q - \lim_{\ell\to\infty} \frac{1}{\ell} \log_q \begin{bmatrix}u\\ v\end{bmatrix}_q \right) \\ &= \sum_v p(v|u) \left(\lim_{\ell\to\infty} \frac{1}{\ell} \log_q \begin{bmatrix}\ell\\ v\end{bmatrix}_q \right) \\ &= \sum_v v\, p(v|u) = \mathrm{E}[\boldsymbol{v}|\boldsymbol{u}=u], \end{aligned}$$

em que a primeira igualdade da última linha é uma consequência de (2.2).

Finalmente, como $\boldsymbol{v} \leq \boldsymbol{r}$, tem-se

$$\mathrm{E}[\boldsymbol{v}|\boldsymbol{u} = u, \boldsymbol{r} = r] \leq r = \mathrm{E}[\boldsymbol{v}|\boldsymbol{u} = n, \boldsymbol{r} = r],$$

para todo $u \in \{0, \ldots, n\}$. Multiplicando os dois lados por $p(r)$ e somando sobre r, obtém-se

$$\mathrm{E}[\boldsymbol{v}|\boldsymbol{u} = u] \leq \mathrm{E}[\boldsymbol{r}] = \mathrm{E}[\boldsymbol{v}|\boldsymbol{u} = n],$$

o que mostra que $\lim_{\ell\to\infty} \bar{C}_u = \mathrm{E}[\boldsymbol{v}|\boldsymbol{u} = u]$ é máximo quando $u = n$, sendo o máximo valor dado por $\mathrm{E}[\boldsymbol{r}]$. □

Para a demonstração do Teorema 3.6, será necessário o seguinte resultado intuitivo.

Lema 3.10. *Tem-se*

$$\lim_{q\to\infty} p(v|u, r) = \begin{cases} 1, & \text{se } v = \min\{u, r\}, \\ 0, & \text{caso contrário.} \end{cases}$$

Demonstração. O teorema é trivialmente verdadeiro se $v > \min\{u, r\}$. Por outro lado, se $v \leq \min\{u, r\}$, tem-se de (2.2) e do Teorema 3.2 que

$$\begin{aligned} q^{v(u-v)} \cdot \gamma_q^{-1} q^{-r(n-r)} \cdot q^{(r-v)(n-u-r+v)} \cdot q^{v(n-u-r+v)} \\ \leq p(v|u, r) = \begin{bmatrix} u \\ v \end{bmatrix}_q \begin{bmatrix} n \\ r \end{bmatrix}_q^{-1} \begin{bmatrix} n-u \\ r-v \end{bmatrix}_q q^{v(n-u-r+v)} \leq \\ \gamma_q q^{v(u-v)} \cdot q^{-r(n-r)} \cdot \gamma_q q^{(r-v)(n-u-r+v)} \cdot q^{v(n-u-r+v)}. \end{aligned}$$

Após simplificação, obtém-se

$$\gamma_q^{-1} q^{-(u-v)(r-v)} \leq p(v|u, r) \leq \gamma_q^2 q^{-(u-v)(r-v)},$$

e o resultado desejado segue porque $\lim_{q\to\infty} \gamma_q = 1$. □

Demonstração do Teorema 3.6. A quantidade $\log_q(\min\{n, m\} + 1)$ no lado direito de (3.11) tende a zero quando $q \to \infty$, de modo que

$$\lim_{q\to\infty} C = \lim_{q\to\infty} C_{u^*},$$

ou seja, entrada de posto constante é suficiente para q assintoticamente grande.

Agora, de (3.10), tem-se

$$\lim_{q\to\infty} C_u = \sum_v \left(\lim_{q\to\infty} p(v|u)\right)\left(\lim_{q\to\infty} \log_q \frac{\begin{bmatrix}\ell\\ v\end{bmatrix}_q}{\begin{bmatrix}u\\ v\end{bmatrix}_q}\right)$$

Para o primeiro parênteses, tem-se do Lema 3.10 que

$$\lim_{q\to\infty} p(v|u) = \sum_{r=0}^{n} p_{\boldsymbol{r}}^{\infty}(r)1[v = \min\{u, r\}].$$

Para o segundo parênteses, tem-se de (2.2) que

$$\lim_{q\to\infty} \log_q \frac{\begin{bmatrix}\ell\\ v\end{bmatrix}_q}{\begin{bmatrix}u\\ v\end{bmatrix}_q} = v(\ell - u).$$

Portanto,

$$\begin{aligned}\lim_{q\to\infty} C_u &= \sum_v \sum_r p_{\boldsymbol{r}}^{\infty}(r)1[v = \min\{u, r\}]v(\ell - u)\\ &= (\ell - u)\sum_r p_{\boldsymbol{r}}^{\infty}(r)\sum_v 1[v = \min\{u, r\}]\ v\\ &= (\ell - u)\sum_r p_{\boldsymbol{r}}^{\infty}(r)\min\{u, r\},\end{aligned}$$

como desejado. □

CAPÍTULO 4

MMC coerente sobre anéis de cadeia finitos

4.1 Introdução

O presente capítulo considera MMCs sobre *anéis de cadeia finitos* (dos quais corpos finitos são um caso particular). A motivação vem de *codificação de rede na camada física*: como visto no Capítulo 1, o canal de comunicação fim-a-fim entre um nó fonte e um nó destino de uma rede sem-fio que emprega codificação de rede na camada física ainda pode ser modelado pela expressão

$$\boldsymbol{Y} = \boldsymbol{G}\boldsymbol{X},$$

em que $\boldsymbol{X} \in R^{n\times\ell}$ é a matriz de entrada, cujas linhas são os n pacotes enviados pelo nó fonte, $\boldsymbol{Y} \in R^{m\times\ell}$ é a matriz de saída, cujas linhas são os m pacotes recebidos pelo nó destino e $\boldsymbol{G} \in R^{m\times n}$ é a matriz de transferência. Nesse caso, as matrizes em questão possuem entradas em um certo anel finito R (induzido pela constelação utilizada), que

Este capítulo é um trabalho conjunto com Chen Feng (*University of Toronto*). O conteúdo deste capítulo foi apresentado no *2013 IEEE International Symposium on Network Coding (NetCod'13)* [40] e no *XXXI Simpósio Brasileiro de Telecomunicações (SBrT'13)* [39].

não necessariamente é um corpo finito.

Na verdade, a discussão apresentada no Capítulo 1 sobre codificação de rede na camada física considerou, por simplicidade, apenas o uso de *modulação não-codificada*, na qual os sinais físicos transmitidos pelos nós podem ser qualquer sequência de ℓ símbolos da constelação. Na camada superior, essa hipótese induziu um "espaço de pacotes" dado por $W = R^\ell$. No entanto, como mostrado recentemente por Nazer e Gastpar [35], o emprego de *modulação codificada* em codificação de rede na camada física, através de *reticulados aninhados*, apresenta diversas vantagens em termos do compromisso entre taxa alcançada e probabilidade de erro. Nesse caso, de acordo com Feng, Silva e Kschischang [12] (que apresentam uma extensão da teoria em [35]), o espaço de pacotes W passa a ser um *módulo finito* sobre um *domínio de ideais principais* T (tipicamente os inteiros, $\mathbb{Z}$, os inteiros gaussianos, $\mathbb{Z}[i]$, ou os inteiros de Eisenstein, $\mathbb{Z}[\omega]$)[1]. Como tal, sabe-se que

$$W \cong T/\langle d_1\rangle \times T/\langle d_2\rangle \times \cdots \times T/\langle d_\ell\rangle,$$

em que $d_1, d_2, \ldots, d_\ell \in T$ são elementos não-nulos e não-inversíveis satisfazendo $d_1 \mid d_2 \mid \cdots \mid d_\ell$ e $\langle d_i\rangle$ representa o ideal gerado pelo elemento d_i. Uma situação particular frequentemente encontrada na prática é aquela na qual os d_is são todos potências de um dado primo de T. Nesse caso, não é difícil mostrar que o espaço de pacotes W também pode ser visto como um R-*módulo finito*, em que $R = T/\langle d_\ell\rangle$ é um *anel de cadeia finito*.

Este capítulo considera MMCs sobre anéis de cadeia finitos nos quais o espaço de pacotes é um módulo finito qualquer. Por simplicidade, é assumido o cenário coerente, em que as instâncias da matriz de transferência $\boldsymbol{G}$ são desconhecidas do transmissor, mas disponíveis ao receptor. Fora isso, não é imposta qualquer restrição às estatísticas de $\boldsymbol{G}$, exceto que essa deve ser independente de $\boldsymbol{X}$.

[1] Módulos são estruturas algébricas que generalizam o conceito de espaço vetorial, de corpos para anéis: se R é um anel comutativo, então um "módulo sobre R" (ou, mais comumente, "R-módulo") é um conjunto no qual está definida a soma de elementos e a multiplicação de elementos por um escalar (isto é, um elemento do anel R), satisfazendo as mesmas propriedades exigidas para espaços vetoriais. O Apêndice A apresenta uma revisão da teoria básica de anéis comutativos e módulos sobre tais anéis.

Inicia-se revisando alguns conceitos básicos acerca de anéis de cadeia finitos (§4.2) e de álgebra linear sobre tais anéis (§4.3). O modelo do canal é então apresentado (§4.4). Como contribuições, é obtida uma expressão fechada para a capacidade do canal (§4.5) e é proposto um esquema de codificação capaz de alcançar essa capacidade em tempo polinomial (§4.6). O esquema combina uma sequência de códigos sobre um corpo finito para obter um código sobre o anel de cadeia finito e é baseado na *expansão π-ádica* de elementos do anel.

Os resultados aqui apresentados estendem (e fazem uso de) alguns daqueles obtidos por Yang et al. [55], o qual lida com o caso de corpos finitos. Vale também mencionar que uma generalização dos resultados de Silva, Kschischang e Kötter [49], de corpos para anéis de cadeia finitos, é apresentada em [11, 10].

4.2 Anéis de cadeia finitos

Apresentam-se aqui alguns conceitos sobre anéis de cadeia finitos. No decorrer deste trabalho, subentende-se pelo termo *anel* um anel comutativo com identidade aditiva 0 e identidade multiplicativa 1, em que $1 \neq 0$. Para mais detalhes, encaminha-se o leitor para [31] ou [37].

Um anel R é dito ser um *anel de cadeia* se, para quaisquer ideais I, J de R, tem-se $I \subseteq J$ ou $J \subseteq I$ (em outras palavras, os ideais formam uma cadeia quando ordenados com respeito a inclusão de conjuntos, $\subseteq$). Uma vez que corpos possuem apenas ideais triviais (isto é, $\{0\}$ e R), todo corpo é um anel de cadeia.

Proposição 4.1. *Seja R um anel finito. Então, R é um anel de cadeia se e somente se R for simultaneamente um* anel de ideais principais *(isto é, um anel no qual todos os ideais são gerados por um único elemento) e um* anel local *(isto é, um anel com um único ideal maximal).*

No que segue, $\langle x \rangle$ denota o ideal gerado por $x \in R$

Exemplo. Existem onze anéis não-isomorfos de ordem 4, incluindo os não-comutativos e os sem identidade multiplicativa [38, 7]. Quatro

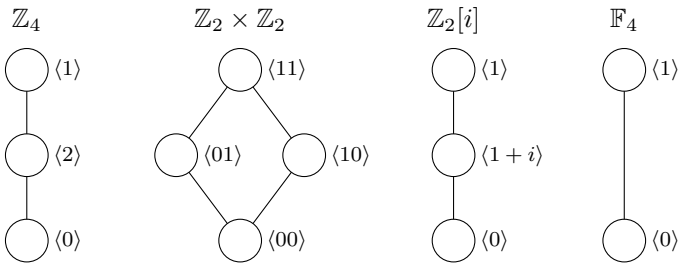


Figura 4.1: Ideais dos anéis comutativos com identidade de ordem 4.

deles são comutativos com identidade multiplicativa, a saber,

$$\begin{aligned} \mathbb{Z}_4 &= \{0,1,2,3\}, \\ \mathbb{Z}_2 \times \mathbb{Z}_2 &= \{(0,0),(0,1),(1,0),(1,1)\}, \\ \mathbb{Z}_2[i] &= \{0,1,i,1+i\}, \\ \mathbb{F}_4 &= \{0,1,\alpha,1+\alpha\}, \end{aligned}$$

em que, em $\mathbb{Z}_2 \times \mathbb{Z}_2$ a adição e a multiplicação são efetuadas entrada-a-entrada, em $\mathbb{Z}_2[i]$ vale $i^2 = 1$ e em $\mathbb{F}_4$ vale $\alpha^2 = 1+\alpha$. Todos eles são anéis de ideais principais, em que

- Em $\mathbb{Z}_4$:

  $\langle 0\rangle = \{0\}, \quad \langle 2\rangle = \{0,2\}, \quad \langle 1\rangle = \langle 3\rangle = \mathbb{Z}_4.$

- Em $\mathbb{Z}_2 \times \mathbb{Z}_2$:

  $\langle 00\rangle = \{00\}, \quad \langle 01\rangle = \{00,01\}, \quad \langle 10\rangle = \{00,10\}, \quad \langle 11\rangle = \mathbb{Z}_2 \times \mathbb{Z}_2.$

- Em $\mathbb{Z}_2[i]$:

  $\langle 0\rangle = \{0\}, \quad \langle 1+i\rangle = \{0,1+i\}, \quad \langle 1\rangle = \langle i\rangle = \mathbb{Z}_2[i].$

- Em $\mathbb{F}_4$:

  $\langle 0\rangle = \{0\}, \quad \langle 1\rangle = \langle \alpha\rangle = \langle 1+\alpha\rangle = \mathbb{F}_4,$

como ilustrado pela Figura 4.1. Da figura, fica claro que $\mathbb{Z}_4$, $\mathbb{Z}_2[i]$ e $\mathbb{F}_4$ são anéis de cadeia, ao passo que $\mathbb{Z}_2 \times \mathbb{Z}_2$ não é. □

Proposição 4.2. *Seja $\pi \in R$ um gerador do ideal máximo de R e seja s o índice de nilpotência de π (isto é, o menor inteiro tal que $\pi^s = 0$). Então, R possui precisamente $s+1$ ideais, a saber,*

$$R = \langle 1 \rangle \supset \langle \pi \rangle \supset \cdots \supset \langle \pi^{s-1} \rangle \supset \langle \pi^s \rangle = \{0\}.$$

Neste trabalho, o parâmetro s é chamado de *profundidade* de R. Adicionalmente, sabe-se que o quociente $R/\langle \pi \rangle$ é um corpo finito, chamado de *corpo residual* de R. Pode-se mostrar que, se q for a ordem do corpo residual de R, então o tamanho de cada ideal de R é $|\langle \pi^i \rangle| = q^{s-i}$, para $0 \leq i \leq s$. Em particular, $|R| = q^s$, de modo que, analogamente ao que ocorre com corpos finitos, o tamanho de um anel de cadeia finito é sempre uma potência de primo. Note que $s = 1$ (isto é, o anel possui apenas os dois ideias triviais) se e somente se R for um corpo finito.

Exemplo. Considere $\mathbb{Z}_8 = \{0, 1, \ldots, 7\}$, o anel dos inteiros modulo 8. Seus ideais são $\langle 1 \rangle = \mathbb{Z}_8$, $\langle 2 \rangle = \{0, 2, 4, 6\}$, $\langle 4 \rangle = \{0, 4\}$ e $\langle 0 \rangle = \{0\}$ (de modo que $s = 3$) e seu corpo residual é $\mathbb{Z}_8/\langle 2 \rangle \cong \mathbb{F}_2$ (de modo que $q = 2$). De maneira mais geral, se p é um número primo e s é um inteiro positivo, então $\mathbb{Z}_{p^s}$ é um anel de cadeia de profundidade s e corpo residual de ordem p. □

Durante o restante deste capítulo, R denota um anel de cadeia com profundidade s e corpo residual de ordem q. Além disso, $\pi \in R$ denota um gerador para o ideal principal de R e $\Gamma \subseteq R$ denota um conjunto fixo de representantes de classes laterais (*cosets*) do corpo residual de R. Sem perda de generalidade, assume-se que $0 \in \Gamma$.

Proposição 4.3. *Todo $x \in R$ pode ser escrito unicamente como*

$$x = \sum_{i=0}^{s-1} x^{(i)} \pi^i,$$

em que $x^{(i)} \in \Gamma$, para $0 \leq i < s$.

A expressão acima é conhecida como *expansão π-ádica de x* (com relação a Γ).

Exemplo. A expansão 2-ádica de $6 \in \mathbb{Z}_8$ com relação a $\Gamma = \{0, 1\}$ é $6 = 0 \cdot 2^0 + 1 \cdot 2^1 + 1 \cdot 2^2$, isto é, a expansão binária usual de 6. □

Note que a unicidade da expansão π-ádica (fixado Γ) permite que sejam definidos os mapeamentos $(\cdot)^{(i)} : R \to \Gamma$, para $0 \leq i < s$. Também define-se

$$x^{\underline{i}} = \sum_{j=0}^{i-1} x^{(j)} \pi^j,$$

para $0 \leq i \leq s$. Pode-se mostrar que

$$x^{\underline{i}} \equiv_{\pi^i} x,$$

para todo $x \in R$, em que $\equiv_a$ denota congruência módulo a (isto é, $x \equiv_a y$ se e somente se $x - y \in \langle a \rangle$). Em particular, $x^{(0)} = x^{\underline{1}} \equiv_\pi x$.

4.3 Álgebra linear sobre anéis de cadeia finitos

Esta seção apresenta alguns resultados fundamentais de álgebra linear sobre anéis de cadeia finitos. Os trabalhos [20] e [4] servem como referência.

4.3.1 Módulos sobre anéis de cadeia finitos

Um *s-shape* é uma lista não-decrescente de s inteiros não-negativos. Seja $\mu = (\mu_0, \ldots, \mu_{s-1})$ um *s-shape*. Define-se

$$R^\mu \triangleq \underbrace{\langle 1 \rangle \times \cdots \times \langle 1 \rangle}_{\mu_0} \times \underbrace{\langle \pi \rangle \times \cdots \times \langle \pi \rangle}_{\mu_1 - \mu_0} \times \cdots \times \underbrace{\langle \pi^{s-1} \rangle \times \cdots \times \langle \pi^{s-1} \rangle}_{\mu_{s-1} - \mu_{s-2}}.$$

Sendo um produto cartesiano de ideais de R, sabe-se que R^μ é um R-módulo. Reciprocamente, tem-se o seguinte.

Proposição 4.4. *Seja M um R-módulo finito. Então,*

$$M \cong R^\mu$$

para um s-shape único μ.

Diz-se então que μ é o *shape* de M, escrevendo-se $\mu = \operatorname{shape} M$.

Observação: No caso de espaços vetoriais, é bem conhecido que, se M é um espaço vetorial finito sobre um corpo finito $\mathbb{F}_q$, então $M \cong \mathbb{F}_q^m$, em que $m = \dim M$ é a *dimensão* de M. Dessa forma, o conceito de

shape de um R-módulo M generaliza o conceito de dimensão de um espaço vetorial. (No caso em que R é um corpo finito, então $s = 1$ e um *s-shape* é simplesmente um único inteiro.)

Proposição 4.5. *Seja μ um s-shape. Então,*

$$|R^\mu| = q^{|\mu|},$$

em que $|\mu|$ é definido como $|\mu| = \mu_0 + \mu_1 + \cdots + \mu_{s-1}$.

Exemplo. Seja $R = \mathbb{Z}_8$ (de modo que $s = 3$) e $\mu = (2, 3, 4)$. Então,

$$\begin{aligned} R^\mu &= \langle 1\rangle \times \langle 1\rangle \times \langle 2\rangle \times \langle 4\rangle \\ &= \{(x_1, x_2, 2x_3, 4x_4) : x_1, x_2, x_3, x_4 \in R\}. \end{aligned}$$

Tem-se $|R^\mu| = 8 \cdot 8 \cdot 4 \cdot 2 = 512 = 2^{2+3+4} = q^{|\mu|}$. □

Seja $\mu = (\mu_0, \mu_1, \ldots, \mu_{s-1})$ um *s-shape*. Define-se $\mu - n = (\mu_0 - n, \mu_1 - n, \ldots, \mu_{s-1} - n)$, que também é um *s-shape*. Além disso, por conveniência, escreve-se o *s-shape* $(m, m, \ldots, m)$ simplesmente como m. De acordo com essa convenção, R^m representa o mesmo objeto, seja m interpretado como um inteiro ou como um *s-shape*.

4.3.2 Matrizes sobre anéis de cadeia finitos

Para qualquer subconjunto $S \subseteq R$, denota-se por $S^{m\times n}$ o conjunto de todas as matrizes $m \times n$ com entradas em S. O conjunto de todas as matrizes $n \times n$ inversíveis sobre R é chamado de *grupo linear geral* de grau n sobre R, denotado por $\mathrm{GL}_n(R)$.

Seja $A \in R^{m\times n}$ e defina $r = \min\{n, m\}$. Uma matriz diagonal (não necessariamente quadrada)

$$D = \mathrm{diag}(d_1, d_2, \ldots, d_r) \in R^{m\times n}$$

é dita ser uma *forma normal de Smith* de A se existirem matrizes $P \in \mathrm{GL}_m(R)$ e $Q \in \mathrm{GL}_n(R)$ tais que $A = PDQ$ e $d_1 \mid d_2 \mid \cdots \mid d_r$. É sabido que matrizes sobre anéis de ideais principais (em particular, anéis de cadeia finitos) sempre possuem forma normal de Smith, a qual é única a menos de multiplicação das entradas da diagonal por elementos inversíveis do anel. Neste trabalho, será exigido que as entradas da

diagonal sejam potências de $\pi \in R$; por conseguinte, a forma normal de Smith torna-se de fato única.

Sejam $\operatorname{row} A$ e $\operatorname{col} A$ os espaços linha e coluna, respectivamente, de $A \in R^{m\times n}$. Utilizando a forma normal de Smith, pode-se mostrar que $\operatorname{row} A$ é isomorfo a $\operatorname{col} A$. Define-se o *shape* da matriz A como $\operatorname{shape} A = \operatorname{shape}(\operatorname{row} A) = \operatorname{shape}(\operatorname{col} A)$.

Observação: Assim como $\operatorname{shape} M$ generaliza a dimensão de um R-módulo M, tem-se que $\operatorname{shape} A$ generaliza o conceito de *posto* de uma matriz A com entradas em R.

Proposição 4.6. *Sejam $A \in R^{m\times n}$. Então, $\mu = \operatorname{shape} A$ se e somente se a forma normal de Smith de A for dada por*

$$\operatorname{diag}(\underbrace{1,\ldots,1}_{\mu_0},\underbrace{\pi,\ldots,\pi}_{\mu_1-\mu_0},\ldots,\underbrace{\pi^{s-1},\ldots,\pi^{s-1}}_{\mu_{s-1}-\mu_{s-2}},\underbrace{0,\ldots,0}_{r-\mu_{s-1}}),$$

em que $r = \min\{n, m\}$.

Exemplo. Considere a matriz

$$A = \begin{bmatrix} 4 & 3 & 6 \\ 6 & 7 & 2 \end{bmatrix}$$

sobre $\mathbb{Z}_8$. Então, $A = PDQ$, em que

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 4 & 3 & 6 \\ 1 & 2 & 6 \\ 5 & 6 & 3 \end{bmatrix},$$

de modo que $\operatorname{shape} A = \operatorname{shape} D = (1, 2, 2)$. □

4.3.3 Matrizes com restrições nas linhas

Seja λ um *s-shape*. Denota-se por $R^{n\times\lambda}$ o subconjunto de matrizes em $R^{n\times\ell}$ cujas linhas são elementos de R^{λ}, em que $\ell = \lambda_{s-1}$. De acordo com a Proposição 4.5, tem-se que $|R^{n\times\lambda}| = q^{n|\lambda|}$.

EXEMPLO. Seja $R = \mathbb{Z}_8$. Em adição, sejam $n = 2$ e $\lambda = (1, 2, 3)$, de modo que $\ell = 3$. Então,

$$R^{n\times\lambda} = \left\{ \begin{bmatrix} x_{11} & 2x_{12} & 4x_{13} \\ x_{21} & 2x_{22} & 4x_{23} \end{bmatrix} : x_{i,j} \in R \right\} \subseteq R^{n\times\ell}.$$

Note que a matriz A do exemplo anterior não pertence a $R^{n\times\lambda}$, ao passo que a matriz D pertence. □

Finalmente, estende-se os mapeamentos $(\cdot)^{(i)}$ de elementos de R para matrizes e vetores com entradas em R, entrada-a-entrada. Assim, $A \in R^{n\times\lambda}$ se e somente se $A^{(i)} = [B_i \;\; 0] \in \Gamma^{n\times\ell}$, para algum $B_i \in \Gamma^{n\times\lambda_i}$, $0 \leq i < s$.

4.4 Modelo do canal

Seja R um anel de cadeia finito com profundidade s e corpo residual de ordem q. Como dito anteriormente, este capítulo estuda MMCs coerentes sobre R nos quais o espaço de pacotes W é um R-módulo finito qualquer. Assim, tendo em vista a Proposição 4.4, pode-se assumir que $W = R^\lambda$ para algum *s-shape* λ. Como consequência, a matriz de entrada $\boldsymbol{X}$ (cujas linhas são n pacotes) e a matriz de saída $\boldsymbol{Y}$ (cujas linhas são m pacotes) podem ser interpretadas como matrizes com restrições nas linhas, isto é, $\boldsymbol{X} \in R^{n\times\lambda}$ e $\boldsymbol{Y} \in R^{m\times\lambda}$.

Seja $p_{\boldsymbol{G}}$ uma distribuição de probabilidade sobre $R^{m\times n}$. De forma análoga à Seção 2.3, define-se o MMC coerente sobre R como um DMC com alfabeto de entrada $\mathcal{X} = R^{n\times\lambda}$, alfabeto de saída $\mathcal{Y} = R^{m\times\lambda} \times R^{m\times n}$ e probabilidade de transição $p_{\boldsymbol{Y},\boldsymbol{G}|\boldsymbol{X}}$, a qual é induzida pela distribuição $p_{\boldsymbol{G}}$. Denota-se o canal em questão por C-MMC$(\boldsymbol{G}, \lambda)$. Também faz-se uso da variável aleatória

$$\boldsymbol{\rho} = \operatorname{shape} \boldsymbol{G},$$

distribuída de acordo com

$$p_{\boldsymbol{\rho}}(\rho) = \sum_{G:\, \operatorname{shape} G = \rho} p_{\boldsymbol{G}}(G).$$

Por fim, é definido $\ell = \lambda_{s-1}$ (comprimento do pacote).

4.5 Capacidade do canal

Inicia-se determinando a capacidade do MMC coerente sobre anéis de cadeia finitos. O seguinte resultado generaliza o Teorema 2.3.

Teorema 4.7. *A capacidade de* C-MMC($\boldsymbol{G}, \lambda$)*, em símbolos q-ários por uso do canal, é dada por*

$$C = \sum_{i=0}^{s-1} \mathrm{E}[\boldsymbol{\rho}_{s-i-1}]\lambda_i,$$

sendo alcançada se a entrada é distribuída uniformemente sobre $R^{n\times\lambda}$*. Em particular, a capacidade depende de* $p_{\boldsymbol{G}}$ *apenas através de* $\mathrm{E}[\boldsymbol{\rho}]$*.*

A prova faz uso do seguinte lema.

Lema 4.8. *Sejam* $\boldsymbol{X} \in R^{n\times\lambda}$ *uma matriz aleatória,* $G \in R^{m\times n}$ *uma matriz determinística e* $\rho = \operatorname{shape} G$*. Defina* $\boldsymbol{Y} = G\boldsymbol{X} \in R^{m\times\lambda}$*. Então,*

$$H(\boldsymbol{Y}) \leq \sum_{i=0}^{s-1} \rho_{s-i-1}\lambda_i,$$

com igualdade se $\boldsymbol{X}$ *for uniformemente distribuída sobre* $R^{n\times\lambda}$*.*

Demonstração. Note que $\boldsymbol{X}$ e $\boldsymbol{Y}$ podem ser expressos como

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_0 & \boldsymbol{X}_1 & \cdots & \boldsymbol{X}_{s-1} \end{bmatrix},$$
$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_0 & \boldsymbol{Y}_1 & \cdots & \boldsymbol{Y}_{s-1} \end{bmatrix},$$

em que $\boldsymbol{X}_i \in \langle\pi^i\rangle^{n\times(\lambda_i-\lambda_{i-1})}$ e $\boldsymbol{Y}_i \in \langle\pi^i\rangle^{m\times(\lambda_i-\lambda_{i-1})}$, para $0 \leq i < s$. Tem-se

$$\boldsymbol{Y}_i = G\boldsymbol{X}_i,$$

de modo que o suporte de cada coluna da matriz $\boldsymbol{Y}_i$ é um subconjunto de $\operatorname{col}\pi^i G$. Uma vez que $\operatorname{shape}\pi^i G = (0,\ldots,0,\rho_0,\ldots,\rho_{s-i-1})$, tem-se, a partir da Proposição 4.5, que $|\operatorname{col}\pi^i G| = q^{\rho_0+\cdots+\rho_{s-i-1}}$. Portanto, o

suporte de $\boldsymbol{Y}$ tem tamanho de no máximo

$$\begin{aligned}\prod_{i=0}^{s-1} |\operatorname{col} \pi^i G|^{\lambda_i - \lambda_{i-1}} &= \prod_{i=0}^{s-1} q^{(\rho_0 + \cdots + \rho_{s-i-1})(\lambda_i - \lambda_{i-1})} \\ &= q^{\sum_{i=0}^{s-1} \rho_{s-i-1} \lambda_i},\end{aligned}$$

de onde a desigualdade segue.

Suponha agora que a matriz $\boldsymbol{X}$ seja uniformemente distribuída sobre $R^{n \times \lambda}$. Isso significa que cada bloco $\boldsymbol{X}_i$ é uniformemente distribuído sobre $\langle \pi^i \rangle^{n \times (\lambda_i - \lambda_{i-1})}$. É possível mostrar que existe $\boldsymbol{X}'_i$ uniformemente distribuído sobre $R^{n \times (\lambda_i - \lambda_{i-1})}$ tal que $\boldsymbol{X}_i = \pi^i \boldsymbol{X}'_i$. Seja $\boldsymbol{y}$ a coluna de $\boldsymbol{Y}_i$, cujo suporte é $\operatorname{col} \pi^i G$. Uma vez que $\boldsymbol{Y}_i = G\boldsymbol{X}_i = \pi^i G \boldsymbol{X}'_i$, tem-se, para todo $y \in \operatorname{col} \pi^i G$, que

$$\begin{aligned}\Pr[\boldsymbol{y} = y] &= \frac{|\{x' \in R^n : \pi^i G x' = y\}|}{|R^n|} \\ &= \frac{|\ker \pi^i G|}{|R^n|} \\ &= \frac{1}{|\operatorname{col} \pi^i G|},\end{aligned}$$

ou seja, a coluna $\boldsymbol{y}$ é uniformemente distribuída sobre seu suporte. Assim, a matriz $\boldsymbol{Y}$ também deve ser uniformemente distribuída sobre seu suporte. Isso conclui a prova. □

Demonstração do Teorema 4.7. A informação mútua do canal é dada por

$$\begin{aligned}I(\boldsymbol{X}; \boldsymbol{Y}, \boldsymbol{G}) &= I(\boldsymbol{X}; \boldsymbol{Y} | \boldsymbol{G}) + I(\boldsymbol{X}; \boldsymbol{G}) \\ &= H(\boldsymbol{Y} | \boldsymbol{G}) - H(\boldsymbol{Y} | \boldsymbol{X}, \boldsymbol{G}) + I(\boldsymbol{X}; \boldsymbol{G}) \\ &= H(\boldsymbol{Y} | \boldsymbol{G}),\end{aligned}$$

em que $H(\boldsymbol{Y} | \boldsymbol{X}, \boldsymbol{G}) = 0$ pois $\boldsymbol{Y} = \boldsymbol{G}\boldsymbol{X}$ e $I(\boldsymbol{X}; \boldsymbol{G}) = 0$ pois $\boldsymbol{X}$ e $\boldsymbol{G}$ são estatisticamente independentes. Portanto,

$$I(\boldsymbol{X}; \boldsymbol{Y}, \boldsymbol{G}) = H(\boldsymbol{Y} | \boldsymbol{G}) = \sum_G p_{\boldsymbol{G}}(G) H(\boldsymbol{Y} | \boldsymbol{G} = G)$$

e o resultado segue do Lema 4.8. □

4.6 Esquema de codificação

Aqui propõe-se um esquema de codificação para o canal, o qual é baseado na expansão π-ádica discutida na Seção 4.2. Por simplicidade de exposição, inicia-se descrevendo o esquema para o caso particular de códigos *one-shot*. O caso geral será discutido em seguida. De agora em diante, seja $\mathbb{F}_q = R/\langle\pi\rangle$.

4.6.1 Resultados auxiliares

Antes de descrever o esquema de codificação proposto, são apresentados dois lemas que tratam da solução de sistemas de equações lineares sobre um anel de cadeia finito. O primeiro deles converte um sistema de equações lineares sobre um anel de cadeia finito em múltiplos sistemas sobre o corpo residual. As demonstrações se encontram na Seção 4.7.

Lema 4.9. *Sejam $y \in R^n$ e $G \in \mathrm{GL}_n(R)$. Seja $x \in R^n$ a solução de $Ax = y$. Então, a expansão π-ádica de x pode ser obtida recursivamente através de*

$$A^{(0)}x^{(i)} \equiv_\pi y^{(i)} - \left(Ax^i\right)^{(i)},$$

para $0 \leq i < s$. [Lembre-se de que $A \in \mathrm{GL}_n(R)$ se e somente se $A^{(0)} \in \mathrm{GL}_n(\mathbb{F}_q)$.]

O segundo problema trata da solução de sistemas diagonais de equações lineares. No que segue, $A_{j:j'}$ denota a submatriz de A contendo a linha j até (mas não incluindo) a linha j', em que as entradas das matrizes são indexadas começando de 0.

Lema 4.10. *Seja $Y \in R^{m\times\lambda}$ e $D \in R^{m\times n}$, em que D está na forma normal de Smith e tem shape ρ. Se $Y = DX$, então*

$$X^{(i)}_{0:\rho_{s-i-1}} = \begin{bmatrix} Y^{(i)}_{0:\rho_0} \\ Y^{(i+1)}_{\rho_0:\rho_1} \\ \vdots \\ Y^{(i+s-1)}_{\rho_{s-i-2}:\rho_{s-i-1}} \end{bmatrix},$$

para $0 \leq i < s$.

EXEMPLO. Seja $R = \mathbb{Z}_8$, com $\pi = 2$ e $\Gamma = \{0, 1\}$. Seja $n = 5$, $m = 4$ e $\lambda = (3, 4, 6)$. Suponha que $\rho = (1, 3, 4)$, de modo que $D = \operatorname{diag}(1, 2, 2, 4) \in \mathbb{Z}_8^{4\times 5}$. Em adição, suponha que

$$Y = \begin{bmatrix} 6 & 7 & 1 & 2 & 0 & 4 \\ 6 & 4 & 2 & 0 & 0 & 0 \\ 0 & 2 & 6 & 4 & 0 & 0 \\ 4 & 0 & 4 & 0 & 0 & 0 \end{bmatrix}.$$

Daí, pode-se concluir que

$$X^{(0)} = \begin{bmatrix} 0\ 1\ 1 \\ 1\ 0\ 1 \\ 0\ 1\ 1 \\ 1\ 0\ 1 \\ *\ *\ * \end{bmatrix}, X^{(1)} = \begin{bmatrix} 1\ 1\ 0\ 1 \\ 1\ 1\ 0\ 0 \\ 0\ 0\ 1\ 0 \\ *\ *\ *\ * \\ *\ *\ *\ * \end{bmatrix}, X^{(2)} = \begin{bmatrix} 1\ 1\ 0\ 0\ 0\ 1 \\ *\ *\ *\ *\ *\ * \\ *\ *\ *\ *\ *\ * \\ *\ *\ *\ *\ *\ * \\ *\ *\ *\ *\ *\ * \end{bmatrix},$$

em que $*$ denota entrada desconhecida e espaço em branco denota entrada igual a zero. Note que as entradas desconhecidas são devido a $\rho = \operatorname{shape} D$, enquanto que as entradas em branco são devido a λ (veja a Seção 4.3). □

4.6.2 Construção do código

Sejam $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{s-1}$ (chamados de *códigos componentes*) códigos matriciais *one-shot* sobre o corpo residual $\mathbb{F}_q$, em que cada $\mathcal{C}_i$, para $0 \leq i < s$, é um código para C-MMC($\boldsymbol{G}_i, \lambda_i$), para alguma matriz de transferência $\boldsymbol{G}_i \in \mathbb{F}_q^{m\times n}$. Os códigos componentes serão combinados para se obter um código matricial $\mathcal{C}$ *one-shot* sobre o anel de cadeia R, projetado para C-MMC($\boldsymbol{G}, \lambda$). O código $\mathcal{C}$ é chamado de *código composto*.

Denota-se por $\varphi : R \to \mathbb{F}_q$ a projeção natural de R sobre $\mathbb{F}_q$ e por $\bar{\varphi} : \mathbb{F}_q \to \Gamma$ o mapeamento seletor do representante da classe lateral, que possui a propriedade de que $\varphi(\bar{\varphi}(x)) = x$ para todo $x \in \mathbb{F}_q$. O

código $\mathcal{C} \subseteq R^{n\times\lambda}$ é definido por

$$\mathcal{C} = \left\{ \sum_{i=0}^{s-1} X^{(i)}\pi^i : X_i \in \mathcal{C}_i, 0 \leq i < s \right\},$$

em que

$$X^{(i)} = \begin{bmatrix} \bar{\varphi}(X_i) & 0 \end{bmatrix} \in \Gamma^{n\times\ell}. \tag{4.1}$$

Verifica-se que $\mathcal{C}$ de fato satisfaz as restrições de $R^{n\times\lambda}$.

4.6.3 Decodificação

O procedimento de decodificação é descrito a seguir, sendo baseado nas ideias da Subseção 4.6.1. Intuitivamente, o decodificador decompõe um único MMC sobre o anel de cadeia em múltiplos MMCs sobre o corpo residual. No que segue, $A_{j\times k}$ denota a submatriz $j \times k$ superior-esquerda de A.

Passo 1. O decodificador, que conhece a matriz de transferência G, inicia computando $D \in R^{m\times n}$, a forma normal de Smith de G. O decodificador também calcula $P \in \mathrm{GL}_m(R)$ e $Q \in \mathrm{GL}_n(R)$ tais que $G = PDQ$.

Passo 2. Seja $\tilde{X} \triangleq QX \in R^{n\times\lambda}$ (que é desconhecida pelo decodificador) e $\tilde{Y} \triangleq P^{-1}Y \in R^{m\times\lambda}$ (que é calculada pelo decodificador), de modo que $Y = GX$ equivale a $\tilde{Y} = D\tilde{X}$. Dessa equação, o decodificador obtém informação parcial sobre $\tilde{X}$. Mais precisamente, o decodificador computa $\tilde{X}^{(i)}_{\rho_{s-i-1}\times\lambda_i}$, para $0 \leq i < s$, de acordo com o Lema 4.10.

Passo 3. De posse de $\tilde{X}^{(i)}_{\rho_{s-i-1}\times\lambda_i}$, para $0 \leq i < s$, o decodificador tentará então decodificar X baseado na equação

$$\tilde{X} = QX,$$

em um estilo similar à decodificação *multi-estágio*. De fato, analogamente ao Lema 4.9, tem-se, para $0 \leq i < s$,

$$\tilde{X}^{(i)} - \left(QX^{\underline{i}}\right)^{(i)} \equiv_\pi Q^{(0)}X^{(i)}.$$

Considerando apenas as ρ_{s-i-1} linhas superiores (pois as linhas restantes são desconhecidas) e mantendo apenas as λ_i colunas da esquerda

(pois sabe-se de antemão que as colunas restantes são nulas), obtém-se

$$\tilde{X}^{(i)}_{\rho_{s-i-1}\times\lambda_i} - \left(Q_{\rho_{s-i-1}\times n} X^{\underline{i}}_{n\times\lambda_i}\right)^{(i)} \equiv_\pi Q^{(0)}_{\rho_{s-i-1}\times n} X^{(i)}_{n\times\lambda_i}.$$

Projetando em $\mathbb{F}_q$ (ou seja, aplicando φ em ambos os lados da equação) e acrescentando suficientes linhas nulas (de modo a se obter um sistema $m \times n$), chega-se finalmente a

$$Y_i = G_i X_i, \tag{4.2}$$

em que $Y_i \in \mathbb{F}_q^{m\times\lambda_i}$ e $G_i \in \mathbb{F}_q^{m\times n}$ são definidas por

$$Y_i = \begin{bmatrix} \varphi\big(\tilde{X}^{(i)}_{\rho_{s-i-1}\times\lambda_i}\big) - \varphi\Big(\big(Q_{\rho_{s-i-1}\times n} X^{\underline{i}}_{n\times\lambda_i}\big)^{(i)}\Big) \\ 0 \end{bmatrix}, \tag{4.3}$$

$$G_i = \begin{bmatrix} \varphi\big(Q_{\rho_{s-i-1}\times n}\big) \\ 0 \end{bmatrix}. \tag{4.4}$$

Note que Y_i só pode ser calculado depois de se conhecer $X_0, X_1, \ldots, X_{i-1}$. Portanto, neste passo, o decodificador obtém, sucessivamente, estimativas de $X_0, X_1, \ldots, X_{s-1}$ de acordo com (4.2). Após isso, o decodificador calcula uma estimativa de X utilizando (4.1) e a expansão π-ádica.

4.6.4 Extensão para o caso *multi-shot*

Por fim, considera-se o caso *multi-shot*. Seja $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{s-1}$ uma sequência de códigos matriciais componentes de comprimento N, em que $\mathcal{C}_i \subseteq (\mathbb{F}_q^{n\times\lambda_i})^N$, para $0 \le i < s$. As palavras-código do código matricial composto $\mathcal{C}$ são então dadas por $(X(1), X(2), \ldots, X(N)) \in (R^{n\times\lambda})^N$, em que $X(j)$ é obtido da j-ésima coordenada das palavras-código dos códigos componentes, de maneira análoga ao caso *one-shot*.

Prosseguindo como nos Passos 1 e 2 acima, o decodificador obtém $\tilde{X}^{(0)}_{\rho_{s-1}\times\lambda_0}(j), \tilde{X}^{(1)}_{\rho_{s-2}\times\lambda_1}(j), \ldots, \tilde{X}^{(s-1)}_{\rho_0\times\lambda_{s-1}}(j)$ e $Q(j)$, para $j = 1, \ldots, N$. O Passo 3 é também análogo, com o detalhe importante de que, antes de prosseguir ao estágio $i+1$, a sequência $(X_i(1), X_i(2), \ldots, X_i(N)) \in \mathcal{C}_i$ é decodificada por completo, com base em $(Y_i(1), Y_i(2), \ldots, Y_i(N))$ e $(G_i(1), G_i(2), \ldots, G_i(N))$, utilizando o decodificador de $\mathcal{C}_i$.

4.6.5 Taxa, probabilidade de erro e complexidade

Do esquema proposto, fica claro que o i-ésimo código componente deve ser projetado para C-MMC($\boldsymbol{G}_i, \lambda_i$), em que a matriz $\boldsymbol{G}_i \in \mathbb{F}_q^{m\times n}$ é definida em (4.4). Em princípio, poderia-se calcular a distribuição de probabilidade de $\boldsymbol{G}_i$ se a distribuição de $\boldsymbol{G}$ for conhecida. No entanto, se for utilizado um dos esquemas de codificação propostos em [55] (veja o Capítulo 2), o conhecimento exato da distribuição de $\boldsymbol{G}_i$ torna-se desnecessário tão logo se saiba o valor esperado de seu posto. De (4.4), tem-se $\operatorname{rank} \boldsymbol{G}_i = \boldsymbol{\rho}_{s-i-1}$, de modo que, nesse caso, apenas é necessário o conhecimento de $\mathrm{E}[\boldsymbol{\rho}]$.

Seja $\mathcal{C}$ o código composto obtido dos códigos componentes $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{s-1}$. Então, $\mathcal{C}$ tem taxa

$$\mathrm{R}(\mathcal{C}) = \mathrm{R}(\mathcal{C}_0) + \mathrm{R}(\mathcal{C}_1) + \cdots + \mathrm{R}(\mathcal{C}_{s-1}),$$

em dígitos q-ários. Além disso, do limitante da união, a probabilidade de erro é limitada superiormente de acordo com

$$\mathrm{P_e}(\mathcal{C}) \leq \mathrm{P_e}(\mathcal{C}_0) + \mathrm{P_e}(\mathcal{C}_1) + \cdots + \mathrm{P_e}(\mathcal{C}_{s-1}).$$

Assim, se cada $\mathcal{C}_i$ alcança a capacidade em C-MMC($\boldsymbol{G}_i, \lambda_i$), tem-se $\mathrm{R}(\mathcal{C}_i)$ arbitrariamente próximo de $\mathrm{E}[\boldsymbol{\rho}_{s-i-1}]\lambda_i$ e $\mathrm{P_e}(\mathcal{C}_i)$ arbitrariamente próximo de zero, para $0 \leq i < s$. Por conseguinte, tem-se $\mathrm{R}(\mathcal{C})$ arbitrariamente próximo de $\sum_i \mathrm{E}[\boldsymbol{\rho}_{s-i-1}]\lambda_i$ (a capacidade do canal) e $\mathrm{P_e}(\mathcal{C})$ arbitrariamente próximo de zero.

A complexidade computacional associada à decodificação do código composto é simplesmente a soma das complexidades individuais das decodificações de cada código componente, acrescida do custo de calcular a forma normal de Smith de G (o que pode ser feito com $O(nm\min\{n,m\})$ operações em R), do custo de calcular $\tilde{Y}$ (o que requer $O(m^2(m+\ell))$ operações) e do custo de $s-1$ multiplicações e adições matriciais em (4.3) (o que exige $O(n^2\ell)$ operações cada).

4.7 Demonstrações omitidas

Os seguintes fatos algébricos serão úteis nas demonstrações à frente.

Lema 4.11. *Sejam $x, y, z \in R$. Então*

(i) $\left(x\pi^i\right)^{(i+j)} = x^{(j)}$, *para* $0 \leq j < s - i$;

(ii) $(x + y\pi^i + z\pi^{i+1})^{(i)} \equiv_\pi x^{(i)} + y^{(0)}$.

Demonstração. A primeira afirmação segue imediatamente da unicidade da decomposição π-ádica. Para a segunda afirmação, tem-se

$$\begin{aligned}
(x + \pi^i y + \pi^{i+1} z)^{(i)} &= \left(\sum_{j=0}^{s-1} \pi^j x^{(j)} + \pi^i \sum_{j=0}^{s-1} \pi^j y^{(j)} + \pi^{i+1} \sum_{j=0}^{s-1} \pi^j z^{(j)}\right)^{(i)} \\
&\overset{(a)}{=} \left(\sum_{j=0}^{i} \pi^j x^{(j)} + \pi^i y^{(0)}\right)^{(i)} \\
&= \left(\sum_{j=0}^{i-1} \pi^j x^{(j)} + \pi^i (x^{(i)} + y^{(0)})\right)^{(i)} \\
&\overset{(b)}{=} \left(\pi^i (x^{(i)} + y^{(0)})\right)^{(i)} \\
&\overset{(c)}{=} \left(x^{(i)} + y^{(0)}\right)^{(0)} \equiv_\pi x^{(i)} + y^{(0)},
\end{aligned}$$

em que (a) segue pois múltiplos de π^{i+1} não contribuem para o valor do i-ésimo termo da expansão π-ádica, (b) é verdadeiro pela unicidade da expansão π-ádica e (c) segue de (i) com $j = 0$. □

Demonstração do Lema 4.9. Para $0 \leq i < s$, tem-se

$$y = Ax = A\sum_{j=0}^{i-1} x^{(j)}\pi^j + Ax^{(i)}\pi^i + A\sum_{j=i+1}^{s-1} x^{(j)}\pi^j,$$

de modo que, do Lema 4.11, deduz-se que

$$y^{(i)} \equiv_\pi \left(Ax^{\underline{i}}\right)^{(i)} + \left(Ax^{(i)}\right)^{(0)}.$$

Após simplificar e rearranjar os termos, segue o resultado. Note que, uma vez que $A^{(0)} \in \mathrm{GL}_n(\mathbb{F}_q)$, é sempre possível calcular, recursivamente, $x^{(0)}, x^{(1)}, \ldots, x^{(s-1)}$. □

Demonstração do Lema 4.10. Note que $Y = DX$ é equivalente a

$$\begin{aligned} Y_{0:\rho_0} &= X_{0:\rho_0}, \\ Y_{\rho_0:\rho_1} &= \pi X_{\rho_0:\rho_1}, \\ &\vdots \\ Y_{\rho_{s-2}:\rho_{s-1}} &= \pi^{s-1} X_{\rho_{s-2}:\rho_{s-1}}. \end{aligned}$$

Pelo Lema 4.11, isso implica em

$$\begin{aligned} X^{(i)}_{0:\rho_0} &= Y^{(i)}_{0:\rho_0}, & 0 \leq i < s, \\ X^{(i)}_{\rho_0:\rho_1} &= Y^{(i+1)}_{\rho_0:\rho_1}, & 0 \leq i < s-1, \\ &\vdots & \vdots \\ X^{(i)}_{\rho_{s-2}:\rho_{s-1}} &= Y^{(i+s-1)}_{\rho_{s-2}:\rho_{s-1}}, & 0 \leq i < 1, \end{aligned}$$

de onde segue o resultado. □

CAPÍTULO 5

Conclusão

Este trabalho considerou canais matriciais multiplicativos sobre corpos e anéis de cadeia finitos, os quais são modelos adequados para a comunicação fim-a-fim em redes que operam de acordo com codificação de rede linear (possivelmente na camada física). O enfoque adotado foi probabilístico, fazendo uso da teoria da informação.

Foram abordados dois problemas. No primeiro deles (Capítulo 3), foram estudados MMCs não-coerentes sobre corpos finitos nos quais a matriz de transferência é uniformemente distribuída dado o valor de seu posto. Foi defendida a aplicação desse modelo de canal em sistemas que operam com codificação de rede linear aleatória cujos enlaces estão sujeitos a apagamentos, uma vez que acredita-se que tal modelo é flexível o suficiente para capturar as características essenciais do sistema, ao mesmo tempo em que mantém tratabilidade matemática. Isso contrasta com outros modelos de canal considerados, os quais são muito restritivos ou muito complexos. Como contribuições, foi mostrado que o problema do cálculo da capacidade do canal pode ser reduzido à solução de um problema de otimização convexa sobre $n+1$ variáveis (ao invés de $q^{n\ell}$), o qual pode ser resolvido por métodos numéricos bem estabelecidos. Os resultados foram então especializados para o impor-

tante caso de entrada de posto constante, no qual foi possível obter uma forma fechada para a capacidade. Para tamanho do corpo ou comprimento do pacote assintoticamente grande, foi mostrado que entrada de posto constante é ótima. Finalmente, foi provado que, mesmo nesse modelo mais geral, codificação de subespaço ainda é suficiente para alcançar a capacidade. Muitos dos resultados obtidos generalizam conclusões obtidas previamente na literatura.

No segundo problema abordado (Capítulo 4), foram investigados MMCs sobre anéis de cadeia finitos, os quais têm aplicações práticas em codificação de rede na camada física baseada em reticulados aninhados. Como contribuições, a capacidade do canal foi determinada, generalizando o resultado correspondente para corpos finitos. Além disso, um esquema de codificação prático que alcança a capacidade foi proposto, combinando vários códigos sobre o corpo residual para obter um novo código sobre o anel de cadeia.

Trabalhos futuros

Como futuras linhas de pesquisa, sugere-se o seguinte.

(*i*) O projeto de esquemas práticos de codificação para o MMC não-coerente sobre corpos finitos. No caso em que o comprimento do pacote, ℓ, é arbitrariamente grande, pode-se adaptar os códigos de Yang et al. [55, 56, 57] para o MMC coerente através do uso de cabeçalhos. No entanto, para ℓ pequeno, o problema parece estar ainda em aberto. Um primeiro passo nessa direção foi o recente trabalho de Yang [52]; no entanto, algoritmos eficientes de codificação e decodificação não foram fornecidos.

(*ii*) A determinação da capacidade do MMC não-coerente sobre anéis de cadeia finitos (em outras palavras, uma extensão simultânea dos resultados obtidos nos Capítulos 3 e 4 desta tese). Nessa linha, a distribuição de probabilidade "uniforme dado o *shape*" parece ser de fundamental importância.

(*iii*) A extensão dos resultados obtidos para o MMC coerente, de anéis de cadeia finitos para *anéis de ideias principais finitos* (dos quais anéis de cadeia finitos são casos particulares). De fato, como mostrado por Feng, Silva e Kschischang [12], o espaço de pacotes

induzido por codificação de rede na camada física via reticulados aninhados é, no caso mais geral, um módulo W sobre um anel de ideais principais finito R.

(*iv*) O cálculo analítico da distribuição de posto em função de uma dada topologia de rede. Este problema aparenta ser desafiador mesmo no caso mais simples no qual os enlaces são livres de apagamento.

(*v*) O estudo do *canal matricial aditivo-multiplicativo* (AMMC, do inglês *additive-multiplicative matrix channel*) não-coerente, o qual é definido pela expressão $\boldsymbol{Y} = \boldsymbol{GX} + \boldsymbol{Z}$. Limitantes superiores e inferiores sobre a capacidade foram apresentados por Silva, Kschischang e Kötter [49], mas um cálculo mais preciso da capacidade ainda está em aberto. Resultados preliminares nessa linha já foram obtidos, em que é assumido que tanto $\boldsymbol{G}$ quanto $\boldsymbol{Z}$ são u.g.r. De fato, o Lema B.4 implica que a probabilidade de transição $p_{\boldsymbol{Y}|\boldsymbol{X}}(Y|X)$ do AMMC depende de X e Y apenas através de $\operatorname{rank} X$, $\operatorname{rank} Y$ e $\dim(\langle X\rangle \cup \langle Y\rangle)$. Com isso, acredita-se que (analogamente ao MMC) o AMMC apresente propriedades de simetria em bloco [43] e que a capacidade também seja atingida por entrada u.g.r.

(*vi*) A generalização, de corpos para anéis de cadeia finitos, dos resultados de Silva e Kschischang [47], em que são determinadas condições necessárias e suficientes sob as quais um código matricial é bem-sucedido em sua tarefa de correção de erros e deficiências de posto em um modelo de canal adversário. Resultados preliminares para o caso sem erros de enlace já foram obtidos.

APÊNDICE A

Anéis e módulos

Este apêndice revisa conceitos básicos da teoria de anéis comutativos e módulos sobre tais anéis. Para mais detalhes, veja qualquer livro de álgebra abstrata, como, por exemplo, [8] ou [21].

A.1 Anéis comutativos

A.1.1 Definições iniciais

Um *anel* é um um conjunto R juntamente com duas operações binárias, *adição*, $+ : R \times R \to R$, $(r, s) \mapsto r + s$ e *multiplicação*, $\cdot : R \times R \to R$, $(r, s) \mapsto rs$, satisfazendo o seguinte:

(*i*) $(a + b) + c = a + (b + c)$ para todo $a, b, c \in R$,

(*ii*) $a + b = b + a$, para todo $a, b \in R$,

(*iii*) Existe $0 \in R$ tal que $a + 0 = a$ para todo $a \in R$,

(*iv*) Para todo $a \in R$, existe $-a \in R$ tal que $a + (-a) = 0$.

(*v*) $(ab)c = a(bc)$, para todo $a, b, c \in R$,

(*vi*) $a(b + c) = ab + ac$ e $(a + b)c = ac + bc$, para todo $a, b, c \in R$.

Note que as propriedades (i)–(iv) acima são equivalentes a dizer que R é um *grupo abeliano* com relação à adição.

Além disso, este trabalho considera apenas anéis com identidade multiplicativa, de modo que uma outra propriedade é exigida:

(vii) Existe $1 \in R$ tal que $1a = a1 = a$ para todo $a \in R$.

Adicionalmente, o anel é dito ser *comutativo* se

$(viii)$ $ab = ba$, para todo $a, b \in R$.

Neste trabalho, a menos que explicitamente indicado, serão considerados apenas anéis comutativos. Além disso, é sempre assumido que $1 \neq 0$.

Como consequências imediatas das definições, tem-se que $0a = 0$ e $(-1)a = -a$ para todo $a \in R$. Além disso, pode-se mostrar que a identidade aditiva (isto é, o elemento 0) e a identidade multiplicativa (isto é, o elemento 1) de um anel são únicas.

A.1.2 Elementos inversíveis e divisores de zero

Seja R um anel (comutativo). Um elemento $r \in R$ é dito ser *inversível* se existir $s \in R$ tal que $rs = 1$. Esse elemento s é unicamente determinado por r e é chamado de *inverso multiplicativo* de r, denotado por $s = r^{-1}$. Um elemento $r \in R$ é dito ser um *divisor de zero* se existir $s \in R$ tal que $s \neq 0$ e $rs = 0$. Dessas definições, conclui-se que 1 é inversível e 0 é um divisor de zero. Pode-se mostrar o seguinte.

Fato A.1. *Em um anel, um elemento não pode ser simultaneamente inversível e divisor de zero.*

Dois elementos $r_1, r_2 \in R$ são ditos *associados* se $r_1 = ur_2$ para algum elemento inversível $u \in R$; escreve-se então $r_1 \approx r_2$. Pode-se mostrar que $\approx$ é uma relação de equivalência em R.

A.1.3 Domínios e corpos

Um *domínio de integridade* (ou, simplesmente, *domínio*) é um anel no qual o único divisor de zero é 0. Um *corpo* é um anel no qual todo elemento não-nulo é inversível. Todo corpo é um domínio.

EXEMPLO. Em $\mathbb{Z}$, o anel dos números inteiros, 1 e -1 são os únicos elementos inversíveis, 0 é o único divisor de zero, e todos os outros elementos não são nem inversíveis nem divisores de zero. Em $\mathbb{Q}$, o anel dos números racionais, 0 é o único divisor de zero, e todos os outros elementos são inversíveis. Assim, $\mathbb{Z}$ é um domínio e $\mathbb{Q}$ é um corpo. Considere agora $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$, o anel dos inteiros modulo 6. Pode-se mostrar que $\{1, 5\}$ são inversíveis e que $\{0, 2, 3, 4\}$ são divisores de zero. Assim, $\mathbb{Z}_6$ não é um domínio (muito menos um corpo). □

A.1.4 Ideais

Um subconjunto não-vazio $I \subseteq R$ é dito ser um *ideal* de R se, para todo $x, y \in I$ e $r \in R$,

(*i*) $x + y \in I$, isto é, I é fechado com relação à adição, e

(*ii*) $rx \in I$, isto é, I é fechado com relação à multiplicação por elementos do anel.

Uma vez que aqui serão considerados apenas anéis com identidade multiplicativa, o fechamento com relação à negação segue do caso particular $r = -1$ acima.

Seja $X \subseteq R$. Denota-se por $\langle X \rangle$ o conjunto de todas as "combinações lineares" (com coeficientes em R) dos elementos de X, isto é,

$$\langle X \rangle \triangleq \{c_1 x_1 + c_2 x_2 + \cdots + c_n x_n : c_i \in R, x_i \in X, n \in \mathbb{N}\}. \qquad \text{(A.1)}$$

Quando $X = \{x_1, \ldots, x_n\}$, abrevia-se $\langle \{x_1, \ldots, x_n\} \rangle$ por $\langle x_1, \ldots, x_n \rangle$. Note que, para um conjunto unitário $X = \{x\}$, (A.1) reduz para

$$\langle x \rangle = \{cx : c \in R\}. \qquad \text{(A.2)}$$

Por definição, $\langle \emptyset \rangle = \{0\}$. Uma vez que $\langle X \rangle$ é claramente um ideal de R, chama-se tal conjunto de *ideal gerado por* X. Reciprocamente, se I é um ideal de R, então $I = \langle X \rangle$ para algum $X \subseteq R$.

Não é difícil mostrar que tanto R quanto $\{0\}$ são sempre ideais de R; esses são chamados de *ideais triviais*. Um ideal I de R é dito ser *próprio* se $I \neq R$, e *não-nulo* se $I \neq \{0\}$. Tem-se $I = R$ se e somente se $1 \in I$. Um anel R é um corpo se e somente se R possuir apenas ideais triviais.

A.1.5 Ideais principais

Ideais da forma em (A.2), os quais são gerados por um único elemento, recebem um nome especial: um ideal I é dito ser *principal* se existir $x \in R$ tal que $I = \langle x \rangle$. A seguir são apresentados alguns fatos relacionados com ideais principais.

(*i*) Se $x \approx y$, então $\langle x \rangle = \langle y \rangle$.

(*ii*) Se $\langle x \rangle \subseteq \langle y \rangle$, então $\langle x, y \rangle = \langle y \rangle$.

Note que os ideais triviais são ideais principais: $R = \langle 1 \rangle$ e $\{0\} = \langle 0 \rangle$.

A.1.6 Anel quociente

Um ideal pode ser usado para construir o que é chamado de *anel quociente*. Seja I um ideal de R. Define-se a relação $\equiv_I$ em R como segue:

$$a \equiv_I b \text{ se e somente se } a - b \in I.$$

Usando as propriedades de ideais, não é difícil mostrar que $\equiv_I$ é uma relação de equivalência; a classe de equivalência de um elemento $a \in R$ é dada por

$$a + I \triangleq \{a + r : r \in I\}$$

e também é chamada de *classe lateral*. O conjunto de todas as classes laterais é denotado por R/I; esse conjunto torna-se um anel, chamado de *anel quociente de R modulo I*, se as operações de adição e multiplicação forem definidas, respectivamente, por

$$(a + I) + (b + I) = (a + b) + I,$$
$$(a + I)(b + I) = (ab) + I,$$

É possível mostrar que tais operações são bem-definidas. A identidade aditiva de R/I é $0 + I = I$ e a identidade multiplicativa é $1 + I$.

Tem-se $R/R = \{0 + R\}$ (um anel trivial) e $R/\{0\} = \{\{r\} : r \in R\}$ (isomorfo ao próprio R).

A.1.7 Maximalidade e primalidade

Um ideal próprio M de R é dito ser um *ideal maximal* se $M \nsubseteq I$ para todo ideal próprio $I \neq M$ de R. Pode-se mostrar (utilizando o Lema

de Zorn) que todo anel tem pelo menos um ideal maximal.

Fato A.2. *Seja M um ideal de R. Então, M é um ideal maximal se e somente se R/M for um corpo. Tal corpo é chamado de* corpo residual.

Para todo $x, y \in R$, diz-se que x *divide* y, escrito como $x \mid y$, se existir $c \in R$ tal que $y = cx$. Note que $x \mid y$ equivale a $y \in \langle x \rangle$, o que, por sua vez, equivale a $\langle y \rangle \subseteq \langle x \rangle$. Um elemento não-nulo e não inversível $p \in R$ é dito ser *primo* se $p \mid ab$ implicar $p \mid a$ ou $p \mid b$. Um ideal próprio P de R é chamado de *ideal primo* se $ab \in P$ implicar $a \in P$ ou $b \in P$. Seja $p \in R$ não-nulo. Então, p é primo se e somente se $\langle p \rangle$ for um ideal primo.

Note que, de acordo com as definições, 0 nunca é um elemento primo, enquanto que $\{0\}$ pode ou não ser um ideal primo. Por exemplo, $\{0\}$ é um ideal primo em $\mathbb{Z}$, mas não é em $\mathbb{Z}_6$. De fato, $\{0\}$ é um ideal primo se e somente se o anel em questão for um domínio. Tal afirmação é um caso particular do fato a seguir.

Fato A.3. *Seja P um ideal de R. Então, P é um ideal primo se e somente se R/P for um domínio.*

Combinando os Fatos A.2 e A.3 com o fato de que todo corpo é um domínio, obtém-se o seguinte.

Fato A.4. *Todo ideal maximal é um ideal primo.*

Observação: Primalidade não deve ser confundida com *irredutibilidade*. Um elemento não-nulo e não inversível, $r \in R$ é dito ser *irredutível* se $r = ab$ implicar a inversível ou b inversível. Em um domínio, se r for primo, então r é irredutível. Em um *domínio de ideais principais* (a ser definido em breve), r é primo se e somente se r for irredutível.

A.1.8 Nilpotência

Um elemento $r \in R$ é dito ser *nilpotente* se existir um inteiro positivo k tal que $r^k = 0$. O menor inteiro positivo k tal que $r^k = 0$, caso exista, é chamado de *índice de nilpotência* de r. Alguns fatos imediatos:

Fato A.5. *O elemento 0 é sempre nilpotente, com índice de nilpotência 1. Todo elemento nilpotente é um divisor de zero. Não existe elemento simultaneamente nilpotente e inversível. Em um domínio, o único elemento nilpotente é 0.*

O conjunto de todos os elementos nilpotentes de um anel é chamado de *nilradical*.

Fato A.6. *O nilradical de um anel é um ideal. O nilradical é dado pela intersecção de todos os ideais primos.*

EXEMPLO. Considere $\mathbb{Z}_{12}$, o anel dos inteiros modulo 12. Os ideais primos são $\langle 2\rangle = \{0, 2, 4, 6, 8, 10\}$ e $\langle 3\rangle = \{0, 3, 6, 9\}$. O nilradical é $N = \{0, 6\}$, o qual é igual a $\langle 2\rangle \cap \langle 3\rangle$. Em $\mathbb{Z}_8$, o anel dos inteiros modulo 8, existe apenas um ideal primo, dado por $\langle 2\rangle = \{0, 2, 4, 6\}$, que é o próprio nilradical. □

A.1.9 Produto de anéis

Sejam R e S dois anéis. Então, o produto cartesiano $R \times S$ torna-se um anel se as operações de adição e multiplicação forem definidas entrada-a-entrada, ou seja,

$$(r_1, s_1) + (r_2, s_2) = (r_1 + r_2, s_1 + s_2),$$
$$(r_1, s_1)(r_2, s_2) = (r_1 r_2, s_1 s_2),$$

para todo $(r_1, s_1), (r_2, s_2) \in R \times S$. A identidade aditiva é $(0_R, 0_S)$, em que 0_R e 0_S são as identidades aditivas de R e S, respectivamente. Analogamente, a identidade multiplicativa é $(1_R, 1_S)$. Essa definição se estende naturalmente para um número arbitrário de anéis.

A.1.10 Anéis locais e anéis de ideais principais

Define-se o seguinte:

(*i*) Um *anel local* é um anel que possui exatamente um ideal maximal (e, portanto, um único corpo residual).

(*ii*) Um *anel de ideais principais* é um anel cujos ideais são todos principais.

A seguir apresenta-se alguns fatos básicos sobre anéis de ideais principais (veja, por exemplo, [59, §IV.15]).

Fato A.7. *O produto de anéis principais é também um anel principal. O quociente de um anel principal é também um anel principal.*

Caso o anel seja um domínio, ele também é chamado de *domínio de ideais principais*. Obviamente, todo corpo é simultaneamente um anel local e um domínio de ideais principais, mas a recíproca é falsa (veja o exemplo a seguir).

Exemplo.

(*i*) $\mathbb{Z}$, o anel dos números inteiros, é um domínio de ideais principais. De fato, pode ser mostrado que $\langle a_1, a_2, \ldots, a_n \rangle = \langle b \rangle$, em que $b \in \mathbb{Z}$ é o máximo divisor comum entre $a_1, a_2, \ldots, a_n \in \mathbb{Z}$. No entanto, $\mathbb{Z}$ não é um anel local, uma vez que os ideais maximais são precisamente os ideais primos $\langle 0 \rangle, \langle 2 \rangle, \langle 3 \rangle, \langle 5 \rangle, \langle 7 \rangle, \langle 11 \rangle$, etc.

(*ii*) $\hat{\mathbb{Z}}_p$, o anel dos inteiros p-ádicos, é um domínio de ideais principais e um anel local. No entanto, $\hat{\mathbb{Z}}_p$ não é um corpo.

(*iii*) $\mathbb{Z}[X]$, o anel dos polinômios com coeficientes inteiros, é um domínio, mas não é nem um anel principal nem um anel local. Pode-se mostrar que $\langle 2, X \rangle$ é um ideal não-principal. Os ideais maximais de $\mathbb{Z}[X]$ são da forma $\langle p, f(X) \rangle$ em que p é um número primo e $f(X) \in \mathbb{Z}[X]$ é irredutível módulo p.

□

A.2 Módulos sobre anéis comutativos

Durante o decorrer desta seção, seja R um anel comutativo com identidade aditiva 0 e identidade multiplicativa 1, em que $1 \neq 0$.

A.2.1 Definições iniciais

Um *R-módulo* é um conjunto U juntamente com uma operação binária, *adição*, $+ : U \times U \to U$, $(u, v) \mapsto u + v$ e uma função, *multiplicação por escalar*, $\cdot : R \times U \to U$, $(r, u) \mapsto ru$, satisfazendo o seguinte:

(*i*) U é um grupo abeliano com relação à adição,

(*ii*) $(r + s)u = ru + su$,

(*iii*) $r(u + v) = ru + rv$,

(*iv*) $(rs)u = r(su)$,

(v) $1u = u$,

para todo $r, s \in R$ e $u, v \in U$.

Consequências imediatas: Para todo $r \in R$ e $u \in U$ tem-se $r0 = 0$, $0u = 0$ e $(-1)u = -u$.

Observação: Note que o símbolo 0 também é usado para denotar a identidade aditiva de U. O contexto tornará claro se está-se referindo a $0 \in R$ ou $0 \in U$.

EXEMPLO. Como casos especiais de módulos, tem-se:

- F-módulos são o mesmo que espaços vetoriais sobre F (em que F é um corpo).
- $\mathbb{Z}$-módulos são o mesmo que grupos abelianos. Nesse caso, multiplicação por escalar é definida como adição repetidas vezes.
- Se I é um ideal de R, então I é um R-módulo. Adição e multiplicação por escalar no módulo são definidos como adição e multiplicação em R, respectivamente. Em particular, R é um R-modulo.

□

A.2.2 Submódulos

Seja U um R-módulo e V um subconjunto não-vazio de U. Diz-se que V é um *submódulo* de U se V for fechado com relação a adição e multiplicação por escalar (ou seja, se $x + y \in V$ e $rx \in V$ para todo $x, y \in V$ e $r \in R$). Obviamente, $\{0\}$ e U são sempre submódulos de U.

EXEMPLO. No caso de espaços vetoriais, submódulos são o mesmo que subspaços. Para grupos abelianos, submódulos são o mesmo que subgrupos. Seja um ideal I de R como um R-módulo. Submódulos de I são precisamente os ideais de R contidos em I. □

A.2.3 Homomorfismo de módulos

Sejam U e V dois R-módulos. Um mapeamento $\phi : U \to V$ é chamado de um *homomorfismo (de R-módulos)*, também chamado de "mapeamento linear", se $\phi(x + y) = \phi(x) + \phi(y)$ e $\phi(rx) = r\phi(x)$, para todo $x, y \in U$ e $r \in R$.

A *imagem* de um homomorfismo ϕ é definida por

$$\operatorname{img}\phi \triangleq \{v \in V : v = \phi(u) \text{ para algum } u \in U\},$$

e o *núcleo* de ϕ é definido por

$$\ker\phi \triangleq \{u \in U : \phi(u) = 0\}.$$

É possível mostrar que $\operatorname{img}\phi$ é um submódulo de V, ao passo que $\ker\phi$ é um submódulo de U.

Fato A.8. *Sejam U e V dois R-módulos e $\phi : U \to V$ um homomorfismo. Então, ϕ é injetor se e somente se* $\ker\phi = \{0\}$.

Um homomorfismo bijetor é chamado de *isomorfismo*. Dois R-módulos U e V são ditos ser *isomorfos* se existir um isomorfismo $\phi : U \to V$; nesse caso, escreve-se $U \cong V$.

A.2.4 Módulo quociente

Seja U um R-módulo e V um submódulo de U. Define-se a relação $\equiv_V$ em U como segue:

$$x \equiv_V y \text{ se e somente se } x - y \in V.$$

Pode-se mostrar que $\equiv_V$ é uma relação de equivalência. A classe de equivalência de um elemento $u \in U$ é dada por

$$u + V \triangleq \{u + v : v \in V\}.$$

O conjunto de todas as classes de equivalência é denotado por U/V; esse conjunto torna-se um módulo, chamado de *módulo quociente*, se as operações de adição e multiplicação por escalar forem definidas, respectivamente, por

$$\begin{aligned}(x + V) + (y + V) &= (x + y) + V,\\ r(x + V) &= (rx) + V,\end{aligned}$$

em que $r \in R$. A identidade aditiva de U/V é $0 + V = V$.

A.2.5 Soma e intersecção de submódulos

Seja W um R-módulo e U e V submódulos de W. A *soma* de U e V é definida por

$$U + V = \{u + v : u \in U, v \in V\},$$

e a *intersecção* de U e V é definida por

$$U \cap V = \{w \in W : w \in U \text{ e } w \in V\}.$$

É possível mostrar que tanto $U+V$ quanto $U\cap V$ são submódulos de W e que $U \cap V \subseteq U, V \subseteq U + V$.

A.2.6 Produto de módulos

Sejam $U_1, \ldots, U_n$ R-módulos. O produto cartesiano $U_1 \times \cdots \times U_n$ torna-se um R-módulo se as operações de adição e multiplicação por escalar forem definidas entrada-a-entrada, ou seja,

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),$$
$$r(x_1, \ldots, x_n) = (rx_1, \ldots, rx_n).$$

para todo $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in U_1 \times \cdots \times U_n$ e $r \in R$. O R-módulo $U_1 \times \cdots \times U_n$ é chamado de *produto* de $U_1, \ldots, U_n$. A identidade aditiva de $U_1 \times \cdots \times U_n$ é $(0, \ldots, 0)$.

A.2.7 Módulos sobre domínios de ideais principais

Suponha que R seja um domínio de ideais principais. Se U for um R-módulo finitamente gerado, então

$$U \cong T/\langle d_1 \rangle \times T/\langle d_2 \rangle \times \cdots \times T/\langle d_\ell \rangle,$$

em que $d_1, d_2, \ldots, d_\ell \in R$ são elementos não-inversíveis satisfazendo $d_1 \mid d_2 \mid \cdots \mid d_\ell$. Adicionalmente, tal representação é única a menos da multiplicação dos d_is por elementos inversíveis.

APÊNDICE B

Resultados auxiliares

B.1 Uma variação do cripto-lema

Os resultados aqui apresentados foram obtidos em conjunto com Chen Feng (*University of Toronto*). Inicia-se revisando o já conhecido *cripto-lema* para o caso de grupos finitos [14].

Lema B.1. *Seja $\mathcal{G}$ um grupo finito, com operação binária $\cdot : \mathcal{G}\times\mathcal{G} \to \mathcal{G}$. Sejam $\boldsymbol{x}$ e $\boldsymbol{g}$ variáveis aleatórias sobre $\mathcal{G}$. Se $\boldsymbol{g}$ for uniforme sobre $\mathcal{G}$ e independente de $\boldsymbol{x}$, então $\boldsymbol{y} = \boldsymbol{g} \cdot \boldsymbol{x}$ também será uniforme sobre $\mathcal{G}$ e independente de $\boldsymbol{x}$, qualquer que seja a distribuição de $\boldsymbol{x}$.*

Lembre-se que uma *ação (à esquerda)* de $\mathcal{G}$ em um dado conjunto $\mathcal{S}$ é uma operação binária $\circ : \mathcal{G} \times \mathcal{S} \to \mathcal{S}$ tal que

(*i*) $(g_1 \cdot g_2) \circ x = g_1 \circ (g_2 \circ x)$, para todo $g_1, g_1 \in \mathcal{G}$ e $x \in \mathcal{S}$, e

(*ii*) $e \circ x = x$, para todo $x \in \mathcal{S}$, em que e é a identidade de $\mathcal{G}$.

Todo grupo $\mathcal{G}$ age em si mesmo ($\mathcal{S} = \mathcal{G}$) por multiplicação (à esquerda), ou seja, através da ação dada por $g \circ x = g \cdot x$. Esta seção generaliza o cripto-lema desse caso especial para o caso de uma ação qualquer de $\mathcal{G}$ sobre um conjunto finito $\mathcal{S}$.

Antes de prosseguir, apresenta-se uma breve revisão sobre ações de grupo (para mais detalhes veja, por exemplo, [8, §4.1]). Para todo $x \in \mathcal{S}$, a *órbita* de $\mathcal{G}$ contendo x é definida por $\mathcal{G} \circ x \triangleq \{g \circ x : g \in \mathcal{G}\}$. A relação em $\mathcal{S}$ dada por

$$x \sim y \qquad \text{se e somente se} \qquad x = g \circ y \text{ para algum } g \in \mathcal{G}$$

é uma relação de equivalência. Tem-se $x \sim y$ se e somente se $\mathcal{G}\circ x = \mathcal{G}\circ y$ se e somente se x e y estiverem na mesma órbita. O tamanho de uma órbita é dado por $|\mathcal{G} \circ x| = |\mathcal{G}|/|\mathcal{G}_{x,x}|$, em que $\mathcal{G}_{x,x} \triangleq \{g \in \mathcal{G} : g \circ x = x\}$ é o *estabilizador* de x em $\mathcal{G}$ (um subgrupo de $\mathcal{G}$). Uma ação é dita ser *transitiva* se possuir apenas uma órbita.

Lema B.2. *Sejam $\mathcal{G}$ um grupo finito, $\mathcal{S}$ um conjunto finito e $\circ : \mathcal{G} \times \mathcal{S} \to \mathcal{S}$ uma ação de $\mathcal{G}$ em $\mathcal{S}$. Sejam $\boldsymbol{x}$ e $\boldsymbol{g}$ variáveis aleatórias sobre $\mathcal{S}$ e $\mathcal{G}$, respectivamente. Se $\boldsymbol{g}$ for uniforme sobre $\mathcal{G}$ e independente de $\boldsymbol{x}$, então $\boldsymbol{y} = \boldsymbol{g} \circ \boldsymbol{x}$ (de modo que $\boldsymbol{x}$ e $\boldsymbol{y}$ estão na mesma órbita) será uniforme por partes sobre as órbitas da ação e condicionalmente independente de $\boldsymbol{x}$ dada a órbita*[1].

Demonstração. Como $\boldsymbol{g}$ é uniforme e independente de $\boldsymbol{x}$, tem-se, para todo $x, y \in \mathcal{S}$,

$$p_{\boldsymbol{y}|\boldsymbol{x}}(y|x) = \frac{|\mathcal{G}_{x,y}|}{|\mathcal{G}|},$$

em que $\mathcal{G}_{x,y} \triangleq \{g \in \mathcal{G} : g\circ x = y\}$. Se $x \sim y$ (de modo que $\mathcal{G}\circ x = \mathcal{G}\circ y$), então pode ser mostrado que $\mathcal{G}_{x,y}$ será um *coset* do estabilizador $\mathcal{G}_{x,x}$, o que implica que $|\mathcal{G}_{x,y}| = |\mathcal{G}_{x,x}|$ e, portanto,

$$p_{\boldsymbol{y}|\boldsymbol{x}}(y|x) = \frac{|\mathcal{G}_{x,x}|}{|\mathcal{G}|} = \frac{1}{|\mathcal{G} \circ x|} = \frac{1}{|\mathcal{G} \circ y|}.$$

Por outro lado, se $x \nsim y$, então $p_{\boldsymbol{y}|\boldsymbol{x}}(y|x) = 0$. Assim,

$$p_{\boldsymbol{y}}(y) = \sum_x p_{\boldsymbol{y}|\boldsymbol{x}}(y|x)p_{\boldsymbol{x}}(x) = \frac{1}{|\mathcal{G} \circ y|} \sum_{x:x\sim y} p_{\boldsymbol{x}}(x) = \frac{\Pr[\boldsymbol{y} \sim y]}{|\mathcal{G} \circ y|},$$

do qual segue o lema. □

[1] Em particular, se a ação for transitiva, então $\boldsymbol{y}$ é uniforme sobre $\mathcal{S}$ (a única órbita) e independente de $\boldsymbol{x}$. Esse é o caso da ação $g \circ x = g \cdot x$, de modo que o Lema B.1 é recuperado.

B.2 Agrupamentos sem perda de informação em DMCs

Uma questão natural no contexto de DMCs diz respeito às condições nas quais duas ou mais letras de entrada (ou saída) podem ser "agrupadas" sem que haja redução na informação mútua do canal. O seguinte resultado (veja [1, §5.9–5.10]) fornece a resposta.

Lema B.3. *Seja $(\mathcal{X}, p_{\boldsymbol{y}|\boldsymbol{x}}, \mathcal{Y})$ um DMC com entrada $\boldsymbol{x}$ e saída $\boldsymbol{y}$. Adicionalmente, sejam $f : \mathcal{X} \to \mathcal{U}$ e $g : \mathcal{Y} \to \mathcal{V}$ funções sobrejetoras e defina $\boldsymbol{u} = f(\boldsymbol{x})$ e $\boldsymbol{v} = f(\boldsymbol{y})$. Então, vale o seguinte:*

(i) $I(\boldsymbol{x};\boldsymbol{y}) = I(\boldsymbol{u};\boldsymbol{y})$ para todo $p_{\boldsymbol{x}}$ se e somente se, para todo par $x, x' \in \mathcal{X}$ tal que $f(x) = f(x')$, valer $p_{\boldsymbol{y}|\boldsymbol{x}}(y|x) = p_{\boldsymbol{y}|\boldsymbol{x}}(y|x')$ para todo $y \in \mathcal{Y}$.

(ii) $I(\boldsymbol{x};\boldsymbol{y}) = I(\boldsymbol{x};\boldsymbol{v})$ para todo $p_{\boldsymbol{x}}$ se e somente se, para todo par $y, y' \in \mathcal{Y}$ tal que $g(y) = g(y')$, existir um número real α tal que $p_{\boldsymbol{y}|\boldsymbol{x}}(y'|x) = \alpha\, p_{\boldsymbol{y}|\boldsymbol{x}}(y|x)$ para todo $x \in \mathcal{X}$.

De forma mais intuitiva, o Lema B.3 afirma o seguinte:

(*i*) Se duas ou mais *linhas* da matriz de probabilidade de transição do canal forem *iguais*, então pode-se manter apenas uma delas, descartando-se as demais.

(*ii*) Se duas ou mais *colunas* da matriz de probabilidade de transição do canal forem *múltiplas* umas das outras, então elas podem ser substituídas pela sua soma.

O exemplo a seguir ilustra isso.

Exemplo. Considere um DMC definido por $(\mathcal{X}, p_{\boldsymbol{y}|\boldsymbol{x}}, \mathcal{Y})$, em que $\mathcal{X} = \{x_1, x_2, \ldots, x_5\}$, $\mathcal{Y} = \{y_1, y_2, \ldots, y_5\}$ e

$$p_{\boldsymbol{y}|\boldsymbol{x}} = \left[\begin{array}{c|ccccc} & y_1 & y_2 & y_3 & y_4 & y_5 \\ \hline x_1 & \frac{1}{18} & \frac{1}{3} & \frac{1}{9} & \frac{1}{3} & \frac{3}{18} \\ x_2 & \frac{1}{18} & \frac{1}{3} & \frac{1}{9} & \frac{1}{3} & \frac{3}{18} \\ x_3 & \frac{1}{6} & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ x_4 & \frac{1}{6} & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ x_5 & \frac{1}{9} & \frac{1}{6} & \frac{2}{9} & \frac{1}{6} & \frac{1}{3} \end{array}\right]$$

As linhas x_1 e x_2 são iguais; o mesmo ocorre com as linhas x_3 e x_4. As colunas y_1, y_3 e y_5 são múltiplas umas das outras; o mesmo ocorre com as colunas y_2 e y_4 (nesse último caso, as colunas são, de fato, iguais). Então, de acordo com o Lema B.3, as funções $f : \mathcal{X} \to \mathcal{U} = \{u_1, u_2, u_3\}$, definida por

$$f(x_1) = f(x_2) = u_1, \quad f(x_3) = f(x_4) = u_2, \quad f(x_5) = u_3,$$

e $g : \mathcal{Y} \to \mathcal{V} = \{v_1, v_2\}$, definida por

$$g(y_1) = g(y_3) = g(y_5) = v_1, \quad g(y_2) = g(y_4) = v_2,$$

são tais que $I(\boldsymbol{x}; \boldsymbol{y}) = I(\boldsymbol{u}; \boldsymbol{v})$, em que $\boldsymbol{u} = f(\boldsymbol{x})$ e $\boldsymbol{v} = g(\boldsymbol{y})$, qualquer que seja a distribuição de entrada $p_{\boldsymbol{x}}$. Assim, é possível simplificar a matriz de probabilidade de transição como segue:

$$\left[\begin{array}{c|ccccc} & y_1 & y_2 & y_3 & y_4 & y_5 \\ \hline x_1 & \frac{1}{18} & \frac{1}{3} & \frac{1}{9} & \frac{1}{3} & \frac{3}{18} \\ x_2 & \frac{1}{18} & \frac{1}{3} & \frac{1}{9} & \frac{1}{3} & \frac{3}{18} \\ x_3 & \frac{1}{6} & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ x_4 & \frac{1}{6} & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ x_5 & \frac{1}{9} & \frac{1}{6} & \frac{2}{9} & \frac{1}{6} & \frac{1}{3} \end{array}\right] \to \left[\begin{array}{c|ccccc} & y_1 & y_2 & y_3 & y_4 & y_5 \\ \hline u_1 & \frac{1}{18} & \frac{1}{3} & \frac{1}{9} & \frac{1}{3} & \frac{3}{18} \\ u_2 & \frac{1}{6} & 0 & \frac{1}{3} & 0 & \frac{1}{2} \\ u_3 & \frac{1}{9} & \frac{1}{6} & \frac{2}{9} & \frac{1}{6} & \frac{1}{3} \end{array}\right] \to \left[\begin{array}{c|cc} & v_1 & v_2 \\ \hline u_1 & \frac{1}{3} & \frac{2}{3} \\ u_2 & 1 & 0 \\ u_3 & \frac{2}{3} & \frac{1}{3} \end{array}\right],$$

em que, no primeiro passo, agrupou-se a entrada e, no segundo passo, a saída. □

B.3 Um resultado de álgebra linear

Esta seção apresenta um resultado básico de álgebra linear. Agradecemos ao Prof. Bill Martin (*Worcester Polytechnic Institute*) pelo auxílio na demonstração.

Lema B.4. *Seja F um corpo. Sejam $X, X' \in F^{n\times\ell}$ e $Y, Y' \in F^{m\times\ell}$ tais que $\operatorname{rank} X = \operatorname{rank} X'$, $\operatorname{rank} Y = \operatorname{rank} Y'$ e $\dim(\langle X\rangle \cap \langle Y\rangle) = \dim(\langle X'\rangle \cap \langle Y'\rangle)$. Então, existem $P \in \mathrm{GL}_n(F)$, $Q \in \mathrm{GL}_m(F)$ e $T \in \mathrm{GL}_\ell(F)$ tais que $X' = PXT$ e $Y' = QYT$.*

Demonstração. Sejam $u = \operatorname{rank} X$, $v = \operatorname{rank} Y$ e $w = \dim(\langle X\rangle \cap \langle Y\rangle)$.

Da álgebra linear, sabe-se que

$$u + v - w = \dim(\langle X \rangle + \langle Y \rangle) \leq \ell$$

e que existem matrizes de posto completo $A_0, A_0' \in F^{w \times \ell}$, $A_1, A_1' \in F^{(u-w)\times \ell}$, $A_2, A_2' \in F^{(v-w)\times \ell}$ e $A_3, A_3' \in F^{(\ell-u-v+w)\times \ell}$ tais que

$$\langle A_0 \rangle = \langle X \rangle \cap \langle Y \rangle, \quad \left\langle \begin{bmatrix} A_0 \\ A_1 \end{bmatrix} \right\rangle = \langle X \rangle, \quad \left\langle \begin{bmatrix} A_0 \\ A_2 \end{bmatrix} \right\rangle = \langle Y \rangle, \quad \langle A \rangle = F^{\ell},$$

e

$$\langle A_0' \rangle = \langle X' \rangle \cap \langle Y' \rangle, \quad \left\langle \begin{bmatrix} A_0' \\ A_1' \end{bmatrix} \right\rangle = \langle X' \rangle, \quad \left\langle \begin{bmatrix} A_0' \\ A_2' \end{bmatrix} \right\rangle = \langle Y' \rangle, \quad \langle A' \rangle = F^{\ell},$$

em que

$$A = \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{bmatrix} \quad \text{e} \quad A' = \begin{bmatrix} A_0' \\ A_1' \\ A_2' \\ A_3' \end{bmatrix}.$$

Como $\langle A \rangle = \langle A' \rangle = F^{\ell}$, tem-se que $A, A' \in \mathrm{GL}_{\ell}(F)$, de modo que existe $T \in \mathrm{GL}_{\ell}(F)$ tal que $A' = AT$. Daí, tem-se $A_i' = A_i T$, para $0 \leq i \leq 3$. Adicionalmente, tem-se

$$X = P_1 \begin{bmatrix} A_0 \\ A_1 \\ 0_{(n-u)\times \ell} \end{bmatrix}, \quad X' = P_2 \begin{bmatrix} A_0' \\ A_1' \\ 0_{(n-u)\times \ell} \end{bmatrix},$$

em que $P_1, P_2 \in \mathrm{GL}_n(F)$ e

$$Y = Q_1 \begin{bmatrix} A_0 \\ A_2 \\ 0_{(m-v)\times \ell} \end{bmatrix}, \quad Y' = Q_2 \begin{bmatrix} A_0' \\ A_2' \\ 0_{(m-v)\times \ell} \end{bmatrix},$$

em que $Q_1, Q_2 \in \mathrm{GL}_m(F)$. Portanto,

$$X' = P_2 \begin{bmatrix} A'_0 \\ A'_1 \\ 0 \end{bmatrix} = P_2 \begin{bmatrix} A_0 T \\ A_1 T \\ 0 \end{bmatrix} = P_2 \begin{bmatrix} A_0 \\ A_1 \\ 0 \end{bmatrix} T = P_2 P_1^{-1} X T,$$

e

$$Y' = Q_2 \begin{bmatrix} A'_0 \\ A'_2 \\ 0 \end{bmatrix} = Q_2 \begin{bmatrix} A_0 T \\ A_2 T \\ 0 \end{bmatrix} = Q_2 \begin{bmatrix} A_0 \\ A_2 \\ 0 \end{bmatrix} T = Q_2 Q_1^{-1} Y T.$$

O resultado agora segue definindo $P = P_2 P_1^{-1}$ e $Q = Q_2 Q_1^{-1}$. □

Como consequência, tem-se o seguinte.

Lema B.5. *Sejam $X, X' \in \mathcal{T}_u(F^{n\times\ell})$ e $Y, Y' \in \mathcal{T}_v(F^{m\times\ell})$ tais que $\langle Y\rangle \subseteq \langle X\rangle$ e $\langle Y'\rangle \subseteq \langle X'\rangle$. Então, existem $P \in \mathrm{GL}_n(F)$, $Q \in \mathrm{GL}_m(F)$ e $T \in \mathrm{GL}_\ell(F)$ tais que $X' = PXT$ e $Y' = QYT$.*

Referências bibliográficas

[1] N. Abramson, *Information Theory and Coding.* McGraw-Hill, 1963.

[2] R. Ahlswede, N. Cai, R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*, 2nd ed. Cambridge University Press, 2004.

[4] W. C. Brown, *Matrices over Commutative Rings*, ser. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., 1992, vol. 169.

[5] P. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proceedings of the 41st Annual Allerton Conference on Communication, Control, and Computing (Allerton'03)*, Monticello, Illinois, Oct. 2003.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.

[7] G. P. Dresden, "Small rings," available at http://home.wlu.edu/~dresdeng/smallrings/.

[8] D. S. Dummit and R. M. Foote, *Abstract Algebra*, 3rd ed. John Wiley and Sons, 2004.

[9] T. Etzion and A. Vardy, "Error-correcting codes in projective space," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 1165–1173, Feb. 2011.

[10] C. Feng, R. W. Nóbrega, F. R. Kschischang, and D. Silva, "Communication over finite-chain-ring matrix channels," Apr. 2013, submitted to the IEEE Transactions on Information Theory. Available at http://arxiv.org/abs/1304.2523.

[11] ——, "Communication over finite-ring matrix channels," in *Proceedings of the 2013 IEEE International Symposium on Information Theory (ISIT'13)*, Istanbul, Turkey, Jul. 2013, pp. 2890–2894.

[12] C. Feng, D. Silva, and F. R. Kschischang, "An algebraic approach to physical-layer network coding," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7576–7596, Nov. 2013.

[13] S. D. Fisher and M. N. Alexander, "Matrices over a finite field," *American Mathematical Monthly*, vol. 73, pp. 639–641, Jun. 1966.

[14] G. D. Forney Jr., "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener," in *Proceedings of the 41st Annual Allerton Conference on Communication, Control, and Computing (Allerton'03)*, Monticello, Illinois, Oct. 2003.

[15] C. Fragouli and E. Soljanin, *Network Coding Applications*, ser. Foundations and Trends® in Networking. Now Publishers Inc., 2007.

[16] ——, *Network Coding Fundamentals*, ser. Foundations and Trends® in Networking. Now Publishers Inc., 2007.

[17] A. J. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[18] T. Ho, R. Koetter, M. Médard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," in *Proceedings of the 2003 IEEE International Symposium on Information Theory (ISIT'03)*, Yokohama, Japan, Jun. 2003, p. 442.

[19] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

[20] T. Honold and I. Landjev, "Linear codes over finite chain rings," *The Electronic Journal of Combinatorics*, vol. 7, 2000.

[21] T. W. Hungerford, *Algebra*, ser. Graduate Texts in Mathematics. Springer, 1974, vol. 73.

[22] M. Jafari Siavoshani, C. Fragouli, and S. Diggavi, "Noncoherent multisource network coding," in *Proceedings of the 2008 IEEE International Symposium on Information Theory (ISIT'08)*, Toronto, Canada, Jul. 2008, pp. 817–821.

[23] M. Jafari Siavoshani, S. Mohajer, C. Fragouli, and S. Diggavi, "On the capacity of non-coherent network coding," in *Proceedings of the 2009 IEEE International Symposium on Information Theory (ISIT'09)*, Seoul, South Korea, Jun. 2009, pp. 273–277.

[24] ——, "On the capacity of non-coherent network coding," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 1046–1066, Feb. 2011.

[25] M. Jafari Siavoshani, S. Yang, and R. W. Yeung, "Non-coherent network coding: An arbitrarily varying channel approach," in *Proceedings of the 2012 IEEE International Symposium on Information Theory (ISIT'12)*, Cambridge, Massachusetts, Jul. 2012, pp. 1672–1676.

[26] A. Khaleghi, D. Silva, and F. R. Kschischang, "Subspace codes," in *Cryptography and Coding*, ser. Lecture Notes in Computer Science, M. G. Parker, Ed. Berlin, Germany: Springer Berlin / Heidelberg, Dec. 2009, vol. 5921, pp. 1–21.

[27] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[28] R. Kötter and F. R. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3579–3591, Aug. 2008.

[29] R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[30] S. C. Liew, S. Zhang, and L. Lu, "Physical-layer network coding: Tutorial, survey, and beyond," *Physical Communication*, vol. 6, pp. 4–42, May 2013.

[31] B. R. McDonald, *Finite Rings with Identity*, ser. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., 1974, vol. 28.

[32] A. Montanari and R. L. Urbanke, "Coding for network coding," *Computing Research Repository (CoRR)*, vol. abs/0711.3935, Nov. 2007, available at http://arxiv.org/abs/0711.3935.

[33] ——, "Iterative coding for network coding," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1563–1572, Mar. 2013.

[34] B. Nazer and M. Gastpar, "Computing over multiple-access channels with connections to wireless network coding," in *Proceedings of the 2006 IEEE International Symposium on Information Theory (ISIT'06)*, Seattle, Washington, Jul. 2006, pp. 1354–1358.

[35] ——, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[36] ——, "Reliable physical layer network coding," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 438–460, Mar. 2011.

[37] A. A. Nechaev, "Finite rings with applications," in *Handbook of Algebra*, M. Hazewinkel, Ed. North-Holland, 2008, vol. 5, pp. 213–320.

[38] C. Nöbauer, "The number of small rings," technical report. Available at ftp://ftp.mathe2.uni-bayreuth.de/axel/papers/noebauer:the_number_of_small_rings.ps.

[39] R. W. Nóbrega, C. Feng, D. Silva, and B. F. Uchôa-Filho, "Canais matriciais multiplicativos sobre anéis de cadeia finitos," in *Proceedings of the XXXI Simpósio Brasileiro de Telecomunicações (SBrT'13)*, Fortaleza, Brazil, Sep. 2013.

[40] ——, "On multiplicative matrix channels over finite chain rings," in *Proceedings of the 2013 IEEE International Symposium on Network Coding (NetCod'13)*, Calgary, Alberta, Jun. 2013.

[41] R. W. Nóbrega, D. Silva, and B. F. Uchôa-Filho, "On the capacity of multiplicative finite-field matrix channels," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 4949–4960, Aug. 2013.

[42] R. W. Nóbrega, B. F. Uchôa-Filho, and D. Silva, "On the capacity of multiplicative finite-field matrix channels," in *Proceedings of the 2011 IEEE International Symposium on Information Theory (ISIT'11)*, Saint Petersburg, Russia, Jul. 2011, pp. 248–252.

[43] J. B. Pedersen and F. Topsøe, "Block symmetry in discrete memoryless channels," in *Proceedings of the 2002 IEEE Information Theory Workshop (ITW'02)*, Bangalore, India, Oct. 2002, pp. 131–134.

[44] P. Popovski and H. Yomo, "The anti-packets can increase the achievable throughput of a wireless multi-hop network," in *Proceedings of the 2006 IEEE International Conference on Communications (ICC'06*, vol. 9, Istanbul, Turkey, Jun. 2006, pp. 3885–3890.

[45] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, Jul. and Oct. 1948.

[46] D. Silva, "Error control for network coding," Ph.D. dissertation, University of Toronto, Toronto, Ontario, Canada, Feb. 2009.

[47] D. Silva and F. R. Kschischang, "On metrics for error correction in network coding," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5479–5490, Dec. 2009.

[48] D. Silva, F. R. Kschischang, and R. Kötter, "A rank-metric approach to error control in random network coding," *IEEE Transac-*

tions on Information Theory, vol. 54, no. 9, pp. 3951–3967, Sep. 2008.

[49] ——, "Communication over finite-field matrix channels," *IEEE Transactions on Information Theory*, vol. 56, no. 2, pp. 1296–1305, Mar. 2010.

[50] D. Silva, F. R. Kschischang, and R. Kötter, "Capacity of random network coding under a probabilistic error model," in *Proceedings of the 24th Biennial Symposium on Communications*, Kingston, Canada, Jun. 2008, pp. 9–12.

[51] B. F. Uchôa-Filho and R. W. Nóbrega, "The capacity of random linear coding networks as subspace channels," *Computing Research Repository (CoRR)*, vol. abs/1001.1021, Jan. 2010, available at http://arxiv.org/abs/1001.1021.

[52] S. Yang, "Superposition coding for linear operator channels over finite fields," in *Proceedings of the 2012 IEEE Information Theory Workshop (ITW'12)*, Lausanne, Switzerland, Sep. 2012, pp. 507–511.

[53] S. Yang, S.-W. Ho, J. Meng, and E.-h. Yang, "Optimality of subspace coding for linear operator channels over finite fields," in *Proceedings of the 2010 IEEE Information Theory Workshop (ITW'10)*, Cairo, Egypt, Jan. 2010, pp. 400–404.

[54] ——, "Capacity analysis of linear operator channels over finite fields," *Computing Research Repository (CoRR)*, vol. abs/1108.4257, Dec. 2012, available at http://arxiv.org/abs/1108.4257.

[55] S. Yang, S.-W. Ho, J. Meng, E.-h. Yang, and R. W. Yeung, "Linear operator channels over finite fields," *Computing Research Repository (CoRR)*, vol. abs/1002.2293, Apr. 2010, available at http://arxiv.org/abs/1002.2293.

[56] S. Yang, J. Meng, and E.-h. Yang, "Coding for linear operator channels over finite fields," in *Proceedings of the 2010 IEEE International Symposium on Information Theory (ISIT'10)*, Austin, Texas, Jun. 2010, pp. 2413–2417.

[57] S. Yang and R. W. Yeung, "Coding for a network coded fountain," in *Proceedings of the 2011 IEEE International Symposium on Information Theory (ISIT'11)*, Saint Petersburg, Russia, Jul. 2011, pp. 2583–2587.

[58] R. W. Yeung, R. Li, N. Cai, and Z. Zhang, *Network Coding Theory*, ser. Foundations and Trends® in Communications and Information Theory. Now Publishers Inc., 2006.

[59] O. Zariski and P. Samuel, *Commutative Algebra*. D. Van Nostrand Company, Inc., 1958, vol. I.

[60] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: Physical-layer network coding," in *Proceedings of the 12th ACM Annual International Conference on Mobile Computing and Networking (MobiCom'06)*, Los Angeles, California, Sep. 2006, pp. 358–365.