On the Representation of Correlated Exponential Distributions by Phase Type Distributions

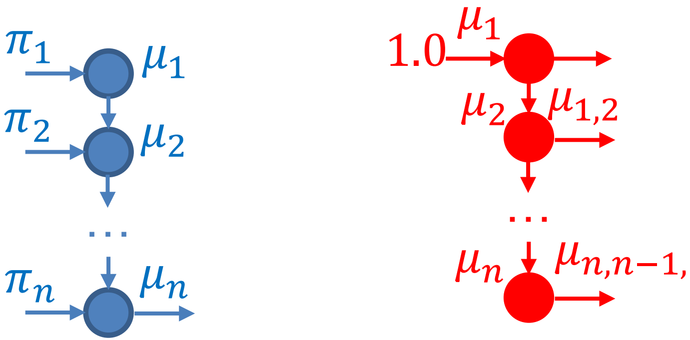
Peter Buchholz
Informatik IV, TU Dortmund

Outline

- Representation of exponential distributions by acyclic phase type distributions (APHDs)
- Correlation between APHDs
- Optimal representations
- Application

Acyclic Phase Type Distributions (APHDs)

Canonical representations



For exponential distribution with $E(X) = \mu^{-1}$ we have Laplace transform

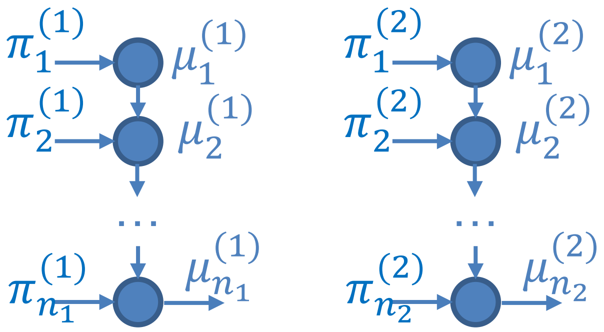
$$L(s) = \frac{\mu}{\mu + s}$$

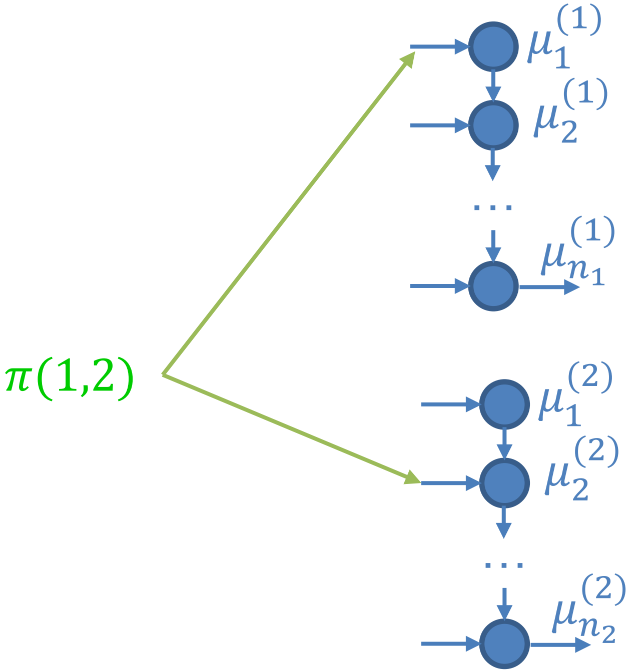
we assume $\mu = 1$ (normalized)

- $\boldsymbol{m}(i)$ sojourn time depending on entry state i
- $\boldsymbol{a}(i)$ sojourn time depending on exit state i
- $\boldsymbol{\pi}$ entry probabilities, $\boldsymbol{\psi}$ exit probabilities

Correlated APHDs

Parallel composition

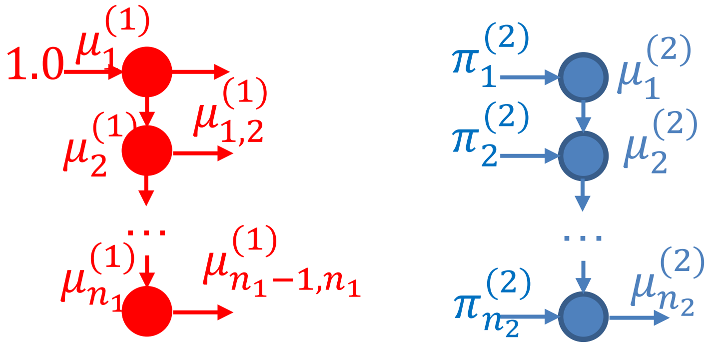


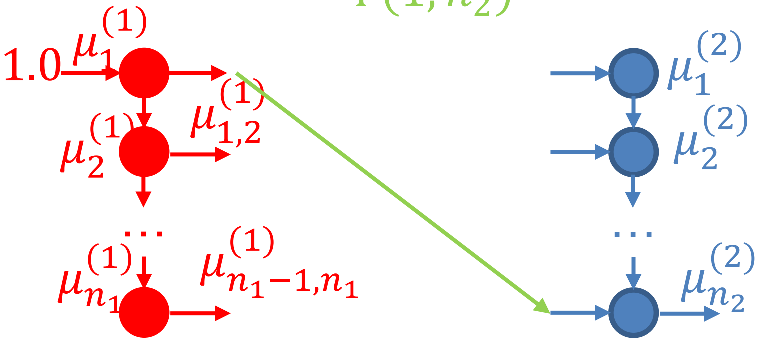


$$\sum_{j=1}^{n_2} \pi(i,j) = \pi_i^{(1)} \text{ and } \sum_{i=1}^{n_1} \pi(i,j) = \pi_j^{(2)}$$

Correlated APHDs

Sequential composition





$$\sum_{i=1}^{n_2} \psi_i^{(1)} \Psi(i,j) = \pi_j^{(2)} \text{ and } \sum_{j=1}^{n_1} \Psi(i,j) = 1$$

Minimal/maximal correlation for given APHDs

Linear program (here sequential composition, similar for parallel)

$$\rho_{X,Y}^{+/-(n_1,n_2)} = \min/\max \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \Psi(i,j) \boldsymbol{a}_1(i) \boldsymbol{m}_2(j)$$

$$s.t. \sum_{i=1}^{n_2} \psi_i^{(1)} \Psi(i,j) = \pi_j^{(2)} \text{ and} \sum_{j=1}^{n_1} \Psi(i,j) = 1$$

But what are optimal APHD representations for given n_1 and n_2?

APHD used by Bladt/Nielson 2010 for sequential composition

- Defined for arbitrary n with $\mu_i = i$ and $\pi_i = {}^1/_n$ or $\mu_{i,i+1} = 1$
- We obtain

$$\rho_{X,X}^{+(n)} = 1 - \frac{1}{n}\sum_{i=1}^{n} \frac{1}{i} \text{ such that } \lim_{n \longrightarrow \infty} \rho_{X,X}^{+(n)} = 1 \text{ (optimal)}$$

$$\rho_{X,X}^{-(n)} = 1 - \sum_{i=1}^{n} \frac{1}{i^2} \text{ such that } \lim_{n \longrightarrow \infty} \rho_{X,X}^{-(n)} = 1 - \frac{\pi^2}{6} \text{ (optimal)}$$

Can we do better?

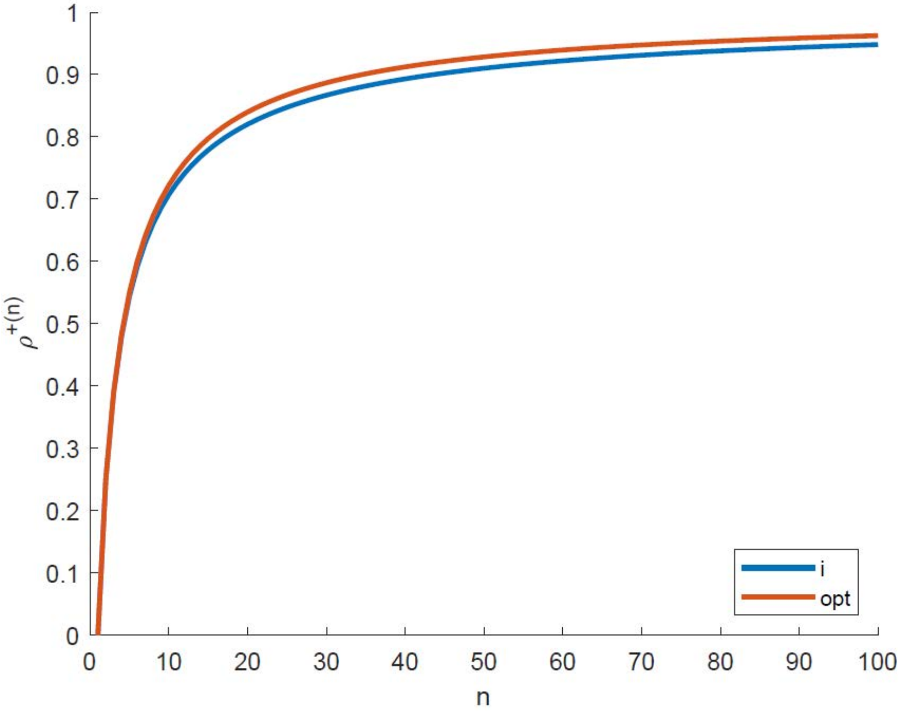
Positive correlation

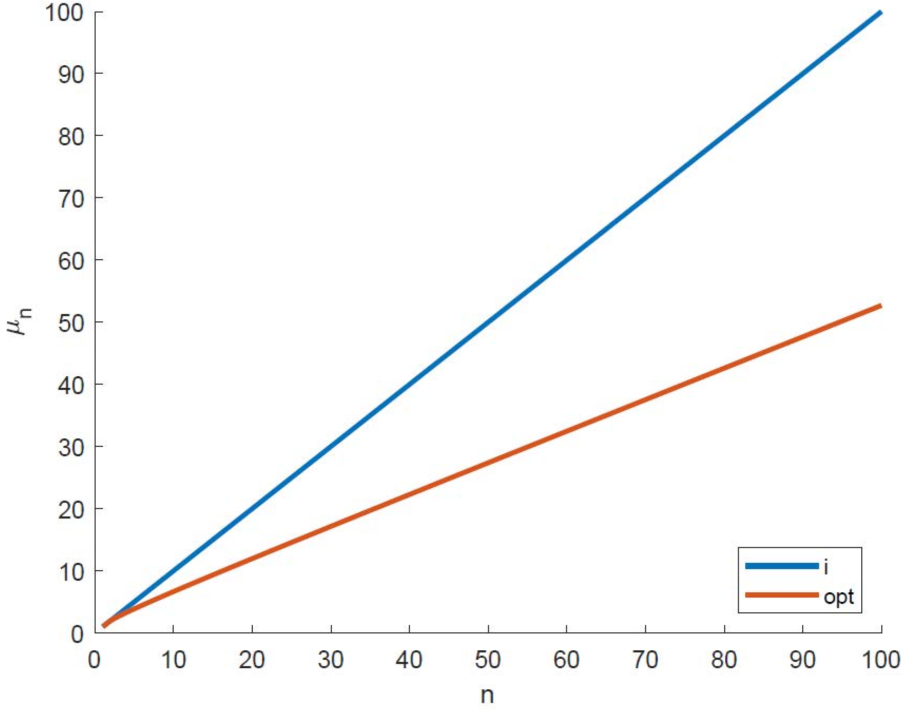
- Algorithm:
  - Start with $n = 1$ and $\mu_1 = 1$
  - Compute consecutively μ_{n+1} from $\mu_1, \ldots, \mu_n$ resulting in

$$\mu_{n+1} = \frac{2}{1 - \rho_{X,X}^{+(n+1)}} \text{ and } \rho_{X,X}^{+(n+1)} = \rho_{X,X}^{+(n)} + 0.25\left(1 - \rho_{X,X}^{+(n)}\right)^2$$

- Faster convergence towards 1 than rates i
- Local conditions for optimality are observed

Positive correlation



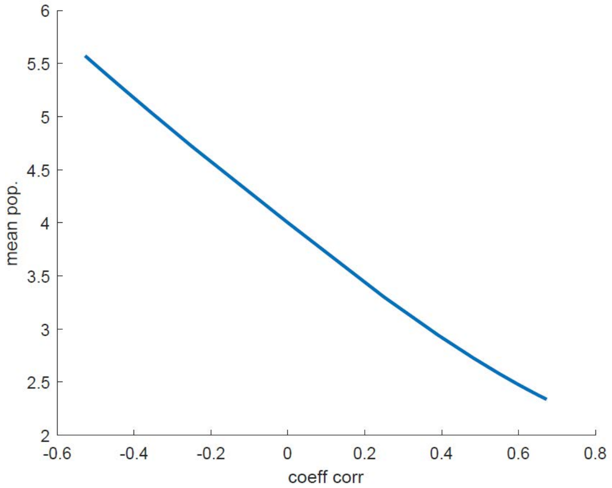


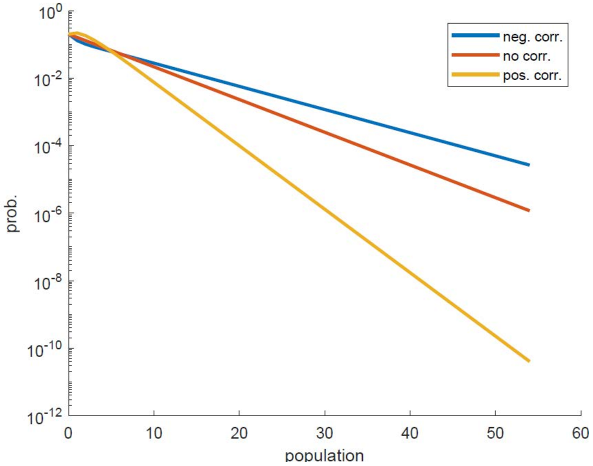
Negative correlation

- Optimization problem with additional degree of freedom
- Consecutive computation of rates does not work
- Rates i observe local conditions for optimality
- Result is optimal for $n = 2$ but for $n = 3$ a better representation could be found

An example

- APHD/APHD/1 (M/M/1) queue where inter-arrival and service times are correlated (matrix geometric solution with block size $O\left(n_{arrival}n_{service}^2\right)$)





Conclusion

- An approach to describe correlated exponential distributions by acyclic phase type distributions
- Improved representation for positive correlation
- More detailed results can be found in https://arxiv.org/pdf/2108.12223
  - Building blocks for general PHDs
  - Extension to Markovian Arrival Processes