Gradus Vol 3, No 2(2016) 13-22
ISSN 2064-8014

Lower tail estimation with Chernoff bound and its application for balancing electricity load by storage admission

Rajmund Drenyovszki[1], Lorant Kovacs[1], Kalman Tornai[2], Andras Olah[2] and Istvan Pinter[1]*

[1]Department of Information Technology, GAMF Faculty of Engineering and Computer Science, Pallas Athene University, Izsaki ut 10, H-6000 Kecskemet, Hungary
[2]Faculty of Information Technology, Pazmany Peter Catholic University, Prater utca 50/a, H-1083 Budapest, Hungary

Keywords:
Large Deviation Theory
Chernoff bound
valley filling
smart grid
Demand Side Management

Article history:
Received 08 Sept 2016
Revised 26 Sept 2016
Accepted 11 Nov 2016

Abstract
In this paper we investigate the applicability of the Chernoff inequality in finding an upper bound on the probability of the lower tail of the aggregate load. The importance of Demand Side Management (DSM) programs in power networks has increased recently, especially because of the new challenges like intensive use of renewable energy sources (wind, photovoltaic) and the expected high penetration of Electric Vehicles (EV). We show that Chernoff bound has the potential to be incorporated in DSM algorithms to integrate energy storage (e.g. batteries) elements into the power grid and facilitate load shifting.

1 Introduction

The main issue in electricity networks is keeping an almost perfect balance between electricity supply and demand. Oversupply means waste of energy, while undersupply causes performance degradation of the grid parameters (e.g. phase, voltage level, etc.). Additionally, there is a need to increase the percentage of renewable energy sources which gives rise to uncertainty in the generation side. The control of the supply side is difficult in many cases because of the large time constants of the base plants (fossile and nuclear); the only feasible solution is to use expensive auxiliary generators (e.g. gas and oil). Hence, an alternative way to keep the balance is to control the demand side. In the literature it is usually referred to as Demand Side Management (DSM) [1]. There are many DSM tecniques, from night-time heating with load switching, through time-of-use pricing, to direct load control, so we can say that in general DSM covers all the activities or programs undertaken by service providers to influence the amount or timing of electricity use. The residential sector accounts for about 30% of total energy consumption [2] and contains flexible appliances. The amount of consumption involved in direct control can eliminate the error between daily prediction based generation and actual demand. Furthermore, storage elements can be involved in electric power applications [3]. Especially electrical energy storages (EES) are promising candidates for the power network integration [4]. The spread of electric vehicles gives an additional impetus to the developement of DSM algorithms. In Europe cars are parked for more than 90% of the their time in average [5]; hence, batteries of electric vehicles may be used as an ancillary storage capacity for the power grid.

In this paper we investigate the applicability of Chernoff bound on the lower tail of the aggregate load for charging storage capacity as part of our previously proposed Consumption Admission Control algorithm [6] for DSM. Smart metering enables us to collect appliance level statistics, hence we can use this additional information in our methods. If we have storage elements in the system and we can

*Corresponding author. E-mail address: drenyovszki.rajmund@gamf.kefo.hu

charge them in low demand periods and use the energy in high demand periods, it will allow us to perform peak shaving and valley filling. For the illustration of the concept of load shifting see Figure 1. We show in this paper that Chernoff bound has the potential to be incorporated in CAC algorithms for DSM to integrate energy storage.

The rest of this paper is organized as follows. Related work is discussed in section 2. We describe our model in section 3. The Chernoff bound is formulated in section 4. Numerical results are explained in section 5 and some notes on the computational complexity in section 6. Finally the conclusions can be read in section 7.

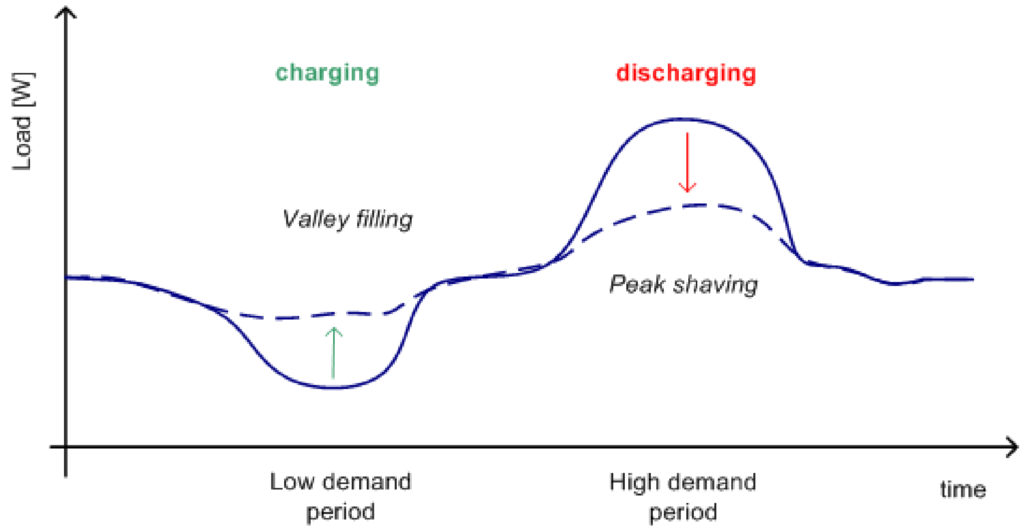

Figure 1. Valley filling and peak shaving with charging and discharging energy storage elements

2 Related work

In the literature we can find many attempts to incorporate energy storage elements into the power grid in many ways. One type of utilization of energy storage is to mitigate the intermittancy of renewable energy sources (RES). The paper [7] reviews the literature of three different kinds of energy storage technologies (pumped hydroelectricity storage, batteries and fuel cells), which have the potential for the integration/management of intermittency. The state of the art energy storage technology options for mitigating wind power intermittency is examined in [8]. Most of the literature investigates the potential of electric vehicles for load shift and renewable energy integration. [9] examined six European mobility studies to identify load shift potentials of electric vehicles, taking into account country-specific driving patterns. The main findings were that possibility to charge at the work place and controlled charging have big potential for load shifting. The vehicle-to-home (V2H) concept is investigated in [10], using the vehicle battery to reduce the peak demand of a household. In their simulation model an on-off controller was used to draw energy from the battery according to a treshold and the charging of the battery was performed in a constant level of 3kWs. The authors of [11] developed a mixed-integer linear program (MILP) to maximize RES utilization, scheduling optimal power and operation time for electric vehicles and appliances. They found that for small residential, solar powered buildings it is possible to schedule appliances and use the batteries of electric vehicles as an energy storage so that renewable energy covers 100% of the charging of the EVs.

3 Model

In a modern power network the smart meter can determine the individual pdfs of the installed appliances. As a consequence, we can build bottom-up consumption models in the smart meters, or based on aggregating the data of the smart meters, for much larger consumption districts as well. Based on the bottom-up model Large Deviation Theory (LDT) bounds can be used for the estimation of the tail probabilities. We use the following appliance model. There are N appliances that are connected to a smart meter. The consumption of the nth appliance at time instant k is $X_n[k]$. The random variables $X_n[k]$ are assumed to be independent for a given k. (Note, that the statistical descriptors are mostly time dependent, but the random variables are independent, e.g. the probability of turning on the light in House A is correlated with the date and hour of the day, but independent from the behavior of House B). In a domestic environment most of the appliances can be modeled by two states (i.e. ON-state and OFF-state), so in addition to the independence, we assume two-state appliance models. For the sake of mathematical convenience we further assume time independence, resulting in a mathematical model for the appliances as a two-state Bernoulli independent identically distributed (iid) random variable sequence. The basic benefit of the Bernoulli iid model is its simplicity; however, it cannot represent the severe auto-correlation of real consumption time series. The aggregate consumption at a given time instant k is

$$X\left[k\right]=\sum_{n=1}^{N}X_n\left[k\right] \tag{1}$$

where N is the number of appliances. Note, that the time dependence will be omitted, if it is not relevant (e.g. in the case of Bernoulli iid sequences).

Electricity networks system operators have optimal operation costs when the load is close to constant (pdf of the aggregate load is close to Dirac delta function). However, we cannot set the load to constant level, as a more realistic goal, we can keep the pdf as narrow as possible, i.e. the mass of the pdf lies between a lower and an upper limit. For the sake of an even more realistic model, we allow the tail probabilities to be non-zero but smaller than a predefined probability. In our model, p is the probability that the aggregate consumption X is greater or equal to the allowed maximum consumption C_{max} level (upper tail), while r is the probability that the aggregate consumption X is less or equal to the allowed minimum consumption C_{min} level (lower tail).

$$p=\Pr\left[X\geq C_{max}\right]=1-\Pr\left[X\leq C_{max}\right] \tag{2}$$

$$r=\Pr\left[X\leq C_{min}\right]=1-\Pr\left[X\geq C_{min}\right] \tag{3}$$

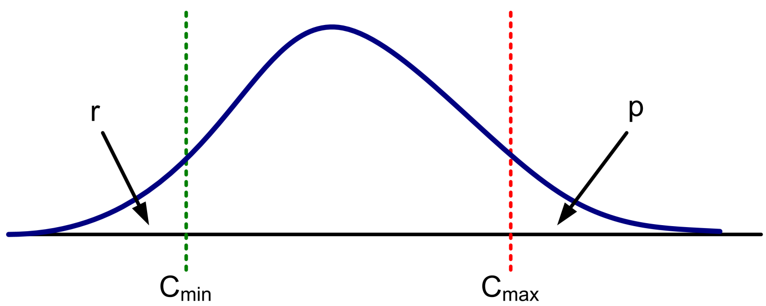

Figure 2. Probability and capacity parameters

The probabilities r and p can be calculated based on the probability density function $f_X\left(x\right)$ of the aggregate consumption. The pdf of the aggregate consumption can be calculated analytically by the convolution of the individual pdfs of all appliances:

$$f_X\left(x\right)=Pr\left(\sum_{i=1}^{M}\sum_{j=1}^{n_i}X_{ij}=x\right)=f_{X_{11}}\left(x\right)*f_{X_{12}}\left(x\right)*f_{X_{13}}\left(x\right)*\ldots*f_{X_{Mn_i}}\left(x\right). \tag{4}$$

where M is the number of appliance classes, and n_i is the number of appliances in class j, and $\sum\limits_{i=1}^{M}n_i$ is the total number of enabled appliances (An appliance class means a set of appliances that have the same statistical descriptors). Because of the independence assumption of the X_{ij} random variables, the expected value of the aggregate consumption X can be expressed as $\mu = E\left\{X\right\} = \sum\limits_{i=1}^{M}\sum\limits_{j=1}^{n_i}\mu_{ij}$; and the variance as $\sigma^2 = E\left\{(X-\mu)^2\right\} = \sum\limits_{i=1}^{M}\sum\limits_{j=1}^{n_i}\sigma_{ij}^2$. The convolution operation in (4) can be very time consuming in the case of high number of appliances and/or classes, so it is suggested to estimate the probability in terms of inequalities of Large Deviation Theory (LDT) bounds. We can define bounds as follows: lower $\breve{L}(C_{min})$ and upper bound $\widehat{L}(C_{min})$ on the probability of the lower tail (r), and lower $\breve{U}(C_{min})$ and upper bound $\widehat{U}(C_{min})$ on the probability of the upper tail (p).

$$\breve{L}(C_{min}) \leq \Pr\left[X \leq C_{min}\right] \leq \widehat{L}(C_{min}) \tag{5}$$

$$\breve{U}(C_{max}) \leq \Pr\left[X \geq C_{max}\right] \leq \widehat{U}(C_{max}) \tag{6}$$

From an application point of view, upper bounds are more frequently used, because service providers try to guarantee the quality of their services (QoS), so a maximum allowable probability for the mass of the lower and upper tail is preferred. In our previous paper [15] we showed the applicability of Markov, Chebisev, Bennett, Hoeffding and Chernoff inequalities for upper bounds of the upper tail. Now in the next section we are going to introduce Chernoff's inequality for the lower tail.

4 Chernoff bound

4.1 Markov's inequality

Markov's inequality gives an upper bound for the probability that a random variable is greater than or equal to an arbitrary positive constant. If X is a non-negative random variable and $\varepsilon > 0$, then

$$\Pr\left[X \geq \varepsilon\right] \leq \frac{E\left[X\right]}{\varepsilon} \tag{7}$$

Proof for discrete case

$$E\left[X\right] = \sum_i x_i p_i \geq \sum_{x_i \geq \varepsilon} x_i p_i \geq \varepsilon \sum_{x_i \geq \varepsilon} p_i = \varepsilon \Pr\left[X \geq \varepsilon\right] \tag{8}$$

In our case, we have the sum of independent random variables as the aggregate load which must be less or equal to the maximum consumption.

4.2 Chernoff's inequality applied for lower tail upper bound

Although Markov's inequality is not enough sharp, hence it cannot be applied for practical problems, it is the base of better inequalities [12]. To sharpen it, we can use the property of Markov's inequality, that it holds for monotonically increasing functions as well:

$$\Pr\left[X \geq C_{max}\right] = \Pr\left[f\left(X\right) \geq f\left(C_{max}\right)\right] = \frac{E\left[f(X)\right]}{f(C_{max})} \tag{9}$$

Based on Markov's inequality using $f\left(x\right) = e^{sx}$ function we can formulate Chernoff's inequality [13]:

$$\Pr\left[e^{sX} \geq e^{sC_{max}}\right] \leq \frac{E\left[e^{sX}\right]}{e^{sC_{max}}} \tag{10}$$

This formula is expressing the inequality for the upper tail:

$$\Pr\left[X \leq C_{min}\right] \leq \widehat{L}(C_{min}) \tag{11}$$

Our aim is to have the sharpest possible upper bound on the lower tail of the pdf. Using the reciprocal of both sides in the inequality inside the probability we have

$$\Pr\left[X \leq C_{min}\right] = \Pr\left[e^{sX} \leq e^{sC_{min}}\right] = \Pr\left[e^{-sX} \geq e^{-sC_{min}}\right] \tag{12}$$

Now using Chernoff's inequality introduced in (10) we can derive the following expression:

$$\Pr\left[e^{-sX} \geq e^{-sC_{min}}\right] \leq \frac{E\left[e^{-sX}\right]}{e^{-sC_{min}}} \tag{13}$$

As X is the sum of the independent random variables:

$$E\left[e^{-sX}\right] = E\left[e^{-s\sum\limits_i X_i}\right] = E\left[\prod_i e^{-sX_i}\right] = \prod_i E\left[e^{-sX_i}\right] \tag{14}$$

The moment generation function is formulated as follows in the case of Bernoulli iid random variables:

$$E\left[e^{-sX_i}\right] = p_i e^{-s1} + (1-p_i)e^{-s0} = p_i e^{-s} + (1-p_i) = 1 - p_i + p_i e^{-s} \tag{15}$$

Substituting the moment generation function into the inequality:

$$\Pr\left[e^{-sX} > e^{-sC_{min}}\right] < \frac{\prod_i \left\{1 - p_i + p_i e^{-s}\right\}}{e^{-sC_{min}}} = e^{log\left\{\frac{\prod_i\left\{1-p_i+p_ie^{-s}\right\}}{e^{-sC_{min}}}\right\}} = \tag{16}$$

$$= e^{log\left\{\prod\limits_i\left\{1-p_i+p_ie^{-s}\right\}-e^{-sC_{min}}\right\}} = e^{log\left\{\prod\limits_i\left\{1-p_i+p_ie^{-s}\right\}\right\}+sC_{min}} = e^{\sum\limits_i log\left\{1-p_i+p_ie^{-s}\right\}+sC_{min}} \tag{17}$$

where log is the natural logarithm. Now we can write the formula in a more compact way by using the logarithmic moment generation function ψ_i:

$$\psi_i\left(-s\right) = log\left\{1 - p_i + p_i e^{-s}\right\} \tag{18}$$

The upper bound on the probability of the lower tail can be expressed as follows. Note that from now on we will refer to it simply as Chernoff bound.

$$\Pr\left[X \leq C_{min}\right] \leq e^{\sum\limits_i \psi_i(-s)+sC_{min}} \tag{19}$$

The tightest bound is accomplished with the optimal s^* parameter that satisfies:

$$s^* : \inf_{s>0} \sum_i \psi_i\left(-s\right) + sC_{min} \tag{20}$$

5 Numerical results

In the numerical experiments the computation of the lower tail of probability functions were on the Chernoff bound (19) and for the sake of comparability by a modified convolution algorithm as well. It is known that the probability function of the sum of random variables can be calculated by the convolution of the individual probability functions (4). This operation is time consuming, however it is possible to speed up the calculation. Assuming two-state appliance model (consisting of ON-state and OFF-state) the vectors of the individual probability functions are sparse and the fact of sparsity can be advantageously used in the convolution algorithm. Our modified convolution algorithm is based on [17], and we will refer to it as Analytic computation from now on.

We have conducted two types of numerical experiments: one with only one appliance type (1000 instances of washer dryer), and the second one with several appliance types. Additionally two scenarios are considered in the second case: first one with 1000 pieces of each appliances and the second one with the number of appliances normalized to the same expected value.

Figure 3 shows the result of 1000 instances of washer dryer, with the parameters of $p_O N = 0.0012$ (ON-state probability) and $h = 800W$ (ON-state consumption). Analytic cdf and Chernoff bound is on top, the error expressed as the difference of Analytic and Chernoff result is on the bottom. The expected value (9600W) is highlighted with a vertical line. The marginal left side of the cdf (depicted in bigger in Figure 4) is the region of our special interest, because Chernoff bound, being as an LDT bound, is effective on the tails only. (Here we do not have enough space to show the experiments with all appliance types, but they show very similar results.)

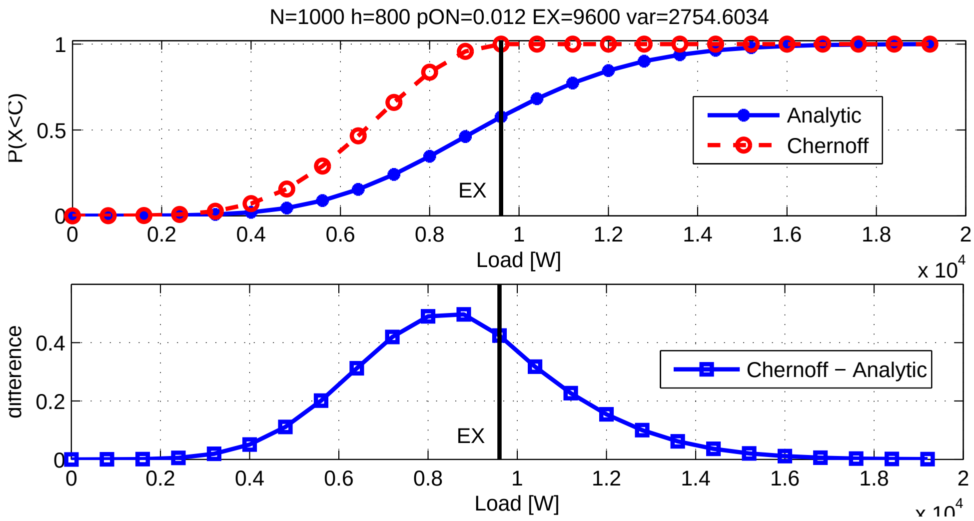

Figure 3. 1000 instances of washer dryer, Analytic cdf and Chernoff bound (top), Error expressed as the difference of Analytic and Chernoff result (bottom)

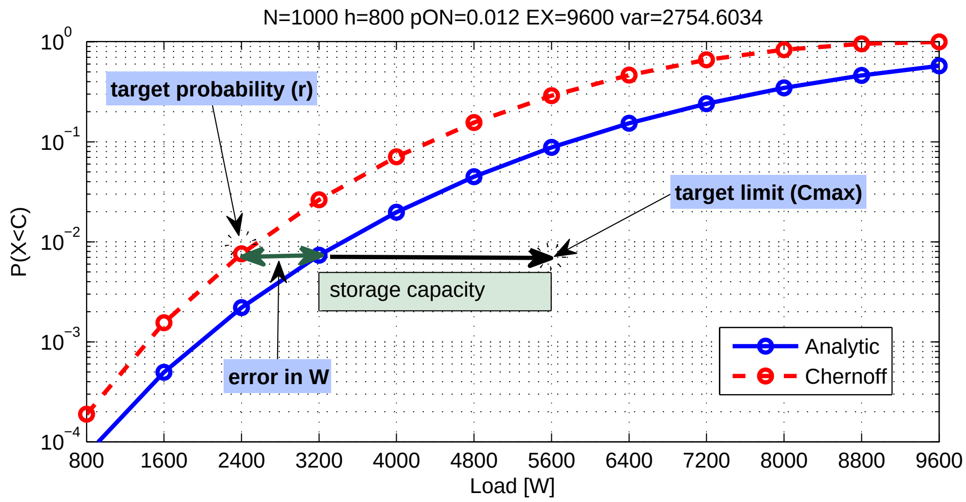

Figure 4. 1000 instances of washer dryer, Analytic cdf and Chernoff bound

As we can see in Figure 4, the difference between the actual cdf value (from Analytic computation) and the estimation (from Chernoff Computation) is less than half of a magnitude. For instance if we are interested in the probability that the consumption is equal or less than 3200W, the Chernoff computation gives 0.02631, but it is actually 0.007348. However, the difference seems a bit big for the first sight, in one hand we can state that according to our knowledge the Chernoff bound is the tightest from all the bounds. In the other hand, examining the result from an engineering perspective, notably the number of appliances needed to satisfy a certain value of probability, the results are promising. For the value of 3200W we found with the Analytic computation that the probability of underconsumption is 7.348×10^{-3}. In this case we had 1000 washer dryers. If we calculate the number of appliances needed to reach the same probability (or less) with Chernoff bound, we find that it is 1153 to reach the probability of 7.339×10^{-3}. It means a 15.3% increase (we need 115.3% of the original number of appliances).

Table 1. Performance of Chernoff bound

C_{min}	*1600W*	*2400W*	*3200W*	*4000W*
No. af appl. Analytic	1000	1000	1000	1000
No. af appl. Chernoff	1113	1134	1153	1172
probability (r) Analyitic	4.962×10^{-4}	2.197×10^{-3}	7.339×10^{-3}	1.981×10^{-2}
probability (r) Chernoff	4.921×10^{-4}	2.194×10^{-3}	7.348×10^{-3}	1.967×10^{-2}

Our purpose is to determine storage capacity in situations when the desired probability of underconsumption r cannot be satisfied in a certain capacity limit C_{min}. Let us explain the concept with the help of Figure 4. If for instance we have a target probability of 1×10^{-2} to keep the consumption above the limit of 2400W, than we can draw the conclusion from both Chernoff and Analytic computation that it is achievable. If the limit is 3200W and the target probability is the same 1×10^{-2}, than the Analytic result is satisfactory but Chernoff bound is not because of the error (labelled with "error in W" in Figure 4). Inversely, if we have a target consumption of 5600W on 1×10^{-2} probability, we can draw the inference that we need 3200W storage capacity.

In the followings two more scenarios are considered: first one with 1000 pieces of five appliance types and the second one with the number of appliances normalized to the same expected value. The appliance types with the parameters are depicted in (Table 1).

Table 2. Appliances types and parameters

	washer dryer	*microwave oven*	*dishwasher*	*refrigerator*	*lighting*
ON-state power consumption	800W	1500W	500W	200W	80W
Probability of ON state	0.012	0.016	0.044	0.254	0.335

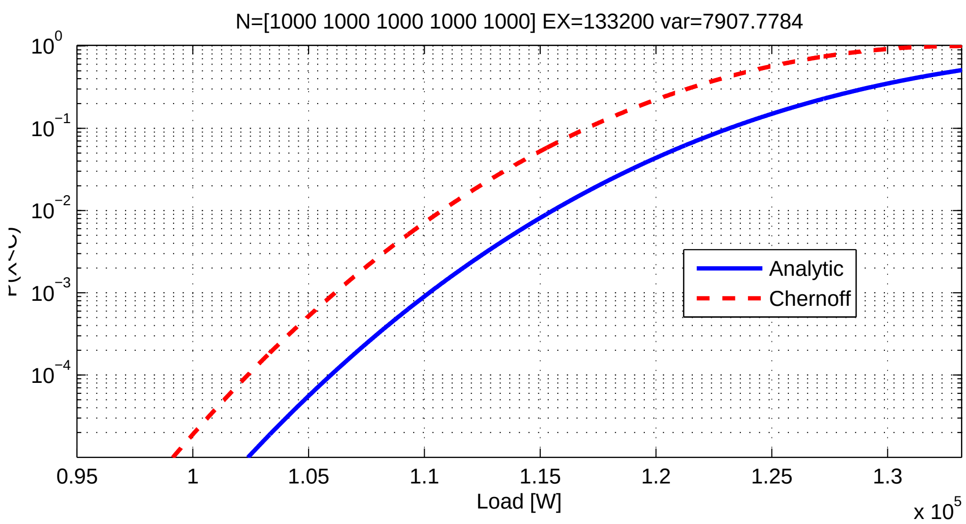

Figure 5. 5 appliance types, 1000 instances for each one, Analytic cdf and Chernoff bound

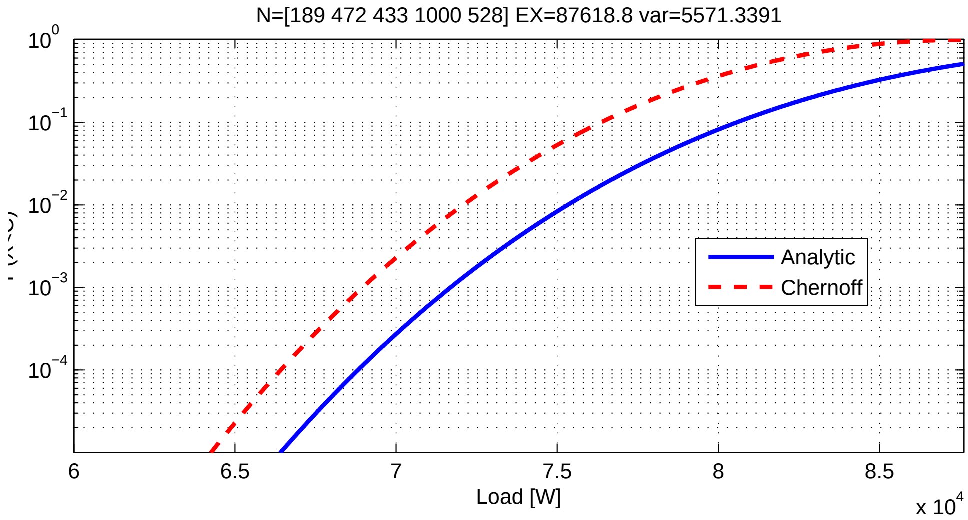

Figure 6. The number of appliances normalized to the same expected value, Analytic cdf and Chernoff bound

6 Computation time

The computation time of discrete convolution to calculate the pdf of the aggregate consumption depends highly on the implementation of the discrete convolution algorithm. However, discrete convolution can be faster with additional implementation tricks (e.g. utilizing more than one CPUs by parallelization parts of the algorithm), it remains a computationally intensive task. Instead of the discrete convolution formula, we used the concept based on the the algorithm presented in [17]. It avoids all of the zero term computations in the construction of the pdf of the aggregate as well as eliminates same computations by constructing binary tree from the individual pdfs and pairing the same types. Figure 7 depicts an experimental result of computation times both with Analytic and Chernoff computation.

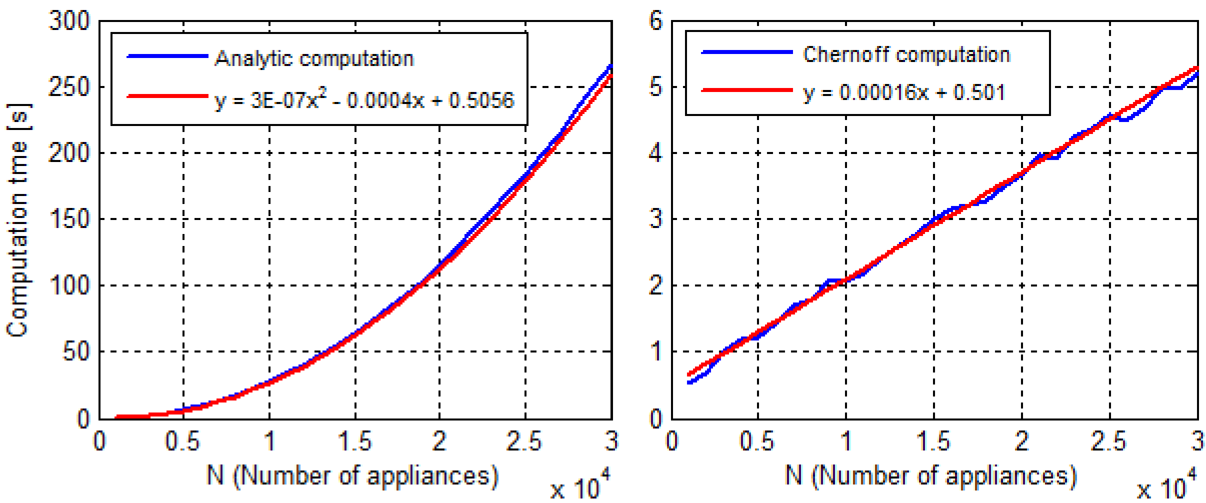

Figure 7. Computation time in the case of Analytic and Chernoff computation

Applying regression analysis on the trends of the computation times, we derived that in the case of Analytic computation the dependence of the computation time on the number of the appliances is quadratic, while in the case of Chernoff estimation it is linear.

7 Conclusion

Utilization of storage technologies in Demand Side Management is a promising way to facilitate load shifting (peak shaving and valley filling). In this paper we formulated the Chernoff inequality in finding an upper bound on the probability of the lower tail of the aggregate load. Instead of computing the probability function of the aggregate load directly with convolution we use Chernoff bound as an estimation because it is computationally more feasible. However we should consider, that the estimation introduces some error, it seems from the numerical results that from an engineering point of view it can be kept under control in storage admission applications. In our future work we would like to incorporate storage admission in our earlier proposed Consumption Admission Control algorithm based on appliance level statistical information and the Chernoff bound introduced in this paper.

Acknowledgement

This research and publication have been supported by the European Union and Hungary and co-financed by the European Social Fund through the project TAMOP-4.2.2.C-11/1/KONV-2012-0004: National Research Center for the Development and Market Introduction of Advanced Information and Communication Technologies. This source of support is gratefully acknowledged.

References

[1] Goran Strbac, Demand side management: Benefits and challenges, Energy Policy, vol. 36, Issue 12, pp. 4419-4426, Dec. 2008.

[2] L. G. Swan, V. I. Ugursal, "Modeling of end-use energy consumption in the residential sector: A review of modeling techniques," Renewable and Sustainable Energy Reviews, vol. 13, Issue 8, pp. 1819-1835, October 2009.

[3] I. Hadjipaschalis, A. Poullikkas, V. Efthimiou, "Overview of current and future energy storage technologies for electric power applications," Renewable and Sustainable Energy Reviews, vol. 13, Issues 6-7, pp. 1513-1522, 2009.

[4] X. Luo, J. Wang, M. Dooner, J. Clarke, "Overview of current development in electrical energy storage technologies and the application potential in power system operation," Applied Energy, vol. 137., pp. 511-536, 1 Jan. 2015.

[5] G. Pasaoglu, D. Fiorello, A. Martino, G. Scarcella, A. Alemanno, A. Zubaryeva, C. Thiel, "Driving and parking patterns of European car drivers - a mobility survey", European Commission, DG JRC, Institute for Energy and Transport, Petten, the Netherlands, 2012, http://publications.jrc.ec.europa.eu/repository/handle/JRC77079

[6] L. Kovacs, J. Levendovszky, A. Olah, R. Drenyovszki, D. Tisza, K. Tornai and I. Pinter: "A probabilistic approach for admission control of smart appliances in Smart Grids," Proceedings of 5th International Scientific and Expert Conference TEAM 2013, pp. 18-21, Presov, Slovakia, November 4-6, 2013.

[7] M. Y.i Suberu, M. W. Mustafa, N. Bashir, "Energy storage systems for renewable energy power sector integration and mitigation of intermittency," Renewable and Sustainable Energy Reviews, vol 35., pp. 499-514, 2014.

[8] T. R. Ayodele, A. S. O. Ogunjuyigbe, "Mitigation of wind power intermittency: Storage technology approach," Renewable and Sustainable Energy Reviews, vol. 44, pp. 447-456, 2015.

[9] S. Babrowski, H.i Heinrichs, P. Jochem, W. Fichtner, "Load shift potential of electric vehicles in Europe," Journal of Power Sources, vol. 255, pp. 283-293, 2014.

[10] G. Haines, A. McGordon, P. A. Jennings, N. Butcher, "The simulation of vehicle-to-home systems - using electric vehicle battery storage to smooth domestic electricity demand", Proceedings of the International Conference on Ecologic Vehicles and Renewable Energies, 2009.

[11] P. Mesaric, S. Krajcar, "Home demand side management integrated with electric vehicles and renewable energy sources, " Energy and Buildings, vol. 108, pp. 1-9, 2015.

[12] Z. Lin, Z. Bai, "Probability inequalities," Springerverlag Berlin Heidelberg, 2010.

[13] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," The Annals of Mathematical Statistics, vol 23, sz 4, o 493-507, 1952.

[14] P. Massart, "Concentration inequalities and model selection," Ecole d'Eté de Probabilités de Saint-Flour 2003. Lecture Notes in Mathematics 1896, Springer Berlin/Heidelberg. 2007.

[15] L. Kovacs, R. Drenyovszki, A. Olah, J. Levendovszky, K. Tornai, I. Pinter, "A probabilistic demand side management approach by consumption admission control," ArXiv e-printsarXiv:1608.00546.

[16] R. Drenyovszki, L. Kovacs, I. Pinter, A. Olah, K. Tornai, J. Levendovszky: "Power system reliability assessment for the residential sector based on Large Deviation Theory bounds," Proceedings of EnergyCon 2016, IEEE International Energy Conference, Leuven, Belgium, April 4-8. 2016.

[17] D.L. Evans, L.M. Leemis, "Algorithms for computing the distributions of sums of discrete random variables," Mathematical and Computer Modelling, vol. 40, Issue 13, pp. 1429-1452, 2004.