Piecewise Linear Approximation of Generators Cost Functions Using Max-Affine Functions

Hamed Ahmadi
José R. Martí
School of Electrical and Computer Engineering
University of British Columbia
Vancouver, BC, Canada
Email: hameda@ece.ubc.ca, jrms@ece.ubc.ca

Ali Moshref
BBA
Vancouver, BC, Canada
Email: ali.moshref@bba.ca

***Abstract*—Nonlinear functions are often encountered in power system optimizations. In this paper, an effective piecewise linear (PWL) approximation technique is introduced which shows promising performance in linearizing the nonlinear functions. This method uses a series of linear functions, called max-affine functions, to linearize a multivariate function over a bounded domain. The important advantage of this method is its ability to decide on the size of the subspaces, which other methods are not capable of. It is also shown that using the PWL approximation, significant efficiency is achievable in computation burden of most power system optimizations, such as unit commitment.**

NOMENCLATURE

$\bar{B}$	Network susceptance matrix.
$\bar{D}$	Matrix product of the susceptance and node-incident matrices.
a, b, c	Quadratic cost function coefficients.
C^{SDn}	Shutdown cost.
C^{SUp}	Startup cost.
C_T	System total cost.
N_b, N_g	Sets of buses and generators, respectively.
N_l, N_t	Sets of lines and time horizon, respectively.
P	Generator active power.
$P^{\max}$, $P^{\min}$	Upper/lower limits on generator active power.
P_d	Active power demand.
P_i^L	Active power flow limit of Line i.
P^{SDn}	Generator minimum power limit at shutdown.
P^{SUp}	Generator maximum power limit at startup.
R^{Dn}	Generator power ramp-down limit.
R^{Up}	Generator power ramp-up limit.
s	Number of PWL partitions.
SDn, SUp	Auxiliary variables for shutdown/startup cost.
u	Generator status (0:'Off', 1:'On').
δ	Bus voltage angle.

I. INTRODUCTION

The increasing demand on real-time operation of power systems has led to numerous research studies on enhancing the power system analysis methods. Due to large-scale problems that need to be solved for real-size power systems, special numerical methods have been developed and utilized by power system experts. Amongst most computationally expensive problems in power system analysis, optimization of system operation is of crucial importance.

Most of the optimization problems in power systems are basically nonlinear, thus nonlinear programming (NLP) techniques have been widely applied to these problems. As an example for continuous optimization, the optimal power flow problem considering AC constraints is an NLP problem, for which a variety of methods have been proposed in the literature, e.g. Interior Point Method [1] and Trust-Region Method [2]. However, due to the nonconvexity of the original problem, most of the methods would probably get trapped in a local minimum. As an example of combinatorial optimization, the unit commitment problem is nonconvex due to the binary variables associated with the generators' status (on/off). Having a quadratic cost function and nonlinear constraints, one has to deal with a mixed-integer nonlinear programming (MINLP) problem, which is NP-hard, and up to now, there is no efficient method for solving large-scale MINLP problems.

Various versions of the unit commitment problem have been formulated in the literature, focusing on linearizing the constraints as well as the objective function, e.g. [3], [4]. In [3], the quadratic cost functions are linearized within the generator's output power limits which leads to $2s + 2$ new constraints and $s + 1$ new variables (s is the number of sections assumed for linearization). Beside increasing the size of the problem, it is not clear how to choose the intervals and, therefore, it might not lead to the best possible piecewise linear (PWL) approximation of the function. In [4], the convex envelope of the quadratic function is obtained and using the *perspective cut*, the linear parts are added to the problem. The problem with this method is that a dynamic constraint has to be added to the original problem which slows down the solution process. In addition, a linearization technique is used in [5] to iteratively solve the MILP problem by replacing the objective by its linear approximation. The method is known as the Kelley's theorem on the cutting plane method for convex programs. However, many iterations and cuts are required to reach to the solution and the number of required cuts is not predictable beforehand.

In a more recent work, a mathematical approach is employed to find the PWL approximation of the quadratic cost functions [6]. Using this method, the length of the intervals can be selected optimally so that a good PWL approximation is obtained. Although this is believed to be a tighter approximation comparing to the previous methods, it is not necessarily the best PWL approximation.

In addition to the quadratic cost functions of generators, there are more instances of nonlinear functions in power system studies, many of which could be a multi-variable function. The problem of fitting a PWL approximation to a multi-dimensional set of data (possibly obtained by evaluating a nonlinear function at different points) has been studied before. The least-squares and Neural Network methods are among the most-used approaches in curve-fitting. In order to find the best PWL approximation with a fixed number of segments available, Magnani and Boyd have proposed a convex PWL fitting technique which is based on the so-called *Max-Affine* function [7]. This method is capable of providing PWL approximations for multivariate functions. This approach has specific application in convex optimization.

To the best of authors' knowledge, this is the first time in power system studies that the Max-Affine functions are used for PWL applications. The capabilities of this method are shown through numerical examples. This is the starting point for the numerous possibilities of these family of PWL approximations in mathematical programming and optimization of power systems. The rest of the paper is organized as follows.

In Section II, the PWL techniques are reviewed and illustrative examples are presented. In Section III the application of the PWL technique in unit commitment problem and its impact on the solution quality and computational efficiency are evaluated. The paper is concluded by highlighting the main contributions and findings of the present study.

II. Piecewise Linear Approximation of Multivariate Functions

Assume a function of multi variables, say $f(x)$ with $x \in \mathbb{R}^n$, is defined within a bound on x, say $x \in \mathcal{D} = \{x|\underline{x} \leq x \leq \overline{x}\}$. Depending on the level of precision required, $f(x)$ can be approximately expressed as PWL functions over small sub-intervals inside $\mathcal{D}$. For instance, considering s intervals, one can derive the following as a PWL approximation of $f(x)$:

$$\hat{f}(x) = \begin{cases} a_1^T x + b_1 \ , & x \in \mathcal{D}_1 \\ a_2^T x + b_2 \ , & x \in \mathcal{D}_2 \\ \quad \vdots & \\ a_s^T x + b_s \ , & x \in \mathcal{D}_s \end{cases} \tag{1}$$

in which $a_i \in \mathbb{R}^n, \ b_i \in \mathbb{R}$ and the following holds:

$$\bigcup_{1 \leq i \leq s} \mathcal{D}_i = \mathcal{D} \quad \text{and} \quad \bigcap_{1 \leq i \leq s} \mathcal{D}_i = \emptyset \tag{2}$$

and on the borders of sequential $\mathcal{D}_i$, the linear segments are connected, which means that $\hat{f}(x)$ is continuous.

Although (1) sounds interesting, there are real challenges in deriving $\hat{f}(x)$. The first challenge is how to mesh $\mathcal{D}$ into its subspaces $\mathcal{D}_i$. The second possible challenge is how to choose the smallest s while still maintaining a good accuracy in the approximation. In order to clarify this and without loss of generality, let the dimension of x be one. In the following, the mentioned challenges are discussed in more details.

A. PWL Approximation for Convex Quadratic Functions

In this section, the dimension of the problem is reduced to one for illustrative purposes. Assume that $f(x)$ is a convex quadratic function of the form

$$f(x) = ax^2 + bx + c \ , \ a > 0, \ \underline{x} \leq x \leq \overline{x} \tag{3}$$

1) Mathematical Background: It is obvious that the most interesting PWL approximation of f is the one with the fewest linear parts and highest accuracy. It is also clear that having more accuracy requires greater number of linear parts. However, if the PWL approximation is going to be used in, for example, an optimization problem, fewer segments is more interesting. Therefore, practical purposes limit the maximum number of segments one can use in the PWL approximation. Now, the question that remains is "having a limit on s, what is the best $\hat{f}(x)$?"

In order to answer the above question, it should be recalled that $\hat{f}(x)$ can be expressed in more compact form as (referred to as Max-Affine function in [7])

$$\hat{f}(x) = \max_{1 \leq i \leq s} \{\alpha_i x + \beta_i\} \tag{4}$$

which surprisingly has no constraints on the subspaces on which the linear approximations are defined. The best $\hat{f}(x)$, with fixed s and m point-wise function evaluations, can then be obtained using the following least-squares problem:

$$\min_{\alpha_i, \beta_i} \quad \sum_{k=1}^{m} \Big(\max_{1 \leq i \leq s} \{\alpha_i x_k + \beta_i\} - f(x_k) \Big)^2 \tag{5}$$

Unfortunately, this problem is not convex [7]. However, an efficient method is proposed in [7] to find the solution of (5). This method is based on choosing the initial subspaces (i.e. $\mathcal{D}_i$) and updating them iteratively to find the best possible mesh on $\mathcal{D}$. Beside that algorithm, there are commercial solvers capable of handling these types of problems, which are usually categorized as *non-smooth* problems. Some instances are CONOPT, MINOS, LGO and IPOPT, all available in GAMS under the option "nonlinear programming with discontinuous derivatives" [8]. The discussion on the algorithms for solving non-smooth optimization problems is beyond the scope of this paper and the reader is referred to the software user manual.

2) Numerical Example: Here, a numerical example for PWL approximation of a cost function for a thermal generation unit is presented. The coefficients in (3) are assumed to be $a = 0.9$, $b = 10$, $c = 200$, $\underline{x} = 10$, $\overline{x} = 200$. Assume that only two segments are allowed, i.e. $s = 2$. Figure 1 shows the original quadratic function, the PWL upper approximation (e.g. used in [3]) obtained by halving the space, and the PWL max-affine approximation. The following remarks are observed:

- Halving the available space is not the optimum way of meshing (Points A and B in Fig. 1 are not equal).
- The PWL upper approximation has no intersection with the original function within the intervals but on the two ends.
- The PWL max-affine approximation has 2 intersections with the original function within each interval (P_1, P_2 in the first interval and P_3, P_4 in the second one in Fig. 1).
- The average of the relative absolute errors for the PWL upper approximation is 32% while this value for the PWL max-affine approximation is 20%.

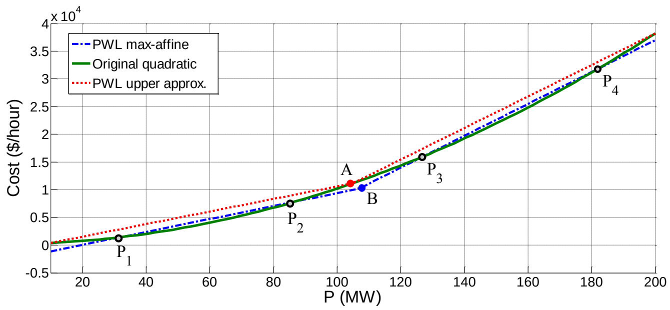

Figure 1. Comparison between the PWL upper and max-affine approximation techniques.

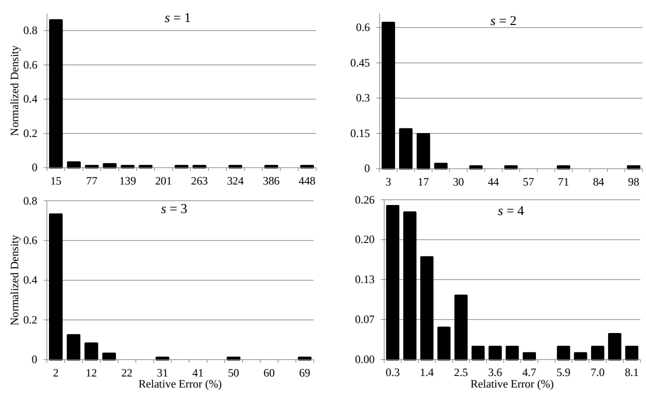

Figure 2. Histograms of relative errors between the linearized and original functions obtained using the PWL max-affine approximation.

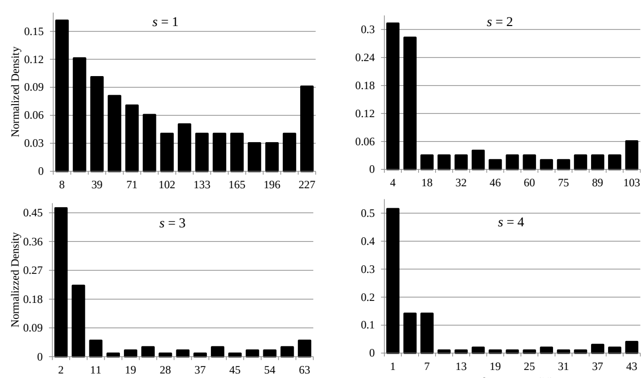

Figure 3. Histograms of relative errors between the linearized and original functions obtained using the PWL upper approximation.

Table I
AVERAGE OF THE RELATIVE ABSOLUTE ERRORS FOR DIFFERENT VALUES OF s OBTAINED BY TWO PWL TECHNIQUES

	PWL Technique	Number of Partitions (s)			
		1	2	3	4
Average of relative absolute errors (%)	Upper Approximation	95.3	31.8	15.5	9.45
	Max-Affine Functions	90.5	20.1	9.2	1.3

It is obvious that the PWL max-affine approximation is more efficient than the other method. Moreover, the method is able to decide the optimal length of the intervals. Table I shows the average of relative absolute errors for different values of s. Also, Figs. 2 and 3 depict the histograms of the relative absolute errors for different values of s using the PWL max-affine and upper approximations, respectively. As can be seen, the average relative error for the PWL max-affine approximation is significantly less than the same values for the PWL upper approximation.

III. APPLICATIONS

A. Min-Max Optimization

If the nonlinear function happens to be in the objective of a minimization programming, the PWL max-affine approximation leads to a linear reformulation of the objective. This type of problems is referred to as *Min-Max* optimization. Mathematically, it is described as

$$\min_{x \in \mathcal{D}} \max_{1 \leq i \leq s} \{\alpha_i^T x + \beta_i\} \tag{6}$$

By introducing a new variable, $z = \max\{\alpha_i^T x + \beta_i\}$, the above problem is reformulated as

$$\min_{x \in \mathcal{D}} z \tag{7a}$$

$$\text{subject to} \quad z \geq \alpha_i^T x + \beta_i \ , \ i = 1, \ldots, s. \tag{7b}$$

Therefore, by introducing one extra variable and s extra inequalities, the problem is reformulated as a linear programming (assuming other constraints to be linear). This has tremendous applications in power system optimization. As an example, this method is applied to the unit commitment problem in the following.

B. Unit Commitment Problem

In this section, the linearization technique is applied to the problem of unit commitment to show the impact of the linearization on computational efficiency and quality. The unit commitment problem is formulated here as follows (without loss of generality, some of the constraints are not considered here for simplicity).

$$\text{Minimize} \quad C_T = \sum_{i \in N_g} \sum_{h \in N_t} \left(z_{i,h} + \text{SDn}_{i,h} + \text{SUp}_{i,h} \right) \tag{8}$$

subject to the following operational constraints:

1) Active Power Flow Equations:

$$P_{i,h} - P_{d_{i,h}} = \sum_{j \in N_b} \bar{B}_{ij}\delta_j \tag{9}$$

2) Line Flow Limits:

$$-P_i^L \le \sum_{j \in N_b} \bar{D}_{i,j}\delta_j \le P_i^L \quad , \quad i \in N_l \tag{10}$$

3) Generation Limits:

$$P_i^{\min} u_{i,h} \le P_{i,h} \le P_i^{\max} u_{i,h} \tag{11}$$

4) Shutdown/Startup Costs:

$$\text{SUp}_{i,h} \ge (u_{i,h} - u_{i,h-1}) C_i^{\text{SUp}} \ , \ \text{SUp}_{i,h} \ge 0 \tag{12}$$

$$\text{SDn}_{i,h} \ge (u_{i,h-1} - u_{i,h}) C_i^{\text{SDn}} \ , \ \text{SDn}_{i,h} \ge 0 \tag{13}$$

5) Ramp Limits:

$$\begin{aligned} P_{i,h} - P_{i,h-1} \le [u_{i,h} - u_{i,h-1}]P_i^{\text{SUp}} + u_{i,h-1}R_i^{\text{Up}} \\ + [1 - u_{i,h}]P_i^{\max} \end{aligned} \tag{14}$$

$$\begin{aligned} P_{i,h-1} - P_{i,h} \le [u_{i,h-1} - u_{i,h}]P_i^{\text{SDn}} + u_{i,h}R_i^{\text{Dn}} \\ + [1 - u_{i,h-1}]P_i^{\max} \end{aligned} \tag{15}$$

6) System Reserve:

$$\sum_{i \in N_g} \left(u_{i,h}P_i^{\max} - P_{i,h} \right) \ge P_h^{Res} \tag{16}$$

7) Auxiliary Constraints: These constraints correspond to the PWL of the quadratic terms in the objective.

$$z_{i,h} \ge \alpha_{i,s}P_{i,h} + \beta_{i,s}u_{i,h} \tag{17}$$

Note that when $u_{i,h} = 0$, it also follows that $P_{i,h} = 0$. Therefore, all the s inequalities turn to be $z_{i,h} \ge 0$, which the solver chooses the zero value. It is trivial to analyze the case of $u_{i,h} = 1$.

C. Numerical Results

In this section, the performance of the PWL approximation (i.e. the quality of the solution and the efficiency of the solution process) is presented through two examples. A Six-bus system and the IEEE 118-bus test system are employed here for which the unit commitment problem is solved. The system data can be found in [9]. In the six-bus system, there are 3 generators and 7 branches. For the IEEE 118-bus system, there are 54 generators and 186 branches. Two approaches are used to solve the unit commitment problem. In the first approach, the problem is formulated using the original quadratic cost functions, which leads to a mixed-integer quadratic programming (MIQP) problem. There are commercial solvers capable of handling MIQP problems, e.g. CPLEX [10]. In the second approach, the problem is formulated using the PWL cost functions and associated extra constraints, as given in Section III-B7. This leads to an MILP problem, for which there are efficient commercial solvers available, e.g. CPLEX [10]. All the problems are formulated in GAMS [8] and solved using CPLEX [10] in this paper.

Table II
UNITS' SCHEDULES OBTAINED BY MIQP FOR THE SIX-BUS SYSTEM.

Hour	1	2	3	4	5	6	7	8
G_1	104.6	100	100	100	100	100	103.7	106
G_2	-	-	-	-	-	-	-	-
G_6	84.2	78.1	71.1	66.8	67.2	73	83.2	85.5
Hour	9	10	11	12	13	14	15	16
G_1	110.9	123.1	146.5	125	128.3	129.1	131.9	135.6
G_2	-	-	-	29.5	32.8	33.6	36.4	40.1
G_6	90.4	100	100	100	100	100	100	100
Hour	17	18	19	20	21	22	23	24
G_1	135.7	130.7	130.3	125.7	125.7	123.2	115.9	115.7
G_2	40.2	35.2	34.8	30.2	30.2	27.7	-	-
G_6	100	100	100	100	100	100	95.4	95.2

Table III
UNITS' SCHEDULES OBTAINED BY MILP FOR THE SIX-BUS SYSTEM.

Hour	1	2	3	4	5	6	7	8
G_1	100	100	100	100	100	100	100	100
G_2	-	-	-	-	-	-	-	-
G_6	88.9	78.1	71.1	66.8	67.2	73	86.9	91.5
Hour	9	10	11	12	13	14	15	16
G_1	101.4	123.1	146.5	124	124	124	124	130
G_2	-	-	-	30.5	37.1	38.6	44.3	46
G_6	100	100	100	100	100	100	100	100
Hour	17	18	19	20	21	22	23	24
G_1	130	124	124	124	124	124	111.2	110.9
G_2	46	42	41.2	31.9	31.8	26.8	-	-
G_6	100	100	100	100	100	100	100	100

For the PWL procedure, four segments are chosen, i.e. $s = 4$. Table II shows the units' schedules obtained considering the exact quadratic cost functions (MIQP). Table III shows the units' schedules for the six-bus system obtained using the linearized objective (MILP). As can be seen, the unit status ("on"/"off") for both methods are identical. Also, the differences between the committed generations are minor. The computational efficiency of the two methods are compared in Table IV (obtained using an Intel Core i7-2600 CPU @ 3400 MHz). The "relative gap" is defined as the relative gap between the best objective achieved up to the current iteration and the best lower bound. For the MIQP case, the solver could not reach to the zero relative gap within the time limit of 1000 s, while for the MILP case, the proven optimal solution has been achieved within a few seconds. The objective value for the six-bus system obtained using the MILP is higher than the one obtained by the MIQP. On the other hand, this is the other way around for the IEEE 118-bus system. These results reveal that both methods would come up with approximately same objective values.

It is worthwhile to compare the problem size with the method proposed in [3]. The number of extra variables required for PWL approximation for each generator in [3] is $s + 1$, while in the min-max formulation, only one extra

Table IV
COMPUTATIONAL EFFICIENCY OF MIQP AND MILP

System	Parameter	MIQP	MILP
Six-bus	CPU Time (s)	0.63	0.13
	Objective Value ($/h)	153880	153975
	Relative Gap	0	0
118-bus	CPU Time (s)	1000	7.8
	Objective Value ($/h)	650433	650387
	Relative Gap	0.026	0

variable is needed. In addition, the number of constraints for each generator in [3] is $2s + 2$, while in the min-max formulation, only s constraints suffice.

The applications of the PWL approximation is not limited to the unit commitment problem. There are many other problems in power system optimization which have a nonlinear objective function subject to some linear constraints. For example, optimal power flow with objectives such as minimizing the cost or active losses is one good example, especially when there is a need for multiple run [11]. As another instance, optimal transmission line switching for congestion management could be another application [12].

IV. CONCLUSION

The nonlinear functions appearing in the objective function of minimization problems are shown to be efficiently linearizable using a piecewise linearization (PWL) technique. The superiority of this method over existing PWL techniques are demonstrated through examples. The main advantages of the introduced approach can be summarized as follows:

- Higher accuracy in linearization is achieved by applying the max-affine PWL technique.
- The size of the subspaces (on which the linear approximations are defined) are selected optimally by this method.
- The method is able to linearize a multivariate function in multi-dimensional space.
- If the nonlinear function is in the objective, the advantage can be taken of by minimizing a linear function subject to a few affine inequalities.
- Significant saving in computation time can be achieved when transforming a mixed-integer nonlinear programming problem (with only the objective being nonlinear) to a linear version.

As future work, further research is undertaken to apply this competent technique to other areas of power system optimization.

REFERENCES

[1] Y.-C. Wu, A. S. Debs, and R. E. Marsten, "A direct nonlinear predictor-corrector primal-dual interior point algorithm for optimal power flows," *IEEE Trans. Power Syst.*, vol. 9, no. 2, pp. 876–883, May 1994.

[2] A. A. Sousa, G. L. Torres, and C. A. Cañizares, "Robust optimal power flow solution using trust region and interior-point methods," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 487–499, May 2011.

[3] M. Carrion and J. M. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1371–1378, Aug. 2006.

[4] A. Frangioni, C. Gentile, and F. Lacalandra, "Tighter approximated MILP formulations for unit commitment problems," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 105–113, Feb. 2009.

[5] A. Viana and J. P. Pedroso, "A new MILP-based approach for unit commitment in power production planning," *Int. Jour. Elect. Power & Energy Syst.*, vol. 44, no. 1, pp. 997–1005, Jan. 2013.

[6] L. Wu, "A tighter piecewise linear approximation of quadratic cost curves for unit commitment problems," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2581–2583, Nov. 2011.

[7] A. Magnani and S. Boyd, "Convex piecewise-linear fitting," *Optimization and Engineering*, vol. 10, no. 1, pp. 1–17, 2009.

[8] "The General Algebraic Modeling System (GAMS)." [Online]. Available: http://www.gams.com

[9] Y. Fu, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with AC constraints*," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1538–1550, Aug. 2005.

[10] "IBM ILOG CPLEX optimization studio." [Online]. Available: http://www-01.ibm.com/software/integration/optimization/cplex-optimization-studio/

[11] H. Ahmadi and H. Ghasemi, "Probabilistic optimal power flow incorporating wind power using point estimate methods," in *10th Int. Conf. Envir. Elec. Eng. (EEEIC)*, Rome, Italy, May 2011, pp. 1–5.

[12] M. Khanabadi, H. Ghasemi, and M. Doostizadeh, "Optimal transmission switching considering voltage security and N-1 contingency analysis," *IEEE Trans. Power Syst.*, vol. Early Access, no. 99, pp. 1–9, 2012.