Concentrated Matrix Exponential distributions with real eigenvalues*

András Mészáros[1], Miklós Telek[1,2]

[1] Department of Networked Systems and Services,
Technical University of Budapest, Budapest, Hungary

[2]MTA-BME Information Systems Research Group, Budapest, Hungary

{meszarosa, telek}@hit.bme.hu

Abstract

Concentrated random variables are frequently used in representing deterministic delays in stochastic models. The squared coefficient of variation (SCV) of the most concentrated phase-type (PH) distribution of order N is $1/N$. To further reduce the SCV, concentrated matrix exponential (CME) distributions with complex eigenvalues were investigated recently. It was obtained that the SCV of an order N CME distribution can be less than $n^{-2.1}$ for odd $N = 2n + 1$ orders, and the matrix exponential distribution, which exhibits such low SCV has complex eigenvalues.

In this paper we consider CME distributions with real eigenvalues (CME-R). We present efficient numerical methods for identifying a CME-R distribution with smallest SCV for a given order n. Our investigations show that the SCV of the most concentrated CME-R of order $N = 2n+1$ is less than $n^{-1.85}$.

We also discuss how CME-R can be used for numerical inverse Laplace transformation (NILT), which is beneficial when the Laplace transform function is impossible to evaluate at complex points.

Keywords: Phase type distribution, Matrix exponential distribution, squared coefficient of variation, eigenvalues, concentrated matrix exponential distributions.

1   Introduction

Concentrated matrix exponential (CME) distributions can be effectively used to approximate distributions with low squared coefficient of variation or to approximate deterministic delay. Compared to phase-type (PH) distributions, where the squared coefficient of variation (SCV) of an order N distribution is not less

---

than $1/N$, the SCV of an order N CME with complex eigenvalues can be less than $n^{-2.1}$, for odd $N = 2n+1$ orders [9]. Recently it has been shown that, due this property, CME distributions with complex eigenvalues can be used for efficient numerical inverse Laplace transformation (NILT), such that the Laplace transform function is evaluated at complex points [10].

In this paper we focus on CME distributions with real eigenvalues only (CME-R distributions), which we also apply for NILT with real values only. This problem was already considered in [6, 3] and the importance of such NILT methods has been recognized also in the field of nuclear magnetic resonance spectroscopy [12, 13], where the Laplace domain description is available only along the real axes. For the investigation of CME-R distributions we follow a similar approach as in [9]: We consider a special subset of CME distributions whose eigenvalues are real and whose non-negativity is ensured by construction, and we derive a formula for computing the SCV efficiently.

Using this efficient computational method for the SCV based on the parameters of the considered subset of matrix exponential distributions, in the first part of the paper, we introduce two optimization procedures for finding CME-R distributions. For orders up to $n = 120$, we optimize all n model parameters, and for higher orders we formulate a heuristic optimization problem which - independently of the order - optimizes 7 parameters.

In the second part of the paper, we discuss the application of CME-R distributions for NILT. Similar to the CME (with complex eigenvalues) based NILT method [10], our method, referred to as CME-R method, also belongs to the Abate-Whitt framework [1], however, an adaptation step is required for CME-R distributions to be used for NILT according to the Abate-Whitt framework.

The main findings of the paper are as follows:

- The SCV of CME-R can be as low as $\sim n^{-1.85}$, which is only slightly higher than the SCV of CME ($\sim n^{-2.1}$) for odd $N = 2n + 1$ orders.

- The problem of NILT transformation with real arithmetic has been considered many times, e.g. [3, 12, 13]. One of the most efficient numerical methods for NILT with real arithmetic is the Gaver method [6, 14, 15]. Compared to the Gaver method, on the one hand, the CME-R method is more sensitive numerically, but on the other hand, it is Gibbs oscillation free, gradually improving with the order, monotonicity and bound preserving. I.e., when computing a probability from its Laplace transform description with the CME-R method, the result is always between zero and one, while the result computed by the Gaver method can be negative or larger than one.

The rest of the paper is organized as follows. We introduce the class of CME distributions with real eigenvalues and one of its subclass in Section 2. To find CME-R distributions with low SCV a general and a heuristic optimization problem is introduced in Section 3 and 4. A CME-R based NILT method is presented in Section 5 and analyzed in Section 6.

2   CME distributions with real eigenvalues

In this section we give a general form for matrix exponential (ME) distributions with real eigenvalues and present a subclass of matrix exponential distributions which is non-negative by constriction. We also give an efficient calculation method for the moments (and thus the SCV) of ME distributions in this subclass.

Definition 1. *Order N ME functions (referred to as ME(N)) are given by*

$$f(t) = \alpha e^{\mathbf{A}t}(-\mathbf{A})\mathbf{1}, \tag{1}$$

where α is a real row vector of dimension N, $\mathbf{A}$ is a real matrix of dimension $N \times N$, and $\mathbf{1}$ is the column vector of ones of dimension N. Furthermore α is such that $\alpha\mathbf{1} > 0$ and $f(t) \geq 0$ for $t \geq 0$.

Definition 2. *Order N ME distributions are distributions whose probability density function (pdf) is a ME function $f(t)$ such that $\int_{t=0}^{\infty} f(t)\mathrm{d}t = 1$.*

Definition 3. *Order N PH distributions are distributions whose probability density function (pdf) is a ME function $f(t)$ such that $\alpha \geq 0$, $\mathbf{A1} \leq 0$, the non diagonal elements of matrix $\mathbf{A}$ are non-negative and $\int_{t=0}^{\infty} f(t)\mathrm{d}t = 1$.*

When the ME distribution has only real eigenvalues (ME-R), the pdf in (1) can be rewritten as

$$f(t) = \sum_{i=1}^{\#\lambda} \sum_{j=1}^{\#\lambda_i} \gamma_{ij} t^{j-1} e^{-\lambda_i t}, \tag{2}$$

where $\#\lambda$ is the number of different eigenvalues, $\#\lambda_i$ is the multiplicity of eigenvalue λ_i, and $N = \sum_{i=1}^{\#\lambda} \#\lambda_i$ is the order of the distribution. In this representation the γ_{ij} coefficients are real and can be negative as well.

Let μ_i be the ith moment of the ME function, that is,

$$\mu_i = \int_{t=0}^{\infty} t^i f(t)\mathrm{d}t. \tag{3}$$

The squared coefficient of variation (SCV) of the ME function is

$$\mathrm{SCV} := \frac{\mu_2 \mu_0}{\mu_1^2} - 1. \tag{4}$$

Unfortunately, the order N ME distribution with minimal SCV is not known for $N > 2$. We refer to ME distributions with numerically minimized SCV as concentrated ME (CME) distribution. CME distributions with complex eigenvalues have been investigated recently in [9]. Here we examine CME distributions with real eigenvalues only (CME-R). Finding CME-R distributions requires the solution of the following constrained non-linear optimization problem:

$$\min_{\substack{\#\lambda \in \{1, \ldots, N\}, \, \#\lambda_i (i \in \{1, \ldots, \#\lambda\}), \\ \lambda_i (i \in \{1, \ldots, \#\lambda\}), \\ \gamma_{ij} (i \in \{1, \ldots, \#\lambda\}, j \in \{1, \ldots, \#\lambda_i\})}} SCV(f(t))$$

$$\text{subject to} \quad f(t) \geq 0, \quad \forall t \geq 0, \quad \sum_{i=1}^{\#\lambda} \#\lambda_i = N$$

where the parameters define $f(t)$ according to (2) and $SCV(f(t))$ is the SCV of $f(x)$ according to (3) and (4).

Unfortunately, it is rather difficult to check if $f(t)$ satisfies the $f(t) \geq 0$ constraint for a given set of parameters and due to that, this optimization problem is very hard. Instead of the solution of this constrained optimization problem, we use a workaround approach, similar to the one in [9]. We restrict our attention to a specific subclass of the ME-R distributions, which is non-negative for $t \geq 0$ by construction, and optimize the SCV only for that subset.

The subset of our interest, which was introduced already in [5], has the form

$$f_n(t) = c \prod_{i=1}^n (\lambda t - \tau_i)^2 \, e^{-\lambda t} \quad \text{with } N = 2n + 1. \tag{5}$$

Using this subset, the optimization problem simplifies to

$$\min_{\tau_1, \ldots \tau_n} SCV(f_n(t)), \tag{6}$$

where the parameters define $f_n(t)$ according to (5) and $f_n(t)$ defines the SCV according to (4). We note that the SCV is insensitive to multiplication and scaling, i.e., $SCV(c_1 f_n(tc_2)) = SCV(f_n(t))$, consequently the optimization problem is independent of c and λ. To simplify the discussion about the optimization of the SCV, for the rest of this section we assume $\lambda = c = 1$.

Theorem 1. *The* SCV *of* $f_n(t) = \prod_{i=1}^n (t - \tau_i)^2 \, e^{-t}$ *is*

$$SCV = \frac{\left(\sum_{i=0}^{2n} a_i(i+2)!\right)\left(\sum_{i=0}^{2n} a_i i!\right)}{\left(\sum_{i=0}^{2n} a_i(i+1)!\right)^2} - 1, \tag{7}$$

where the a_i coefficients are the coefficients of the order $2n$ polynomial

$$\prod_{i=1}^n (t - \tau_i)^2 = a_{2n} t^{2n} + a_{2n-1} t^{2n-1} + \ldots + a_1 t + a_0. \tag{8}$$

Proof. Using that $\int_{t=0}^{\infty} t^k e^{-t} dt = k!$ for $k = \{0, 1, \ldots\}$, we have

$$\mu_k = \int_{t=0}^{\infty} t^k f_n(t) dt = \int_{t=0}^{\infty} t^k \sum_{i=0}^{2n} a_i t^i e^{-t} dt = \sum_{i=0}^{2n} a_i(i+k)! \, , \tag{9}$$

from which

$$SCV = \frac{\mu_2 \mu_0}{\mu_1^2} - 1 = \frac{\left(\sum_{i=0}^{2n} a_i(i+2)!\right)\left(\sum_{i=0}^{2n} a_i i!\right)}{\left(\sum_{i=0}^{2n} a_i(i+1)!\right)^2} - 1.$$

$\square$

Remark 1. *The computation of the $a_0, \ldots, a_{2n}$ coefficients from $\tau_1, \ldots, \tau_n$ according to (8) is a non-trivial task. The following procedure implements this in a computationally efficient iterative way based on the following relation*

$$\prod_{i=1}^{n}(t - \tau_i)^2 = \left(\prod_{i=1}^{n-1}(t - \tau_i)^2\right)(t^2 - 2\tau_n t + \tau_n^2).$$

Let $\mathbf{a}^{(n)}$ be the vector of the a_i coefficients, i.e., $\mathbf{a}^{(n)} = [a_0, a_1, \ldots, a_{2n}]$. Then $\mathbf{a}^{(n)}$ can be obtained by the following procedure

$\mathbf{a}^{(0)} = [1];$
$\text{For}(i = 1, i \leq n, i{+}{+})$
$\quad \mathbf{a}^{(i)} = [0, 0, \mathbf{a}^{(i-1)}] - 2\tau_i[0, \mathbf{a}^{(i-1)}, 0] + \tau_i^2[\mathbf{a}^{(i-1)}, 0, 0];$

where $[0, 0, \mathbf{a}^{(i-1)}]$ is a vector of dimension $2i + 1$ whose first two elements are zero and the next $2i - 1$ elements are from $\mathbf{a}^{(i-1)}$. $[0, \mathbf{a}^{(i-1)}, 0]$ and $[\mathbf{a}^{(i-1)}, 0, 0]$ are defined similarly.

The $f_n(t)$ function of CME-R with minimal SCV has the general structure as shown in Figure 1 for $n = 9$: Between the zeros of $f_n(t)$ (which are at $t = \tau_i$, indicated by downward arrows) there are small peaks, and there is a large main peak between two zeros, which are further apart. This structure is valid for orders $n \geq 7$. For orders $n < 7$, $f_n(t)$ with minimal SCV is such that all zeros are left of the main peak and right of the main peak $f_n(t)$ decays due to the exponentially decaying multiplier $e^{-\lambda t}$. In Figure 1, $f_n(t)$ is scaled such that $\mu_0 = \mu_1 = 1$.

3 Optimizing the τ_i parameters

In this section we discuss the numerical solution of the optimization problem in (6). This numerical optimization can be performed with various optimization methods. For low orders ($n \leq 12$), Wolfram Mathematica can solve this optimization problem in less than 5 minutes. Table 1 contains the obtained results. The 1/SCV column of the table indicates the minimal order of Phase-type distribution needed, to obtain such a low SCV, since the minimal SCV of an order N' Phase-type distribution is $1/N'$. E.g., the SCV of CME-R with order $N = 25$ is less than the SCV of any Phase-type distribution of order $N' = 84$.

For higher orders, efficient optimization tools, allowing high precision computation, are required to perform the optimization. In the numerical optimization of the parameters, we had success with *evolution strategies*. We found the covariance matrix adoption evolution strategy (CMA-ES) [7] to be the most efficient method for this optimization problem.

More precisely, even more accurate evolution strategy methods, e.g., the BIPOP-CMA-ES with restarts [8] failed to find the global optimum even for low order cases, because they distribute the τ_i points on the left and the right of the main peak of $f_n(t)$ sub-optimally.
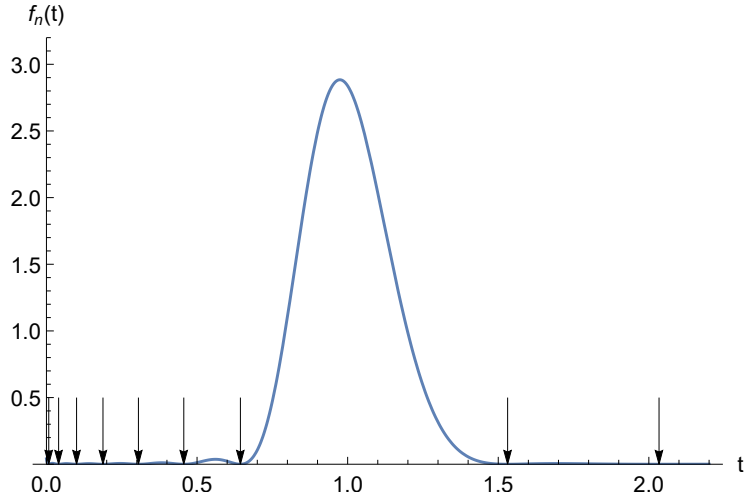
Figure 1: The peaks and zeros of $f_n(t)$ for $n = 9$ with 2 zeros on the right of the main peak

n	$N = 2n + 1$	SCV	1/SCV
1	3	0.2765	3.61
2	5	0.138	7.22
3	7	0.08	11.61
4	9	0.06	16.65
5	11	0.044	22.26
6	13	0.035	28.37
7	15	0.028	35.21
8	17	0.0228	43.81
9	19	0.0188	53.04
10	21	0.0159	62.85
11	23	0.01366	73.21
12	25	0.01189	84.06

Table 1: The SCV and its inverse for low order CME-R

6

With appropriate initial guess, CMA-ES keeps the number of points left and right to the main peak fixed and optimizes the position of the left and right points much faster than the BIPOP-CMA-ES method. Generally, the SCV optimized by the CMA-ES method decreased smoothly with increasing order. For the few SCV values that were out of trend, we repeated the optimization with different initial guess, which resulted in in-trend values in each case.

The computation time of the CMA-ES method is around one day at $n = 120$. The method provided CME-R distributions with SCV values as shown in Figure 2. In the figure, the $PH(2n+1) = 1/(2n+1) = 1/N$ curve represents the minimal SCV of PH distributions [4], while the dashed line represents the SCV of matrix exponential distributions with complex eigenvalues (CME) of order $N = 2n+1$ [9].

3.1 Required precision of the floating point arithmetic

For higher orders, the computation of the SCV based on $\tau_1, \ldots, \tau_n$ might cause numerical issues if the applied floating point arithmetic is not sufficiently accurate. Indeed the computation of the $a_0, \ldots, a_{2n}$ coefficients based on (8) and then the computation of the SCV based on (7) require high precision arithmetic.

The precision loss in computing the SCV increases with increasing order, as both the maximum value of the a_i parameters and the $i!$ factorials increases rapidly in (7), but the (unnormalized) moments in (9) increase relatively slowly, because the a_i parameters have alternating sign. Using the `Precision` function of Mathematica we made numerical investigations in order to characterize the required floating point arithmetic as a function of the n. The results are shown in Figure 3. We found that the precision loss in computing the SCV is approximately $0.95n$ digits.

That is, computing the SCV for $n = 100$ (i.e., order $N = 201$) results in 95 digits precision loss, thus to obtain the SCV in standard double precision the computation of (8) and (7) needs to be performed with extended floating point arithmetic whose precision is 95 digits larger than the one of double precision. We assume that the number of accurate digits in standard double precision numbers is approximately 20, and in the implementation of the optimization procedure, for order $2n+1$, we computed the SCV with $20+0.95n$ digit precision arithmetic.

We note that the input data of this procedure $(\tau_1, \ldots, \tau_n)$ and its results (SCV) need to be represented in only standard double precision.

4 Heuristic optimization

Since the direct applicability of CMA-ES for the optimization of all τ_i ($i = 1, \ldots, n$) parameters is limited to $n \leq 120$ due to the high number of parameters to optimize, for higher orders we needed to simplify the optimization problem.

To reduce the complexity of the optimization procedure, first we investigated the location of the optimal τ_i parameters, which are summarized in Figure 4.
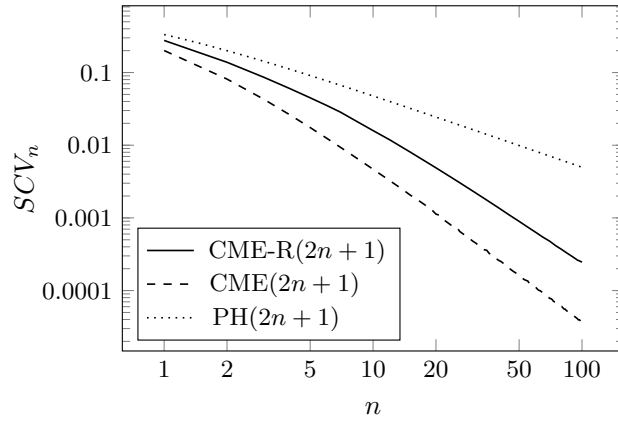
Figure 2: The minimal SCV computed by the CMA-ES method as a function of the order in log-log scale
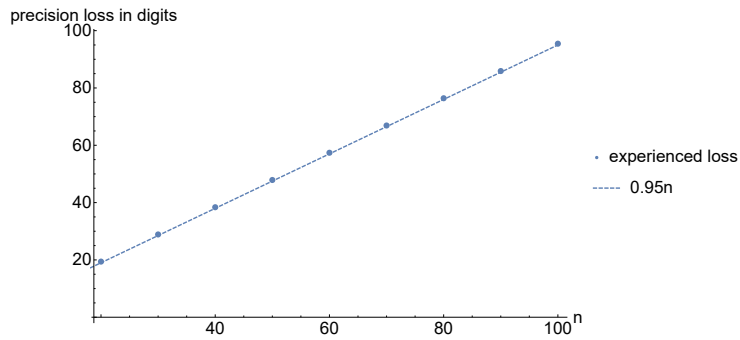

Figure 3: Precision loss during the computation of the SCV as a function of the order
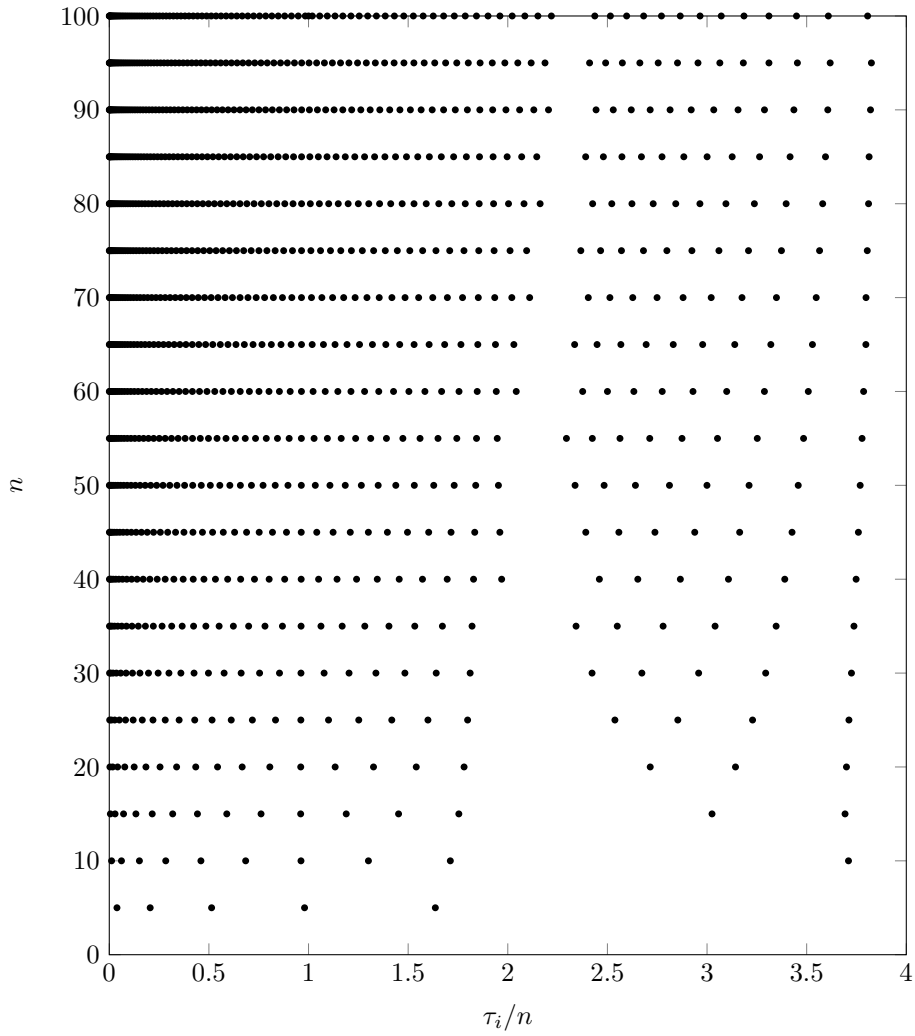
Figure 4: Results of the CMA-ES method, which optimizes n parameters for order n. The plot depicts the location of the zeros of $f_n(t)$, denoted by τ_i ($i = 1, \ldots, n$), normalized by the order for different orders up to $n = 100$.

According to the figure, the main properties of the optimal τ_i $(i = 1, \ldots, n)$ parameters are as follows:

- the τ_i parameters are located in the interval $(0, 3.8n)$,

- the largest gap between the τ_i parameters represents the main peak of the function and is located at around $2.3n$,

- the number of τ_i parameters right of the main peak is gradually increasing with n and it is in the range of $0.1n - 0.2n$,

- the τ_i parameters are dense around zero and the distance between the consecutive τ_i values increases gradually for larger values, (aside from the main gap, c.f. Figure 5).

Based on these properties of the τ_i values we developed a heuristic optimization procedure which intends to mimic the same properties based on a limited number of parameters.

We approximate the location of the τ_i parameters with two polynomial curves, one for τ_i values left to the main peak and one for right to it, which intuition is gained from Figure 5 and related numerical experiments. Thus

$$\tau_i = \begin{cases} \theta(i - \phi)^a & \text{if } i \leq i^*, \\ \Theta(i - \Phi)^A & \text{if } i^* < i \leq n, \end{cases} \tag{10}$$

where i^* denotes the number of τ_i parameters left to the main peak and a and A are the shape parameters left and right to the main peak. We could optimize the i^*, θ, Θ, ϕ, Φ, a, A parameters of the polynomial curves in (10) directly, however we transform the optimization problem to the following set of more expressive parameters: i^*, p_{min}, p_{max}, p, w, a, A, where p_{min} and p_{max} are the smallest and largest τ_i parameters, and p and w are the location of the main peak and its width.

The relation between the optimization parameters and the parameters of the polynomial curves are provided by the following equations

$$\tau_1 = p_{min}, \quad \tau_{i^*} = p - w/2, \quad \tau_{i^*+1} = p + w/2, \quad \tau_n = p_{max}, \tag{11}$$

where we apply the following constraints:

$$i^* < n - 1, \quad 0 < p_{min} < p - w/2 < p + w/2 < p_{max}, \quad a > 0, \quad A > 0.$$

Using (10) and (11) we get

$$\phi = \frac{(p - w/2)^{1/a} - i^* p_{min}^{1/a}}{(p - w/2)^{1/a} - p_{min}^{1/a}}, \quad \theta = p_{min} (1 - \phi)^{-a}, \tag{12}$$

$$\Phi = \frac{n(p + w/2)^{1/A} - (i^* + 1)p_{max}^{1/A}}{(p + w/2)^{1/A} - p_{max}^{1/A}}, \quad \Theta = p_{max} (n - \Phi)^{-A}. \tag{13}$$
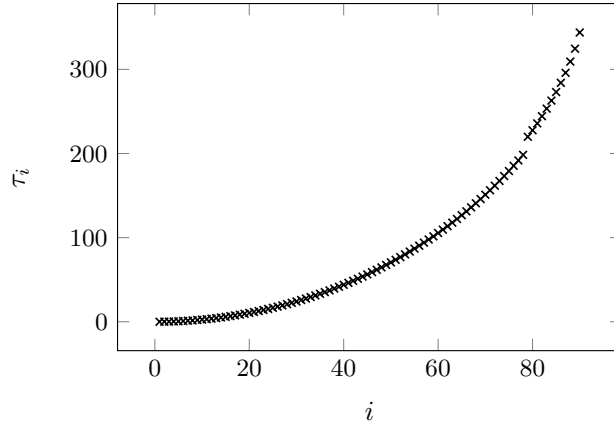
Figure 5: The optimal location of the zeros of $f_n(t)$ for $n = 90$ obtained by CMA-ES optimization method. The zeros are denoted as $\tau_1, \ldots, \tau_{90}$.
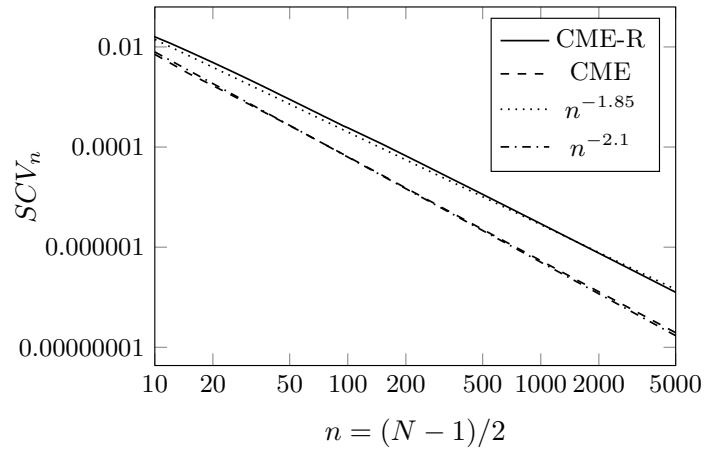

Figure 6: The minimal SCV of the CME and the CME-R distributions of order n as a function of n in log-log scale. To indicate the decay rates, the $n^{-1.85}$ and the $n^{-2.1}$ curves are plotted as well.

Applying (12) and (13), the optimization problem modifies to

$$\min_{i^*, p_{min}, p_{max}, p, w, a, A} \frac{\left(\sum_{i=0}^{2n} a_i (i+2)!\right)\left(\sum_{i=0}^{2n} a_i i!\right)}{\left(\sum_{i=0}^{2n} a_i (i+1)!\right)^2} - 1,$$

where $i^*, p_{min}, p_{max}, p, w, a, A$ defines $\tau_1, \ldots, \tau_n$ according to (12), (13) and (10) and $\tau_1, \ldots, \tau_n$ defines $a_0, \ldots, a_{2n}$ according to (8).

The number of parameters of this optimization problem is 7 for any order $N = 2n + 1$. For $n > 50$ the computational complexity of this heuristic optimization procedure is significantly less than for the CMA-ES method, when optimizing all τ_i parameters. Due to the reduced complexity of the optimization procedure we could optimize CME-R distributions for up to $n = 5000$.

For order $N = 2n+1$, we start the optimization with the following initial values of the parameters, which are good approximates of the optimal parameters according to our numerical experiences.

i^*	p_{min}	p_{max}	p	w	a	A
$\lfloor 0.85n \rfloor$	$1/n$	$3.8n$	$2.5n$	10	2.5	3

The SCV values obtained from the 7 parameter optimization procedure are depicted in Figure 6 for different orders up to $n = 5000$. In the figure, the solid line indicates the SCV values obtained by the CMA-ES optimization of the heuristic (7 parameter) procedure. To demonstrate the decay of the SCV values the $n^{-1.85}$ curve is plotted next to it. The decay of the CME-R curve drops below the $n^{-1.85}$ curve at $n = 2000$ and decays faster from that point on. Additionally, the figure contains the CME curve, which is the SCV of the concentrated matrix exponential distribution with complex eigenvalues and its supporting decay curve $n^{-2.1}$. At $n = 2000$, the SCV of the CME distribution is one order of magnitude smaller than the SCV of the CME-R distribution.

5 Numerical ILT with CME-R

In this section we discuss the application of CME-R for numerical Inverse Laplace transformation (NILT) in the Abate-Whitt framework [1]. The Laplace transform of a real or complex valued function $h(t)$ is defined as

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

The goal of NILT is to obtain $h(t)$ from $h^*(s)$. The NILT methods in the Abate-Whitt framework use the approximation

$$h(T) \approx \sum_{k=1}^{M} \frac{\eta_k}{T} h^* \left(\frac{\beta_k}{T}\right), \tag{14}$$

where η_k and β_k are real or complex numbers that depend on M, but not on the transform function h^* or the argument T. M is called the order of the approximation, and indicates the number of h^* evaluations needed to calculate the ILT.

The Abate-Whitt framework has an integral-based interpretation [10] using that

$$
\begin{aligned}
h(T) \approx \sum_{k=1}^{M} \frac{\eta_k}{T} h^* \left(\frac{\beta_k}{T} \right) &= \sum_{k=1}^{M} \frac{\eta_k}{T} \int_{t=0}^{\infty} e^{-\frac{\beta_k}{T}t} h(t) \mathrm{d}t \\
&= \sum_{k=1}^{M} \frac{\eta_k}{T} \int_{t=0}^{\infty} e^{-\frac{\beta_k}{T}t} h(t) \mathrm{d}t = \int_{t=0}^{\infty} h(tT) \sum_{k=1}^{M} \eta_k e^{-\beta_k t} \mathrm{d}t \\
&= \int_{t=0}^{\infty} h(tT) g_M(t) \mathrm{d}t,
\end{aligned}
\tag{15}
$$

where

$$
g_M(t) = \sum_{k=1}^{M} \eta_k e^{-\beta_k t}.
\tag{16}
$$

If $g_M(t)$ was the Dirac impulse at $t = 1$, then (15) would provide $h(T)$ exactly. When $g_M(t)$ is a good approximate of the Dirac impulse, we assume that (15) closely approximates $h(T)$.

Similar to [10], using the fact that $g_M(t)$ is non-negative, we measure the quality of the approximation of the Dirac impulse function with the SCV of $g_M(t)$. The SCV of the Dirac impulse is 0, therefore we assume that the lower the SCV of $g_M(t)$, the better it approximates the Dirac impulse function.

As it is discussed in the previous section, the pdf of CME-R is

$$
f_n(t) = c \prod_{i=1}^{n} (\lambda t - \tau_i)^2 \; e^{-\lambda t} = \sum_{i=0}^{2n} a_i t^i e^{-\lambda t}.
\tag{17}
$$

Assuming that c and λ are set such that $\mu_0 = \mu_1 = 1$, the above pdf is concentrated around 1. Additionally, the SCV of $f_n(t)$ can be set to be low, in which case $f_n(t)$ is a close approximate of the Dirac impulse.

To apply CME-R distributions for NILT in the Abate-Whitt framework according to (14) we still need to approximate $f_n(t)$ (defined in (17)) with $g_M(t)$ (defined in (16)) such that the η_j and β_j parameters are real.

To this end we set the β_j parameters to be equidistant around λ and set the η_j parameters such that the set of the first M coefficients of the Taylor series of $f_n(t)$ and $g_M(t)$ are identical.

That is, for $-m \le j \le m$ we choose $\beta_j = \lambda + j\delta$ and solve

$$
Taylor \left(t^i e^{-\lambda t}, 2m+1 \right) = Taylor \left(\sum_{j=-m}^{m} \nu_{ij} e^{-(\lambda+j\delta)t}, 2m+1 \right)
\tag{18}
$$

13

for the unknowns ν_{ij}. In (18), the nth coefficient of the Taylor series of $f(t)$ is $\lim_{t \to 0} \frac{d^n}{dt^n} \frac{f(t)}{n!}$ and $Taylor(f(t), k)$ denotes the vector of the $\{0, 1, \ldots, k\}$ coefficients of the Taylor series of $f(t)$.

It is worth noting that the ν_{ij} coefficients, computed according to (18), are in close connection with the *finite difference coefficients* from numerical differentiation [11], as $\nu_{ij} = \frac{c_{ij}}{(-\delta)^i}$, where c_{ij} are the aforementioned finite difference coefficients. The problem of approximating $f_n(t)$ in (17) in the form of $g_M(t)$ in (16) can be formulated using Laplace transformation and the *finite difference method* as it is shown in A.

Based on $t^i e^{-\lambda t} \approx \sum_{j=-m}^{m} \nu_{ij} e^{-(\lambda + j\delta)t}$ we have

$$f_n(t) = \sum_{i=0}^{2n} a_i t^i e^{-\lambda t} \approx \sum_{j=-m}^{m} \sum_{i=0}^{2n} a_i \nu_{ij} e^{-(\lambda + j\delta)t} = \sum_{j=-m}^{m} \eta_j e^{-\beta_j t} \triangleq f_{m,n}(t),$$
(19)

with

$$\eta_j = \sum_{i=0}^{2n} a_i \nu_{ij} = \sum_{i=0}^{2n} \frac{a_i c_{ij}}{(-\delta)^i} \quad \text{and} \quad \beta_j = \lambda + j\delta.$$
(20)

That is, $f_{m,n}(t)$ is the order m approximation of the order n CME-R distribution $f_n(t)$.

The $f_{m,n}(t)$ expression is in the form (16), which allows the application of the Abate-Whitt NILT framework according to

$$h(T) \approx h_{m,n}(T) = \sum_{j=-m}^{m} \frac{\eta_j}{T} h^* \left(\frac{\beta_j}{T} \right).$$
(21)

The order of the NILT approximation is $M = 2m + 1$, i.e., we need to evaluate $h^*(s)$ in $2m + 1$ real points to obtain $h_{m,n}(t)$.

A possible setting of parameter δ is for example $\delta = \frac{\lambda}{20m}$, which means that $\beta_j \in [0.95\lambda, 1.05\lambda]$. According to our numerical investigations $m \approx 1.3n$ gives an accurate approximation of the derivatives in this case. The effect of these parameters are discussed below, and the consideration about the numerical precision of the computations in the next section, is obtained using this setting of δ and m.

5.1 Required precision of the floating point arithmetic

Similar to Section 3.1 the precision related considerations of this section are obtained using the `Precision` function of Wolfram Mathematica.

We emphasize again that in (17) the τ_i parameters can be represented with standard double precision for $n \leq 100$, but the a_i parameters need to be represented in higher precision. In Section 3.1, we reported that for computing the SCV with approximately 20 digit precision, the required precision of a_i was

$20 + 0.95n$. Unfortunately, the alternating sign of ν_{ij} makes the computation of η_j in (20) even more sensitive numerically. To obtain an appropriate accuracy for $f_{m,n}(t)$ (such that the accuracy of the SCV computed from (17) is approximately 20 digits), the precision of the a_i parameters have to be set to $50 + 6n$ digits. This requirement was obtained based on a similar precision loss analysis as the one in Figure 3.

Starting from a_i parameters of $50 + 6n$ digit precision and ν_{ij} parameters of at least the same precision, and then computing the η_j parameters according to (20) results in a numerical accuracy of approximately $20 + 3.4m$ digits (for $m = 1.3n$ and $\delta = \lambda/20m$, as discussed below).

This $50 + 6n$ digit precision of the a_i parameters and the $20 + 3.4m$ digit precision of the η_i parameters are sufficient for NILT. Sufficient means that the precision of $\sum_{j=-m}^{m} \eta_j$ is approximately 20 digits, where the η_j parameters have large absolute value and alternating sign and $\sum_{j=-m}^{m} \eta_j = f_{m,n}(0)$ is a small positive number, since $f_{m,n}(t)$ approximates the Dirac impulse. That is, for NILT the η_j parameters can be pre-computed and stored with $20 + 3.4m$ digit precision.

The β_j parameters are very simple to obtain from λ and δ. It is a nice feature of the NILT method based on (21) that the β_j parameters and the $h^*(\beta_j/T)$ transform function values can be computed in standard double precision arithmetic and only the summation in (21) needs to be computed with high ($20 + 3.4m$ digit) precision.

In the following we summarize the steps and the respective precision requirements for NILT with CME-R. We can pre-compute the η_j and β_j parameters and store them in a file. In the NILT procedure we use these stored η_j and β_j parameters to calculate $h(T)$ from $h^*(s)$.

Steps of pre-computation:

- Starting from the c, λ, $\tau_1, \ldots, \tau_n$ parameters given in standard precision, transform the parameters to $50 + 6n$ digit precision and compute the a_i parameters based on (17) with $50 + 6n$ digit precision.

- Compute the ν_{ij} coefficients with at least $50 + 6n$ digit accuracy. (As long as δ is rational, the ν_{ij} parameters are rational numbers which can be obtained in "integer divided by integer" form without any precision loss).

- Compute the η_j parameters according to (20) with $50 + 6n$ digit accuracy.

- Store the η_j parameters with $20 + 3.4m$ digit accuracy.

- Compute the β_j parameters according to (20) with standard accuracy.

- Store the β_j parameters with standard accuracy.

NILT procedure:

- Read η_j with $20 + 3.4m$ digit accuracy.

- Read β_j with standard accuracy.

- Compute $h^* \left(\beta_j / T \right)$ with standard accuracy (if it does not incur significant (> 5 digits) precision loss due to h^*).

- Transform the resulting $h^* \left(\beta_j / T \right)$ to $20 + 3.4m$ digit accuracy representation

- Compute $\sum_{j=-m}^{m} \frac{\eta_j}{T} \, h^* \left(\frac{\beta_j}{T} \right)$ with $20 + 3.4m$ digit accuracy.

Due to the large absolute value and the alternating sign of the η_j parameters the final accuracy of the summation is going to be approximately 20 digits.

5.2 Accuracy of the NILT procedure

Approximating $h(T)$ by $h_{m,n}(T)$ contains 2 approximation errors, since $h(T) \approx h_n(T)$ and $h_n(T) \approx h_{m,n}(T)$, where $h_n(T) \triangleq \int_{t=0}^{\infty} h(tT) f_n(t) \mathrm{d}t$. The difference between $h(T)$ and $h_n(T)$ is due to the fact that $f_n(t)$ (the pdf of the chosen CME-R) differs from the Dirac impulse. We can measure the accuracy of this approximation independent of $h(T)$ using the SCV of $f_n(t)$. As discussed in Section 4, the SCV of $f_n(t)$ changes according to $n^{-1.85}$. The exact values of the SCV for $n = 1, \ldots, 12$ are given in 1 and for higher orders in Figure 6

The difference between $h_n(T)$ and $h_{m,n}(T)$ is due to the approximation of $f_n(t)$ by $f_{m,n}(t)$. Since $f_n(t)$ is such that $\int_t f_n(t) \mathrm{d}t = \int_t t f_n(t) \mathrm{d}t = 1$, we used $| \int_t t f_{m,n}(t) \mathrm{d}t - 1|$ as an error measure of the approximation and considered $| \int_t t f_{m,n}(t) \mathrm{d}t - 1| < 10^{-4}$ to be appropriate accuracy for the $h_n(T) \approx h_{m,n}(T)$ approximation. For a given n, the accuracy of the $h_n(T) \approx h_{m,n}(T)$ approximation is determined by δ and m. According to our numerical experiences, for fixed n and δ, the modification of m has two contradicting effects. When m increases, the $| \int_t t f_{m,n}(t) \mathrm{d}t - 1|$ measure of the error decreases, but the precision loss of the computation increases.

The choice of δ is also a trade-off. A lower δ value requires a lower m/n ratio, which means a more efficient NILT, because the order of the NILT is $M = 2m + 1$ (the Laplace transform function needs to be evaluated in $2m + 1$ points), while the SCV of $f_{m,n}(t)$ is characterized by n. That is, decreasing δ is beneficial for efficient NILT, but using a lower δ value also increases the numerical precision loss.

Since the quantitative effect of m and δ depends on multiple factors, we used numerical investigations to determine the highest δ value that provides the required accuracy ($| \int_t t f_{m,n}(t) \mathrm{d}t - 1| < 10^{-4}$). Table 2 summarizes the results. Our numerical investigations indicate that for a given m value the best choice of δ is the one provided in the table, because using a lower m violates the $| \int_t t f_{m,n}(t) \mathrm{d}t - 1| < 10^{-4}$ inequality and consequently the $h_n(T) \approx h_{m,n}(T)$ approximation is poor, while using larger δ increases precision loss, thus the required precision exceeds the values reported in the previous section.

Our recommended values of the parameters, $\delta = \frac{\lambda}{20m}$ and $m \approx 1.3n$, offer a potential compromise between efficiency and precision loss.

6 Numerical experiments

To investigate the properties of the proposed NILT method using CME-R, we compared it with other methods of the Abate-Whitt framework, namely the CME [10], Euler [2], and Gaver-Stehfest [14] methods. The CME method is an obvious choice, as the proposed method can be thought of as its real eigenvalue counterpart. The Euler method was chosen since, aside from the CME method, it had the best overall performance in the tests of [10]. Finally, the Gaver-Stehfest method (referred to as Gaver method in the following) was included because, unlike the CME and Euler method (and similar to the CME-R method) it can be implemented using real arithmetic. A short summary of these methods can be found in [10].

First, in Figure 7, we examine the $g_M(t)$ functions (defined in (16)) corresponding to the different NILT methods, to assess some of their characteristics. In each of the Abate-Whitt framework methods, if $g_M(t)$ approximates the Diract impulse at $t = 1$ well, then we expect the NILT method to provide a close approximation of $h(t)$. Figure 7 shows that the main peak of the Euler and Gaver methods is more concentrated than that of the CME and CME-R methods, respectively. On the other hand, contrary to the CME and CME-R methods, the $g_M(t)$ of the Euler and Gaver methods oscillate heavily and have high negative peaks, which leads to Gibbs oscillation for $h(t)$ functions with discontinuity. A more in-depth discussion of these characteristics and their consequences can be found in [10].

To compare the previously mentioned NILT methods we performed NILT using the above methods with orders $M = 30, 50, 100$ on the functions in Table 3. We considered two performance metrics: the 1-norm of the error of the NILT over the $[0, T]$ time interval, and the required numerical precision in digits for the calculations. To approximate the 1-norm of the error we used that

$$e_1 = \|h - h_n\|_1 = \int_0^T |h(t) - h_n(t)| \mathrm{d}t \approx \frac{1}{K} \sum_{k=1}^{K} \left| h\left(\frac{kT}{K}\right) - h_n\left(\frac{kT}{K}\right) \right|, \quad (22)$$

where $h(t)$ is the original time domain function and $h_n(t)$ is its NILT approximation. We used $T = 5$ for the interval and $K = 100$ points to approximate the integral.

To evaluate the required precision, we determined what is the p_{min} minimal precision which guarantees that the error of the calculation is affected only by the error of the NILT method and not by the numerical precision. To do so we performed the calculations with a sufficiently high p_{calc} initial precision, and for every NILT method and order we subtracted from it the p_{norm} precision of the 1-norm error, i.e.,

$$p_{min} = p_{calc} - p_{norm} + 2,$$

where the extra 2 digits ensure that the first two digits of the 1-norm error are accurate. The results of the NILT computations can be found in Table 4. We can see that for smooth functions (e^{-t}, $\sin t$, i.e., those without discontinuity), the

m	$2.1n$	$1.5n$	$1.35n$	$1.3n$	$1.2n$
δ	$\lambda/2m$	$\lambda/5m$	$\lambda/10m$	$\lambda/20m$	$\lambda/50m$

Table 2: The required number of points and their distances for accurate derivative approximation. The derivative approximation is assumed to be accurate when $|\int_t t f_{m,n}(t)\mathrm{d}t - 1| < 10^{-4}$
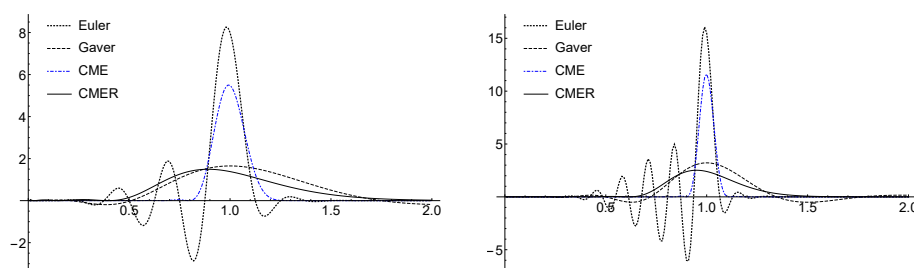

Figure 7: The $g_M(t)$ function of the CME-R method, which intends to approximating the Dirac impulse function and the associated functions of the Euler, Gaver and CME methods for order 10 and 20

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

Table 3: The set of test functions used for the numerical evaluation of the NILT methods.

Gaver and the Euler method have the lowest 1-norm error (indicated by bold-face numbers), and it also decreases rapidly with increasing order. The CME and the CME-R method are significantly worse in this regard, but if extremely high accuracy is not required, both are usable alternatives. For functions with discontinuity, the CME method gives the lowest 1-norm error, the Gaver and the CME-R method have similar performance, while the Euler method has slightly smaller error than the Gaver and CME-R methods for smaller orders, but the error increases for higher orders.

For every computation, the CME method requires the lowest precision. When using the CME method, standard double precision is always sufficient in the examined cases. The Euler method has slightly higher precision requirements. It can be applied using double precision for moderate orders - if extremely high accuracy is not required. If we want to attain high accuracy (e.g., 10^{-16} for $h(t) = e^{-t}$ with order 50 NILT), naturally, we need to apply the appropriately high precision, as shown in Table 4. Double precision is not enough in case of the Gaver method, not even for lower orders. The required precision is around $20 - 30$ digits for order 10 and $60 - 70$ digits for order 100 NILT, even if only lower accuracy is needed. The CME-R method demands the highest precision: approximately 50 digits for order 10 and as high as 170 digits for order 100 NILT.

Based on the above, the CME method is the best overall NILT method, but the Euler method can be used with lower error, if it is known that the time domain function is smooth. However, both require complex arithmetic and the ability to evaluate the Laplace domain function in complex points. If only real arithmetic is available, or the Laplace domain function cannot be evaluated in complex points, then the Gaver method is more efficient than the CME-R method, because it requires lower precision and has smaller error for smooth functions. The CME-R method should be used if the CME and Euler methods cannot be used, and we want to avoid overshoot.

7   Conclusion

In this paper we investigated concentrated matrix exponential distributions with real eigenvalues (CME-R). We gave an efficient numerical method for generating CME-R for high orders and obtained that for order $N = 2n + 1$, the squared coefficient of variation is lower than $n^{-1.85}$ (for sufficiently high orders).

Using CME-R distributions, we also proposed a numerical inverse Laplace transformation method (referred to as CME-R method). To summarize the presented numerical results, complex analysis based NILT methods are superior to the real analysis based ones. If only real analysis based NILT is available, then the Gaver method is less intensive numerically (lower precision arithmetic is required for its use than for CME-R), while due to the non-negativity of the matrix exponential function $g_M(t)$, the CME-R method is free of Gibbs oscillation, gradually improving with the order, monotonicity and bound preserving.

Our C++ implementation for CME-R optimization and Mathematica imple-

		Gaver	Euler	CME	CME-R
order		$h(t) = e^{-t}$			
30	error	**1.49E−11**	2.14E−11	9.31E−05	2.86E−03
	min. prec	31.6	18.1	10.5	50.1
50	error	**6.05E−20**	1.27E−16	2.99E−05	9.96E−04
	min. prec	53.4	26.7	11.8	85.3
100	error	**1.32E−46**	1.53E−31	7.29E−06	2.66E−04
	min. prec	113.6	50.0	13.2	169.7
order		$h(t) = \sin t$			
30	error	7.36E−05	**8.08E−11**	1.49E−03	4.18E−02
	min. prec	24.6	17.3	9.2	48.6
50	error	5.22E−10	**2.09E−17**	4.89E−04	1.57E−02
	min. prec	42.9	27.1	10.4	83.4
100	error	1.22E−28	**1.85E−31**	1.20E−04	4.42E−03
	min. prec	94.8	49.2	11.6	167.4
order		$h(t) = \mathbb{1}(t > 1)$			
30	error	1.82E−02	1.32E−02	**3.45E−03**	2.00E−02
	min. prec	21.7	9.0	8.67	48.3
50	error	1.20E−02	1.79E−02	**1.97E−03**	1.17E−02
	min. prec	34.5	11.7	9.66	81.8
100	error	5.50E−03	9.82E−02	**3.45E−04**	6.00E−03
	min. prec	66.5	17.9	11.2	162.2
order		$h(t) = \mathbb{1}(t > 1)e^{1-t}$			
30	error	1.84E−02	1.32E−02	**3.67E−03**	2.44E−02
	min. prec	21.5	8.9	8.6	48.1
50	error	1.21E−02	1.80E−02	**2.04E−03**	1.35E−02
	min. prec	34.4	11.7	9.6	81.7
100	error	5.54E−03	9.82E−02	**3.61E−04**	6.60E−03
	min. prec	66.5	17.8	11.2	162.1
order		$h(t) = \lfloor t \rfloor$			
30	error	1.69E−01	7.58E−02	**4.48E−02**	1.80E−01
	min. prec	20.7	8.5	7.8	47.4
50	error	1.18E−01	9.73E−02	**2.62E−02**	1.35E−01
	min. prec	33.5	11.3	8.9	80.8
100	error	6.86E−02	5.24E−01	**1.22E−02**	7.53E−02
	min. prec	65.4	17.5	10.0	161.1
order		$h(t) = \lfloor t \rfloor \bmod 2$			
30	error	1.58E−01	8.70E−02	**4.49E−02**	2.29E−01
	min. prec	20.7	8.5	7.8	47.2
50	error	1.12E−01	9.12E−02	**2.62E−02**	1.42E−01
	min. prec	33.5	11.3	8.9	80.8
100	error	6.93E−02	5.13E−01	**1.22E−02**	7.59E−02
	min. prec	65.4	17.5	10.0	161.1

Table 4: 1-norm error of the ILT methods for the test functions. "p. inf." (practically infinite) stands for values larger than 1000, and boldface number indicates the method with minimal 1-norm error in each case.

20

mentation for CME-R based NILT are available at `http://webspn.hit.bme.hu/~telek/tools.htm`.

References

[1] J. Abate and W. Whitt. A Unified Framework for Numerically Inverting Laplace Transforms. *INFORMS Journal on Computing*, 18(4):408–421, Fall 2006.

[2] Joseph Abate, Gagan L. Choudhury, and Ward Whitt. An introduction to numerical transform inversion and its application to probability models. In *Computational Probability*, pages 257–323. Springer US, Boston, MA, 2000.

[3] Ruben G. Airapetyan and Alexander G. Ramm. Numerical inversion of the laplace transform from the real axis. *Journal of Mathematical Analysis and Applications*, 248(2):572 – 587, 2000.

[4] D. Aldous and L. Shepp. The least variable phase type distribution is Erlang. *Stochastic Models*, 3:467–473, 1987.

[5] T Éltető, S Rácz, and M Telek. Minimal coefficient of variation of matrix exponential distributions. In *2nd Madrid Conference on Queueing Theory, Madrid, Spain (July 2006)*, 2006.

[6] D. P Gaver. Observing stochastic processes and approximate transform inversion. *Oper. Res.*, 14:444–459, 1966.

[7] Nikolaus Hansen. The CMA evolution strategy: a comparing review. In *Towards a New Evolutionary Computation*, pages 75–102. Springer, 2006.

[8] Nikolaus Hansen. Benchmarking a BI-population CMA-ES on the BBOB-2009 function testbed. In *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers*, pages 2389–2396. ACM, 2009.

[9] Gábor Horváth, Illés Horváth, and Miklós Telek. High order concentrated matrix-exponential distributions. *Stochastic Models*, 36(2):176–192, 2020.

[10] Illés Horváth, Gábor Horváth, Salah Al-Deen Almousa, and Miklós Telek. Numerical inverse laplace transformation by concentrated matrix exponential distributions. *Performance Evaluation*, 2019.

[11] H. B. Keller and V. Pereyra. Symbolic generation of finite difference formulas. *Mathematics of Computation*, 32(144):955–971, 1978.

[12] Harri Koskela, Ilkka Kilpeläinen, and Sami Heikkinen. Evaluation of protein ^{15}N relaxation times by inverse Laplace transformation. *Magnetic Resonance in Chemistry*, 42(1):61–65, 2004.

[13] Yusuke Nishiyama, Michael H. Frey, Sseziwa Mukasa, and Hiroaki Utsumi. ^{13}C solid-state NMR chromatography by magic angle spinning ^1H T_1 relaxation ordered spectroscopy. *Journal of Magnetic Resonance*, 202(2):135 – 139, 2010.

[14] H. Stehfest. Algorithm 368: Numerical Inversion of Laplace Transforms [D5]. *Commun. ACM*, 13(1):47–49, January 1970.

[15] P. Valko and J. Abate. Comparison of sequence accelerators for the Gaver method of numerical Laplace transform inversion. *Comput. Math. Appl.*, 48:629–636, March 2004.

A Finite difference method

Starting from (17), to get a form similar to (14) we use the following steps:

$$h(T) \approx h_n(T) \triangleq \int_{t=0}^{\infty} h(tT)f_n(t)\mathrm{d}t = \int_{t=0}^{\infty} h(tT) \sum_{i=0}^{2n} a_i t^i e^{-\lambda t}\mathrm{d}t$$

$$= \int_{u=0}^{\infty} h(u) \sum_{i=0}^{2n} a_i \frac{u^i}{T^i} e^{-\lambda u/T} \frac{\mathrm{d}u}{T} = \sum_{i=0}^{2n} \frac{a_i}{T^{i+1}} \int_{u=0}^{\infty} h(u)u^i e^{-\lambda u/T}\mathrm{d}u$$

$$= \sum_{i=0}^{2n} (-1)^i \frac{a_i}{T^{i+1}} \frac{d^i}{ds^i} h^*(s) \bigg|_{s=\lambda/T}, \tag{23}$$

because from $h^*(s) = \int_{u=0}^{\infty} h(u)e^{-su}\mathrm{d}u$ we obtain

$$\frac{d^i}{ds^i} h^*(s) \bigg|_{s=\lambda/T} = (-1)^i \int_{u=0}^{\infty} h(u)u^i e^{-\lambda u/T}\mathrm{d}u.$$

Compared to the form of the Abate-Whitt approximation in (14), $h_n(T)$ in (23) contains the derivatives of $h^*(s)$.

Based on the finite difference method of numerical differentiation the ith derivative of $h^*(s)$ can be obtained as

$$\frac{d^i}{ds^i} h^*(s) \bigg|_{s=\lambda} \approx \sum_{j=-m}^{m} \frac{c_{ij}}{\Delta^i} h^*(\lambda + j\Delta),$$

where $i < 2m + 1$ and the c_{ij} coefficients can be obtained as discussed in [11].

Based on this approximation of the derivatives and assuming $\Delta = \delta/T$ and

$m \geq n$ we have

$$h(T) \approx h_n(T) = \sum_{i=0}^{2n}(-1)^i \frac{a_i}{T^{i+1}} \frac{d^i}{ds^i} h^*(s)\Big|_{s=\lambda/T}$$

$$\approx h_{m,n}(T) \triangleq \sum_{i=0}^{2n}(-1)^i \frac{a_i}{T^{i+1}} \sum_{j=-m}^{m} \frac{c_{ij}}{\delta^i/T^i} h^*(\lambda/T + j\delta/T)$$

$$= \sum_{j=-m}^{m} h^*(\underbrace{(\lambda + j\delta)}_{\beta_j}/T) \underbrace{\sum_{i=0}^{2n}(-1)^i \frac{a_i c_{ij}}{\delta^i}}_{\eta_j} /T = \sum_{j=-m}^{m} \frac{\eta_j}{T} \, h^*\left(\frac{\beta_j}{T}\right). \quad (24)$$