POWER

PWP-031

Coordinated Multilateral Trades for Electric Power Networks: Theory and Implementation

Felix F. Wu and Pravin Varaiya

June 1995

This paper is part of the working papers series of the Program on Workable Energy Regulation (POWER). POWER is a program of the University of California Energy Institute, a multicampus research unit of the University of California, located on the Berkeley campus.

University of California Energy Institute
2539 Channing Way
Berkeley, California 94720-5180
www.ucei.berkeley.edu/ucei

$$N_k(\bar{\boldsymbol{q}} + \Delta\bar{\boldsymbol{q}}(t)) = \left(\frac{\partial \boldsymbol{F}}{\partial\theta}(\theta(t))\right)^{-T}\left(\frac{\partial \boldsymbol{g}_k}{\partial\theta}\right)(\theta(t)) \tag{91}$$

and $\theta(t)$ is the solution of the power flow equations

$$\boldsymbol{F}(\theta(t)) = \bar{\boldsymbol{q}} + \Delta\bar{\boldsymbol{q}}(t) \tag{92}$$

With these modifications, let us run down the list of previous results to check their validity. Under the convexity assumption of the feasible set for nonlinear flows, Lemma 2 will be valid. Using the modified definition of FD trades, Lemma 3 follows from the property of the geodesic curve. The proof of Lemma 4 is based on Kuhn-Tucker optimality conditions and is valid if $\boldsymbol{n}_\mathrm{k}$ is replaced by $\boldsymbol{N}_\mathrm{k}$. Lemma 5 uses only the property of the cost function c and is therefore valid for nonlinear flows. Lemmas 6 and 7 are the results of convexity and are therefore true. Consequently, Theorem 1 is still true for the nonlinear power flows.

Remark. The above definition of a FD trade is not practical from computational point of view. We may use linear approximation at the operating point to define a “practical” FD trade, as a trade along the tangent plane at $\boldsymbol{q}$ of the manifold (or surface) defined by eq. (73). More precisely, we say a (multilateral) trade $\Delta\boldsymbol{q} = (\Delta q_0, \Delta\bar{q})$ is a feasible-direction (FD) trade at $\boldsymbol{q}$, if

$$<\boldsymbol{N}_k(\bar{\boldsymbol{q}}), \Delta\bar{\boldsymbol{q}}> \le 0 \quad \text{for } k = k_1, k_2, \ldots, k_m \tag{93}$$

where

$$\boldsymbol{N}_k(\bar{\boldsymbol{q}}) = \left(\frac{\partial \boldsymbol{F}}{\partial\theta}(\theta)\right)^{-T}\left(\frac{\partial \boldsymbol{g}_k}{\partial\theta}\right)(\theta) \tag{94}$$

and θ is the solution of the power flow equations $\boldsymbol{F}(\theta) = \bar{\boldsymbol{q}}$

Strictly speaking, as a result of the trade $\Delta\boldsymbol{q}$ the new operating point will move outside of the feasible set. There are two practical fixes to this problem. The first one is to rely on curtailment to enforce feasibility. First we argue that the result is still “feasible” from practical point of view if the trade $\Delta\boldsymbol{q}$ is small. As stated earlier, the constraint equations are usually derived with built-in safety margins, they are soft constraints. Therefore, a FD trade, if it is small, may result in an operating point that is practically acceptable. If a trade is too large, resulting in significant violation of the line flow limit, it will be curtailed and a new loading sensitivity vector $\boldsymbol{N}_\mathrm{k}$ will be used. This is in fact a approach using piecewise linear approximation of the nonlinear equations. The second one is to modify the definition of a FD trade to force it to move inward to the feasible set by requiring

$$<\boldsymbol{N}_k(\bar{\boldsymbol{q}}), \Delta\bar{\boldsymbol{q}}> \le \varepsilon \quad \text{for some } \varepsilon < 0, \quad k = k_1, k_2, \ldots, k_m \tag{95}$$

The amount ε is estimated from the errors of the linear approximation to the power flow equations.

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{bmatrix} = \begin{bmatrix} \boldsymbol{T}(\boldsymbol{T}(\boldsymbol{u}, \boldsymbol{v}), \boldsymbol{u}) \\ \boldsymbol{T}(\boldsymbol{u}, \boldsymbol{v}) \end{bmatrix}$$

If the function Ψ is a contraction mapping, then the Loss Adjustment Process converges. Intuitively, if the losses are small and the quadratic approximation of the power flow equations is reasonably good, the process should converge. Both requirements depend on the R/X (resistance-reactance ratio) of the transmission system. We believe that the following conjecture is true, similar to the case in the convergence of power flow solutions.[29]

Conjecture Under reasonable conditions on transmission system R/X ratio, the Loss Allocation Process converges to a set of loss-included trades.

8.7 Nonlinear Flows

The assumption that the line flows are linear functions of the nodal voltage angles was actually relaxed in the last section on transmission losses. There are two other issues related to nonlinearity of power flow equations where modifications are necessary. One is purely technical, that is, the assumption of convexity of the constraint set and the second one involves the definition of the feasible direction. We will show in this section one theoretical and two practical solutions to the second issue.

In the nonlinear case, we need to assume that the feasible set S defined by eqs. (5)-(6) is convex. The convexity assumption is fundamental in all our derivations. Examples of nonconvex power flow equations under certain operating conditions may be found. However, we shall make the convexity assumption nonetheless to make the analytic study tractable.[30]

With nonlinear power flow eq. (6), the definition of FD trades needs modification. A precise definition of a FD trade is possible, using the idea of geodesics on the constraint manifold.[31]

Definition. Suppose that $\boldsymbol{q} \in \partial S$, i.e., eqs. (5)-(6) are satisfied and

$$g_k(\theta) = l_k \quad \text{for } k = k_1, k_2, \ldots, k_m \tag{88}$$

A (multilateral) trade $\Delta\boldsymbol{q} = (\Delta q_0, \Delta\bar{q})$ is a feasible-direction (FD) trade at $\boldsymbol{q}$, if there exists a curve

$$\Delta\boldsymbol{q}(t), 0 \le t \le 1, \text{ such that } \Delta\boldsymbol{q}(0) = \boldsymbol{0} \text{ and } \Delta\boldsymbol{q}(1) = \Delta\boldsymbol{q} \text{ and} \tag{89}$$

$$<\boldsymbol{N}_k(\bar{\boldsymbol{q}} + \Delta\bar{\boldsymbol{q}}(t)), \Delta\dot{\bar{\boldsymbol{q}}}(t)> \le 0 \quad \text{for } k = k_1, k_2, \ldots, k_m \tag{90}$$

where

[29] Similar techniques have been used for the convergence of power flow solutions, see: Wu, F. F., "Theoretical study of the convergence of the fast decoupled load flow," *IEEE Power App and Systems*, vol. 96, 1977, pp. 268-275.

[30] See: A. Arapostathis and P. Varaiya, "Behaviour of three-node power networks," *Electric Energy and Power Systems*, vol 5 (1), 1983, pp. 22-30.

[31] The basic idea is from the "gradient projection method" in nonlinear programming, see: Luenberger, D. G., *Linear and Nonlinear Programming*, Second Edition, Addison-Wesley, 1984, pp. 330-359.

(i) The broker can increase the generation of the generator or reduce the demand of the load already in the group. So, $\left(\alpha_k^m\right)_i$ is equal to zero if $q_i^m[k]$ is zero, and $\left(\alpha_k^m\right)_i$ and $q_i^m[k]$ have the same sign if $q_i^m[k]$ is nonzero.

(ii) The total adjustment in generation is equal to 100%, i.e., $<\alpha_k^m, \mathbf{1}> = 1$

With this rule set by each broker, the new injections for each trade become,

$$\Delta\boldsymbol{q}^m[k] = \alpha_k^m \Delta q_L[k] \tag{83}$$

$$\Delta\boldsymbol{q}^{m+1} = \sum_k \Delta q^{m+1}[k] \tag{84}$$

set m=m+1 and go to Step 2.

Step 4. (Termination)

Stop.

It is easy to see that if the Loss Allocation Process converges, the resulting set of trades will be a set of loss-included trades. The corresponding set of injections will be consistent and each trade takes care of its share of the losses. The convergence of the Loss Allocation Process can be studied using the standard technique of Fixed Point Theorem. Let us view the above procedure as a function or mapping ***T*** defined on the 2x(n+1)xK Euclidean space, where K is the total number of trades. First we let:

$$\boldsymbol{x} = \left[\boldsymbol{q}[1] \ \ldots\ldots \ \boldsymbol{q}[K]\right] \tag{85}$$

and we define the mapping ***T*** as follows:

$$\boldsymbol{x}^{m+1} = \boldsymbol{T}\left(\boldsymbol{x}^m, \boldsymbol{x}^{m-1}\right) \tag{86}$$

$$\boldsymbol{q}^{m+1}[k] = \boldsymbol{q}^m[k] + \alpha_k^m \tau[k] \tag{87}$$

where

$$\tau[k] = \left(<\boldsymbol{L}(\boldsymbol{F}^{-1}(\bar{\boldsymbol{q}}^{m-1})), \bar{\boldsymbol{q}}^m[k] - \bar{\boldsymbol{q}}^{m-1}[k]> \quad \frac{f_0(\boldsymbol{F}^{-1}(\bar{\boldsymbol{q}}^m)) - q_0^m}{<\boldsymbol{L}(\boldsymbol{F}^{-1}(\bar{\boldsymbol{q}}^{m-1})), \sum_k (\bar{\boldsymbol{q}}^m[k] - \bar{\boldsymbol{q}}^{m-1}[k])>} \right)$$

Consider the mapping Ψ: (**y**, **z**) = Ψ (**u**, **v**), defined through ***T*** as follows:

Let us first note that the loss allocation formula Eq. (70) was derived from the quadratic expansion of the power flow equations. The total losses to be allocated are $\sum_k \langle \boldsymbol{L}(\theta), \bar{\boldsymbol{q}}[k] \rangle$, which is indeed equal to the discrepancy in slack bus generation $f_0(\theta) - \sum_k q_0[k]$, if power flow equations were truly quadratic. In other words, the loss allocation scheme can be viewed as a way to allocate the slack bus discrepancy. When there is no discrepancy (i.e., when the set of injections is consistent), the allocation is complete. We are going to use the precise power flow equations for checking consistency of injections and use the quadratic approximations for loss allocation. We will be allocating the discrepancy in slack bus generation pro-rata to individual trades according to the formula Eq. (70), derived from the quadratic approximation.

Loss Adjustment Process

Step 1. (Initialization)

Set $\boldsymbol{q} = [0 \; .... \; 0]$, $\theta = [0 \; .... \; 0]$, and $\Delta\boldsymbol{q}^0[k] = \boldsymbol{q}^0[k]$ from the lossless trade $\boldsymbol{q}^0[k]$. Set $\Delta\boldsymbol{q}^0 = \sum_k \boldsymbol{q}^0[k]$ and m = 0.

Step 2. (Consistency)

$$\boldsymbol{q}^{m+1} = \boldsymbol{q}^m + \Delta\boldsymbol{q}^m \tag{79}$$

Find a power flow solutions and calculate total system loss discrepancy.

$$\boldsymbol{F}(\theta^{m+1}) = \bar{\boldsymbol{q}}^{m+1} \tag{80}$$

$$\Delta q_L^{m+1} = f_0(\theta^{m+1}) - q_0^{m+1} \tag{81}$$

If $\Delta q_L^{m+1} < \varepsilon$, go to Step 4, otherwise go to Step 3.

Step 3. (Allocation)

Allocation system loss discrepancy to individual trades according to

$$q_L^m[k] = \langle \boldsymbol{L}(\theta^m), \bar{\boldsymbol{q}}^m[k] \rangle \frac{\Delta q_L^{m+1}}{\langle \boldsymbol{L}(\theta^m), \sum_k \bar{\boldsymbol{q}}^m[k] \rangle} \tag{82}$$

The allocated losses for each trade is assigned or distributed to the generators and/or load participating in the trade by the broker using some rule. This may be described by a vector α_k^m as follows:

The closer L is to 1, the more accurate is the approximation. The following loss allocation formula allocates q_L to individual trade $\boldsymbol{q}[k]$ according to Eq. (73).

$$q_L[k] = (\langle R(0), \bar{\boldsymbol{q}}[k] \rangle + (\sum \bar{\boldsymbol{q}}[m])^T \boldsymbol{Q}(0) \bar{\boldsymbol{q}}[k]) L \tag{76}$$

We summarize the foregoing discussion into a corollary.

Corollary 3. If the total system losses is known, then the allocation scheme of Eq. (76) is exact when the loss approximation factor $L = 1$ otherwise it is an approximation. Let $\boldsymbol{q}[k]$ and $\boldsymbol{q}[m]$ be two simultaneous trades in a set of trades on the system. The trade $\boldsymbol{q}[m]$ causes a change in transmission losses allocated to the trade $\boldsymbol{q}[k]$ by an amount which is approximately equal to $\frac{1}{2}(\bar{\boldsymbol{q}}[m])^T \boldsymbol{Q}(0) \bar{\boldsymbol{q}}[k] L$ and the losses caused by $\boldsymbol{q}[k]$ alone is approximately equal to $\left(\langle R(0), \bar{\boldsymbol{q}}[k] \rangle + \frac{1}{2}(\bar{\boldsymbol{q}}[k])^T \boldsymbol{Q}(0) \bar{\boldsymbol{q}}[k] \right) L$

Similarly, if the loss allocation formula Eq. (70) is used and the total losses q_L is known, we may define a loss approximation factor

$$L^1 = \frac{q_L}{\sum_k \langle \boldsymbol{L}(\theta), \bar{\boldsymbol{q}}[k] \rangle} \tag{77}$$

and allocate the total losses q_L according to

$$q_L[k] = \langle \boldsymbol{L}(\theta), \bar{\boldsymbol{q}}[k] \rangle L^1 \tag{78}$$

Remark. The Hessian matrices are symmetrical, therefore for two trades $\boldsymbol{q}[k]$ and $\boldsymbol{q}[m]$, the effect on losses are mutual and equal. Another interesting implication of this approximate formula is that the order in which the trades are entered is immaterial. Corollary 2 provides a foundation for parties to negotiate compensation for the losses that one trade inflicts on others.

3. Loss-included trades

From a practical point of view, given a set of planned (lossless) trades, we can use the loss allocation formula in Theorem 2 (or Corollary 1) to determine the share of losses for each trade. The broker of each trade then arranges to have its share of losses be taken care of by either increasing the generation or reducing the load. The result should be satisfactory. From a theoretical point of view, however, once you adjust the generation and/or the load, the transmission losses in the network change and the set of injections may be not be consistent any more. An interesting theoretical question arises: Is it really possible to define a loss-included trade? We are going to answer this question of theoretical interest in this section. We will suggest an iterative scheme to tackle this problem.

$$\boldsymbol{q}_L = \langle R(0), \sum \bar{\boldsymbol{q}}[k] \rangle + \frac{1}{2}\theta^T \boldsymbol{H}(0)\left[\frac{\partial \boldsymbol{F}}{\partial \theta}(0)\right]^{-1} \sum \bar{\boldsymbol{q}}[k] \tag{72}$$

Eq. (70) follows immediately.

Remark. 1). We shall refer to the vector $\boldsymbol{L}(\theta)$ defined in Eq. (71) as the Loss vector.

2). Since the loss allocation formula Eq. (70) depends only on $\bar{\boldsymbol{q}}[k]$, it can be applied directly to lossless trades without the need for actually carrying out the modification of the slack bus injection as described in the beginning of this subsection.

Instead of applying Eq. (58), we may use the alternative approximation Eq. (48) and obtain a slightly different result, which we state as a corollary.

Corollary 2. The transmission losses caused by a trade can be further approximated by

$$\boldsymbol{q}_L[k] = \langle R(0), \bar{\boldsymbol{q}}[k] \rangle + \left(\sum \bar{\boldsymbol{q}}[m]\right)^T \boldsymbol{Q}(0)\, \bar{\boldsymbol{q}}[k] \tag{73}$$

The significance of Corollary 2 is two-fold. First, this approximation of losses caused by a particular trade in this formula does not depend on the operating point θ any more. It is calculated using only data on network configuration and line parameters. The approximation is a little less accurate, but much easier to implement. Second, it shows clearly the interactions of trades in terms of their effect on losses. Eq. (73) has two terms, $\langle R(0), \bar{\boldsymbol{q}}[k] \rangle + (\bar{\boldsymbol{q}}[k])^T \boldsymbol{Q}(0)\, \bar{\boldsymbol{q}}[k]$, which depend solely on the trade itself and the other (K-1) terms, $(\bar{\boldsymbol{q}}[m])^T \boldsymbol{Q}(0)\, \bar{\boldsymbol{q}}[k]$, which depend on the other (K-1) trades. The latter terms represent added burden or benefit, as the case may be, other trades place on the losses allocated to the trade k.

We shall refer to the matrix $\boldsymbol{Q}(0)$ as the quadratic loss matrix. A further approximation of Eq. (73) can be done by ignoring the first term, which is almost 0, and the estimation of losses involves only the quadratic loss matrix.

Eq.(73) can be viewed as serving two functions. First, it estimates the total system losses

$$q_L = \sum_k q_L[k] = \sum_k \left(\langle R(0), \bar{\boldsymbol{q}}[k] \rangle + \left(\sum_m \bar{\boldsymbol{q}}[m]\right)^T \boldsymbol{Q}(0)\, \bar{\boldsymbol{q}}[k] \right) \tag{74}$$

Second, it divides this total loss among the individual trades. Using Eq. (74), which is independent of the operating point, to estimate system losses is an approximation whose accuracy may not be good enough in cases, for example, where the operating point θ is significantly different from 0. When the exact losses are known, we may utilize Eq. (73) purely for its allocation function. Let us define a loss approximation factor

$$L = \frac{q_L}{\sum_k \left(\langle R(0), \bar{\boldsymbol{q}}[k] \rangle + \left(\sum_m \bar{\boldsymbol{q}}[m]\right)^T \boldsymbol{Q}(0)\, \bar{\boldsymbol{q}}[k] \right)} \tag{75}$$

$$F(\theta) = \bar{q} \tag{65}$$

The total system transmission losses are given by

$$q_L = \sum_i q_i \tag{66}$$

For loss-included trades, the losses caused by each trade is given by

$$q_L[k] = \sum_i q_i[k] \tag{67}$$

The transmission losses due to an individual trade, $q_L[k]$, is a function of both the trade $\boldsymbol{q}[k]$ and the operating point θ, i.e.,

$$q_L[k] = q_L(\boldsymbol{q}[k], \theta) \tag{68}$$

If we can find a function $q_L(\boldsymbol{q}[k], \theta)$ satisfying Eq. (68) and such that it also results in the following relation

$$q_L = \sum_k q_L(\boldsymbol{q}[k], \theta) \tag{69}$$

then the function $q_L(\boldsymbol{q}[k], \theta)$ is a way to allocate transmission losses to individual trades and we say Eq. (68) is a formula for transmission loss allocation. Using the quadratic approximation, we can indeed derive a loss allocation scheme, as stated in the following theorem. If the power flow equations were exactly quadratic, the loss allocation scheme suggested here would be the only one and there is no other way of allocating the losses. Since the quadratic approximation for losses is fairly accurate, we believe the proposed allocation scheme is practically the only reasonable one. Any other reasonable allocation scheme will not be very much different from the proposed one. The difference occurs only in the amount one allocates for the losses attributed to approximation error.

Theorem 2. The transmission losses caused by a trade as part of the total injections under the operating point θ is approximately equal to

$$q_L[k] = \langle \boldsymbol{L}(\theta), \bar{\boldsymbol{q}}[k] \rangle \tag{70}$$

where

$$\boldsymbol{L}(\theta) = \boldsymbol{R}(0) + \frac{1}{2}\left(\frac{\partial \boldsymbol{F}}{\partial \theta}(0)\right)^{-T} \boldsymbol{H}(0)\,\theta \tag{71}$$

Proof. Applying Eq.(58) in Lemma 8 with the operating point $\boldsymbol{q}$=0 and θ=0 and the “additional” trade $\sum_k q[k]$, we obtain

The incremental loss formula Eq. (48) that we derived provides a superior approximation in most cases. The quadratic term is particularly significant in situations where (i) the additional trade $\Delta\boldsymbol{q}[k]$ is not very small, and (ii) the operating point θ used in the formula is itself an approximation.

Remark. From the point of view of practical implementation for large systems, matrix sparsity is a crucial consideration. Without sparsity, computation may become simply infeasible. Both the Jacobian matrix $\frac{\partial \boldsymbol{F}}{\partial \theta}$ and the Hessian matrices $\boldsymbol{H}_i(\theta)$ are sparse for power systems. Inverting a large matrix such as the one in Eq. (49)-(50), however, may not be practical. But that is really not necessary, as the second term in Eq. (49) can be obtained by the solution of the equation $\frac{\partial \boldsymbol{F}}{\partial \theta}\boldsymbol{x} = \frac{\partial \boldsymbol{f}_0}{\partial \theta}(\theta)$ and the solution of $\frac{\partial \boldsymbol{F}}{\partial \theta}\boldsymbol{y} = \Delta\bar{\boldsymbol{q}}$ may be used in computing the quadratic term in Eq. (48).

Lemma 8 is valid for any set of additional power injections $\Delta\boldsymbol{q} = (\Delta q_0, \Delta\bar{\boldsymbol{q}})$, as long the resulting set of injections $(q_0 + \Delta q_0, \bar{\boldsymbol{q}} + \Delta\bar{\boldsymbol{q}})$ is consistent. Therefore it is also applicable to calculating incremental losses for a loss-included trade. However, in such a case, the problem of determining incremental losses and the problem of defining loss-included trades are related. We suggest an iterative process to solve the two problems simultaneously. The proposed approach has two steps. The first step is to start with a lossless trade and determine the losses it causes, which is then added to the trade. The second step is to check discrepancy in the slack bus generation (consistency of the power flow equations). The necessary adjustment as a result of additional losses is added to the trade again. This iterative process continues until no more discrepancy is present. A more general scheme for loss-included trade, using a similar approach, is proposed later. More details are provided there. As in the general case, we actually expect the result of the first step to give a fairly good approximation for most practical cases and the iterative process included here is more for theoretical completeness.

2. Loss allocation

We will formulate the loss allocation problem more generally for any set of loss-included trades. To apply the result to lossless and other types of trades, we modify the trades to make them loss-included by (i) augmenting the slack bus generator as a participant in every trade, (ii) if the original trade does not include the slack bus generator, the slack bus injection is equal to the losses the trade causes, and (iii) if the original trade includes the slack bus, its injection is modified to include the losses the trade causes.

Consider a set of loss-included trades $\sum_k q[k]$ resulting in a set of consistent injections $\boldsymbol{q} = \sum_k q[k]$. The power balance equations Eq.(43)-(44) are satisfied, i.e.,

$$f_0(\theta) = q_0 \tag{64}$$

Substituting Eqs. (52)-(55) into Eq. (47), one obtains

$$\Delta \boldsymbol{q}_L = \left\langle \frac{\partial f_0}{\partial \theta}(\theta), \Delta\theta \right\rangle + \frac{\partial \boldsymbol{F}}{\partial \theta}(\theta)\,\Delta\theta + \frac{1}{2}(\Delta\theta)^T \boldsymbol{H}(\theta)\,\Delta\theta \tag{56}$$

Now using the linear approximation of $\Delta\theta$ from

$$\frac{\partial \boldsymbol{F}}{\partial \theta}\Delta\theta = \Delta\bar{\boldsymbol{q}} \tag{57}$$

into the first term in Eq. (56), one has

$$\Delta \boldsymbol{q}_L = \langle R(\theta), \Delta\bar{\boldsymbol{q}} \rangle + \frac{1}{2}(\Delta\theta)^T \boldsymbol{H}(\theta)\,\Delta\theta \tag{58}$$

Applying the linear approximation of $\Delta\theta$ again in the second term of Eq. (58), one obtains Eq. (48).

Remark. The vector $\boldsymbol{R}$ defined in eq. (49) is the contribution to transmission losses from the trade using only linear approximation. It is directly related to the penalty factors in the classical economic dispatch. Without transmission losses the optimality condition requires that all generators operate at the same marginal cost. The penalty factor L_i for generator i in a lossy system is defined with respect to the reference bus 0 (the slack bus) as a weighting factor added to the generator marginal cost to “penalize” generators for their effect on losses,

$$L_i \frac{\partial c_i}{\partial q_i} = \frac{\partial c_0}{\partial q_0} \qquad i = 1, 2, ... \tag{59}$$

Comparing Eq. (59) and Eq. (16), it follows that $L_i = p_0/p_i$. Applying the optimality condition Eq. (17) with all μ_k equal to zero (no congestion), the penalty factors can be calculated by first solving

$$\frac{\partial \boldsymbol{F}^T}{\partial \theta}\beta = -\frac{\partial \boldsymbol{f}_0}{\partial \theta} \tag{60}$$

and then setting

$$L_i = \frac{1}{\beta_i} \qquad i = 1, 2, ... \tag{61}$$

The vector $\boldsymbol{R}$ is therefore equal to

$$\boldsymbol{R} = \mathbf{1} - \beta \;\; or \tag{62}$$

$$R_i = 1 - \frac{1}{L_i} \qquad i = 1, 2, ... \tag{63}$$

1. Incremental losses

Given a set of consistent injections $(q_0, \bar{\boldsymbol{q}})$, consider an additional trade $\Delta\boldsymbol{q} = (\Delta q_0, \Delta\bar{\boldsymbol{q}})$ such that $(q_0 + \Delta q_0, \bar{\boldsymbol{q}} + \Delta\bar{\boldsymbol{q}})$ is consistent. Assume that the system was initially operating at the operating point θ and the total system transmission losses were

$$q_L = \sum_i q_i = q_o + \langle \mathbf{1}, \bar{\boldsymbol{q}} \rangle \tag{46}$$

The additional trade results in incremental transmission losses

$$\Delta\boldsymbol{q}_L = \Delta q_0 + \langle \mathbf{1}, \Delta\bar{\boldsymbol{q}} \rangle \tag{47}$$

The following lemma provides a formula to calculate the incremental losses.

Lemma 8. The incremental loss caused by a trade $\Delta\boldsymbol{q}$ for a system operated at θ can be approximated by

$$\Delta\boldsymbol{q}_L = \langle R(\theta), \Delta\bar{\boldsymbol{q}} \rangle + \frac{1}{2}(\Delta\bar{\boldsymbol{q}})^T \boldsymbol{Q}(\theta)\Delta\bar{\boldsymbol{q}} \tag{48}$$

where

$$\boldsymbol{R}(\theta) = \mathbf{1} + \left(\frac{\partial \boldsymbol{F}}{\partial\theta}(\theta)\right)^{-T}\left(\frac{\partial \boldsymbol{f}_0}{\partial\theta}(\theta)\right) \tag{49}$$

$$\boldsymbol{Q}(\theta) = \left(\frac{\partial \boldsymbol{F}}{\partial\theta}(\theta)\right)^{-T}\boldsymbol{H}(\theta)\left(\frac{\partial \boldsymbol{F}}{\partial\theta}(\theta)\right)^{-1} \tag{50}$$

$$\boldsymbol{H}(\theta) = \boldsymbol{H}_0(\theta) + \boldsymbol{H}_1(\theta) + \dots + \boldsymbol{H}_n(\theta) \tag{51}$$

and $\boldsymbol{H}_i(\theta)$ is the Hessian matrix of the i-th power flow equation $f_i(\theta) = q_i$

Proof: After $\Delta\boldsymbol{q}[k]$ is introduced, the operating point moves to $\theta + \Delta\theta$ where

$$f_0(\theta + \Delta\theta) = q_0 + \Delta q_0 \tag{52}$$

$$\boldsymbol{F}(\theta + \Delta\theta) = \bar{\boldsymbol{q}} + \Delta\bar{\boldsymbol{q}} \tag{53}$$

Taking the first three terms of the Taylor expansion of the above two equations, one has

$$f_0(\theta + \Delta\theta) = f_0(\theta) + \left\langle \frac{\partial f_0}{\partial\theta}, \Delta\theta \right\rangle + \frac{1}{2}(\Delta\theta)^T \boldsymbol{H}_0(\theta)\Delta\theta \tag{54}$$

$$f_i(\theta + \Delta\theta) = f_i(\theta) + \left\langle \frac{\partial f_i}{\partial\theta}, \Delta\theta \right\rangle + \frac{1}{2}(\Delta\theta)^T \boldsymbol{H}_i(\theta)(\Delta\theta), \ i = 1, 2, \dots n \tag{55}$$

generator included, hence the optimality of the coordinated multilateral trading process can no longer be guaranteed with this approach. However, the simplicity of this traditional approach still makes it worth including in the discussion.

A more sophisticated multilateral trading arrangement is to have each individual trade take care of its own losses. This means that each trade is required to have its total generation minus total load be equal to the transmission losses it causes. Moreover, the resulting set of injections must be consistent. This definition of a trade requires the knowledge of the transmission losses it causes. A mechanism to do that needs to be developed. The advantage of this approach is that once a trade is so defined, the rest of the derivation in the previous sections will still be valid and the optimality of the coordinated multilateral trading process is guaranteed.

It should be pointed out that this more general model includes the previous one with slack bus handling losses as a special case if we include slack bus as a participant in every trade. To help distinguishing the two cases in future discussion, we propose the following definitions.

Definition. A trade $\boldsymbol{q}[k]$ is called a *lossless trade* if $\sum_i q_i[k] = 0$. A trade $\boldsymbol{q}[k]$ is called a *loss-included trade* if the total generation minus the total load involved in the trade is equal to the transmission losses, $q_L[k]$, it causes, i.e., $\sum_i q_i[k] = \sum_{i \in G} q_i[k] - \sum_{i \in D} q_i[k] = q_L[k]$. In other words, a lossless trade is one without the transmission losses accounted for and a loss-included one is one with the transmission losses properly accounted for.

We shall first derive a formula for incremental losses caused by a trade and use the result to derive a loss allocation formula. These formulas are in fact applicable to any set of injections that are consistent. We then use them to propose an iterative process to account for losses in a loss-included trade.

Exact computation of transmission losses for a particular case is possible by solving the power flow equations. But our goal is to derive an explicit expression for the losses in order to develop a mechanism to determine losses for any arbitrary trade. The complexity of the nonlinear power flow equations precludes a closed form expression. One has to resort to some form of approximations to the power flow equations. Linear approximations have generally been used in economic dispatch and other applications involving transmission losses. It is known that the transmission loss along a line is proportional to the square of the current through the line and is therefore predominately a function of the square of the angle difference across the line. The transmission losses, as a function of the angles θ is thus at least second-order. Linear approximation will induce unacceptable errors in some cases. We shall use a quadratic approximation in our derivations.

is small, however, in terms of accumulated effect on revenue, it is significant. Any proposal for restructuring without a solution to the problems due to losses is incomplete and unacceptable.

The lossless assumption was used in Lemma 1 as a foundation for the definition of multilateral trades. From the viewpoint of mathematical modeling, the lossless assumption affects the rest of the derivation in two ways. First, it makes the power balance equation of the reference bus Eq. (13) redundant. One may then focus on the power flow equations Eq. (14) and the injections $\bar{\boldsymbol{q}}$ and $\bar{\boldsymbol{q}}[k]$. Second, in a lossless system, individual trades can easily be defined (the total generation must be equal to the total load). In a lossy system, if these two interrelated problems can be taken care of, the rest of the derivation, dealing only with the injections $\bar{\boldsymbol{q}}$ and $\bar{\boldsymbol{q}}[k]$, will be valid. We will show in this section how these two problems can be handled.

For a lossy system, first comes the problem of power flow balance or the consistency of the power flow equations (5). Let’s separate the power flow equations for the reference bus and the rest as follows:

$$f_0(\theta) = q_0 \tag{43}$$

$$\boldsymbol{F}(\theta) = \bar{\boldsymbol{q}} \tag{44}$$

where

$$\boldsymbol{F}(\theta) = \begin{bmatrix} f_1(\theta) \\ \dots \\ f_n(\theta) \end{bmatrix} \tag{45}$$

Conventional power system analysis starts with the specification of injections in all nodes except the reference node 0 (or the slack bus), i.e. specifying the injections, $\bar{\boldsymbol{q}}$ and use Eq. (44) to solve for the operating point θ. The “slack” bus generation is then determined from Eq. (44). In other words, slack bus generation is used to balance power to match injections, including the resulting losses, at other buses. We shall call a set of injections $(q_0, \bar{\boldsymbol{q}})$ satisfying Eq. (43)-(44) a consistent set. The conventional approach is one way of ensuring consistency of injections.

A multilateral trading model may be implemented based on this conventional approach. We define trades the same as in the lossless case and rely on the slack generator to provide for the losses. In other words, the slack bus generator becomes part of the trade. In such a model, the power flow balance or the consistency problem disappears. One still needs to determine the amount of losses attributed to individual trades, however. Because of interactions of power flows and nonlinear dependence of losses on power flows, the transmission losses due to an individual trade cannot be specified in isolation. To allocate the slack bus generation of total losses to individual trades is therefore operating point dependent and is nontrivial. But a more serious problem is introduced by including slack bus generator in every trade, that is, the existence of a profitable FD trade in Lemma 4 may not be a trade with slack bus

Theorem 1. Under the assumptions that all participants make reasonable decisions (assumptions (i) and (ii) above), the Coordinated Multilateral Trading Process converges to the solution of the optimal dispatch $\boldsymbol{q}^*$.

Proof: The feasibility of the initial point $\boldsymbol{q}$ is established by Lemma 2. If $\boldsymbol{q}$ is not the optimal point $\boldsymbol{q}^*$, Lemmas 4 and 5 imply that there are opportunities for parties to make profitable trades that are consistent with feasibility, i.e., if the parties follow the rules guaranteeing feasible directions. (Remember that a curtailed FD trade is also a FD trade.) Suppose such a profitable trade is identified and is made. The resulting flows on the network may cause further congestion on other lines. The power system operator will have to curtail such a profitable FD trade to make it feasible (Lemma 6). Lemma 7 establishes that the curtailed trade is still profitable. The total cost goes down as a result. On the other hand, if a profitable trade does not cause any more congestion, the total cost will of course go down by definition of being profitable. In any case, if the resulting operating point is not the optimal point, Lemmas 4 and 5 are applicable again and the process continues.

We will show that the above process will converge to the optimal solution $\boldsymbol{q}^*$. We do this in two steps. First, we claim that, under assumption (i) and (ii), the process will terminate at a point $\boldsymbol{q}^N$ which is no more than "ε away" from the optimal, i.e., $(c(\boldsymbol{q}^N) - c(\boldsymbol{q}^*)) < \varepsilon$. Next, since assumption (i) is valid for any ε, we let $\varepsilon \to 0$, then $\boldsymbol{q}^N \to \boldsymbol{q}^*$.

To prove the claim, note that the fact that the process terminates implies that

$$c(\boldsymbol{q}^N) - c(\boldsymbol{q}^N + \Delta\boldsymbol{q}^N) \le \varepsilon \qquad (42)$$

$$\text{for all } \Delta\boldsymbol{q}^N \text{ satisfying } \quad \langle \boldsymbol{n}_k, \Delta\boldsymbol{q}^N \rangle \le 0$$

Since the set $X := \{\boldsymbol{q}: \ \boldsymbol{q} = \boldsymbol{q}^N + \Delta\boldsymbol{q}^N, \ \ \langle \boldsymbol{n}_k, \Delta\boldsymbol{q}^N \rangle \le 0\}$ is convex, the convex cost function c achieves a minimum at, say, $\boldsymbol{q}^m$. Eq. (42) implies that

$$c(\boldsymbol{q}^m) \ge c(\boldsymbol{q}^N) - \varepsilon$$

The feasible set has more constraints than X has and is therefore a subset of X, $S \subseteq X$. The optimal $\boldsymbol{q}^*$ is in the feasible set, hence it must satisfy $c(\boldsymbol{q}^*) \ge c(\boldsymbol{q}^m)$. This implies that $c(\boldsymbol{q}^*) \ge c(\boldsymbol{q}^N) - \varepsilon$, i.e., $c(\boldsymbol{q}^N)$ is ε-away from the optimal. The claim is thus proved. Since assumption (i) is valid for any ε, we let $\varepsilon \to 0$, then $\boldsymbol{q}^N \to \boldsymbol{q}^*$. The validity of the Theorem is thus established.

8.6 Transmission Losses

In this section, we relax the assumption that transmission lines are lossless. Small losses are always present in transmission lines and transformers due to resistances. Total losses in a transmission system typically only amounts to 2-4% of the total generation. Percentage-wise, the amount of losses

$$\frac{n_2\left(\frac{dc_1}{dq_1}(q_1 + \Delta q_1)\right) - n_1\left(\frac{dc_2}{dq_2}(q_2 + \Delta q_2)\right)}{n_2 - n_1} = \frac{n_3\left(\frac{dc_1}{dq_1}(q_1 + \Delta q_1)\right) - n_1\left(\frac{dc_3}{dq_3}(q_3 + \Delta q_3)\right)}{n_3 - n_1} \quad (41)$$

8.5 Coordinated Multilateral Trading Process

Now we are ready to state the proposed multilateral trading process.

Coordinated Multilateral Trading Process

Step 1. (Initialization)

Brokers arrange trades $\boldsymbol{q}[k]$. Let $\boldsymbol{q}^0 = \sum_k \boldsymbol{q}[k]$.

Step 2. (Curtailment)

If $\boldsymbol{q}^0$ is not feasible, the power system operator (PSO) curtails the trades to a point where the resulting injections $\boldsymbol{q}$ are feasible.

Step 3. (Announcement)

If lines $k_1,\ldots,k_m$ are congested at $\boldsymbol{q}$, the system operator announces the Loading vectors, $\boldsymbol{n}_k,\ k = k_1, \ldots, k_m$.

Step 4. (Trading)

If a profitable trade in the feasible direction is found, a broker arranges it. The broker uses $\boldsymbol{n}_k$ to determine whether a trade is in the feasible direction. If no profitable trade is found, go to Step 6.

Step 5. (Feasibility)

If the trade is infeasible, let the power system operator curtail the trade and go to Step 3. If the trade is feasible, let PSO fulfill it and go to Step 4.

Step 6. (Termination)

Stop.

We are going to show that the coordinated multilateral trading process converges to the optimal solution under the assumption that all participants make reasonable decisions. By that we mean whenever there is a worthwhile profitable trade, the participants will carry it out. More precisely, let us call a trade $\Delta\boldsymbol{q}$ with profit less than ε, i.e., $(c(\boldsymbol{q}) - c(\boldsymbol{q} + \Delta\boldsymbol{q})) < \varepsilon$, an ε-unworthy trade and one with profit greater than ε, $(c(\boldsymbol{q}) - c(\boldsymbol{q} + \Delta\boldsymbol{q})) \geq \varepsilon$, an ε-worthy trade. We assume that (i) For any $\varepsilon > 0$ (sufficiently small), any ε-unworthy trade in the feasible direction will not be arranged and any ε-worthy trade will eventually be identified and arranged. (ii) Once a worthy profitable trade is identified, the parties involved are willing to carry it out. The following theorem asserts that the coordinated multilateral trading process converges to the solution of the optimal dispatch under such assumptions.

arrangement of today's integrated utility arrangement, the problem formulated above then becomes the so-called transmission constrained economic dispatch in today's advanced control centers.

For the case with only one transmission congestion, further guidelines for finding a feasible trilateral trade can be derived. For a trilateral trade involving, say generator 1, generator 2, and consumer 3, such that $\Delta q_1 + \Delta q_2 = \Delta q_3$, we want to derive guidelines for a broker to spot profitable trades that are feasible. The transmission system security constraint derived from the Loading vector (n_1, n_2, n_3,...) can be rewritten as:

$$\frac{\Delta q_1}{\Delta q_2} \le \frac{n_2 - n_3}{n_3 - n_1} \tag{38}$$

For a profitable trade to exist, the marginal benefit of additional MW for consumer 3 must be higher than the marginal cost of one of the generators, say generator 1, i.e., $MC_3 > MC_1$., where $MC_i = \frac{dc_i}{dq_i}(q_i)$. Suppose that generator 1 is also serving as the broker. If $n_3 \ge n_1$, $(n_1 - n_3)\Delta q_1 \le 0$ for any amount of Δq_1 MW. As the trade from node 1 to node 3 helps relieve the congestion, any profitable trade can be carried out and there is no need to solicit the participation of another generator. A more interesting situation is when $n_3 < n_1$. Assertion 3 below follows immediately upon examining the signs in Eq. (38).

Assertion 3. Suppose $MC_3 > MC_1$ and $n_3 < n_1$, hence profitable bilateral trade between generator 1 and consumer 3 is not feasible. The broker should look for a generator, say generator 2, that satisfies either of the following conditions for possible trades that are profitable. (i) $n_3 > n_2$, generator 2 to generate. (ii) $n_3 < n_2$, generator 2 to back down.

In the special case of a single transmission congestion, the amount of generation and consumption in a trilateral trade can be determined if the cost and benefit functions are explicit. Assertion 4 states the result.

Assertion 4. For a trilateral trade with one transmission congestion, the optimal trade $(\Delta q_1, \Delta q_2, \Delta q_3)$, is the solution of the equations:

$$\Delta q_1 + \Delta q_2 = \Delta q_3 \tag{39}$$

$$\frac{\Delta q_1}{\Delta q_2} = \frac{n_2 - n_3}{n_3 - n_1} \tag{40}$$

subject to

$$\sum_{i = I} \Delta q_i = 0 \tag{33}$$

$$\sum_{i = I} (\boldsymbol{n}_k)_i \Delta q_i \le 0 , \qquad k \in K = \{k_1, k_2, \ldots k_m\} \tag{34}$$

The optimization problem formulated above (32)-(34) is a version of the transmission constrained economic dispatch (TCED). If the cost function is quadratic, the resulting TCED is a simple quadratic programming problem for which efficient solution algorithms exist. The optimality conditions for the optimization problem stated above are:

$$\frac{dc_i}{dq_i}(q_i + \Delta q_i) = \lambda + \sum_{k \in K} \mu_k (\boldsymbol{n}_k)_i \qquad i \in I, i \ne 0 \tag{35}$$

$$\frac{dc_0}{dq_0}(q_0 + \Delta q_0) = \lambda \qquad i \in I, i = 0 \tag{36}$$

where λ and μ are the Lagrange multipliers of Eqs. (33)and (34), respectively. Several immediate conclusions follow from the above results, which are stated below as assertions.

Assertion 1. If (i) the optimal solution occurs at a point involving only those constraints in *K* and (ii) the additional trade involves all participants, i.e., $I = \{0, 1, 2, \ldots\ldots, n\}$, then the solution to the TCED is precisely the solution of the optimal dispatch and the marginal cost at the buses correspond to the optimal solution. The value of the optimal marginal cost at bus i relative to the marginal cost at the slack bus 0 is equal to

$$\frac{dc_i}{dq_i} - \frac{dc_0}{dq_0} = \sum_{k \in K} \mu_k (\boldsymbol{n}_k)_i \tag{37}$$

Proof: Combining Eqs. (35) and (36) we obtain Eq. (37). Comparing Eq. (37) with the optimality condition Eqs. (17) and (18), ignoring losses, the assertion follows.

A very important conclusion that can be drawn from the above formulation is on the number of parties necessary to construct a profitable trade in the feasible direction. It is stated below.

Assertion 2. If there is only one transmission congestion, i.e, m=1, a trilateral trade may be necessary in order to construct a profitable trade in the feasible direction.

Proof: At the optimal solution of the TCED, Eqs. (33)-(36) must all be satisfied. If the trade involves only two parties, Eqs (33)-(34) become two equations with two unknowns. If there is a solution, it may not satisfy Eqs. (35)-(36). Therefore, a trilateral trade may be necessary.

Of course, if more parties are involved, the better the chance a profitable solution can be found. The extreme case is that every generator and load is involved in the trading, which is precisely the

We propose the following definition.

Definition We call $\Delta\boldsymbol{q}$ a *profitable* multilateral trade at $\boldsymbol{q} \in S$ if it reduces the total cost, $c(\boldsymbol{q} + \Delta\boldsymbol{q}) < c(\boldsymbol{q})$ (i.e., increase total welfare).

If we start from a set of multilateral trades that is feasible, the parties involved (generators and consumers) then make a series of profitable trades that are always in the feasible direction, the total cost would go down at every step, until the optimal is reached. There is, however, one potential problem: the result of a profitable FD trade may not be feasible. This is because Lemma 3 only guarantees that the original constraints are satisfied, but the new trade may cause power flow in other lines to be overloaded. Therefore, further curtailment of profitable trade may be necessary to enforce feasibility.

Lemma 6 A profitable trade $\Delta\boldsymbol{q}$ at $\boldsymbol{q}$ in the feasible direction can always be made feasible by curtailment.

Proof: Suppose that additional constraints

$$\langle \boldsymbol{g}_m, \theta \rangle \le l_m, \quad m = m_1, \ldots, m_k \tag{31}$$

are violated at $(\boldsymbol{q} + \Delta\boldsymbol{q})$. Using the same argument as in Lemma 2, we can show that there exists a curtailment schedule γ such that $\gamma\Delta\boldsymbol{q}$ is on the boundary of the set defined by eq. (31).

Lemma 7 After curtailment, a profitable trade is still profitable.

Proof: Let the original profitable trade be $\Delta\boldsymbol{q}$ and the curtailed one be $(\gamma\Delta\boldsymbol{q})$. The convexity of C guarantees that

$$c(\boldsymbol{q} + \gamma\Delta\boldsymbol{q}) - c(\boldsymbol{q}) \le \gamma[c(\boldsymbol{q} + \Delta\boldsymbol{q}) - c(\boldsymbol{q})] < 0 .$$

Hence, $(\gamma\Delta\boldsymbol{q})$ is still profitable.

8.4 Trading Arrangement

We now elaborate the trading arrangement process between the generators and the consumers, with the help of a broker. The decision-making by a broker to find a profitable trade $\Delta\boldsymbol{q}$ in the feasible direction can be formulated analytically as an optimization problem. Let I be the index set of the generators and loads engaged in the trade $\Delta\boldsymbol{q}$ arranged by the broker, i.e., $I = \{i ; \Delta q_i \neq 0\}$. The objective of the broker is to maximize the net profit of the trade, or to maximize the total cost reduction $\sum_{i \in I} c_i(q_i) - c_i(q_i + \Delta q_i)$, subject to the constraints that (i) the amount of generation and load balances out, and (ii) the loading constraints (feasible direction) set by the PSO is observed. Mathematically, this can be written as

$$min \sum_{i \in I} c_i(q_i + \Delta q_i) - c_i(q_i) \tag{32}$$

(ii) $B_2 = \sum_{i \in G^-} -[c_i(q_i + \Delta q_i) - c_i(q_i)]$

Again B_2 is positive and Δq_i must be negative. B_2 is the cost saved by this group of generators by backing down generation Δq_i MW each. It is the maximum amount this group of generators is willing to contribute if they are allowed to back down their generation.

(iii) $C_3 = \sum_{i \in G^+} [c_i(q_i + \Delta q_i) - c_i(q_i)]$

C_3 is positive and Δq_i is positive. This is the cost to generate additional power by the generators in G^+. C_3 is the minimum payment they would accept to generate additional Δq_i.

(iv) $C_4 = \sum_{i \in D^+} [c_i(q_i + \Delta q_i) - c_i(q_i)]$

C_4 is positive and Δq_i is positive or the demand at bus i is decreased by $|\Delta q_i|$ MW. C_4 represents reduced benefit for the group of consumers in D^+ and is the minimum payment they would accept to reduce their consumption.

We shall demonstrate by way of an example that there is a profit to be made by whoever arranges such a trade. Let us call the person who arranges trades a broker. A broker could be a generator owner or a consumer involved in the trade, or could be the representative of all parties involved. Suppose that the consumers in D^- are willing to pay $B_1 - \frac{B}{6}$, for the additional supply. The broker would take this amount and pay the generators in G^+ the amount $C_3 + (B/6)$ to generate, and the consumers in D^+ the amount $C_4 + (B/6)$ to reduce their consumption. Furthermore, the broker would also arrange with the generators in G^- to reduce their generation and contribute $B_2 - (B/6)$ from their cost saving of B_2. With this arrangement the broker would profit $\frac{B}{3}$.

Remarks. 1). The example in the proof of Lemma 5 is used to show that there is an opportunity for a broker to make a profit if a trade $\Delta \boldsymbol{q}$ results in $c(\boldsymbol{q} + \Delta \boldsymbol{q}) < c(\boldsymbol{q})$. “Equitable” distribution of the surplus among the parties to the trade is a secondary issue and is beyond the scope of this paper.

2). It should be pointed out that there is no need for the generators in G^- to be concerned that they had been involved in previous trades to generate the current amount. Suppose that generator 1 had a trade with consumer 2 to supply q_1 MW of power. Now comes this profitable trade Δq. If consumer 2 is not part of this profitable trade, i.e., $\Delta q_2 = 0$, it will still receive q_1 MW. If consumer 2 is a party in the trade and is asked to back down, the consumer will be adequately compensated.

$$\boldsymbol{v} = -\frac{\partial c}{\partial \bar{\boldsymbol{q}}} - \boldsymbol{F}^{-T}\boldsymbol{f_0}\left(\frac{\partial c_0}{\partial q_0}\right) \tag{27}$$

To prove our claim, first of all, one notes that

$$<-\frac{\partial c}{\partial \boldsymbol{q}}, \Delta\boldsymbol{q}> = -\frac{\partial c_0}{\partial q_0}\Delta q_0 + <-\frac{\partial c}{\partial \bar{\boldsymbol{q}}}, \Delta\bar{\boldsymbol{q}}> \tag{28}$$

by definition. Since $<\boldsymbol{f_0},\Delta\theta> = \Delta q_0$, $\boldsymbol{F}\Delta\theta = \Delta\bar{\boldsymbol{q}}$, substituting $<\boldsymbol{f}_0, \boldsymbol{F}^{-1}\Delta\bar{\boldsymbol{q}}> = \Delta q_0$ into eq. (28) and then applying eq. (25), we obtain eq. (26).

Applying Farkas' lemma to (23) and (26), we obtain

$$v = \sum_{k = k_1}^{k_m} \mu_k \boldsymbol{n}_k, \quad \mu_k \geq 0\,. \tag{29}$$

Substituting (27) into (29) and multiplying the result by $\mathbf{F}^{\mathrm{T}}$, we obtain

$$\boldsymbol{f}_0\left(\frac{\partial c_0}{\partial q_0}\right) + \boldsymbol{F}^T\frac{\partial c}{\partial \bar{\boldsymbol{q}}} + \sum \mu_k \boldsymbol{g}_k = 0\,. \tag{30}$$

This implies that $\boldsymbol{q}$ satisfies the optimality conditions (16-18), i.e., $\boldsymbol{q} = \boldsymbol{q}^*$, which is the contradiction we seek.

Lemma 5 For a FD trade $\Delta\boldsymbol{q}$ at $\boldsymbol{q} \in S$ that reduces the total cost, $c(\boldsymbol{q} + \Delta\boldsymbol{q}) < c(\boldsymbol{q})$, there is profit to be made in arranging such a trade.

Proof Suppose that Δq involves a set of generator buses G and a set of demand buses D, and $\Delta q_i = 0$ for $i \notin G \cup D$. Furthermore, let $G = G^+ \cup G^-$ and $D = D^+ \cup D^-$ where for $i \in G^+$ or $i \in D^+$, $c_i(q_i + \Delta q_i) - c_i(q_i) > 0$ and for $i \in G^-$ or $i \in D^-$, $c_i(q_i + \Delta q_i) - c_i(q_i) < 0$. We may decompose $B = -[c(\boldsymbol{q} + \Delta\boldsymbol{q}) - c(\boldsymbol{q})] > 0$ into four parts, $B = B_1 + B_2 - C_3 - C_4$.

(i) $B_1 = \sum_{i \in D^-} -[c_i(q_i + \Delta q_i) - c_i(q_i)]$

B_1 is positive. Since c_i is strictly increasing, Δq_i must be negative. In our convention this means that the demand at bus i for this group of consumers is increased by $|\Delta q_i|$ MW and B_1 represents the additional benefit of the increased demand to this group. B_1 is the maximum amount this group of consumers is willing to pay for the additional demand.

power flow through line k when the operating point is at $\bar{\theta}$, the solution of the power flow equations when the power injected to the network is $\boldsymbol{e}_i$, i.e., one unit power injected into node i. The inner product $<\boldsymbol{n}_k, \Delta\bar{\boldsymbol{q}}>$ is thus the net amount of power flow on line k as a result of the trade $\Delta\boldsymbol{q}$. The requirement for a trade to be in the feasible direction (Eq.(20)) is that the trade should result in reducing the net power flowing through the congested line. We shall refer to the vector $\boldsymbol{n}_k$ as the Loading Sensitivity Vector corresponding to line k. Those who are familiar with optimization theory will recognize $\boldsymbol{n}_k$ as the normal vector to the manifold (affine subspace, in the linear case) defined by the constraints $<\boldsymbol{g}_k, \theta> = l_k$.

Lemma 3 Let $\Delta\boldsymbol{q}$ be a FD trade at $\boldsymbol{q}$. Then $(\boldsymbol{q} + \Delta\boldsymbol{q})$ satisfies the power balance equations (13)-(14) and the line flow constraints (15) for $k = k_1, k_2, \ldots, k_m$.

Proof Let the corresponding power flow solutions be $(\theta + \Delta\theta)$, i.e., $\boldsymbol{F}(\theta + \Delta\theta) = \bar{\boldsymbol{q}} + \Delta\bar{\boldsymbol{q}}$, or $\Delta\theta = \boldsymbol{F}^{-T}\Delta\bar{\boldsymbol{q}}$. Then

$$
\begin{aligned}
<\boldsymbol{g}_k, (\theta + \Delta\theta)> &= <\boldsymbol{g}_k, \theta> + <\boldsymbol{g}_k, \Delta\theta> \\
&= l_k + <\boldsymbol{n}_k, \Delta\bar{\boldsymbol{q}}> \le l_k
\end{aligned}
\tag{22}
$$

8.3 Profitable trades

We now present two key results.

Lemma 4. Let $\boldsymbol{q} \in S$ and $\boldsymbol{q} \neq \boldsymbol{q}^*$. There exists a FD trade $\Delta\boldsymbol{q}$ at $\boldsymbol{q}$ that reduces the total cost, i.e., $c(\boldsymbol{q} + \Delta\boldsymbol{q}) < c(\boldsymbol{q})$.

Proof: The proof is by contradiction. Suppose that the assertion is false, i.e., for all $\Delta\boldsymbol{q} = (\Delta q_o, \Delta\bar{\boldsymbol{q}})$

$$<\boldsymbol{n}_k, \Delta\bar{\boldsymbol{q}}> \le 0 \quad \text{for } k = k_1, k_2, \ldots, k_m \tag{23}$$

$$c(\boldsymbol{q} + \Delta\boldsymbol{q}) \ge c(\boldsymbol{q}) \tag{24}$$

Since $c(\boldsymbol{q} + \Delta\boldsymbol{q}) = c(\boldsymbol{q}) + <\frac{\partial c}{\partial \boldsymbol{q}}, \Delta\boldsymbol{q}> + o(\Delta\boldsymbol{q})$, this implies that

$$<-\frac{\partial c}{\partial \boldsymbol{q}}, \Delta\boldsymbol{q}> \le 0 \tag{25}$$

We claim that Eq. (25) is equivalent to:

$$<\boldsymbol{v}, \Delta\bar{\boldsymbol{q}}> \le 0 \tag{26}$$

where

Definition A multilateral trade is represented by an injection vector $\boldsymbol{q}[k]$ such that $\sum_{i=0}^{n} q_i[k] = 0$, where k is used to index a set of multilateral trades. The set of all multilateral trades in the system results in an injection vector $\boldsymbol{q} = \sum_k \boldsymbol{q}[k]$. A multilateral trade $\boldsymbol{q}[k]$ is said to be curtailed if $\boldsymbol{q}[k]$ is replaced by $\gamma_k \boldsymbol{q}[k]$, $0 \le \gamma_k \le 1$.

Lemma 2. Suppose a set of multilateral trades $\boldsymbol{q}^0 = \sum_m \boldsymbol{q}^0[m]$ is not feasible (i.e., it violates eq. (11) for some $k = 1, 2, \ldots, L$; note by Corollary 1, eq. (10) is satisfied by a multilateral trade.) There exists a curtailment schedule γ_m such that $\boldsymbol{q} = \sum_m \gamma_m \boldsymbol{q}^0[m]$ is on the boundary of the feasible set.

Proof: Note that $\boldsymbol{q} = 0$ is in the feasible set. Consider $\gamma\boldsymbol{q}^0$: for $\gamma = 0$, $\gamma\boldsymbol{q}^0 \in S$ and for $\gamma = 1$, $\gamma\boldsymbol{q}^0 \notin S$. Since S is convex, there is a (unique) $0 < \bar{\gamma} < 1$ such that $\bar{\gamma}\boldsymbol{q}^0 \in \partial S$.

Remark In the proof of Lemma 2, a specific curtailment scheme is invoked, in which all trades are curtailed uniformly. Other curtailment schemes that are easier to implement may actually be used in practice.

Definition Suppose that $\boldsymbol{q} \in \partial S$, i.e., eqs. (10)-(11) are satisfied and

$$<\boldsymbol{g}_k,\theta> = l_k \quad \text{for } k = k_1, k_2, \ldots, k_m \text{ and} \tag{19}$$

$$<\boldsymbol{g}_k,\theta> < l_k \quad \text{otherwise}$$

A (multilateral) trade $\Delta\boldsymbol{q} = (\Delta q_0, \Delta\bar{q})$ is a *feasible-direction* (FD) trade at $\boldsymbol{q}$, if

$$<\boldsymbol{n}_k,\Delta\bar{\boldsymbol{q}}> \le 0 \quad \text{for } k = k_1, k_2, \ldots, k_m \tag{20}$$

where

$$\boldsymbol{n}_k = \boldsymbol{F}^{-T}\boldsymbol{g}_k \tag{21}$$

and **F** is defined in eq. (14).

Remark. The physical meaning of the vector $\boldsymbol{n}_k$ defined in eq. (21) is the following: the i-th element of $\boldsymbol{n}_k$ is equal to the amount of MW flows on line k if 1 MW is injected into the i-th node of the network (and taken out at the reference node). To see this, we first express the i-th element of $\boldsymbol{n}_k$, $(\boldsymbol{n}_k)_i$, using a vector $\boldsymbol{e}_i$, whose elements are all zeros except the i-th element which is equal to 1. We have $(\boldsymbol{n}_k)_i = <\boldsymbol{n}_k,\boldsymbol{e}_i> = <g_k, \bar{\theta}>$ where $F\bar{\theta} = \boldsymbol{e}_i$. The second inner product implies that $(\boldsymbol{n}_k)_i$ is the amount of

$$<\boldsymbol{g}_k,\theta> \le l_k \quad k = 1, 2, \ldots, L \tag{15}$$

where $\bar{\boldsymbol{q}} = (q_1, \ldots, q_n)^T$ and $\boldsymbol{F} = \begin{bmatrix} — \boldsymbol{f}_1^T — \\ \vdots \\ — \boldsymbol{f}_n^T — \end{bmatrix}$. Note that $\boldsymbol{F}$ is invertible. Problem (12)-(15) is a convex programming problem with linear constraints. Therefore, $\boldsymbol{q}^*$ is optimal if and only if it satisfies (7)-(9):

$$\frac{\partial c}{\partial \boldsymbol{q}}(\boldsymbol{q}^*) = \boldsymbol{p}^* \tag{16}$$

$$p_0^* \boldsymbol{f}_0^T + (\bar{\boldsymbol{p}}^*)^T \boldsymbol{F} + \sum_{k=1}^{L} \mu_k \boldsymbol{g}_k^T = 0 \tag{17}$$

$$\begin{cases} \mu_k (<g_k,\theta> - l_k) = 0 \\ \mu_k \ge 0 \end{cases} \tag{18}$$

Using the standard terminology of nonlinear programming, any $\boldsymbol{q}$ satisfying (10)-(11), or (13)-(15) is said to be feasible. The set of feasible $\boldsymbol{q}$, denoted by S, is a convex set in a (n+1)-dimensional space. Assume that S has interior, the boundary of S is denoted by ∂S.

Lemma 1. For a lossless system, $-\boldsymbol{F}^{-T}\boldsymbol{f}_0 = \mathbf{1}$, where **1** is the vector whose components are all 1's.

Proof. From the definition of the terms in the vector $\mathbf{f}_0$ and the matrix **F**, it can be verified that the sum of all elements in a column of **F** plus the corresponding element in $\mathbf{f}_0$ is equal to zero.[28] Hence the lemma follows.

Corollary 1. Any injection vector $\boldsymbol{q}$ satisfies the power flow balance equations (13) and (14) if and only if the sum of nodal injections is equal to zero, $\sum_i q_i = 0$.

Physically, Corollary 1 states that in a lossless system, the total generation must be equal to the total consumption.

8.2 Feasible Trades

We shall define multilateral trades. A bilateral trade between a generator at node *i* and a consumer at node *j* for α MW to be generated at node *i* and α MW to be delivered at node *j* is represented by an injection vector $\boldsymbol{q}$ whose *i*th component is α and *j*th component is (-α) In general, a trade may involve multiple parties and will be called a multilateral trade.

[28] These terms can be obtained from the partial derivatives of the power flow equations described in the footnote 27.

$$min \;\; c(\boldsymbol{q}) \;=\; \sum_{i=0}^{n} c_i(q_i) \tag{4}$$

subject to

$$\boldsymbol{f}(\theta) \;=\; \boldsymbol{q} \tag{5}$$

$$\boldsymbol{g}(\theta) \le \boldsymbol{l} \tag{6}$$

The Kuhn-Tucker optimality conditions for the nonlinear programming problem (4)-(6) are:

$$\frac{\partial c}{\partial \boldsymbol{q}} \;=\; \boldsymbol{p} \tag{7}$$

$$\boldsymbol{p}^T\frac{\partial \boldsymbol{f}}{\partial\theta} + \mu^T\frac{\partial \boldsymbol{g}}{\partial\theta} \;=\; 0 \tag{8}$$

$$\begin{cases} \mu_k(g_k(\theta) - l_k) = 0 \\ \mu_k \ge 0 \end{cases} \tag{9}$$

The optimal dispatch problem (4)-(6) stated here is the standard welfare maximization problem in economic theory. It is also a generalization of the standard optimal power flow (OPF) problem in power systems analysis, with consumer benefit included.

We now make two simplifying assumptions. The first is to assume that the lines are lossless and the second is to assume that the linear approximation of the line flows (2) and (3) is valid. We will revisit and relax these two assumptions in Sec. 8.5 and Sec. 8.6, respectively. It should be pointed out that linear approximations are often used in power flow analysis. The approximations are fairly good for real power flows.

Under the linearity assumption, the power flow equations (2) and (3) can be written as:

$$<\boldsymbol{f}_i,\theta> = q_i \qquad i \;=\; 0, 1, \ldots, n \tag{10}$$

$$<\boldsymbol{g}_k,\theta> \le l_k \qquad k \;=\; 1, 2, \ldots, L \tag{11}$$

where the inner product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, $<\boldsymbol{a}, \boldsymbol{b}>$, is defined as $\boldsymbol{a}^T\boldsymbol{b}$.

In this case, the optimal dispatch problem becomes:

$$min \;\; c(\boldsymbol{q}) \tag{12}$$

subject to

$$<\boldsymbol{f}_0,\theta> \;=\; q_0 \tag{13}$$

$$\boldsymbol{F}\theta \;=\; \bar{\boldsymbol{q}} \tag{14}$$

VIII. Appendix: Theory

8.1 Optimal Dispatch

Consider a power network consisting of *(n + 1)* nodes or buses operating in sinusoidal steady-state. The voltage phasor at bus *i* is denoted

$$V_i e^{j\theta_i} \quad i = 0, 1, \ldots, n \tag{1}$$

An arbitrary bus, say bus 0, is chosen as the reference bus for the voltage phasor angles, i.e., $\theta_0 = 0$. Let us assume that the voltages V_i are kept constant by adjusting reactive powers at the buses and we focus on real power flows. The real power flow balance at each bus is expressed by the (real) power balance equation[27]

$$f_i(\theta) = q_i \quad i = 0, 1, \ldots, n \tag{2}$$

where $\theta = (\theta_1, \ldots, \theta_n)^T$, q_i is the net power injection at bus *i*. We adopt the sign convention that q_i is positive (negative) if there is net generation (consumption) at bus *i*. The vector $\boldsymbol{q} = \left[q_0, q_1, \ldots\ q_n\right]^T$, the array of all nodal injections, is called an injection vector.

The (real) power flow through any transmission line (or transformer) can be expressed as a function of θ. Indeed, the power flow through a set of lines or between two regions can always be expressed as a function of θ. Let the line flow limit on line *i* or the transfer limit between two regions be denoted by l_i; then the transmission loading constraints are

$$g_k(\theta) \le l_k \quad k = 1, 2, \ldots, L \tag{3}$$

where *L* is the total number of such constraints.

Let $c_i(q_i)$ be the cost function if bus *i* is a net generation bus, $q_i > 0$, and let $c_i(q_i)$ be the negative of the consumer benefit function if bus *i* is a net consumption or load bus, $q_i < 0$. We assume that the functions c_i are convex and strictly increasing.

The optimal dispatch problem is to maximize total consumer benefit and to minimize total generation cost such that the power flows are balanced and the line flow constraints are satisfied. If consumer benefit is treated as negative cost, then the problem becomes

---

27. Let the reciprocal of the impedance between line km be $y_{km} = 1/z_{km} = g_{km} - jb_{km}$. Then the k-th equation is $f_k(\theta) = \sum_m g_{km}V_k^2 + \sum_m -g_{km}V_kV_m\cos(\theta_k - \theta_m) + \sum_m b_{km}V_kV_m\sin(\theta_k - \theta_m) = q_k$. See textbooks such as Bergen, A. R. *Power Systems Analysis*, Prentice-Hall, 1986, and Wood, A. J., and B. F. Wollenberg, *Power Generation, Control, and Operation*. New York: John Wiley and Sons, 1984.

management systems. It uses on-line load flow or optimal power flow to examine the consequence of potential disturbances, with all the loading and meteorological data available, rather than working with off-line postulated conditions. Significant research has been accomplished in recent years in furthering the security assessment into system dynamic responses to disturbances. Deployment of on-line steady-state and dynamic security assessments can greatly enhance effective utilization of existing transmission capacity. Coordinated multilateral trading model is compatible with such technological development. We believe feasibility conditions under on-line security assessment scenarios can be developed.

For ease of presentation, the rigorous mathematical derivations of the coordinated multilateral model in the appendix include only the real power flows, assuming voltages are held constants. The extension to include reactive power flows and voltage variations is necessary and will be presented in a future publication. The theory presented in this paper is a general methodology which is applicable to any nonlinear systems of this type. Therefore the extension to include reactive power flows is conceptually straightforward.

In this paper we focus on how to operate the system to achieve economically efficient generation and consumption in a network where the integrity of the network operation has to be maintained. The network is treated as a shared resource used by the parties connected to it. The owner of the network and the operator of the network provide a service to the generators and consumers using the network and should be compensated. The question of transmission charges or compensation has not been addressed. The merchandising surplus in the coordinated multilateral trading model, i.e., the difference between the total benefit for consumers and the total cost for generators in the whole network, represent the profit made by the participants and should be used to compensate the transmission network owner and operator. This is a subject of further research.

Acknowledgment

The authors thank George Gross, Carl Imparato, John Kaye, Shmuel Oren, Hugh Outhred, and Pablo Spiller for many insightful discussions. We also thank Hoyt Sze for his help in making the paper readable.

[26.] See for example Wu, F. F., "Analysis techniques for power system security assessment and optimization: research needs and emerging tools," *Proc. Wksp on Power System Security Assessment*, Iowa State University, April 1988. Direct methods for transient stability have been proposed as a tool for on-line dynamic security assessment, see: Fouad, A. A., and V. Vittal, *Power System Transient Stability Analysis Using the Transient Energy Function Method*, Prentice-Hall, 1992, and also: Varaiya, P., R. L. Chen, and F. F. Wu, "Direct methods for transient stability analysis of power systems: recent results," *Proc. of the IEEE*, vol. 73, 1985, pp. 1703-1715.

not. The proposed model can readily be implemented with existing communications, computing and control infrastructure.

Thus the proposed model achieves short term efficiency. The separation within our model between the technical security and loss allocation functions carried out by the PSO and the financial transactions carried out in the free market, leads to great flexibility. First, consumers and generators can create arbitrarily differentiated commodities of electric power (based, for example, on reliability, interruptibility, temporal constraints, etc.) in order to take advantage of diversity in consumer choice and generation technology. Second, over time, even the limited functions of the PSO, can be provided by alternatives. Thus, for example, allocated losses can be provided by independent generators, as can back-up generation. Thus the proposed model provides incentives for long term innovation. The poolco and bilateral models, by installing the centralized PSO at the core of their proposals, preclude the possibility of such innovation. Moreover, the latter models require a permanent regulatory body that is even more sophisticated, resourceful, and vigilant than it is today. The coordinated multilateral trading model, by contrast, envisions a steady reduction in the need for regulation.

At the equilibrium operating point under coordinated multilateral trading model, the marginal cost (price) at each node of the network will be equal to what a central dispatch authority would determine[24] if true cost/benefit functions can be, and are in fact, provided. The coordinated multilateral trading model thus attains all the advantages of optimal nodal pricing. The optimal nodal prices send correct economic signals to the generators and consumers in the system to encourage efficient competition in generation, as well as efficient transmission expansion. As a result of transmission congestion, the marginal cost of generation at some locations will be lower than the marginal price the consumers are willing to pay at some other locations. That gives incentives for generators, consumers or brokers to engage in network upgrading or expansion to facilitate more profitable trading. The incentives are shared by all the generators and consumers affected by the congestion and the incentive for network enhancement is not limited to the congested line.

A vibrant financial market facilitates competition.[25] It has been proposed to establish locational (nodal) electricity forward markets throughout the network. Such markets will facilitate the development of optimal nodal pricing. The coordinated multilateral trading model is compatible with such nodal markets. One important consequence of having nodal forward markets is that they help hedging the price difference between the supplier and the consumer due to transmission congestion.

Advances in 3C technologies have moved the security function in operation from using simple guidelines on transfer limit set by off-line computer simulations to real-time assessment.[26] On-line steady-state security assessment software have been introduced and used in some advanced energy

---

24. Optimal nodal pricing is discussed in: Schweppe, F. C., R. D. Tabors and R. E. Bohn, *Electricity Spot Pricing*, Kluwer Academic Press, 1988.

25. See Outhred, H., and R. J. Kaye, "The nodal auction model for implementing competition in a bulk electricity industry," U of New South Wales Report DEPE 941012, Oct. 1994 and also Oren, S. S., P. T. Spiller, P. Varaiya, and F. F. Wu, "Nodal prices and transmission rights: a critical appraisal," *The Electricity Journal*, vol. 8, April 1995, pp. 14-23.

Within the boundary of a PSO's operation, generators and consumers can form coalitions of any size, from any location in the system, and at any time. A coalition becomes a group of generators and consumers engaged in a joint dispatch. In other words, in the coordinated multilateral trading model, economic dispatch (or more generally, transmission constrained economic dispatch) can be provided as a service and generators and consumers have a choice whether to purchase the service and from whom. There is also no limitation to the possibility of carrying out trades between generators and consumers belonging to different coalitions. They should be encouraged to promote economic efficiency.

A coalition is basically a "pool" in the coordinated multilateral trading model with the pool operator serving as the broker for the trades within the pool. The pool jointly dispatches its generation to serve the aggregate load in the pool. The pool operator may operate the system by accepting bids for generation and consumption at different locations as proposed in poolco model. However, unlike in the poolco model, participation in the pool is totally voluntary.

The function of the PSO is to coordinate various parties to ensure the integrity of network operation and the operation is mechanical and transparent. Large coalitions with hundreds of generators and loads will be treated no differently by the PSO than a coalition of one generator and one load. Comparability of service is a non-issue. The PSO in the coordinated multilateral trading model, by definition, provides non-discriminatory service to all. This is especially useful during the transitional period moving from traditional electric utility structure to direct access structure, when both systems coexist, both utility customers and the direct access customers are guaranteed to receive comparable service from the PSO.

VII. Conclusion

As the electric power industry moves into an era of supply competition and consumer choice, a new operating paradigm for the transmission network connecting suppliers and consumers is needed. The new operating paradigm must be compatible with the economic principles of the new era. The two alternative transmission restructuring proposals, bilateral model and poolco model, adhering to the old operating paradigm developed for regulated monopolies, are trying to sculpture the foot to fit the shoe. They restructure the wrong end of the problem! The proposed changes are unnecessary, unworkable, and undesirable. We have designed a new operating paradigm in which coordinated private multilateral trades, each of which benefits all parties to the trade, lead to overall welfare maximization. The proposed coordinated multilateral trading model achieves better economic efficiency because there is no need to provide the power system operator or anybody else the individual's explicit cost and benefit function. All cost/benefit information are private and they are used only for the purpose of negotiating contracts between willing parties. The same level of service reliability and system security as in today's centralized operation can be achieved in the proposed model through proper design of the information structure. All generators and consumers have open access to the transmission network as long as the parties observe their fair share of responsibility for the shared resource of the transmission network. Non-discriminatory transmission service is offered to everyone, whether part of a coalition or nication requirement for PSO broadcast. A Loading vector or a Loss vector contains a few thousand bytes of data. The communications in today's energy management systems typically uses a 9600 bits per second (bps) line. In that case, the PSO can broadcast the vector within a couple of seconds. If a 64 kbps digital channel is used, which is the smallest denomination in modern wide-area communication networks, the broadcast will take even less time. Another option is to use the FM or UHF broadcast radio currently employed for direct load control in demand-side management.

The communication requirement can be drastically reduced using the fact that the vectors sent by the PSO are as insensitive to transmission loading conditions as the penalty factors. Just like in today's control centers where several sets of pre-calculated penalty factors related to seasonal or peak/off peak patterns are used, the Loading vectors and the Loss vectors can similarly be pre-calculated for several classes of major loading patterns. They can be sent to the brokers off-line and the brokers store the data in their own computers. In real-time operation, the PSO only need announce which relevant vector is in effect and should be used for calculating congestion management or loss allocation. Only a few bytes of data now need to be broadcast on-line.

6.4 Data requirements

To carry out its responsibility, the PSO needs to find out the loading on each facility in the interconnected network. The data required to perform this function is precisely the data required to run a load flow. Two types of data are needed. One is the fixed set of data of the transmission network, such as transmission line and transformer impedances, network configuration, etc., and the second type of data is the data pertaining to the current operating point, including generation and load level at each bus. The former can be stored and the latter must be sent on-line.

The process of calculating the Loading vector and the Loss vector is straightforward and transparent. Since load flow is so standard and software running on personal computers are widely available, the result can be checked by almost everybody and is impossible to manipulate. As a matter of fact, the PSO's function in normal situations is so mechanical, it can be automated. The PSO's function is really in handling emergency situations.

The PSO in the coordinated multilateral trading model does not hold any data pertaining to the profitability of the trades, which are private. There is no danger of PSO manipulating the market and there is no need for intensive scrutiny by the regulators.

6.5 Organizational requirements

The domain of the PSO's operation is the part of the interconnected power network for which the PSO is responsible: network reliability, security and power balance. Because power flows interact in a network, the domain of operation must be contiguous. Other than that, there are no physical, technical or organizational constraints on the size and boundary of PSO's operation. Generators and consumers in neighboring systems can engage in profitable trades with parties within the domain of a PSO's operation through power transfers at the boundary buses.

To ensure that additional trades will not result in transfer limit violation, the PSO calculates and broadcasts the Loading vector. The calculation of a Loading vector involves the solution of a set of linear equations. The coefficients of these equations are elements of the Jacobian matrix of the power flow equations. They are available as the result of a load flow calculation. The matrix and the vector involved in the solution of the Loading vector are all sparse (i.e., only a very small percentage of their elements are nonzero); extremely efficient solution algorithms exist and for a network with one or two thousand buses; and it takes a few seconds 0f computer time to find a solution. As a matter of fact, existing software for computing penalty factors in economic dispatch can be easily modified to calculate the Loading vector. Penalty factors are routinely calculated in today's energy management systems.

To allocate transmission losses to individual trades, the PSO calculates and broadcasts the Loss vector. The Loss vector has two components, the first component is directly related to the penalty factors and the second component uses the second derivatives of the load flow equations (elements of the Hessian matrix). The required terms are again readily available from the result of a load flow. The total computational effort in calculating the Loss vector is about twice that of solving a set of linear equations, which is well within the capability of existing energy management systems. The calculation of the quadratic loss matrix Q by straightforward methods requires the number of multiplications on the order of n^3, where n is the number of buses in the network. Special sparse matrix techniques will be needed and are available for handling this problem for large networks.[23]

The basic computational engine in the coordinated multilateral trading model is the load flow, whereas the poolco and the bilateral models both require an on-line optimal power flow. The on-line load flow has been routinely performed in energy management systems for almost twenty years, whereas experience with successful on-line optimal power flows for large networks is still limited.

The data required to compute the Loading vectors and the Loss vector are basically the same as that required for the penalty factor calculation. Extensive practical experience has shown that penalty factors need not be updated every hour; rather, it suffices to update penalty factors when major changes in the generation and consumption pattern occur. This is because the coefficients in the equations are not very sensitive to small changes in transmission loading. Therefore, the frequency of computation required for updating the Loading vectors and the Loss vector is similarly expected to be low.

6.3 Communications requirements.

The data exchange between the PSO and brokers runs in both directions. The brokers report to the PSO information concerning MW input and output schedule of their trades. The PSO broadcasts the Loading vectors and the Loss vector. Let us assume all trades are arranged at the substation (generation and load) level and consider for example a network with a few hundred substations with the existing communication infrastructure of the energy management system. The communication requirement from the brokers to the PSO is insignificant and can easily be accommodated. Let's look at the commu-

23. See the second remark after lemma 8 in appendix.

ever, this requirement of lead time is only limited by technological capabilities, i.e., the computation turnaround time. With today's control center capabilities, this lead time can be less than half an hour before real-time.

The real-time market, operated by the PSO with the help of the AGC facility, ensures power balance and frequency and voltage regulation in order to maintain the integrity of transmission network operation. Any trade whose real-time supply and demand deviate from the scheduled amount is charged, if the deviation is compensated by the PSO. The broker of the trade can avoid the charge either by enforcing power balance within the trade, or by contracting the service to a third party. This may stimulate technological innovation, a long-term benefit of competition.

6.2 Computation requirements

1. Broker

A broker arranges trades and bears responsibility for arranging generation to compensate for losses and to insure that the trading is feasible. The broker calculates the share of losses based on the Loss vector broadcast by the PSO. The calculation involves a few multiplications and additions, and is trivial. If it is desired to calculate various components of the losses, i.e., the losses imposed by other trades, the broker uses the Q loss matrix, again broadcast by the PSO, and the computation is still insignificant.

When there is congestion in the transmission network, the broker needs to determine the feasibility of trades based on the Loading vector broadcasted by the PSO. For a trilateral trade, the conditions for profitable trades that are feasible are spelled out in section 4.3 and section 8.4 in the appendix and the computation involves solving a set of three simple algebraic equations. When the broker wants to engage more parties in a complex multilateral trade, the amount of computation for checking feasibility of a trade is still trivial. To find an optimal solution that is profitable and feasible in the general case, however, requires the solution of a transmission constrained economic dispatch (TCED) problem. Efficient solution algorithms for TCED exist and software is available. A TCED software can be efficiently executed in a PC for a group of several dozen generators and loads.

2. Power System Operator

The PSO functions in the coordinated multilateral trading model are primarily for the maintenance of network reliability and security and real-time power balance. It is the PSO’s job to ensure that the power flows in the network are within their loading limits and that the collective losses are fairly allocated.

The PSO checks feasibility of the proposed trades and curtails trades if necessary. To check if trades are feasible, the PSO solves a load flow, based on the generation and consumption data from the proposed trades. Standard packages for load flow are available for computers ranging from personal computers to large mainframes. In today's energy management systems, on-line load-flows are routinely performed and it takes less than a minute of computer time to solve a load flow for a network having more than a thousand buses.

The accuracy of the estimate is equal to 98.05% (0.0198/0.0202) in this case. Eq. (72) can be used to correct the error in the estimation, which results in:

Table 4: Components of Transmission Losses (MW)

	TA	TB	TC	Total
TA	0.62	0.094	0.32	1.03
TB	0.094	0.097	0.11	0.30
TC	0.32	0.11	0.21	0.65
Total				1.98

A power flow calculation will verify that if one trade is present, its corresponding losses is indeed equal to the diagonal elements in the above table. Moreover, if only two trades TA and TB are present, the sum of all the components corresponding to these two trades (the four blocks on the upper left corner) is equal to 0.91 MW, which is again exactly the result obtained from a load flow calculation. The incremental losses caused by TC can be calculated by using the formula Eq. (48) in appendix, which gives 1.07 MW, matching the load flow results and the results obtained from Table 4 (1.98-0.91= 1.07MW).

VI. Implementation Requirements

In this section, we will discuss the implementation of the coordinated multilateral trading model. We will demonstrate that the new operating paradigm for the coordinated multilateral trading model can readily be implemented using today's communication, computing and control infrastructure.[22] The organizational issues are also discussed and we will show that the coordinated multilateral trading model has flexible boundary and provides non-discriminatory services to parties connected to the network, whether they are part of a coalition or not.

6.1 Scheduled vs. real-time markets

The coordinated multilateral trading model can be implemented in the scheduled power market. The difference between a scheduled and real-time power markets is the lead time provided in the former. Lead time gives the PSO a chance to check security before the trades are executed. If transmission security is not threatened, execution proceeds. If security is threatened, the PSO curtails certain trades. Some lead time is required to curtail certain trades to avoid threatening system security. How-

22. A good reference for the status of present computer and communication systems for real-time power system control can be found in: Wu, F. F., and R. D. Masiello (ed.), *Computers in Power System Operations*, Special Issue of the *Proceedings of the IEEE*, vol. 75, Dec. 1987, a brief description of an integrated 3C infrastructure for both electric utility business and engineering applications is in: Lun, S. M., F. F. Wu, N. Xiao, and P. Varaiya, “NetPlan: an integrated network planning environment”, *Proc. Int. Conf. on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, San Diego, Jan 1993, and Wu, F. F., “The integrated information system-gains and obstacles”, (Invited talk) *Stockholm Power Tech*, June 1995.

The accuracy of the estimate for this example is 98.85% (0.0198/0.0200). To allocate the total losses of 1.98 MW, PSO would add a correction factor and use Eq. (76), resulting in the following allocation:

Table 2: Loss Allocation (MW)

Trading	TA	TB	TC	Total
Loss Alloc	1.03	0.30	0.65	1.98

We have also derived formulas to estimate (and calculate) the additional losses (may be positive or negative, in general) caused by other trades. For that purpose, the PSO calculates and broadcasts the Q loss matrix (Q for quadratic). The calculation involves the Jacobian matrix and the Hessian matrix of the power flow equations (see Eq. (50)) and, in this case, it is equal to:

$$Q = \begin{bmatrix} 0.0589 & 0.0207 & 0 \\ 0.0207 & 0.0496 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

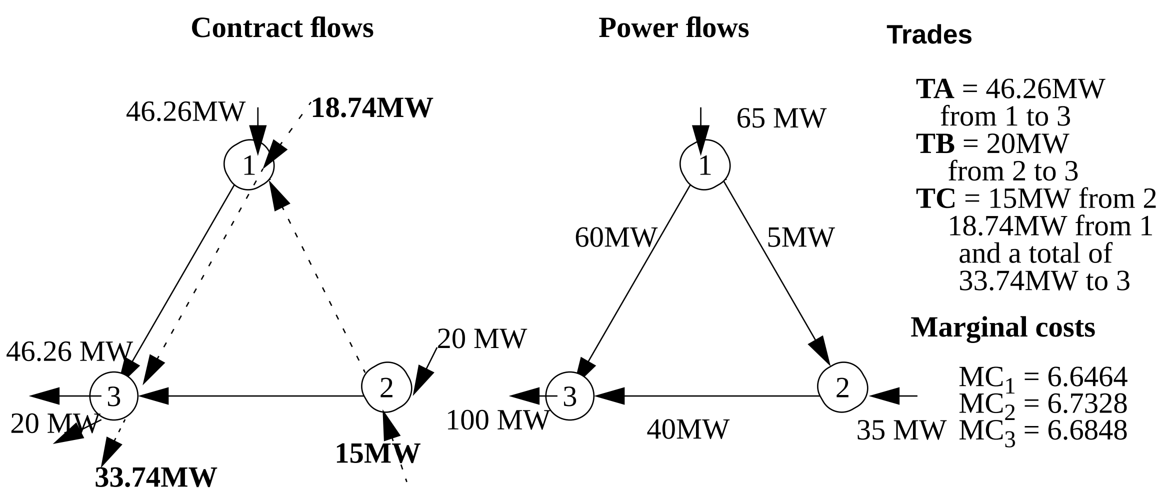
Based on the Q matrix, brokers can estimate the losses of their trade and its components. The results are tabulated in the following table. To illustrate how the elements in the table is obtained, take for example, the one corresponds to row TB and column TA, it is the result of the product (TB)'Q(TA)/2, where (TB)' represents the transpose of the vector (TB). The product is simply equal to (0.2 x 0.0207 x 0.4626)/2 = 0.00096pu = 0.096MW.

Table 3: Estimated Loss Component (MW)

	TA	TB	TC	Total
TA	0.63	0.096	0.33	1.05
TB	0.096	0.099	0.11	0.31
TC	0.33	0.11	0.22	0.66
Total				2.02

Table 3 should be read this way. The number corresponding to TA/TA, 0.63 MW, is the losses caused by trade A alone, i.e., if there is no other trades going on in the network. The number corresponding to TA/TB, 0.096 MW, is the additional losses attributed to trade A due to the presence of trade B. Notice the additional losses attributed to trade B due to the presence of trade A is also 0.096 MW.

Fig. 4 Additional trade



2. Loss allocation

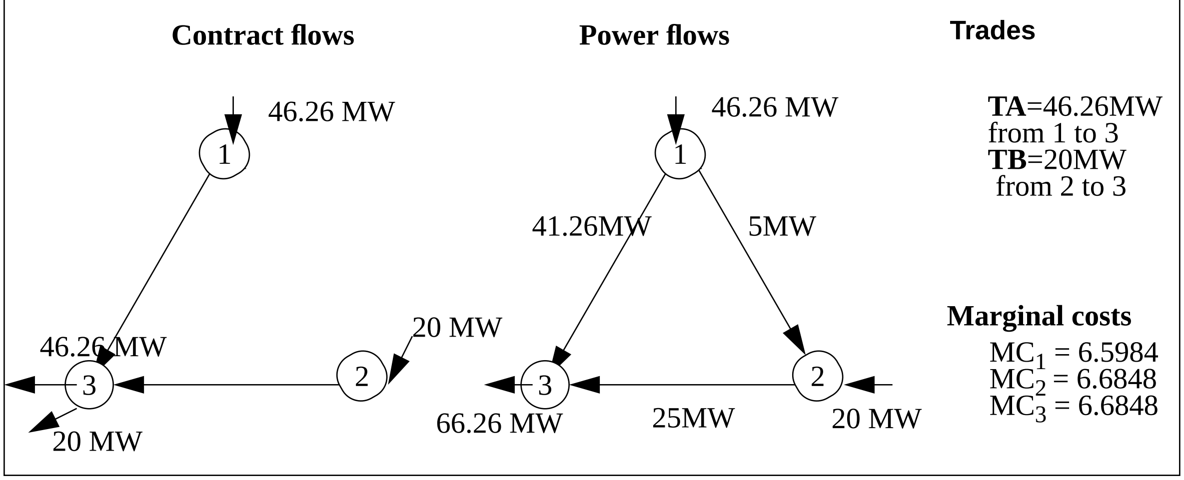
The three lossless trades in this example are TA=(0.4626, 0, -0.4626), TB=(0, 0.2, -0.2) and TC=(0.1874, 0.15, -0.3374). The total losses caused by the these trades can be calculated by a load flow solution and is equal to Tloss= 0.0198pu = 1.98 MW.

The PSO calculates the Loss vector using the power flow Jacobian matrix and broadcast the results to all parties. The Loss vector is equal to L=(0.0226, 0.0153, 0). Based on the Loss vector, the parties engaged in trades can estimate their share of the losses using Eq. (70). For example, an estimated losses for trade A is equal to 0.0226 x 0.04626 + 0.0153 x 0 + 0 x (-0.04626) = 0.0104 pu (i.e., 1.04MW). The results are shown in Table 1.

Table 1: Estimated Losses (MW)

Trading	TA	TB	TC	Total
Est Losses	1.04	0.31	0.65	2.00

Fig. 3 Curtailed trades



C. *Additional trades*.

At this moment, the marginal costs at the three buses are: MC_1= 6.5984, MC_2=6.6848 and MC_3=6.6848. Since $MC_1 < MC_3$, generator 1 notices that there may be an opportunity for additional profitable trade. But the Loading vector indicates that any bilateral trade between the two would further aggravate the congestion on line 1-2, and therefore, would be infeasible. A trilateral trade is needed (assertion 2, appendix).

The next question is: how to spot a profitable trilateral trade that is feasible. By examining the sign of the elements in the Loading vector, generator 1 would be able to generate more if generator 2 would increase its output. Since the average of the marginal costs of generators 1 and 2 is lower than what consumer 3 is willing to pay, a profitable trade is possible (see also assertion 3, appendix).

To determine how much power generator 1 wants generator 2 to produce, the marginal cost information is used. In general this should left to the parties to negotiate. However, in the case where the cost and benefit functions are explicit, assertion 4 in the appendix provides a solution. Solving Eqs. (39)-(41), one obtains:

$$\Delta q_1 = 0.1874$$

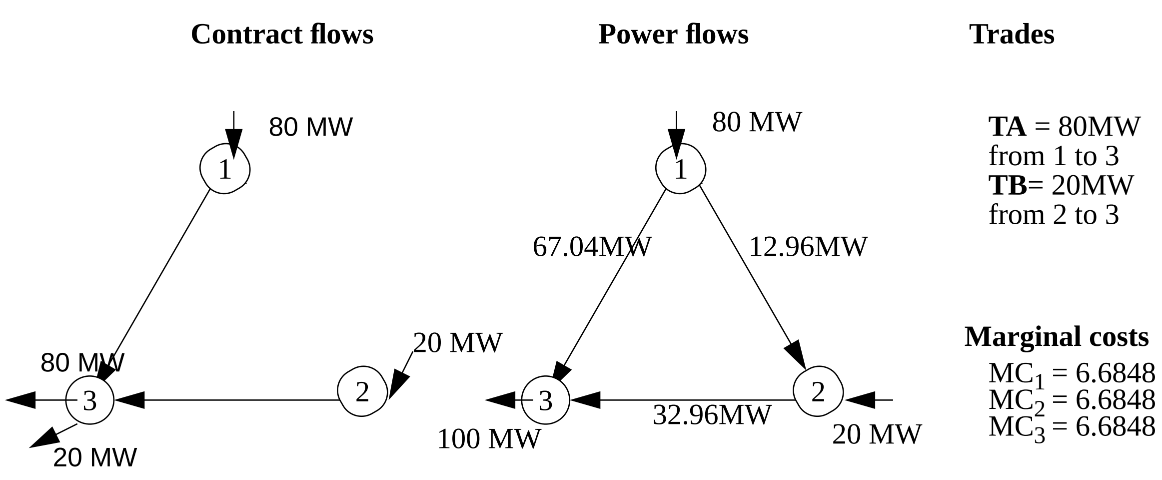
$$\Delta q_2 = 0.1500$$

Let us represent this additional trade as TC = (0.1874, 0.1500, -0.3374). With TC in place, the power flow in line 1-2 still remains 0.05 and the marginal costs at the buses are: MC_1= 6.6464, MC_2=6.7328 and MC_3=6.6848, which are exactly the optimal marginal costs, as can be checked from the solution of the optimal dispatch problem.

cate where the trade is from and to whom it is sent, let us use the vector notation TA=(0.8, 0, -0.8). The positions in the vector TA correspond to the three buses (in order), a positive number represents the amount of power injected into the network and a negative number represents the amount of power taken from the network. Thus TB=(0, 0.2. -0.2).

The resulting power flow of these two trades will overload transmission line 1-2, as the result of a power flow calculation shows that the power flow on line 1-2 would be 12.94 MW (0.1294 pu), exceeding its limit of 5 MW or 0.05pu.

Fig. 2 Initial trades



B. *Curtailment*

The PSO must curtail the trades to the secure operating level, i.e., to relieve the congestion. First, the PSO calculates the Loading vector by solving a linear equation using the power flow Jacobian matrix (Eq. (21) in appendix). The Loading vector in this example is: **n** = (0.2353, -0.2941, 0). The Loading vector indicates that curtailing trade A would relieve congestion and the amount of curtailment can be calculated using the sensitivity information provided by the Loading vector. Let Δq_1 be the amount of curtailment required, it can be obtained from the following equation:

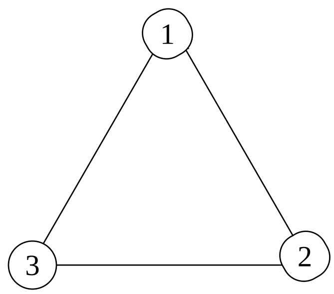
$$0.2353 \times \Delta q_1 = 0.05 - 0.1294$$

which gives $\Delta q_1 = -0.3374$. After this curtailment the trades are modified to: TA= (0.4626, 0, -0.4626) (since 0.4626 = 0.8-0.3374), and TB= (0, 0.2, -0.2) the PSO then broadcasts the Loading vector to all parties.

V. Illustrative Example

The coordinated multilateral trading model has been tested on IEEE test systems. The results for more realistic systems will be reported in the future. A simple 3-bus example system is used here to illustrate the concepts of the model. The data pertaining to the 3-bus system are shown below.[21]

Fig. 1 Data for the 3-bus example



Per unit: 1 pu = 100 MW
Transmission line data

	r	x
1-2	0.0600	0.4000
2-3	0.0315	0.2500
3-1	0.0400	0.2000

Cost functions ($/hr)

$$c_1(q_1) = 6480q_1 + 128q_1^2$$

$$c_2(q_2) = 6620.8q_2 + 160q_2^2$$

$$c_3(q_3) = 6684.8q_3$$

1. Congestion management

A. *Initial trades*

For simplicity, this part of the example ignores losses. Suppose initially generator 1 and generator 2 arrange independent contracts with consumer 3, with no consideration for the transmission network. The amount of the contracted power for each trade is set at the point where the marginal cost of generation is equal to the marginal benefit of consumption. Therefore, trade A from generator 1 to consumer 3 will be TA= 0.8 pu (80MW), when the marginal cost of generator 1 and the marginal benefit of consumption of consumer 3 are both equal to 6.6848 (c/kwh) and trade B from generator 2 to consumer 3 will be TB=0.2 pu (20MW), when the marginal cost of generator 2 is also at 6.6848 (c/kwh). To indi-

[21]. Data for this example are derived from an example in: Wood, A. J., and B. F. Wollenberg, *Power Generation, Control, and Operation*. New York: John Wiley and Sons, 1984. For convenience, we shall use the standard per unit calculation and translate the result in MW quantities at the end.

A simple estimate of a profitable trade can be obtained by using the marginal cost (benefit) of generation (consumption) MC_i.[20] A profitable trade is one that reduces the total cost, using a linear approximation, this means:

$$MC_1 q_1 + MC_3 q_3 \le MC_5 q_5$$

By examining the general case as to how to satisfy these equations for different combinations, we obtain answers to the questions above, which are stated as assertions in the appendix and repeated here.

If there is only one transmission congestion, a trilateral trade may be necessary in order to construct a profitable trade that is feasible (Assertion 1). If two transmission limits are congested, a quadrilateral trade may be necessary. Of course, if more parties are involved, the better the chance a more profitable solution can be found.

Consider, for simplicity, the case with only one transmission congestion and ignore transmission losses. To find a profitable trilateral trade involving, say generator 1, generator 3, and consumer 5, such that $q_1 + q_3 = q_5$, the marginal benefit of additional MW for consumer 5 must be higher than the marginal cost of one of the generators, say generator 1, i.e., $MC_5 > MC_1$. Suppose that generator 1 is also serving as the broker. If $n_5 \ge n_1$, $(n_1 - n_5)\, q_1 \le 0$ for any amount of q_1 MW from bus 1 to bus 5. The trade from node 1 to node 5 helps relieve the congestion, so any profitable trade can be used and there is no need to solicit the participation of another generator. A more interesting situation is when $n_5 < n_1$, so that a profitable bilateral trade between generator 1 and consumer 5 is not feasible. A simple rule (Assertion 3) obtains in this case to assist the broker to look for a generator, say generator 3, to negotiate for a feasible trade that is profitable. The rule says that if $n_5 > n_3$ the broker should negotiate for generator 2 to generate, whereas if $n_3 < n_2$, the broker should negotiate for generator 2 to back down.

Finally, in the case of a single transmission congestion, the optimal amount of generation and consumption in a trilateral trade can be determined if the cost benefit functions are explicit. The solution to a set of three algebraic equations Eq. (39)-(41) gives the optimal amount of generation and consumption in the trade.

In general, multiple parties may be involved in a trade. The trading arrangement problem can be formulated analytically as a transmission constrained economic dispatch problem, for which efficient solution algorithms exist. In the extreme case every generator and load is involved in the trade. This is also the case of the traditional integrated utility arrangement, and the transmission constrained economic dispatch becomes identical to the one performed in today's advanced energy management systems.

---

20. The marginal cost at a node is the derivative of the cost function, $MC_i = dc_i/dq_i$.

In fact, the coordinated multilateral trading model achieves better economic efficiency if the cost and benefit functions can not be explicitly written down as functions of generation or consumption as required for any centralized economic dispatch. In such a case, there is potential welfare loss for the centralized dispatch model, but not in the coordinated multilateral model. Thus the coordinated multilateral trading model is superior to any scheme that relies on centralized dispatch both on grounds of incentive compatibility and information efficiency.[18]

4.3 Trading arrangements

We have proved in the appendix that if the current operating point is not optimal, participants can always find a trade that is feasible and profitable (Lemma 4). We discuss in this subsection how to discover and arrange such trades. Specifically, the following questions are to be answered:

- How many parties need to be involved in the trade?
- Who should be contacted for negotiation?
- What is the optimal level of generation (or consumption) for each party in the trade?

Section 8.4 in the appendix provides detailed answers to these questions. Before getting to the specifics, let us demonstrate how the information provided by the Loading vector can be utilized. For a network, for example, with 8 nodes, the Loading vector for a particular transmission congestion is an 8-dimensional vector, $(n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8)$. The number n_i is the amount of additional MW power flow on the congested line (or lines) if 1 MW of power is injected into bus i. For simplicity, assume there is only one congestion. Consider a trade, say involving generation at nodes 1 and 3, and consumption at node 5. Let the amount of generation at node 1 be q_1 MW, at node 3 be q_3 MW and the amount of consumption at node 5 be q_5 MW.[19] First of all, the total generation of the trade must equal to the sum of the consumption and the losses, i.e.,

$$q_1 + q_3 - q_5 = q_{loss}$$

The transmission system security constraint is expressed using the Loading vector as:

$$n_1 q_1 + n_3 q_3 - n_5 q_5 \leq 0$$

The two equations above impose constraints on q_1, q_3 and q_5. A broker must find a profitable trade within these constraints.

---

18. The literature on the design of coordination mechanism is concerned with reducing the amount of information transfer and ensuring the compatibility of the incentives of the individual participants and the goals of coordination. See: Hurwicz, L. "On informational decentralization and efficiency in resource allocation mechanisms", in S. Reiter (ed), *Studies in Mathematical Economics*, Mathematics Association of America, 1986, and Reichelstein, S. "Incentive compatibility and informational requirements", Journal of Economic Theory, vol 32, pp. 384-390, 1984.

19. To avoid confusion with negative numbers we depart from the sign convention used in the appendix. In the same spirit, here MC_5 represents marginal benefit.

tional overload. The broker can simply calculate the net increase in transmission loading [10.3 x n_3 - 10 x n_5] and if this is less than zero, there is no problem. To summarize, the coordinated multilateral trading process goes like this: The PSO calculates the Loading vector and broadcasts this set of numbers to everyone. (Data required for calculating the Loading vector is basically what is required for a load flow study, i.e., transmission network data and power generation and consumption data for the trades.) Based on the Loading vector, brokers can privately arrange profitable trades that are feasible, i.e., trades that do not result in overload. A more detailed explanation of how to arrange trades is provided in the next subsection.

Since power flow relations are nonlinear, the Loading vector based on local linear approximation of the nonlinear functions is valid if the magnitude of the trade is small. A nonlinear version of constructing feasible trades by way of geodesic curves is derived in the appendix. The required computations, however, may become intractable and the result may be only of academic interest. In practice, however, it is well-known that the difference between the non-linear and linear approximation for real power flows (MW flows) is usually small. Furthermore, as discussed previously, transmission limits are soft and there is usually an inherent safety margin in it. Small technical breaches of the limits for a short duration may not necessarily lead to actual security threat. If on the other hand, linear approximation is not acceptable, a practical alternative exists, where the PSO requires that a safety margin be added to the calculation to account for the error in linear approximation. Suppose that x is the estimated percentage of error in linear approximation, say x is equal to 2%, the requirement for the above example becomes [10.3 x n_3 - 10 x n_5] < - 0.2.

It is possible that a profitable trade between participants can result in overload of another transmission limit besides the original congestion. If the trade potentially overloads an additional transfer limit, the PSO must solve a set of equations similar to those concerning the original transmission congestion to ensure that the additional transfer limit is not overloaded. In other words, the PSO must curtail the trade to insure that this transfer limit is not overloaded and then broadcast the Loading vector corresponding to this constraint. The PSO thus sends to participants two vectors: a Loading vector for the original transmission congestion and a Loading vector for the additional congestion. With these two vectors, participants can negotiate trades that are profitable while not resulting in overload for either the original congested transmission or the additional transfer limit. In the general case, if several of transmission transfers are congested, the PSO can provide Loading vectors, one for each transfer limit and participants can use these vectors to ensure that their trade does not overload any of them.

We have proved that, even without a central authority, participants in the coordinated multilateral trading model reach an optimal operating point. As shown in the appendix, if the current operating point is not optimal, participants can always find a trade that is feasible and profitable, i.e., all parties involved in the trade will profit. Participants have clear incentives to find such trades. The sequence of coordinated multilateral trades continues until there is no more profit to be made and an optimal solution is reached. (See appendix, theorem 1).

4.2 Feasible trades

A trade involves suppliers and consumers. Cost of generation is associated with supply and benefit is associated with consumption. A *profitable* trade is a trade in which the total benefit of all loads involved in the trade is greater than the cost of all generators. In such a case, a broker can profit by arranging the trade.[16] We assume that a broker's principal objective is to arrange profitable trades.

Suppose that the security of the transmission network is not a constraint to the supply/demand market. Since transmission losses are allocated and included in the trades, the transmission network may be ignored. The situation is simply a competitive supply and demand market and the standard economic theory applies. Assuming no participant has significant market power, an equilibrium or optimal will be reached. At the optimal point, the marginal cost of all generators and the marginal benefit of all consumers are equal, if losses are ignored, otherwise they differ by the "penalty factors" for losses. In other words, if brokers arrange profitable trades for the participants and if the trades do not result in overloading the network, an optimal solution will eventually be reached, without the intervention of the PSO.

If the trades result in transmission loading exceeding a transfer limit, in our coordinated multilateral trading model, the PSO, whose primary responsibility is to maintain security, curtails the amounts demanded in the trades to a point at which overload is no longer possible. The curtailment scheme, i.e., whose generation is reduced and by how much, does not rely on the cost-benefit information of the participants since the PSO is only concerned with system security, and the PSO's only function is to relieve overload, not to satisfy economic concerns. The curtailment scheme is fairly arbitrary and a reasonable one is perhaps to select the most effective way to reduce overload. Because economic decisions are separate, the participants will make additional trades to improve the economic efficiency of the system after being curtailed by the PSO.

To insure that these additional trades do not result in overloading the transmission system requires coordination by the PSO. The PSO provides guidance to all participants as to what trades are allowed. This is not done on a case by case basis; rather, the PSO gives out sufficient public information based on which brokers can privately construct profitable trades that are compatible with secure operation of the network. We claim that this can be accomplished by the use of *sensitivities*, i.e., the increase in transmission loading (on the congested transmission line or the aggregate transfer between two regions) if one megawatt of power is injected into the network at a node. The PSO has the responsibility to obtain this data for every bus of the network. We call this set of numbers the Loading vector.[17] The kth component of the Loading vector, n_k, is the additional loading on the congestion if one MW of power is injected into node k. With Loading vector publicly known, a broker can easily check whether a planned trade, say, 10.3MW generated from bus 3 and 10MW delivered to bus 5, will cause addi-

---

16. The broker has to pay for the use of the transmission network, which is the subject of transmission pricing and will be discussed in Sec VII.

17. Network reduction mentioned in footnote 14 to reduce the dimension of network calculation, hence the size of the Loading vector, is more effective for security analysis, especially for real-power flows.

order to relieve congestion or to ensure security of operation, it is *essential* to have coordinated trades involving three or more parties.

4.1 Loss-included trades

In general, each trade should include its share of transmission losses. In other words, the net generation should equal to the sum of the demand and the transmission losses caused by the trade. The total transmission losses caused by all the trades on the network can either be measured or calculated. But the allocation of the total losses to individual trades is nontrivial. It is argued that due to nonlinear nature of the loss relationship, it is impossible to allocate losses to individual trades in a theoretically correct way.[13] If the transmission loss were a linear function of the generation, the allocation would be possible. Linear approximation of losses has been used, but it can be shown that linear approximations always over-estimate, hence over-charge, the losses caused by the trades.

We show in the appendix that, whereas unattainable for general nonlinear functions, an exact allocation method can be obtained if the loss relation is a special type of nonlinear function, namely, a quadratic function. It is well-known that the relation for transmission loss is almost quadratic. Indeed the error in quadratic approximation is usually no more than a few percentage points. Using the quadratic approximation, we have derived a formula enabling a broker to estimate the transmission losses for any planned trade, and a formula to estimate the effect of one trade on losses (increase or decrease) to another trade. For the exact allocation of transmission losses to individual trades we have derived an allocation method based on the percentage allocation derived from quadratic relationship.

A broker estimates the transmission losses caused by a trade based on the Loss vector calculated and broadcast by the PSO. For an n-node network, a Loss vector is a set of n numbers.[14] Suppose that the planned trade, ignoring losses, is 10MW from bus 3 to bus 5 and the 3rd and 5th positions of the Loss vector are 0.041 and 0.011, respectively.[15] The estimated allocated losses then are 10 x 0.041 - 10 x 0.011 = 0.3 MW. If it is desired for the brokers to estimate the effect of other trades on their share of losses, the PSO calculates and broadcasts a quadratic loss matrix, based on which the brokers can estimate these effects. An example is shown in Sec. V.

---

13. Moreover, even if that were possible, it is argued that it is still theoretically impossible to balance power generation to exactly match the allocated losses in a power network, because if the broker of the trade adjusts the generation level to account for the losses, this changes the generation level and affects the total losses. We show in the appendix that there is a solution to this problem too. It is possible to derive a loss-included trade by an iterative process. The process goes as follows: (1) participants plan a proposed multilateral trade and initially ignore losses; (2) participants use the loss allocation scheme to calculate their share of losses for the proposed trade; (3) participants adjust generation levels to account for the losses; and (4) if necessary, recalculate the loss adjustment and undergo an iterative process. We have derived conditions under which the process converges. Our tests show that the iterative process is more of academic interest. Usually after a single pass, the discrepancy is negligible.

14. Several methods have been proposed for network reduction, see for example the survey paper: Wu, F. F., and A. Monticelli, “A critical review of external network modeling for on-line security analysis,” *Electrical Power and Energy Systems*, vol. 5, 1984, pp. 623-636. If network reduction can effectively construct an equivalent lower order network model the computational effort may be reduced. However, none of the existing equivalents gives a good approximation of losses. Further research is needed.

15. We shall call a node in a power network a bus--a commonly used terminology in power system engineering.

In a competitive market, if each trade can determine how much losses it causes, preferably at the moment while economic decision is being made, the one who arranges the trade can then decide whether to cover these losses through its own additional generation or through external purchase. We stress the fact that it is important that the customer of any transmission service must have a choice in a competitive environment.

Power system operation also requires the maintenance of real-time power balance even if all scheduled powers are balanced because of uncertainty, generation and load fluctuations, small amounts of loss adjustments and possible generation and transmission facility failure. Generation reserve and AGC (automatic generation control) are means to achieve real-time power balance in today's centralized power system operation. To distribute the decision-making responsibility of real-time power balance to individual trade requires three conditions: (1) the exact amount of losses caused by each trade is known, (2) metering capability is available on all participants in a trade, and (3) real-time control devices are available to match generation to the load and the losses. The first condition can be met by theoretical analysis, and is the subject of discussion here, whereas the other two require technological solutions, that are functions of market incentives and technology innovations.

3.4 Requirement summary

We propose a new operating paradigm which allows suppliers and consumers primarily to seek profit on their own, while the PSO guarantees security. Only when security is threatened, does the PSO intervene to make a decision on curtailment. Participants collaborate in the responsibility for maintaining security with the help of information provided by PSO. For the purpose of maintaining security, the PSO passes on information to the generators and loads based on which the generators and loads structure their trades so as not to cause security problems. Because maintaining security is a community effort, all security information is open to the public.

Suppliers and consumers themselves carry out the economic function, making the decisions concerning the price and other terms and the amounts to buy and sell. The pricing information relied upon by participants for their economic decisions is all private. Since all trades are independently arranged, methods to calculate and allocate transmission losses are required.

Ultimately, the goal of this new paradigm is to: (1) organize arrangements so as to achieve economic efficiency; and (2) encourage search for alternatives and innovations for any function that requires a centralized authorized.

IV. Coordinated Multilateral Trading Model

A multilateral trade is a trade involving two or more parties in which the sum of generation minus the sum of consumption loads is equal its share of losses. The party that arranges the trade is called a broker. The broker may be a generator or consumer involved in the trade but may also be an unrelated third party. A multilateral trade is a generalization of a bilateral trade. But, as will be shown later, in

3.2 Separating security and economy

As competition is introduced in generation, thus breaking up the monopoly, we believe that the decision-making for efficiency and reliability can and should be decentralized and it can be done with a proper information structure. First of all, it is helpful to analyze security and economic generation separately by examining the differences in decision-making and information structure requirements for the two functions.

In analyzing the task of economy, we must remember the goal of a market based economy. In a free market, decision-making authority should be decentralized and decisions should be made by participating suppliers and consumers. The information structure in a free market should facilitate economic efficiency. Furthermore, the information structure should not result in a situation in which there is market dominance or anti-competitive gaming opportunity. To achieve that, it is necessary to keep the cost/benefit information completely private.

The security of the transmission system is a shared responsibility among all generators and consumers. With today's technology, there is a delay in communication and control that makes it impossible or undesirable to mitigate a security violation after it occurs. Therefore, in system operation, the security concerns must be incorporated while economic decisions are being made. This is consistent with the principle of current practice in which security is treated as an operating constraint in economic decisions. The challenge, however, is to determine the degree of centralization and decentralization in the decision-making process and to design an information structure such that sufficient information can be made available to all potential trades to assess security feasibility and economic viability. Moreover, for fair competition, public information should be shared and transparent, i.e., the method by which the data is obtained and security constraints are calculated must be obvious and readily reproducible by anyone.

3.3 Separating scheduled and real-time power balance

As explained before, the task of maintaining power balance after resources are in place can be divided into two stages: scheduled and real-time. Security/reliability consideration is treated as a constraint in the maintenance of power balance. In a lossless transmission system where balance power simply means supply equal to demand, the decision-making authority for scheduled power balance can be totally decentralized, as long as each trade meets the condition that its supply equal to its demand. It becomes a problem if the transmission losses in a shared network must be considered because the losses caused by a trade depend on other trades on the network. Typically transmission losses are about three or four percent of the total generation. This may not sound significant, but the cumulative effect of losses on a year-long contract can mean millions of dollars.[12]

[12]. An EPRI study clearly indicates the importance of estimating transmission losses in traditional utility transmission planning, see: Scientific Systems, Inc., “Bulk transmission system loss analysis: theory and practice”, *EPRI Report* El-6814, 1990 and also: Nadira, R., F. F. Wu, D. J. Maratukulam, E. P. Weber, and C. L. Thomas, “Bulk transmission system loss analysis”, *IEEE Trans. Power Systems*, vol. 8, 1993, pp. 405-413.

fied version of optimal power flow, called *transmission constrained economic dispatch*, has been implemented in some of today's advanced EMS. Instead of using a full-fledged power flow model in the formulation, transmission constrained economic dispatch uses network sensitivities, which are quantities relating the changes in generation to the loading in transmission. Therefore, unlike ordinary economic dispatch, which adjusts generation level with an eye only for economic efficiency, transmission constrained economic dispatch adjusts generation levels while considering both economics and transmission loading limits.

III. Requirements for a New Operating Paradigm

3.1 Coordination

For the transmission system to properly support "shared" services such as maintaining security and providing transmission losses, coordination among all parties is required. It is a gross simplification to equate coordination with centralization. As a matter of fact, an examination of coordination should focus on ways to distribute information (who knows what) and control (who does what) for all parties: generators, consumers, system operator, regulator, etc., to achieve the goals. In the traditional paradigm, both information and decision-making authority are centralized to achieve the three main operating objectives: power balance, security/reliability, and economy. As we move toward increased competition, the way in which the traditional paradigm relies on centralized authority may not be necessary or desirable to attain coordination. It is possible to reach the same level of coordination through different information and decision-making structures. With this objective in mind, we seek to (i) separate each objective (power balance, security/reliability, and economy) and (ii) design information and decision mechanisms for each objective. We can then evaluate the alternative operating paradigm in terms of economic efficiency, reliability and implementability.[11] The goal is to find a new operating paradigm that achieves at least the same economic efficiency and the same level of reliability as the centralized paradigm does, and hopefully more, and which is reliable and implementable using existing computer, communication and control (3C) infrastructure.

---

10. The optimal power flow (OPF) was formulated by Carpentier, J. W., "Contribution a l'etude du dispatching economique", *Bull. Soc. Fr. Elec.*, vol. 3, 1962. Among the early implementations of OPF on large system off-line studies are: Dommel, H. W., and W. F. Tinney, "Optimal power flow solutions", *IEEE Trans Power App. Syst*., vol. 87, 1968, pp. 1866-1876. Wu, F. F., G. Gross, J. F. Luini, and P. M. Look, "A two-stage approach to solving large-scale optimal power flows", *Proc. of Power Industry Computer Applications*, Cleveland, 1979, pp. 126-136. Practising engineers are still trying to solve many detailed OPF implementation problems, see for example: Papalexopoulos, A. D., C. F. Imparato, and F. F. Wu, "Large-scale optimal power flow: effects of initialization, decoupling and discretization", *IEEE Trans. Power Systems*, vol. 4, 1989, pp. 748-759. A survey of developments in OPF and TCED can be found in: Wu, F. F., "Real-time network security monitoring, assessment and optimization," *Electric Power and Energy Systems*, vol. 10, 1988, pp. 83-100, and also in: Stott, B., O. Alsac, and A. Monticelli, "Security analysis and optimization", in Wu, F. F., and R. D. Masiello (ed.), *Computers in Power System Operations*, Special Issue of the *Proceedings of the IEEE*, vol. 75, Dec. 1987.

11. Several alternative information structures have been proposed: Kaye, R. J., F. F. Wu, P. Varaiya, "Pricing for system security," IEEE PES Winter Meeting, New York, Jan. 1992, Baldick, R., R. J. Kaye, and F. F. Wu, "Electricity tariffs under imperfect knowledge of participant benefits," *IEEE Tran. Power Systems*, vol. 7, 1992, pp. 1471-1480.

bances, called credible contingencies, establishing all loading levels and running simulations to ensure system security.[9] In operation, when a transmission line or the power transfer between two regions reaches its loading limit, the transmission network is said to be *congested* and the congestion must be relieved. The fact that these limits must be established beforehand leads to subjectivity in determining those limits. The PSO must select which conditions to study, i.e., the credible contingencies and supply/demand levels, and translate the simulation results into transfer limits. Because the selection of conditions and interpretation of results depend on the PSO's judgment, the transfer limits can be somewhat imprecise and subjective.

2.3 Economic generation

The PSO in the traditional paradigm also has the responsibility of insuring economic generation. The PSO possesses the fuel-cost curve of each generator. Based on these curves, the PSO performs *economic dispatch* to insure that the overall system production cost is minimized. Classical economic dispatch does not consider transmission limits, but does consider transmission losses in the dispatch. Transmission losses in the dispatch are incorporated by the use of *penalty factors* against generators. In the absence of losses, economic dispatch requires all generators to operate at the same marginal cost. With losses, each generator's marginal cost is multiplied by a penalty factor, a number that reflects the relative contribution of that generator to transmission losses.

The operation of economic dispatch is fairly standard in today's control center or EMS (energy management system). The PSO runs economic dispatch every few minutes using real-time data on generation and load demand. Penalty factors against generators are calculated in one of two ways: (1) the factors are calculated in advance using planning data and, typically, several sets of penalty factors are stored in the EMS computer; or (2) the factors are calculated by using real-time data and on-line load-flow calculations. Real-time data on the generation level of each generator in the system and the aggregate load demand at each substation are pooled throughout the system every two to four seconds. This data proceeds through a data collection system called SCADA (supervisory control and data acquisition system) to the control center. Real-time data are also used to maintain instantaneous power balance through AGC (automatic generation control), which calculates the deviation of actual from scheduled frequency (an indication of power imbalance) and interchange schedule and then resets the generation in the system according to a set of *participation factors* derived from the result of economic dispatch.

Traditionally, security considerations override economic considerations. In practice, this means that when overload occurs on a transmission line or power flow between two regions exceeds transfer limit, generation is curtailed with little consideration to economy. Recently, however, attempts have been made to combine security and economics by including transmission limits in a generalized economic dispatch called the *optimal power flow*. Computation of such an optimal power flow is involved and time-consuming and there are many practical problems in real-time implementation.[10] A simpli-

[9]. The studies include transient stability, dynamic stability, load flow etc. NERC, established after the 1965 Northeast blackout, and its member regional reliability councils set the guidelines for credible contingencies.

2.2 Security and reliability

In power system operation, in addition to the requirement that power must be balanced for the present situation, power balance must be maintained to guarantee continuity of supply even after a disturbance in the system such as the failure of a generator or the outage of a transmission line. The continuity of service to consumers is referred to as the *reliability* of the system. To insure reliability, prudent measures must be taken from both the generation side and the transmission side. More generation capacity than necessary for the forecasted load is scheduled. In other words, the traditional paradigm requires reserved generation, unused but readily available, for unforeseen events of generation failure or sudden demand surge.

The transmission system adds another dimension for the maintenance of reliability of system operation. There are steady state operating limits, generally referred to as thermal limits, for all transmission facilities, i.e., transmission lines and transformers. Kirchhoff lows dictate how power flows in the network. The individual generator has very limited control over power flow through a particular line. The whole network, together with all generation and load consumption connected to it, determines the loading of a transmission line.

Any power system is subject to unpredictable disturbances, e.g., lightning strikes and disables a transmission line or a generator fails. Such changes disturb the power balance and, as a result, power must be redistributed. Such redistribution sometimes results in sudden surges of power on other transmission lines. A protective system is installed that guards against overloading any facility in the system by disconnecting the affected device. Another possible result of the power surge is that, upon overloading one line, the redistributed power may overload another line and the protective system will disable the second line. Such repeated overloading and disabling is called the cascading effect and can eventually lead to total system black-out.[8] From the system operator's point of view, the ability to avoid these cascading outages as a result of disturbances in the power system is called system *security.* As we have mentioned previously, from the consumer's point of view, continued service is the main concern and is called *reliability.* Therefore, security and reliability concerns of a transmission network are identical, although the concerns are seen from different vantage points.

As explained above, disturbances result in redistribution of power. During this redistribution transient, the system must be stable and it must eventually reach a new power balance in which no facilities exceed their thermal limits. The ability of a power system to achieve that depends on the distribution of supply and demand of power in a network, or the loading of the transmission system. To maintain security, the PSO must make sure that the loading on a specific transmission line or a set of lines between two regions in the system is not too high. We shall call such limit as *transmission loading limit* or *transfer limit.* In practice, the setting of the transmission loading or transfer limit, must be performed in advance. The PSO does a set of studies ahead of time by postulating a set of distur-

[8]. Such cascading effects have been attributed to the two major blackout in recent history: 1965 Northeast blackout and the 1977 New York city blackout. A clear exposition of the 1977 New York blackout is described in: Wilson, G. L., and P. Zarakas, “Anatomy of a blackout”, *IEEE Spectrum,* vol. 15, 1978.

II. Power System Operation

In this section, we will review the conventional paradigm for power system operation and the methods by which it achieves three main operating objectives: power balance, security/reliability, and economy.[7]

2.1 Power balance

In a power system, power balance must be maintained, i.e., load demand must be balanced by generation supply at all times. In the traditional paradigm, the PSO makes a series of coordinated decisions to insure power balance, which breaks down into three basic stages. First, the PSO performs *resource acquisition*. This stage involves choosing the types of supply-side or demand-side resources to meet forecasted load demand, often years in advance. This load forecast is crude. The second stage, taking place a day, week, or year in advance, is called *scheduling*. Scheduling involves several functions: (1) hydro scheduling: the PSO determines the schedule for next year's or next week's reservoir level, i.e., the amount of water to run and the time at which to run it; (2) maintenance scheduling: the PSO determines when to shut down the thermal power plants for maintenance; (3) schedule interchange between companies; (4) unit commitment: the scheduling decision concerning which generator to fire up or shut down for the next week. These two stages involve pre-planning well in advance of real time; thus, although it is true that instantaneous demand fluctuates, the average demand is not entirely unpredictable and a power system can anticipate a large component of the demand fluctuation and schedule resources in advance to meet the demand.

The third power balance stage is *dispatch,* which is done in real time. Dispatch includes the following PSO functions: (1) economic dispatch: the PSO raises or lowers generation levels according to fuel costs, typically performed in minute intervals; and (2) automatic generation control: the PSO balances power instantaneously on a second-by-second basis.

To a lesser extent (in terms of the level of complexity and sophistication, certainly not in terms of importance), similar scheduling and dispatch functions have been developed for reactive power to maintain the voltage level throughout the network within an acceptable range.

It is important to note that taking care of transmission losses is included in the function of power balance. In traditional power system operation, losses are considered a part of total generation necessary to meet the aggregate customer demand, and are supplied from a pool of generators belonging to the same company, and therefore, do not impose any problem. The allocation of transmission losses may impose a problem in a competitive generation market, as we shall see later.

---

[7]. Standard textbooks in power system analysis include: Bergen, A. R. *Power Systems Analysis*, Prentice-Hall, 1986. Wood, A. J. and B. F. Wollenberg, *Power Generation, Control, and Operation*. New York: John Wiley and Sons, 1984. Anderson, P. M. and A. A. Fouad, *Power System Control and Stability*, IEEE Press, 1994. Survey papers on recent advances include: Wu, F. F., “Stability, security, and reliability of interconnected power systems,” *Large Scale Systems*, vol. 7, 1984, pp. 99-113. Wu, F. F., “Real-time network security monitoring, assessment and optimization,” *Electric Power and Energy Systems*, vol. 10, 1988, pp. 83-100. For computer-controlled power system operation in an energy management environment, see Wu, F. F., and R. D. Masiello (ed.), *Computers in Power System Operations*, Special Issue of the *Proceedings of the IEEE*, vol. 75, Dec. 1987.

poolco model -- were fundamentally flawed in that they assumed the operating paradigm as a given constraint and then attempt to construct a market around it. The current operating paradigm, however, developed in a different era for a different industry structure, namely, the regulated monopoly. There is no inherent incentive in the traditional operating paradigm to promote competition; it was never design to do so. Because both the bilateral model and the poolco model embrace the traditional paradigm, both models require regulation in order to force the PSO to perform its duties. Therefore, it is not surprising that both models result in significant gain in the requirement for regulation, contrary to the original intent.

We take a completely different approach. We propose to develop a new operating paradigm that is compatible with the competitive market structure. The crowning achievement of the traditional operating paradigm is its seamless coordination of transmission operation. Coordination does not require centralization. Coordination among various parties has two dimensions: information structure and decision-making authority.[6] In the traditional operating paradigm, both the information structure and decision-making are centralized. We have examined alternative information and decision-making structures for a new operating paradigm to achieve the same level of coordination and have developed a model which we call the Coordinated Multilateral Trading model. We have solved the two fundamental problems, namely ensuring system reliability/security and allocating losses, that weaken the bilateral model. Moreover, our model will achieve the same economic efficiency as the poolco model ideally achieves. But the PSO has no hand in the economic decisions. Efficiency is attained through the invisible hand of the market. The proposed model does not require explicit description of cost/benefit functions as other models do- such requirement in an economic system can lead to welfare loss. Therefore the coordinated multilateral trading model will achieve higher economic efficiency.

Section II of this paper reviews the conventional paradigm for power system operation and Section III discusses the requirements for a new operating paradigm that is compatible with a competitive market arrangement. In Section IV we will describe the basic process of our proposed coordinated multilateral trading model. The complete mathematical theory underlying the coordinated multilateral trading process and a loss allocation scheme is presented in the appendix. In Section V we work out an illustrative example of the multilateral process. The coordinated multilateral trading model can be implemented with today's computers, communication and control infrastructure. These implementation issues are discussed in Section VI. It is also shown that the coordinated multilateral trading model can coexist with today’s vertically integrated utility model and provides non-discriminatory services to all. Furthermore, voluntary pools can be operated within the proposed model. A summary conclusion is included in Section VIII.

---

6. These concepts are borrowed from the theory of coordination. Coordination theory studies the specification of the “rules of the game” by which the decisions of various agents can be coordinated in order to achieve a common goal.

be included in power balance. The energy loss in the transmission network is a function of the aggregate trades, and therefore the trades are not really independent if they have to account for their own losses. In the traditional vertically integrated utility, coordination is achieved by the PSO who ensures instantaneous power balance without overloading transmission facilities and maintains the network's ability to withstand disturbances.

Faced with the need for coordination of trades to ensure system safety and power balance, bilateral model advocates resort to the traditional model. They propose to grant the PSO the authority to determine the safe level of generation (if the aggregate of individual trades is not safe) and to compensate for losses. But with the introduction of this centralized authority, the bilateral model is metamorphosed into poolco---the model to which it was originally counterposed.

In the poolco model all utilities combine to form a “super-utility” in the form of a poolco, and the market structure is altered to suit this super-utility. Suppliers and consumers offer price and quantity bids to poolco for traditional bundled services and the Pool System Operator (PSO) keeps the traditional responsibilities, such as ensuring instantaneous power balance, maintaining network reliability and security, and coordinating transmission access and services. In place of the free bilateral trades between generator and consumer, every “trade” is now essentially required to be with the poolco, or rather with the centralized PSO. The PSO determines which trades to accept and execute and which trades to reject so that the system is safe, and sets the price at which trades are settled so as to promote economic efficiency. Thus to the PSO's traditional role in the vertically integrated utilities, poolco would add the powerful role of marketmaker. With its dictatorial power, the Pool System Operator can, in principle, enforce any of a large number of operating points. Advocates refer to this obvious point when they assert that the poolco can operate efficiently. However, the PSO has no incentive to operate efficiently. Because of its increased authority, the PSO must be regulated---indeed regulated more closely than utilities are today.

Advocates of poolco also liken it to the New York Stock Exchange. This analogy could not be more misleading. The reason that stock prices at the NYSE are competitive is because anyone who owns a share can sell it to anyone who wants to buy it, as in the bilateral model, and this guarantees competition. By contrast, in the poolco model, anyone who wants to buy power must buy it from poolco, and anyone who wants to sell power must sell it to poolco. Thus poolco is a (regulated) monopolist that consumers face, and a (regulated) monopsonist that generators face. In other words, poolco is a cartel. Because of the extreme complexity of power system operation, the ability of the regulators to regulate effectively today’s utilities has been called to question, their ability to regulate a much more powerful utility cartel- poolco- is doubtful.[5]

We believe that any restructuring model, no matter how noble its goal, if its proper functioning relies on external enforcement, rather than internal economic incentives, is in danger of being derailed in practice. The two approaches that arose from the CPUC proposal -- the bilateral model and the

5. In its remarkably candid Yellow Book the CPUC recognizes its inability to regulate effectively today's utilities; yet, by embracing the poolco model, the CPUC also proclaims its ability to regulate a much more powerful utility cartel.

These recent moves to open up the electric power industry demonstrate the new general agreement on three principles: (1) pressure for greater competition should continue; (2) consumer choice should be enhanced; and (3) access to transmission services should be arranged in ways necessary to accommodate consumer choice and supply competition. These new competitive pressures and emerging market forces exert an increasing influence over the future structure of the industry.

When the CPUC issued its proposal promising direct access and open competition by the year 2002, it touched off a debate over how the transmission system should be restructured in order to meet that goal. The opposing sides of this debate are now commonly represented as the *bilateral model* and the *poolco model.*[2]

The bilateral model is based on the principle that free market competition is a route to economic efficiency. In this model suppliers and consumers independently arrange trades, setting by themselves the amount of generation and consumption and the corresponding financial terms, with no involvement or interference by the power system operator (PSO).[3] Economic incentives will lead generators to find the best-paying customers and consumers to find the cheapest generators. So long as consumers or generators do not have significant market power, these trades will lead to short term economic efficiency. Perhaps even more important in the long run, generators and other providers will find it profitable to support innovations that consumers want. However, the bilateral model faces two fundamental problems which detract considerably from its ability to promote free market competition.

First, the lack of coordination among the independent trades can lead to a violation of transmission network constraints. The network constraints arise from loading limits on transmission equipment and from the requirement that the network be operated in a secure state.[4] Coordination is necessary because physical laws dictate how power goes from the generators to the consumers in a transmission network. An individual generator has very limited control over transmission loading on a particular facility. Second, power flows must be balanced throughout the network and transmission losses must

---

2. Several articles in the Electricity Journal, September 1994, articulate both sides of the debate very well. Poolco model is largely based on the system being implemented in England and Wales, a detailed description of the basic structure of the system can be found in: White, A., “The electricity industry in England and Wales,” James Capel & Co., Feb. 1990. For the conceptual foundation of poolco, see Hogan, W. W., “Contract networks for electric power transmission,” *J. Regulatory Economics,* vol. 4, 1992, pp. 211-242, and Ruff, L., “Stop wheeling and start dealing: resolving the transmission dilemma,” *The Electricity Journal,* June 1994, pp. 24-43. The bilateral model has been implemented in Norway, see: Moen, J., “Electric utility regulation, structure and competition. Experiences from the Norwegian electric supply industry,” *Norwegian Water Resources and Energy Administration,* Aug. 1994. A critical appraisal of both models can be found in: Wu, F. F., P. Varaiya, P. Spiller, and S. Oren, “Folk theorems on transmission open access: proofs and counterexamples,” *POWER report* PWP-23, U of Calif Energy Inst., Oct 1994, and Oren, S. S., P. T. Spiller, P. Varaiya, and F. F. Wu, “Nodal prices and transmission rights: a critical appraisal,” *The Electricity Journal*, vol. 8, April 1995, pp. 14-23. There are variations in different bilateral and poolco proposals; we consider the basic models here. In the basic poolco model, nodal prices are combined with “transmission contracts” that give rights to transfer power from one node to another. In a very interesting paper, Chao, H-P. and S. Peck, “Market mechanisms for electric power transmission,” May 28, 1995, preprint, show that transmission contracts can be replaced by spot markets for capacity on individual lines. To transfer power from node 1 to node 2, a generator must purchase sufficient capacity on all links to accommodate all the “parallel” paths in the network along which the power will flow.

3. The term PSO used here is rather general, it represents the organization with facilities to carry out the necessary functions.

4. There is also need to coordinate the use of other shared resources, e.g., back-up generators.

I. Introduction

Recently, regulated industries such as telecommunications, trucking, airline and gas have experienced major changes resulting from reduced regulation and increased competition. The electric utility industry is the last major regulated monopoly to undergo change. The traditional regulated monopoly is justified only in an environment in which (1) economy of scale exists in the industry, and (2) the pace and magnitude of technological advancement remains moderate and predictable. As social acceptance and financial viability of large generators have declined and possibilities for innovations in electrical, as well as supporting information technologies, have soared, the continued monopolization of the electric utility industry becomes untenable.

The Public Utility Regulatory Policy Act (PURPA) of 1978 essentially ended the complete monopoly of electric generation. PURPA introduced *qualifying facilities* (QFs) which resulted in competition in the generation sector. Today, the non-utility enterprises constitute a multi-billion dollar industry and are important players in the electric supply business. The Energy Policy Act (EPAct), enacted in 1992, profoundly altered federal policies governing generation and sale of electric power in the wholesale market. EPAct embraced the prevailing ideas of increased competition and a greater reliance on the market mechanism as the preferred means to develop, deliver and market energy services in the United States. The California Public Utilities Commission (CPUC) in April 1994 issued a bold proposal to restructure California's electric power industry. Under the CPUC proposal, the electric utilities would be required to provide retail wheeling (direct access) to all customers who choose to do so by the year 2002.[1]

The initial focus of competition in the electric energy industry was generation. However, it soon became clear that a chief impediment to increasing competition in electric systems was access to the transmission system. Electric utility companies owned the transmission system, and the non-utility generators needed transmission access to be competitive. Eventually, non-utility generators demanded open transmission access. EPAct reformed federal transmission policy and broadened the power of the Federal Energy Regulatory Commission (FERC) to mandate wheeling by transmission owning utilities (TOUs). Under EPAct, the TOUs were required to respond to requests for transmission services by any electricity wholesaler, and FERC could issue transmission orders, provided that such orders served the public interest and did not unnecessarily impair the continued reliability of the TOUs' system. The recent FERC Mega NOPR asks utilities to provide comparable services to their own customers and to wheeling customers and to propose tariffs for various transmission services.

---

1. An excellent account of the changes in electric utility industry with a California focus can be found in: CPUC Division of Strategic Planning, "California's Electric Services Industry: Perspectives on the Past, Strategies for the Future," February 1993 (commonly called the Yellow Book). The CPUC proposal is in: Order Instituting Rulemaking and Order Instituting Investigation on the Commissions Proposed Policies Governing Restructuring Californias Electric Services Industry and Reforming Regulation, R.94-04-031 & I.94-04-032. April, 1994 (commonly called the Blue Book). The long-waited CPUC decision, announced in May 1995, however, was considered anticlimactic by some. The headline in the *San Francisco Chronicle* expressed this sentiment: "PUC passes plan to pool power supply: interim proposal would delay utility competition, critics say," (May 25, 1995, front page).

Coordinated Multilateral Trades for Electric Power Networks: Theory and Implementation

Felix F. Wu and Pravin Varaiya

Department of Electrical Engineering and Computer Sciences
University of California
Berkeley, CA 94720-1770

Internet: ffwu@eecs.Berkeley.EDU, varaiya@eecs.Berkeley.EDU

Abstract

Recent moves to open up electric power transmission networks to foster generation competition and customer choice have touched off a debate over how the transmission system should be restructured in order to meet the goal. The opposing sides of this debate are now commonly represented by the bilateral model and the poolco model. Both models resort to conventional centralized operation in dealing with the shared resources of an integrated transmission network. The conventional operating paradigm was developed in a different era for electric utilities operated as regulated monopolies. A new operating paradigm is needed for a restructured industry that encourages efficient competition and at the same time maintains necessary coordination to guarantee a high standard of reliability. We propose a new operating paradigm in which the decision mechanisms regarding economics and reliability (security) of system operation are separated. Economic decisions are carried out by private multilateral trades among generators and consumers. The function of reliability is coordinated through the power system operator who provides publicly accessible data based upon which generators and consumers can determine profitable trades that meet the secure transmission loading limits. We prove that any sequence of such coordinated private multilateral trades, each of which benefits all parties to the trade, leads to efficient operations, i.e., maximizes social welfare. The coordinated multilateral trading model achieves all the benefits of a centralized pool operation without the visible, heavy hand of a pool operator in economic decisions. It is shown that the existing communications, computing and control infrastructure is adequate to support the proposed model. It is also shown that the coordinated multilateral trading model can coexist with the traditional model and provides non-discriminatory service to both utility customers and direct-access customers.

August 19, 1995