Concentrated matrix-exponential distributions and their applications

Miklós Telek[1,2]

[1]Budapest University of Technology and Economics, Hungary
[2]MTA-BME Information Systems Research Group, Hungary

Joint work with Gábor Horváth, Illés Horváth, András Mészáros, Salah Al-Deen Almousa, Nail Akar, Omer Gursoy, ...

Mathematical modelling seminar
Oct. 27, 2020

Outline

Deterministic delay is hard to represent in Markovian models.

Minimal SCV of phase type distribution: $1/N$
Minimal SCV of matrix exponential distribution: $\mathcal{O}(1/N^2)$

How to obtain them?
What to use them for?

Matrix exponential (ME) distributions

- Distributions with density function

$$f(x) = \alpha e^{\boldsymbol{A}x}(-\boldsymbol{A})\mathbb{1},$$

  where α is a row vector, $\boldsymbol{A}$ is a square matrix and $\mathbb{1}$ is the column vector of ones
  (of size $1 \times N$, $N \times N$ and $N \times 1$, respectively)
- *No sign constraints on α and $\boldsymbol{A}$*
  $\rightarrow$ *lack of stochastic interpretation*.
- It is hard to check (in general) if $f(x)$ is non-negative.

Matrix exponential (ME) distributions

- The kth moment of the ME with $f(x) = \alpha e^{\boldsymbol{A}x}(-\boldsymbol{A})\mathbb{1}$ is

$$\mu_k = \int_x x^k f(x) dx = k! \alpha (-\boldsymbol{A})^{-k} \mathbb{1},$$

- and its squared coefficient of variation (SCV) is

$$SCV = \frac{\mu_0 \mu_2}{\mu_1^2} - 1.$$

The SCV is insensitive to multiplication and scaling, i.e. $SCV(f(x)) = SCV(c_1 f(c_2 x))$.

Phase type (PH) distributions

- Same as ME, but sign constraints apply for α and $\boldsymbol{A}$:
  - α is non-negative,
  - $\boldsymbol{A}$ has negative diagonal and non-negative off-diagonal elements, such that $\boldsymbol{A}\mathbb{1} \leq 0$.
- α and $\boldsymbol{A}$ can be interpreted as the initial vector and the generator of a transient continuous time Markov chain $\rightarrow$ time to absorption.
- ME distributions of order N is a superset of phase type PH distributions of order N.

Properties

- Bounded SCV (Aldous-Shepp , O'Cinneide): $SCV \geq \frac{1}{N}$.
- Equality is provided by Erlang (Gamma) distribution.

Minimal SCV of matrix exponential distributions

Since it is hard to check if $f(x) = \alpha e^{\boldsymbol{A}x}(-\boldsymbol{A})\mathbb{1}$ is non-negative. The constrained non-linear optimization problem:

$$\min_{\alpha,\boldsymbol{A}} SCV(f(x))$$

$$\text{subject to} \;\; f(x) \geq 0, \;\; \forall x \geq 0,$$

is hard to solve.

The solution of this problem is not known for $N > 2$!!

Conjectures are available based on the following workaround approach.

A non-negative subset

Workaround approach:
Search for the minimum in a special subset, which is non-negative by construction.

A promising candidate set is:

$$f^{+}(t) = ce^{-\eta t} \prod_{i=1}^{n} \cos^2\left(\frac{\omega t - \phi_i}{2}\right),$$

which is non-negative by construction.

Persisting (unproven) conjecture:
ME($2n+1$) with minimal SCV is an $f^{+}(t)$ of order n.

Simplified non-linear optimization problem

The fact that $f^+(t)$ has a size $N = 2n+1$ matrix exponential representation is due to the trigonometric – exponential relation

$$f^+(t) = c e^{-\eta t} \prod_{i=1}^{n} \cos^2\left(\frac{\omega t - \phi_i}{2}\right) = \sum_{k=0}^{2n} \eta_k e^{-\beta_k t} = \pi e^{\boldsymbol{A}t}(-\boldsymbol{A})\mathbb{1},$$

with

$$\pi = \left[\frac{\eta_1}{\beta_1}, \ldots, \frac{\eta_N}{\beta_N}\right] \text{ and } \boldsymbol{A} = -\mathbf{diag}\{\beta_1, \ldots, \beta_N\}$$

Simplified non-linear optimization problem

The non-linear optimization problem associated with $f^+(t)$ is:

$$\min_{\omega,\phi_1,\ldots,\phi_n} SCV(f^+(x)).$$

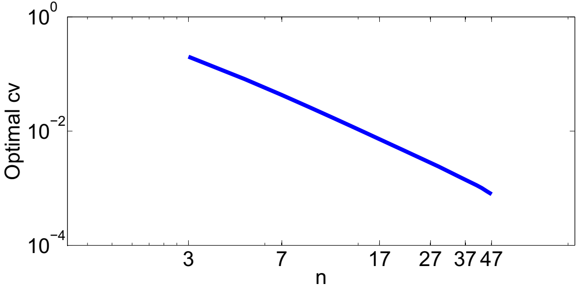
Unfortunately, it does not offer nice symbolic solution
$\rightarrow$ numerical optimization is required.

Results for CME distributions

Numerical solutions of the non-linear problem

N	SCV	$1/SCV$	date	optimization
3	0.20090	4.9776	2006	Mathematica built in
5	0.081264	12.306		
$\vdots$	$\vdots$	$\vdots$		
15	0.0093128	107.38		
17	0.0072074	138.75		
19	0.0057368	174.31	2016	Evolution strategy
21	0.0046708	214.10		
$\vdots$	$\vdots$	$\vdots$		
45	0.00088322	1132.2		
47	0.00078490	1274.0		

Concentrated ME distributions



- SCV decays with $\sim 2/N^2$ instead of $1/N$.

Numerical optimization

For efficient numerical minimization of the SCV for $N > 47$ we need accurate and efficient computation methods with low computational complexity for

i) for computation of the SCV based on the parameters,

ii) for minimization of the SCV.

Concentrated ME distributions

Main bottleneck of the optimization:

- inefficient computation of the moments based on $ce^{-\eta t}\prod_{i=1}^{n}\cos^2\left(\frac{\omega t-\phi_i}{2}\right)$.
- $\rightarrow$ efficient moments computation based on $\sum_{k=0}^{2n}\eta_k e^{-\beta_k t}$, since

$$\mu_i=\int_{t=0}^{\infty}t^i\sum_{k=0}^{2n}\eta_k e^{-\beta_k t}dt=\sum_{k=0}^{2n}\frac{i!\eta_k}{\beta_k^{i+1}}.$$

Hyper-trigonometric form of $f^+(t)$

Theorem

An exponential cosine-square function can be transformed to the following hyper-trigonometric form

$$f^+(t) = e^{-t} \prod_{i=1}^{n} \cos^2\left(\frac{\omega t - \phi_i}{2}\right) = \sum_{k=0}^{2n} \eta_k e^{-\beta_k t}$$

$$= c^{(n)} \cdot e^{-t} + e^{-t} \sum_{k=1}^{n} a_k^{(n)} \cos(k\omega t)$$

$$+ e^{-t} \sum_{k=1}^{n} b_k^{(n)} \sin(k\omega t),$$

where $c^{(n)} = \frac{1}{2} a_0^{(n)}$ and the coefficients $a_k^{(n)}$, $b_k^{(n)}$ can be calculated recursively.

Hyper-trigonometric representation of $f^+(t)$

- for $n=1$: $\quad a_0^{(1)}=1, \;\; b_0^{(1)}=0, \;\; a_1^{(1)}=\frac{1}{2}\cos\phi_1, \;\; b_1^{(1)}=\frac{1}{2}\sin\phi_1,$
- for $k>n, n\geq 1$: $\quad a_k^{(n)}=b_k^{(n)}=0,$
- for $k=0, n\geq 1$:

$$
\begin{aligned}
a_0^{(n)} &= \frac{1}{2}a_0^{(n-1)}+\frac{1}{2}a_1^{(n-1)}\cos\phi_n+\frac{1}{2}b_1^{(n-1)}\sin\phi_n,\\
b_0^{(n)} &= 0,
\end{aligned}
$$

- for $1\leq k\leq n, n\geq 2$:

$$
\begin{aligned}
a_k^{(n)} &= \frac{1}{2}a_k^{(n-1)}+\frac{1}{2}\frac{a_{k-1}^{(n-1)}+a_{k+1}^{(n-1)}}{2}\cos\phi_n+\frac{1}{2}\frac{b_{k+1}^{(n-1)}-b_{k-1}^{(n-1)}}{2}\sin\phi_n,\\
b_k^{(n)} &= \frac{1}{2}b_k^{(n-1)}+\frac{1}{2}\frac{b_{k-1}^{(n-1)}+b_{k+1}^{(n-1)}}{2}\cos\phi_n+\frac{1}{2}\frac{a_{k-1}^{(n-1)}-a_{k+1}^{(n-1)}}{2}\sin\phi_n.
\end{aligned}
$$

Hyper-trigonometric representation of $f^+(t)$

Based on the $a_k^{(n)}$, $b_k^{(n)}$ coefficients SCV is obtained efficiently.

Corollary

The $\mu_i, i = 0, 1, 2$ moments of the exponential cosine-square function are

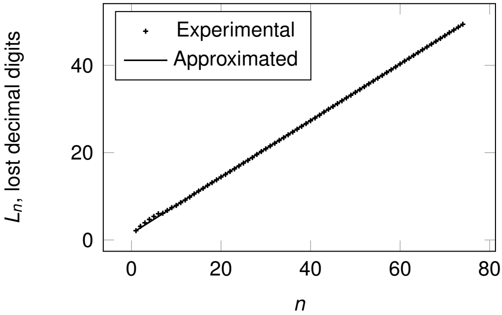
$$\mu_0 = c^{(n)} + \sum_{k=1}^{n} \frac{a_k^{(n)} + b_k^{(n)} k\omega}{1 + (k\omega)^2},$$

$$\mu_1 = c^{(n)} + \sum_{k=1}^{n} \frac{a_k^{(n)} + 2b_k^{(n)} k\omega - a_k^{(n)} (k\omega)^2}{(1 + (k\omega)^2)^2},$$

$$\mu_2 = 2c^{(n)} + \sum_{k=1}^{n} \frac{2a_k^{(n)} + 6b_k^{(n)} k\omega - 6a_k^{(n)} (k\omega)^2 - 2b_k^{(n)} (k\omega)^3}{(1 + (k\omega)^2)^3}.$$

Floting point computation of the coefficients

The computation of $a_k^{(n)}$, $b_k^{(n)}$ is prone to numerical errors



The precision loss in the computation is

$$L_n \approx 1.487 + 0.647n$$

decimal digits.

Floting point computation with the coefficients

Only the $a_k^{(n)}$, $b_k^{(n)}$ coefficients need to be computed with appropriate *high precision*,

standard double precision is sufficient for all other computations.

Optimization methods

For finding the parameters (ω, ϕ_i) which provide the minimal SCV we got success with *evolution strategies* (ES):

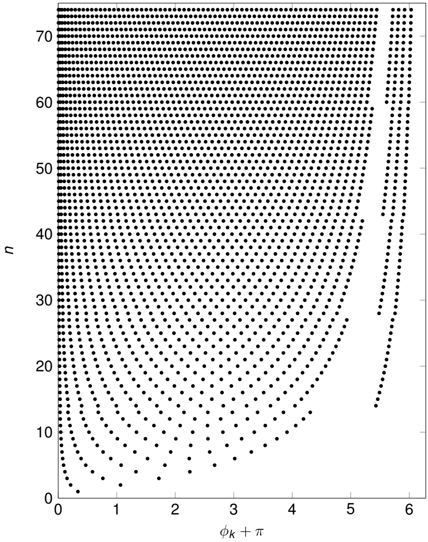
- (1+1)-ES or Rechenberg method
- CMA-ES: covariance matrix adoption ES
- BIPOP-CMA-ES: CMA-ES with smart resart policies

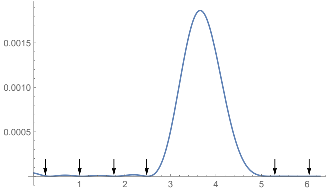
Numerical experiences:

$$T_{\text{CMA-ES}} < T_{\text{(1+1)-ES}} << T_{\text{BIPOP-CMA-ES}},$$
$$Q_{\text{CMA-ES}} \sim Q_{\text{(1+1)-ES}} < Q_{\text{BIPOP-CMA-ES}}.$$

I. Rechenberg. Evolutionsstrategien. In Simulationsmethoden in der Medizin und Biologie, pages 83-114, 1978.
N. Hansen. The CMA evolution strategy: a comparing review. In Towards a New Evolutionary Computation, pages 75-102. Springer, 2006.
N. Hansen. Benchmarking a BI-population CMA-ES on the BBOB-2009 function testbed. 11th Ann. Conf. on Genetic and Evolutionary Computation, pages 2389-2396. ACM, 2009.

Optimal parameters





Complexity as a function of n

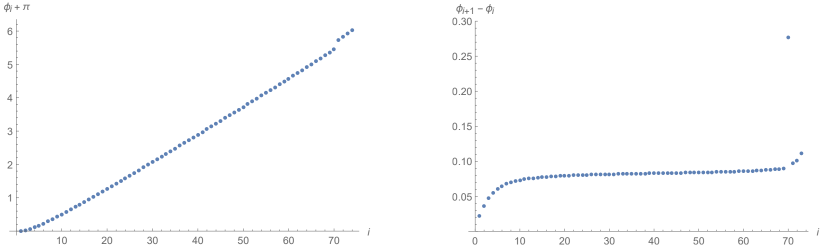
The number of parameters to optimize is $n + 1$,

- $\rightarrow$ complexity of optimization increases super linearly with n.
- $\rightarrow$ complexity of the fastest method (CMA-ES) gets prohibitive around $n = 180$.
- $\rightarrow$ low complexity suboptimal minimum is needed for $n > 180$.

Heuristic approach

$f^+(t)e^t = \prod_{i=1}^n \cos^2\left(\frac{\omega t - \phi_i}{2}\right)$ is periodic with $\phi_i \in (-\pi, \pi)$.

Motivated by the location of the optimized ϕ_i values



Main idea:

- approximate the ϕ_i parameters with a function,
- different function parameters below and above the spike.

Heuristic approach

Assuming that i^* zeros are below the main spike and $i^* < n-1$, we apply two polynomial functions to approximate the location of the ϕ_i parameters:

$$\phi_i = \begin{cases} \theta(i-\gamma)^a & \text{if } i \le i^*, \\ \Theta(i-\Gamma)^A & \text{if } i^* < i \le n, \end{cases} \tag{1}$$

where the auxiliary parameters, $\gamma, \theta, \Gamma, \Theta$, are obtained from the equations

$$\tau_1 = p_{min}, \quad \tau_{i^*} = p - w/2, \quad \tau_{i^*+1} = p + w/2, \quad \tau_{n+1} = 2\pi, \tag{2}$$

and the following constraints apply:

$$0 \le p_{min} < p - w/2 < p + w/2 < 2\pi, \quad a > 0, \quad A > 0.$$

Heuristic approach

Using (1) and (2) we get

$$\gamma = \frac{(p - w/2)^{1/a} - i^* p_{min}^{1/a}}{(p - w/2)^{1/a} - p_{min}^{1/a}} ,$$

$$\theta = p_{min} (1 - \gamma)^{-a} ,$$

$$\Gamma = \frac{(n+1)(p + w/2)^{1/A} - (i^* + 1)(2\pi)^{1/A}}{(p + w/2)^{1/A} - (2\pi)^{1/A}} ,$$

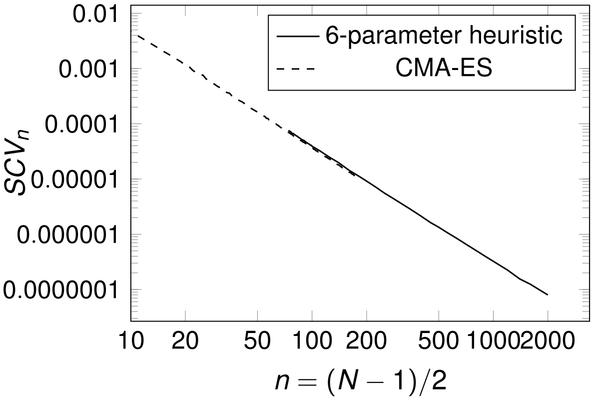
$$\Theta = 2\pi (n + 1 - \Gamma)^{-A} .$$

Heuristic approach

The heuristic procedure sets the ϕ_i parameters based on the following parameters

- i^*: the number of ϕ_i parameters left to the main peak of the function,
- p_{min}: the smallest ϕ_i parameter,
- p, w: the location of the main peak and its width,
- a, A: shape parameters defining the distribution of ϕ_i left and right to the main peak.

Heuristic approach



- The heuristic minimum follows the trend for $n > 100$.
- ME distributions with SCV$< 10^{-6}$ can be obtained.

Main messages about CME

- While the min SCV of *phase type* distribution decays *linearly* $(1/N)$, **the min SCV of matrix exponential distribution decays faster than quadratically** $(< 1/N^2)$.
- We can compute **concentrated matrix exponential distribution** up to order 10^4 **with SCV** $<$ $\mathbf{10^{-8}}$.
- The precision loss of the computation is moderate. Both representations of $f^+(t)$ can be used with **standard double precision.**

$$f^+(t) = ce^{-\eta t} \prod_{i=1}^{n} \cos^2\left(\frac{\omega t - \phi_i}{2}\right) = \sum_{k=0}^{2n} \eta_k e^{-\beta_k t},$$

The main concept

The transient analysis of some stochastic models is much harder than their stationary analysis.

In this cases:

- Extend the model with a random clock as follows
  - Run the model until the clock expire
  - When the clock expire reset the model to its initial state
  - Repeat the cycle
- Compute the *stationary distribution* of the extended model
- Obtain the transient measure from the stationary distribution.

If the clock has very low variance with mean T

$\rightarrow$ approximate transient analysis at time T.

Houdt BV, Blondia C: Approximated transient queue length and waiting time distributions via steady state analysis. *Stochastic Models* 21(2-3):725–744 (2005)

Simple application example (CTMC)

For a CTMC with generator $\boldsymbol{Q}$ and initial distribution π the transient probability vector at time T is

$$\boldsymbol{p}(T) = \pi e^{\boldsymbol{Q}T},$$

We are interested in the transient probability vector at a random time θ, that is $E(\boldsymbol{p}(\theta))$,

where the pdf of θ is

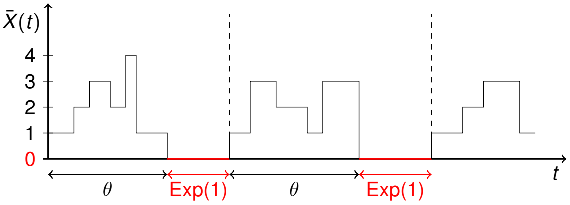
$$f(x) = \alpha e^{\boldsymbol{A}x}\boldsymbol{a},$$

with $\boldsymbol{a} = -\boldsymbol{A}\mathbb{1}$.

Simple application example (CTMC)

We can compute $E(\boldsymbol{p}(\theta))$ based on the *stationary analysis* of the extended Markov chain with generator

$$\bar{\boldsymbol{Q}} = \begin{array}{|c|c|}\hline -1 & \boldsymbol{\pi} \otimes \boldsymbol{\alpha} \\ \hline \mathbb{1} \otimes \boldsymbol{a} & \boldsymbol{Q} \oplus \boldsymbol{A} \\ \hline \end{array}\,.$$



Simple application example (CTMC)

This extended Markov chain has

- a block of size one which ”resets the CTMC to the initial state”
- a block of size $|Q| \times |A|$ which accounts for the evolution of the CTMC and the clock together.

The stationary distribution of this extended Markov chain is $[c, \gamma]$, which is the solution of

$$c(\boldsymbol{\pi} \otimes \boldsymbol{\alpha}) + \gamma(\boldsymbol{Q} \oplus \boldsymbol{A}) = \mathbf{0},$$

from which γ is

$$\gamma = -c(\boldsymbol{\pi} \otimes \boldsymbol{\alpha})(\boldsymbol{Q} \oplus \boldsymbol{A})^{-1}.$$

Simple application example (CTMC)

$E(\boldsymbol{p}(\theta))$ can be obtained by conditioning on the occurrence of a transition concluding the PH distributed θ long period according to vector $\boldsymbol{a}$, that is

$$E(\boldsymbol{p}(\theta)) =$$

$$\lim_{t\to\infty}\lim_{\delta\to 0}\frac{\text{the CTMC is in state } i \text{ and the clock expires in } (t, t+\delta)}{\text{the clock expires in } (t, t+\delta)} =$$

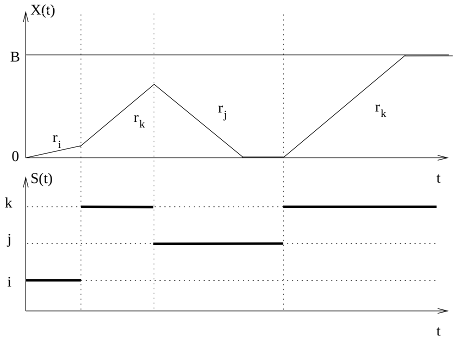
$$\frac{\gamma(\boldsymbol{I}\otimes\boldsymbol{a})}{\gamma(\mathbb{1}\otimes\boldsymbol{a})} = \frac{-(\pi\otimes\alpha)(\boldsymbol{Q}\oplus\boldsymbol{A})^{-1}(\boldsymbol{I}\otimes\boldsymbol{a})}{-(\pi\otimes\alpha)(\boldsymbol{Q}\oplus\boldsymbol{A})^{-1}(\mathbb{1}\otimes\boldsymbol{a})}.$$

If $E(\theta) = T$ and $SCV(\theta)$ is small, then $E(\boldsymbol{p}(\theta)) \approx \boldsymbol{p}(T)$.

Markov fluid queue (MFQ)

The state of a CTMC, $S(t)$, governs the evolution of a continuous variable $X(t)$ which is bounded by 0 and B

$\rightarrow$ $\{S(t), X(t)\}$ is a Markov process.



Markov fluid queue

During a sojourn of the CTMC in state i ($S(t) = i$) the fluid level ($X(t)$) increases at rate r_i when $0 < X(t) < B$:

$$\frac{d}{dt}X(t) = r_i \quad \text{if } S(t) = i, 0 < X(t) < B.$$

When $X(t) = 0$ the fluid level can not decrease:

$$\frac{d}{dt}X(t) = \max(r_i, 0) \quad \text{if } S(t) = i, X(t) = 0.$$

When $X(t) = B$ the fluid level can not increase:

$$\frac{d}{dt}X(t) = \min(r_i, 0) \quad \text{if } S(t) = i, X(t) = B.$$

Markov Fluid queue

That is

$$\frac{d}{dt}X(t) = \begin{cases} r_{S(t)}, & \text{if } 0 < X(t) < B, \\ \max(r_{S(t)}, 0), & \text{if } X(t) = 0, \\ \min(r_{S(t)}, 0), & \text{if } X(t) = B. \end{cases}$$

Model description:

- generator matrix of the CMTC $\boldsymbol{Q}$
- diagonal matrix of the fluid rates $\boldsymbol{R}$

Transient and stationary measures

$$\boldsymbol{f}^{(i,a)}(t,x) = \left[f_1^{(i,a)}(t,x) \quad f_2^{(i,a)}(t,x) \quad \cdots \quad f_n^{(i,a)}(t,x) \right],$$

where for $0 < x < B, 1 \le k \le n$

$$f_k^{(i,a)}(t,x) = \frac{d}{dx} \Pr\{X(t) \le x,\ S(t) = k \mid X(0) = a, S(0) = i\},$$

and

$$\boldsymbol{f}(x) = \lim_{t\to\infty} \boldsymbol{f}^{(i,a)}(t,x).$$

Transient analysis of MFQ

Extended fluid level dependent MFQ is defined as

$$\bar{\boldsymbol{Q}}(x) = \left[\begin{array}{c|c} 0 & 0 \\ \hline \boldsymbol{a} \otimes \mathbb{1} & \boldsymbol{A} \oplus \boldsymbol{Q}(x) \end{array}\right], \text{ if } x \neq a,$$

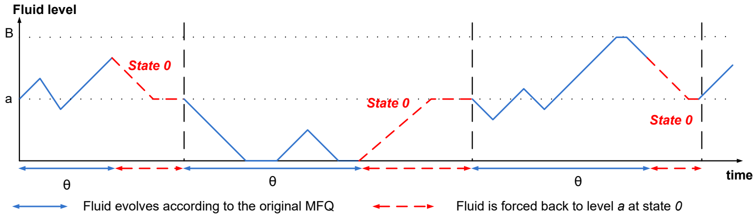
$$\bar{\boldsymbol{R}}(x) = \begin{cases} \textbf{diag}\{1, \boldsymbol{I} \otimes \boldsymbol{R}(x)\} & \text{if } x < a, \\ \textbf{diag}\{-1, \boldsymbol{I} \otimes \boldsymbol{R}(x)\} & \text{if } x > a, \end{cases}$$

The extended MFQ has an extra boundary at $x = a$ with parameters

$$\bar{\boldsymbol{Q}}(a) = \left[\begin{array}{c|c} -1 & \boldsymbol{\alpha} \otimes \boldsymbol{e_i} \\ \hline \boldsymbol{a} \otimes \mathbb{1} & \boldsymbol{A} \oplus \boldsymbol{Q}(a) \end{array}\right], \tag{3}$$

$$\bar{\boldsymbol{R}}(a) = \textbf{diag}\{0, \boldsymbol{I} \otimes \boldsymbol{R}(a)\}, \tag{4}$$

Transient analysis of MFQ



Solution method

- Stationary analysis of the extended MFQ ($\bar{\boldsymbol{Q}}(x)$, $\bar{\boldsymbol{R}}(x)$),
- Derivation of the stationary measure.

Transient analysis of MFQ

Stationary measures of this extended MFQcan be obtained by efficient numerical methods.

Let $\bar{\boldsymbol{f}}(s,x)$ and $\bar{\boldsymbol{c}}(s,x)$ by the stationary density and boundary probability:

$$\bar{f}(s,x) = \frac{d}{dx} \lim_{t\to\infty} \Pr\{\bar{X}(t) \le x, \bar{S}(t) = s\},$$
$$\bar{c}(s,x) = \lim_{t\to\infty} \Pr\{\bar{X}(t) = x, \bar{S}(t) = s\}, x \in \{0, B\},$$

for $s = 0$ or $s = (k,\ell), 1 \le k \le N, 1 \le \ell \le n$,
The vector of size $N \times n$ (for $s = (k,\ell), 1 \le k \le N, 1 \le \ell \le n$) composed by these elements are

$$\bar{\boldsymbol{f}}(x) = [\bar{f}(s,x)]\ , \quad \bar{\boldsymbol{c}}(x) = [\bar{c}(s,x)]$$

Transient analysis of MFQ

Theorem

The transient measures at clock expiration are

$$\boldsymbol{f}_{\Theta}^{a,i}(x) = \frac{\bar{\boldsymbol{f}}(x)(\boldsymbol{I} \otimes \boldsymbol{a})}{\left(\int_{y=0}^{B} \bar{\boldsymbol{f}}(y)dy + \bar{\boldsymbol{c}}(0) + \bar{\boldsymbol{c}}(B)\right)(\mathbb{1} \otimes \boldsymbol{a})},$$

$$\boldsymbol{c}_{\Theta}^{a,i}(x) = \frac{\bar{\boldsymbol{c}}(x)(\boldsymbol{I} \otimes \boldsymbol{a})}{\left(\int_{y=0}^{B} \bar{\boldsymbol{f}}(y)dy + \bar{\boldsymbol{c}}(0) + \bar{\boldsymbol{c}}(B)\right)(\mathbb{1} \otimes \boldsymbol{a})}, x \in \{0, B\}$$

Clocks with low SCV

From

$$f^+(t) = \sum_{k=1}^{N} \eta_k e^{-\beta_k t} = -\pi e^{\boldsymbol{A} t}(-\boldsymbol{A})\mathbb{1}$$

with low SCV, we have

$$\pi = [\frac{\eta_1}{\beta_1}, \dots, \frac{\eta_N}{\beta_N}] \text{ and } \boldsymbol{A} = -\textbf{diag}\{\beta_1, \dots, \beta_N\},$$

which require complex arithmetic or real representation of the conjugate complex pairs.

Laplace transformation

Laplace transform is defined as

$$h^*(s) = \int_{t=0}^{\infty} e^{-st} h(t) \mathrm{d}t. \tag{5}$$

The inverse transform problem is to find an approximate value of h at point T (i.e., $h(T)$) based on the complex function $h^*(s)$.

Assumptions

- $\int_{t=0}^{\infty} e^{-st} h(t) \mathrm{d}t$ is finite for $\mathrm{Re}(s) > 0$,
- $h^*(s)$ is not available for $\mathrm{Re}(s) \leq 0$.
- $h(t)$ is real
  $\rightarrow h^*(\bar{s}) = \bar{h}^*(s)$ and $h^*(\bar{s}) + h^*(s) = 2\mathrm{Re}(h^*(s))$.

Inverse Laplace transformation methods

There are several approaches for ILT.
Recently, the commonly used ones

- Euler,
- Talbot,
- Gaver-Stehfest,
- ...

belong to the Abate-Whitt framework.

W. Whitt J. Abate., A unified framework for numerically inverting Laplace transforms. *INFORMS Journal on Computing*, 18(4):408–421, 2006.

Abate-Whitt framework

The idea is to approximate h by a finite linear combination of the transform values, via

$$h(T) \approx h_n(T) := \sum_{k=1}^{n} \frac{\eta_k}{T} h^*\left(\frac{\beta_k}{T}\right), \; T > 0, \tag{6}$$

where the nodes β_k and weights η_k are (potentially) complex numbers, which depend on n, but not on the transform $h^*()$ or the time argument T.

Gaver-Stehfest method

Only for even n!
For $1 \leq k \leq n$

$$\beta_k = k \ln(2),$$

$$\eta_k = \ln(2)(-1)^{n/2+k} \sum_{j=\lfloor (k+1)/2 \rfloor}^{\min(k,n/2)} \frac{j^{n/2+1}}{(n/2)!} \binom{n/2}{j} \binom{2j}{j} \binom{j}{k-j},$$

where $\lfloor x \rfloor$ is the greatest integer less than or equal to x.

Euler method

Only for odd n!
For $1 \leq k \leq n$

$$\beta_k = \frac{(n-1)\ln(10)}{6} + \pi i(k-1),$$
$$\eta_k = 10^{(n-1)/6}(-1)^k \xi_k,$$

where

$$\xi_1 = \frac{1}{2}$$
$$\xi_k = 1, \quad 2 \leq k \leq (n+1)/2$$
$$\xi_n = \frac{1}{2^{(n-1)/2}}$$
$$\xi_{n-k} = \xi_{n-k+1} 2^{-(n-1)/2} \binom{(n-1)/2}{k} \text{ for } 1 \leq k < (n-1)/2.$$

Talbot method

For $2 \leq k \leq n$

$$\beta_1 = \frac{2n}{5}$$

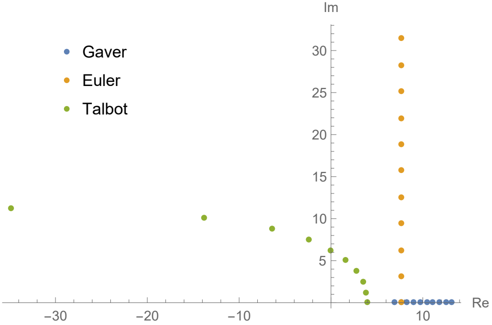
$$\beta_k = \frac{2(k-1)\pi}{5}\left(\cot\left(\frac{(k-1)\pi}{n}\right) + i\right),$$

$$\eta_1 = \frac{1}{5}e^{\beta_1}$$

$$\eta_k = \frac{2}{5}\left[1 + i\frac{(k-1)\pi}{n}\left(1+\left[\cot\left(\frac{(k-1)\pi}{n}\right)\right]^2\right) - i\cot\left(\frac{(k-1)\pi}{n}\right)\right]e^{\beta_k}.$$

Location of nodes

Location of β_k nodes on the complex plane for Gaver ($n = 10$), Euler ($n = 11$), Talbot ($n = 10$) methods



Integral interpretation

For $\mathrm{Re}(\beta_k) > 0, \forall k$, we reformulate the Abate–Whitt framework as

$$h_n(T) = \sum_{k=1}^{n} \frac{\eta_k}{T} h^*\left(\frac{\beta_k}{T}\right) = \sum_{k=1}^{n} \frac{\eta_k}{T} \int_0^\infty e^{-\frac{\beta_k}{T}t} h(t)\mathrm{d}t$$
$$= \sum_{k=1}^{n} \eta_k \int_0^\infty e^{-\beta_k t} h(tT)\mathrm{d}t = \int_0^\infty h(tT) f_n(t)\mathrm{d}t,$$

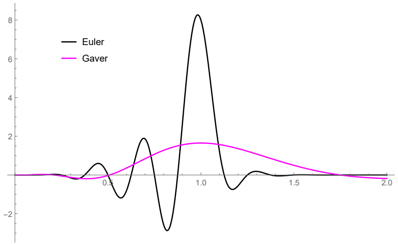
i.e. the numerical approximation of the Laplace inverse at point T is obtained as the integral of a scaled version of the original function, $h(tT)$, with

$$f_n(t) = \sum_{k=1}^{n} \eta_k e^{-\beta_k t}.$$

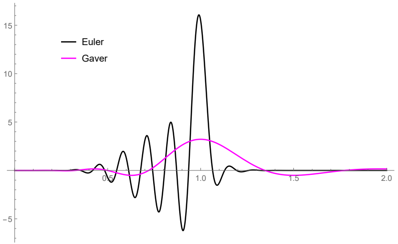
If $f_n(t)$ was the Dirac impulse function at one then the Laplace inversion would be perfect, that is $h_n(T) = h(T)$.

Properties of $f_n(t)$

But $f_n(t)$ differs from the Dirac impulse function depending on the order of the approximation (n), the applied inverse transformation method (weights η_k, nodes β_k).



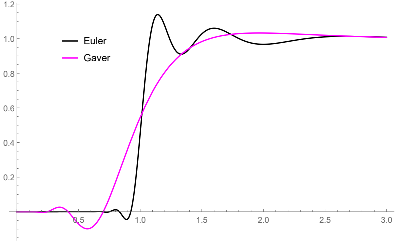
Gaver ($n = 10$), Euler ($n = 11$)



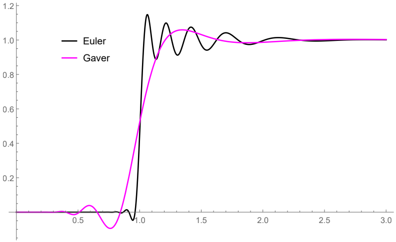
Gaver ($n = 22$), Euler ($n = 23$)

Consequence in ILT

The ILT of the unit step function at 1, $h^*(s) = \frac{e^{-s}}{s}$, with Gaver and Euler methods



Gaver ($n = 10$), Euler ($n = 11$)



Gaver ($n = 22$), Euler ($n = 23$)

Martix exponential distributions with low CV

As discussed above CME distributions can be obtained in the form:

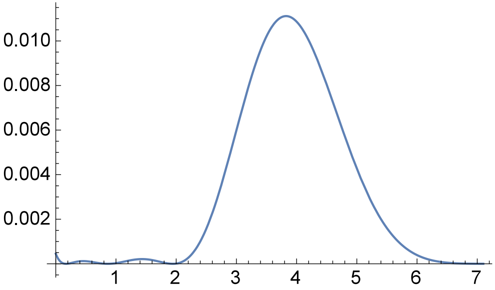
$$f_{\mathrm{ME}}(t) = c\,\mathrm{e}^{-\lambda t} \prod_{i=0}^{(N-1)/2} \cos^2(\omega t - \phi_i) = \sum_{i=1}^{N} \eta_i e^{-\beta_i t},$$

which is compatible with the form of the Abate–Whitt framework and closely approximates the unit impulse function.

Martix exponential distributions with low CV

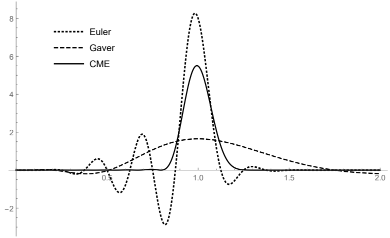
For example for $n = 4$

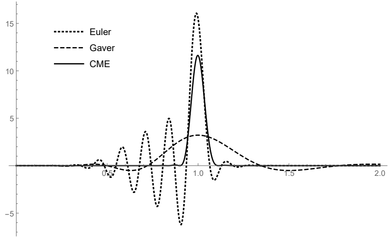
i	1	2	3	4
η_i	38.5032	$-18.9855 - 23.2984i$	$-2.70326 + 13.374i$	$2.47829 - 1.37694i$
β_i	-3.93763	$-3.93763 + 3.48448i$	$-3.93763 + 6.96896i$	$-3.93763 + 10.4534i$



CME versus Euler and Gaver

$f_n(t)$ for order 10 and 20





For the CME method $f_n(t) \geq 0$!!

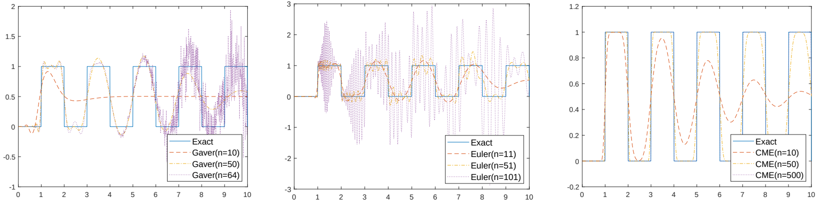
Numerical experiment

Set of test functions:

	exp	sin	heaviside	shifted exp	staircase	square wave
$h(t)$	e^{-t}	$\sin t$	$\mathbb{1}(t>1)$	$\mathbb{1}(t>1)\ e^{1-t}$	$\lfloor t \rfloor$	$\lfloor t \rfloor \mod 2$
$h^*(s)$	$\frac{1}{1+s}$	$\frac{1}{s^2+1}$	$\frac{1}{s}e^{-s}$	$\frac{e^{-s}}{1+s}$	$\frac{1}{s}\frac{1}{e^s-1}$	$\frac{1}{s}\frac{1}{e^s+1}$

Numerical experiment

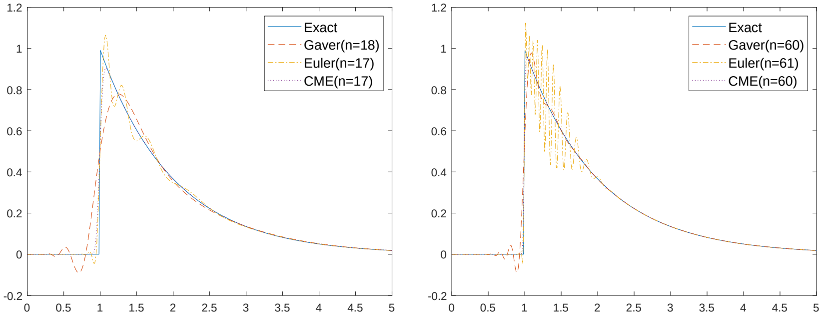
Inverse Laplace transformation of the $h(t) = \lfloor t \rfloor \mod 2$ function



No overshoot/undershoot for the CME method due to $f_n(t) \geq 0$!!

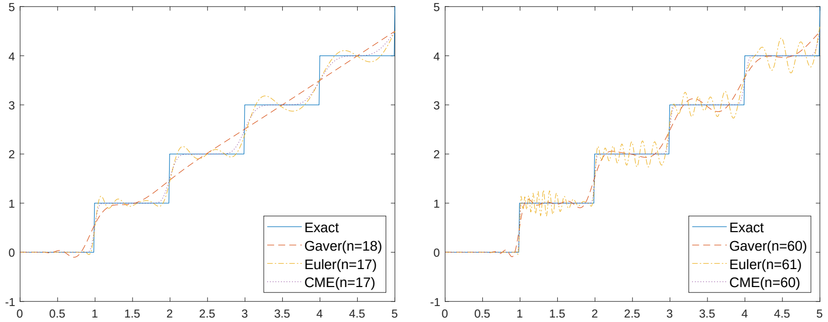
ILT of the shifted exponential function

$h(t) = \mathbb{1}(t > 1)e^{1-t}$ with order 17 and 60



ILT of the staircase function

$h(t) = \lfloor t \rfloor$ with order 17 and 60



$\|\cdot\|_1$ errors ($T=5, M=100$)

$$\|h-h_n\|_1 = \int_0^T |h(t)-h_n(t)|\mathrm{d}t \approx \frac{1}{M}\sum_{m=1}^{M}\left|h\left(\frac{mT}{M}\right)-h_n\left(\frac{mT}{M}\right)\right|.$$

	Gaver	Euler	Talbot	P–W	Laguerre	CME
order	$h(t)=\sin t$					
10	$1.34E-01$	**3.06E−04**	$1.40E-03$	$2.12E-01$	$1.65E-01$	$1.68E-02$
30	$7.37E-05$	**8.09E−11**	$2.30E-17$	$8.92E-02$	$7.61E-02$	$2.10E-03$
50	$7.29E-10$	**2.09E−17**	$2.30E-17$	$5.62E-02$	$3.43E-02$	$7.40E-04$
100	p. inf.	**2.33E−26**	$2.30E-17$	$2.92E-02$	$7.94E-02$	$1.80E-04$
500	p. inf.	p. inf.	p. inf.	p. inf.	p. inf.	**6.47E−06**
order	$h(t)=\mathbb{1}(t>1)e^{1-t}$					
10	$4.70E-02$	$2.03E-02$	p. inf.	$7.78E-02$	$8.02E-02$	**1.37E−02**
30	$1.84E-02$	$1.32E-02$	p. inf.	$4.07E-02$	$5.98E-02$	**4.45E−03**
50	$1.21E-02$	$1.80E-02$	p. inf.	$2.99E-02$	$4.93E-02$	**2.65E−03**
100	p. inf.	$9.82E-02$	p. inf.	$1.97E-02$	$3.69E-02$	**8.36E−04**
500	p. inf.	p. inf.	p. inf.	p. inf.	p. inf.	**8.69E−07**
order	$h(t)=\lfloor t\rfloor$					
10	$2.18E-01$	**1.28E−01**	p. inf.	$2.19E-01$	$8.37E+00$	$1.39E-01$
30	$1.69E-01$	$7.58E-02$	p. inf.	$2.02E-01$	p. inf.	**5.37E−02**
50	$1.18E-01$	$9.73E-02$	p. inf.	$1.89E-01$	p. inf.	**3.28E−02**
100	p. inf.	$5.24E-01$	p. inf.	$1.68E-01$	p. inf.	**1.58E−02**
500	p. inf.	p. inf.	p. inf.	p. inf.	p. inf.	**5.44E−03**
order	$h(t)=\lfloor t\rfloor \mod 2$					
10	$3.64E-01$	**1.21E−01**	$3.64E-01$	$3.88E-01$	$4.17E-01$	$1.48E-01$
30	$1.58E-01$	$8.70E-02$	$1.34E-01$	$3.11E-01$	$3.00E+01$	**5.37E−02**
50	$1.12E-01$	$9.12E-02$	$2.50E-01$	$2.62E-01$	$2.05E+01$	**3.28E−02**
100	p. inf.	$5.13E-01$	$7.81E-02$	p. inf.	p. inf.	**1.58E−02**
500	p. inf.	p. inf.	p. inf.	p. inf.	p. inf.	**5.44E−03**

Summary of NILT

We proposed a NILT method based on CME distributions.

- The CME method is a member of the Abate–Whitt framework and inherits many if its properties
  - + simple and cheap computation,
- but it differs in other properties
  - + improves with increasing order,
  - + $f_n(t)$ is non-negative $\rightarrow$ no overshooting,
  - + numerically stable
    works fine with double precision arithmetic up to $n = 1000$,
  - - no explicit expressions for the nodes β_k and weights η_k
    a priori numerical optimization & stored parameters.

Implementation and technical details:

`https://inverselaplace.org`

Summary

- Concentrated matrix exponential distributions are available with SCV$< 1/N^2$ up to $n = 10^3$ (with SCV$< 10^{-6}$).

  Gabor Horvath, Illes Horvath, and Miklos Telek. High order concentrated matrix-exponential distributions. *Stochastic Models*, doi:10.1080/15326349.2019.1702058

- Stochastic models augmented with a matrix exponential distribution can be used to compute transient measures.

  Nail Akar, Omer Gursoy, Gabor Horvath and Miklos Telek. Transient and First Passage Time Distributions of First and Second-order Multi-regime Markov Fluid Queues via ME-fication. *Methodology and Computing in Applied Probability*, doi:10.1007/s11009-020-09812-y

- Concentrated matrix exponential distributions can be efficiently used in
  - various stochastic models to represent deterministic delays,
  - numerical inverse Laplace transformation.

  Illes Horvath, Gabor Horvath, Salah Al-Deen Almousa, and Miklos Telek. Numerical inverse Laplace transformation using concentrated matrix exponential distributions. *Performance Evaluation*, doi:10.1016/j.peva.2019.102067