Modern Digital Satellite Television: How It Works

Matthew C. Valenti

Lane Department of Computer Science and Electrical Engineering
West Virginia University
U.S.A.

Nov, 1, 2011
Outline

1. Satellite Television Standards
2. DVB-S2 Modulation
3. LDPC Coding
4. Constellation Shaping
5. Conclusion
Outline

1 Satellite Television Standards

2 DVB-S2 Modulation

3 LDPC Coding

4 Constellation Shaping

5 Conclusion
Digital Satellite Television in the United States

DirecTV
- Spinoff of Hughes Network Systems.
- Began operations in 1994.
- 23,000 employees in U.S. and Latin America.
- $33.6 billion market cap.

Dish Network.
- Spinoff of EchoStar.
- Began operations in 1996.
- 14.1 subscribers in 2010.
- 22,000 employees.
- $10.9 Billion market cap.
DVB is a family of open standards for digital video broadcasting.
- Maintained by 270-member industry consortium.
- Published by ETSI.

Modes of transmission
- Satellite: DVB-S, DVB-S2, and DVB-SH
- Cable: DVB-C, DVB-C2
- Terrestrial: DVB-T, DVB-T2, DVB-H
- Modulation: QPSK with $\alpha = 0.35$ rolloff.
- Channel coding: Concatenated Reed Solomon and convolutional.
DVB-S2 was introduced in 2003 with the following goals:

- Improve spectral efficiency by 30% through better modulation and coding.
  - Modulation: QPSK, 8PSK, 16/32 APSK.
  - Channel coding: LDPC with outer BCH code.
- Offer a more diverse range of services.
  - HDTV broadcast television.
  - Backhaul applications, e.g., electronic news gathering.
  - Internet downlink access.
  - Large-scale data content distribution, e.g., electronic newspapers.

Ratified 2005.
Adaptive Internet Downlink
Outline

1. Satellite Television Standards
2. DVB-S2 Modulation
3. LDPC Coding
4. Constellation Shaping
5. Conclusion
Why Not Use QAM?

- Higher spectral-efficiencies require larger signal constellations.
- However, nonlinear satellite channels are not well suited to square QAM.
The DVB-S2 Signal Constellations

The primary objective of DVB-S2 was to bring 8PSK within reach of consumer-size satellite dishes and the increase in spectrum efficiency that 8PSK brings. With this work came the acceptance that real-world transmission facts should be taken into account for the new system design.

Rather than considering only the standard linear channel as with the previous DVB-S and DVB-D satellite specifications, the DVB-S2 specification recognises the following effects:

- Phase noise
- Non-linear magnitude and phase characteristics of a saturated transponder
- The fact that the transponder is power limited
- Group delay effects

This work led to the defined constellations to be optimised for the above conditions. The constellations that were chosen are shown below:

- QPSK
- 8PSK
- 16APSK
- 32APSK

Figure 1: DVB-S2 Constellations

16APSK and 32APSK were chosen over the more familiar QAM constellations because their round shape makes them more power efficient in the power-limited channel that is saturated transponder. It is interesting to note that the ratio of the radius of the concentric circles for 16- and 32APSK changes slightly depending upon the FEC that is used in order to achieve maximum performance.
DVB-S2 uses a tighter root RC-rolloff filter.

- \( B = R_s(1 + \alpha) \)
- Assuming a 6 MHz transponder channel...
- DVB-S Example:
  - \( \alpha = 0.35 \).
  - QPSK: \( R_b = 2R_s \)
  - \( 2(6)/(1.35) = 8.9 \) Mbps
- DVB-S2 Example:
  - \( \alpha = 0.20 \).
  - 32-APSK: \( R_b = 5R_s \)
  - \( 5(6)/(1.2) = 25 \) Mbps
BER in AWGN

![Graph showing BER vs Es/No in dB for different modulation schemes (32APSK, 16APSK, 8PSK, QPSK). The x-axis represents Es/No in dB, while the y-axis represents BER. The graph shows the performance of each modulation scheme at various Es/No values, with 32APSK having the best performance, followed by 16APSK, 8PSK, and QPSK in that order.](image-url)
Outline

1. Satellite Television Standards
2. DVB-S2 Modulation
3. LDPC Coding
4. Constellation Shaping
5. Conclusion
Performance can be improved by using error control coding.

Gains are limited by the modulation-constrained capacity.

LDPC codes are capable of approaching capacity.
Available Code Rates

- The encoder maps length-$k$ messages to length-$n$ codewords.
- The code rate is $R = k/n$.
- Useful bit rate is $R_u = R \log_2(M)$.
- Two codeword lengths:
  - 16, 200.
  - 64, 800.
DVB-S2 vs. Shannon

Dotted lines: modulation constrained Shannon limit.

- QPSK
- 8PSK
- 16APSK
- 32APSK
- DVB-S
- DVB-DSNG

Graph showing the channel capacity $R_c$ (in bits per unit symbol) versus the signal-to-noise ratio $C/N$ (in $R_s$ units) for various modulation schemes.
Consider the following rate $R = 5/6$ single parity-check code:

$$c = [1 \ 0 \ 1 \ 0 \ 1 \ 1]$$

One error in any position may be detected:

$$c = [1 \ 0 \ X \ 0 \ 1 \ 1]$$

Problem with using an SPC is that it can only detect a single error.
Product Codes

- Place data into a $k \times k$ rectangular array.
  - Encode each row with a SPC.
  - Encode each column with a SPC.
  - Result is a rate $R = \frac{k^2}{(k + 1)^2}$ code.

Example $k = 2$.

$$
\begin{array}{ccc}
  c_1 = u_1 & c_2 = u_2 & c_3 = c_1 \oplus c_2 \\
  c_4 = u_3 & c_5 = u_4 & c_6 = c_4 \oplus c_5 \\
  c_7 = c_1 \oplus c_4 & c_8 = c_2 \oplus c_5 & c_9 = c_3 \oplus c_6
\end{array}
= \begin{bmatrix}
1 & 0 & 1 \\
1 & 1 & 0 \\
0 & 1 & 1
\end{bmatrix}
$$

- A single error can be corrected by detecting its row and column location.
The example product code is characterized by the set of five linearly-independent equations:

\[
\begin{align*}
    c_3 & = c_1 \oplus c_2 & \Rightarrow & & c_1 \oplus c_2 \oplus c_3 = 0 \\
    c_6 & = c_4 \oplus c_5 & \Rightarrow & & c_4 \oplus c_5 \oplus c_6 = 0 \\
    c_7 & = c_1 \oplus c_4 & \Rightarrow & & c_1 \oplus c_4 \oplus c_7 = 0 \\
    c_8 & = c_2 \oplus c_5 & \Rightarrow & & c_2 \oplus c_4 \oplus c_8 = 0 \\
    c_9 & = c_3 \oplus c_6 & \Rightarrow & & c_3 \oplus c_6 \oplus c_9 = 0
\end{align*}
\]

In general, it takes \((n - k)\) linearly-independent equations to specify a linear code.
Parity-check Matrices

- The system of equations may be expressed in matrix form as:

\[ cH^T = 0 \]

where \( H \) is a parity-check matrix.

Example:

\[
\begin{align*}
    c_1 \oplus c_2 \oplus c_3 &= 0 \\
    c_4 \oplus c_5 \oplus c_6 &= 0 \\
    c_1 \oplus c_4 \oplus c_7 &= 0 \\& \Leftrightarrow \quad H = \\
    c_2 \oplus c_4 \oplus c_8 &= 0 \\
    c_3 \oplus c_6 \oplus c_9 &= 0 \\
\end{align*}
\]

System of equations

\[
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

Parity-check matrix
The parity-check matrix may be represented by a *Tanner* graph.

Bipartite graph:
- Check nodes: Represent the $n - k$ parity-check equations.
- Variable nodes: Represent the $n$ code bits.

If $H_{i,j} = 1$, then $i^{th}$ check node is connected to $j^{th}$ variable node.

Example: For the parity-check matrix:

\[
H = \begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1
\end{bmatrix}
\]

The Tanner Graph is:
Observations:
- To achieve capacity, a long code is needed.
- The decoder’s complexity depends on the number of edges in the Tanner graph.
- The number of edges is equal to the number of zeros in $H$.
- It is desirable to have a code that is long, yet has a sparse $H$.

Low-density parity-check codes:
- An LDPC code is characterized by a sparse parity-check matrix.
- The row/column weights are independent of length.
- Decoder complexity grows only linearly with block length.

Historical note:
- LDPC codes were the subject of Robert Gallager’s 1960 dissertation.
- Were forgotten because the decoder could not be implemented.
- Were “rediscovered” in the mid-1990’s after turbo codes were developed.
Example LDPC Code

- A code from MacKay and Neal (1996):

\[ H = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
\end{bmatrix} \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
\end{bmatrix} \]

- The code is *regular* because:
  - The rows have constant weight (check-nodes constant degree).
  - The columns have constant weight (variable-nodes constant degree).
- This is called a \((3, 4)\) *regular* LDPC code because the variable nodes have degree 3 and the check nodes have degree 4.
The binary erasure channel is a conceptual model used to explain the operation of LDPC codes.

The BEC has two inputs (data 0 and data 1) and three outputs (data 0, data 1, and erasure $e$).

A bit is erased with probability $\epsilon$.

A bit is correctly received with probability $1 - \epsilon$. 

![Diagram of the binary erasure channel](image)
Several erasures may be corrected by iteratively decoding the SPC on each row and column.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>1</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>e</td>
<td>1</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>1</td>
<td>e</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>e</td>
<td>e</td>
<td>0</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>

Received word

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>1</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>e</td>
<td>1</td>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>e</td>
<td>e</td>
<td>0</td>
<td>e</td>
<td>1</td>
</tr>
</tbody>
</table>

Row decoding

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>1</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Column decoding

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>1</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Row decoding
Decoding can be performed on the Tanner graph.

- Load the variable nodes with the observed code bits.
- Each check node $j$ sends a message to each of its connected variable nodes $i$.
  - The message is the modulo two sum of the bits associated with the connected variable nodes other than $i$ (if none are erased).
  - If a check node touches a single erasure, then it will become corrected.
- Iterate until all erasures corrected or no more corrections possible.
Decoding can be performed on the Tanner graph.

- Load the variable nodes with the observed code bits.
- Each check node \( j \) sends a *message* to each of its connected variable nodes \( i \).
  - The message is the modulo two sum of the bits associated with the connected variable nodes *other* than \( i \) (if none are erased).
  - If a check node touches a *single* erasure, then it will become corrected.
- Iterate until all erasures corrected or no more corrections possible.
Decoding can be performed on the Tanner graph.

- Load the variable nodes with the observed code bits.
- Each check node $j$ sends a message to each of its connected variable nodes $i$.
  - The message is the modulo two sum of the bits associated with the connected variable nodes other than $i$ (if none are erased).
  - If a check node touches a single erasure, then it will become corrected.
- Iterate until all erasures corrected or no more corrections possible.
Decoding can be performed on the Tanner graph.

- Load the variable nodes with the observed code bits.
- Each check node $j$ sends a message to each of its connected variable nodes $i$.
  - The message is the modulo two sum of the bits associated with the connected variable nodes other than $i$ (if none are erased).
  - If a check node touches a single erasure, then it will become corrected.
- Iterate until all erasures corrected or no more corrections possible.
Decoding can be performed on the Tanner graph.

- Load the variable nodes with the observed code bits.
- Each check node $j$ sends a message to each of its connected variable nodes $i$.
  - The message is the modulo two sum of the bits associated with the connected variable nodes other than $i$ (if none are erased).
  - If a check node touches a single erasure, then it will become corrected.
- Iterate until all erasures corrected or no more corrections possible.
Decoding can be performed on the Tanner graph.

- Load the variable nodes with the observed code bits.
- Each check node $j$ sends a message to each of its connected variable nodes $i$.
  - The message is the modulo two sum of the bits associated with the connected variable nodes other than $i$ (if none are erased).
  - If a check node touches a single erasure, then it will become corrected.
- Iterate until all erasures corrected or no more corrections possible.
A stopping set $\mathcal{V}$ is a set of erased variable nodes that cannot be corrected, regardless of the state of the other variable nodes.

Let $G$ be the neighbors of $\mathcal{V}$. Every check node in $G$ touches at least two variable nodes in $\mathcal{V}$. The minimum stopping set $\mathcal{V}_{\text{min}}$ is the stopping set containing the fewest variable nodes. Let $d_{\text{min}} = |\mathcal{V}_{\text{min}}|$ be the size of the minimum stopping set.

- There exists at least one pattern of $d_{\text{min}}$ erasures that cannot be corrected.
- The erasure correcting capability of the code is $d_{\text{min}} - 1$, which is the maximum number of erasures that can always be corrected.
A **stopping set** $\mathcal{V}$ is a set of erased variable nodes that cannot be corrected, regardless of the state of the other variable nodes.

Let $G$ be the neighbors of $\mathcal{V}$.

Every check node in $G$ touches at least two variable nodes in $\mathcal{V}$.

The *minimum* stopping set $\mathcal{V}_{\text{min}}$ is the stopping set containing the fewest variable nodes.

Let $d_{\text{min}} = |\mathcal{V}_{\text{min}}|$ be the size of the minimum stopping set.

- There exists at least one pattern of $d_{\text{min}}$ erasures that cannot be corrected.
- The erasure correcting capability of the code is $d_{\text{min}} - 1$, which is the maximum number of erasures that can always be corrected.
Density Evolution

For a \((d_v, d_c)\) regular code, the probability that a variable-node remains erased after the \(\ell^{th}\) iteration is

\[
\epsilon_{\ell} = \epsilon_0 \left( 1 - (1 - \epsilon_{\ell-1})^{d_c-1} \right)^{d_v-1}
\]

where \(d_v\) is the variable-node degree, \(d_c\) is the check-node degree, and the initial condition is \(\epsilon_0 = \epsilon\).

The above result assumes independent messages, which is achieved when the girth of the Tanner graph is sufficiently large.

If \(\epsilon_{\ell} \to 0\) as \(\ell \to \infty\) for a particular channel erasure probability \(\epsilon\), then a code drawn from the ensemble of all such \((d_v, d_c)\) regular LDPC codes will be able to correctly decode.

The threshold \(\epsilon^*\) is the maximum \(\epsilon\) for which \(\epsilon_{\ell} \to 0\) as \(\ell \to \infty\).

For the \((3, 6)\) regular code, the threshold is \(\epsilon^* = 0.4294\)
Proof of (1), Part I/II

- Decoding involves the exchange of messages between variable nodes and check nodes.
  - Let $p_{\uparrow}$ denote the probability of an erased message going up from the variable nodes to the check nodes.
  - Let $p_{\downarrow}$ denote the probability of an erased message going down from the check nodes to the variable nodes.

- Consider the degree $d_c$ check node.
  - An outgoing message sent over a particular edge is a function of the incoming messages arriving over the other $d_c - 1$ edges.
  - For the outgoing message to be correct, all $d_c - 1$ incoming messages must be correct.
  - The outgoing message will be an erasure if any of the $d_c - 1$ incoming messages is an erasure.
  - The probability of the check node sending an erasure is:

$$p_{\downarrow} = 1 - (1 - p_{\uparrow})^{d_c-1}$$  \(2\)
Proof of (1), Part II/II

- Consider the degree $d_v$ check node.
  - An outgoing message sent over a particular edge is a function of the incoming messages arriving over the other $d_v - 1$ edges.
  - An outgoing message will be an erasure if the variable node was initially erased \textit{and} all of the arriving messages are erasures.
  - The probability of the variable node sending an erasure is:
    
    \[ p_{\uparrow} = \epsilon_0 p_{\downarrow}^{d_v - 1} \]  \hspace{1cm} (3)

- Letting $\epsilon_\ell$ equal the value of $p_{\uparrow}$ after the $\ell^{th}$ iteration, and substituting (2) into (3) yields the recursion given by (1):

  \[ \epsilon_\ell = \epsilon_0 \left( 1 - (1 - \epsilon_{\ell-1})^{d_c - 1} \right)^{d_v - 1} \]
Density Evolution result for (3,6) LDPC code over BEC

\[
\varepsilon = 0.7
\]

\[
\varepsilon = 0.6
\]

\[
\varepsilon = 0.5
\]

\[
\varepsilon = 0.4
\]

\[
\varepsilon = 0.3
\]

Fig. 5. Density Evolution result for regular (3,6) LDPC code over BEC

Substituting for \(q_l - 1\):

\[
p_l = \varepsilon \lambda \left(1 - \rho \left(1 - p_l - 1\right)\right) \tag{18}
\]

Recursion equation (18) is the DE for irregular LDPC codes.

D. Threshold

The object of density evolution is to determine which channel parameters (e.g. \(\varepsilon\) in BEC) the message-passing decoder is likely to correct all of the error bits. We can find the maximum \(\varepsilon\) for one LDPC ensemble over BEC by Equation (18) with the iteration \(l\) approximating \(\infty\), assuming that the graphs are cycle free.
Code Realization

- Density evolution only describes the asymptotic performance of the ensemble of LDPC codes.
- Implementation requires that an $H$ matrix be generated by drawing from the ensemble of all $(d_v, d_c)$ LDPC codes.
- Goals of good $H$ design:
  - High girth.
  - Full rank.
  - Large minimum stopping set.
- If the girth is too low, the short cycles invalidate the iterative decoder.
- High girth achieved through girth conditioning algorithms such as progressive edge growth (PEG).
- If $H$ is not full rank, then the rate will be reduced according to the number of dependent equations.
- Small stopping sets give rise to an error floor.
- A database of good regular LDPC codes can be found on MacKay’s website.
Performance of an Actual Code

Simulation result of length 8000 regular (3,6) ldpc code over BEC

Through the results in Fig. 10, irregular code has better threshold while yields more complexity during the decode complexity. Generally speaking, regular code has smaller threshold but also smaller gap between the simulation results and theoretical result. Other algorithms or longer code length may result in better result for irregular LDPC code.

IX. CONCLUSION

In this report we introduced one kind of error correcting code: LDPC code, and some of the encoding and decoding algorithms. My work is mainly about finding the optimum degree distribution code ensembles using the judgement of Density Evolution, then tried to generate the actual code and test its performance over BEC. The PEG algorithm is a well-known algorithm which can maximize the girth of the parity-check matrix, but through the simulation result, it is not the best one. First, the algorithm involves lots of searching computation to find suitable position of 1's, so it takes long time to run the program; Second, the PEG algorithm cannot guarantee the parity check matrix is full rank, so the actual code rate may be smaller than the designed code rate; PEG only average 1's in different rows so that it cannot handle other cases.
Irregular LDPC Codes

- Although regular LDPC codes perform well, they are not capable of achieving capacity.
- Properly designed irregular LDPC codes are capable of achieving capacity.
  - The degree distribution of the variable nodes is not constant.
  - The check-node degrees are often still constant (or close to it).
  - Here “designing” means picking the proper degree distribution.
Degree Distribution

- Edge-perspective degree distributions:
  - $\rho_i$ is the fraction of edges touching degree $i$ check nodes.
  - $\lambda_i$ is the fraction of edges touching degree $i$ variable nodes.

- For example, consider the Tanner graph:

  - 15 edges.
  - All are connected to degree-3 check nodes, so $\rho_3 = \frac{15}{15} = 1$.
  - Four are connected to degree-1 variable nodes, so $\lambda_1 = \frac{4}{15}$.
  - Eight are connected to degree-2 variable nodes, so $\lambda_2 = \frac{8}{15}$.
  - Three are connected to the degree-3 variable node, so $\lambda_3 = \frac{3}{15}$. 
DE for Irregular LDPC

- The degree distributions are described in polynomial form:
  - \( \rho(x) = \sum_i \rho_i x^{i-1} \) for check nodes.
  - \( \lambda(x) = \sum_i \lambda_i x^{i-1} \) for variable nodes.

- For an irregular code, the probability that a variable-node remains erased after the \( \ell^{th} \) iteration is

\[
\epsilon_\ell = \epsilon_0 \lambda (1 - \rho (1 - \epsilon_{\ell-1}))
\]

The proof follows from the Theorem on Total Probability.

- Convergence:
  - Error-free decoding requires that the erasure probability goes down from one iteration to the next.
  - Define the related function:
    \[
    f(\epsilon, x) = \epsilon \lambda (1 - \rho (1 - x))
    \]
  - Error-free decoding is possible iff \( f(\epsilon, x) \leq x \) for all \( 0 \leq x \leq \epsilon \).
Convergence

For a regular LDPC code ensemble \((d_v, d_c)\), the minimum of equation (24) can be solved by considering the derivative of \(\varepsilon(x)\). Omitting the algebra process, the threshold is:

\[
\varepsilon^* = 1 - \frac{s}{d_c - \frac{1}{d_v}} - \frac{1}{(d_v - 1)(d_c - 1) - 1}\sum_{i=0}^{\infty} \frac{1}{y_i}
\]

(26)

\(s\) is the positive real root of following equation:

\[
\left((d_v - 1)(d_c - 1) - 1\right)y^{d_c - 2} - y^{d_c - 3} - \sum_{i=0}^{\infty} y_i = 0
\]

(27)

The threshold of the \((3,6)\) regular ensemble over BEC is \(\varepsilon^* = 0.4294\), which can be approximately seen from Figure 5, 6. This code has the rate \(R = 0.5\). The capacity limit for the BEC is \(1 - r\), which is 0.5 in this case.

The general methodology of density evolution is outlined here again as follows:

Step 1: The channel is specified by a single parameter \(\sigma\), and the decoding algorithm is specified.

Step 2: Assume the corresponding code graph is cycle-free; that is, every neighborhood can be represented by a tree graph, in which case we can assume that the messages are independent.
Optimization

- The threshold is
  \[ \epsilon^* = \sup\{ \epsilon : f(\epsilon, x) < x, \forall x, 0 < x \leq \epsilon \} \]

- Solving \( f(\epsilon, x) = x \) for \( \epsilon \)

  \[
  x = f(\epsilon, x) = \epsilon \lambda (1 - \rho (1 - x))
  \]

  \[
  \epsilon = \frac{x}{\lambda (1 - \rho (1 - x))}
  \]

  Which is a function of \( x \), and henceforth expressed as \( \epsilon(x) \).

- This allows the threshold to be rewritten as:

  \[ \epsilon^* = \min\{ \epsilon(x) : \epsilon(x) \geq x \} \]
Optimization with Linear Programming

- Our goal is to find the degree distribution which yields maximum threshold

\[
\max_{\varepsilon^*} \{ \varepsilon^* = \min(\varepsilon(x) : \varepsilon(x) \geq x) \};
\]

- Several Constraints

\[
\frac{\int_0^1 \rho(x)dx}{\int_0^1 \lambda(x)dx} = 1 - R
\]

\[
\sum_{i \geq 2} \lambda_i = 1; \sum_{i \geq 2} \rho_i = 1;
\]

\[
x \in [0, 1]
\]

- Which can be modeled as an optimization problem using linear programming
  - Can use Matlab’s Optimization Toolbox.
Optimization Results ($\epsilon^* = 0.49596$)

Figure 7. Density Evolution for LDPC codes with code rate $R=0.5$

- The independence among each bits, which will achieve a closer value to the theoretically case.

Figure 9 is the decode results for length 2048 irregular LDPC code generated by PEG algorithms. The maximum column weight or the maximum variable node degree is 15, check nodes have degrees of 8 or 7. The specific degree distribution is:

$$\lambda(x) = 0.218x^4 + 0.0312x^3 + 0.0871x^2 + 0.1587x + 0.24x + 0.2648x,$$

$$\rho(x) = 0.4453x^7 + 0.5547x^6.$$

The threshold for the code ensemble is 0.47786, and the generated 2048 PEG code in 9 has the threshold near 0.45. The distance may be decrease with the increasing of code length or using other methods to generate the parity-check matrix.

The decode results of both regular and irregular LDPC code with the same length are compared in Fig. 10, the irregular code has the following degree distribution with the theoretical threshold 0.4694:

$$\lambda(x) = 0.3793x^4 + 0.0156x^6 + 0.1215x^4 + 0.0347x^3 + 0.2104x^2 + 0.2385x,$$

$$\rho(x) = x^7.$$
Encoding LDPC Codes

Encoding of LDPC codes is not necessarily straightforward.

- “Systematic-form” \( H \)
  - Using Gaussian elimination, find \( H = [P \ I] \).
  - Then \( c = uG \) where \( G = [I \ P^T] \).
  - However, \( P \) is likely to be high-density (complex encoding).

- Back-substitution.
  - If \( H \) is in an appropriate form, then \( c \) can be encoded using back substitution.
  - Example, \( cH^T = 0 \), where

\[
H = \begin{bmatrix}
1 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 1 & 1 \\
\end{bmatrix}
\]

- The LDPC code in the DVB-S2 standard allows for back substitution.
Iterative Demodulation and Decoding

- Conventional receivers first demodulation, then decode.
- Performance is improved by iterating between the demodulator and decoder.
- BICM-ID: bit-interleaved modulation with iterative decoding.
AWGN Performance of 16APSK with BICM-ID

![Graph showing BER performance vs. Es/N0 for different BICM-ID rates and LDPC codes.]

- Rate 3.2 BICM (4by5 LDPC)
- Rate 3.2 BICM-ID (4by5 LDPC)
- Rate 3 BICM-ID (3by4 LDPC)
- Rate 3 BICM-ID (3by4 LDPC)
AWGN Performance of 32APSK with BICM-ID

![Graph showing BER vs. Es/N0 for different rates and LDPC codes.](image-url)
Outline

1. Satellite Television Standards
2. DVB-S2 Modulation
3. LDPC Coding
4. Constellation Shaping
5. Conclusion
Constellation Shaping

- Capacity curves assume equiprobable signaling.
- It is possible to increase capacity by transmitting higher-energy signals less frequently than lower-energy signals.

**Figure:** Uniform 32APSK ○ vs. shaped 32APSK ○. Both constellations have the same energy.

**Figure:** The capacity of shaped 32APSK is about 0.3 dB better than uniform 32APSK.
Sub-constellations

- The 32APSK is partitioned into two equal-sized sub-constellations.
- A *shaping bit* selects the sub-constellation, while the remaining bits select a symbol from the chosen sub-constellation.
- The lower-energy sub-constellation is selected more frequently.

Figure: 32APSK w/ 2 sub-constellations

Figure: 32APSK symbol-labeling map
The shaping encoder should produce more zeros than ones.

Example: \((n_s, k_s) = (5, 3)\)

<table>
<thead>
<tr>
<th>3 input data bits</th>
<th>5 output codeword bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 0</td>
<td>0 0 0 0 0 0</td>
</tr>
<tr>
<td>0 0 1</td>
<td>0 0 0 0 1</td>
</tr>
<tr>
<td>0 1 0</td>
<td>0 0 0 1 0</td>
</tr>
<tr>
<td>0 1 1</td>
<td>0 0 1 0 0</td>
</tr>
<tr>
<td>1 0 0</td>
<td>0 1 0 0 0</td>
</tr>
<tr>
<td>1 0 1</td>
<td>1 0 0 0 0</td>
</tr>
<tr>
<td>1 1 0</td>
<td>0 0 0 1 1</td>
</tr>
<tr>
<td>1 1 1</td>
<td>1 0 1 0 0</td>
</tr>
</tbody>
</table>

- \(p_0 = 31/40\): fraction of zeros.
- \(p_1 = 9/40\): fraction of ones.
Additional complexity relative to BICM-ID due to shaping decoder.

MAP shaping decoder compares against all $2^{k_s}$ shaping codewords.
BER of Shaping in AWGN

BER of 32-APSK in AWGN at rate $R=3$ bits/symbol

![Graph showing BER of 32-APSK in AWGN at rate $R=3$ bits/symbol]
Outline

1. Satellite Television Standards
2. DVB-S2 Modulation
3. LDPC Coding
4. Constellation Shaping
5. Conclusion
DVB-S2 is a highly efficient system, thanks to
- APSK modulation.
- Tight RC-rolloff filtering.
- Capacity-approaching LDPC codes.

The performance of DVB-S2 can be improved by
- BICM-ID.
- Constellation shaping.

Future work:
- Using density evolution to optimize degree distribution of LDPC-coded APSK with shaping.
- Extension to 64APSK and beyond.
Thank You.