Concentrated matrix exponential distributions

Illés Horváth[1], Orsolya Sáfár[2], Miklós Telek[3], and Bence Zámbó[4]

[1] MTA-BME Information Systems Research Group
[2] Department of Analysis, Budapest University of Technology and Economics
[3] Department of Networked Systems and Services, Budapest University of Technology and Economics
[4] Institute of Mathematics, Budapest University of Technology and Economics

Abstract. We revisit earlier attempts for finding matrix exponential (ME) distributions of a given order with low coefficient of variation (cv). While there is a long standing conjecture that for the first non-trivial order, which is order 3, the cv cannot be less than 0.200902 but the proof of this conjecture is still missing.
In previous literature ME distributions with low cv are obtained from special subclasses of ME distributions (for odd and even orders), which are conjectured to contain the ME distribution with minimal cv. The numerical search for the extreme distribution in the special ME subclasses is easier for odd orders and previously computed for orders up to 15. The numerical treatment of the special subclass of the even orders is much harder and extreme distribution had been found only for order 4.
In this work, we further extend the numerical optimization for subclasses of odd orders (up to order 47), and also for subclasses of even order (up to order 14). We also determine the parameters of the extreme distributions, and compare the properties of the optimal ME distributions for odd and even order.
Finally, based on the numerical results we draw conclusions on both, the behavior of the odd and the even orders.
Keywords: matrix exponential distributions, minimal coefficient of variation

1 Introduction

The class of matrix exponential (ME) distributions, along with the subclass of phase type (PH) distributions are widely used for a number of reasons. The class PH is widespread in Markovian modeling: it can be efficiently used for approximating non-Markovian stochastic models (where random durations are non-exponential) by Markovian models, while maintaining a simple analytic formula [2].

The class ME exhibits an even wider range of behavior, allowing for even more efficient approximations; the cost is that a simple stochastic interpretation with background Markov chain is no longer available. Still, simple analytic formulas are retained, so it is still useful for efficient calculations.

The class PH is known to approximate deterministic delay poorly. That is why it is an important application of either class to approximate the deterministic distribution; in other words, we are looking for PH and ME distributions which are highly concentrated. A usual measure of concentration is the coefficient of variation (cv). For the class PH, it is known that the minimal cv depends only on the order n, it is $1/n$ and is obtained by the Erlang distribution [4].

However, for the class ME, only partial results are available. It has been known that ME(2) $\equiv$ PH(2) [12], that is the class of order 2 ME distributions is identical with the class of order 2 PH distributions.

For higher order ME, numerical investigations indicate that the minimal cv can be significantly lower compared to the class PH of the same order [10], [5]. No analytic results are available for the minimal value of cv in the class ME(n) for any $n \geq 3$. [5] lists some conjectures for the minimal value of cv; later, [10] numerically optimizes cv for a convenient subclass of ME(n) for odd values of n up to $n \leq 15$ and even values of n up to $n \leq 4$.

In the present paper, we extend the previously used approaches to larger values of n, for both odd and even orders and based on the numerical results we draw conclusions on the trends of their behavior with increasing orders.

The rest of the paper is organized as follows. In Section 2, we give the necessary mathematical background and notations are introduced. In Section 3, we introduce special subsets of ME(n) distributions for both odd and even values of n, detail the techniques and calculations necessary for finding distribution with minimal cv in the respective subsets, and present the results of the numerical optimization (some of which are put in the Appendix for better readability).

2 Preliminaries

Definition 1 *Let X be a non-negative continuous random variable with cumulative distribution function (cdf)*

$$F_X(t) = \Pr(X < t) = 1 - \alpha e^{\mathbf{A}t}\mathbf{1}, \qquad t \geq 0 \tag{1}$$

where α is a row vector of length n, $\mathbf{A}$ is a matrix of size $n \times n$ and $\mathbf{1}$ is a column vector of ones of size n. Then X is matrix exponentially distributed with representation $(\alpha, \mathbf{A})$, or shortly X is ME$(\alpha, \mathbf{A})$ distributed, where α is referred to as the initial vector and n as the order.

The probability density function (pdf) of X is then

$$f_X(t) = -\alpha \mathbf{A} e^{\mathbf{A}t}\mathbf{1}, \qquad t \geq 0. \tag{2}$$

We note that, as the above terminology suggests, α and $\mathbf{A}$ is not unique to F_X; the same function F_X may have several different α, A pairs referred to as representations. Further more, not even the order n of the representation is unique. There are ME$(\alpha_1, \mathbf{A}_1)$ of order m and ME$(\alpha_2, \mathbf{A}_2)$ of order n such that $m \neq n$ and ME$(\alpha_1, \mathbf{A}_1)$ and ME$(\alpha_2, \mathbf{A}_2)$ have the same distribution function, see e.g. [3].

Definition 2 *The class ME(n) contains matrix exponential distributions which have a representation of order at most n.*

Definition 3 *If X is ME($\alpha, \mathbf{A}$)-distributed, and α and $\mathbf{A}$ satisfies the following assumptions:*

- $\alpha_i \geq 0$,
- $A_{i,j} \geq 0$ *for* $i \neq j$, $A_{j,j} \leq 0$,
- $A\mathbf{1} \leq 0$

then we say X is phase type (PH) distributed, or shortly PH(α,$\boldsymbol{A}$) distributed. A representation $(\alpha, \mathbf{A})$ satisfying the above properties is called Markovian.

Definition 4 *The class PH(n) contains matrix exponential distributions which have a Markovian representation of order at most n.*

Based on the Jordan decomposition of $\mathbf{A}$ in (2) the probability density function (pdf) of an ME distribution has the following general form:

$$f(t) = \sum_{i=1}^{k} \sum_{j=0}^{N_i-1} c_{i,j} t^j e^{\lambda_i t}, \tag{3}$$

where $\lambda_1, \ldots, \lambda_k$ are eigenvalues of $\mathbf{A}$, and λ_i has multiplicity N_i. Some of the eigenvalues may be complex (in which case they come in complex conjugate pairs). Some eigenvalues of $\mathbf{A}$ may not be present in f; the eigenvalues which are present in the pdf are referred to as *contributing* eigenvalues. All contributing eigenvalues must have negative real parts, and among the contributing eigenvalues there is always a dominant real eigenvalue with maximal real part ([6], [8]). λ_1 denotes this dominant eigenvalue. In the class ME, there may be complex eigenvalues with real part equal to λ_1; in the class PH, this is not possible. Further differences between the classes ME and PH can be found in, e.g., [9, 7, 6].

To keep the subsequent discussion simple, we sometimes calculate with *unnormalized* pdfs, that is, $\int_0^\infty f(t)\mathrm{d}t = 1$ is not required, only

$$0 < \int_0^\infty f(t)\mathrm{d}t = m_0 < \infty.$$

Of course, this means that $\frac{f(t)}{m_0}$ is the proper normalized pdf corresponding to X. Then the moments can be calculated from an unnormalized f as

$$\mathbf{E}(X^n) = \frac{m_n}{m_0},$$

where

$$m_n = \int_0^\infty t^n f(t)\mathrm{d}t.$$

Definition 5 *The coefficient of variation* (cv) *of* $X \sim ME(\alpha, \mathbf{A})$ *is*

$$\mathrm{cv}(X) = \frac{\mathbf{E}(X^2) - (\mathbf{E}(X))^2}{(\mathbf{E}(X))^2} = \frac{m_2 m_0}{m_1^2} - 1.$$

The notation $\mathrm{cv}(f)$ *will be used as well.*

The coefficient of variation is a widely used measure of probability concentration of positive random variables. It is invariant to scaling of the variable (e.g., multiplying with a positive number, or changing the unit the random variable is expressed in); that is, $\mathrm{ME}(\alpha, \mathbf{A})$ and $\mathrm{ME}(\alpha, c\mathbf{A})$ (or, correspondingly, $f(t)$ and $f(t/c)/c$) have the same cv. This property also allows us to conveniently scale the considered distributions; for example, the dominant eigenvalue may be assumed to be -1.

For the class $\mathrm{PH}(n)$, the minimal cv is known.

Theorem 1 *[4] For* $X \in PH(n)$, $\mathrm{cv}(X) \geq \frac{1}{n}$, *and the minimum is obtained for the Erlang distribution with parameters* (n, λ) *where* $\lambda > 0$ *is arbitrary.*

We note that in accordance with our previous remark on scaling, λ (the dominant eigenvalue) does not affect the minimal cv.

However, an analytical result similar to Theorem 1 for $\mathrm{argmin}\{\mathrm{cv}(f) : f \in \mathrm{ME}(n)\}$ is available only for $n \leq 2$. $\mathrm{ME}(1) = \mathrm{PH}(1)$ is just the family of exponential distributions with $\mathrm{cv} = 1$, while $\mathrm{ME}(2) = \mathrm{PH}(2)$ with $\mathrm{cv} \geq 1/2$. For $n \geq 3$, [10] numerically optimizes cv for a convenient subclass of $\mathrm{ME}(n)$ for odd values of the order n up to $n = 15$, and conjectures that the minimal value of cv is indeed obtained within the given subclass.

3 Numerical optimization of cv for ME(n)

3.1 Optimization for odd n

Following [5, 10], for odd n, we look to minimize coefficient of variation in the subclass containing ME probability distribution functions of the following form (unnormalized pdf):

$$f(t) = \mathrm{e}^{-t} \prod_{i=0}^{(n-1)/2} \cos^2(\omega t - \phi_i). \tag{4}$$

This is the same subclass presented in [10], where the technology for computing the optimal parameters of (4) is the following: the moments and cv can be calculated analytically by Laplace-transform with Mathematica, then a numerical optimization is carried out for the variables $\omega, \phi_1, \ldots, \phi_{(n-1)/2}$. For more details, see [10]. The main difficulty for this approach is that as n increases, the analytical expression for cv gets more difficult to compute.

In the present paper, we use a different approach with only numerical steps. The moments are calculated by numerical integration for specific values of

$\omega, \phi_1, \ldots, \phi_{(n-1)/2}$. We found the minimum implementing an evolution strategy. Starting with one feasible solution, the parameter values were changed by adding a normally distributed random value for each parameter (mutating the solution). The deviation was changed according to Rechenberg's 1/5 rule; if the success ratio (that is, the cv has decreased for the new parameter values) after 20 steps is over 1/5, then the deviation is sligthly increased, in order to ensure the exploration of the search space. If the ratio is below 1/5 then the deviation is decreased in order to get close to the local optimum. Due to the numerous local optima and the high number of parameter values to be optimised, a population size of one was used, exploring the search space by the self-adaptation of the deviation. We changed to the new modified parameter values only if the cv decreased, otherwise we tried a new mutation. The best result from several runs was selected for each n. See [11] for more on the theory and [1] for the code in Matlab.

This approach is feasible for significantly higher values of n. We calculate the optimum for up to $n \leq 47$ and we also list the argmin parameters.

Non-negativity of $f(t)$ for $t \geq 0$ in (4) is guaranteed from its form in (4). The structure of the pdf is the following: the part $\prod_{i=0}^{(n-1)/2} \cos^2(\omega t - \phi_i)$ is periodic with period ω/π, and has $(n-1)/2$ zeros within a period. The exponential term in (4) introduces a decay that renders the part of the pdf from the second period on negligible. Within the first period, the optimal arrangement of zeros is best seen the following: the zeros are roughly (but not exactly) equidistant, starting from slightly above 0, but leaving a large gap between the last zero and the end of the period; see Figure 1. The closely situated zeros bring the pdf close to 0, while the large gap results in a relatively large, concentrated bump.

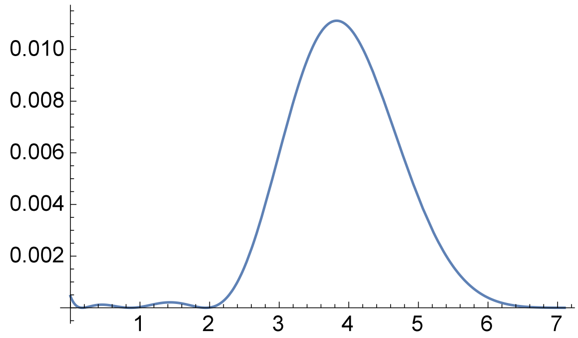
The form of (4) is yet again slightly different from the form given in (2) or (3); in Appendix A it is shown how (4) can be converted back to (3).

Table 2 contains the parameters which provide the minimal cv in ME(n). Due to readability, only values for $n \leq 17$ are displayed here; the rest of the list of optimal parameter values for n up to 47 can be found in Table 4 in Appendix B. Table 1 contains the minimal values of cv for various values of n; even values of n are also included in the table for easier comparison. Note that for odd n (even $n+1$), the difference between the optimal cv for n and $n+1$ is rather small, especially for larger values of n. The optimum in ME(n) and ME($n+1$) is further compared in Section 3.2 in more detail, and a detailed analysis of even values of n follows in Subsection 3.3.

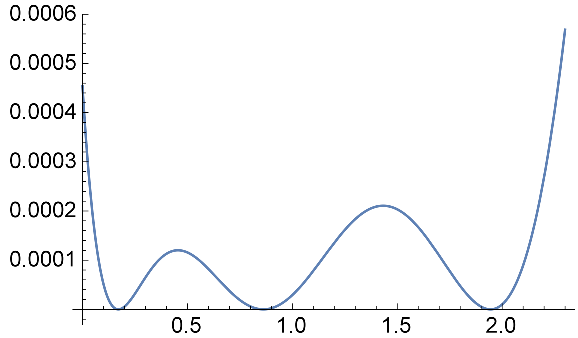
3.2 ME(n) for even n

For even values of n, we look to minimize coefficient of variation in the subclass containing ME probability distribution functions of the following form:

$$f(t) = a_3 \mathrm{e}^{t\lambda_2} + \mathrm{e}^{-t}(a_2 + \prod_{i=0}^{(n-2)/2} \cos^2(\omega t - \phi_i)) \tag{5}$$



Entire period



Zoomed in on zeros

Fig. 1: Optimal f for $n = 7$

Again, this subclass was already present in [5] but no optimization was carried out except for $n = 4$.

We present an intuitive argument that explains the subclass (5) and sheds some light on the difference between the optimal function in ME(n) and ME($n+1$) for odd n.

Compared to (4), the novelty in (5) is the term $(a_3 \mathrm{e}^{t\lambda_2} + a_2 \mathrm{e}^{-t})$. To understand the formulas better, we note that in (5),

$$\prod_{i=0}^{(n-1)/2} \cos^2(\omega t - \phi_i) \tag{6}$$

is periodic with period π/ω, and it has $(n-1)/2$ zeros for each period. Each zero effectively serves to make the function (6) more concentrated by bringing it close to 0 on a large portion of the interval $[0, \pi/\omega]$ (see Figure 1), while the exponential term e^{-t} in (4) brings a decay that renders the concentrated parts from the second interval on irrelevant.

Now for $n+1$, a new zero within the period is not possible, as an additional $\cos^2$ term would correspond to $n+2$ in the order. So instead we can use it to change some other part of the function instead. To have the most effect on the value of cv, we use this additional term $(a_3 \mathrm{e}^{t\lambda_2} + a_2 \mathrm{e}^{-t})$ to decrease the value

n	cv	$1/$ cv	n	cv	$1/$ cv
3	0.20090	4.9776	4	0.14981	6.6752
5	0.081264	12.306	6	0.075532	13.2394
7	0.042880	23.321	8	0.041349	24.1845
9	0.026157	38.231	10	0.025589	39.079
11	0.017494	57.163	12	0.017237	58.015
13	0.012470	80.195	14	0.012337	81.060
15	0.0093128	107.38			
17	0.0072074	138.75			
19	0.0057368	174.31			
21	0.0046708	214.10			
23	0.0038745	258.10			
25	0.0032646	306.31			
27	0.0027874	358.75			
29	0.0024053	415.76			
31	0.0020760	481.70			
33	0.0018094	552.66			
35	0.0015907	628.64			
37	0.0014092	709.64			
39	0.0012568	795.66			
41	0.0011278	886.71			
43	0.0010051	994.94			
45	0.00088322	1132.2			
47	0.00078490	1274.0			

Table 1: Minimal values of cv for various values of n

of the function around 0 to as small as possible. This leads naturally to the assumption $f(0) = 0$ for even n (see Figure 2). Note that the addition of the term (6) lifts all the zeros of the function (4) to a slightly positive value – except one. Numerical results show that the minimal cv is obtained when the first zero remains unlifted. See Figure 2.

The reason the cv value changes little when going from odd n to even $n+1$ is that the gain by the change $f(0) = 0$ is smaller in scale than what is gained in concentration by the introduction of a new $\cos^2$ term.

To summarize, we look for an optimal $f(t)$ in the form (5) with the additional properties

- $f(0) = 0$;
- $f(x) = 0$ and $f'(x) = 0$ for some $x > 0$.

Order	ω	$\phi_1, \phi_2, \phi_3, \ldots$
3	1.03593	0.337037
5	0.474055	1.67698; 2.10333
7	0.442459	1.64632; 1.95221; 2.4311
9	0.418775	1.62842; 1.86379; 2.23549; 2.69568
11	0.400272	1.61684; 1.80633; 2.10839; 2.48269; 2.90758
13	0.385334	1.60882; 1.7663; 2.01958; 2.33446; 2.69071; 3.0794
15	0.372959	0.07936; 1.60297; 1.73697; 1.95429; 2.22554; 2.53226; 2.86521
17	0.362491	0.19780; 1.59854; 1.90442; 1.71468; 2.14228; 2.41145; 2.70297 3.012594372647776

Table 2: Optimal parameter values for odd $n \leq 17$

3.3 Optimization for even n

The difficulty in optimizing $f(t)$ in (5) for an even n is that nonnegativity of $f(t)$ for $t \geq 0$ is not guaranteed and is difficult to check in general. To compute $f(t)$ with parameters with minimal cv, we reparametrize $f(t)$ in the following way.

We eliminate a_2 by solving $f(0) = 0$; thus

$$a_2 = -a_3 - \prod_{i=0}^{(n-2)/2} \cos^2(\phi_i).$$

Next we replace the variable a_3 by x where x is the zero of f; that is, solve the equation

$$f(x) = 0$$

for a_3. The solution is explicit for a given n:

$$a_3 = \frac{\prod_{i=0}^{(n-2)/2} \cos^2(\phi_i) - \prod_{i=0}^{(n-2)/2} \cos^2(\omega x - \phi_i)}{e^{x(\lambda_2+1)} - 1}.$$

Next we eliminate ϕ_1 from the equation

$$f'(x) = 0 \tag{7}$$

in the following way. Rewriting the terms containing ϕ_1 using the formula

$$\cos^2(\phi_1) = \frac{1}{2}(\cos(2\phi_1) + 1),$$

(7) leads to an equation that is linear in $\cos(2\phi_1)$ and $\cos(2(\omega x - \phi_1))$, which in turn leads to a quadratic equation for $\cos(2\phi_1)$. The solution is explicit but prohibitively long – we omit the explicit formula. We note that only one of the solutions corresponds to a proper nonnegative $f(t)$ in (5).

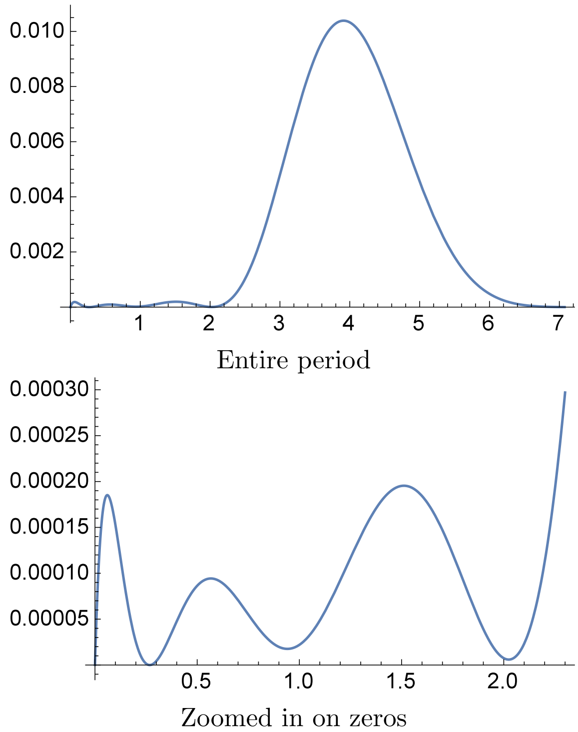


Fig. 2: Optimal f for $n = 8$

After these transformations, f (and thus cv) is parametrized by $x, \omega, \phi_2, \dots, \phi_{n/2-1}$. With these parameters, f is guaranteed to be nonnegative.

A numerical optimization for these parameters is now feasible; after optimization, the original parameters (a_2, a_3, ϕ_1) can be calculated explicitly. The results of the numerical optimization are presented in Table 3.

3.4 Rate of decay of cv

We numerically determine the rate of decay of the optimal value of cv as a function of n (the optimal value of cv for order n is denoted by $\text{cv}(n)$). Based on the cv values in Table 1 and due to the different behavior of $\text{cv}(n)$ for odd and even values of n as discussed in Section 3.2, we only consider odd values of n. Table 1 suggests a polynomial decay in n and this trend is tested by plotting $\text{cv}(n)$ against n with log-log scaling in Figure 3. The figure shows that the decay is very close to linear on log-log scale, with gradient -2.03. Least square fitting gives the approximation

$$\text{cv}(n) \approx \frac{2.175}{n^{2.03}},$$

but for a simple yet relatively accurate formula, one can also use

$$\text{cv}(n) \sim \frac{2}{n^2}.$$

Poles	λ_2	a_2	a_3	ω	$\phi_1, \phi_2, \phi_3, ...$
	Parameters				
4	1.94907	0.224603	-0.589603	0.519765	2.2195
6	-8.34112	0.00074	-0.01551	0.477957	1.78917; 2.1665
8	-13.917	4.50119×10^{-5}	-1.93645×10^{-3}	0.44404	1.70917; 1.98567 2.46912
10	-20.01365	4.53552×10^{-6}	-3.15042×10^{-4}	0.419619	1.66911; 1.88536 2.25959; 2.72108
12	-26.64365	5.89469×10^{-7}	-5.89972×10^{-5}	0.40078	1.64553 1.82172; 2.12538 2.5004; 2.9258
14	-33.78365	8.89442×10^{-8}	-1.20051×10^{-5}	0.385688	1.63023; 1.77798 2.03236; 2.34773 2.704288; 3.09330

Table 3: Optimal parameter values for even n

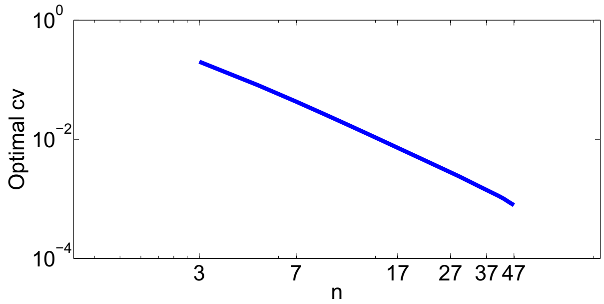


Fig. 3: Decay of cv(n) on a log-log scale

4 Conclusion

Following [10], we have made further numerical investigations for concentrated higher order ME distributions. We have obtained numerical results for a certain subclass conjectured to contain the ME distribution with minimal cv of ME distributions of odd order up to order 47 and for even order up to order 14. We also expanded and improved the optimization methods. We also provided the parameters of the extreme distributions, and compared the (numerical and abstract) properties of the optimal ME distributions for odd and even order.

As main conclusions we found that

– the minimal cv of even order n ME distributions gets to be very close to the one of order $n-1$ as n is increasing, and
– the minimal cv of odd order n ME distribution is close to $\frac{2}{n^2}$.

References

1. Code for numerical optimization in Matlab. `http://webspn.hit.bme.hu/~illes/mincvnum.zip`. Accessed: 2016-07-27.
2. N. G. Bean and B. F. Nielsen. Quasi-birth-and-death processes with rational arrival process components. *Stochastic Models*, 26(3):309–334, 2010.
3. C. Commault and S. Mocanu. Phase-type distributions and representations: some open problems for system theory. *Int. J. Control*, 76(6):566–580, 2003.
4. Aldous David and Shepp Larry. The least variable phase type distribution is Erlang. *Stochastic Models*, 3(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. abstract.
6. I. Horváth and M. Telek. A constructive proof of the phase-type characterization theorem. *Stochastic Models*, 31(2):316–350, 2015.
7. R. S. Maier. The algebraic construction of phase-type distributions. *Commun. Stat., Stochastic Models*, 7(4):573–602, 1991.
8. S. Mocanu and C. Commault. Sparse representations of phase-type distributions. *Commun. Stat., Stochastic Models*, 15(4):759–778, 1999.
9. Colm Art O'Cinneide. Characterization of phase-type distributions. *Communications in Statistics. Stochastic Models*, 6(1):1–57, 1990.
10. A. Horvath, P. Buchholz and M. Telek. Stochastic Petri nets with low variation matrix exponentially distributed firing time. *International Journal of Performability Engineering*, 7:441–454, 2011.
11. I. Rechenberg. *Evolutionstrategie: Optimierung technisher Systeme nach Prinzipien des biologischen Evolution.* Frommann-Hollboog Verlag, Stuttgart, 1973.
12. A. van de Liefvoort. The moment problem for continuous distributions. Technical report, University of Missouri, WP-CM-1990-02, Kansas City, 1990.

A Various forms of matrix exponential functions

This section is dedicated to show the equivalence of the various forms of matrix exponential functions throughout the paper. Specifically, we show the equivalence of the forms (2) and (3) and also show how (4) (and also (5)) can be brought to a form consistent with (2) and (3).

As mentioned in Section 2, (2) can be converted directly into (3) using the Jordan-decomposition of A.

From a pdf given in the form (3), one can reconstruct a matrix-vector representation in (2) in the following manner: A will be in block-diagonal form, with each block corresponding to either a single real eigenvalue λ_j or a pair of complex eigenvalues $\lambda_j, \lambda_{j+1} = a \pm \mathcal{I}b$, where $\mathcal{I} = \sqrt{-1}$.

If λ_j is real, then the block in A is

$$[\lambda_j] \qquad \text{and} \qquad \begin{bmatrix} \lambda_j & 1 & 0 & \dots & 0 \\ 0 & \lambda_j & 1 & \dots & 0 \\ \vdots & & \ddots & & \\ & & & \lambda_j & 1 \\ 0 & & \dots & 0 & \lambda_j \end{bmatrix}$$

for multiplicity 1 and $N_j > 1$, respectively.

For a complex pair of eigenvalues $\lambda_j, \lambda_{j+1} = a \pm \mathcal{I}b$, the block is

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \qquad \text{or} \qquad \begin{bmatrix} a & b & 1 & 0 & & & & 0 \\ -b & a & 0 & 1 & 0 & & \dots & 0 \\ 0 & 0 & a & b & 1 & 0 & & 0 \\ 0 & 0 & -b & a & 0 & 1 & 0 & 0 \\ 0 & & & & & & & \\ \vdots & & & & \ddots & & & \vdots \\ & & & & & & a & b \\ 0 & & & \dots & & & -b & a \end{bmatrix}$$

for multiplicity 1 or $N_j > 1$ respectively (the matrix on the right is size $2N_j \times 2N_j$). Once A is constructed, α can be obtained by solving a system of linear equations.

Finally, (4) can be represented in a form consistent with (3); the eigenvalues are $-1, (-1 \pm 2\mathcal{I}\omega), \dots, (-1 \pm (n-1)\mathcal{I}\omega)$. We demonstrate this for $n = 5$; for higher odd values of n, it follows a similar structure, albeit with more terms:

$$\begin{aligned} f(t) =& \mathrm{e}^{-t} \cos^2(\omega t - \phi_1) \cos^2(\omega t - \phi_2) \\ =& \frac{1}{8}\mathrm{e}^{-t}\big(2 + \cos(2\phi_1 - 2\phi_2)\big) + \frac{1}{8}\mathrm{e}^{-t}\big(\cos(2\phi_1)\cos(2t\omega) \\ &+ \sin(2\phi_1)\sin(2t\omega) + \cos(2\phi_2)\cos(2t\omega) + \sin(2\phi_2)\sin(2t\omega)\big) \\ &+ \frac{1}{8}\mathrm{e}^{-t}\big(\cos(2\phi_1 + 2\phi_2)\cos(4t\omega) + \sin(2\phi_1 + 2\phi_2)\sin(4t\omega)\big) \\ =& \frac{1}{8}\mathrm{e}^{-t}(2 + \cos(2\phi_1 - 2\phi_2)) + \\ &+ \frac{1}{8}(\mathrm{e}^{2i\phi_1} + \mathrm{e}^{2i\phi_2})e^{(-1-2i\omega)t} + \frac{1}{8}(\mathrm{e}^{-2i\phi_1} - \mathrm{e}^{2i\phi_2})e^{(-1+2i\omega)t} + \\ &\frac{1}{16}\mathrm{e}^{2i(\phi_1+\phi_2)}\mathrm{e}^{2t(-1+4i\omega)} + \frac{1}{16}\mathrm{e}^{-2i(\phi_1+\phi_2)}\mathrm{e}^{2t(-1-4i\omega)}. \end{aligned}$$

Representing (5) in a form consistent with (3) is essentially the same, just with one extra real eigenvalue.

B Optimal parameter values

19	0.353490	0.29829; 1.59508; 1.69721; 1.86521; 2.07671; 2.31637; 2.57570 2.85029; 3.13827
21	0.345640	0.10490; 0.38461; 1.59232; 1.68321; 1.83365; 2.02382; 2.23966 2.47319; 2.72005; 2.97806
23	0.338715	0.19898; 0.45955; 1.59006; 1.67176; 1.80774; 1.98031; 2.17653 2.38888; 2.61318; 2.84711; 3.08961
25	0.332545	0.04609; 0.281467; 0.52523; 1.58820; 1.66224; 1.78614; 1.94394 2.12371; 2.31837; 2.52389; 2.73798; 2.95940
27	0.327002	0.13289; 0.35436; 0.58327; 1.58663; 1.65422; 1.767866; 1.91313 2.07891; 2.25855; 2.44819; 2.64559; 2.84945; 3.05913
29	0.31348	0.03935; 0.24334; 0.45271; 1.14455; 1.58818; 1.65855; 1.77085 1.91037; 2.06735; 2.23615; 2.41354; 2.59761; 2.78723; 2.98175
31	0.308959	0.11601; 0.30925; 0.50731; 1.16023; 1.58673; 1.65148; 1.75543 1.88509; 2.03128; 2.18865; 2.35408; 2.52570; 2.70239; 2.88346 3.06855
33	0.304832	0.005513; 0.18531; 0.36880; 0.55668; 1.17491; 1.58548; 1.64538 1.74205; 1.86308; 1.99983; 2.14712; 2.30215; 2.46291; 2.62834 2.79775; 2.97073
35	0.301040	0.07689; 0.24812; 0.42288; 0.60154; 1.18864; 1.58439; 1.64006 1.73034; 1.84377; 1.97218; 2.11069; 2.25642; 2.40761; 2.56315 2.72235; 2.88477; 3.05017
37	0.297540	0.14201; 0.30558; 0.47218; 0.64248; 1.20148; 1.58344; 1.63540 1.72002; 1.82669; 1.94769; 2.07833; 2.21585; 2.35853; 2.50531 2.65547; 2.80859; 2.96438; 3.12272;
39	0.294289	0.04762; 0.20164; 0.35809; 0.51732; 0.67999; 1.21350; 1.58260 1.63127; 1.71086; 1.81149; 1.92585; 2.04945; 2.17962; 2.31470 2.45364; 2.59576; 2.74061; 2.88789; 3.03744
41	0.291265	0.10875; 0.25643; 0.40632; 0.55879; 0.71449; 1.62759; 1.22476 1.58186; 1.70267; 1.79788; 1.90627; 2.02352; 2.14707; 2.27531 2.40722; 2.54211; 2.67955; 2.81924; 2.96098; 3.10469
43	0.286709	0.017601; 0.15905; 0.29899; 0.44498; 0.59176 0.74070; 1.23092 1.58084; 1.62373; 1.69468; 1.78474; 1.88730; 1.99909; 2.12208 2.29038; 2.30290; 2.47864; 2.61584 2.74717; 2.88339; 3.02186
45	0.276478	0.08675; 0.22262; 0.35787; 0.498054; 0.63799; 1.08335; 1.25082 1.58080; 1.62305; 1.69202; 1.77581; 1.87951; 1.98293; 2.09473 2.20973; 2.32692; 2.45504; 2.57721; 2.70603; 2.83471; 2.96499 3.09616
47	0.283839	0.11500; 0.23946; 0.37625; 0.51191; 0.64863; 1.07858; 1.23819 1.57945; 1.61595; 1.65715; 1.76319; 1.85467; 1.95591; 2.06011 2.16775; 2.29640; 2.37101; 2.56629; 2.62701; 2.70473; 2.92734; 2.94086; 3.11827

Table 4: Optimal parameter values for odd values $19 \leq n \leq 47$