



Efficient Algorithms for Computation of
Equilibrium/Transient Probability Distribution of
Finite Markov Chains: Potential Lower Bound on
Computational Complexity

Rama Murthy Garimella

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

October 4, 2018

EFFICIENT ALGORITHMS FOR COMPUTATION OF EQUILIBRIUM/TRANSIENT PROBABILITY DISTRIBUTION OF FINITE CONTINUOUS TIME MARKOV CHAINS: POTENTIAL LOWER BOUND ON COMPUTATIONAL COMPLEXITY

Garimella Rama Murthy,

Professor,

Department of Computer Science & Engineering,

Mahindra Ecole Centrale, Hyderabad, India

ABSTRACT

In this research paper, efficient algorithms for computation of equilibrium as well as transient probability distribution of arbitrary finite state space Continuous Time Markov Chains are proposed. The effective idea is to solve a structured systems of linear equations efficiently. The algorithms potentially achieve lower bound on the computational complexity (of solving structured system of linear equations arising in the equilibrium/transient analysis of Continuous Time Markov Chains). It is realized that the approach proposed in this research paper generalizes to infinite state space Continuous Time Markov Chains. The ideas of this research paper could be utilized in solving structured system of linear equations that arise in other applications.

1. Introduction:

Markov chains provide interesting stochastic models of natural/artificial phenomena arising in science and engineering. The existence of equilibrium behavior enables computation of equilibrium performance measures. Thus, researchers invested considerable effort in EFFICIENTLY computing the equilibrium probability distribution of Markov chains and thus, the equilibrium performance measures.

Traditionally, computation of transient probability distribution of Continuous Time Markov Chains (CTMCs) was considered to be a difficult open problem. It requires computation of MATRIX EXPONENTIAL associated with the generator matrix of CTMC. Even in the case of finite state space CTMCs, efficient computation of transient probability distribution was considered to be a difficult open problem. In [Rama1], an interesting approach for recursive computation of transient probability distribution of arbitrary finite state space CTMCs was proposed. In the case of infinite state space, Quasi-birth-and Death (QBD) processes, matrix geometric recursion for transient probability distribution was found in [ZhC].

It is well known that computation of equilibrium distribution of CTMCs reduces to solution of linear system of equations. The approach proposed in [Rama 2] reduces the computation of transient probability

distribution of finite state space CTMCs to solving linear system of equations in the transform (Laplace Transform) domain. Thus, an interesting question that remained deals with efficient solution of such STRUCTURED linear system of equations in the Laplace transform domain. In fact, a more interesting problem is to design an algorithm which meets a LOWER BOUND on the computation of solution of structured system of linear equations arising in the transient/equilibrium analysis of CTMCs.

This research paper is organized as follows. In Section 2, algorithm for efficient computation of equilibrium probability mass function of Continuous Time Markov Chains (CTMCs) is discussed. In Section 3, algorithm for efficient computation of transient probability mass Function of Continuous Time Markov Chains is discussed. In section 4, we briefly discuss how the proposed algorithm potentially achieves the lower bound on computation of equilibrium/transient Probability Mass Function. The research paper concludes in Section 5.

2. Efficient Computation of Equilibrium Probability Mass Function of Continuous Time Markov Chains:

It is well known that computation of equilibrium probability vector (PMF) $\bar{\pi}$ of a Homogeneous Continuous Time Markov Chain (CTMC) with generator \bar{Q} reduces to the solution of following linear system of equations (i.e. computation of vector $\bar{\pi}$ in the left null space of the generator matrix \bar{Q}) i.e.

$$\bar{\pi} \bar{Q} \equiv \bar{0}.$$

In computational linear algebra, there are efficient algorithms to solve linear system of equations, [Gall], [Sto], [Wil]. Volker Strassen made fundamental contributions to this problem [Str1], [Str2]. In fact, "COMPUTATIONALLY OPTIMAL ALGORITHMS" are designed to reduce the computational complexity (in terms of elementary operations such as additions, multiplications etc.). But, from the point of view of computing $\bar{\pi}$, they do not take into account the structure of generator matrix \bar{Q} .

Thus, the goal of this research paper is to design efficient algorithm for computing $\bar{\pi}$, taking into account the structure of generator matrix \bar{Q} . Some related effort

dealing with Laplacian system of equations is discussed in [DGA]. In fact, we would like to design a COMPUTATIONALLY OPTIMAL ALGORITHM for computation of $\bar{\pi}$ (*in terms of Computational Complexity*).

We illustrate the essential idea with a Continuous Time Markov chain (CTMC) with 4 states. First consider the generator matrix of a CTMC with 4 states (i.e. 4×4 matrix) partitioned using blocks of size 2×2 (i.e. there are 4 such 2×2 matrices in the generator) [RaR1], [RaR2]. Specifically, we have

$$Q = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \text{where } \{A_{ij}; i, j \in \{1,2\}\} \text{ are}$$

2×2 square matrices in the 4×4 block matrix Q .

Claim: For a positive recurrent (recurrent non-null) CTMC with generator matrix Q ,

it is well known that the matrices $\{A_{11}, A_{22}\}$ are non-singular.

Hence, we have that

$$A_{21} = A_{22} X \quad \text{with } X = A_{22}^{-1} A_{21}.$$

Thus, by means of elementary column operations, the generator matrix Q can be converted into the following form i.e.

$$\tilde{Q} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & A_{22} \end{bmatrix} \quad \text{i.e. block upper triangular matrix.}$$

In fact, let $X = [f_1 \ f_2]$. Hence, column operations to arrive at \tilde{Q} (from Q) are determined by the column vectors $\{f_1 \ f_2\}$. Also, it follows from linear algebra, that the equilibrium probability vector $\bar{\pi}$ is unaffected by such column operations. Hence, with $\bar{\pi} = [\bar{\pi}_1 \ \bar{\pi}_2]$, we have the following system of linear equations i.e.

$$\bar{\pi}_1 \tilde{A}_{11} = \bar{0} \quad \text{i.e. boundary system of two equations}$$

$$\bar{\pi}_1 \tilde{A}_{12} + \bar{\pi}_2 A_{22} = \bar{0}. \quad \text{Thus, we have } \bar{\pi}_2 = -\bar{\pi}_1 \tilde{A}_{12} A_{22}^{-1}.$$

The boundary system of linear equations leads to a single linear equation in 2 variables i.e. denoting $\bar{\pi}_1 = [\pi_1^1 \ \pi_1^2]$, we have a linear equation of the form

$$\pi_1^1 \alpha + \pi_1^2 \beta = 0. \quad \text{Thus, } \pi_1^2 = -\pi_1^1 \frac{\alpha}{\beta}.$$

Further, since $\bar{\pi}_2$ can be expressed in terms of $\bar{\pi}_1$, we utilize the normalizing equation $(\bar{\pi}_1 + \bar{\pi}_2) \bar{e} = 1$, where \bar{e} is a column vector of ones,

to determine π_1^1 and hence all the other equilibrium probabilities.

Remark 1:

We realize that the inverse of a 2×2 matrix can be computed by inspection and essentially only requires computation of the determinant. For instance, let us consider

a 2 x 2 matrix B i.e. $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$. It is well known that $B^{-1} = \frac{1}{\Delta} \begin{bmatrix} b_{22} & -b_{12} \\ -b_{21} & b_{11} \end{bmatrix}$, where $\Delta = \text{determinant}(B) = (b_{11}b_{22} - b_{12}b_{21})$.

COMPUTATIONAL COMPLEXITY: Now, we determine the computational complexity of our algorithm:

- The number of arithmetic operations required to convert the generator matrix Q into block upper triangular matrix \tilde{Q} needs to be included in calculating the computational complexity of the algorithm
- π_1^2 computation requires ONE division.
- $\overline{\pi_2}$ computation requires inversion of 2 x 2 matrix A_{22} and one matrix multiplication (of 2 x 2 matrices \tilde{A}_{12} and A_{22}).
- Also, the normalizing equation to determine π_1^1 requires 3 additions and ONE division.
- Finally, to determine $\overline{\pi_1}, \overline{\pi_2}$ in terms of π_1^1 , we require 3 multiplications.

Now, we generalize the above algorithm for the case, where the number of states, $N = 2m$ where $m > 1$. In such case, the generator matrix is of the following form:

$$Q = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1m} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2m} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ A_{m1} & A_{m2} & A_{m3} & \dots & A_{mm} \end{bmatrix} \text{ where } A'_{ij} \text{ s are } 2 \times 2 \text{ matrices.}$$

Lemma 1: For a positive recurrent CTMC, the following sub-matrices of generator i.e. $\{A_{ii}: 1 \leq i \leq m\}$ are all non-singular matrices.

Proof: Refer [Rama1]. The proof utilizes the fact that strictly diagonally dominant matrix is non-singular.

Hence, as in the $m=1$ case, by means of elementary column operations on the generator matrix Q, we arrive at the following block upper triangular matrix \tilde{Q} .

$$\tilde{Q} = \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} & \widetilde{A}_{13} & \cdots & \widetilde{A}_{1\ m-1} & \widetilde{A}_{1m} \\ 0 & \widetilde{A}_{22} & \widetilde{A}_{23} & \cdots & \widetilde{A}_{2\ m-1} & \widetilde{A}_{2m} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \widetilde{A}_{m-1\ m-1} & \widetilde{A}_{m-1\ m} \\ 0 & 0 & 0 & \dots & 0 & \widetilde{A}_{mm} \end{bmatrix}.$$

Thus, the equilibrium probability vector $\bar{\pi}$ satisfies the following linear system of equations:

$$\begin{aligned} \bar{\pi}_1 \widetilde{A}_{11} &\equiv \bar{0} \\ \bar{\pi}_1 \widetilde{A}_{12} + \bar{\pi}_2 \widetilde{A}_{22} &\equiv \bar{0} \\ &\vdots \\ \bar{\pi}_1 \widetilde{A}_{1\ m-1} + \bar{\pi}_2 \widetilde{A}_{2\ m-1} + \cdots + \bar{\pi}_{m-1} \widetilde{A}_{m-1\ m-1} &\equiv \bar{0} \\ \bar{\pi}_1 \widetilde{A}_{1m} + \bar{\pi}_2 \widetilde{A}_{2m} + \cdots + \bar{\pi}_m \widetilde{A}_{mm} &\equiv \bar{0}. \end{aligned}$$

The above system of linear equations is recursively solved to compute the equilibrium probability vector i.e.

$$\begin{aligned} \bar{\pi}_1 \widetilde{A}_{11} &\equiv \bar{0} \\ \bar{\pi}_2 &= -\bar{\pi}_1 \widetilde{A}_{12} \widetilde{A}_{22}^{-1}. \\ &\vdots \\ \bar{\pi}_j &= -\bar{\pi}_1 \widetilde{A}_{1j} \widetilde{A}_{jj}^{-1} - \bar{\pi}_2 \widetilde{A}_{2j} \widetilde{A}_{jj}^{-1} - \cdots - \bar{\pi}_{j-1} \widetilde{A}_{j-1\ j} \widetilde{A}_{jj}^{-1} \text{ for } 2 \leq j \leq (m-1). \\ \bar{\pi}_m &= -\bar{\pi}_1 \widetilde{A}_{1m} \widetilde{A}_{mm}^{-1} - \bar{\pi}_2 \widetilde{A}_{2m} \widetilde{A}_{mm}^{-1} - \cdots - \bar{\pi}_{m-1} \widetilde{A}_{m-1\ m} \widetilde{A}_{mm}^{-1}. \end{aligned}$$

Note: As in the 2 x2 case, using non-singular matrices on the diagonal of generator matrix, Q, it can be converted to a block upper-triangular matrix. Since inverse of 2 x 2 matrices on the diagonal of Q can be computed efficiently, such an algorithm may be more efficient than the well known algorithms (in numerical linear algebra) to convert Q into an upper triangular matrix. We choose the efficient algorithm from among those two choices.

Definition: Finite Memory Recursion of order "L" for the equilibrium probability vector is of the following form

$$\bar{\pi}(\mathbf{L} + \mathbf{1}) = \bar{\pi}(1) W_1 + \bar{\pi}(2) W_2 + \cdots + \bar{\pi}(L) W_L, \quad \text{where } W_i\text{'s are}$$

recursion matrices. We call such a recursion as “forward finite memory recursion” [RaC]. Such a forward Finite Memory Recursion holds true for even “infinite state space” Continuous Time Markov Chains [Rama1], [Rama2], [RaC].

Remark 2:

Since, \widetilde{A}_{jj} 's are all non – singular 2×2 matrices, their inverse can be efficiently computed.

We now compute the “computational complexity” of the algorithm to Determine equilibrium probabilities in most general case.

- The number of arithmetic operations required to convert the generator matrix Q into block upper triangular matrix \widetilde{Q} needs to be included in calculating the computational complexity of the algorithm

(I) 2×2 Matrix inversions: We require (m-1) inversions of 2×2 matrices. Each 2×2 matrix inversion requires (a) Determinant computation i.e. 2 multiplications and one subtraction (b) Division of elements by determinant i.e. 4 divisions

(II) 2×2 Matrix Multiplications: In the above system of linear equations, we require

$1 + 2 + 3 + \dots + (m-1) = \frac{m(m-1)}{2}$ multiplications of two 2×2 matrices. Each such multiplication (of two 2×2 matrices) requires atmost 8 multiplications and 4 additions. To reduce the complexity, we can utilize Strassen’s multiplication algorithm. The number of multiplications required to compute the product of 2×2 matrices reduces to 7 (using Strassen’s algorithm). But with Strassen’s algorithm we require 18 additions/substractions

(III) Vector Matrix Multiplications: We require

$$1 + 2 + 3 + \dots + (m-1) = \frac{m(m-1)}{2}$$

multiplications of 1×2 and 2×2 matrices (i.e. multiplication of row vector and 2×2 matrix).

(IV) Normalization: It requires $(2 m - 1)$ additions and ONE division to determine π_1^1 .

(V) Finally, we require $(2 m - 1)$ multiplications to determine the equilibrium probabilities using π_1^1 .

Thus, in summary, we need the following number of arithmetic operations:

- (i) Additions / subtractions : $10m^2 - 7m - 2$
- (ii) Multiplications: $\frac{11m^2 - 7m - 4}{2}$
- (iii) Divisions: $4m - 3$,

where $m = \frac{N}{2}$ where the generator matrix Q is $N \times N$ matrix.

Note: In the case of Quasi-Birth-and- Death (QBD) processes, Latouche et.al utilize Gaussian Elimination method to compute the equilibrium probability distribution [GJL]. In fact they provide probabilistic interpretation of the method.

Remark 3:

Suppose, the number of states of the CTMC is an ODD number. Then, we consider the 3×3 boundary system of linear equations (i.e. the initial probability vector is of dimension 3) and utilize the above idea to compute the equilibrium probability vector efficiently.

Remark 4:

It should be noted that using similar idea (as discussed above), by means of elementary column operations, the generator matrix can be converted into a block lower triangular matrix. In this case, the boundary system of linear equations is at the trailing boundary. By solving for the last probability vector, we recursively compute the equilibrium probabilities. We call such a recursion as the “backward finite memory recursion” [Rama1].

3. Efficient Computation of Transient Probability Mass Function of Continuous Time Markov Chains:

It is well known that the transient probability mass function (PMF) of a homogeneous CTMC (generator matrix doesnot depend on time t, unlike non-homogeneous CTMC) satisfies the following vector matrix differential equation i.e.

$$\frac{d}{dt} \bar{\pi}(t) = \bar{\pi}(t) Q .$$

Taking Laplace-Transform on both sides, we have that

$$s \bar{\pi}(s) - \bar{\pi}(0) = \bar{\pi}(s) Q .$$

Equivalently, we have the following expression for $\bar{\pi}(s)$:

$$\bar{\pi}(s) = -\bar{\pi}(0) [Q - s I]^{-1} .$$

These constitute a “structured” system of linear equations. Our goal is to efficiently solve such system of equations for $\bar{\pi}(s)$ and compute the inverse Laplace Transform of $\bar{\pi}(s)$ to arrive at the time dependent (transient) PMF of CTMC.

Lemma 2: For a positive recurrent (recurrent non-null) CTMC, in the Region of Convergence (ROC) (of Laplace Transform) the sub-matrices on the diagonal of $(Q - sI)$ are non-singular

Proof: As in the equilibrium case, such matrices are all strictly diagonally dominant and hence are all non-singular. Q.B.D.

Thus, as in the case of computation of equilibrium PMF, $\bar{\pi}(s)$ can be determined efficiently for the values of ‘s’ lying in the Region of Convergence. Detailed Duplication of equations is avoided for brevity.

4. Lower Bound on Computation of Equilibrium / Transient Probability Mass Function:

From the above discussion, it is clear that the computation of transient/equilibrium Probability Mass Function (PMF) reduces to solving structured system of linear equations. We exploit the fact that inversion of 2×2 matrices (lying on the diagonal of the generator matrix of CTMC) can be carried out in a computationally efficient manner. Using that idea, based on Finite Memory Recursions for the equilibrium/transient PMF, such probabilities are determined with minimum number of arithmetic operations.

Remark 5:

If the generator matrix of a CTMC is arbitrary (without any specific structure as in the case of say QBD process), the above algorithm MOSTLY achieves the lower bound on computational complexity (i.e. number of arithmetical operations are minimum possible in number). In computational linear algebra, there are well known algorithms to achieve the LOWER BOUND on solving an arbitrary linear system of equations [Gall], [Sto], [Wil] . We borrow those ideas alongwith the approach proposed in this research paper to efficiently compute the equilibrium/transient PMF of an arbitrary CTMC. On the other hand the ideas of this research paper could be utilized in solving structured system of linear equations arising in other applications.

Remark 6: Consider an arbitrary system of linear equations of the following form:

$$AX = b, \text{ where } A \text{ is } N \times N \text{ matrix and } X \text{ is } N \times 1 \text{ vector.}$$

We convert the above system of linear equations into homogeneous system in the following manner. Let $\tilde{X}^T = [X^T: 1]$, $\tilde{A} = \begin{bmatrix} A & -b \\ \bar{0} & 0 \end{bmatrix}$. We thus arrive at the following homogeneous system of linear equations

$$\tilde{A} \tilde{X} = \bar{0}.$$

If the matrices on the diagonal of \tilde{A} are all non-singular as in the case of CTMC, the above recursions can be used to compute the vector in the null space of \tilde{A} .

[Remark 7]: The method proposed in this paper can be applied to solving system of linear equations, where the coefficient matrix has the structure similar to a generator matrix (i.e. Structured Diagonally Dominant matrix).

We now provide some numerical results for transient probability mass function of an arbitrary CTMC.

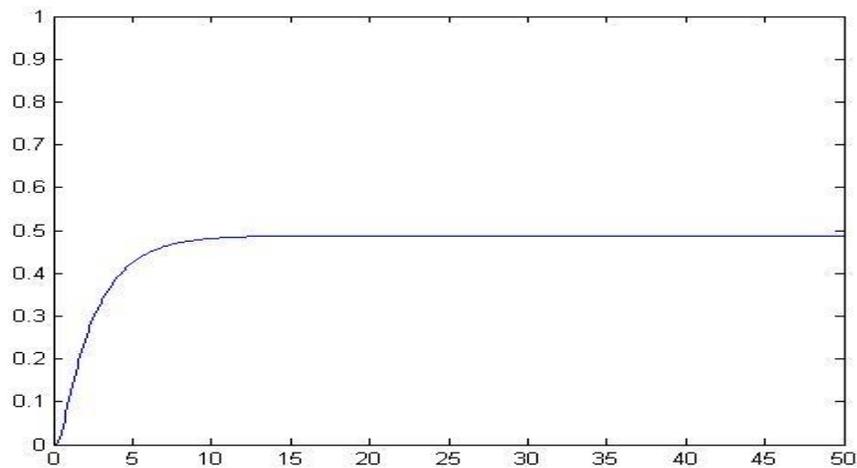


Figure 1. Transient Probability of CTMC being in State 10, as a function of time, given that it starts out in state '1'

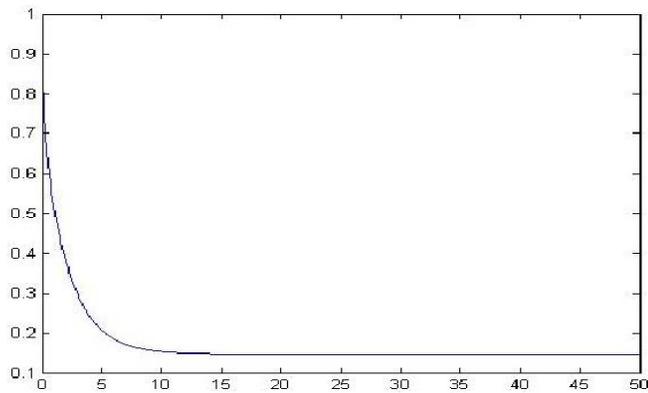
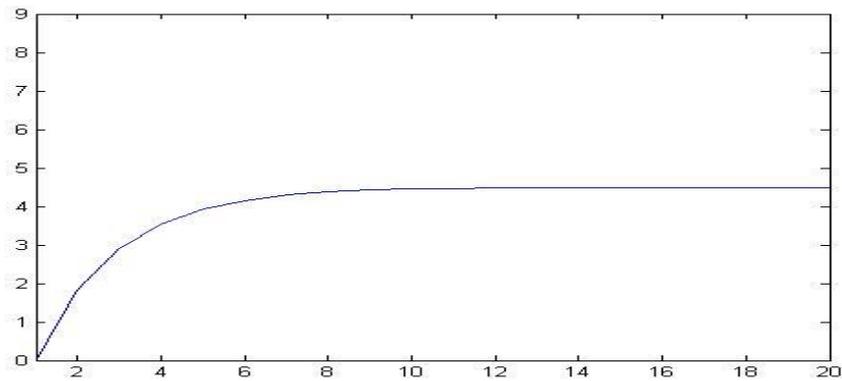
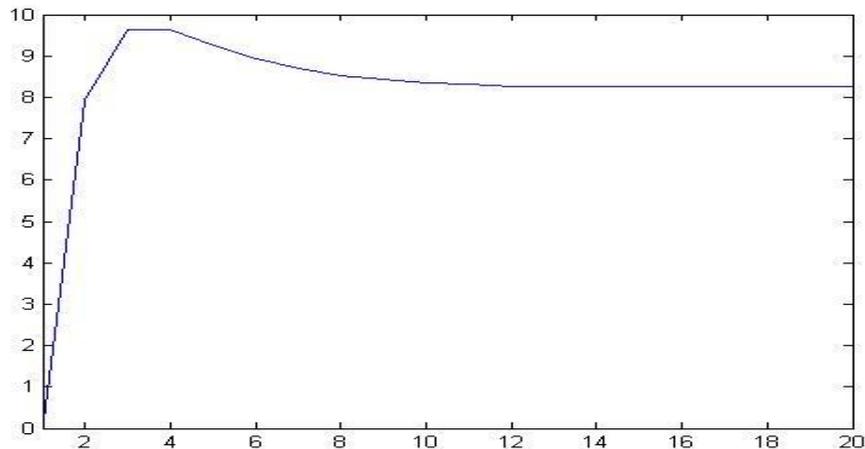


Figure 2: Transient Probability of the CTMC being in state '1' as a function of time, given that the CTMC starts out in state '1'.

- We can plot the expected value of the transient probability distribution and variance of the transient probability distribution (as a function of time) , as follows:



- Variance of the transient probability distribution can also be obtained (as a function of time):



5. Conclusions:

In this research paper, efficient algorithms for computing the equilibrium and transient probability distribution of an arbitrary finite state space Continuous Time Markov Chain are discussed. The algorithms effectively solve a structured system of linear equations efficiently. The algorithms potentially achieve the lower bound on computational complexity.

REFERENCES:

- [DGA] Dean Doron, François Le Gall and Amnon Ta-Shma, " **Probabilistic Logarithmic-Space Algorithms for Laplacian Solvers**," *Proceedings of the 21st International Workshop on Randomization and Computation (RANDOM 2017)*, pp. 41:1-41:20, 2017.
- [Gall] François Le Gall "Powers of Tensors and Fast Matrix Multiplication" **Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation (ISSAC 2014)**, pp. 296-303, 2014.
- [GJL] D. P. Gaver, P. A. Jacobs and G. Latouche, "Finite Birth-and-Death Models in Randomly Changing Environments," **Advances in Applied Probability, Vol. 16, No. 4 (Dec., 1984), pp. 715-731**
- [Rama 1] G. Rama Murthy, "Transient and Equilibrium Analysis of Markov Chains: Forward and Backward Finite memory Recursions: Canonical Forms," **Proceedings of the Princeton Annual Conference on Information Sciences and Systems, pp.574-579, Vol.2, March 1990.**
- [Rama2] G. Rama Murthy," Transient and Equilibrium Analysis of Computer Networks: Finite Memory and Matrix Geometric Recursions," **Ph.D thesis, 1989, Purdue University, West Lafayette, USA**

[RaC] G. Rama Murthy and E.J Coyle, " Finite Memory Recursive Solutions in Stochastic Models: Equilibrium and Transient Analysis," **Technical Rept. no.29, Purdue University, May 1989.**

[RaR1] G. Rama Murthy and Alexander Rumyantsev," On an exact solution of the rate matrix of Quasi-Birth-Death process with small number of phases" **Proceedings of 31st European Conference on Modelling and Simulation (ECMS 2017), Budapest, Hungary .**

[RaR2] G. Rama Murthy and Alexander Rumyantsev," On an exact solution of the rate matrix of G/M/1-type Markov Process with small number of phases," Accepted for Publication in **Journal of Parallel and Distributed Computing, Elsevier Science, Vol. 119, pp. 172-178, September 2018**

[Sto] A. J. Stothers,"On the Complexity of Matrix Multiplication," **Ph. D Thesis, University of Edinburgh, 2010**

[Str1] V. Strassen,"Gaussian Elimination is not optimal," **Numer. Math, 13, pp.354-356, 1969**

[Str2] V. Strassen,"The asymptotic spectrum of tensors and the exponent of matrix multiplication," **In FOCS, pp.49-54, 1986**

[Wil] Williams, Virginia Vassilevska (2012-05-19). "[Multiplying matrices faster than coppersmith-winograd](#)". ACM: 887–898. doi:[10.1145/2213977.2214056](#). ISBN [9781450312455](#).

[ZhC] J. Zhang and E. J. Coyle, "Transient Analysis of Quasi-Birth-and-Death Processes," **Communications in Statistics—Stochastic Models, Vol. 5, No. 3, August-September 1989**