A Polymatroid Approach to Separable Convex Optimization with Linear Ascending Constraints

P. T. Akhil ECE Department Indian Institute of Science Bangalore 560012, India Email: akhilpt@gmail.com Rahul Singh
ECE Department
Texas A&M University
College Station, Texas 77801, USA
Email: rahulsiitk@gmail.com

Rajesh Sundaresan ECE Department Indian Institute of Science Bangalore 560012, India Email: rajeshs@ece.iisc.ernet.in

Abstract—We revisit a problem studied by Padakandla and Sundaresan [SIAM J. Optim., August 2009] on the minimization of a separable convex function subject to linear ascending constraints. The problem arises as the core optimization in several resource allocation problems in wireless communication settings. It is also a special case of an optimization of a separable convex function over the bases of a specially structured polymatroid. We give an alternative proof of the correctness of the algorithm of Padakandla and Sundaresan. In the process we relax some of their restrictions placed on the objective function.

I. Introduction

In this paper we consider the following separable convex optimization problem with linear inequality constraints. This optimization problem arises in a wide variety of resource allocation problems in communication settings. We first describe the abstract problem before saying a few words about its applicability in communication settings.

Let $x=(x(1),x(2),\cdots,x(n))\in\mathbb{R}^n$. Let the functions $w_e:(a_e,b_e)\to\mathbb{R},\ e=1,2,\cdots,n$ be strictly convex and continuously differentiable (\mathcal{C}^1) functions. Assume $a_e<0< b_e$. Let $W:\mathbb{R}^n\to\mathbb{R}$ be the separable function $\sum_e w_e$.

Problem Π :

$$\begin{array}{ll} \text{Minimize} & W(x) := \sum_{e=1}^n w_e(x(e)) \\ \text{subject to} & x(e) \in [0,\beta(e)], \qquad e = 1,2,\cdots,n, \qquad (1) \\ & \sum_{e=1}^l x(e) \geq \sum_{e=1}^l \alpha(e), \quad l = 1,2,\cdots,n-1, \ (2) \\ & \sum_{e=1}^n x(e) = \sum_{e=1}^n \alpha(e). \end{array}$$

We assume $\beta(e) \in (0,b_e]$ for $e=1,2,\cdots,n$, and $\alpha(e) \geq 0$ for $e=1,2,\cdots,n$. The inequalities in (1) impose positivity and upper bound constraints. The inequalities in (2) impose a sequence of *ascending constraints* with increasing heights $\sum_{e=1}^{l} \alpha(e)$ indexed by l. We also assume that

$$\sum_{e=1}^{l} \alpha(e) \le \sum_{e=1}^{l} \beta(e), \tag{4}$$

978-1-4799-2361-8/14/\$31.00 © 2014 IEEE

a necessary condition for the feasible set to be nonempty.

We now discuss some settings where Problem Π arises. If $w_e(t) = \log(1 + \frac{t}{a(e)})$ and $\alpha(1) = \alpha(2) = \cdots = \alpha(n-1) = 0$, but $\alpha(n) = P$, we have the classical power allocation problem on the vector gaussian channel with the familiar waterfilling solution. See Patriksson [9] for a survey of such problems with a single sum constraint. Problem Π also arises in a multiple access physical layer power and sequence allocation problem where mobiles have a rate requirement (Padakandla and Sundaresan [8]). The allocations should be such that total system power, as measured by the sum of the mobiles' powers, is minimized. See Viswanath and Anantharam [11] for a problem on sum rate maximization subject to power constraints, Zacharias and Sundaresan [12], [13] for two related problems on power optimization in a wireless sensor network. These communication theoretic applications of Problem II have motivated us to send our work to this conference. Problem Π arises in other settings as well. See Bellman and Dreyfus [1, p.105] for a problem on smoothing, Dantzig [2] for a network flow problem, and Veinott Jr. [10] for an extension of Dantzig's problem.

Morton et al. [6] studied the special case when $w_e(t) = \lambda(e)t^p, e=1,2,\cdots,n$, where p>1. They characterized the constraint set as the *bases* of a *polymatroid*; both terms will be defined shortly. Problems of optimizations over polymatroids arise frequently in communications, networks, and queues. Fujishige [3] provided an algorithm to find the lexicographically optimal base of a polymatroid. Groenevelt [5] subsequently extended it to find the base of a polymatroid that minimized a separable convex objective function. Fujishige [4, Ch. 8] provided a decomposition algorithm for minimizing over the bases of a more general submodular system. These algorithms are generic and apply to any polymatroid (or submodular system). They do not exploit the structure of the special polymatroid arising from the constraints in (1)-(3).

Morton et al. [6] gave another algorithm that did exploit the special structure arising from the constraints in (1)-(3), but for the special case of $w_e(t) = \lambda(e)t^p$, p > 1. Padakandla and Sundaresan [7] provided an extension of this algorithm

¹Intuitively, the components are most balanced.

²The exact definition of a submodular system need not concern us here.

to strictly convex \mathcal{C}^1 functions that satisfy certain slope conditions. Their proof of optimality is via a verification of the Karush-Kuhn-Tucker (KKT) conditions. In this paper, our goal is to give a polymatroidal interpretation to the proof of correctness of the algorithm of Padakandla and Sundaresan [7], and in the process relax some restrictions assumed in [7].

II. PRELIMINARIES

In this section, we state preliminary results that reduce Problem Π to an optimization over the bases of an appropriate polymatroid. This reduction is due to Morton et al. [6] and is given here for completeness. We then state a result due to Groenevelt [5] for polymatroids which was subsequently generalized to submodular functions by Fujishige [4, Ch. 8]. Groenevelt's result will suffice to provide a necessary and sufficient condition for optimality in Problem Π. The next section provides the algorithm of Padakandla and Sundaresan [7], rewritten in our notation, to arrive at a candidate solution that satisfies the necessary and sufficient condition of Groenevelt [5]. We begin with some relevant definitions.

Let $E = \{1, 2, ..., n\}$. Let $f: 2^E \to \mathbb{R}_+$ be a *rank* function, i.e., a nonnegative real function on the set of subsets of E satisfying

$$f(\emptyset) = 0 \tag{5}$$

$$f(A) \leq f(B), (A \subseteq B \subseteq E)$$
 (6)

$$f(A) + f(B) \ge f(A \cup B) + f(A \cap B), (A, B \subseteq E).$$
 (7)

Equation (5) says f is normalized, (6) says f is increasing, and (7) says f is submodular. The pair (E, f) is called a *polymatroid* with ground set E. For an $x \in \mathbb{R}_+^E$ and $A \subseteq E$ define the convenient notation

$$x(A) := \sum_{e \in A} x(e).$$

An $x \in \mathbb{R}_+^E$ is called an *independent* vector if $x(A) \leq f(A)$ for every $A \subseteq E$. Let P(f), called the *polymatroidal polyhedron*, denote the set of all independent vectors of (E, f). The bases of the polymatroid (E, f), denoted B(f), is defined as

$$B(f) := \{x \in P(f) : x(E) = f(E)\}.$$

These are the maximal elements of P(f) with respect to the partial order " \leq " on \mathbb{R}_+^E defined by component-wise domination $(x \leq y \text{ if and only if } x(e) \leq y(e) \text{ for every } e \in E$)

For an $x \in B(f)$, a base, and an $e \in E$, define

$$dep(x, e, f) = \bigcap \{A \mid e \in A \subset E, x(A) = f(A)\},\$$

which in words is the smallest subset among those subsets A of E containing e for which x(A) equals the upper bound f(A). Fujishige [4] shows that dep(x,e,f) is made up of all those elements $u \in E$ from which a small amount of mass can be moved from x(u) to x(e) yet keeping the new vector independent (indeed, the new vector remains a base). Thus (u,e) is called an *exchangeable pair* if $u \in dep(x,e,f) - \{e\}$.

For a $\beta \in \mathbb{R}_+^E$, define the set function

$$f_{\beta}(A) = \min_{D \subseteq A} \{ f(D) + \beta(A - D) \} \quad (A \subset E).$$

We now state without proof an interesting property of the subset of independent vectors of a polymatroid that are dominated by β . See Fujishige [3] for a proof.

Proposition 1: The set function f_{β} is a rank function and (E, f_{β}) is a polymatroid. Furthermore, $P(f_{\beta})$ is given by $P(f_{\beta}) = \{x \in P(f) : x \leq \beta\}.$

We now relate the constraint set in Problem Π to the bases of a polymatroid, as done by Morton et al [6]. Define

$$c(0) := 0$$
 $c(j) := \sum_{e=1}^{j} \alpha(e) \quad (1 \le j \le n)$ (8)

and further define

$$\zeta(A) := \max \{ c(j) : \{1, 2, \dots, j\} \subseteq A, 0 \le j \le n \}, \ (A \subseteq E)$$
$$f(A) := \zeta(E) - \zeta(E - A) = c(n) - \zeta(E - A), \ (A \subseteq E). \tag{9}$$

Proposition 2:

- The f in (9) is a rank function and therefore (E,f) is a polymatroid.
- The set of all $x \in \mathbb{R}_+^E$ that satisfy the ascending constraints (2)-(3) equals the bases B(f) of the polymatroid (E, f).
- The set of $x \in \mathbb{R}_+^{\mathbb{E}}$ that satisfy the ascending constraints (2)-(3) and the domination constraint (1) equals the bases $B(f_{\beta})$ of the polymatroid (E, f_{β}) .

See Morton et al. [6] for a proof. Incidentally, this is shown by recognizing that (E,ζ) is a related object called the contrapolymatroid, that the set of all vectors meeting the constraints above are bases of the contrapolymatroid, and that the bases of contrapolymatroid (E,ζ) and the bases of the polymatroid (E,f) coincide. The above proposition thus says that Problem Π is simply a special case of Problem Π_1 below with $g=f_\beta$.

Let (E,g) be a polymatroid.

Problem
$$\Pi_1:$$
 Minimize $\sum_{e\in E} w_e(x(e))$ (10) subject to $x\in B(g)$

We next state a necessary and sufficient condition for a base in B(g) to be optimal. For each $e \in E$, define w'_e to be the derivative of w_e .

Theorem 1 (Groenevelt [5]): A base $x \in B(g)$ is an optimal solution of Problem Π_1 if and only if for each exchangeable pair (u,e) associated with base x, i.e., $u \in \text{dep}(x,e,g) - \{e\}$, we have $w'_e(x(e)) \ge w'_u(x(u))$.

For a proof and generalization to submodular systems, see Fujishige [4, Th. 8.1] and [4, Th. 8.2]. We have been at pains to state this result in a precise fashion because the proof of our main result hinges on Theorem 1.

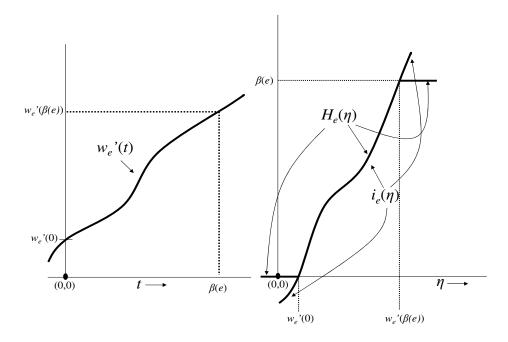


Fig. 1. The functions w'_e , i_e and H_e .

III. MAIN RESULTS

Fujishige [4, Ch. 8] provides the so-called *decomposition* algorithm, a general algorithm for submodular systems, for finding the optimal solution. This proceeds by identifying a chain of subsets of E with certain properties and thence the optimal solution. In this section, however, we extend an algorithm of Morton et al. [6] to arrive at the chain, and the optimal solution, in a direct fashion. The algorithm is a rewording of the algorithm of Padakandla and Sundaresan [7], and reduces to that of Morton et al. [6] for the case when the objective function's constituents are $w_e(\zeta) = \lambda(e)\zeta^p$, $e \in E$, where p > 1.

Define $i_e := (w'_e)^{-1}$ and

$$H_e(\eta) := \begin{cases} 0 & \eta < w'_e(0) \\ i_e(\eta) & w'_e(0) \le \eta \le w'_e(\beta(e)) \\ \beta(e) & \eta > w'_e(\beta(e)). \end{cases}$$

See Figure 1 for an illustration of the functions w'_e , i_e , and H_e . The function H_e is i_e but truncated at zero and saturated at $\beta(e)$.

Algorithm 1 (Sundaresan and Padakandla [7]):

Let s(0) = 0 and define $s(j) \in \{1, 2, \dots, n\}$ recursively for $j = 1, 2, \dots$ as follows. Let a_l^j be the smallest solution of

$$\sum_{e=s(j-1)+1}^{l} H_e(\eta) = c(l) - c(s(j-1)), \quad (s(j-1) < l \le n),$$

if it exists. Let

$$\Gamma_j := \max\{a_l^j : s(j-1) < l \le n\},$$
(12)

and let s(j) be the largest index in $s(j-1) < l \le n$ where the maximum is attained. Since $0 = s(0) < s(1) < \cdots \le n$, the process terminates after say p steps so that s(p) = n. Set

$$T_j := \{s(j-1) + 1, \dots, s(j)\} \ (1 \le j \le p).$$

Finally, set
$$x(e) = H_e(\Gamma_i), e \in T_i, 1 < j < p.$$

The following should be shown to ensure that the algorithm terminates and generates the desired allocation.

- 1) The set whose maximum is taken in (12) should be nonempty at each iteration step.
- 2) The output of the algorithm should be feasible.
- The output should satisfy the sufficient condition of Theorem 1 for optimality.

These were done by Padakandla and Sundaresan [7] via a direct verification of the KKT conditions without exploiting the polymatroid properties. Their proof assumed some conditions on the slopes $w_e^\prime(0)$. Our main result is the following.

Theorem 2: (Correctness of the Algorithm) If the feasible set satisfying constraints (1)-(3) is nonempty, Algorithm 1 runs to completion and puts out the unique optimal vector.

Observe that the only hypothesis is that the feasible set is nonempty; no conditions on the slopes $w_e'(0)$ are needed. The next section is devoted to the proof of this theorem.

IV. PROOF OF CORRECTNESS AND OPTIMALITY

We begin by proving that a solution to (11) exists for all iterations and all relevant indices l. This will establish that the set whose maximum is taken in (12) is nonempty.

Lemma 1: If the feasible set satisfying constraints (1)-(3) is nonempty, then (11) has a solution for all indices l satisfying $s(j-1) < l \le n$ and for all iterations $j = 1, 2, \dots, p$.

Proof: Let the feasible set be nonempty. We claim that for $1 \le j \le p$, the following hold:

(a)
$$0 \leq c(l) - c(s(j-1))$$

$$= \sum_{e=s(j-1)+1}^{l} \alpha(e)$$

$$\leq \sum_{e=s(j-1)+1}^{l} \beta(e), \quad (s(j-1) < l \le n), (13)$$

(b) A solution to (11) exists for $s(j-1) < l \le n$.

We prove the claim by induction on j.

For j=1, s(j-1)=s(0)=0, and therefore claim (a) holds on account of being a necessary condition for feasibility. Claim (b) for j=1 immediately follows for each l by considering

$$\underline{\eta} = \min \{ w'_e(0) : s(j-1) < e \le l \},$$
 $\overline{\eta} = \max \{ w'_e(\beta(e)) : s(j-1) < e \le l \},$

the definition of H_e , and its continuity. Indeed

$$\sum_{e=s(j-1)+1}^{l} H_e(\underline{\eta}) = 0$$

and

$$\sum_{e=s(j-1)+1}^{l} H_e(\overline{\eta}) = \sum_{e=s(j-1)+1}^{l} \beta(e),$$

and since $\sum_{e=s(j-1)+1}^{l} \alpha(e)$ is in between, there is an η that attains it, thus solving (11). This establishes the claim above for j=1.

Suppose that the claim holds for a particular j such that j . We shall now prove its validity for <math>j + 1. By the choice of a_l^j in iteration j, the smallest η solving (11), we have

$$\sum_{e=s(j-1)+1}^{l} H_e\left(a_l^j\right) = c(l) - c(s(j-1)), \quad (s(j) < l \le n),$$
(14)

and

$$\sum_{e=s(j-1)+1}^{s(j)} H_e\left(a_{s(j)}^j\right) = c(s(j)) - c(s(j-1)). \tag{15}$$

Furthermore $a_l^j < a_{s(j)}^j$ for $s(j) < l \leq n$, because of the choice of s(j) as the largest index attaining the maximum in

(12). As $H_e(\cdot)$ are increasing functions, we deduce from (14), (15), and $a_l^j < a_{s(j)}^l$ for $s(j) < l \le n$ that

$$\sum_{s(j-1)+1}^{s(j)} H_e\left(a_l^j\right) < c(s(j)) - c(s(j-1)).$$
 (16)

Strict inequality holds in (16) for otherwise, under equality, a_l^j for an l>s(j) is a strictly smaller solution to (11) than $a_{s(j)}^j$ in iteration j, which contradicts the choice of $a_{s(j)}^j$. Now subtract (14) from (16) and use the positivity of c(l)-c(s(j)) for all l>s(j) to get

$$0 \leq c(l) - c(s(j))$$

$$= \sum_{s(j)+1}^{l} \alpha(e)$$

$$< \sum_{s(j)+1}^{l} H_e\left(a_l^j\right)$$

$$\leq \sum_{s(j)+1}^{l} \beta(e), \quad (s(j) < l \leq n),$$

where the last inequality follows from the definition of $H_e(\cdot)$. This proves (a) of the claim. Part (b) of the claim holds as in the case of j=1 with $\overline{\eta}$ replaced by a_l^j . This completes the proof of the claim. Claim (b) is the statement of the lemma.

The following lemma shows feasibility and optimality of Algorithm 1 to complete the proof of Theorem 2.

Lemma 2: The output of Algorithm 1 is the unique optimal solution to Problem Π .

Proof: We first claim that $\Gamma_j > \Gamma_{j+1}, \ j=1,2,\cdots,p-1$. Indeed, we have s(j+1)>s(j) and therefore in iteration j, we must have $a^j_{s(j)}>a^j_{s(j+1)}$. The increasing property of H_e implies that

$$c(s(j+1)) - c(s(j-1))$$

$$= \sum_{e=s(j-1)+1}^{s(j+1)} H_e\left(a_{s(j+1)}^j\right)$$

$$< \sum_{e=s(j-1)+1}^{s(j+1)} H_e\left(a_{s(j)}^j\right)$$

$$= \sum_{e=s(j-1)+1}^{s(j)} H_e\left(a_{s(j)}^j\right) + \sum_{e=s(j)+1}^{s(j+1)} H_e\left(a_{s(j)}^j\right)$$

$$= c(s(j)) - c(s(j-1)) + \sum_{e=s(j)+1}^{s(j+1)} H_e\left(a_{s(j)}^j\right)$$
(19)

Strict inequality holds in (18) because otherwise *saturation* should have occurred for all e in the sum, i.e.,

$$H_e\left(a_{s(j+1)}^j\right) = H_e\left(a_{s(j)}^j\right) = \beta(e), \quad s(j-1) < e \le s(j),$$

implying that the algorithm would have picked $a_{s(j+1)}^{\jmath}$ or some number smaller as a solution to (11) instead of $a_{s(i)}^{j}$ for l = s(j) in iteration j, a contradiction. From (19), we

$$\sum_{e=s(j)+1}^{s(j+1)} H_e\left(a_{s(j+1)}^{j+1}\right) = c(s(j+1)) - c(s(j))$$

$$< \sum_{e=s(j)+1}^{s(j+1)} H_e\left(a_{s(j)}^{j}\right).$$

Using the fact that H_e is increasing, we conclude that

 $a_{s(j+1)}^{j+1} < a_{s(j)}^{j} \text{ or } \Gamma_{j} > \Gamma_{j+1}.$ We next prove that the output of the algorithm is feasible. Since the range of H_e is restricted between 0 and $\beta(e)$, and $x(e) = H_e(\Gamma_i)$ for $s(j-1) < e \le s(j)$, the constraint (1) is automatically satisfied. Next observe that since $\Gamma_j = a_{s(j)}^j$, we have

$$\sum_{e=s(j-1)+1}^{s(j)} x(e) = \sum_{e=s(j-1)+1}^{s(j)} H_e(\Gamma_j)$$

$$= \sum_{e=s(j-1)+1}^{s(j)} H_e\left(a_{s(j)}^j\right)$$

$$= c(s(j)) - c(s(j-1))$$

so that $\sum_{e=1}^{s(j)} x(e) = c(s(j))$ by induction on j. Thus equality holds for indices s(j), $j = 1, 2, \dots, p$. To show x is feasible, it is now sufficient to show that for any fixed j, we have

$$\sum_{e=s(j-1)+1}^{l} x(e) \ge c(l) - c(s(j-1)), \quad (s(j-1) < l \le s(j)).$$

But this holds because $a_{s(i)}^j \ge a_l^j$ for all such l, and x(e) = $H_e(\Gamma_j) = H_e\left(a_{s(j)}^j\right)$, which imply

$$\sum_{e=s(j-1)+1}^{l} x(e) = \sum_{e=s(j-1)+1}^{l} H_e \left(a_{s(j)}^j \right)$$

$$\geq \sum_{e=s(j-1)+1}^{l} H_e \left(a_l^j \right)$$

$$= c(l) - c(s(j-1)).$$

Thus x is feasible.

We now use Theorem 1 to prove the optimality of x. Recall that f is as given in (9) and $g = f_{\beta}$. Define $A_0 = \emptyset$ and set $A_j = \bigcup_{i=1}^{j} T_{p-i+1}$ for $j = 1, 2, \dots, p$. Observe that $A_p = E$; it is then an immediate consequence that $x(A_i) = f(A_i)$. Next observe that $dep(x, e, f) \subseteq A_i$ for every $e \in A_j - A_{j-1}, j = 1, \dots, p$, and $dep(x, e, g) \subseteq$ dep(x, e, f). Furthermore, $dep(x, e, g) = \{e\}$ for every e satisfying $x(e) = \beta(e)$, and a $u \neq e$ with x(u) = 0 cannot belong to dep(x, e, g). This implies that for any $u \in dep(x, e, g)$,

 $u \neq e$, we must have $x(e) < \beta(e)$ and 0 < x(u). We then claim that

$$w'_{e}(x(e)) = w'_{e}(H_{e}(\Gamma_{j})) \stackrel{(a)}{\geq} \Gamma_{j} > \Gamma_{j+1}$$

$$\stackrel{(b)}{\geq} w'_{u}(H_{u}(\Gamma_{j+1}))$$

$$= w'_{u}(x(u)).$$

Inequality (a) above clearly holds with equality if x(e) > 0, and with inequality if $x(e) = H_e(\Gamma_j) = 0$. Similarly (b) clearly holds with equality if $x(u) < \beta(u)$, and with inequality if $x(u) = H_u(\Gamma_i) = \beta(u)$. The sufficient condition of Theorem 1 for optimality holds, and the proof is complete.

Proof of Theorem 2: The validity of Theorem 2 follows immediately from Lemma 1 and Lemma 2.

ACKNOWLEDGMENT

This work was supported by the Department of Science & Technology, Government of India.

REFERENCES

- [1] R. E. Bellman and S. R. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, NJ, 1962.
- G. B. Dantzig, "A control problem of Bellman", Management Science, 17 (1971), pp. 542-546. S. Fujishige, "Lexicographically optimal base of a polymatroid with
- respect to a weight vector," Mathematics of Operations Research, vol. 5, no. 2, pp. 186-196, May 1980.
- [4] S. Fujishige, Submodular Functions and Optimization, 2nd ed., ser. Annals of Discrete Mathematics, Elsevier, 2005, vol. 58.
- H. Groenevelt, "Two algorithms for maximizing a separable concave function over a polymatroid feasible region," European Journal of Operational Research, 54 (1991) 227-236.
- [6] G. Morton, R. Von Randow, K. Ringwald, "A greedy algorithm for solving a class of convex programming problems and its connection with polymatroid theory," Mathematical Programming, 32 (1985), pp. 238-241.
- [7] A. Padakandla and R. Sundaresan, "Separable Convex Optimization Problems with Linear Ascending Constraints," SIAM Journal on Optimization, vol. 20, no. 3, pp. 1185-1204, 19 August 2009.
- A. Padakandla and R. Sundaresan, "Power minimization for CDMA under colored noise," IEEE Transactions on Communications, vol. 57, no.10, pp.3103-3112, October 2009.
- [9] M. Patriksson, "A survey on the continuous nonlinear resource allocation problem," European Journal of Operational Research, 185 (2008), pp.
- [10] A. F. Veinott Jr., "Least d-majorized network flows with inventory and statistical applications," Management Science, 17 (1971), pp. 547-567.
- [11] P. Viswanath and V. Anantharam, "Optimal sequences for CDMA with colored noise: A Schur-saddle function property," IEEE Transactions on Information Theory, IT-48 (2002), pp. 1295-1318.
- [12] L. Zacharias and R. Sundaresan, "An analog MVUE for a wireless sensor network," Proceedings of the 2008 National Conference on Communication, NCC 2008, IIT Mumbai, February 2008.
- L. Zacharias and R. Sundaresan, "Decentralized sequential change detection using physical layer fusion," IEEE Transactions on Wireless Communications, vol. 7, no. 12, pp. 4999-5008, December 2008.