Scheduling Inefficient Storage

Rozanna N. Jesudasan, Lachlan L. H. Andrew, Hai L. Vu

***Abstract*—Energy storage is becoming increasingly important, both to mitigate intermittency in renewable generation and to reduce peak demand. However, storage remains expensive and so must be managed optimally. This paper considers the optimal management of storage that is subject to inefficiency in charging/discharging and to self-discharge, with the objective of minimizing energy costs. Notably, it shows that the less efficient the storage is, the less capacity is required to achieve the maximum peak-shaving benefit.**

I. Introduction

As increasing demands are placed on an aging grid, there is an increased need to incorporate large scale energy storage to smooth the peak loads placed on the generation, transmission and distribution systems.

To minimize demands on the network, the storage should be placed near the most variable elements, be they wind turbines, photovoltaic systems or loads. We will consider a storage device placed near the load, and managed by the customer. It is unclear who should pay for such storage; the primary beneficiary of reduced transmission peaks is not the customer but the load serving entity (LSE). Thus, we assume that the storage is purchased by the LSE, but managed by customer to minimize the customer's electricity costs. This type of arrangement is being explored by Australian LSE SP Ausnet in response to serious fires on Black Saturday in 2010 caused by overloaded distribution circuits. It is facilitated by the roll-out of smart meters in Australia.

This paper investigates the structural properties of the optimal charging schedule for such storage. Prior work [1], [2] typically assumes that batteries are lossless. In contrast, this paper will consider the effects of two types of losses. The first is inefficiency in charging, such as electrolysis of the solvent in a chemical battery or pumping inefficiency in a pumped hydro storage system. The second is self-discharge, such as friction losses in a flywheel. These two types of losses have different effects, but both significantly reduce the ability of the storage to reduce the peak load.

The optimal action at any point in time depends on the load and price at future instants. Look-ahead is assumed by many algorithms such as [3] for distributed computation in demand-response settings, in which different customers predict their daily demands and the LSE calculates corresponding prices. The computer science on-line algorithms community has developed techniques that achieve almost optimal results for models similar to this [2], [4]. These algorithms motivate us to investigate the structure of the optimal solution, since they do not themselves give insight into that structure. A contribution of this paper is to demonstrate that, in many practical situations, only a limited amount of future knowledge is required to make optimal charging decisions. This is due to the presence of quasi-renewal points beyond which future conditions do not affect the current optimal action.

After introducing the model in Section II, we investigate the structure of the optimal solution for ideal batteries in Section III. We then consider charging inefficiency in Section IV and self-discharge in Section V.

II. Model

The current electricity grid relies on the balance between supply and demand. This means that without energy storage, the draw from the grid, $g(t)$, should be able to satisfy the demand (load) $D(t) \geq 0$ at all time t, i.e., $g(t) = D(t)$. We consider demand to be inelastic (independent of changes in price). Now consider a situation where an energy storage system with maximum capacity $B > 0$ is installed between the generator and the load. This would enable the user to either consume energy straight from the grid, or to store some energy $b(t)$, by drawing power at a rate of $c(t)$ during low energy prices, and consume the stored energy at a rate of $d(t)$ when the energy price is high. The storage device has a maximum rate at which it can charge and discharge given by $C_{\max}, D_{\max}$ respectively. Further the storage has a charging efficiency of $\eta \leq 1$ and also loses energy over time such that, if no charging occurs, then $b(t) = Sb(t-1)$ for some $S \in (0, 1]$.

The cost of drawing power $g(t)$ from the grid is assumed to have the form $P(t)N(g(t))$ for some positive price function $P(\cdot)$ and some nonlinearity $N(\cdot)$, which is strictly convex increasing, which models the fact that peak grid power increases the cost for the utility and increases the strain on the grid.

This gives rise to an objective that would require the utility to minimise its generation cost, assuming that the cost savings by the utility will be passed on to the user by reducing the electricity bill of the users who shift their demands on the grid. In particular, we would like to know the optimal charging and grid use schedule under arbitrary prices and arbitrary demands. That is, the user seeks to minimize the energy cost over a horizon T, subject to the final storage level being F:

$$\arg\min_{g,b,c,d} \sum_{t=1}^{T} P(t)N(g(t)) \tag{1}$$

subject to,

$$
\begin{aligned}
b(t) - Sb(t-1) - \eta c(t) + d(t) = 0 & \qquad [\phi(t)] \quad (2a)\\
D(t) + c(t) - d(t) - g(t) = 0 & \qquad [\theta(t)] \quad (2b)\\
g(t) \geq 0 & \qquad [\lambda(t)] \quad (2c)\\
B - b(t) \geq 0 \qquad b(t) \geq 0 & \qquad [\bar{\beta}(t), \underline{\beta}(t)] \quad (2d)\\
C_{\max} - c(t) \geq 0 \qquad c(t) \geq 0 & \qquad [\bar{\chi}(t), \underline{\chi}(t)] \quad (2e)\\
D_{\max} - d(t) \geq 0 \qquad d(t) \geq 0 & \qquad [\bar{\delta}(t), \underline{\delta}(t)] \quad (2f)\\
b(T) - F = 0 & \qquad [\xi] \quad (2g)
\end{aligned}
$$

with $b(0) = 0$, where the variables in square brackets are the Lagrange dual variables corresponding to each constraint. Note that capital letters denote parameters of the problem instance, lower case letters denote decision variables and Greek letters (except η) denote Lagrange multipliers.

If $\eta = 1$ the solutions for c and d need not be unique. We consider only the solution in which $\min(c(t), d(t)) = 0$ for all t, which corresponds to the storage never charging and discharging simultaneously.

When dimensioning a storage facility, it is useful to consider long time horizons, and so it is useful for the solution to be well defined in the limit of large T. Let $(g_F^T, b_F^T, c_F^T, d_F^T)$ be a solution to (1)–(2) for a given F. Under mild conditions, the limit for large T exists.

Theorem 1. *Consider a system and a time t for which one of the following conditions holds:*

1) *there exists a $T > t$ such that $b_B^T(t) = 0$;*
2) *there exists a $T > t$ such that $b_0^T(t) = B$.*

Then $g^*(t) = \lim_{T\to\infty} g_F^T(t)$ and $b^*(t) = \lim_{T\to\infty} b_F^T(t)$ exist.

Indeed, the first main structural result, Theorem 3, will show that exact convergence occurs for finite T, which allows us to focus on the finite horizon formulation.

III. Structure of Optimal Solution

First, let us explore structural properties for the optimal solution, that will later allow us to study the behaviour of the optimal charging and generation schedules.

We will start with some preliminary results. As may be expected, requiring the storage to end at a higher state of charge never reduces the storage level at an earlier time. The following lemma is proved in the appendix.

Lemma 2. *For any $F_1, F_2 \in [0, B]$ with $F_1 < F_2$, we have that $b_{F1}^T(t) \le b_{F2}^T(t)$ for all $t \in [0, T]$.*

The first main structural result, proven in the appendix, shows that the optimal solution can often be found with knowledge of conditions only a finite distance into the future.

Theorem 3. *If $b_B^T(t) = 0$ or $b_0^T(t) = B$, for some $t \in [0, T]$, then for all $T' \ge T$, all $t' \le t$ and all $k \in [0, B]$, $b_k^{T'}(t') = b_0^{T'}(t') = b_B^{T'}(t') = b^*(t')$, and consequently $g_k^{T'}(t') = g_0^{T'}(t') = g_B^{T'}(t') = g^*(t')$.*

We call t a *quasi-renewal* point. In a stochastic process, a renewal point is a point $\hat{t}$ such that the process after $\hat{t}$ is independent of the process before $\hat{t}$ given the value at time $\hat{t}$. Similarly, any future demand or prices beyond T will not affect the optimal schedule prior to the quasi-renewal point t. Diurnal patterns mean that, when the storage capacity is small, T is often less than a day ahead of t. This means that demand-response schemes such as [3] do not need predictions substantially more than a day ahead. However, if the storage is large, then $T - t$ can be multiple days, as illustrated in Fig. 1. This example, and others in the paper, uses total demand from the Australian state of New South Wales (NSW) in March 2012 [5]. This is a highly aggregated load, with regular diurnal variation. We have also produced the results in this paper for the much burstier demand of a single house, and found that our qualitative conclusions still hold, but space does not allow the figures to be included.

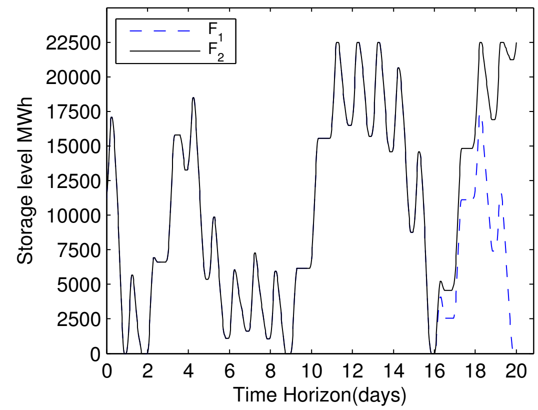

Figure 1. NSW Renewal point for 20 days, with a 22500MWh, 90% efficient storage facility.

Provided that the charging schedule results in renewal points, let us first consider the case of ideal storage ($\eta = S = 1$, $C_{\max} = D_{\max} = \infty$), to obtain insight into the structure of the solution. A consequence of these conditions is the following

Proposition 4. *Consider an ideal storage facility and any interval $[t_1, t_2]$ in which the storage is partially filled ($b^*(t) \in (0, B)$). Then the optimal incremental grid power cost $P(t)N'(g^*(t))$ is constant on $[t_1, t_2]$.*

Moreover, $P(\cdot)N'(g^*(\cdot))$ increases from time t to $t+1$ only if the storage is full $b^*(t) = B$, and decreases from t to $t+1$ only if the storage is empty $b^*(t) = 0$.

This is a consequence of the more general result Theorem 5 in Section IV. Note that this is in contrast to the structure observed in [1], in which the storage level increased monotonically, then remained constant, and then decreased monotonically over the horizon. That is because the model of [1] imposed a penalty for the storage being less than completely full. This penalty accumulates over time, and so there is a greater incentive to charge the storage early, and less incentive nearer to the horizon.

IV. Impact of Efficiency

We now investigate the effect that inefficiency has on the optimal charging schedule. At first, it may seem intuitive that the loss of energy would simply require that the rate of charging of the storage be greater when the system is inefficient, but that the structure would be otherwise unchanged. It turns out that inefficiency causes a substantially greater effect, as explained in the following theorem.

Theorem 5. *Consider an interval $[t_1, t_2]$ on which $b^*(t) \in (0, B)$. There exists a constant*

$$R = \xi^T + \sum_{\tau=t}^{T-1} S^\tau \left(\underline{\beta}^T(\tau) - \bar{\beta}^T(\tau) \right) \tag{3}$$

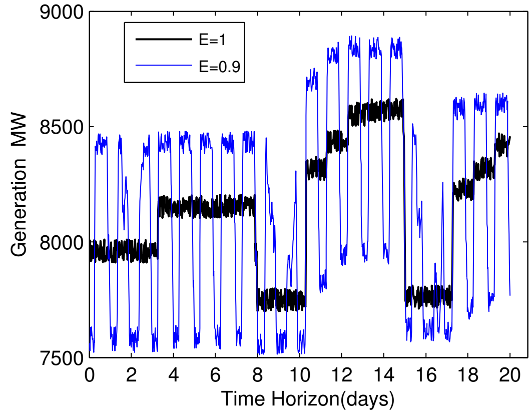

Figure 2. Generation schedules for 22500MWh storage devices with 100% and 90% efficiency.

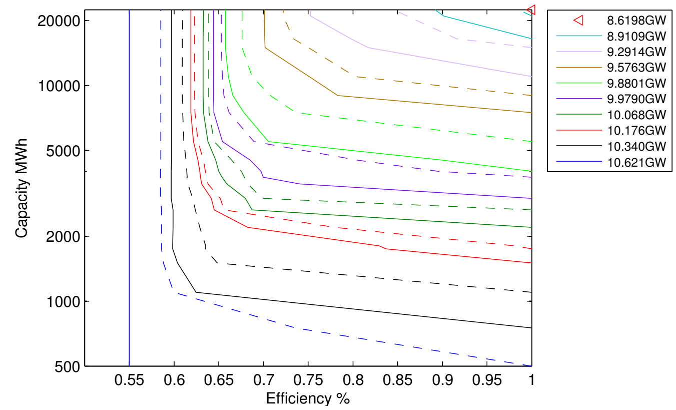

Figure 3. Capacity vs Efficiency for the March 2012 NSW demand.

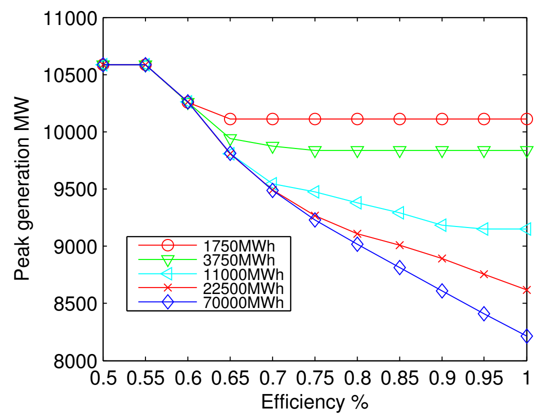

Figure 4. NSW Maximum generation vs efficiency.

such that, when the storage is neither charging at maximum rate nor discharging at maximum rate (i.e., $d < D_{\max}$ and $c < C_{\max}$) then

$$P(t)N'(g^T(t))S^t \in [\eta R^+, R^+] \quad (4)$$

where $R^+ = \max(R, 0)$. The left hand side is at the lower limit of this interval when the storage is charging, $c(t) > 0$, and the upper limit when the storage is discharging, $d(t) > 0$.

Moreover, when the storage is discharging at maximum rate, $P(t)N'(g^*(t))S^t \geq R$ and when it is charging at maximum rate, $P(t)N'(g^*(t))S^t \leq \eta R^+$. When $P(t)N'(g^*(t))S^t$ is in the interior of the interval, the storage is neither charging nor discharging.

This theorem demonstrates that the marginal generation cost is no longer constant on intervals in which the storage is partially charged. If the charging and discharging rates are low, the fluctuations are by a factor of η; if the maximum charge or discharge rate is reached, the fluctuations are even higher. Note also that the storage only charges when the marginal cost of grid power is low compared to other times in $[t_1, t_2]$ and only discharges when the cost is high.

This structure is illustrated in Figure 2, which shows the fluctuations in the generation required to meet the NSW demand with a storage of 22500MWh that is either 90% or 100% efficient,with the price function $P(.) = 1$.

This fluctuation reduces the ability of the storage device to shave the demand peaks. As a result, the less efficient a storage facility is, the larger it must be to shave the peak demand to a given level. This is illustrated in Fig. 3. For a fixed budget, the optimal design for peak shaving will need to balance the choice of an expensive storage technology with high efficiency versus a cheaper technology for which a higher capacity can be deployed.

However, this natural intuition (that lower efficiency storage should be larger) only applies for small capacities. Remarkably, the reverse is true for high capacity storage. This is illustrated in Figure 4, which shows the peak generation required as a function of the efficiency for several storage capacities. For high efficiencies, the larger storage capacities provide more smoothing and so reduce the peak demand. The diagonal slope on the left of the graph indicates that, at lower efficiency, having storage larger than a threshold does not provide further shaving. Moreover, that threshold of storage capacity is *smaller* as the efficiency drops.

To understand why, consider fully efficient storage smoothing a given demand. Once the capacity is large enough to supply all the peaks without fully discharging, increasing the capacity does not provide further smoothing. For inefficient storage, the charging “dead zone” identified above means that it is not optimal to charge or discharge unless the difference in marginal costs differs sufficiently, resulting in a smaller range between the maximum and minimum state of charge, even without capacity constraints. Thus the size of storage required to accommodate the maximum range of states of charge decreases as the efficiency decreases.

Another striking feature of Fig 4 is that, for small capacities, increasing the efficiency beyond a certain point does not further reduce the peaks. The explanation is the counterpart to that of the above. The peak generation is determined by the single largest peak. To obtain optimal smoothing, all that is required is that the storage be fully charged before that peak and fully discharged after it[1]. Even an inefficient storage facility will fully charge and fully discharge provided that the

[1] Note that our definition of capacity is the amount of energy that the storage can deliver, not the amount required to charge it fully

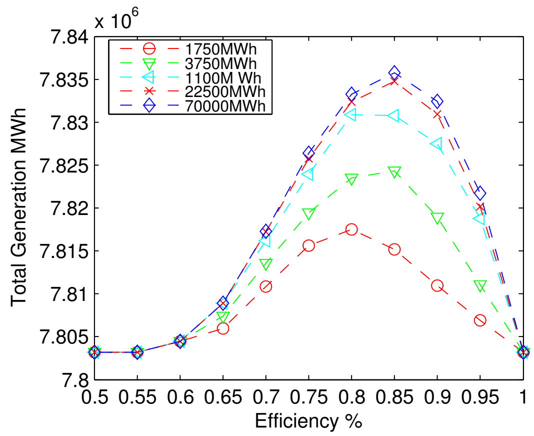

Figure 5. NSW total generation vs efficiency for a 20 day period.

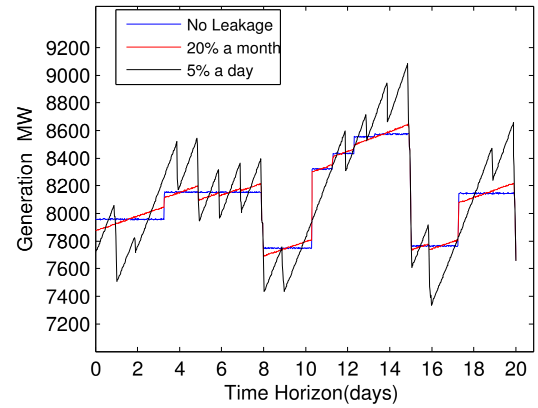

Figure 6. Optimal generation with leakage.

difference in prices is sufficient.

This may suggest that the transmission and distribution providers are indifferent between different efficiencies, provided the capital costs are the same. However, if the storage is distributed to individual customers, the *retailer* sometimes has a perverse incentive to install *lower* efficiency storage. That is because the customer must pay for the energy lost due to lower efficiency. However, this effect does not provide an incentive to reduce the efficiency to zero, since the customer's optimal policy will then be to ignore the storage completely. Figure 5 shows the total energy generated as a function of the storage efficiency for the NSW load.

V. Impact of self-discharge

Let us now investigate the effect of self-discharge on the optimal charging schedule. This is important for some storage technologies such as flywheels, which lose up to 50% of their stored energy in 24 hours [6]. As can be seen from (4), if the fraction of energy retained from one time step to the next is not $S = 1$, then we no longer have the qualitative conclusion that the marginal cost is constant on intervals during which the storage is either being charged or discharged. Instead, the charging rate will be such that the incremental cost increases exponentially. This is because the earlier energy is stored, the more of it will have leaked away by the time it is needed. Note that this may be an artefact of decision to model leakage as a constant fraction of the current charge, rather than as a fixed loss of energy per unit time.

Since the only way S appears in the Theorem 5 is as this multiplicative factor of S^t, it is tempting to assume that the optimal charging schedule for $S < 1$ will be the same as that for $S = 1$ with the exception that intervals on which the marginal cost is flat are replaced by exponentially increasing periods. Figure 6 shows that this is not the case. The increased leakage causes the storage to empty and fill more frequently. As a result, intervals over which the marginal cost is constant with $S = 1$ are split into multiple intervals of increasing marginal cost, separated by downward jumps.

It appears that, whenever the price decreases for a low rate of leakage, the price also decreases for a higher rate of leakage. Conversely, whenever the price jumps *up* for a *high* rate of leakage, it also jumps up for a lower rate. This is because these decreases occur when the storage is empty, and the storage is empty with high leakage whenever it is with low leakage.

VI. Conclusion and Acknowledgement

We have demonstrated that optimal management of storage must consider inefficiency. Inefficient storage shares the useful "renewal" structure of ideal storage, which allows exact calculation of the optimal schedule knowing load only a finite time into the future. However inefficiency introduces conflict between the goals of reducing peak load and reducing total energy consumption, and also competing arguments for both larger and smaller storage capacities. We hope that this encourages future theoretical studies to consider the non-trivial implications of imperfections in storage technology.

This work was funded by ARC grant FT0991594. We thank Steven Low for valuable feedback on drafts of this paper.

References

[1] K. Chandy, S. Low, U. Topcu, and H. Xu, "A simple optimal power flow model with energy storage," in *Proc. IEEE Conference on Decision and Control (CDC)*, pp. 1051 –1057, Dec. 2010.

[2] L. Huang, J. Walrand, and K. Ramchandran, "Optimal demand response with energy storage management," *preprint arXiv:1205.4297*, 2012.

[3] N. Li, L. Chen, and S. Low, "Optimal demand response based on utility maximization in power networks," in *IEEE Power and Energy Society General Meeting*, pp. 1–8, 2011.

[4] R. Urgaonkar, B. Urgaonkar, M. Neely, and A. Sivasubramaniam, "Optimal power cost management using stored energy in data centers," in *Proc. ACM SIGMETRICS*, pp. 221–232, ACM, 2011.

[5] Australian Energy Market Operator. http://www.aemo.com.au/Electricity/Data/Price-and-Demand/Aggregated-Price-and-Demand-Data-Files/Aggregated-Price-and-Demand-2011-to-2015#2012.

[6] EPRI, *Handbook of Energy Storage for Transmission or Distribution Applications*, 2002. [Online].

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

Appendix A
Structure of finite horizon optimum

We first investigate the structure of the finite horizon problem (1), using Lagrange duality [7]. In particular, we identify the Karush-Kuhn-Tucker (KKT) conditions.

The Lagrangian of the above optimisation problem L is then

$$L = \sum_{t=1}^{T} \Big[P(t)N(g(t)) + \phi(t)\left[b(t) - Sb(t-1) - \eta c(t) + d(t)\right] + \theta(t)\left[D(t) + c(t) - d(t) - g(t)\right] - \lambda(t)g(t) - \underline{\beta}(t)b(t) - \bar{\beta}(t)(B - b(t)) + \xi(b(T) - F) - \underline{\chi}(t)c(t) - \bar{\chi}(t)(C_{\max} - c(t)) - \underline{\delta}(t)d(t) - \bar{\delta}(t)(D_{\max} - d(t)) \Big]$$

Stationarity with respect to $g(t)$, $b(t)$, $c(t)$ and $d(t)$ gives

$$P(t)N'(g^T(t)) - \theta^T(t) - \lambda^T(t) = 0 \quad (5a)$$

$$\phi^T(t) + (-S\phi^T(t+1) - \underline{\beta}^T(t) + \bar{\beta}^T(t))\mathbf{1}_{t<T} + \xi^T \mathbf{1}_{t=T} = 0 \quad (5b)$$

$$\theta^T(t) - \eta\phi^T(t) - \underline{\chi}^T(t) + \bar{\chi}^T(t) = 0 \quad (5c)$$

$$\phi^T(t) - \theta^T(t) - \underline{\delta}^T(t) + \bar{\delta}^T(t) = 0 \quad (5d)$$

where $\mathbf{1}_A = 1$ if A is true and 0 otherwise. The complementary slackness conditions are

$$b^T(t)\underline{\beta}^T(t) = 0 \qquad (B - b^T(t))\bar{\beta}^T(t) = 0$$
$$c^T(t)\underline{\chi}^T(t) = 0 \qquad (C_{\max} - c^T(t))\bar{\chi}^T(t) = 0$$
$$d^T(t)\underline{\delta}^T(t) = 0 \qquad (D_{\max} - d^T(t))\bar{\delta}^T(t) = 0$$
$$g^T(t)\lambda^T(t) = 0$$

Next by subtracting (5c) from (5d) and making $\theta^T(t)$ the subject we get,

$$\theta^T(t) = \frac{1}{2}[(1+\eta)\phi^T(t) + \underline{\chi}^T(t) - \bar{\chi}^T(t) - \underline{\delta}^T(t) + \bar{\delta}^T(t)] \quad (6)$$

Then solving (5b) iteratively gives

$$\phi^T(t) = \xi^T + \sum_{\tau=t}^{T-1} S^{\tau-t}\left[\underline{\beta}^T(\tau) - \bar{\beta}^T(\tau)\right] \quad (7)$$

Finally by substituting (6) and (7) in (5a) and applying the primal feasibility condition $g^T(t) \geq 0$ to eliminate $\hat{\lambda^T}(t)$, we get the optimal solution

$$P(t)N'(g^T(t)) = \frac{1}{2}[(1+\eta)\left(\xi^T + \sum_{\tau=t}^{T-1} S^{\tau-t}\left(\underline{\beta}^T(\tau) - \bar{\beta}^T(\tau)\right)\right) + \underline{\chi}^T(t) - \bar{\chi}^T(t) - \underline{\delta}^T(t) + \bar{\delta}^T(t)]^+ \quad (8)$$

Note in particular that under the ideal storage assumption,

$$P(t)N'(g^T(t)) = \left[\sum_{\tau=t}^{T-1}(\underline{\beta}^T(\tau) - \bar{\beta}^T(\tau)) + \xi^T\right]^+ \quad (9)$$

Proof of Theorem 5: Add (5c) and (5d) and substitute (7) to get

$$(1-\eta)\,(\xi^T + \sum_{\tau=t}^{T-1} S^{\tau-t}\left[\underline{\beta}^T(\tau) - \bar{\beta}^T(\tau)\right]) = \underline{\chi}^T(t) - \bar{\chi}^T(t) - \underline{\delta}^T(t) + \bar{\delta}^T(t) \quad (10)$$

Next by substituting either $\underline{\delta}^T(t) - \bar{\delta}^T(t)$ or $\underline{\chi}^T(t) - \bar{\chi}^T(t)$ from (10) into (8) and substituting (3) we get respectively

$$P(t)N'(g^T(t)) = \left[\eta S^{-t}R + \underline{\chi}^T(t) - \bar{\chi}^T(t)\right]^+ \quad (11)$$

$$P(t)N'(g^T(t)) = \left[S^{-t}R - \underline{\delta}^T(t) + \bar{\delta}^T(t)\right]^+ \quad (12)$$

When the storage is charging ($c(t) > 0$), then $\underline{\chi}^T(t) = \bar{\delta}^T(t) = 0$. Thus (11) implies $P(t)N'(g^T(t)) \leq \left[\eta S^{-t}R - \bar{\chi}^T(t)\right]^+$ which establishes that $P(t)N'(g^T(t)) \leq \eta S^{-t}R^+$ with inequality only if $\bar{\chi}^T(t) > 0$ which, by (2e) only occurs if $c(t) = C_{\max}$.

When the storage is discharging ($d(t) > 0$), then $\underline{\delta}^T(t) = \bar{\chi}^T(t) = 0$. Thus (12) implies $P(t)N'(g^T(t)) \geq \left[S^{-t}R + \bar{\delta}^T(t)\right]^+$ which establishes that $P(t)N'(g^T(t)) \geq S^{-t}R^+$ with inequality only if $\bar{\delta}^T(t) > 0$ which, by (2f) only occurs if $d(t) = D_{\max}$.

Since the sum in (3) starts at t, R increases when $\bar{\beta}^T(t) > 0$ and decreases when $\underline{\beta}^T(t) > 0$. ■

Next, consider the following lemmas.

Lemma 6. *For any $F_1, F_2 \in [0, B]$, if $b^T_{F1}(\tau) = b^T_{F2}(\tau)$ for some $\tau \in [0, T]$, then $b^T_{F1}(t) = b^T_{F2}(t)$ for all $t \in [0, \tau]$.*

Proof: Let $A = b^T_{F1}(t) = b^T_{F2}(t)$. Since the costs and constraints at different times are only coupled by (2a), for all $t \in [0, \tau]$, both $b^T_{F1}(t)$ and $b^T_{F2}(t)$ are equal to the solution $b^{\tau}_A(t)$ to the problem with T replaced by τ and (2g) replaced by $b(\tau)$= A. ■

Lemma 7. *For any $F_1, F_2 \in [0, B]$ with $F_1 < F_2$ and any $\tau \in [0, T]$, if there is no such $t \in [\tau, T]$, that $b^T_{F1}(t) = b^T_{F2}(t)$, then $b^T_{F1}(t) < b^T_{F2}(t)$ for all $t \in [\tau, T]$.*

We can now prove our main monotonicity result.

Proof of Lemma 2: Since $F_1 < F_2$, (2g) gives $b^T_{F1}(T) < b^T_{F2}(T)$. Let $\tau \in [0, T]$ be the last time that $b^T_{F1}(\tau) = b^T_{F2}(\tau)$; this exists since $b^T_{F1}(0) = b^T_{F2}(0) = 0$. It follows from Lemma 6 that $b^T_{F1}(t) = b^T_{F2}(t)$ for all $t \in [0, \tau]$. By definition there is no $t \in (\tau, T]$ such that $b^T_{F1}(t) = b^T_{F2}(t)$, and so Lemma 7 states that $b^T_{F1}(t) < b^T_{F2}(t)$ for all $t \in (\tau, T]$. ■

Since the prefix on $[0, \tau]$ of an optimal solution is itself optimal (i.e., $b^T_{F1}(t) = b^{\tau}_{F2}(t)$ where $F_2 = b^T_{F_1}(\tau)$), Lemma 2 implies that, for a given t, $b^T_0(t)$ is monotonic increasing in T and $b^T_B(t)$ is monotonic decreasing.

Lemma 8. *If at some time t we have $b^T_B(t) = 0$ (resp. $b^T_0(t) = B$), then for any $j \in [0, B]$ the optimal solution to any problem b^T_j, where $j \in [0, B)$, also saturates below at time t, and $b^T_B(t') = b^T_j(t')$ for all time $t' \leq t$.*

Proof: If at some time $t \in [0, T]$ we have $b^T_B(t) = 0$, then at time t, $b^T_j(t) = 0$ by Lemma 2 (i.e., due to monotonicity). Then Lemma 6 implies that $b^T_B(t') = b^T_j(t')$ for all $t' \leq t$. ■

Proof of Theorem 3: We will prove the case that $b^T_B(t) = 0$; the case for $b^T_0(t) = B$ is analogous.

Choose an arbitrary $k \in [0, B]$. By Theorem 2 and Lemma 6, $b^T_k(t') = b^T_B(t')$ for all $t' \leq t$.

Next consider a horizon $T' \geq T$. Let $A = b^{T'}_B(T) \leq B$. Then $b^{T'}_B(t) = b^T_A(t)$ by Lemma 6, whence $b^{T'}_B(t) = 0$ by Theorem 2. Then again $b^{T'}_k(t') = b^{T'}_B(t')$ for all $t' \leq t$. ■