Improved bounds on the word error probability of RA(2) codes with linear programming based decoding

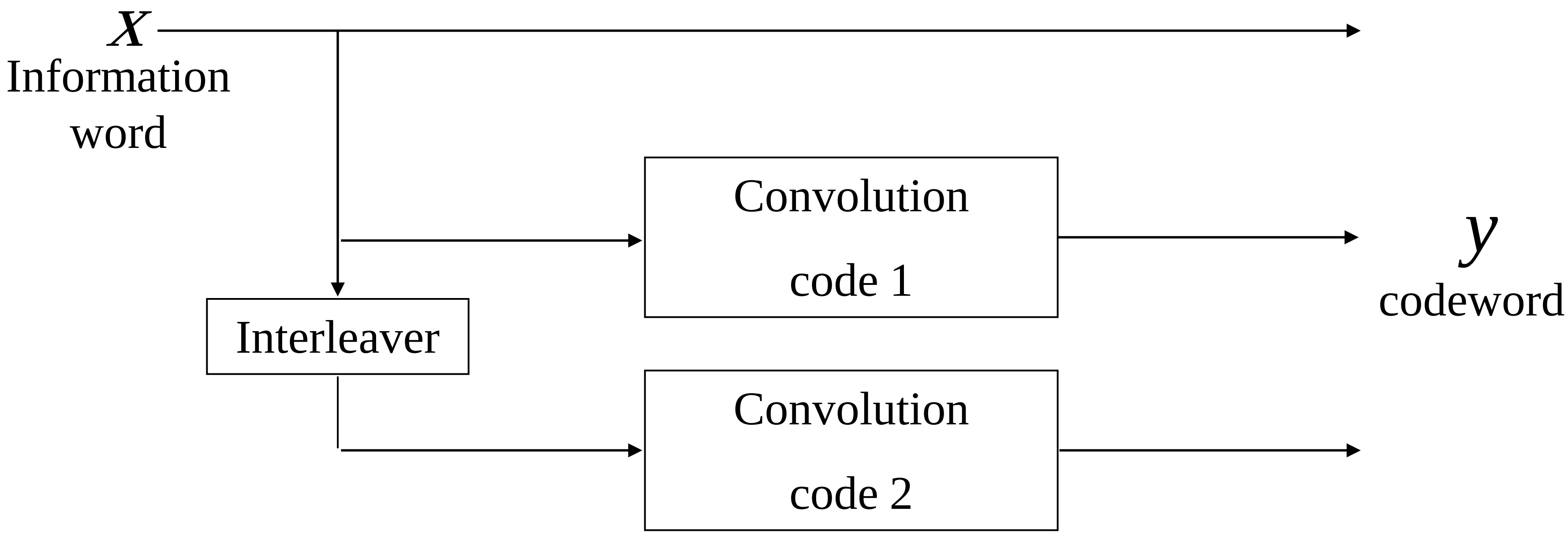
Nissim Halabi
Tel-Aviv University
Joint work with Guy Even

Outline

- “Turbo-like” codes.
  - classic Turbo codes & “turbo-like” codes.
  - Repeat Accumulate codes.
- Auxiliary graphs and promenades.
- RALP decoding.
- Characterization of RALP failure.
  - Non-positive cost minimal promenades.
  - Skeleton graphs and skeleton promenades.
- Algorithms for error bounds.
- Experimental results.

Turbo Codes

[Berrou, Glavieux, Thitimajashima, 1993]:

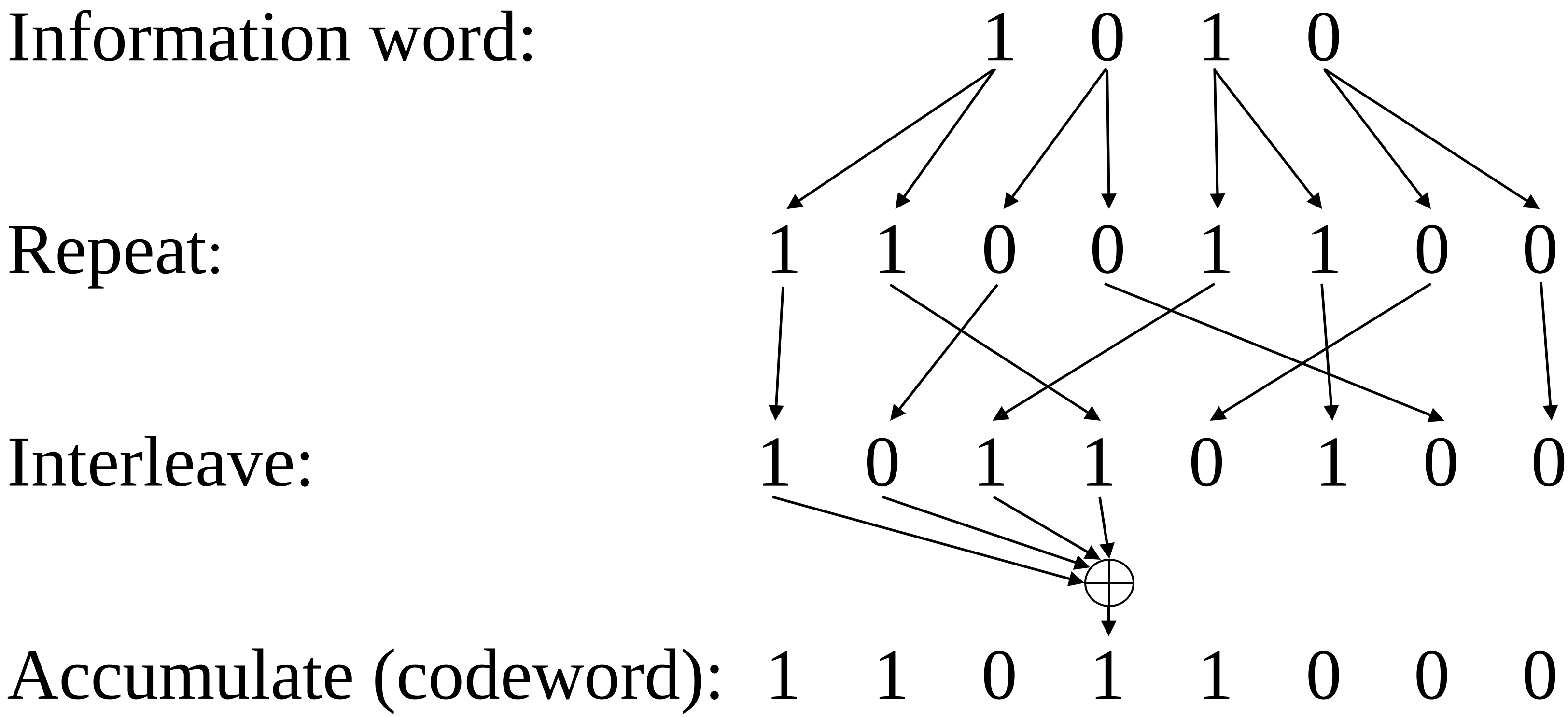


- "Turbo-like" codes [Divsalar, Jin, McEliece, 1998]: parallel, serial concatenated convolutional codes with interleavers.

Repeat Accumulate Codes *RA(q)*

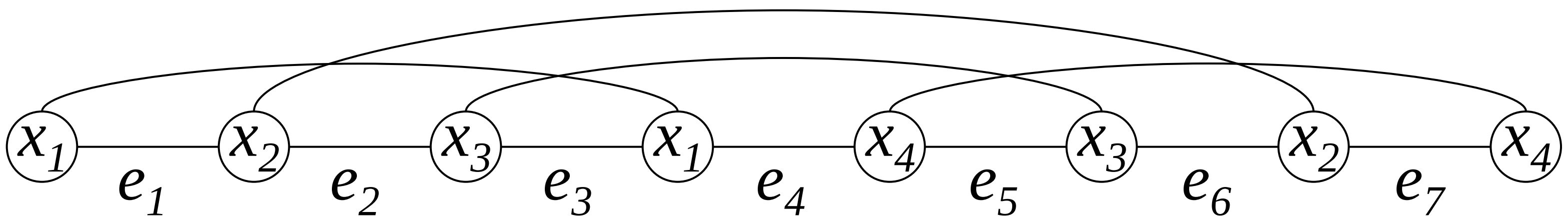
[Divsalar, Jin, McEliece, 1998]

RA(2) code:



Auxiliary Graphs of RA(2) codes

- Model for RA(2) codes [Bazzi *et al*. 2001].
- Undirected graph: path + matching.
  - Vertices: codeword bits
  - Matching edges: interleaver;
  - Path edges: also called *Hamiltonian* edges.



- Theorem [BMMS01]: code distance = graph's girth (shortest cycle).
- Construct maximal distance RA(2) codes: cubic graphs with girth $\Theta(\log n)$ [Erdös & Sachs, 1963][BMMS01].

Auxiliary Graphs of RA(2) codes (cont.)

- Error word $\mapsto$ costs to Hamiltonian edges [Feldman & Karger, 2002].
  - BSC:

$$c[e_i] = \begin{cases} -1 & \text{if bit } i \text{ is flipped by the channel} \\ +1 & \text{otherwise} \end{cases}$$

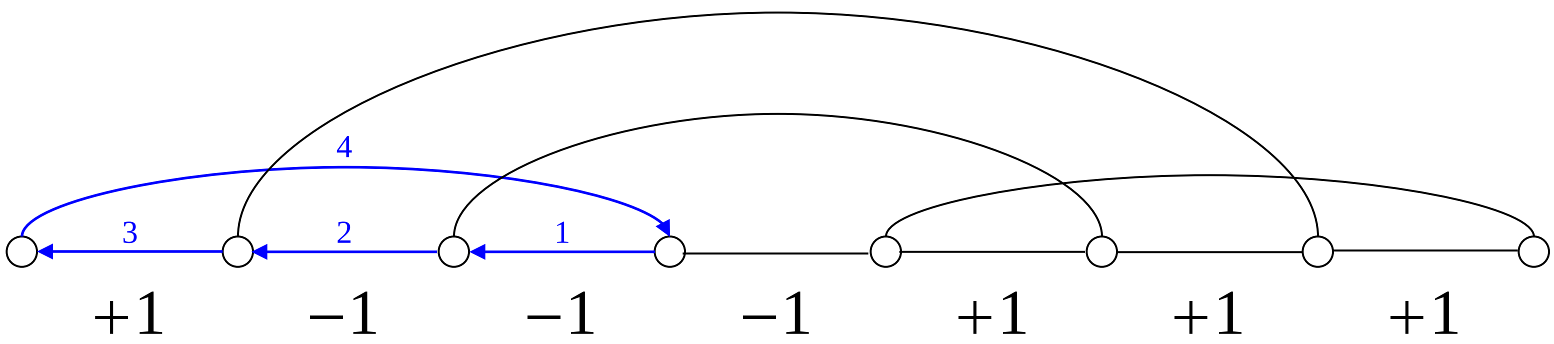
  - AWGN channel:

$$c[e_i] = 1 + \varphi_i \quad \text{where} \quad \varphi_i \sim \mathcal{N}\left(0, \tfrac{N_0}{2}\right)$$

- Matching edge: cost $\equiv 0$.

Promenade [FK]

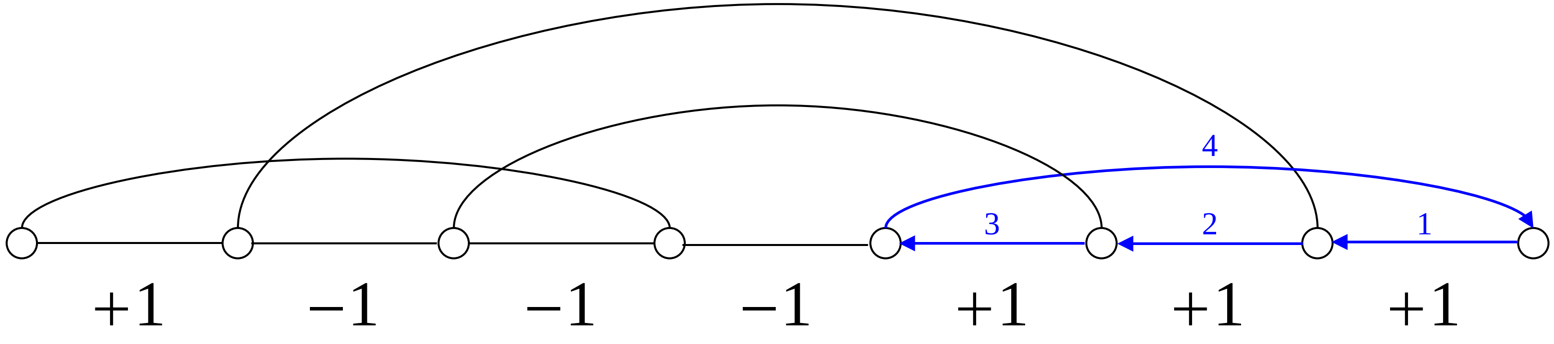
- A *Promenade* is a closed walk that does not traverse an edge twice in a row.
- The *cost* of a promenade is the sum of the costs of the edges traversed by the promenade.
- Infinitely many promenades.
- At least every second edge is Hamiltonian.



$$c[M] = -1$$

Promenade [FK]

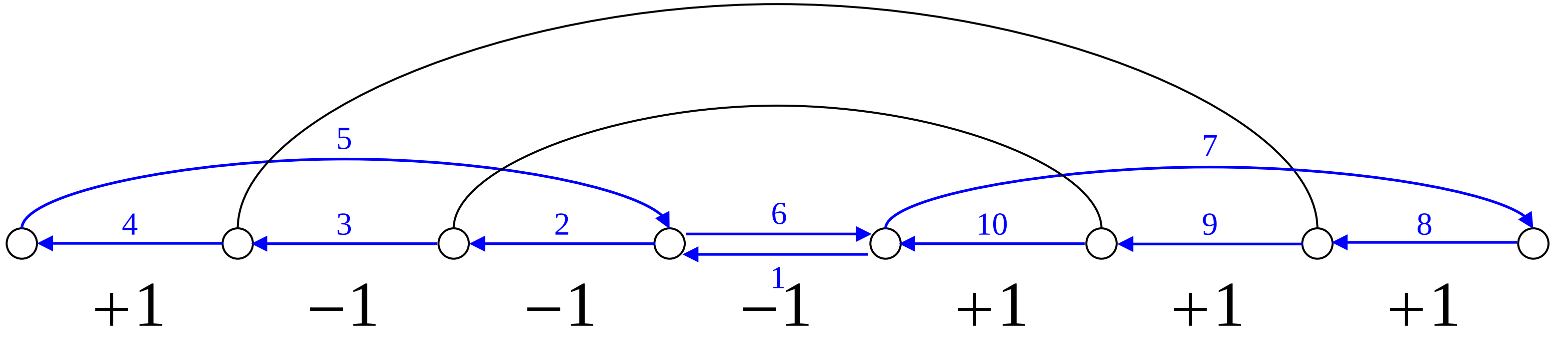
- A *Promenade* is a closed walk that does not traverse an edge twice in a row.
- The *cost* of a promenade is the sum of the costs of the edges traversed by the promenade.
- Infinitely many promenades.
- At least every second edge is Hamiltonian.



$$c[M] = +3$$

Promenade [FK]

- A *Promenade* is a closed walk that does not traverse an edge twice in a row.
- The *cost* of a promenade is the sum of the costs of the edges traversed by the promenade.
- Infinitely many promenades.
- At least every second edge is Hamiltonian.



$$c[M] = 0$$

RALP decoding [FK]

- A provably polynomial time algorithm.
- Agrees with ML decoding; or outputs “error”.
- Theorem [FK02]: The RALP decoder succeeds if all promenades have positive cost. The RALP decoder fails if there is a promenade with negative cost.
  - *Success*: output the original information word.
- $\Pr\{fail\} \leq \Pr\{\exists \, \text{promenade } M : c[M] \leq 0\}$
- Theorem [FK02]: $$\Pr(fail) \leq \frac{1}{poly(n)}$$
  - Specific, deterministically constructible codes.
  - Every code length.

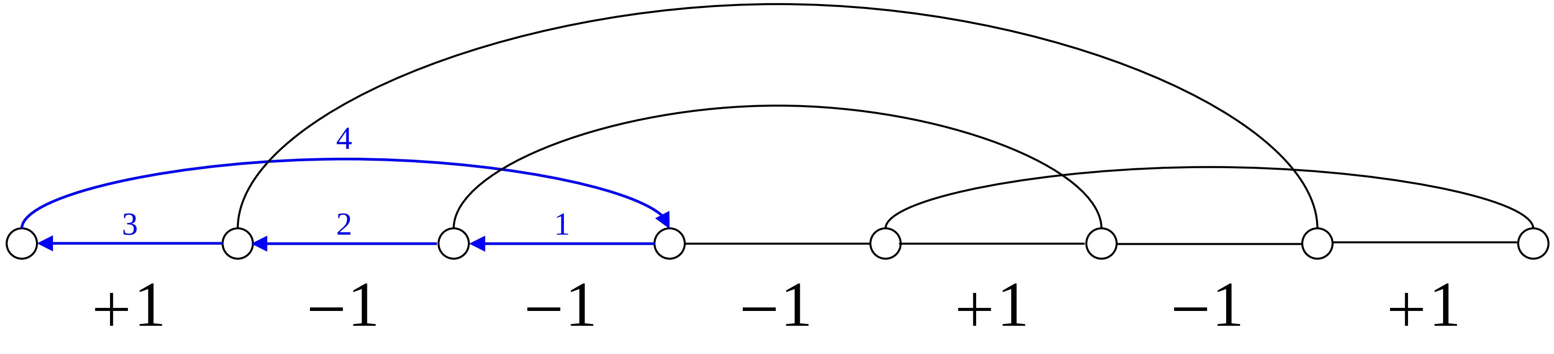
Our Results

- New structural theorem that characterizes the event that RALP fails.
- Present polynomial time algorithms that, given an RA(2) code, compute upper and lower bounds on P_w.
- Experiments demonstrate an improvement for bounds on P_w.

NPCM-Promenades

- Non-Positive Cost Minimal Promenade:

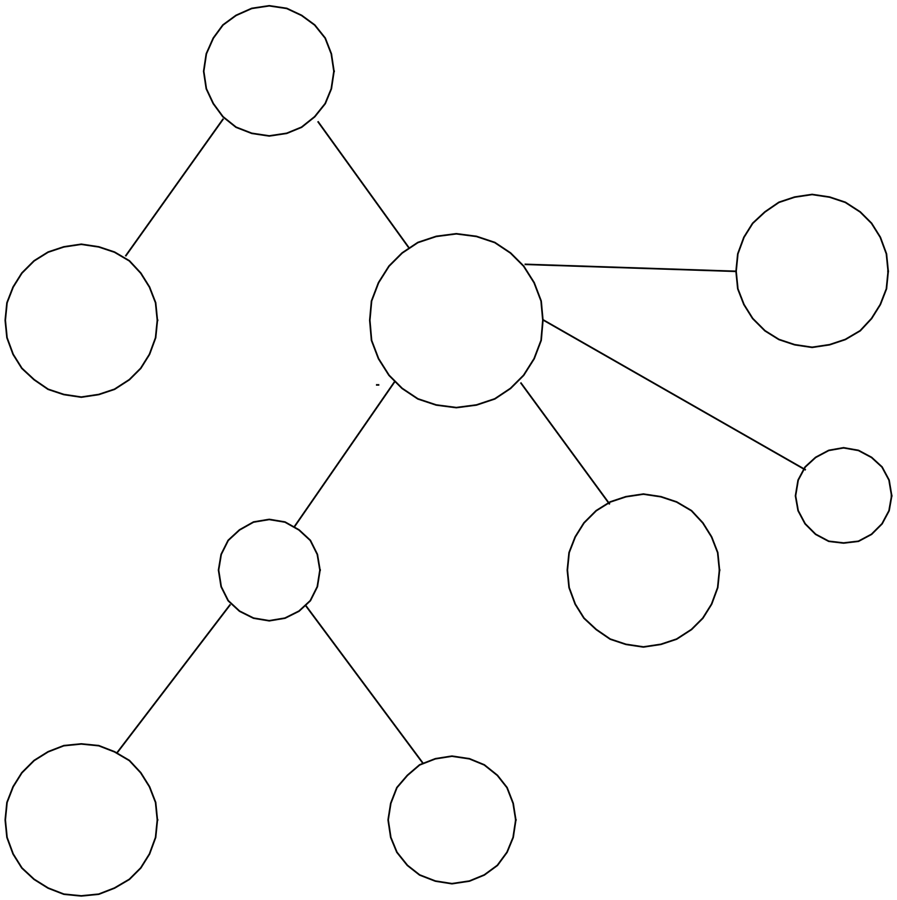
  A promenade with:
  - Non-positive cost
  - Minimal with respect to inclusion.
- Observation: ∃ NPCM-promenade ⇔ ∃ non-positive cost promenade.
- The number of NPCM-promenades is finite.



$$c[M] = -1$$

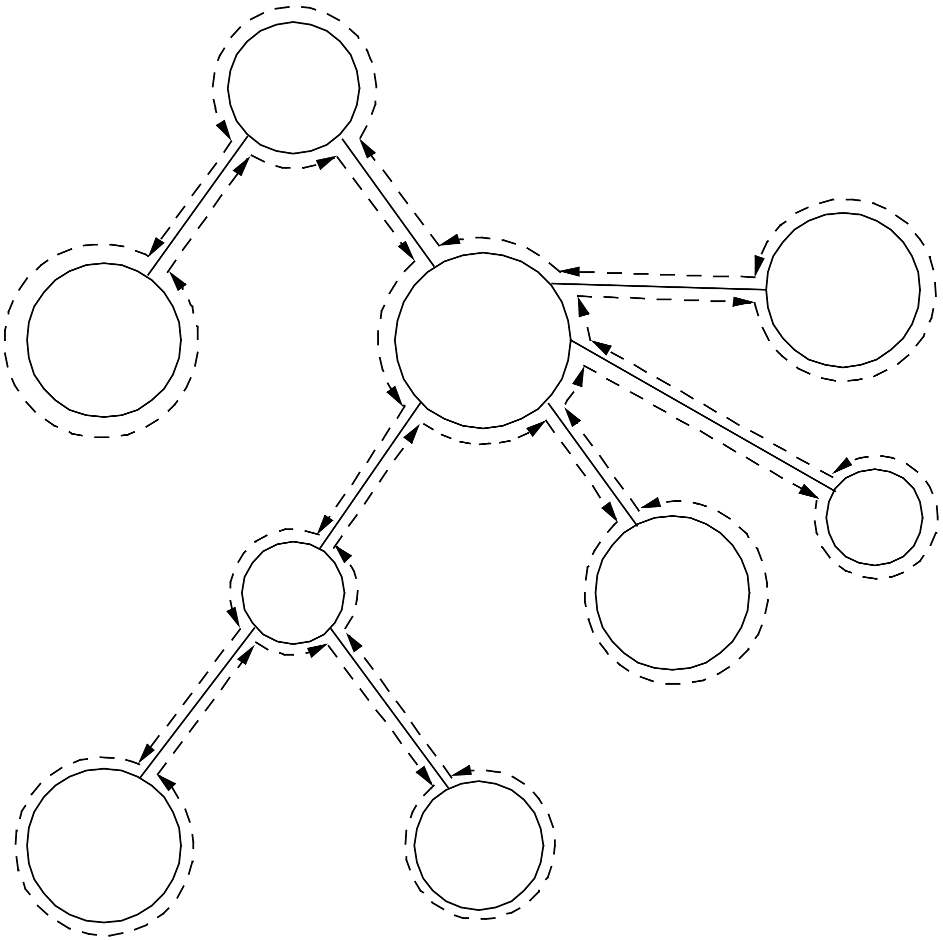
Skeleton Graphs and Promenades

- A *skeleton graph* has the structure of a “tree of cycles”.



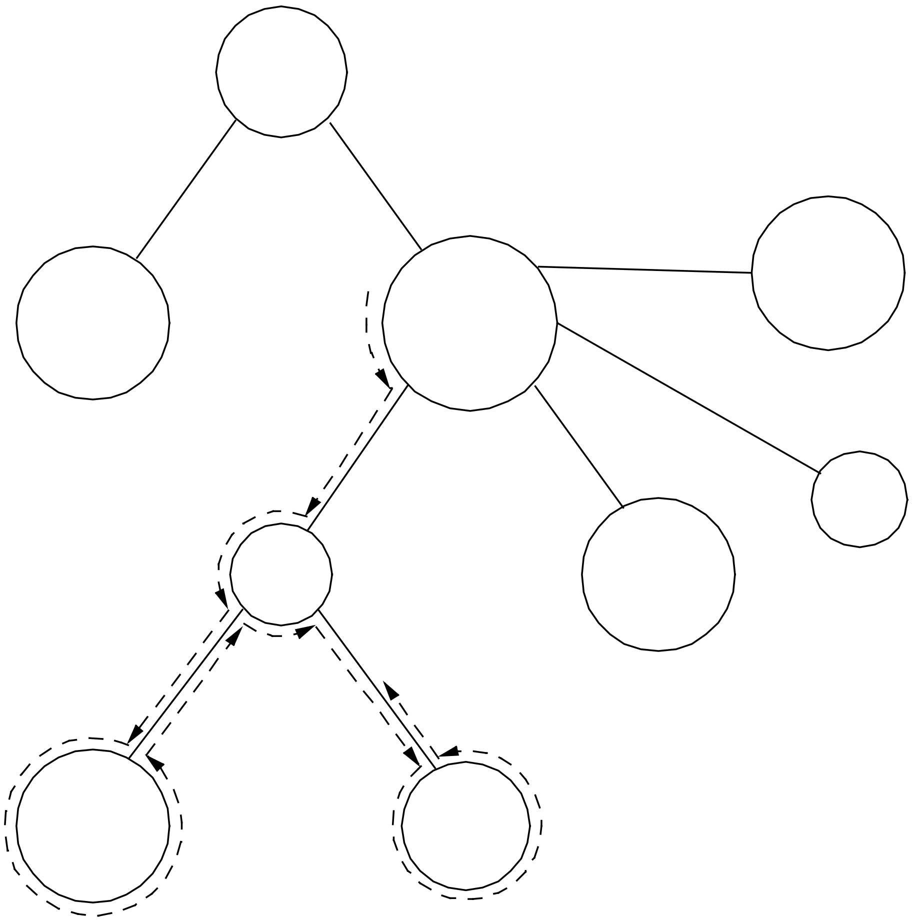
Skeleton Graphs and Promenades

- A *skeleton graph* has the structure of a “tree of cycles”.
- A *skeleton promenade* is a closed Eulerian tour induced by a tree of cycles



Skeleton Graphs and Promenades

- A *skeleton graph* has the structure of a “tree of cycles”.
- A *skeleton promenade* is a closed Eulerian tour induced by a tree of cycles
- A *skeleton walk* is a sub-walk of a skeleton promenade.



Characterization of RALP-failure

Theorem: Every NPCM-promenade is a skeleton promenade.

- characterization → bound

$$\Pr\{fail\} \le \Pr\{\exists \text{skeleton promenade } M \ \& \ c[M] \le 0\}$$

- $g \triangleq$ girth (=log n)
- Distinction between two types of promenades:
  - Short promenades: length $< 2g + 2$ → P_{short}
  - Long promenades: length $\geq 2g + 2$ → P_{long}

Short and Long promenades - Intuition

$$\Pr\{c[e_i] = -1\} = p \quad << \quad 1$$

- Promenade is simple ⇒

$$E\{c[prom.]\} \geq \tfrac{1}{2} \cdot |prom.|(1-2p) \quad >> \quad 0$$

- Chernoff bounds ⇒
  - Long promenades are "easy"
  - Short promenades are "hard"

Problem: <u>repetitions</u> (dependency).

Short NPCM-Promenades (length < 2g+2)

- Few Hamiltonian edges; Few errors ⇒ non-positive cost
- Claim: every short NPCM-promenade is a simple cycle, namely:

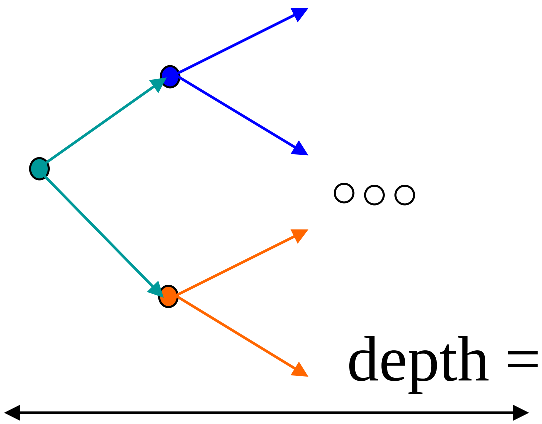
$$P_{short} = \Pr\{\exists \text{ simple short cycle } C \ : \ c[C] \le 0\}$$

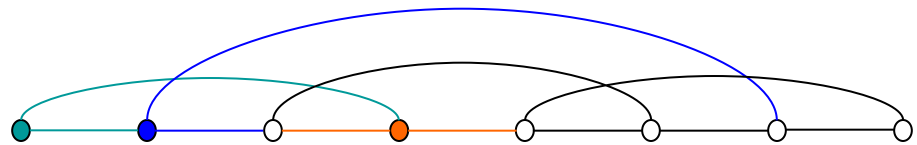
- For a cycle C with h Hamiltonian edges:

$$\Pr\{c[C] \le 0\} = \underbrace{\sum_{i=\lceil \frac{h}{2} \rceil}^{h} \binom{h}{i} p^i (1-p)^{h-i}}_{\text{Majority of Hamiltonian edges are negative.}}$$

Short NPCM-Promenades (Cont.)

- One can enumerate all short simple cycles in polynomial time.





depth = $\log n$ $\rightarrow$ $2^{\log n} = n$ paths

- Lower bound:
  - Consider cycles with fewest Hamiltonian edges.
  - Deal with intersections of cycles: Compute $\Pr\{\exists \text{ cycle } C : c[C] \leq 0\}$ using inclusion-exclusion principle.

Long NPCM-Promenades (length $\geq 2g+2$)

- Lemma: If there exists a long NPCM-promenade M, then there exists a non-positive cost skeleton walk that contains $g + 1$ Hamiltonian edges (with repetitions).

$$P_{long} \leq \Pr\{\exists \text{ skeleton walk W } : c[W] \leq 0 \ \& \ ham(W) = g+1\}$$

- Computed similarly to the tree-bound of Feldman *et al*. [FKW02].

Experimental Results

$n = 1024$, $g = 10$; values in log scale ($\log_{10}$)

p	-2	-3	-4	-5
skeleton-bound	-1.42	-6.39	-9.51	-12.52
P_w Lower Bound	-3.53	-6.52	-9.52	-12.52
P_{long} Upper Bound	-1.43	-7.19	-12.49	-17.53
P_{short} Upper Bound	-3.13	-6.47	-9.51	-12.52
Tree-bound [FKW02]	No Bound	-2.75	-5.75	-8.75

Experimental Results

$n = 1024$, $g = 10$; values in log scale ($\log_{10}$)

p	-2	-3	-4	-5
skeleton-bound	-1.42	-6.39	-9.51	-12.52
P_w Lower Bound	-3.53	-6.52	-9.52	-12.52
P_{long} Upper Bound	-1.43	-7.19	-12.49	-17.53
P_{short} Upper Bound	-3.13	-6.47	-9.51	-12.52
Tree-bound [FKW02]	No Bound	-2.75	-5.75	-8.75

- NPCM-promenades characterization → Improve previous bounds by ~×1000.

Experimental Results

n = 1024, g =10 ; values in log scale ($\log_{10}$)

p	-2	-3	-4	-5
skeleton-bound	-1.42	-6.39	-9.51	-12.52
P_w Lower Bound	-3.53	-6.52	-9.52	-12.52
P_{long} Upper Bound	-1.43	-7.19	-12.49	-17.53
P_{short} Upper Bound	-3.13	-6.47	-9.51	-12.52
Tree-bound [FKW02]	No Bound	-2.75	-5.75	-8.75

- Upper and lower bounds are close ($p \rightarrow 0$) .

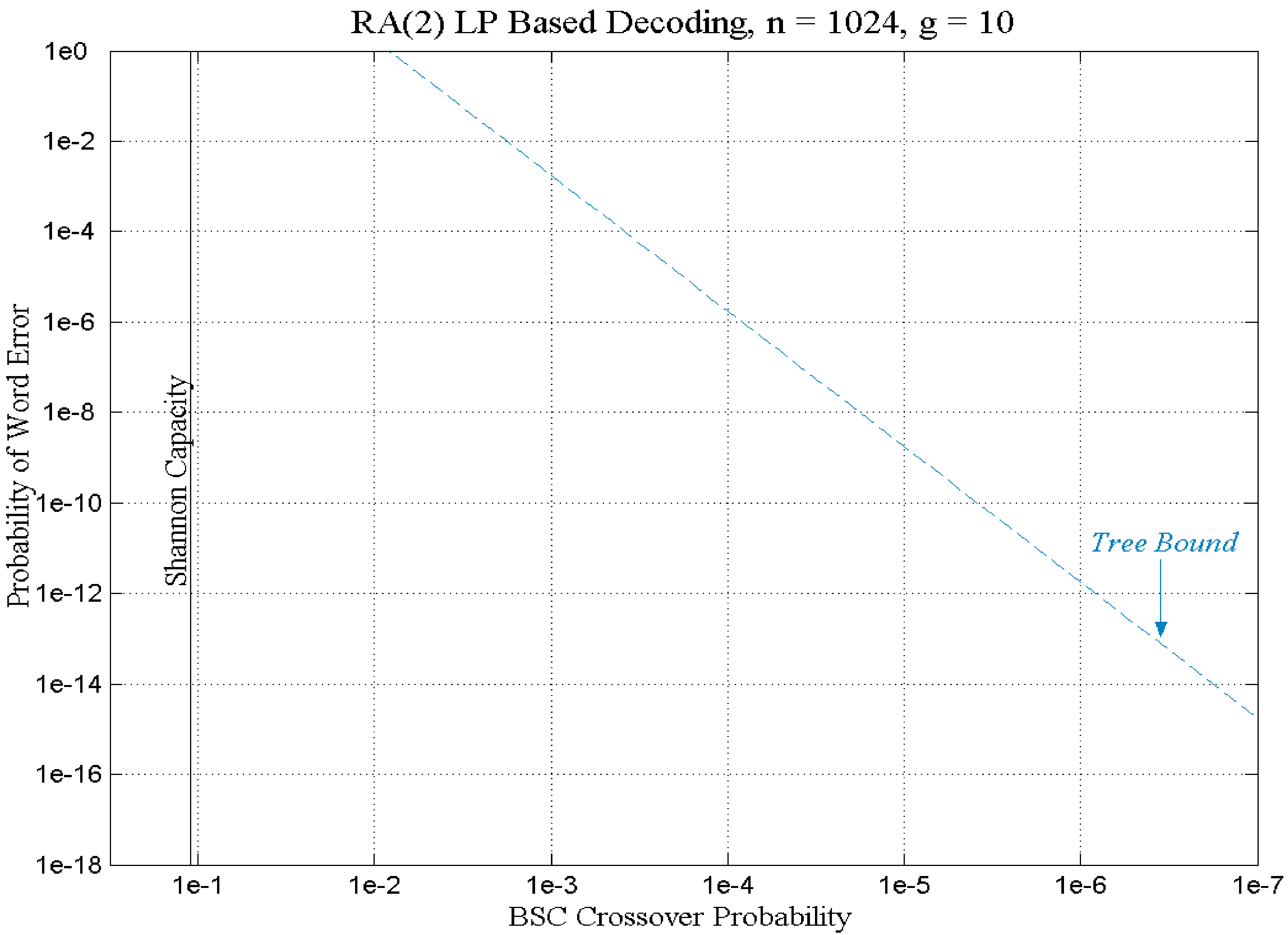
Experimental Results

$n = 1024,\ g = 10$; values in log scale ($\log_{10}$)

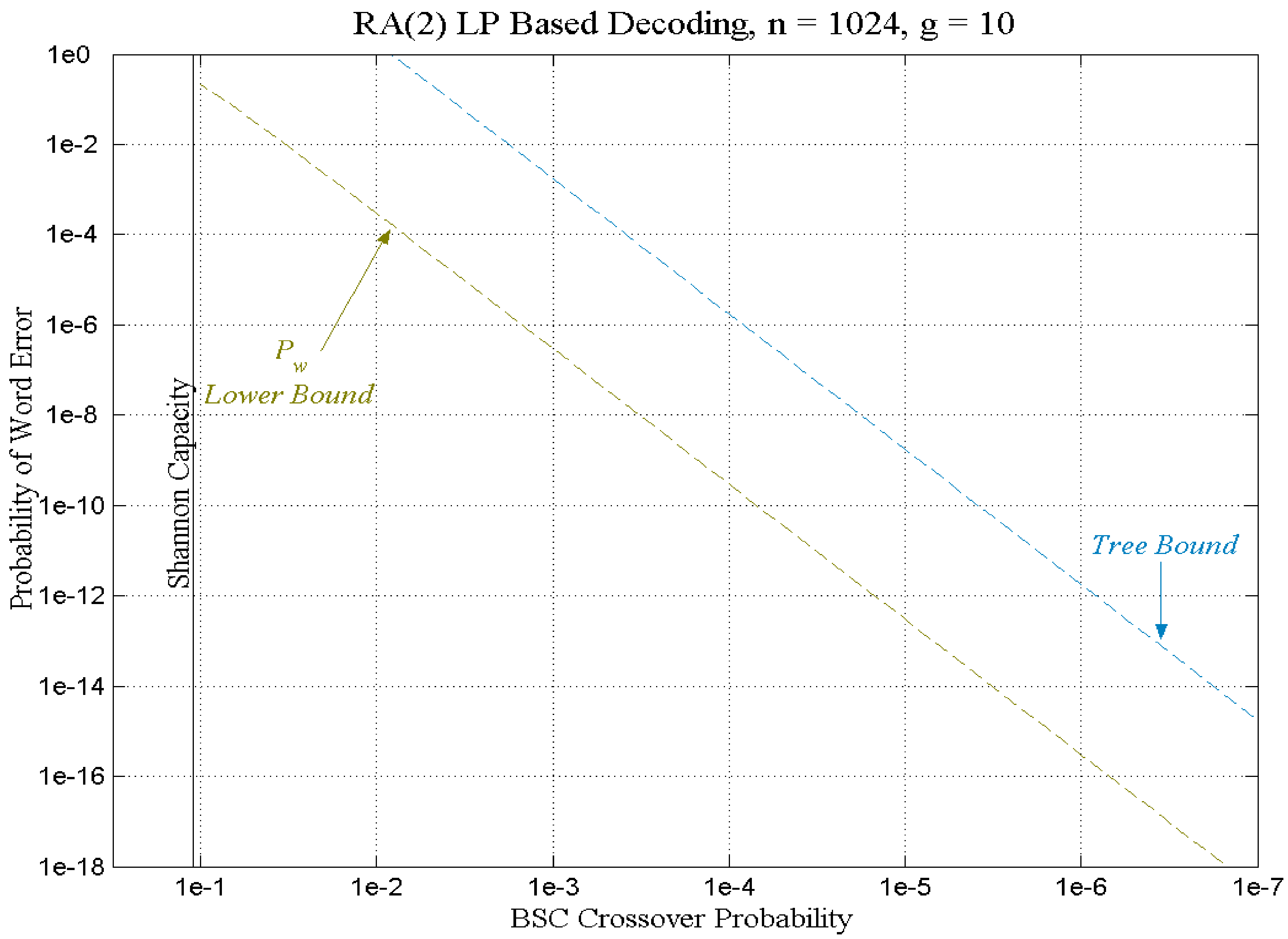
p	-2	-3	-4	-5
skeleton-bound	-1.42	-6.39	-9.51	-12.52
P_w Lower Bound	-3.53	-6.52	-9.52	-12.52
P_{long} Upper Bound	-1.43	-7.19	-12.49	-17.53
P_{short} Upper Bound	-3.13	-6.47	-9.51	-12.52
Tree-bound [FKW02]	No Bound	-2.75	-5.75	-8.75

- Short promenades determine P_w. $\left(P_{short} \xrightarrow[p \to 0]{} P_w\right)$

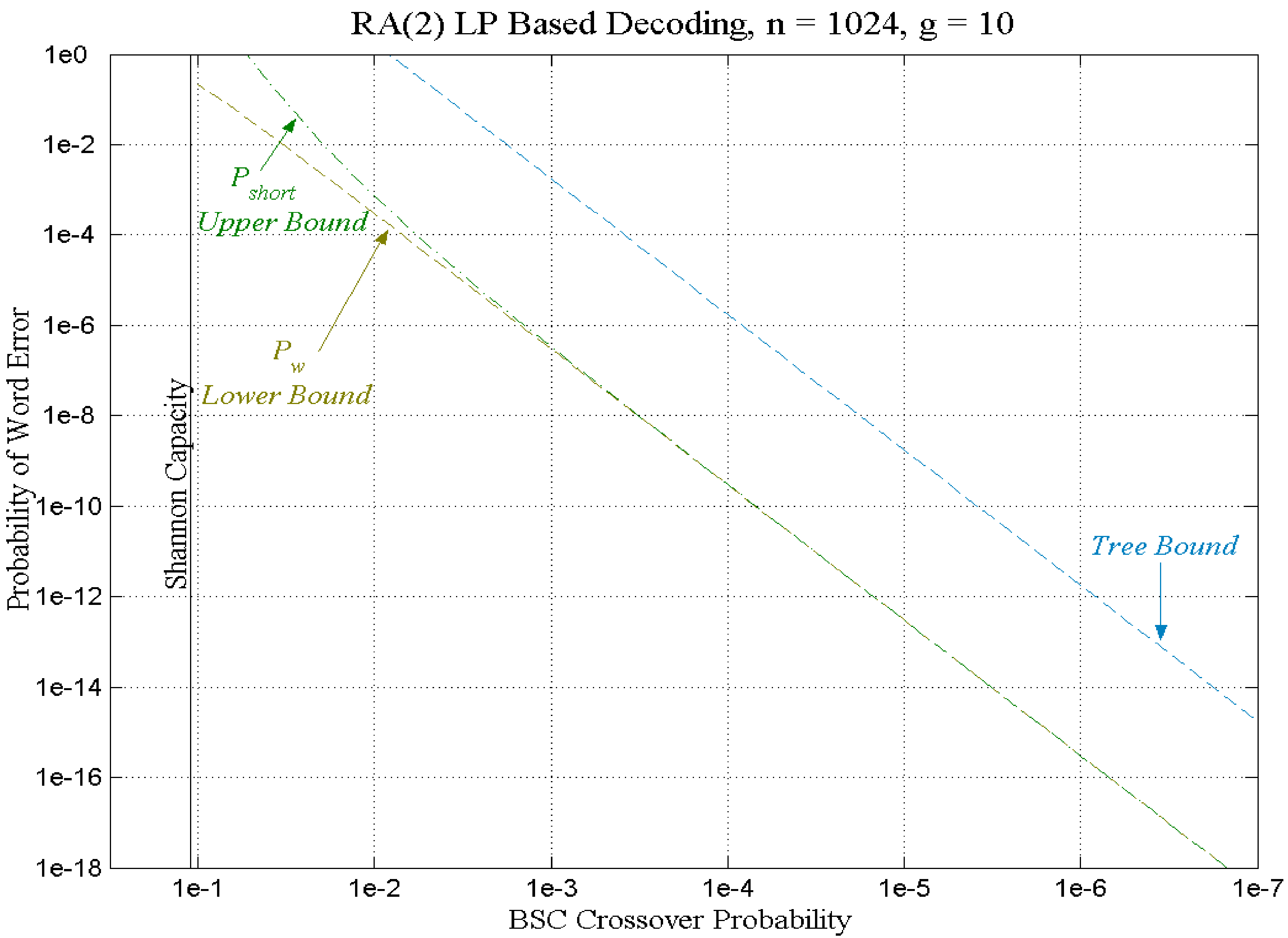
Experimental Results



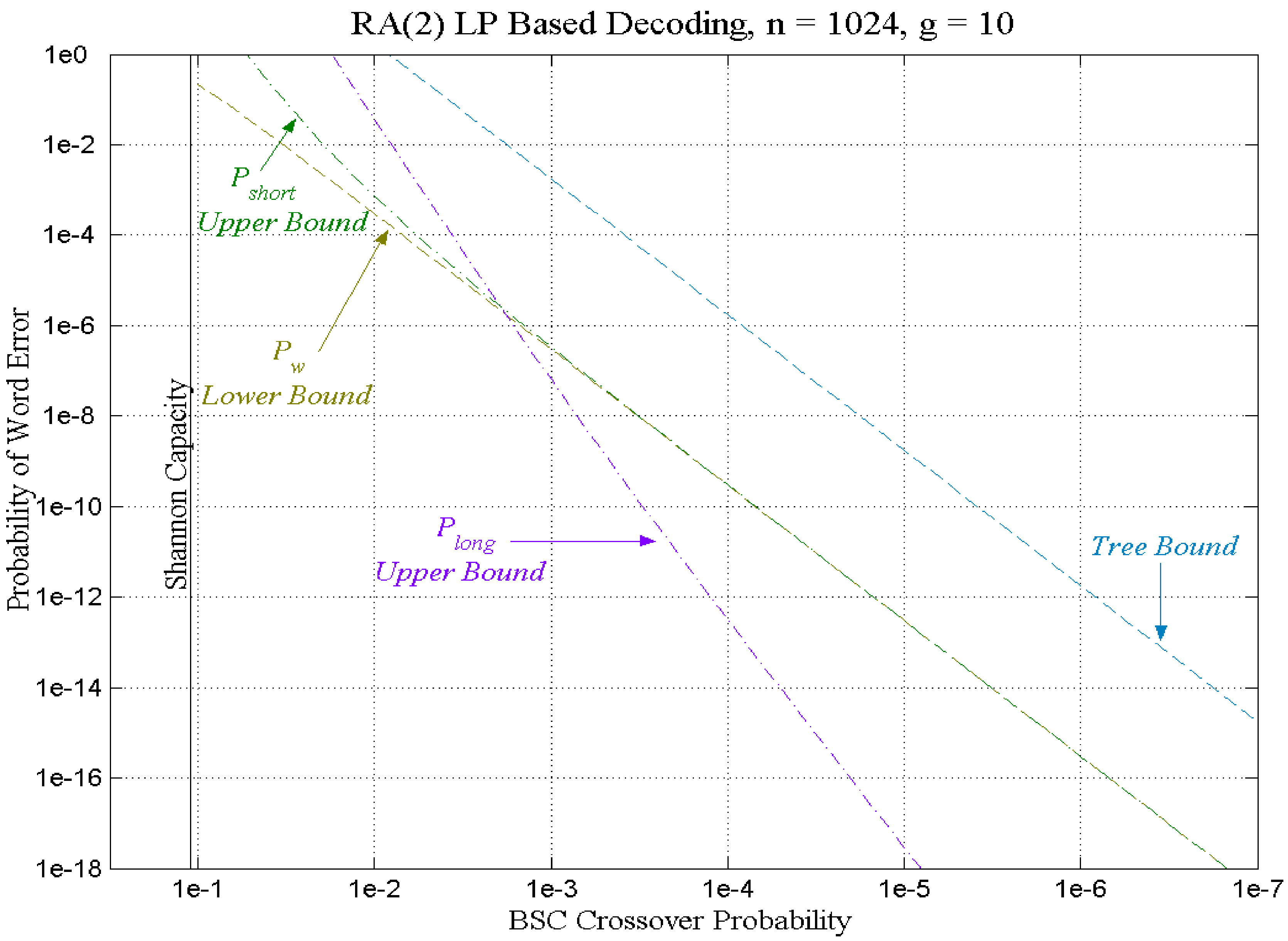
Experimental Results



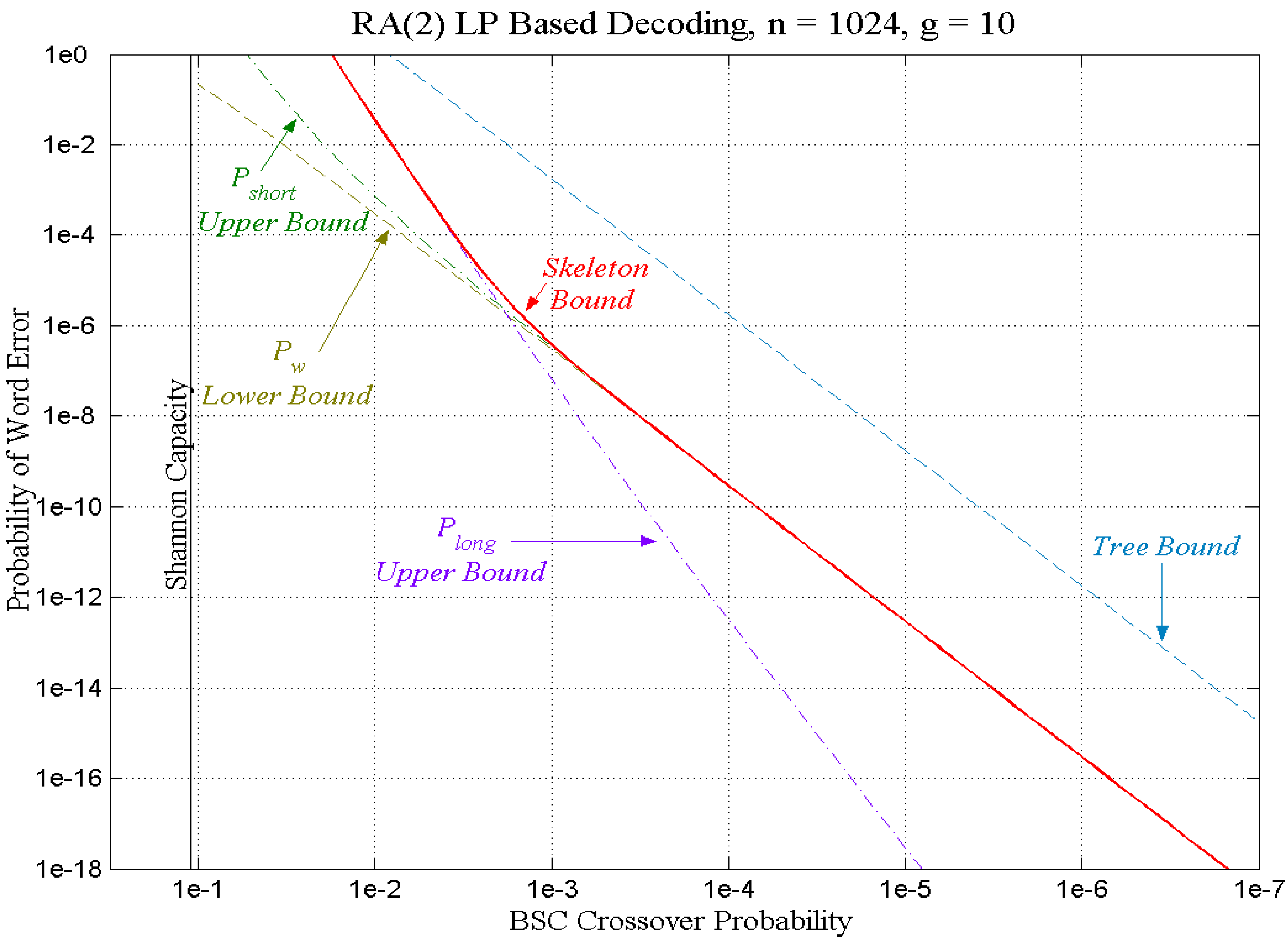
Experimental Results



Experimental Results



Experimental Results



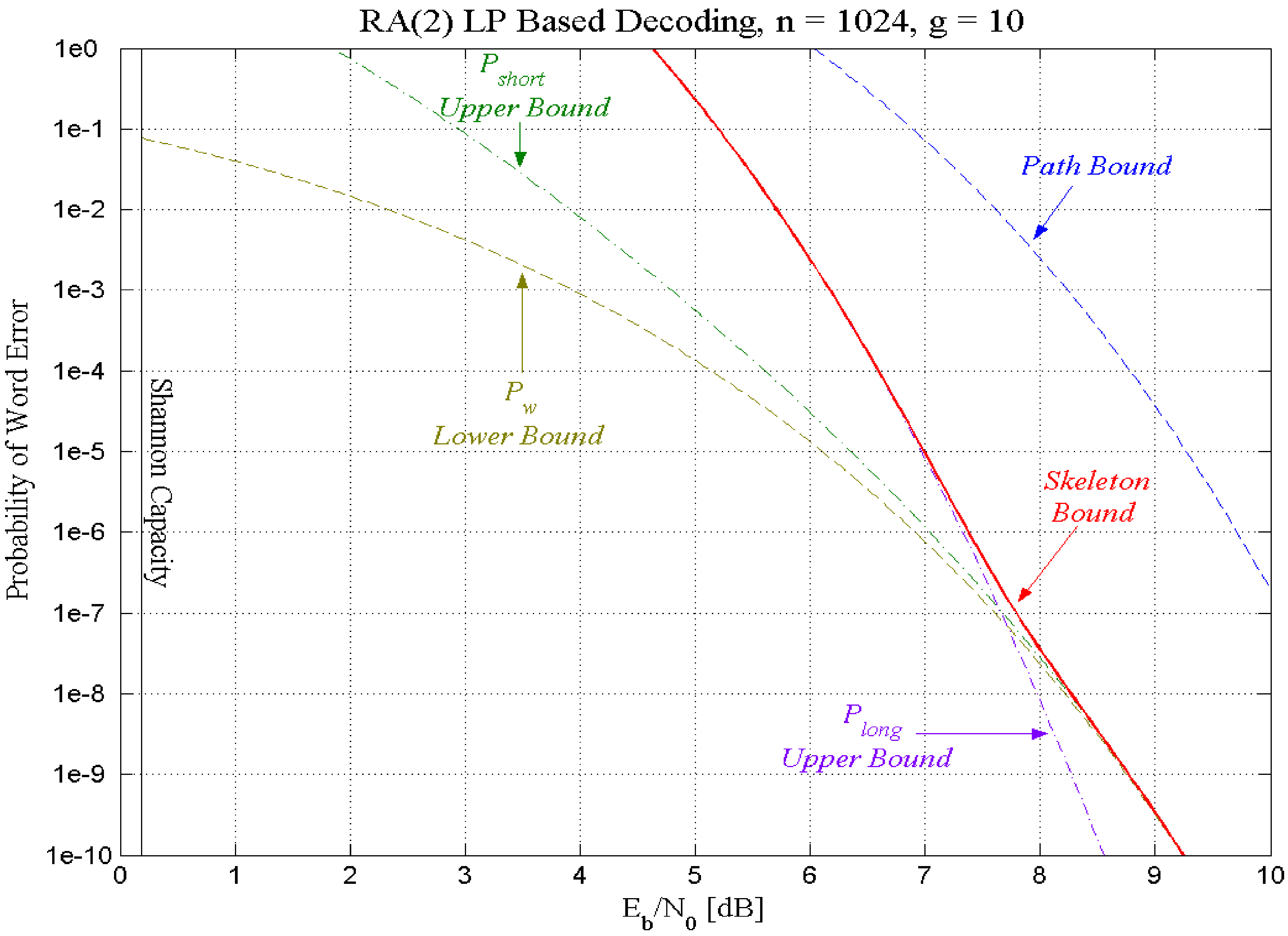
Conclusion

- New characterization of RALP decoding failure.
- Efficient algorithms for computing upper- & lower-bounds on P_w.
- Experimental results:
  - P_w smaller by ~×1000.
  - Lower bound close to upper bound.

Open problems

- Bound for specific RA(3) codes.
- Coding theorem for RA(3).

Experimental Results



Experimental Results

$n = 1024$, $g = 10$; values in log scale ($\log_{10}$)

p	-2	-3	-4	-5
skeleton-bound	-1.42	-6.39	-9.51	-12.52
P_w Lower Bound	-3.53	-6.52	-9.52	-12.52
P_{long} Upper Bound	-1.43	-7.19	-12.49	-17.53
P_{short} Upper Bound	-3.13	-6.47	-9.51	-12.52
Tree-bound [FKW02]	No Bound	-2.75	-5.75	-8.75
exp-tree-bound	No Bound	-2.93	-5.93	-8.93

- Applying universal bounds to specific RA(2) codes → Minor improvement.

Experimental Results