

Enhanced optimization of high order concentrated matrix-exponential distributions*

Salah Al-Deen Almousa^a, Miklós Telek^{ab}

^aDepartment of Networked Systems and Services, Technical University of Budapest, Budapest, Hungary

^bMTA-BME Information Systems Research Group, Budapest, Hungary
almousa@hit.bme.hu, telek@hit.bme.hu

*Conference on Information Technology and Data Science
November 6–8, 2020
Debrecen, Hungary*

Abstract

This paper presents numerical methods for finding high order concentrated matrix-exponential (ME) distributions, whose squared coefficient of variation (SCV) is very low. Due to the absence of symbolic construction to obtain the most concentrated ME distributions, non-linear optimization problems are defined to obtain high order concentrated matrix-exponential (CME) distributions. The number of parameters to optimize increases with the order in the “full” version of the optimization problem. For orders, where “full” optimization is infeasible ($n > 184$), a “heuristic” optimization procedure, optimizing only 3 parameters independent of the order, was proposed in [6].

In this work we present an enhanced version of this heuristic optimization procedure, optimizing only 6 parameters independent of the order, which results in CME distributions with lower SCV than the existing 3-parameter method. The SCV gain of the new procedure compared to the old one is

*This work is partially supported by the OTKA K-123914 and the NKFIH BME NC TKP2020 projects.

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

approximately 1.66 and it is almost independent of the order. The range of the applicability of the heuristic optimization methods extends until order $n = 5000$.

To further extend the range of available CME distributions, we also propose a parameter extrapolation approach, which provides CME distributions until order $n = 20000$. The SCV of the obtained order 20000 CME distribution is $\approx 10^{-9}$.

Keywords: squared coefficient of variation, optimization, concentrated matrix exponential distributions, extrapolation.

1. Introduction

Highly concentrated matrix-exponential functions are useful in many research areas, for example, in numerical inverse Laplace transform (NILT) methods [5], as well in numerical inverse Z-transform (NIZT) methods [7]. Recently, Akar *et al.* [1], proposed the ME-fication technique, in which a concentrated matrix exponentiation distribution replaces the Erlang distribution for approximating deterministic time horizons.

Concentrated ME distributions of order N , with $N = 2n + 1^1$, are abbreviated as CME(N). CME distributions successfully constructed in [6] in the range of $N = 369, \dots, 2001$ based on a heuristic numerical optimization procedure optimizing 3 parameters independent of the order. This preliminary result indicated that the minimal SCV of CME(N) is less than $1/N^2$. The reasons for applying a heuristic approach are that there is no symbolic construction available to obtain the most concentrated ME distribution, and the full numerical optimization-based approaches (i.e., where the number of parameters to optimize is increasing with N) gets to be prohibitively complex for $N > 369$ according to [6]. In this work, we aim at improving the heuristic optimization procedure presented in [6], which we refer to as *3-parameter optimization*. The proposed enhanced optimization procedure optimizes 6 parameters (independent of the order) and we will refer to it as *6-parameter optimization* method.

The rest of the paper is organized as follows. In Section 2, we provide a brief introduction of ME distributions and discuss the definition of SCV and the optimization problem to obtain its minimum. In Section 3, we review the optimization methods proposed for SCV minimization in the literature and discuss their applicability. Section 4 introduces the proposed enhanced SCV optimization procedure with 6 parameters and Section 5 discusses its numerical properties. Section 6 presents the parameter extrapolation approach to extend the availability of CME distributions up to order $n = 20000$. Finally, Section 7 concludes the paper.

Both of these two order definitions are present in the related literature. N , “the cardinality of the describing matrix”, is more commonly used in phase type and matrix exponential distribution related literature, while n , “the number of complex conjugate eigenvalue pairs” is more commonly used in NILT related literature.

2. Matrix exponential distributions

Definition 2.1. Order N ME functions (referred to as $\text{ME}(N)$) are given by

$$f(t) = \underline{\alpha} e^{\mathbf{A}t} (-\mathbf{A}) \mathbf{1}, \quad (2.1)$$

where $\underline{\alpha}$ is a real row vector of size N , \mathbf{A} is a real matrix of size $N \times N$ and $\mathbf{1}$ is the column vector of ones of size N .

Definition 2.2. If $f(t) \geq 0, \forall t \geq 0$, and $\underline{\alpha}$ is such that $\underline{\alpha} \mathbf{1} = 1$ then $f(t)$ is the probability density function of a ME distribution of order N .

According to (2.1), vector $\underline{\alpha}$ and matrix \mathbf{A} define a matrix exponential function. We refer to the pair $(\underline{\alpha}, \mathbf{A})$ as *matrix representation* in the sequel.

An ME distribution is said to be concentrated when its squared coefficient of variation

$$SCV(f(t)) = \frac{\mu_0 \mu_2}{\mu_1^2} - 1, \quad (2.2)$$

is low. In (2.2), μ_i denotes the i th moment, defined by $\mu_i = \int_{t=0}^{\infty} t^i f(t) dt$. We note that the SCV according to (2.2) is insensitive to multiplication and scaling, i.e. $SCV(f(t)) = SCV(cf(\lambda t))$.

The optimization problem to obtain the minimal SCV of $\text{ME}(N)$ can be formulated as

$$\begin{aligned} & \min_{\underline{\alpha}, \mathbf{A}} SCV(f(t)) \\ & \text{subject to } f(t) \geq 0, \quad \forall t > 0. \end{aligned}$$

Although matrix-exponential functions have been used for many decades, there are still many questions open regarding their properties. Such an important question is how to decide efficiently if a matrix-exponential function is non-negative for $\forall t > 0$. In general, $f(t) \geq 0, \forall t > 0$ does not necessarily hold for given $(\underline{\alpha}, \mathbf{A})$ representation, and it is rather difficult to check. A potential numerical solution for checking this property is proposed in [9].

Due to the difficulty of checking the constraints of the above constrained optimization problem, its solution is an open problem currently.

3. Concentrated ME distributions

A possible way to simplify the constrained optimization problem is to search for the minimum in a special subset of $\text{ME}(N)$, which is non-negative by construction. Horváth *et al.* in [6] suggest such a subset which is characterized by

$$f(t) = cf^+(\lambda t), \quad (3.1)$$

where $f^+(t)$ is an exponential cosine-square function with order n defined as

$$f^+(t) = e^{-t} \prod_{j=1}^n \cos^2 \left(\frac{\omega t - \phi_j}{2} \right), \quad (3.2)$$

where $\omega \geq 0$ and $0 \leq \phi_j < 2\pi$ for $j \in \{1, \dots, n\}$.

In [6] the authors conjectured that the density function of the most concentrated ME distribution of order N belongs to this special class of $\text{ME}(N)$, but the validity of this conjecture is not proved even for the smallest non-obvious case, $N = 3$.

An exponential cosine-square function is a non-negative (due to its construction) matrix exponential function and [8, Appendix A] presents how to obtain the matrix representation of size $N = 2n + 1$ associated with $f^+(t)$ in (3.2). Consequently, the set of exponential cosine-square functions of order n is a special subset of $\text{ME}(N)$ (where $N = 2n + 1$).

In this paper, we make use the fact that exponential cosine-square functions can also be represented in the following hyper-exponential form [6]

$$f^+(t) = e^{-t} \prod_{j=1}^n \cos^2 \left(\frac{\omega t - \phi_j}{2} \right) = \sum_{k=0}^{2n} \eta_k e^{-\beta_k t}, \quad t \geq 0, \quad (3.3)$$

where the η_k, β_k ($k = 0, \dots, 2n$) coefficients contain complex conjugate pairs. Generally, calculating the μ_0, μ_1, μ_2 moments based on (3.2), is not an easy task due to computational complexity caused by the product of the cosine square terms. Instead calculating the μ_0, μ_1, μ_2 moments based on (3.3) is much easier since

$$\mu_i = \int_{t=0}^{\infty} t^i \sum_{k=0}^{2n} \eta_k e^{-\beta_k t} dt = \sum_{k=0}^{2n} \frac{i! \eta_k}{\beta_k^{i+1}}, \quad (3.4)$$

3.1. Full optimization of the $f^+(t)$ parameters

In the sequel, we utilize the fact that multiplication and scaling (with c and λ in (3.1)) does not effect the SCV and optimize the SCV of $f^+(t)$ instead of $f(t)$. $f^+(t)$ in (3.2) is defined by $n + 1$ parameters: the frequency ω and the zeros ϕ_j for $j = 1, \dots, n$. Unfortunately, $\text{SCV}(f^+(t))$ is not a simple function of the parameters. For a given n to find $f^+(t)$ with minimal SCV, i.e.

$$\min_{\omega, \phi_1, \dots, \phi_n} \text{SCV}(f^+(t))$$

is still a hard non-linear optimization problem, where the number of parameters to optimize is $n + 1$.

Numerical methods for the solution of this problem are discussed in [6]. The main findings reported there are that evolution strategy based optimization provided the best numerical results. The solution of the problem with the *CMA-ES*

method [4] is fast, but does not find the best optimum compared to the *BIPOP-CMA-ES* method [3], which is much slower. The applicability of the two methods are $n \leq 74$ ($N \leq 149$) in case of the BIPOP-CMA-ES method and $n \leq 184$ ($N \leq 369$) in case of the CMA-ES method. For these orders the respective the optimization procedures take several days to terminate on an average PC. The computational complexity of these procedures increases super linearly with the order n , which inhibits the application of these procedure for higher orders.

3.2. Heuristic optimization of the $f^+(t)$ parameters

To go beyond order $n = 184$, [6] proposed a sub-optimal, 3-parameter heuristic optimization procedure, that reduce the complexity of the optimization problem by reducing the number of parameters to optimize to three, independent of the order.

Figure 2 displays the location of the the ϕ_j parameters obtained by the full optimization method for $n = 74$. As it is visible in the figure, there is a gap between the ϕ_j parameters at around $p \approx 5.2$ and the size of that gap, which is the maximum value in Figure 1 is around $w \approx 0.28$. The heuristic optimization procedure proposed in [6] assumes that the ϕ_j parameters are equidistant below and above that gap. Figure 1 and 2 display how good this assumption is compared to the fully optimized ϕ_j parameters.

Since the ϕ_j parameters are located between 0 and 2π and the number of parameters are n , this assumption allows to determine the ϕ_j parameters based on p and w according to the following expression

$$\phi_j = \begin{cases} (j - 1/2)d & \text{if } j \leq i, \\ (j - 1/2)d + w & \text{if } j > i. \end{cases} \quad (3.5)$$

where

$$d = \frac{2\pi - w}{n}, \quad i = \left\lfloor \frac{p - w/2}{d} + \frac{1}{2} \right\rfloor. \quad (3.6)$$

With the use of (3.5), $f^+(t)$ is defined by the parameters ω , p and w and the related optimization problem is

$$\begin{aligned} & \min_{\omega, p, w} SCV(f^+(t)), \\ & \text{subject to: } 0 < p - w/2 < p + w/2 < 2\pi, \omega > 0. \end{aligned}$$

The solution of this optimization problem is computed by the CMA-ES method up to $n = 1000$ in [6]. Moreover, in this work we expand this solution up to $n = 5000$.

4. Enhanced heuristic optimization of $SCV(f^+(t))$ with 6 parameters

Figure 1 and 2 suggests that the equidistant location of the ϕ_j parameters according to (3.5) is not flexible enough to obtain similar low SCV as obtained with the full optimization method. Starting from this assumption, we try to locate the ϕ_j parameters in a more flexible way. To this end we introduce two different power functions below and above the gap of the ϕ_j parameters as follows

$$\phi_j(a_1, b_1, a_2, b_2, i, \gamma, \delta) = \begin{cases} a_1 + b_1 j^\gamma & \text{for } 1 \leq j \leq i, \\ a_2 + b_2 j^\delta & \text{for } i + 1 \leq j \leq n. \end{cases} \quad (4.1)$$

The auxiliary parameters, a_1 , b_1 , a_2 , b_2 , can be transformed to a set of more expressive parameters based on the following relations

$$\phi_1 = 0, \quad \phi_i = p - w/2, \quad \phi_{i+1} = p + w/2, \quad \phi_{n+1} = 2\pi. \quad (4.2)$$

Substituting these relations into (4.1) results in the following function for the ϕ_j parameters

$$\phi_j(p, w, i, \gamma, \delta) = \begin{cases} \frac{(p-w/2)j^\gamma}{(i^\gamma-1)} - \frac{p-w/2}{(i^\gamma-1)} & \text{for } 1 \leq j \leq i, \\ 2\pi - \frac{(j^\delta - (n+1)^\delta)(2\pi - p - w/2)}{((i+1)^\delta - (n+1)^\delta)} & \text{for } i + 1 \leq j \leq n. \end{cases} \quad (4.3)$$

In (4.3) the parameters are constrained by

$$0 < p - w/2 < p + w/2 < 2\pi, \quad \gamma > 0, \quad \delta > 0.$$

The intuitive meaning of the parameters in (4.3) are as follows: the meaning of p , w , ω , and i are the same as in the 3-parameter optimization method, i.e.

- i : is the number of ϕ_j parameters left to the gap,
- p , w : are the midpoint of the gap and its width,

while γ and δ are shape parameters defining the power series of the ϕ_j parameters below and above the gap.

Based on (4.3), which defines the ϕ_j parameters based on 5 parameters, the optimization of $SCV(f^+(t))$ for a given order n is the following 6-parameter optimization problem

$$\begin{aligned} & \min_{\omega, p, w, i, \gamma, \delta} SCV(f^+(t)) \\ & \text{subject to: } 0 < p - w/2 < p + w/2 < 2\pi, \gamma > 0, \delta > 0, \omega > 0, \end{aligned}$$

where p , w , i , γ , δ define the ϕ_j parameters according to (4.3) and the ϕ_j parameters and ω define $f^+(t)$ according to (3.2).

To solve this optimization problem we propose to compute the SCV according to Algorithm 1 and obtain the optimum by the CMA-ES method using Algorithm

1 as the objective function. The procedure to obtain η_i , β_i and the required high precision arithmetic are detailed in [6]. Here we only recall that all computations can be performed with standard double precision arithmetic except the ones indicated to be “high precision”. In those cases, to obtain results in 16 digits precision, the required numerical precision is $0.647n + 17.478$ digits for $ME(2n + 1)$.

Algorithm 1 The objective function of the heuristic method

```

1: procedure COMPUTESCV( $\omega, p, w, i, \gamma, \delta$ )
2:   Obtain  $\phi_j$  for  $j \in \{1, \dots, n\}$  by (4.3)
3:   Compute  $\eta_i, \beta_i$  (high precision) by (3.3)
4:   Compute  $\mu_i$  (high precision) by (3.4)
5:   Compute  $SCV$  by (2.2)
6:   return  $SCV$ 
7: end procedure

```

5. Numerical properties

The behaviour of the ϕ_j parameters obtained by the proposed 6-parameter heuristic method is also depicted in Figure 1 and 2. The figures suggest, that the ϕ_j parameters obtained by the 6-parameter optimization method better approximate the behaviour of the ϕ_j parameters obtained by full optimization than the ones of the 3-parameter method.

Moreover, Figure 2 displays how the distribution of ϕ_j locations are influenced by the shaping parameters γ , δ , and getting closer (compared to the 3-parameter case) to the fully optimized ones. This improvement in the positioning of the ϕ_j parameters leads to a significant SCV reduction compared to the 3-parameter method as illustrated in Figure 3. In the depicted range the gain (the ratio of the SCV obtained by the two methods) is approximately 1.66 and it is almost independent of the order. The proposed heuristic optimization resulted in almost the same SCV values as the ones obtained by the full optimization method in the range where full optimization is feasible and beyond that order ($n > 184$) the 6-parameter optimization results seem to follow the same decay trend.

We believe with some confidence in the possibility of expanding the heuristic optimization for orders larger than $n = 5000$, using a more powerful computing device.

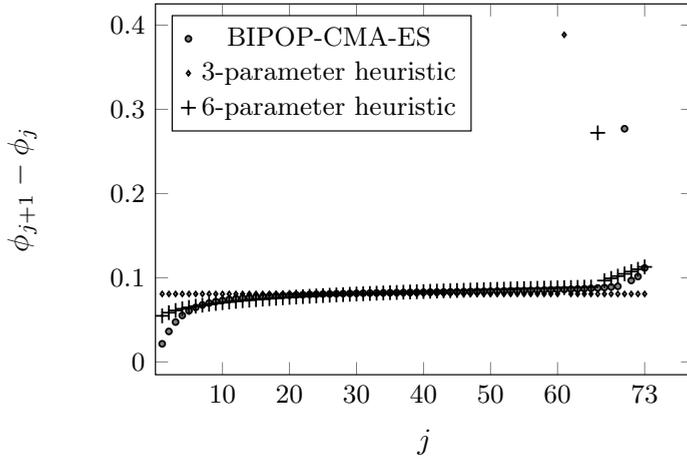


Figure 1: Difference of consecutive ϕ_j values obtained by full optimization, 3-parameter optimization and the proposed 6-parameter optimization for order $n = 74$

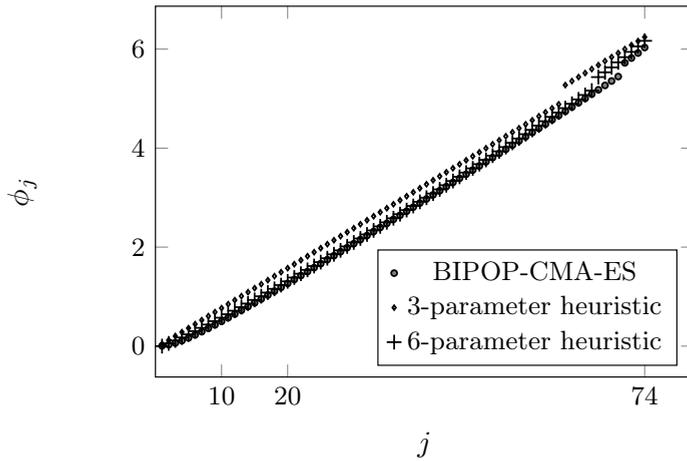


Figure 2: The location of the ϕ_j parameters obtained by full optimization, 3-parameter optimization and the proposed 6-parameter optimization for order $n = 74$

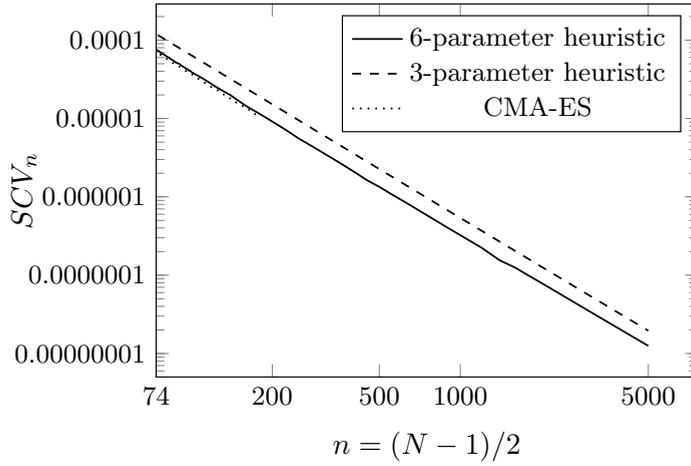


Figure 3: The minimal SCV values obtained by full optimization, 3-parameter optimization and the proposed 6-parameter optimization as a function of order n in log-log scale

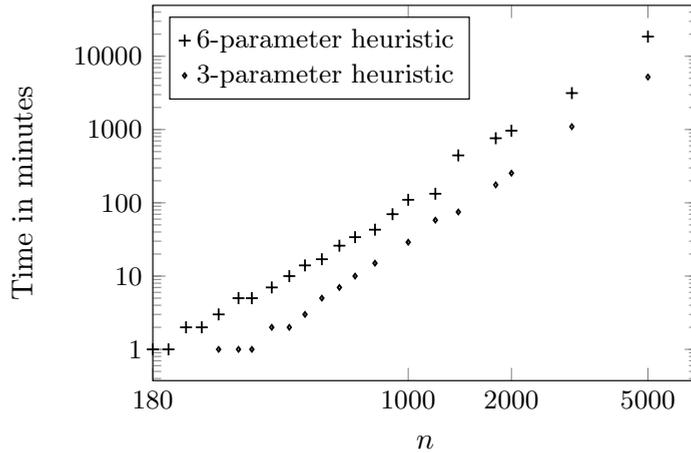


Figure 4: Running time of the heuristic parameter optimization procedures for different orders in log-log scale

Figure 4 depicts the running time of the heuristic optimization procedure on an average PC clocked at 2.9 GHz as a function of the order n .

6. Extrapolation of the parameters of heuristic optimization

The high computational costs of the heuristic optimization methods, plotted in Figure 4, inhibits their application for orders higher than $n = 5000$. In this section, we intend to obtain CME distributions for orders $n > 5000$ by extrapolating parameters of the heuristic optimization procedures.

Let $\mathbf{v}(n)$ denote the parameter values obtained from the heuristic optimization method for order n , that is, for the 3-parameter method $\mathbf{v}(n) = \{\omega(n), p(n), w(n)\}$ and for the 6-parameter method $\mathbf{v}(n) = \{\omega(n), p(n), w(n), i(n), \gamma(n), \delta(n)\}$. Furthermore, let $\mathcal{N} = \{n_1, n_2, \dots, n_K\}$ be the set of K orders for which the parameter is available (the heuristic optimization is performed) $n_K = 5000$ in our case and the other evaluated orders are visible in Figure 4.

6.1. Extrapolation methods

To extrapolate the $\mathbf{v}(n)$ vector for $n > 5000$ we considered the following extrapolation approaches.

- Element-wise extrapolation of $\mathbf{v}(n)$
In this set of methods the elements of $\mathbf{v}(n)$ are extrapolated independent of each others.

- Polynomial extrapolation ($k + 1$ parameters):

$$\hat{v}_i(n) = a_i + b_i n + c_i n^2 + \dots + z_i n^k, \quad (6.1)$$

where $v_i(n)$ is the i th element of $\mathbf{v}(n)$ and $i \in \{1, 2, 3\}$ in case of the 3-parameter method and $i \in \{1, 2, \dots, 6\}$ in case of the 6-parameter method.

- Power function extrapolation (3 parameters):

$$\hat{v}_i(n) = a_i n^{b_i} + c_i. \quad (6.2)$$

- Exponential extrapolation (3 parameters):

$$\hat{v}_i(n) = a_i e^{b_i n} + c_i. \quad (6.3)$$

In the element-wise extrapolation, we apply the following distance measure for $\hat{v}_i(n)$

$$\mathcal{D}_i = \frac{1}{K} \sum_{n \in \mathcal{N}} |v_i(n) - \hat{v}_i(n)|. \quad (6.4)$$

That is, in power function and exponential extrapolation, the optimal extrapolation parameters are obtained as

$$\{a_i^*, b_i^*, c_i^*\} = \arg \min_{\{a_i, b_i, c_i\}} \mathcal{D}_i, \quad (6.5)$$

and in polynomial extrapolation, the $\{a_i, b_i, \dots, z_i\}$ parameters are obtained similarly.

- Vector-wise extrapolation of $\mathbf{v}(n)$
 - Vector polynomial extrapolation ($(m+k)m$ parameters):

$$\hat{\mathbf{v}}(n) = (\mathbf{a} + \mathbf{b}n + \mathbf{c}n^2 + \dots + \mathbf{z}n^k) \mathbf{G}, \quad (6.6)$$

where each row sum of \mathbf{G} is one (and this way \mathbf{G} contains $(m-1)m$ free parameters), m is the number of elements of \mathbf{v} . $m = 3$ or 6 depending on the applied heuristic optimization.

- Matrix power function extrapolation ($(m+2)m$ parameters):

$$\hat{\mathbf{v}}(n) = \mathbf{a} \text{Diag}\langle n^{b_1}, \dots, n^{b_m} \rangle \mathbf{G} + \mathbf{c}. \quad (6.7)$$

- Matrix exponential extrapolation ($(m+2)m$ parameters):

$$\hat{\mathbf{v}}(n) = \mathbf{a} \text{Diag}\langle e^{b_1 n}, \dots, e^{b_m n} \rangle \mathbf{G} + \mathbf{c}. \quad (6.8)$$

In case of vector-wise extrapolation we apply the L^2 vector norm as the distance measure

$$\mathcal{D} = \frac{1}{K} \sum_{n \in \mathcal{N}} \|\mathbf{v}(n) - \hat{\mathbf{v}}(n)\|_2 = \frac{1}{K} \sum_{n \in \mathcal{N}} \sqrt{\sum_{i=1}^m (\mathbf{v}_i(n) - \hat{\mathbf{v}}_i(n))^2}. \quad (6.9)$$

That is, in Matrix power and Matrix exponential extrapolation, the optimal parameters are obtained as

$$\{\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*, \mathbf{G}^*\} = \arg \min_{\{\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{G}\}} \mathcal{D} \quad (6.10)$$

and the Matrix polynomial case is optimized similarly according to its parameters.

We note that the results obtained by any of these methods are sensitive for \mathcal{N} , the set of orders which are considered in the parameter estimations. That is, different extrapolation parameters are obtained by the same extrapolation procedure for different \mathcal{N} sets. Generally, we used the optimization results between orders 400 and 5000, that is, $400 \leq n \leq 5000$ for $\forall n \in \mathcal{N}$.

The goodness of an extrapolation approach can be judged by computing the SCV obtained from the extrapolated parameters and checking if the trend of decay for the given order $n > 5000$ follows the trend obtained by the heuristic method for order $n \leq 5000$ and plotted in Figure 3. Based on this goodness measure, we found all extrapolation approaches inappropriate except the element-wise power function extrapolation for all parameters, whose results are presented in the following subsection.

6.2. Element-wise power function extrapolation

Below, we present the results of the element-wise power function extrapolation method which we obtained by the *CF tool* of Matlab Curve Fitting Toolbox [2].

6.2.1. Extrapolation for the 3-parameter heuristic method

As discussed in [6] and in subsection 3.2, the 3-parameter heuristic optimization procedure minimizes the SCV as a function of ω , p , w . The procedure can be applied with reasonable computation time (c.f. Figure 4) up to order $n = 5000$. Beyond this limit we apply the element-wise power function extrapolation according to (6.2), (6.4), and (6.5).

Table 1 summarizes the results for all the three parameters and Figure 5 demonstrate the quality of the obtained result for the ω parameter.

	a_i^*	b_i^*	c_i^*	\mathcal{D}_i
$\hat{w}(n)$	25.03	-1.017	0	2.045E-06
$\hat{p}(n)$	-2.691	-0.2467	6.029	3.876E-03
$\hat{\omega}(n)$	0.8919	-0.2399	0.1737	1.377E-05

Table 1: The optimal extrapolation parameters for ω , p and w

Based on the extrapolation parameters in Table 1 and the associated extrapolation model in (6.2) we can extrapolate the $\omega(n)$, $p(n)$, $w(n)$ for orders larger than 5000. Using those extrapolated $\hat{\omega}(n)$, $\hat{p}(n)$, $\hat{w}(n)$ values, Figure 6 and Table 2 present the associated SCV as a function of the order up to $n = 20000$. Figure 6 and Table 2 indicate that the SCV values obtained by the extrapolation method follow the same decay trend of the heuristic optimization. For orders less than 5000, Table 2 also compares the $\omega(n)$, $p(n)$, $w(n)$ values obtained from the 3-parameter heuristic method, and the $\hat{\omega}(n)$, $\hat{p}(n)$, $\hat{w}(n)$ values provided by the power function extrapolation method. For those orders the SCV value computed by the $\omega(n)$, $p(n)$, $w(n)$ and the $\hat{\omega}(n)$, $\hat{p}(n)$, $\hat{w}(n)$ parameters are identical in their first 3 digits.

Based on Figure 6 and Table 2 we conclude that the extrapolation of the $\hat{\omega}(n)$, $\hat{p}(n)$, $\hat{w}(n)$ parameters with the element-wise power function extrapolation provide fairly concentrated matrix exponential distributions up to order 20000, whose SCV follows the same decay trend for as the one of the 3-parameter heuristic method up to order 5000.

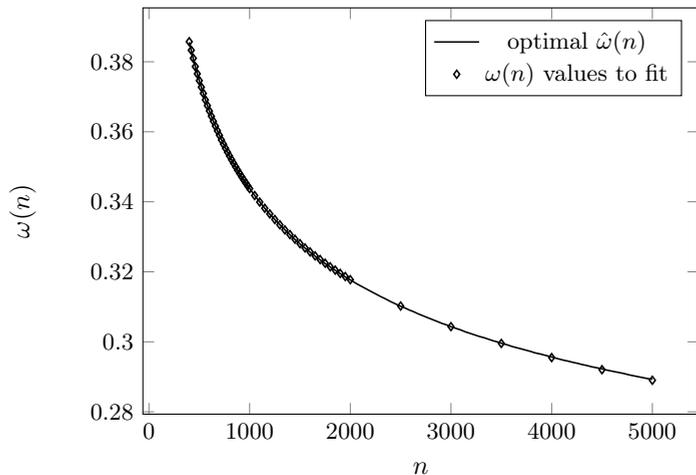


Figure 5: Curve fitting for $\omega(n)$ using a power function according to (6.2)

n	Heuristic Optimization				Extrapolation			
	$w(n)$	$p(n)$	$\omega(n)$	SCV	$\hat{w}(n)$	$\hat{p}(n)$	$\hat{\omega}(n)$	SCV
400	0.0563612	5.4162	0.385725	3.5945E-06	0.0565153	5.41526	0.3855761	3.5947992E-06
800	0.0278559	5.51193	0.353088	8.53737E-07	0.0279266	5.51173	0.3531175	8.538201E-07
1200	.0184659	5.56361	0.336537	3.69091E-07	0.0184899	5.56096	0.3364873	3.691452E-07
1500	0.014703	5.58534	0.328039	2.32831E-07	0.014735	5.58603	0.328002	2.32849E-07
2000	0.0109708	5.61548	0.317745	1.28656E-07	0.010997	5.61638	0.317712	1.28668E-07
2500	0.00874565	5.63895	0.310221	8.12596E-08	0.008765	5.63848	0.310205	8.12665E-08
3000	0.0072661	5.65621	0.304338	5.58463E-08	0.007281	5.65565	0.304363	5.58511E-08
3500	0.00621258	5.66986	0.299541	4.06814E-08	0.006225	5.66958	0.299619	4.06848E-08
4000	0.00542217	5.68131	0.295500	3.09232E-08	0.005434	5.68123	0.295650	3.09269E-08
4500	0.0048099	5.69091	0.292034	2.42819E-08	0.004821	5.69119	0.292252	2.42852E-08
5000	0.00432189	5.69975	0.289008	1.95615E-08	0.004331	5.69986	0.289293	1.9564E-08
10000	-	-	-	-	0.002140	5.75159	0.271585	4.73132E-09
15000	-	-	-	-	0.001417	5.77800	0.262512	2.06641E-09
20000	-	-	-	-	0.001057	5.79519	0.256589	1.14904E-09

Table 2: Original and extrapolated parameters and the associated SCV for the 3-parameter heuristic optimization

6.2.2. Extrapolation for the 6-parameter heuristic method

We applied the same extrapolation approach for the parameters of the 6-parameter heuristic method using the element-wise power function approximation according to (6.2), (6.4), and (6.5). The obtained optimal extrapolation parameter values are summarized in Table 3. Using the associated $\hat{w}(n)$, $\hat{p}(n)$, $\hat{\omega}(n)$, $\hat{i}(n)$, $\hat{\gamma}(n)$, $\hat{\delta}(n)$ functions, we also computed the SCV up to order 15000. The results are plotted in Figure 6. Unfortunately, the SCV values obtained by this 6-parameter extrapolation methods do not follow the same decay as the one of the 6-parameter

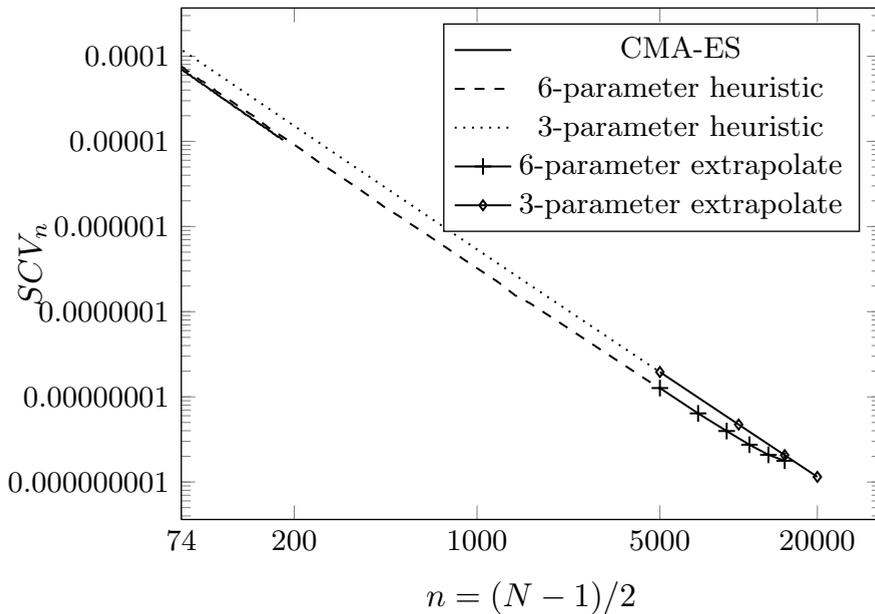


Figure 6: The approximated and the heuristics SCV as a function of order n in log-log scale

heuristic method up to $n = 5000$. At around, $n = 15000$ the SCV obtained from the 6-parameter extrapolation gets to be as high as the one obtained from the 3-parameter extrapolation. Most probable, the reason for this behaviour is the instability caused by the higher number of extrapolated parameters.

	a_i^*	b_i^*	c_i^*	\mathcal{D}_i
$\hat{w}(n)$	21.59	-1.017	0	1.958E-03
$\hat{p}(n)$	-26.35	-0.9762	5.653	1.963E-02
$\hat{\omega}(n)$	0.8952	-0.2458	0.1318	4.484E-03
$\hat{i}(n)$	0.9001	1	-0.7137	1.569
$\hat{\gamma}(n)$	3.631	-0.76579	0.9988	3.98E-03
$\hat{\delta}(n)$	10.48	-0.4475	0.8504	1.393E-01

Table 3: Optimal extrapolation parameters based on the 6-parameter heuristic optimization method according to Equation 6.2

Figure 7 plots the time to compute the SCV as a function of the order on a regular PC. The computation time is practically identical for both methods because the most expensive step of the computation is to transform the cosine-square form into the hyper-exponential form according to (3.3), which is need in both cases.

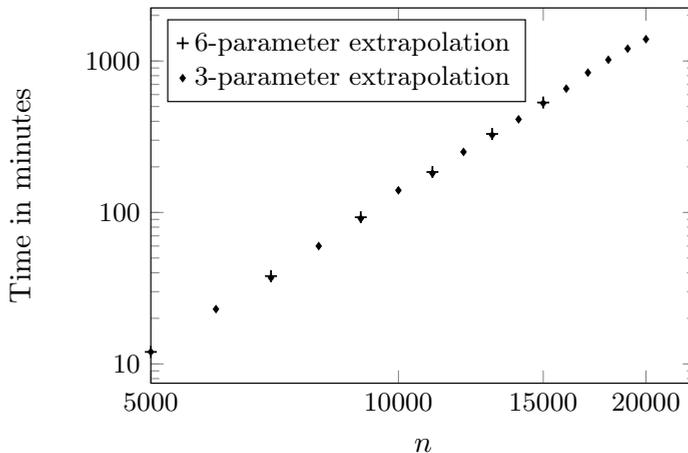


Figure 7: Running time of the parameter approximation procedures for different orders in log-log scale

Our full C++ implementation for the 3 and 6-parameter heuristic optimization methods is reachable at webspn.hit.bme.hu/~almousa/tools/CME_heur_approx.zip. The procedure uses extended floating point arithmetic when needed and it also contain the CMA-ES method, which is the optimization engine applied in the 3 and 6-parameter heuristic optimization methods

7. Conclusion

We propose an efficient 6-parameter heuristic SCV optimization procedure for concentrated matrix exponential distributions. The SCV values resulted by this 6-parameter optimization procedure are rather close to the ones obtained by the full optimization methods when both methods are feasible to compute, and seem to follow the same SCV decay trend for larger orders. Due to the exponential increase of the computation time as a function of the order, the applicability of the proposed heuristic optimization method extends until order $n = 5000$.

For larger orders, we also propose a parameter extrapolation approach, which allowed us to obtain CME distributions up to order 20000, such that the decay of the SCV follows the same trend as the one of the optimization procedures up to order 5000.

References

- [1] NAIL AKAR, OMER GURSOY, GABOR HORVATH, MIKLOS TELEK: *Transient and First Passage Time Distributions of First-and Second-order Multi-regime Markov Fluid Queues via ME-fication*, in: Methodology and Computing in Applied Probability (2020), pp. 1–27, DOI: 10.1007/s11009-020-09812-y.

- [2] *Curve Fitting Toolbox*, [for Use with MATLAB® : User’s Guide. Natick, MA :MathWorks, 2001.]
- [3] 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, ACM, 2009, pp. 2389–2396, DOI: 10.1145/1570256.1570333.
- [4] NIKOLAUS HANSEN: *The CMA evolution strategy: a comparing review*, in: Towards a New Evolutionary Computation, Springer, 2006, pp. 75–102, DOI: 10.1007/3-540-32494-1_4.
- [5] GÁBOR HORVÁTH, ILLÉS HORVÁTH, SALAH AL-DEEN ALMOUSA, MIKLÓS TELEK: *Numerical inverse Laplace transformation using concentrated matrix exponential distributions*, in: Performance Evaluation 137 (2020), p. 102067, DOI: 10.1016/j.peva.2019.102067.
- [6] GÁBOR HORVÁTH, ILLÉS HORVÁTH, MIKLÓS TELEK: *High order concentrated matrix-exponential distributions*, in: Stochastic Models 36.2 (2020), pp. 176–192, DOI: 10.1080/15326349.2019.1702058.
- [7] ILLÉS HORVÁTH, ANDRÁS MÉSZÁROS, MIKLÓS TELEK: *Numerical Inverse Transformation Methods for Z-Transform*, in: Mathematics 8.4 (2020), p. 556, DOI: 10.3390/math8040556.
- [8] ILLÉS HORVÁTH, ORSOLYA SÁFÁR, MIKLÓS TELEK, BENCE ZÁMBÓ: *Concentrated matrix exponential distributions*, in: European Workshop on Performance Engineering, Springer, 2016, pp. 18–31, DOI: 10.1007/978-3-319-46433-6_2.
- [9] P. REINECKE, M. TELEK: *Does a given vector-matrix pair correspond to a PH distribution?*, in: Performance Evaluation 80.0 (2014), pp. 40–51, DOI: 10.1016/j.peva.2014.08.001.