List Decoding: Geometrical Aspects and Performance Bounds

Maja Lončar

Department of Information Technology
Lund University, Sweden

Summer Academy: Progress in Mathematics for Communication Systems

Bremen, Germany, July 2007

Outline

- Maximum-likelihood decoding
- List decoding
- List configuration matrix
- List distance
- List error probability for a worst-case list
- New bound for the list decoding error probability
- Summary

Maximum-Likelihood Decoding

- Let $\mathcal{S} = \{\boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_{M-1}\}$ be a set of M distinct signal vectors used to communicate over the AWGN channel.

- Let $\boldsymbol{s} \in \mathcal{S}$ denote the transmit signal. Then the received signal is $\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{n}$.

- Optimum decision strategy at the receiver:

  minimize the error probability

$$\hat{\boldsymbol{s}}_{\text{MAP}} = \arg\min_{\hat{\boldsymbol{s}} \in \mathcal{S}} \left\{ \Pr(\hat{\boldsymbol{s}} \neq \boldsymbol{s}) \right\} = \arg\max_{\hat{\boldsymbol{s}} \in \mathcal{S}} \left\{ \Pr(\hat{\boldsymbol{s}} = \boldsymbol{s}) \right\} = \arg\max_{\hat{\boldsymbol{s}} \in \mathcal{S}} \left\{ \sum_{\boldsymbol{r}} \Pr(\hat{\boldsymbol{s}} = \boldsymbol{s} | \boldsymbol{r}) p(\boldsymbol{r}) \right\}$$

  that is, for every $\boldsymbol{r}$ decide in favour the most probable signal $\boldsymbol{s} \in \mathcal{S}$

$$\boxed{\hat{\boldsymbol{s}}_{\text{MAP}} = \arg\max_{\boldsymbol{s} \in \mathcal{S}} \{ p(\boldsymbol{s} | \boldsymbol{r}) \}}$$

This is the maximum *a posteriori* probability decoder.

From Bayes' formula we have

$$\hat{\boldsymbol{s}}_{\text{MAP}} = \arg\max_{\boldsymbol{s} \in \mathcal{S}} \left\{ \frac{p(\boldsymbol{r}|\boldsymbol{s}) p(\boldsymbol{s})}{p(\boldsymbol{r})} \right\} = \boxed{\arg\max_{\boldsymbol{s} \in \mathcal{S}} \left\{ p(\boldsymbol{r}|\boldsymbol{s}) p(\boldsymbol{s}) \right\}}$$

Maximum-Likelihood Decoding

- If the signals $\boldsymbol{s} \in \mathcal{S}$ are *a priori* equiprobable: $p(\boldsymbol{s}) = 1/M$.

  Then the MAP decoder reduces to $$\boxed{\hat{\boldsymbol{s}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{s}\in\mathcal{S}}\{p(\boldsymbol{r}|\boldsymbol{s})\}}$$

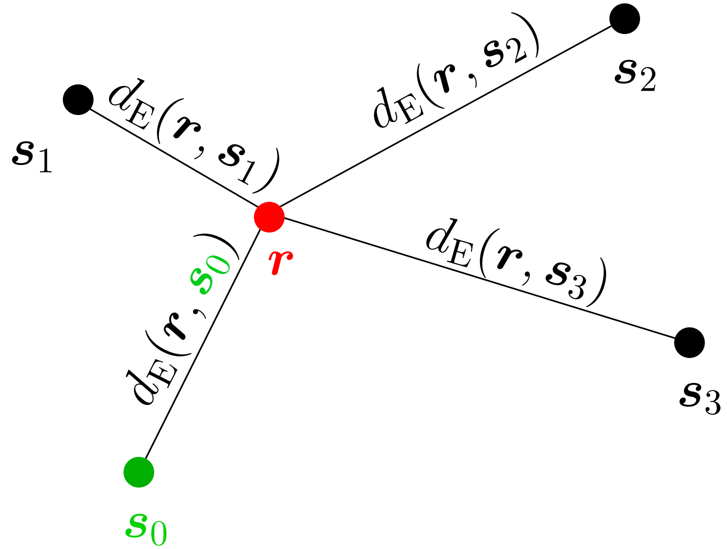
  This is the maximum likelihood decoder.

- For the AWGN channel

$$p(\boldsymbol{r}|\boldsymbol{s}) \propto e^{-\frac{1}{N_0}||\boldsymbol{r}-\boldsymbol{s}||^2}$$

$$\boxed{\hat{\boldsymbol{s}}_{\mathrm{ML}} = \arg\min_{\boldsymbol{s}\in\mathcal{S}}\{||\boldsymbol{r}-\boldsymbol{s}||^2\}}$$

  ML decoder is the minimum Euclidean distance decoder. It chooses the point with the smallest $d_{\mathrm{E}}(\boldsymbol{r}, \boldsymbol{s}) = ||\boldsymbol{r}-\boldsymbol{s}||$.



- Decoding error occurs if $d_{\mathrm{E}}(\boldsymbol{r}, \boldsymbol{s}_i) \leq d_{\mathrm{E}}(\boldsymbol{r}, \boldsymbol{s}_0)$, for some i. Let $\varepsilon_i|\boldsymbol{s}_0$ denote this event.

- Union bound on the ML error probability:

$$P_{\mathrm{e}|\boldsymbol{s}_0} \leq \sum_{i=1}^{M-1} \Pr(\varepsilon_i|\boldsymbol{s}_0)$$

List Decoding

- Generalization of ML decoding for $L \geq 1$ most likely codewords.

- List decoder finds a list of the L best codewords. For Gaussian channel, these are L codewords $\boldsymbol{s}_i$ closest to the received vector $\boldsymbol{r}$.

- List decoding error occurs if transmitted signal $\boldsymbol{s}_0$ is not on the list, that is, if,

$$d_{\mathrm{E}}(\boldsymbol{r}, \boldsymbol{s}_i) \leq d_{\mathrm{E}}(\boldsymbol{r}, \boldsymbol{s}_0), \quad i = 1, 2, \ldots, L$$

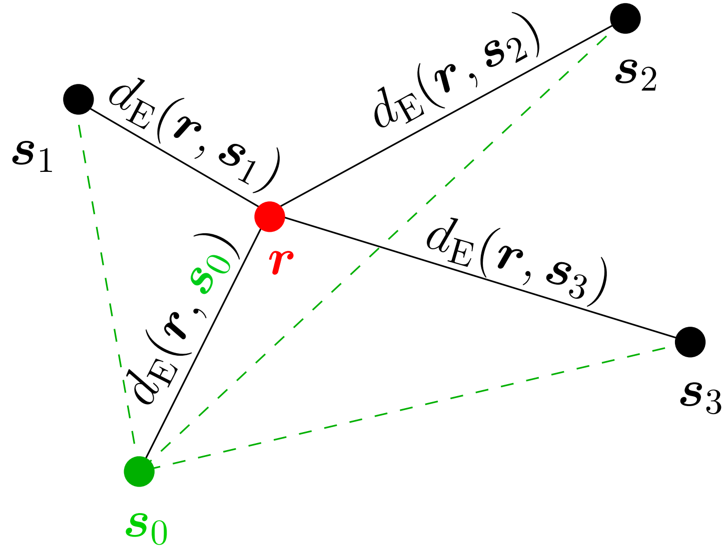
or, equivalently, if projections of the noise $\boldsymbol{n}$ along $(\boldsymbol{s}_i - \boldsymbol{s}_0)$ are larger than $d_{\mathrm{E}}(\boldsymbol{s}_i, \boldsymbol{s}_0)/2$

$$\underbrace{\langle \boldsymbol{n}, \boldsymbol{s}_i - \boldsymbol{s}_0 \rangle}_{t_i} \geq \underbrace{d^2_{\mathrm{E}}(\boldsymbol{s}_i, \boldsymbol{s}_0)/2}_{d^2_{\mathrm{E}0i}/2}, \quad i = 1, 2, \ldots, L$$

that is,

$$(t_1 \; t_2 \; \ldots \; t_L) \geq (d^2_{\mathrm{E}01} \; d^2_{\mathrm{E}02} \; \ldots \; d^2_{\mathrm{E}0L})/2$$

$$\boxed{\boldsymbol{t} \geq \boldsymbol{w}/2}$$



List Configuration Matrix

- Components of $\boldsymbol{t}$, $\langle \boldsymbol{n}, \boldsymbol{s}_i - \boldsymbol{s}_0 \rangle$, are Gaussian distributed with covariance matrix

$$\mathrm{E}[\boldsymbol{t}^{\mathrm{T}}\boldsymbol{t}] = \sigma^2 \boldsymbol{W}$$

where $\boldsymbol{W}$ is the Gram matrix of the vectors $(\boldsymbol{s}_i - \boldsymbol{s}_0)$, $i = 1, 2, \ldots, L$ with elements

$$w_{ij} = \langle \boldsymbol{s}_i - \boldsymbol{s}_0, \boldsymbol{s}_j - \boldsymbol{s}_0 \rangle = (d^2_{\mathrm{E}0i} + d^2_{\mathrm{E}0j} - d^2_{\mathrm{E}ij})/2$$

$$\boldsymbol{W} = \begin{pmatrix} d^2_{\mathrm{E}01} & (d^2_{\mathrm{E}01} + d^2_{\mathrm{E}02} - d^2_{\mathrm{E}12})/2 & \ldots & (d^2_{\mathrm{E}01} + d^2_{\mathrm{E}0L} - d^2_{\mathrm{E}1L})/2 \\ (d^2_{\mathrm{E}01} + d^2_{\mathrm{E}02} - d^2_{\mathrm{E}12})/2 & d^2_{\mathrm{E}02} & \ldots & (d^2_{\mathrm{E}02} + d^2_{\mathrm{E}0L} - d^2_{\mathrm{E}2L})/2 \\ \vdots & \vdots & \ddots & \vdots \\ (d^2_{\mathrm{E}01} + d^2_{\mathrm{E}0L} - d^2_{\mathrm{E}1L})/2 & (d^2_{\mathrm{E}02} + d^2_{\mathrm{E}0L} - d^2_{\mathrm{E}2L})/2 & \ldots & d^2_{\mathrm{E}0L} \end{pmatrix}$$

- $\boldsymbol{W}$ is a list configuration matrix. It determines the list error probability for a given list:

$$P_{\mathrm{e}L}(\boldsymbol{W}) = \Pr(\boldsymbol{t} \geq \boldsymbol{w}/2),$$

- Union-type bound for list decoding error probability:

$$P_{\mathrm{e}L} \leq \sum_{\boldsymbol{W}} N(\boldsymbol{W}) P_{\mathrm{e}L}(\boldsymbol{W})$$

List Distance

- List radius R_L for a given signal set $\mathcal{S}_L = \{\boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\}$ and a given reference (transmit) signal $\boldsymbol{s}_0$, is the point of the list-error region that is closest to $\boldsymbol{s}_0$.

  This is a radius of the smallest sphere $\mathbb{S}$ that contains (encompasses) all the signal points from $\mathcal{S}_L$ (points on or inside the sphere).

- Euclidean list distance for a given signal set $\mathcal{S}_L$ is $\boxed{d_{\mathrm{E}L} = 2R_L}$.

- *Theorem 1:* The radius $\widetilde{R}_L$ of the sphere $\widetilde{\mathbb{S}}$ that passes through all the points of $\mathcal{S}_L$ (all the points on the sphere) is given by

$$\boxed{\widetilde{R}_L(\boldsymbol{W}) = \frac{1}{2}\sqrt{\boldsymbol{w}\boldsymbol{W}^{-1}\boldsymbol{w}^{\mathrm{T}}}}$$

- *Theorem 2:* The list radius R_L of the smallest sphere $\mathbb{S}$ that passes through $\boldsymbol{s}_0$ and encompasses the points $\boldsymbol{s}_i$, $i = 1, 2, \ldots, L$, is given by

$$\boxed{R_L(\boldsymbol{W}) = \max_{\mathcal{I}\,:\,\boldsymbol{w}_{\mathcal{I}}\,\mathrm{adj}(\boldsymbol{W}_{\mathcal{I}}) \geq \boldsymbol{0}} \left\{ \frac{1}{2}\sqrt{\boldsymbol{w}_{\mathcal{I}}\boldsymbol{W}_{\mathcal{I}}^{-1}\boldsymbol{w}_{\mathcal{I}}^{\mathrm{T}}} \right\}}$$

List Distance

- Example:

$L = 2$

$\boldsymbol{s}_0 = (0\ 0)$

$\boldsymbol{s}_1 = (0\ 2)$

$\boldsymbol{s}_2 = (1\ 3)$

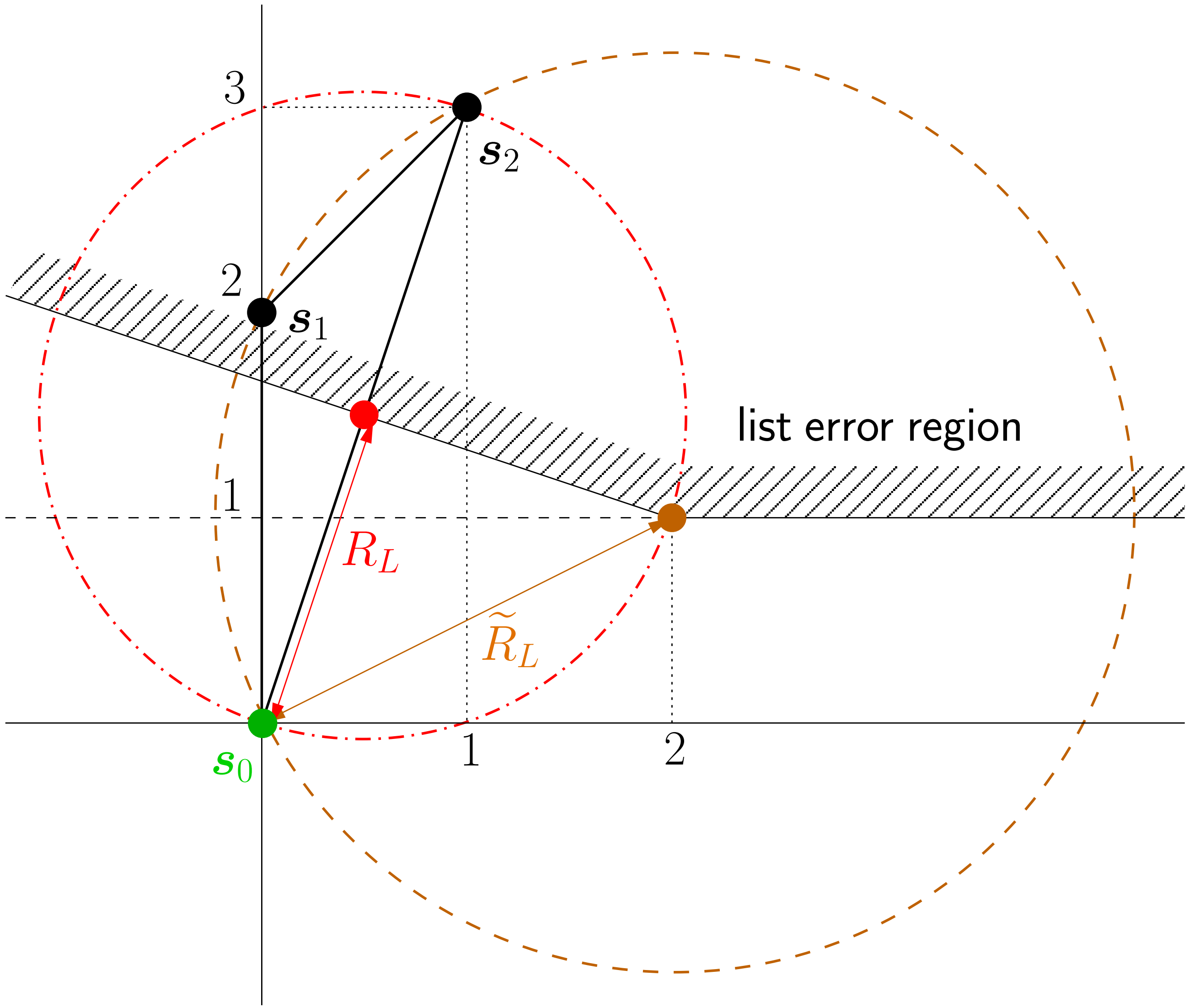
$$\boldsymbol{W} = \begin{pmatrix} 4 & 6 \\ 6 & 10 \end{pmatrix}$$

$\widetilde{R}_L = \frac{1}{2}\sqrt{\boldsymbol{w}\boldsymbol{W}^{-1}\boldsymbol{w}^{\mathrm{T}}} = \sqrt{5}$

$\boldsymbol{w}\,\mathrm{adj}(\boldsymbol{W}) = (-20\ 16)$

$R_L = \sqrt{10}/2$

$d_{\mathrm{E}L} = 2R_L = \sqrt{10}$



List Distance

- Example:

  $L = 2$

  $\boldsymbol{s}_1 = (0\ 0)$
  $\boldsymbol{s}_0 = (0\ 2)$
  $\boldsymbol{s}_2 = (1\ 3)$

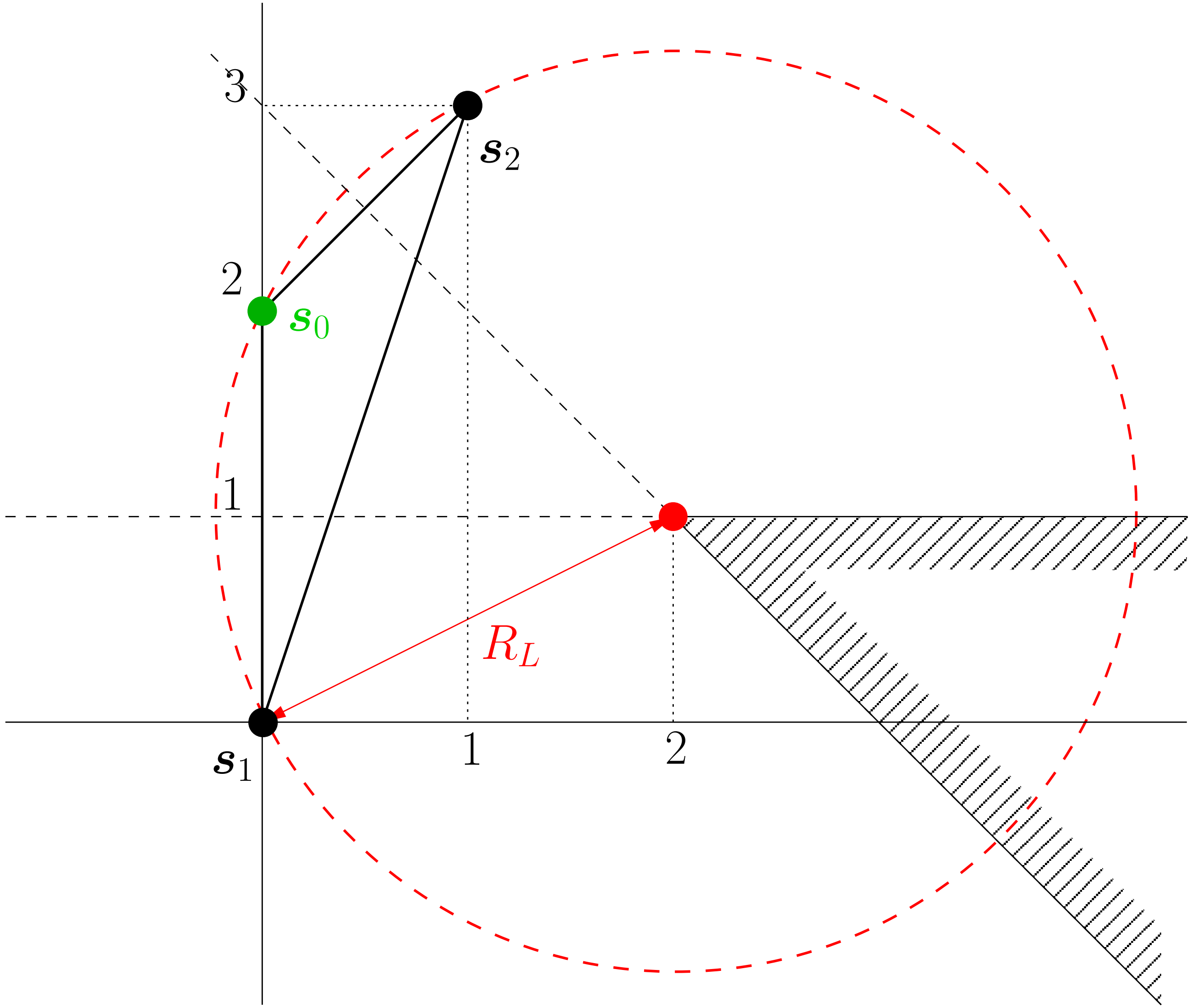
  $$\boldsymbol{W} = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix}$$

  $\widetilde{R}_L = \frac{1}{2}\sqrt{\boldsymbol{w}\boldsymbol{W}^{-1}\boldsymbol{w}^{\mathrm{T}}} = \sqrt{5}$

  $\boldsymbol{w}\,\mathrm{adj}(\boldsymbol{W}) = (12\ 16)$

  $R_L = \widetilde{R}_L = \sqrt{5}$

  $d_{\mathrm{E}L} = 2R_L = \sqrt{20}$



Minimum List Distance

- Minimum list radius R_L for a list size L is

$$R_{L\,\min} = \min_{\boldsymbol{W}}\{R_L(\boldsymbol{W})\} = \min_{\boldsymbol{W}} \max_{\mathcal{I}\,:\,\boldsymbol{w}_{\mathcal{I}}\,\mathrm{adj}(\boldsymbol{W}_{\mathcal{I}})\geq \boldsymbol{0}} \left\{ \frac{1}{2}\sqrt{\boldsymbol{w}_{\mathcal{I}}\boldsymbol{W}_{\mathcal{I}}^{-1}\boldsymbol{w}_{\mathcal{I}}^{\mathrm{T}}} \right\}$$

- Minimum Euclidean list distance for a list size L is $\boxed{d_{\mathrm{E}L\mathrm{min}} = 2R_{L\,\min}}$.

- For binary linear codes and BPSK signaling, $d_{\mathrm{E}ij}^2 = 4E_s d_{\mathrm{H}ij}$.

  Minimum Hamming list distance of a code, for list size L is $\boxed{d_{\mathrm{H}L\mathrm{min}} = d_{\mathrm{E}L\mathrm{min}}^2/(4E_s)}$.

- Minimum list distance determines the performance of the list decoder at higher SNR, in the same way as the minimum distance determines the performance of the ML decoder.

- *Theorem 3:* For any binary linear code the minimum Hamming list distance is

$$\boxed{d_{\mathrm{H}L\mathrm{min}} \geq \frac{2L}{L+1} d_{\mathrm{Hmin}}}$$

  For any binary linear code with odd d_{Hmin}, we have

$$\boxed{d_{\mathrm{H}L\mathrm{min}} \geq \frac{2L}{L+1} d_{\mathrm{Hmin}} + \frac{L-1}{L+1}}$$

Minimum List Distance

- For **even** d_{Hmin}

Worst-case list configuration yielding minimum list distance

$$d_{\text{H}L\text{min}} = \frac{2L}{L+1} d_{\text{Hmin}}$$

consists of L codewords of weight d_{Hmin} at pairwise distances d_{Hmin}:

$$\boldsymbol{W}_{\text{H}} = \begin{pmatrix} d_{\text{Hmin}} & \frac{d_{\text{Hmin}}}{2} & \frac{d_{\text{Hmin}}}{2} & \dots & \frac{d_{\text{Hmin}}}{2} \\ \frac{d_{\text{Hmin}}}{2} & d_{\text{Hmin}} & \frac{d_{\text{Hmin}}}{2} & \dots & \frac{d_{\text{Hmin}}}{2} \\ \frac{d_{\text{Hmin}}}{2} & \frac{d_{\text{Hmin}}}{2} & d_{\text{Hmin}} & \dots & \frac{d_{\text{Hmin}}}{2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{d_{\text{Hmin}}}{2} & \frac{d_{\text{Hmin}}}{2} & \frac{d_{\text{Hmin}}}{2} & \dots & d_{\text{Hmin}} \end{pmatrix}$$

The codewords form an L-dimensional simplex.

Minimum List Distance

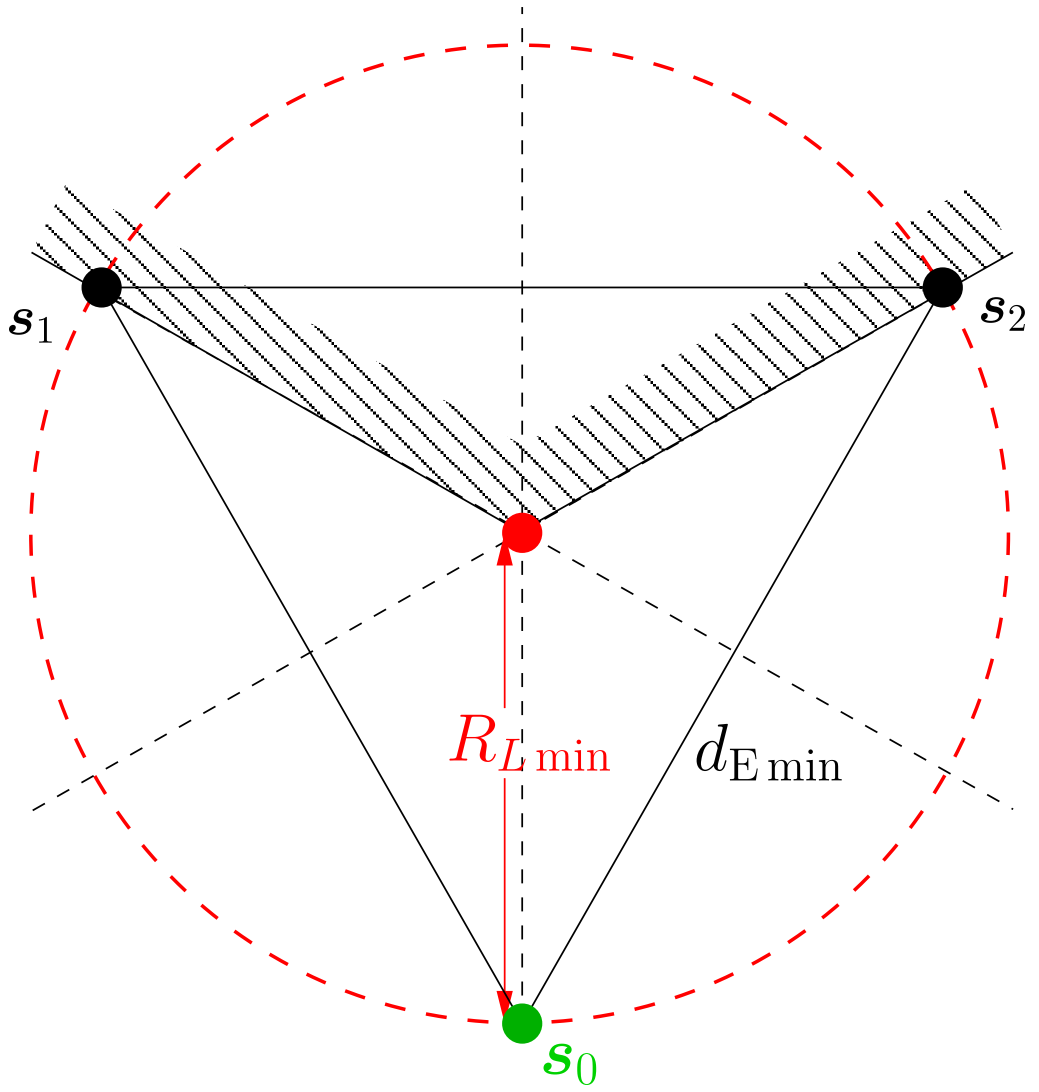
- Example:

  even d_{Hmin}

  $L = 2$

  $R_{L\,\mathrm{min}} = \frac{1}{\sqrt{3}} d_{\mathrm{Emin}}$

  $d_{\mathrm{E}L\mathrm{min}} = \frac{2}{\sqrt{3}} d_{\mathrm{Emin}}$



Minimum List Distance

- For **odd** d_{Hmin}

  Worst-case list configuration yielding minimum list distance

$$d_{\mathrm{H}L\mathrm{min}} = \frac{2L}{L+1} d_{\mathrm{Hmin}} + \frac{L-1}{L+1}$$

consists of

$(L+1)/2$ codewords of weight d_{Hmin}

$(L-1)/2$ codewords of weight $d_{\mathrm{Hmin}}+1$

at pairwise distances d_{Hmin} and $d_{\mathrm{Hmin}}+1$. For example, for $L=5$:

$$\boldsymbol{W}_{\mathrm{H}} = \left(\begin{array}{ccc|cc} d_{\mathrm{Hmin}} & \frac{d_{\mathrm{Hmin}}-1}{2} & \frac{d_{\mathrm{Hmin}}-1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} \\ \frac{d_{\mathrm{Hmin}}-1}{2} & d_{\mathrm{Hmin}} & \frac{d_{\mathrm{Hmin}}-1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} \\ \frac{d_{\mathrm{Hmin}}-1}{2} & \frac{d_{\mathrm{Hmin}}-1}{2} & d_{\mathrm{Hmin}} & \frac{d_{\mathrm{Hmin}}+1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} \\ \hline \frac{d_{\mathrm{Hmin}}+1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} & d_{\mathrm{Hmin}}+1 & \frac{d_{\mathrm{Hmin}}+1}{2} \\ \frac{d_{\mathrm{Hmin}}+1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} & \frac{d_{\mathrm{Hmin}}+1}{2} & d_{\mathrm{Hmin}}+1 \end{array} \right)$$

List Error Probability for Worst-Case List

- Worst-case list configuration determines the performance of list decoder at high SNR

  Therefore, we want to estimate the list error probability $\Pr(\boldsymbol{t} \geq \boldsymbol{w}/2)$ for the worst-case list configuration

- Problem: Finding $\Pr(\boldsymbol{t} \geq \boldsymbol{w}/2)$ involves L-dimensional integration over the PDF of $\boldsymbol{t}$

- Solution: Consider instead the orthogonalized vector

$$\boldsymbol{q} = \boldsymbol{t}\boldsymbol{V}$$

  The components of $\boldsymbol{q}$ are uncorrelated and its PDF breaks up into a product of L 1-dimensional PDFs

  By estimating the integration limits for $\boldsymbol{q}$ we obtain an upper bound on $\Pr(\boldsymbol{t} \geq \boldsymbol{w}/2)$.

List Error Probability for Worst-Case List

- For **even** d_{Hmin}

  Lemma 1: The list-error probability $\Pr(\boldsymbol{t} \geq \boldsymbol{w}/2)$, for a worst-case list configuration, for a code with even d_{Hmin} is upper-bounded by

$$\Pr(\boldsymbol{t} \geq \boldsymbol{w}/2) \;\leq\; \int_{\frac{\alpha L}{\sqrt{\sigma_1}}}^{\infty} f(y) \prod_{l=2}^{L} \left(\int_{v_l(y)}^{u_l(y)} f(x)\mathrm{d}x \right) \mathrm{d}y$$

  with equality for $L \leq 2$. $f(x)$ and $f(y)$ denote Gaussian $\mathcal{N}(0,1)$ PDF.

  The integration intervals are determined by the eigenvalues of $\boldsymbol{W}$.

- For **odd** d_{Hmin}

  Lemma 2: The list-error probability $\Pr(\boldsymbol{t} \geq \boldsymbol{w}/2)$, for a worst-case list configuration, for a code with odd d_{Hmin} is upper-bounded by

$$\Pr(\boldsymbol{t} \geq \boldsymbol{w}/2) \;\leq\; \int_{\frac{\phi(\alpha,\eta)}{\sqrt{\sigma_L}}}^{\infty} f(y_1) \int_{g(y_1)}^{h(y_1)} f(y_2)\mathrm{d}y_2 \prod_{l=1}^{\frac{L-1}{2}} \left(\int_{v_l(y_1)}^{u_l(y_1)} f(x)\mathrm{d}x \right) \prod_{l=\frac{L+1}{2}}^{L-2} \left(\int_{w_l(y_1)}^{z_l(y_1)} f(x)\mathrm{d}x \right) \mathrm{d}y_1$$

Generalized Tangential Bound on List Decoding Error Probability

- We want to improve the union bound on the list-error probability P_{eL} for binary codes
- Tangential-bound approach:

Decompose noise vector $\boldsymbol{n}$ into

- one radial component x, along transmitted signal $\boldsymbol{s}_0$:

$$x = \langle \boldsymbol{n}, \boldsymbol{s}_0 \rangle$$

- L components y_l orthogonal to the radial component:

$$y_l = \langle \boldsymbol{n}, \boldsymbol{s}_l \rangle - (1 - d_{\mathrm{H}0l}/N)x, \quad l = 1, 2, \ldots, L$$

- List decoding error probability $P_{eL} \equiv \Pr(\varepsilon)$ is upper-bounded by (*few*–*many errors*)

$$\boxed{\begin{aligned} P_{eL} &= \Pr(\varepsilon, x \le T) + \Pr(\varepsilon, x > T) \\ &\le \int_{-\infty}^{T} \Pr(\varepsilon|x) f(x)\mathrm{d}x + \Pr(x > T) \end{aligned}}$$

This yields

$$P_{eL} \le \min_T \left(\int_{-\infty}^{T} \min\left\{1, \sum_{\boldsymbol{K}_y} N(\boldsymbol{K}_y) P_{eL}(\boldsymbol{K}_y, x)\right\} f(x)\mathrm{d}x + Q(T) \right)$$

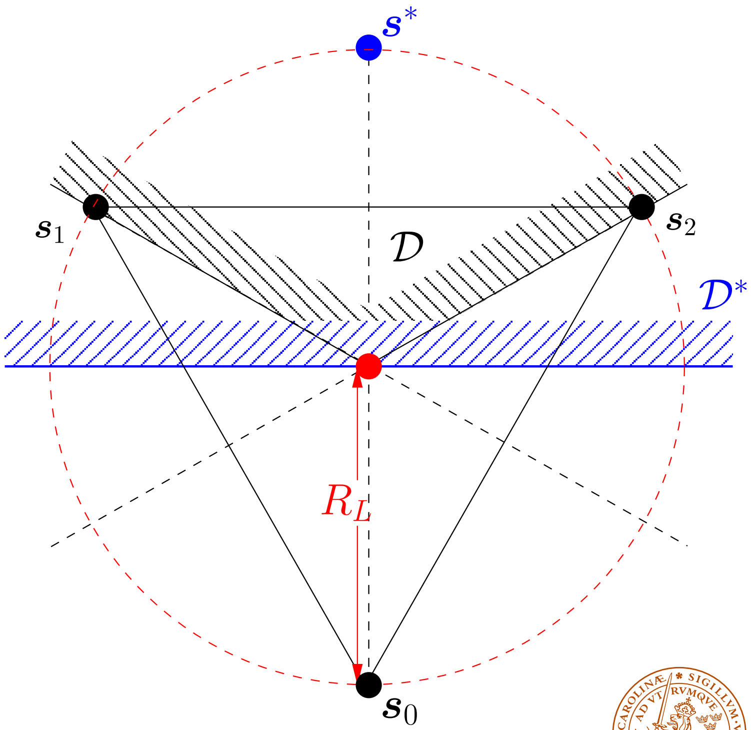
Generalized Tangential Bound on List Decoding Error Probability

- New generalized tangential union bound for list decoding error probability

$$P_{eL} \le \min_{T} \left(\int_{-\infty}^{T} \min\left\{ 1, N(\boldsymbol{K}_y) P_{eL}(\boldsymbol{K}_y, x) + \sum_{\boldsymbol{W}} N(\boldsymbol{W}) P_{eL}(d_{HL}(\boldsymbol{W}), x) \right\} f(x) \mathrm{d}x + Q(T) \right)$$

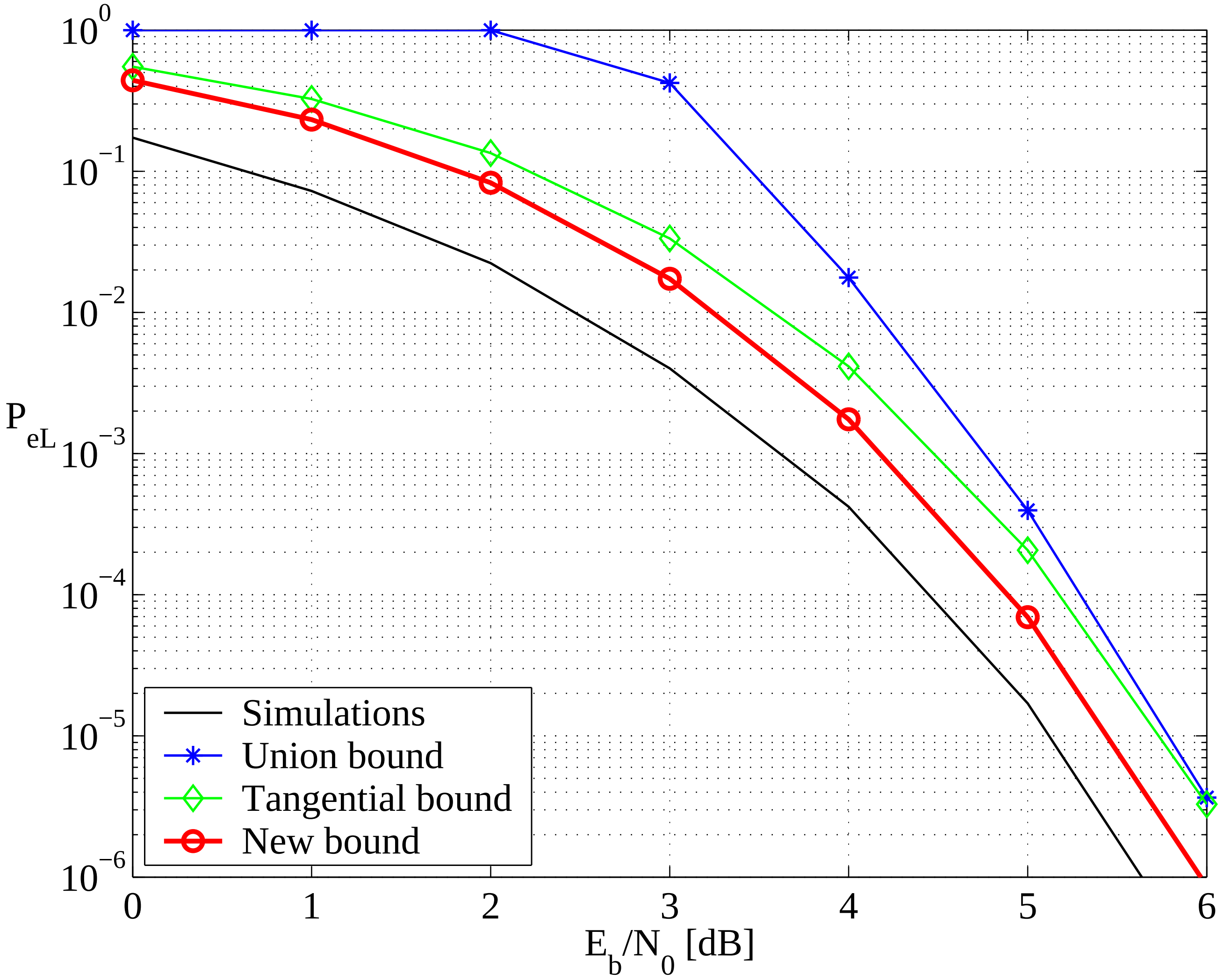
• The dominant term is estimated using the upper bound on worst-case-list error probability

• The remaining terms are upper-bounded only by using the list distance $d_{HL}(\boldsymbol{W})$, which is equivalent to *replacing the codeword sets with list configuration matrix* $\boldsymbol{W}$ *by an "average codeword"* $\boldsymbol{s}^*$ *at distance* $d_{HL}(\boldsymbol{W})$ *from the transmitted codeword.*



Generalized Tangential Bound on List Decoding Error Probability

- Comparison of bounds for $(24, 12, 8)$ Golay code with list size $L = 2$



Summary

- List configuration matrix describes the list geometry and list error probability
- List radius and list distance are defined via the list configuration matrix
- Minimum list distance is determined by the worst-case list configuration
- Upper bound on the list error probability for the worst-case lists is derived
- New generalized tangential union bound for the list decoding error probability is derived