Approximating distributions and transient probabilities of Markov chains by Bernstein expolynomial functions

András Horváth[1], Lorenzo Ridi[2], Enrico Vicario[3]

[1]Dipartimento di Informatica - Università di Torino - horvath@di.unito.it
[2]Dip. Sistemi e Informatica - Università di Firenze - ridi@dsi.unifi.it
[3]Dip. Sistemi e Informatica - Università di Firenze - enrico.vicario@unifi.it

***Abstract*—In this extended abstract we consider the use of Bernstein polynomials (BPs) for the approximation of distributions and transient probabilities of continuous time Markov chains (CTMCs). We show that while standard BPs are not appropriate to this purpose it is possible to derive from them a family of functions, called in the sequel Bernstein expolynomials (BEs), which enjoys those properties that are necessary for these approximations. For what concerns distribution fitting, BEs correspond to a subset of the family of matrix exponential distributions and hence they are of interest in the field of matrix analytic methods. For what concerns transient probabilities of CTMCs, BEs can provide closed-form approximations which are useful in the analysis of models where the process subordinated to a possibly non-Markovian period is described by a CTMC. The application of BEs for approximating both distributions and transient probabilities will be illustrated through several numerical examples.**

***Index Terms*—continuous time Markov chain, Bernstein polynomial, matrix exponential distribution**

I. Introduction

A typical problem in stochastic modeling is to approximate the behaviour of a random process by an appropriate choice from a family of distributions or from a family of functions. It is often the case that the approximation is used as a building block of a model and for this reason such distributions and functions are convenient to use which facilitate (or at least do not make impossible) the analysis of the whole model.

A natural choice is to use polynomials and we will consider Bernstein polynomials (BPs) to approximate distributions and transient probabilities of CTMCs. Even if BPs have a number of favourable properties they are not feasible for such approximations in their standard form. We will show that by a change of variable it is possible to derive from them a family of functions which maintains the favourable features of BPs and provides the missing ones. We will refer to this family of functions as Bernstein expolynomials (BEs).

BEs can be used to fit distributions and the provided approximations are in the class of matrix exponential (ME) distributions [1]. The class of ME distributions is a proper superset of the well-known and widely-used class of phase type (PH) distributions [2] and it has been shown recently in [3] and [4] that many methods developed for models with PH distributions can be adopted to models with ME distributions.

We can use BEs also to approximate transient probabilities of CTMCs which provides a useful ingredient for a recently proposed technique for the analysis of non-Markovian processes. This technique, called the method of stochastic state classes, is based on grouping into classes those states of the stochastic process that share a common future [5]. An extension of this approach has been proposed in [6] which takes advantage of subordinated Markovian periods and requires a closed-form description of the transient probabilities of the subordinated CTMCs.

The paper is organised as follows. In Section II we give a brief description of BPs and introduce BEs. In Section III and Section IV we illustrate the application of BEs to distribution fitting and to approximation of transient probabilities of CTMCs, respectively. Conclusions are drawn in Section V.

II. Bernstein polynomials and expolynomials

The degree-n Bernstein polynomial [7] approximating a function f on $[a,b]$ is defined as

$$B_n(x) = \sum_{i=0}^{n} f\left(a + \frac{i}{n}(b-a)\right)\binom{n}{i}\frac{(x-a)^i(b-x)^{n-i}}{(b-a)^n} \tag{1}$$

Approximations based on BPs exhibit the following favourable properties. *Globality:* the approximant is global in the sense that a single polynomial is used to approximate the whole interval $[a,b]$. *Positivity:* since all the polynomials $(x-a)^i(b-x)^{n-i}, i = 0,1,\ldots,n$, are positive in the interval $[a,b]$ and we assume that also the function f is positive, the polynomial in (1) is positive on $[a,b]$. *Simplicity of derivation:* a BP is determined in a straightforward manner by the samples of the approximated function without any optimisation process. *Convergence of the approximation:* if the approximated function is continuous then the degree-n BP converges uniformly to it as n is increased. Moreover, the approximation error can be bounded by a Lipschitz inequality [7].

Conversely, the following requirements are not guaranteed by BPs but can be accommodated by a suitable adaptation. *Divergence of the approximant:* a BP is either a constant function or it diverges to plus or minus infinity. For this reason it cannot be used to approximate infinite support distributions and it is not feasible for the approximation of transient probabilities of CTMCs if the required interval overlaps with the steady state regime. *Unit-measure:* BP approximation does not guarantee that the approximant preserves the integral of the approximated function. This is clearly a problem when fitting distributions because the approximant is expected to have unit-measure.

The last problem can be managed simply by normalising the approximant. The problem of divergence can be solved instead by applying the change of variable $x \rightarrow e^{-x}$ to (1) resulting in the following sum of exponentials with negative exponents

$$E_n(x) = \sum_{i=0}^{n} f\left(-\log\left(e^{-a} - \frac{i}{n}\left(e^{-a} - e^{-b}\right)\right)\right) \times \binom{n}{i} \frac{(e^{-a} - e^{-x})^i (e^{-x} - e^{-b})^{n-i}}{(e^{-a} - e^{-b})^n} \qquad (2)$$

which converges as x tends to infinity. The functions defined in (2), which we call Bernstein expolynomials, maintain the favourable properties of BPs.

Both BPs and BEs catch exactly the approximated function at the limits of the considered interval, $[a, b]$. The important difference between BPs and BEs is that the latter can be used to approximate a function on the interval $[0, \infty]$. Applying (2) with $b = \infty$ the limiting value of the BE will catch exactly the limiting value of the approximated function.

Another difference between the two approximations is the way they sample the approximated function. BPs result in an equidistant sampling of the considered interval, i.e., the value of the approximated function is taken at the values $a + \frac{i}{n}(b-a), i = 0, 1, \ldots, n$. In order to consider the sampling applied by the expolynomial approximation, let us assume $a = 0, b = \infty$. In this case the approximated function is sampled at the values $-\log(i/n), i = n, n-1, \ldots, 0$. This implies that more samples are taken from the beginning of the considered interval. This effect can be mitigated by approximating a scaled version of the function, namely $f(x/c)$, and bringing back then the approximant to the original scale. Without going into details, we illustrate this effect in two figures. In Fig. 1 we depict for different values of c, assuming that $n = 10$ and the interval to be approximated is [0,10], the positions where the value of the approximated function, $f(x)$ itself, is taken. The larger c, the closer the sampling points are to a (for except the last sampling point which is always at b). Instead, as c tends to 0 the sampling tends to be equidistant. In Fig. 2 the effect of c for $n = 10, a = 0, b = \infty$ is shown.

When the functions in (2) are applied to fit the probability density function (pdf) of a distribution, the interval of the approximation must be set to [0,∞]. This guarantees that the limiting value of the approximant is 0 which is a necessary condition for a pdf. It is easy to check that the approximating pdf is in the class of ME distributions.

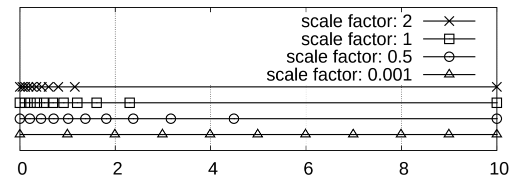

Fig. 1. Sampling points for different scale factors if $n = 10$ and the interval to fit is [0,10].

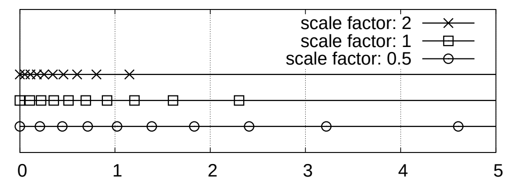

Fig. 2. Sampling points for different scale factors if $n = 10$ and the interval to fit is [0,∞] (also the value of the function at infinity is used).

III. Application of BEs to distribution fitting

In this section we illustrate the application of BEs to distribution fitting. We consider the distributions that are part of the benchmark proposed in [8] for PH fitting. Fig. 3-11 depict the pdf of the distributions and the pdf of the approximations for various degrees of the BE. For all the cases we have chosen such scalings of the original pdf which are favourable for the goodness of fitting. It can be seen that, as it is foreseen by the theory of BPs, increasing the degree the approximant is closer to the original pdf.

As mentioned earlier, the approximations provided by BEs capture the value of the approximated function at the borders of the considered interval. When it comes to distribution fitting, this property can be destroyed by the necessity of normalising the approximant. This is the case, for example, in Fig. 4 where the pdf corresponding to the degree-5 approximation has a rather differing value at 0.

It is important to have in mind that these approximations are provided with extremely low computational cost: for a degree-n approximation, it is sufficient to compute the value of the original pdf at $n + 1$ points.

IV. Application of BEs to transient probabilities

In this section we experiment the approximation of transient probabilities of CTMCs by BEs. We define the CTMC under study by a Petri net which is depicted in Fig. 12. In the net we have three non-exponential transitions, $t0$, $t1$ and $t2$, while the other transitions are with exponential firing time with parameter $1/2$. The exponential transitions give rise to a CTMC which is subordinated to the activity period of the non-exponential transitions. In order to analyse the model by the theory of stochastic state classes [5], [6], it is necessary to have a closed-form approximation of the transient behaviour of this subordinated CTMC.

This CTMC describes the interaction of producers and consumers exchanging products through a buffer that may

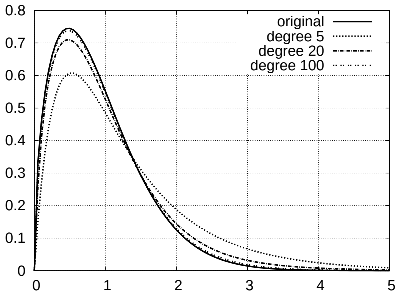

Fig. 3. Approximation of the Weibull(1,1.5) distribution.

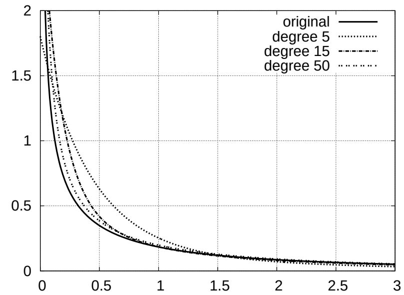

Fig. 4. Approximation of the Weibull(1,0.5) distribution.

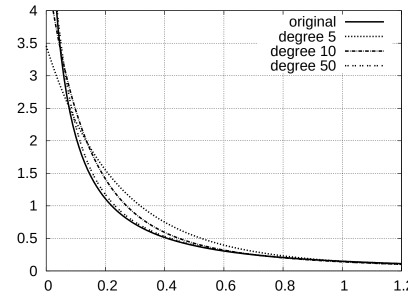

Fig. 5. Approximation of the lognormal(1,1.8) distribution.

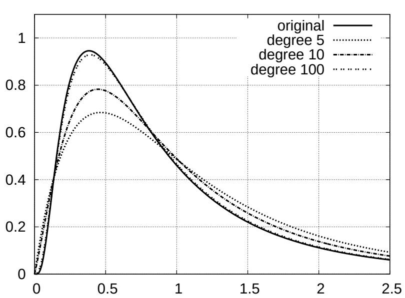

Fig. 6. Approximation of the lognormal(1,0.8) distribution.

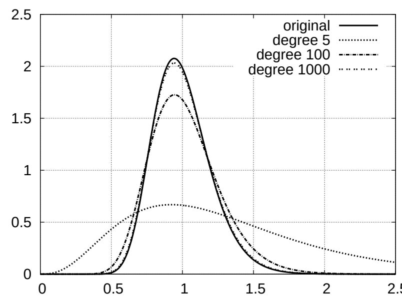

Fig. 7. Approximation of the lognormal(1,0.2) distribution.

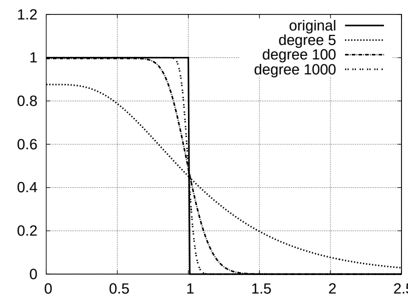

Fig. 8. Approximation of the uniform(0,1) distribution.

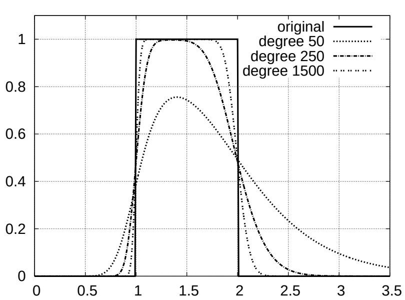

Fig. 9. Approximation of the uniform(1,2) distribution.

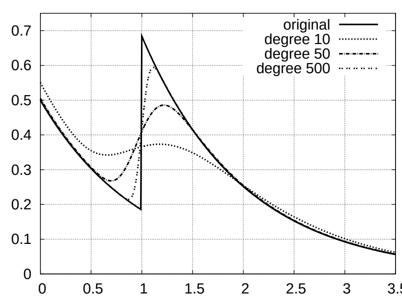

Fig. 10. Approximation of the shifted exponential distributions given in [8].

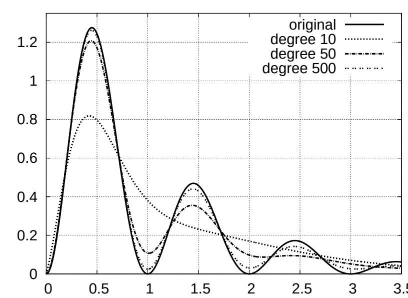

Fig. 11. Approximation of the matrix exponential distributions given in [8].

fail and be repaired. After each successful consume action (transition te), a choice is made between transition tf which restarts a new consume cycle and transition $tout$ by which the consumer leaves. We start the model with 4 tokens in place pa, 4 tokens in place $free$ and 4 tokens in place pd. With this initial marking the subordinated CTMC has 1218 states.

In Fig. 13-16 we have depicted some transient probabilities of the subordinated CTMC and the corresponding approximations for various degrees. We denote by $[\exp(tQ)]_{i,j}$ the probability of being in marking j at time t assuming that the chain started in marking i. The considered markings are: marking 1 which is the already described initial marking, marking 2 where the token distribution (not listing empty places) is $\#pa = 3, \#pb = 1, \#free = 3, \#pd = 4$, marking 10 with $\#pb = 4, \#pd = 4$ and marking 1000 with $\#pa = 3, \#pc = 1, \#pe = 2, \#pf = 2, \#failed = 1, \#busy = 1$. The corresponding transient probabilities cover a wide range of cases. To illustrate the effect of changing the sampling pattern by scaling, in Fig. 13 and 15 we give approximations for two different scalings. In particular, $[\exp(tQ)]_{1,1}$ is approximated well by a degree-1 BE (one exponential term plus a constant) if the scaling is chosen well. It is worth to observe that BEs are able to catch the probabilities in the steady state regime as well.

V. Conclusions

In this extended abstract we have proposed the family of Bernstein expolynomials for the approximation of distributions and transient probabilities of CTMCs. When applied to fitting distributions, Bernstein expolynomials results in ME distributions while, when approximating probabilities of CTMCs, they can provide compact, closed-form expressions. In both cases the computational cost of the approximation is extremely low.

The following aspects have not been covered and will be included in the full paper. The goodness of the approximations has to be verified both through statistical measures (like

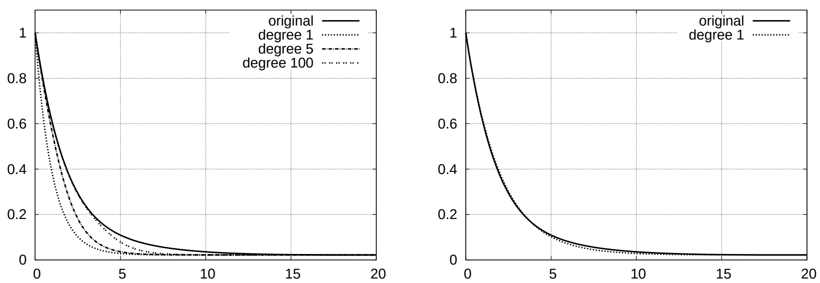

Fig. 13. Fitting $[\exp(tQ)]_{1,1}$ without scaling and with convenient scaling on the left and right, respectively.

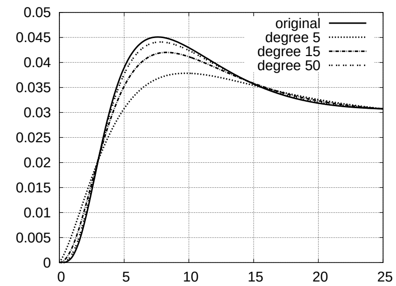

Fig. 14. Fitting $[\exp(tQ)]_{1,10}$ with scalings that are convenient for the whole interval

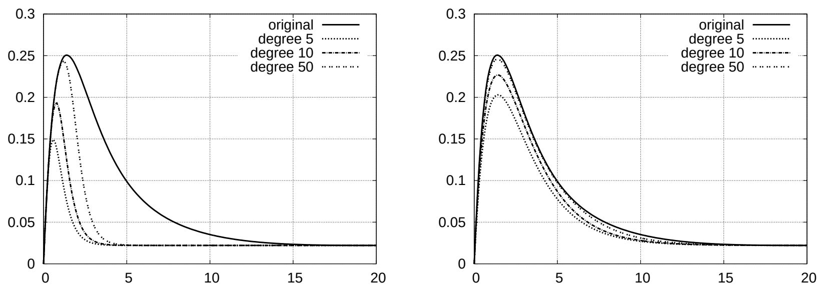

Fig. 15. Fitting $[\exp(tQ)]_{1,2}$ with a scaling that concentrates sampling points close to 0 (left) and with a scaling that is convenient for the whole interval (right)

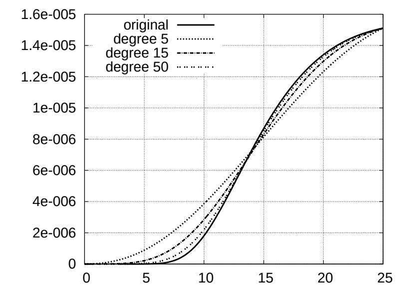

Fig. 16. Fitting $[\exp(tQ)]_{1,1000}$ with scalings that are convenient for the whole interval

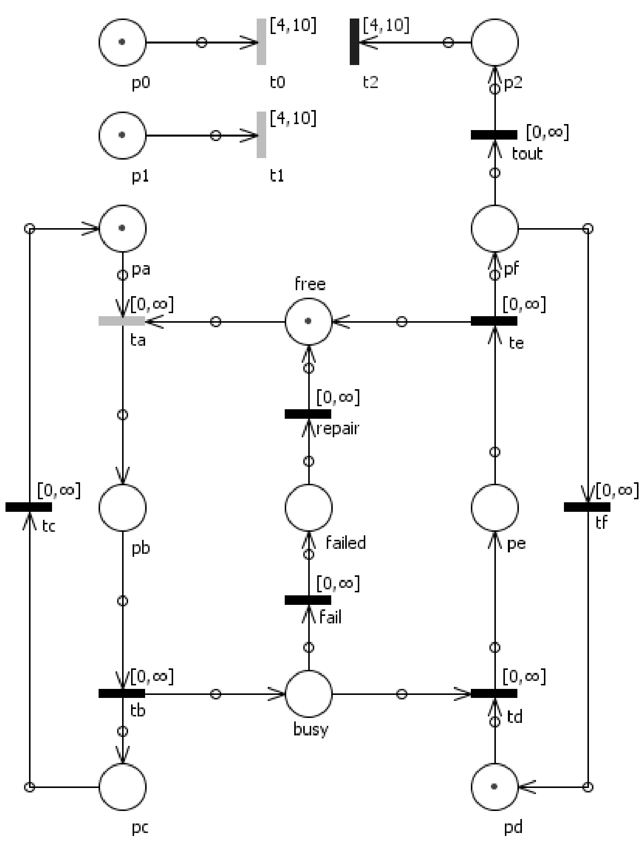

Fig. 12. Petri net whose subordinated CTMC is considered for experimenting BEs

moments) and by plugging them into applications. Bernstein polynomial and Bernstein expolynomial can be applied to multivariate functions as well and this can be necessary when the approximations are applied in the theory of stochastic state classes. For distributions, Bernstein expolynomial approximations result in ME distributions with a fixed pole-structure. This restriction can be mitigated by extending the class of Bernstein expolynomials but the approximations then require numerical optimisation.

REFERENCES

[1] L. Lipsky, *Queueing Theory: A Linear Algebraic Approach*. Springer, 2008.

[2] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approachg*. Dover, 1981.

[3] N. Bean and B. Nielsen, "Quasi-birth-and-death processes with rational arrival process components," Tech. Rep. IMM-Technical Report-2007-20, Technical University of Denmark, 2007.

[4] P. Buchholz and M. Telek, "Stochastic Petri nets with matrix exponentially distributed firing times." submitted for publication, 2009.

[5] L. Carnevali, L. Grassi, and E. Vicario, "State-density functions over DBM domains in the analysis of non-Markovian models," *IEEE Trans. on Soft. Eng.*, vol. 35, pp. 178 – 194, March-April 2009.

[6] A. Horváth and E. Vicario, "Aggregated stochastic state classes in quantitative evaluation of non-markovian stochastic Petri nets," in *Proc. of 6th International Conference on the Quantitative Evaluation of Systems (QEST)*, (Budapest, Hungary), Sept 2009.

[7] G. Lorentz, *Bernstein Polynomials*. University of Toronto Press, 1953.

[8] A. Bobbio and M. Telek, "A benchmark for PH estimation algorithms: results for Acyclic-PH," *Stochastic Models*, vol. 10, pp. 661–677, 1994.