On the Overhead Factors of Optimized LDGM-Staircase and LDGM-Triangle FEC Codes

Dejan Vukobratović, Vojin Šenk, Dragana Bajić
Department of Communications and Signal Processing
University of Novi Sad

Overview

- Introduction
- DF vs. LDPC FEC Schemes
- Overhead Factor
- LDGM-type FEC Codes
- Simulation Results
- Conclusions

Introduction

- Large Scale Content Distribution systems:
  - Increasingly popular Internet application
  - To distribute a large file from one source to many nodes in the network
- Possible Scenarios:
  - Uncoded block transmission over Overlay (P2P) Networks (e.g. Bittorrent)
  - Coding allowed only at the source node
    - Digital Fountain (DF) approach
    - LDPC-based FEC approach
  - (Network) Coding allowed at intermediate nodes (e.g. Microsoft Avalanche)

DF vs. LDPC FEC approach

- DF approach:
  - Chop original file into K equal-length packets
  - Produce infinite amount of encoded packets by randomly XOR-ing subsets of original packets
  - Critical Part: Design of Degree Distributions
  - Any slightly more than K encoded packets enough for successful file recovery at the destinations
  - Universally nearly optimal

DF vs. LDPC FEC approach

- LDPC FEC approach
  - Finite approximation of DF
  - Select appropriate rate R = K/N LDPC code to handle worsest channel
  - Encode original K packets into length N codeword
  - Permute and send encoded packets
  - “Greedy” BP decoder at receiver
  - Hope for enough encoded packets to arrive at the destinations (Minimize Overhead Factor)

LDPC Code Design

- Design of LDPC Codes for BEC Channel
  - Theoretically mature
  - Asymptotically solved
    - Use capacity-approaching optimized (λ, ρ)-irregular LDPC ensemble
  - Finite-length design
    - From a given ensemble of LDPC codes select one with the size of the Minimal Stopping Set maximized
    - Unsolved.

Overhead Factor

- Overhead Factor F
  - Performance metric of interest in Content Distribution applications
  - Quantifies average amount of excessive encoded packets needed for the file reconstruction
  - $K' = (1 + F) \cdot K$
  - $K' > K$ is average number of encoded packets needed for file recovery
  - We got used to BER Curves, Waterfall and Error Floor Regions, DE Thresholds… Any Connections?

Asymptotic Overhead Factor

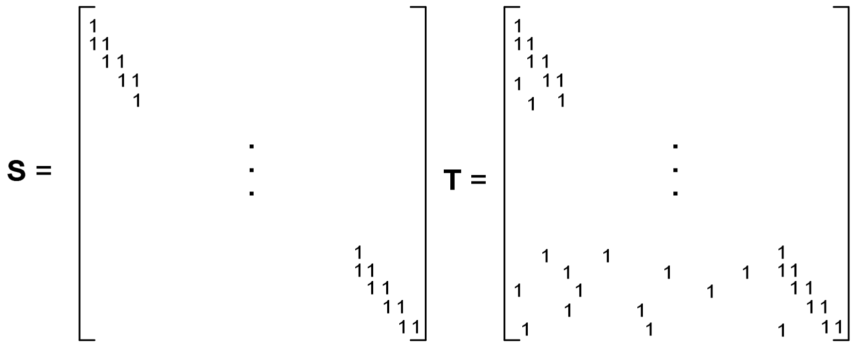
- Asymptotic Overhead Factor F_∞
  - $\epsilon^{Th}(\lambda, \rho)$: DE Threshold for (λ, ρ) LDPC ensemble
  - Ensemble rate $R = 1 - \epsilon^{Th}(\lambda, \rho)$ over BEC($\epsilon^{Th}(\lambda, \rho)$)
  - As $N \to \infty$ number of erased packets sharply concentrate around $N \cdot \epsilon^{Th}(\lambda, \rho)$ (Typicality)
  - Since we're on an ensemble capacity limit $N \cdot (1 - \epsilon^{Th}(\lambda, \rho))$ received packets is enough to drive error probability $\to 0$
  - Equalizing this with $(1 + F_\infty) \cdot K$
  - $F_\infty = \dfrac{1-\varepsilon^{Th}}{R} - 1$

Overhead Factor on Finite Lengths

- Finite-Length Overhead Factor
  - Should be related to Error-Floor performance?
  - Optimizing Error-Floor behavior:
    Maximizing Minimal Stopping Sets
  - Applying State-of-the-art design algorithms that optimizes Error-Floor behavior and check what happens with an Overhead Factor at short-to-medium code lengths

LDGM-type Codes

- Low-Density Generator Matrix Codes
  - Class of LDPC Codes with parity-check matrix suitable for efficient encoding purposes
  - H=[P I], where P is sparse
  - Derivates: LDGM-Staircase and LDGM-Triangle LDPC Codes
  - Replace I with



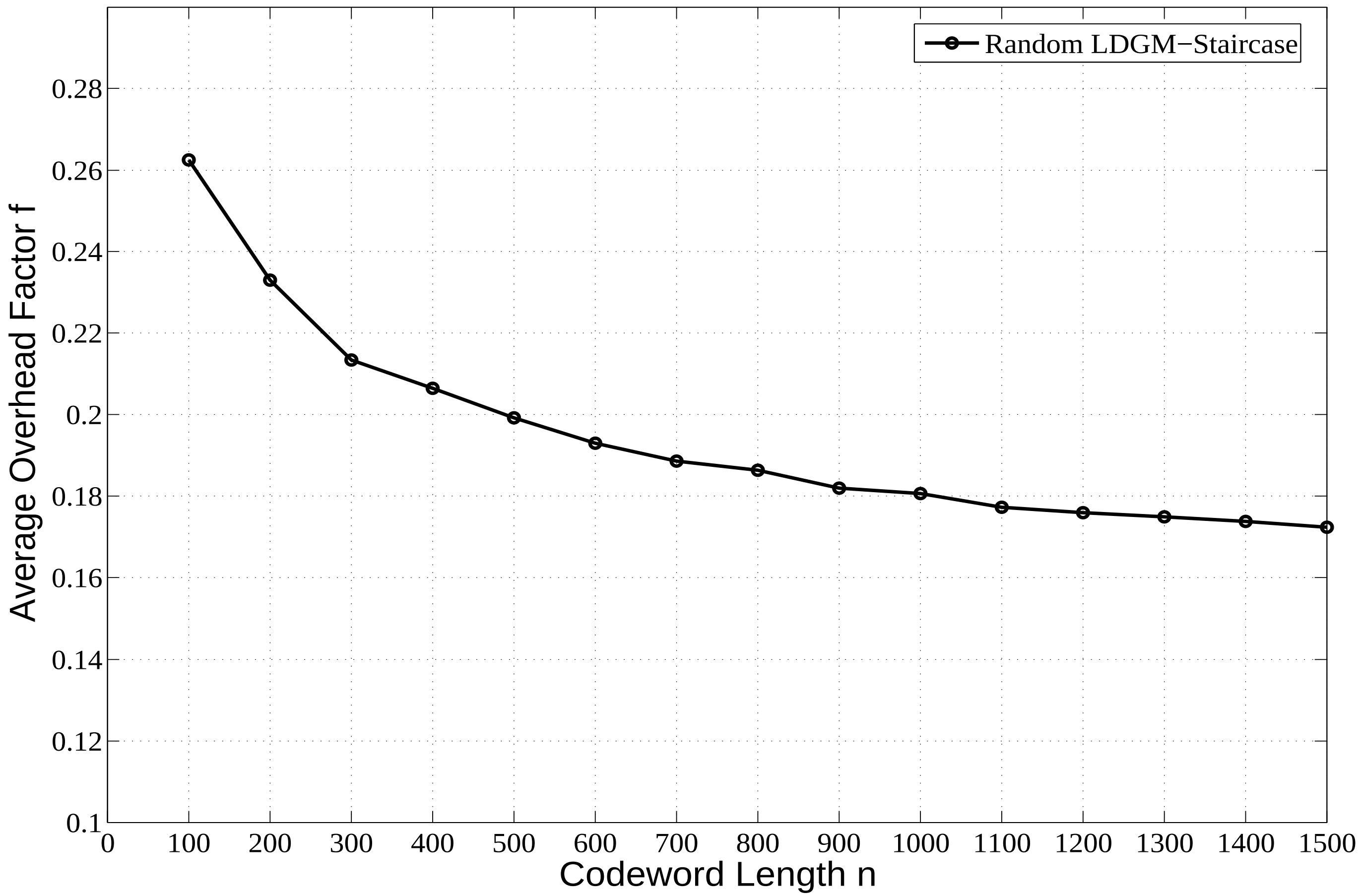
Large Content Distribution Systems

- Standardization work
  - DF Raptor Codes as a core mechanism for
    - 3GPP MBMS (Multimedia Broadcast/Multicast System)
    - DVB-H IP Content Delivery
  - Possibilities for Patent-Free Codecs are left open
  - LDGM-type design codes are suggested
    - V. Roca and C. Neumann “Low Density Parity Check (LDPC) Forward Error Correction”, June 2005. IETF Draft Work in Progress: <draft-roca-rmt-ldpc-00.txt>

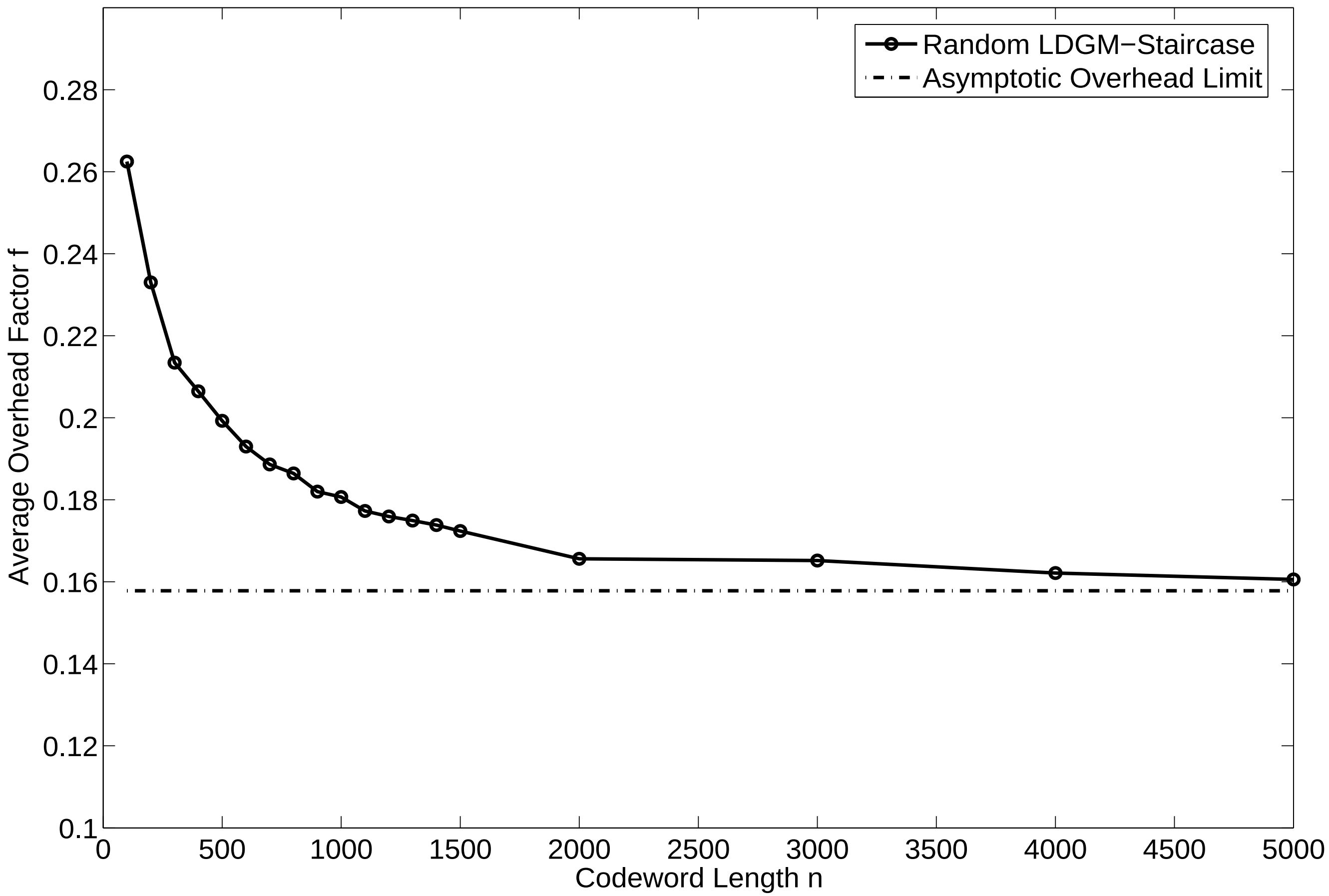
Simulation Results

- Simulation Setting
  - Overhead Factor F is examined by MC simulation for R = ½ left-quasi regular LDGM-Staircase codes of increasing length from 100-1500 with step 100
  - For each length 1000 experiments with random selection of H and random permutation of codeword symbols before transmission
  - All-zero codeword assumption
  - Based on all generated H matrices, average degree distribution pair (λ, ρ) and corresponding $\epsilon^{Th}(\lambda, \rho)$ is calculated

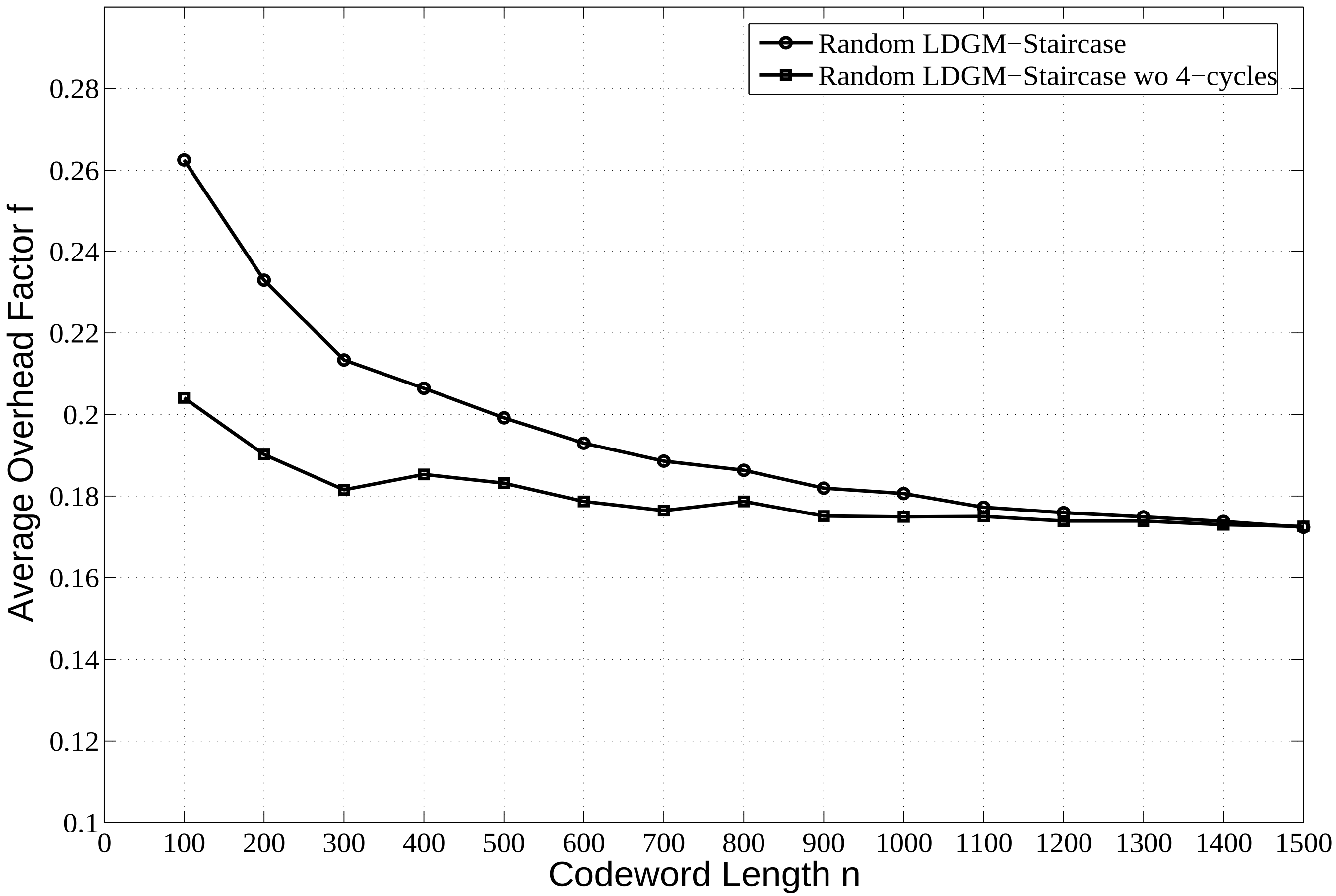
Random LDGM Staircase Design



Asymptotic Overhead Factor



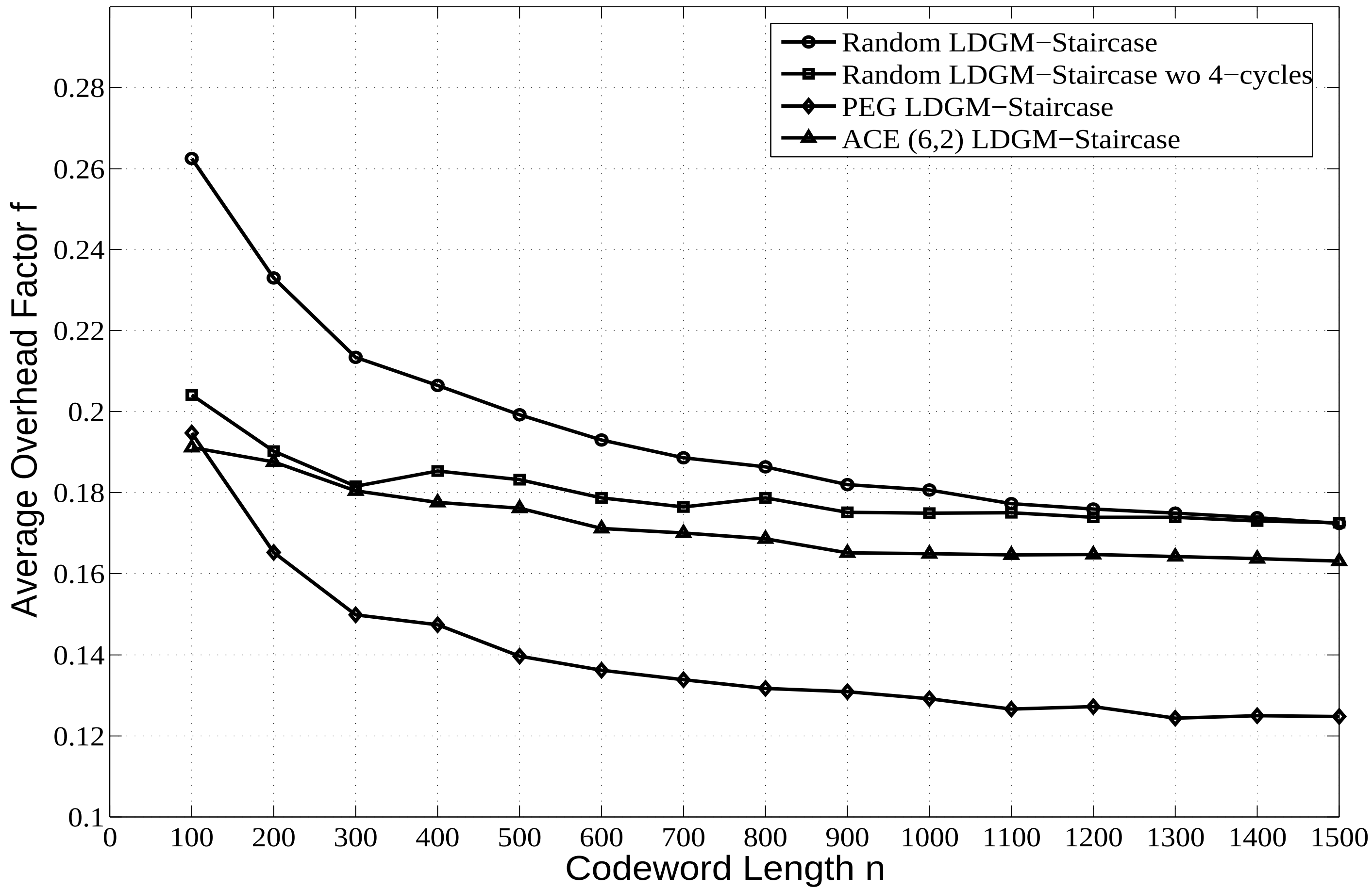
Random LDGM Staircase Design without 4-cycles



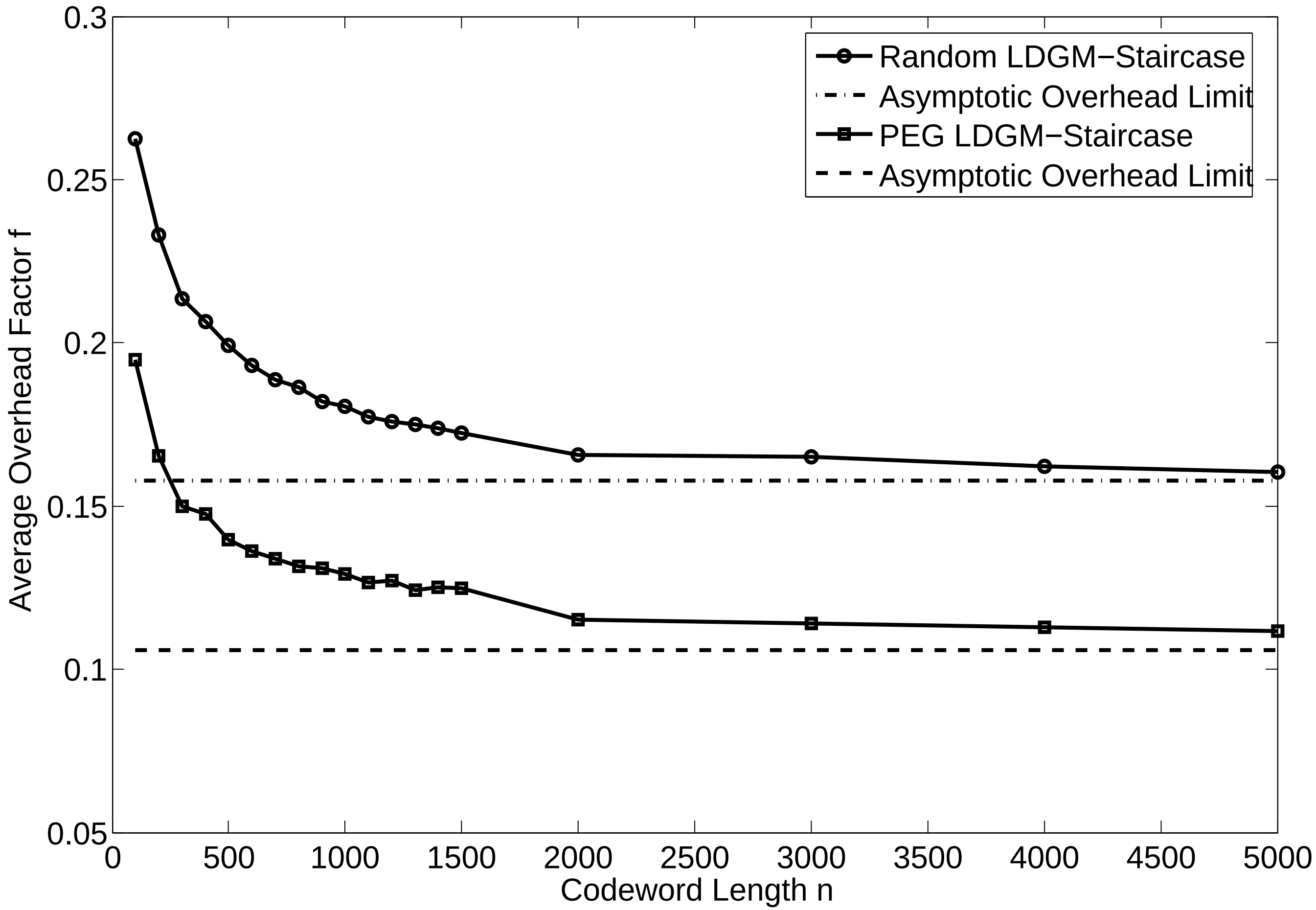
Improved LDGM-Staircase Design

- Two well-known algorithms
  - Progressive Edge Growth (PEG) Algorithm
    - [Hu, Eleftheriou, Arnold (2002)]
    - Heuristic algorithm for large girth graph design
  - ACE Constraining Algorithm
    - [Tian, Jones, Villasenor, Wesel (2004)]
    - Associates ACE metric (measure of cycle connectivity with the rest of the graph) for each cycle in graph
    - (d_{max},η) ACE constrained LDPC codes
    - Each cycle of length less than or equal to $2d_{max}$ has ACE value larger or equal than η

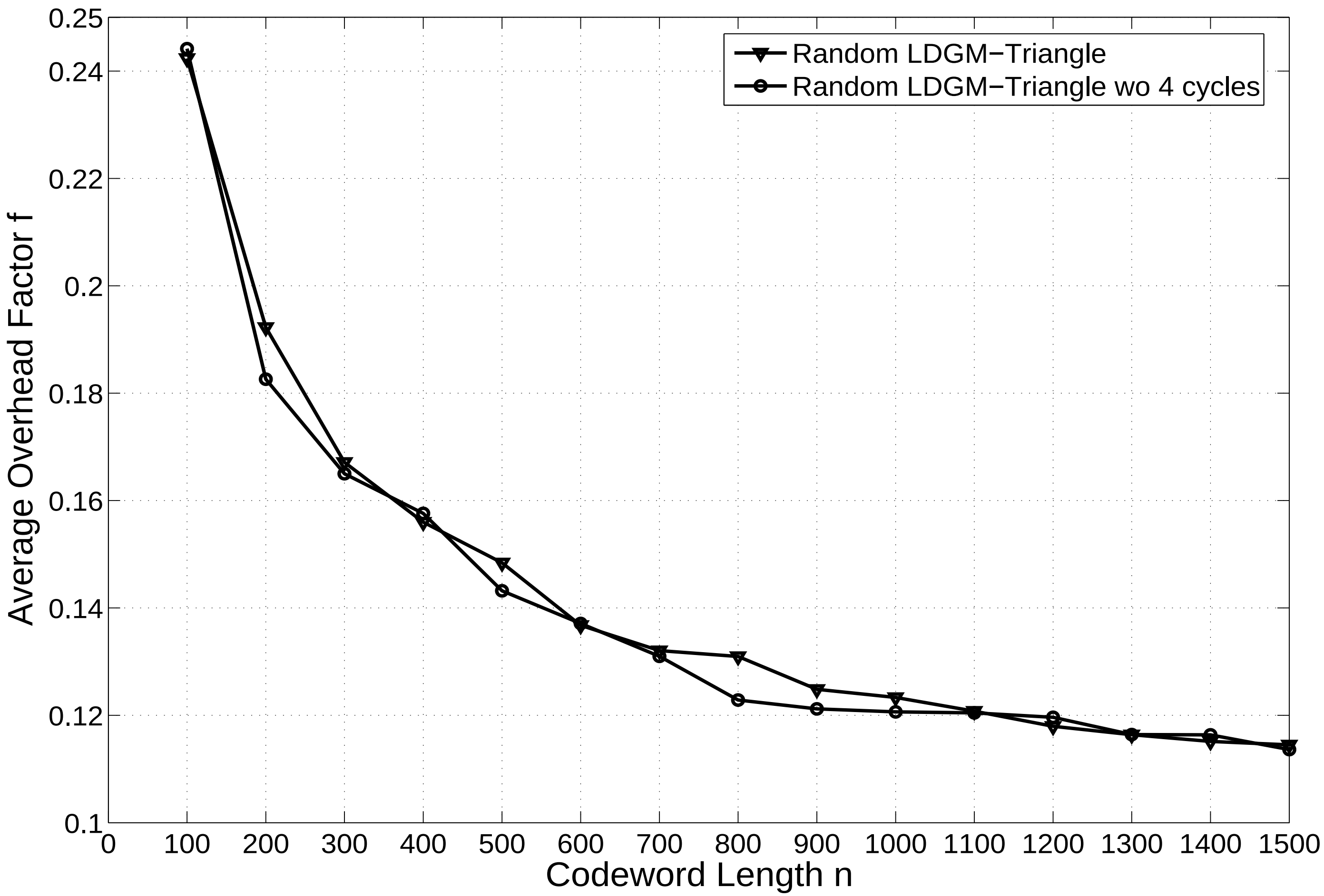
Improved LDGM-Staircase Design



Asymptotic Overhead Factor



LDGM-Triangle Design



Conclusions and Future Work

- Overhead Factor
  - Performance measure of interest for Large Content Distribution Applications
  - Not much of a recent work
  - Patent-free constrained codes
  - Significant improvements over currently suggested codes are possible
- Future Work
  - Generally, extensive design analysis and simulation study is needed