Real-Time Co-Optimization: Interdependent Reserve Types for Primary Frequency Response *

Manuel Garcia†
University of Texas at Austin
Austin, Texas

Ross Baldick‡
University of Texas at Austin
Austin, Texas

ABSTRACT

As renewable energy penetration increases and system inertia levels drop, primary frequency response is becoming a critical concern. To address this problem the Electric Reliability Council of Texas (ERCOT) is introducing a Fast Frequency Responsive (FFR) reserve product to the electricity market that helps to arrest frequency decline in the event of a large generator outage. This reserve type is required to be capable of nearly instant deployment of allocated reserve, which differs from traditional Primary Frequency Responsive (PFR) reserve that exhibits ramping limitations on power output. This paper derives a real-time co-optimization problem from first principles that ensures the reserve allocation is sufficient to arrest frequency decline before reaching some critical frequency threshold. The reserve constraints that are introduced capture the coupling between PFR reserve, FFR reserve, and system inertia, which is assumed constant in the context of the real-time market. Numerical results illustrate this coupling via a realistically large representation of the Texas power system.

CCS CONCEPTS

• **Applied computing** → **Multi-criterion optimization and decision-making**; • **Mathematics of computing** → *Nonconvex optimization*; • **Networks** → Network economics.

KEYWORDS

Frequency Response, Economic Dispatch, Swing Equation, Electricity Markets, Real-time Co-optimization

ACM Reference Format:
Manuel Garcia and Ross Baldick. 2019. Real-Time Co-Optimization: Interdependent Reserve Types for Primary Frequency Response . In *Proceedings of the Tenth ACM International Conference on Future Energy Systems (e-Energy '19), June 25–28, 2019, Phoenix, AZ, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3307772.3335319

1 INTRODUCTION

Increasing penetration of renewable energy has raised concern for a potential future degradation of system frequency response. Wind and solar generation have two key features that raise this concern, namely they lack inertia and they exhibit random/intermittent power output [4]. In response Independent System Operators (ISOs) throughout North America have shown interest in introducing products in the electricity market that intend to improve the frequency response of the system. Some of these products aim to contribute to primary frequency response by arresting frequency decline [5]. Other existing products aim to contribute to secondary frequency response by restoring nominal frequency after the frequency has been arrested [13]. There has been recent interest in a Fast Frequency Responsive (FFR) reserve type that can be fully deployed almost instantly as a means of improving primary frequency response [11]. This new reserve type exploits the fast response of either battery storage or load shedding. FFR reserve contrasts with standard Primary Frequency Responsive (PFR) reserve that exhibits ramping limits imposed by turbine governor dynamics. This paper studies the interaction between FFR reserve, PFR reserve and inertia. A reserve constraint is derived that can be included in a real-time co-optimization problem and couples these three services. The proposed constraint ensures that the frequency will be arrested before reaching some critical frequency threshold in response to a specified maximum loss of generation.

The Electric Reliability Council of Texas (ERCOT), the ISO in Texas, is a single balancing authority that has no synchronous connections to other balancing regions. Recently having reached over 50% instantaneous wind penetration [3], ERCOT is faced with the task of maintaining adequate primary frequency response. To address this issue, ERCOT plans to introduce an FFR product intended to accommodate both battery storage and under-frequency load shedding as described in the Nodal Protocol Revision Request (NPRR) 863 [6]. ERCOT additionally plans to introduce co-optimization into the real-time market as opposed to the current real-time market design that optimizes reserve separately from generation dispatch [7, 12]. Recent work has developed a co-optimization problem that couples FFR and PFR reserve using an equivalency ratio that they denote α [10, 11]. This co-optimization problem treats 1 unit of PFR reserve equivalently to α units of FFR reserve and determines the equivalency ratio via simulation.

The distinguishing characteristic between FFR and PFR reserve is their difference in ramping ability. Although the equivalency ratio in [11] intends to capture this key characteristic, ramping ability is not directly represented in their proposed co-optimization problem. In contrast, we derive a co-optimization problem from first principles that directly accounts for the ramping ability of both reserve types. Our work extends that of [2], which introduces a constraint into the co-optimization problem that limits each generator's PFR reserve allocation based on their individual ramping ability. We extend this work by introducing FFR reserve into the co-optimization

*Corrections have been made to this paper. Figure 3 and equation (1) have been updated.
†Ph.D. candidate in the Electrical and Computer Engineering department
‡Professor in the Electrical and Computer Engineering department

e-Energy '19, June 25–28, 2019, Phoenix, AZ, USA

ACM ISBN 978-1-4503-6671-7/19/06…$15.00
https://doi.org/10.1145/3307772.3335319

problem. In this context each generator's PFR reserve allocation is limited by an increasing function of the total FFR reserve procured. We additionally specify the pricing implications and show that our formulation results in an FFR reserve price that is lower bounded by the PFR reserve price.

The remainder of the paper is organized as follows. Section 2 details our model of the system frequency response to a large generator outage, highlighting inertia, FFR reserve, and PFR reserve as the three main contributors to arresting frequency. Section 3 derives a sufficient condition for reserve procurement that guarantees that the frequency nadir lies above some critical minimum frequency threshold in response to a specified maximum generator outage. Section 4 places this sufficient condition in the context of a real-time co-optimization problem that treats inertia as being fixed. Section 5 numerically illustrates the coupling of PFR and FFR reserve using a large test case representing the Texas system.

2 MODELING THE FREQUENCY RESPONSE TO A LARGE GENERATOR OUTAGE

Many ISOs require sufficient reserve to cover a specific amount of generation loss. For example, ERCOT requires reserve that can withstand an outage of the two largest generators [11]. This section outlines a simplified model of the system response to a large loss of generation in the amount L. We begin by characterizing the contributions of inertia, PFR reserve, and FFR reserve. We then model the aggregate response of the system as a whole.

2.1 Three Contributors to Arresting Frequency

The three main contributors to arresting frequency are inertia, PFR reserve, and FFR reserve. Each of these three contributors to the frequency response will now be modeled in detail. Throughout the section we will emphasize the conservative nature of each model that results in an underestimate of the frequency trajectory that would be realized.

2.1.1 Inertia and Frequency Dynamics. The voltage frequency at time t is modeled as being the same at each generator in the system and is denoted $f(t)$. The total post-outage inertia in the system is M and represents the sum of inertia values for all generators that are still in service after the outage. The simplified system dynamics are represented by the *swing equation*, which is expressed as follows:

$$\frac{df(t)}{dt} = \frac{f_0}{2M}(\mathbf{1}^\dagger m(t) - e(t)), \tag{1}$$

where $m(t) \in \mathbb{R}^n$ represents the vector of mechanical power input from the turbine governor of each generator in the system and $e(t) \in \mathbb{R}$ represents the total electrical demand in the system. The number of generators in the system is denoted n, the vector of ones is denoted $\mathbf{1}$, and a superscript dagger $\dagger$ represents the transpose operator. This model makes the common assumption that there is no system damping. The *nominal frequency* is denoted f_0 and will be assumed to be the frequency just prior to the time of the generator outage.

2.1.2 Primary Frequency Responsive Reserve. Generators providing PFR reserve typically respond to local frequency via droop control. During droop control the reference mechanical power output of each generator's turbine governor is set to be proportional to the generator's local frequency deviation and imposes a dead-band of $\Delta_1 := f_0 - f_1$ where $f_1 < f_0$ represents the low end of the dead-band. The reference mechanical power output then traverses the turbine governor dynamics of the generator to produce realized mechanical power output. These turbine governor dynamics can be very complicated and are not detailed in our work, see [9].

Fortunately, we need only characterize each generator's turbine governor response to the very specific situation where a large loss of generation occurs. Such a response is similar to that of a step response in the reference mechanical power output because of the fast frequency drop. This type of response is illustrated in figure 1, where the amount of PFR reserve for generator i is denoted by r_i and the mechanical power output of the turbine governor for generator i at time t is denoted $m_i(t)$. The dispatched generation for generator i is denoted p_i and the mechanical power output of generator i is assumed to match its dispatched generation p_i at the time of the generator outage $t = 0$, e.g. $m_i(0) = p_i$.

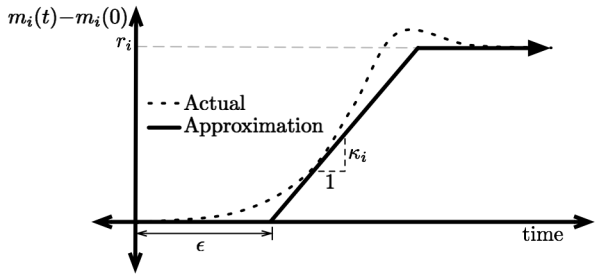

Figure 1: Turbine governor response to generator outage.

Similar to [2], we will represent the response in the mechanical power using a constant ramp rate κ_i and a constant time delay ϵ. We will assume ϵ is the same for each generator. On the other hand κ_i represents a generator specific ramping constant that is chosen in a way that underestimates the mechanical power output of the turbine governor. Since this model underestimates the mechanical power output of the turbine governor, the modeled frequency trajectory will underestimate the realized frequency trajectory.

2.1.3 Fast Frequency Responsive Reserve. Fast Frequency Responsive (FFR) reserve can be fully deployed instantaneously and represents either deployment of battery storage or load-shedding. Furthermore, the FFR is triggered when the frequency drops below a frequency threshold of $f_2 < f_1$. Note that this frequency threshold is typically significantly lower than the frequency f_1 corresponding to the dead-band of droop control. In fact, FFR reserve is considered a reserve type that is deployed only during emergencies as opposed to PFR reserve which is used regularly during droop control. Furthermore, the amount of FRR reserve for FFR unit i is denoted b_i and is non-negative. When deployed, the FFR reserve decreases the electrical demand in the system $e(t)$. We additionally introduce the non-negative constant $\Delta_2 := f_1 - f_2$.

2.2 System-Wide Frequency Response Model

This subsection provides a simple model of the frequency response exhibited by the system in response to a large generator outage. The response in system frequency $f(t)$ and power imbalance $\mathbf{1}^\dagger m(t) - e(t)$ is shown by the solid black trajectories in figure 2. These trajectories satisfy the swing equation (1) and so the frequency $f(t)$ is proportional to the integral of the energy imbalance curve.

As shown in figure 2, the PFR reserve is deployed at time t_1 when the frequency falls below the dead-band threshold of $f_1 = f_0 - \Delta_1$ but the ramp of mechanical power $\mathbf{1}^\dagger m(t)$ begins after a time delay of ϵ. The frequency is $f_0 - \Delta_1 - \delta_1$ at the time the ramp begins $t_1 + \epsilon$. The PFR reserve is modeled as ramping with constant *aggregate ramp rate* K. While the PFR reserve is ramping up the FFR reserve is deployed. The FFR is instantaneously deployed at time $t_1 + \epsilon + t_2$ when the frequency falls below the dead-band threshold of $f_2 = f_0 - \Delta_1 - \Delta_2$. Subsequently, the ramp continues until the mechanical power input of the turbine governors meet the electric power demand of the system at time t_{NAD}, at which point the frequency nadir is realized, denoted f_{NAD}.

Remark 1. Based on the model from the previous subsection the summed mechanical power output of PFR generators $\mathbf{1}^\dagger m(t)$ will not exhibit a constant aggregate ramp rate, K. In fact, some generators may deploy all PFR reserve before the frequency nadir is reached, in which case the aggregate ramp rate will effectively decrease over time. Section 3.2 will explain why using a constant aggregate ramp rate K is a conservative model of the aggregate behavior of the PFR reserve.

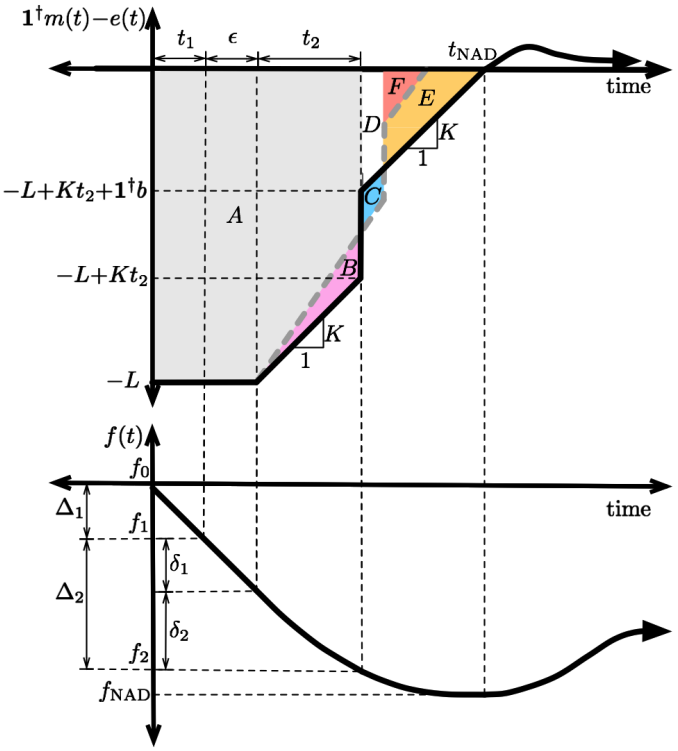

Figure 2: This plot is not drawn to scale. The top plot shows the energy imbalance over time. The main power trajectory is shown as a solid black line. The dashed gray line represents a trajectory with a slightly increased aggregate ramp rate K. Six non-overlapping regions are colored and labeled A through F. The bottom plot shows the frequency trajectory corresponding to the main power trajectory.

Notice that ϵ, Δ_1, Δ_2, f_0, f_1, and f_2 are all non-negative constant parameters. The swing equation (1) allows us to derive expressions for the other parameters by computing the integral of the power imbalance curve in figure 2. In this way it can be shown that:

$$\delta_1 = \tfrac{f_0}{2M}\epsilon L \quad \text{and} \quad \delta_2 = \Delta_2 - \tfrac{f_0}{2M}\epsilon L, \tag{2}$$

$$t_1 = \tfrac{2M\Delta_1}{Lf_0} \quad \text{and} \quad t_2 = \tfrac{1}{K}\left(L - \sqrt{L^2 - \tfrac{4M}{f_0}K\delta_2}\right). \tag{3}$$

We will additionally impose a few assumptions regarding the response of the system. Specifically, we assume the deployment of FFR reserve occurs in the middle of the PFR reserve ramp. This seems to be a good assumption because PFR reserve is typically deployed at a dead-band threshold Δ_1 that is much tighter than the FFR reserve dead-band threshold $\Delta_1 + \Delta_2$. Furthermore, we assume there is enough reserve to restore power balance and we assume that the power imbalance $\mathbf{1}^\dagger m(t) - e(t)$ remains non-positive after the FFR reserve is deployed. These assumptions are stated as follows:

Assumption 1. We assume the constant parameters are such that the FFR is deployed during the PFR ramp:

$$\delta_1 = \tfrac{\epsilon L f_0}{2M} \le \Delta_2 \quad \text{and} \quad \Delta_1 + \Delta_2 \le f_0 - f_{\text{NAD}} \tag{4}$$

We assume there is sufficient reserve to restore power balance and the power imbalance remains non-positive immediately after the FFR reserve is deployed:

$$\mathbf{1}^\dagger b + \mathbf{1}^\dagger r \ge L \quad \text{and} \quad Kt_2 + \mathbf{1}^\dagger b \le L \tag{5}$$

As in ERCOT NPRR 863 [6], the frequency thresholds are set to $f_0 = 60$Hz, $f_1 = 59.9833$Hz, and $f_2 = 59.8$Hz. These parameters will be used in all numerical results in this paper along with a PFR time delay of $\epsilon = 0.5$ seconds. Furthermore, L is set to 2750MW to represent the loss of the two largest generators in ERCOT. With these parameters, the assumption from equation (4) holds.

Remark 2. Future work should focus on situations where assumption 1 does not hold. For example, we could accommodate FFR reserve being triggered before or after the PFR ramp. We could also accommodate the situation where FFR reserve deployment immediately results in a positive energy imbalance.

3 REQUIREMENTS FOR ADEQUATE RESERVE PROCUREMENT

The main purpose of primary frequency response is to arrest the frequency decline in response to a large generator outage before the frequency drops past a critical threshold $f_{\min}$ that could cause damage to generators and/or trigger large scale shedding of firm load. This subsection derives a sufficient condition for PFR and FFR reserve to withstand a specific generator outage of size L.

3.1 Minimum Aggregate Ramp Rate

The frequency nadir increases with the aggregate ramp rate K. Equivalently, the frequency deviation $\Delta f := f_0 - f_{\text{NAD}}$ decreases with the aggregate ramp rate. To see this first notice that the frequency deviation corresponding to the main power trajectory illustrated in figure 2 can be expressed as $\Delta f := \frac{f_0}{2M}(A+B+D+E+F)$, where capital letters represent the shaded area indicated in the figure. Now consider a slight increase in aggregate ramp rate K, resulting in the energy imbalance curve illustrated by the dashed gray line. Notice that the FFR reserve is now triggered later in time because the frequency does not decline as fast. The frequency deviation can

now be expressed as $\Delta f' := \frac{f_0}{2M}(A+C+D+F)$. Since the FFR reserve is always deployed instantaneously and fully at the same frequency f_2, the area under the curve prior to the FFR deployment will remain constant for all aggregate ramp rates. As a result we must have $B = C + D$ and thus the following holds:

$$\Delta f = \tfrac{f_0}{2M}(A + C + 2D + E + F) > \tfrac{f_0}{2M}(A + C + D + F) = \Delta f' \quad (6)$$

This analysis implies that the frequency deviation Δf is strictly monotonically decreasing in the aggregate ramp rate K. Equivalently, the frequency nadir f_{NAD} is strictly monotonically increasing with the aggregate ramp rate K. As a result there exists a unique ramp rate $K_{\min}$ such that the frequency nadir is $f_{\text{NAD}} = f_{\min}$. Furthermore, all aggregate ramp rates greater than this *minimum aggregate ramp rate* $K_{\min}$ will satisfy the minimum frequency threshold, e.g. $f_{\text{NAD}} \geq f_{\min}$. The following result provides an expression for $K_{\min}$ where the sum of all FFR reserve is denoted $\tilde{b} := \mathbf{1}^\dagger b$ and a constant is introduced as $\Delta_3 := f_0 - \Delta_1 - \Delta_2 - f_{\min}$.

Proposition 1. *Under assumption 1, the frequency nadir satisfies the minimum frequency threshold $f_{NAD} \geq f_{min}$ if the aggregate ramp rate satisfies $K \geq K_{min}$ where:*

$$K_{min} = \frac{\left(\tilde{b}\sqrt{\Delta_3} - \sqrt{(\Delta_2 + \Delta_3 - \frac{f_0}{2M}\epsilon L)L^2 - (\Delta_2 - \frac{f_0}{2M}\epsilon L)\tilde{b}^2}\right)^2}{\frac{4M}{f_0}(\Delta_2 + \Delta_3 - \frac{f_0}{2M}\epsilon L)^2} \quad (7)$$

Proof: The nadir frequency can be related to the aggregate ramp rate using the swing equation (1). From the integral of the energy imbalance curve in figure 2, this relationship is as follows:

$$\tfrac{2M}{f_0}(f_0 - f_{\text{NAD}}) = L(t_1+\epsilon) + (L - Kt_2)t_2 + \tfrac{1}{2}t_2^2 K + \tfrac{1}{2K}(L - Kt_2 - \tilde{b})^2 \quad (8)$$

$$= \tfrac{2M}{f_0}\Delta_1 + L\epsilon + \tfrac{1}{2K}(L - \tilde{b})^2 + \tilde{b}t_2 \quad (9)$$

$$= \tfrac{2M}{f_0}\Delta_1 + L\epsilon + \tfrac{1}{2K}(L - \tilde{b})^2 + \tilde{b}\tfrac{1}{K}\left(L - \sqrt{L^2 - \tfrac{4M}{f_0}K\delta_2}\right) \quad (10)$$

Step (8)-(9) uses $t_1 = \frac{2M\Delta_1}{Lf_0}$ and performs additional algebra. Step (9)-(10) expresses t_2 as $t_2 = \frac{1}{K}L - \frac{1}{K}\sqrt{L^2 - \frac{4M}{f_0}K\delta_2}$. We will now set $f_{\text{NAD}} = f_{\min}$ and solve (10) for $K_{\min}$. Note that there is only one such $K_{\min}$ and any $K > K_{\min}$ will result in a nadir frequency $f_{\text{NAD}} > f_{\min}$ because f_{NAD} is strictly monotonically increasing in K.

First, let's introduce the constant $\omega = f_0 - f_{\min} - \Delta_1 - \frac{f_0}{2M}\epsilon L$ to simplify notation. From equation (10) we have the following:

$$\tilde{b}\sqrt{L^2 - \tfrac{4M}{f_0}K\delta_2} = \tfrac{1}{2}L^2 + \tfrac{1}{2}\tilde{b}^2 - \tfrac{2M}{f_0}K\omega \quad (11)$$

Note that both sides of this equation are real and non-negative, else t_2 from equation (3) would be complex. Squaring both sides and rearranging gives the following quadratic equation in K:

$$\tfrac{4M^2}{f_0^2}\omega^2 K^2 + (\tfrac{4M}{f_0}\tilde{b}^2\delta_2 - (L^2 + \tilde{b}^2)\tfrac{2M}{f_0}\omega)K + \tfrac{1}{4}(L^2 + \tilde{b}^2)^2 - \tilde{b}^2 L^2 = 0 \quad (12)$$

This equation has two solutions, $K_+^\star$ and $K_-^\star$, written as follows:

$$\frac{\frac{2M}{f_0}\omega(L^2+\tilde{b}^2) - \frac{4M}{f_0}\tilde{b}^2\delta_2 \pm \sqrt{(\frac{4M}{f_0}\tilde{b}^2\delta_2 - \frac{2M}{f_0}\omega(L^2+\tilde{b}^2))^2 - \frac{4M^2}{f_0^2}\omega^2((L^2+\tilde{b}^2)^2 - 4\tilde{b}^2L^2)}}{\frac{8M^2}{f_0^2}\omega^2}$$

Algebraically rearranging the descriminant results in the following equivalent expression for $K_+^\star$ and $K_-^\star$:

$$K_\pm^\star = \frac{\omega(L^2+\tilde{b}^2) - 2\tilde{b}^2\delta_2 \pm 2\tilde{b}\sqrt{\omega - \delta_2}\sqrt{\omega L^2 - \delta_2\tilde{b}^2}}{\frac{4M}{f_0}\omega^2} \quad (13)$$

Under the assumption that $K_{\min} = K_-^\star$ where $\pm$ takes the sign $-$, the result follows by factoring the numerator of the previous expression and replacing $\omega = \Delta_2 + \Delta_3 - \frac{f_0}{2M}\epsilon L$ and $\delta_2 = \Delta_2 - \frac{f_0}{2M}\epsilon L$.

Let's now show that $\pm$ cannot take the sign $+$. Suppose it can take the sign $+$. Then $K_+^\star$ solves equation (11) and thus results in a real non-negative LHS and RHS of (11). Furthermore, $K_-^\star \leq K_+^\star$ and so the LHS and RHS of (11) remain real and non-negative when evaluated at $K_-^\star$. Since $K_-^\star$ solves (12) and results in a real non-negative LHS and RHS of (11), $K_-^\star$ must also solve (11). Thus $K_-^\star$ and $K_+^\star$ result in the same nadir frequency. This contradicts the fact that the nadir frequency is strictly monotonically increasing in the aggregate ramp rate K. □

3.2 Sufficient Condition for Satisfying Frequency Threshold

As stated in remark 1, the system model of a constant aggregate ramp rate does not fully capture the response of each individual generator. In fact, each individual generator will exhibit a ramp rate of κ_i until their reserve has been fully deployed as explained in section 2.1.2. Some generators may deploy all PFR reserve before the frequency nadir is reached, in which case the aggregate ramp rate will effectively decrease over time. That being said, the frequency threshold is guaranteed to be satisfied if we assume the generators are capable of fully deploying all of their PFR reserve before the time $t_{\min}$, which represents the time of the frequency nadir at the minimum aggregate ramp rate $K_{\min}$. This is implied by the intuitive fact that the frequency nadir rises when a generator ramps faster than expected. This intuitive fact was illustrated in section 3.1 but is not formally proven. This logic leads to the following result:

Proposition 2. *Under assumption 1, the frequency nadir satisfies the minimum frequency threshold $f_{NAD} \geq f_{min}$ if the following holds:*

$$r_i \leq \kappa_i h(M, \mathbf{1}^\dagger b) \;\; \forall i \in [1, \ldots, n] \quad (14)$$

where the limit function $h(M, \tilde{b})$ is as follows:

$$h(M, \tilde{b}) := \frac{\frac{4M}{f_0}(\Delta_2 + \Delta_3 - \frac{f_0}{2M}\epsilon L)^2(L - \tilde{b})}{\left(\tilde{b}\sqrt{\Delta_3} - \sqrt{(\Delta_2 + \Delta_3 - \frac{f_0}{2M}\epsilon L)L^2 - (\Delta_2 - \frac{f_0}{2M}\epsilon L)\tilde{b}^2}\right)^2} \quad (15)$$

Sketch of Proof: Following from the discussion, the frequency threshold will be met if each generator is capable of deploying all of its PFR reserve before the time $t_{\min}$. This requirement is mathematically written as follows:

$$r_i \leq \kappa_i(t_{\min} - t_1 - \epsilon) = \kappa_i \tfrac{L - \mathbf{1}^\dagger b}{K_{\min}} \quad (16)$$

Substituting $K_{\min}$ with its expression from (7) gives the result. □

The function $h(M, \tilde{b})$ is convex in its second argument, which makes constraint (14) a non-convex constraint. Figure 3 provides example plots of $h(M, \tilde{b})$ as it varies in its second argument $\tilde{b}$ for different values of inertia M. Notice that this function is increasing in M and $\tilde{b}$. As a result constraint (14) allows more PFR reserve to be allocated to a generator if the inertia is higher, if the total FFR reserve allocation is higher, or if it has a larger ramp rate κ_i.

4 REAL-TIME CO-OPTIMIZATION

The real-time market in ERCOT will soon introduce FFR reserve and retain the already existing PFR reserve. Furthermore, ERCOT

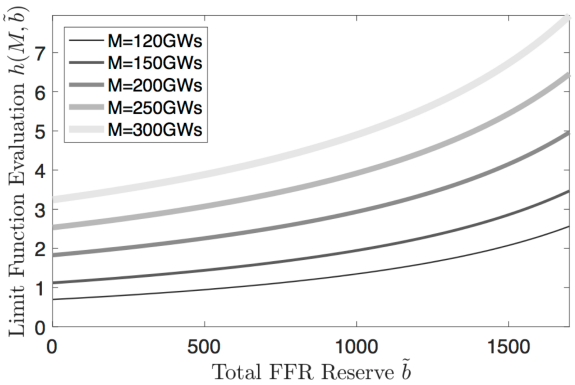

Figure 3: Function $h(M,\tilde{b})$ with parameters from section 5.

is changing its current practice of optimizing the reserve allocation separately from dispatch decisions. Specifically, ERCOT will soon implement real-time co-optimization by incorporating PFR and FFR reserve into the economic dispatch problem. This section proposes a real-time co-optimization problem that incorporates the sufficient condition for reserve allocation from the previous section.

4.1 Problem Formulation

The the real-time co-optimization problem maximizes social welfare, or equivalently minimizes the total system cost, subject to all system constraints. Without loss of generality we assume there is one generator and one FFR unit located at each bus in the transmission system. The non-negative decision variables represent the vector of nodal dispatched generation $p \in \mathbb{R}^n_+$, nodal PFR reserve $r \in \mathbb{R}^n_+$ and nodal FFR reserve $b \in \mathbb{R}^n_+$, where n is the number of nodes (or buses) in the system. The vector of fixed nodal demand is denoted $d \in \mathbb{R}^n$. The convex generator cost function is denoted $c(p)$. Convex cost functions pertaining to PFR and FFR reserve are denoted $c_1(r)$ and $c_2(b)$ respectively. The real-time co-optimization problem is written as follows:

$$\min_{b\in\mathbb{R}^n_+, p\in\mathbb{R}^n_+, r\in\mathbb{R}^n_+} \quad c(p)+c_1(r)+c_2(b) \tag{17}$$

$$st: \mathbf{1}^\dagger(p-d) = 0 \tag{17a}$$

$$H(p-d) \le \bar{T} \tag{17b}$$

$$L \le 1^\dagger r + 1^\dagger b \tag{17c}$$

$$r_i \le \kappa_i h(M, 1^\dagger b) \;\; \forall i \in [1,\dots,n] \tag{17d}$$

$$p + r \le \bar{p} \tag{17e}$$

$$b \le \bar{b} \tag{17f}$$

$$r \le \bar{r} \tag{17g}$$

Constraint (17a) represents the overall system power balance constraint. Constraint (17b) represents the line limits, where H is the matrix of shift factors. The sufficient condition from the previous section is incorporated into constraints (17c) and (17d). Constraint (17e) ensures that generators providing PFR have sufficient headroom to produce the reserve they are allocated. Constraints (17f) and (17g) enforce the offered limits of FFR reserve and PFR reserve. Note that this formulation assumes $Kt_2 + \mathbf{1}^\dagger b \le L$ as in assumption 1. This additional linear constraint should be enforced if the FFR reserve offers $\bar{b}$ are significantly large.

4.2 Economic Implications

Many electricity markets today set reserve prices based on the Lagrange multiplier of a reserve requirement constraint that is analogous to constraint (17c), which we will denote λ. Constraint (17d) allows the co-optimization problem to capture marginal costs associated with the ramping abilities of FFR reserve and PFR reserve. According to this formulation, the PFR reserve price for generator i is $\lambda - \gamma_i$ and the FFR reserve price for all FFR units is $\lambda + \gamma^\dagger \kappa \nabla h(M, \mathbf{1}^\dagger b^\star)$ where $\gamma \ge \mathbf{0}$ is the vector of Lagrange multipliers for constraint (17d). These reserve prices ensure that PFR generators and FFR units are dispatched at values that maximize their profit under the price taking assumption.

4.3 Accommodating non-convexity

There are many potential methods of accommodating the non-convexity presented by constraint (17d). First, iterative algorithms can be used to solve the non-convex economic dispatch problem to a local minimum. This is the method used in the numerical results of this paper. Second, the function $h(M,\cdot)$ can be approximated using Taylor expansions around some base-case FFR allocation. This method would result in a linear program under the assumption that cost functions are linear. Third, the function $h(M,\cdot)$ can be conservatively approximated by a piecewise linear function. The resulting problem can be formulated as a Mixed Integer Program (MIP) under the assumption that the cost is linear. Furthermore, a good approximation of $h(M,\cdot)$ only requires a few integers to be introduced, resulting in a MIP that is very easy to solve. Finally, the co-optimization problem can be made convex under the assumption that the cost function for FFR reserve is zero, e.g. $c_2(b) = 0$. This convex reformulation introduces a decision variable $\hat{b} := h(M, \mathbf{1}^\dagger b)$. Future work will investigate each of these options.

5 NUMERICAL RESULTS

This section studies the coupling between FFR and PFR reserve in the co-optimization problem (17). A realistically large 2000 bus test case is used that represents the electric power system in Texas described in [1] and [14]. We intend to illustrate the effect of increasing FFR reserve offers $\mathbf{1}^\dagger \bar{b}$ where the FFR reserve cost is assumed to be zero $c_1(b) = 0$. Since the FFR reserve cost is zero, the FFR reserve is always fully dispatched. The interior point algorithm from the MATLAB package TOMLAB [8] is used to solve problem (17) to a local minimum. This algorithm requires approximately 25 seconds to converge.

The 50 natural gas generators with the largest capacity are selected to provide PFR reserve and their maximum PFR reserve offer is set to 20 percent of their capacity, e.g. $\bar{r}_i = .2\bar{p}_i$. The ramp rate $\kappa_i = 20$MW/s and the delay $\epsilon = 0.5$s are the same for each PFR reserve generator i and were determined using a dynamic simulation of a loss of the two largest generators amounting to about 2750MW.

The frequency thresholds are set to values that match the ERCOT NPRR 863 [6]. Specifically, the PFR threshold is $f_1 = 59.9833$Hz, the FFR threshold is $f_2 = 59.8$Hz, and the minimum frequency threshold is $f_{\min} = 59.4$Hz, which represents the threshold at which firm load is shed. We additionally analyze a loss of generation in the amount of $L = 2750$MW, which represents the two largest nuclear plants in Texas. A post-outage inertia of $M = 300$GWs is assumed and is represented in the plot of the limit function in figure 3.

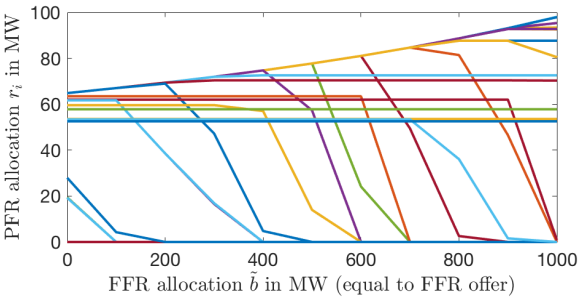

Figure 4: PFR allocation for increasing FFR offers.

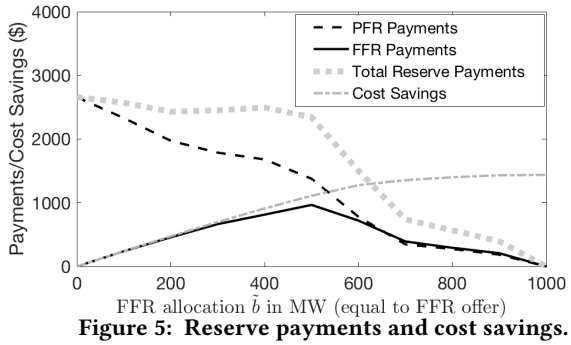

Figure 5: Reserve payments and cost savings.

Figure 4 plots the PFR allocation to each PFR reserve generator as the total FFR offer $\mathbf{1}^\dagger \bar{b}$ increases from 0 to 1000MW. When the FFR offer increases past 1000MW the price of reserve becomes zero and the co-optimization problem has multiple solutions, making it difficult to analyze. Notice that assumption 1 is satisfied with the selected parameters and for the selected range of FFR offers. As the total FFR reserve offer increases it replaces the most expensive PFR reserve. As a result the PFR reserve allocation does not decrease uniformly among all generators, but instead sparsely decreases to zero for only one or two generators at a time.

The upper bound on the PFR reserve allocation from constraint (17d) increases as the total FFR reserve offer $\mathbf{1}^\dagger \bar{b}$ increases. Inexpensive generators that are operating at this upper bound are allocated more PFR reserve as the FFR reserve offer increases. As a result we see PFR reserve allocation increasing for many generators from approximately 65MW to approximately 100MW. On the other hand, inexpensive generators that are limited by their PFR offer $\bar{r}_i$ do not experience this increase in PFR allocation.

Figure 5 shows the trajectory of reserve payments as the FFR reserve offer $\mathbf{1}^\dagger \bar{b}$ increases. Notice that the payments to FFR reserve initially increase because the FFR reserve allocation is increasing. However, when the FFR reserve offer $\mathbf{1}^\dagger \bar{b}$ reaches 500 MW the payments to FFR reserve begin to decrease because the reserve price is dropping. On the other hand, PFR reserve experiences a steady decline of payments. The total reserve payments (including both FFR and PFR reserve payments) also decreases as more FFR reserve is introduced.

Increasing FFR reserve also has the benefit of reducing total system costs, or equivalently increasing social welfare. Figure 5 additionally plots the cost savings, which are increasing. Notice that FFR increases cost savings in two ways. First, it replaces expensive PFR reserve. Second, it allows more inexpensive PFR reserve to be allocated by increasing the PFR reserve limit from constraint (17d).

6 CONCLUSIONS

This paper derives a real-time co-optimization problem from first principles that incorporates two reserve types that contribute to primary frequency response. Specifically, a new FFR reserve type is considered in addition to the standard PFR reserve type. This FFR reserve type has nearly instant ramping capabilities, as is the case for battery storage and load shedding. The constraints of the co-optimization problem couple the FFR and PFR reserve allocation and include a sufficient condition that guarantees the reserve allocation is capable of arresting the frequency before it reaches some critical frequency threshold in response to a specified maximum generator outage. Numerical results illustrate the effect of increasing FFR reserve using a 2000 bus test case that represents the Texas power system. We show that increasing FFR reserve offers decreases system costs not only by replacing expensive PFR reserve but also by allowing inexpensive PFR reserve to be dispatched higher. Increasing FFR reserve offers additionally decreases total reserve payments.

ACKNOWLEDGMENTS

The authors were supported, in part, by the National Science Foundation under grant ECCS-1406894.

The authors would like to thank Dr. Shams Siddiqi for comments on this work.

REFERENCES

[1] Adam B Birchfield, Ti Xu, Kathleen M Gegner, Komal S Shetye, and Thomas J Overbye. 2017. Grid structural characteristics as validation criteria for synthetic networks. *IEEE Transactions on power systems* 32, 4 (2017), 3258–3265.

[2] Héctor Chávez, Ross Baldick, and Sandip Sharma. 2014. Governor rate-constrained OPF for primary frequency control adequacy. *IEEE Transactions on Power Systems* 29, 3 (2014), 1473–1480.

[3] Wesley Cole and A Will Frazier. 2018. Impacts of increasing penetration of renewable energy on the operation of the power sector. *The Electricity Journal* 31, 10 (2018), 24–31.

[4] Pengwei Du and Julia Matevosyan. 2018. Forecast system inertia condition and its impact to integrate more renewables. *IEEE Transactions on Smart Grid* 9, 2 (2018), 1531–1533.

[5] Erik Ela, Aidan Tuohy, Michael Milligan, Brendan Kirby, and Daniel Brooks. 2012. Alternative approaches for incentivizing the frequency responsive reserve ancillary service. *The Electricity Journal* 25, 4 (2012), 88–102.

[6] ERCOT. 2018. *NPRR 863: Creation of Primary Frequency Response Service Product and Revisions to Responsive Reserve.* Technical Report. ERCOT. http://www.ercot.com/mktrules/issues/reports/nprr

[7] ERCOT. 2018. *Study of the Operational Improvements and Other Benefits Associated with the Implementation of Real-Time Co-optimization of Energy and Ancillary Services.* Technical Report. ERCOT. 10 pages. http://www.ercot.com/content/wcm/lists/144930/Study_of_the_Benefits_of_Real-Time_Co-optimization_FINAL.pdf

[8] Kenneth Holmström, Anders O Göran, and Marcus M Edvall. 2010. Users Guide for TOMLAB 7. *Tomlab Optimization Inc* (2010).

[9] Prabha Kundur, Neal J Balu, and Mark G Lauby. 1994. *Power system stability and control.* Vol. 7. McGraw-hill New York.

[10] Weifeng Li, Pengwei Du, and Ning Lu. 2018. Design of a new primary frequency control market for hosting frequency response reserve offers from both generators and loads. *IEEE Transactions on Smart Grid* 9, 5 (2018), 4883–4892.

[11] Cong Liu and Pengwei Du. 2018. Participation of load resources in day-ahead market to provide primary-frequency response reserve. *IEEE Transactions on Power Systems* 33, 5 (2018), 5041–5051.

[12] Stephen Reedy. 2018. *Simulation of Real-Time Co-Optimization of Energy and Ancillary Services for Operating Year 2017.* Technical Report. Potomac Economics. 8 pages. http://www.ercot.com/content/wcm/lists/144930/IMM_Simulation_of_Real-Time_Co-optimization_for_2017.pdf

[13] Bolun Xu, Yury Dvorkin, Daniel S Kirschen, Cesar A Silva-Monroy, and Jean-Paul Watson. 2016. A comparison of policies on the participation of storage in US frequency regulation markets. In *2016 IEEE Power and Energy Society General Meeting (PESGM).* IEEE, 1–5.

[14] Ti Xu, Adam B Birchfield, and Thomas J Overbye. 2018. Modeling, tuning, and validating system dynamics in synthetic electric grids. *IEEE Transactions on Power Systems* 33, 6 (2018), 6501–6509.