COMMUNICATIONS IN INFORMATION AND SYSTEMS
Vol. 1, No. 2, pp. 163-180, April 2001

003

AN OPTIMIZATION ON A MANIFOLD APPROACH FOR SOLVING AN ANTENNA ARRAY PROBLEM*

JOHN B. MOORE†, JEREMY S. THORNE†, AND TAT M. LOK†

Abstract. There is a fundamental existence and construction question which arises in many array signal processing problems. Although there is a wealth of experience and intuition about this question, a fundamental theorem and proof is only now here documented. The question concerns existence and calculation of power gains $p_i \geq 0$ which satisfy the set of equalities $p_i g_i^H (\sum p_i g_i g_i^H)^{-1} g_i = x_i$ for given vectors g_i and scalars $1 > x_i > 0$. The results will have application in antenna array optimization and direct sequence CDMA communication system design.

1. Introduction. In the optimization of certain antenna array problems and the design of direct sequence CDMA communication systems, there is a fundamental existence and construction question which arises. The first challenge is to tackle the issues of singularity inherent in the problem to achieve an existence result, and understand the meaning of the necessary and sufficient conditions for applications. The next challenge is to provide practical algorithms, including decentralized algorithms, for achieving on-line optimization for adaptive power control. For this, it also important to understand when there is guaranteed convergence, and to know the rates of convergence.

1.1. Antenna Array. Consider a communication system with m users. During each symbol interval, the ith user, for $1 \leq i \leq m$, sends a data symbol q_i with zero mean and unit variance to the base station. At the base station, an antenna array of n elements is used to capture the signal. The received signal is, therefore, an n-dimensional vector of the form

$$\mathbf{q} = \sum_{i=1}^{m} \sqrt{p_i} q_i \mathbf{g}_i + \mathbf{n}_{therm}, \tag{1.1}$$

where p_i represents the transmitted power of the signal of the ith user. The elements of $\mathbf{g}_i$ represent the gains from the transmitter of the ith user to the antennas at the base station. The vector $\mathbf{n}_{therm}$ represents the contribution of thermal noise.

We consider a decentralized linear receiver for each user at the base station. Therefore, the decision statistic for the symbol of the ith user is of the form $\mathbf{w}_i^H \mathbf{q}$. The signal-to-noise ratio (SNR) for the ith user is given by

$$\mathrm{SNR}_i = \frac{\left|\mathbf{w}_i^H \sqrt{p_i} \mathbf{g}_i\right|^2}{\mathbf{w}_i^H \left(\sum_{j \neq i} p_j \mathbf{g}_j \mathbf{g}_j^H + \sigma^2 I_n\right) \mathbf{w}_i}, \tag{1.2}$$

where σ^2 is the variance of the thermal noise, and I_n is the n-dimensional identity matrix. We pick $\mathbf{w}_i$ to maximize the SNR. Equivalently, we pick $\mathbf{w}_i$ to maximize the signal-to-total-power (STR) ratio

*Received Dec 11, 2000; accepted for publication Mar 30, 2001

† Department of Information Engineering, The Chinese University of Hong Kong. John B. Moore is on leave from the Department of Systems Engineering, the Australian National University, Canberra, Australia. Jeremy S. Thorne was from the same department. John B. Moore's work was partially supported by a grant (RGC Ref. No. CUHK 4227/00E). Email: jmoore@ie.cuhk.edu.hk, tjeremy@ie.cuhk.edu.hk, tmlok@ie.cuhk.edu.hk

$$\text{STR}_i = \frac{\left|\mathbf{w}_i^H \sqrt{p_i}\mathbf{g}_i\right|^2}{\mathbf{w}_i^H \left(\sum_{j=1}^m p_j \mathbf{g}_j \mathbf{g}_j^H + \sigma^2 I_n\right) \mathbf{w}_i} = \frac{\text{SNR}_i}{1+\text{SNR}_i}. \tag{1.3}$$

It can be shown [1] that $\mathbf{w}_i$ should be chosen as

$$\mathbf{w}_i = \left(\sum_{j=1}^m p_j \mathbf{g}_j \mathbf{g}_j^H + \sigma^2 I_n\right)^{-1} \mathbf{g}_i, \tag{1.4}$$

and the resulting STR is given by

$$p_i \mathbf{g}_i^H \left(\sum_{j=1}^m p_j \mathbf{g}_j \mathbf{g}_j^H + \sigma^2 I_n\right)^{-1} \mathbf{g}_i. \tag{1.5}$$

Written in a compact form, the STR for the ith user is the ith diagonal element of

$$\mathbf{PG}\left(\mathbf{G}^H \mathbf{PG} + \sigma^2 I_n\right)^{-1} \mathbf{G}^H. \tag{1.6}$$

Suppose that for satisfactory performance, each user requires a certain target SNR. Equivalently, we can specify the requirement for the ith user as a target STR of x_i where $0 < x_i < 1$. We would like to know whether the set of m users can be supported with suitable choices of p_i. While the application of antenna arrays in wireless communication systems has been a topic of intense research interests [2, 3, 4, 5, 6, 7], this existence problem has not been solved. We consider this existence problem and are interested in the case where $m > n$. Since then the system is interference limited, we would neglect the effect of thermal noise. The STR's become the diagonal elements of

$$\mathbf{PG}\left(\mathbf{G}^H \mathbf{PG}\right)^{-1} \mathbf{G}^H. \tag{1.7}$$

1.2. Direct Sequence CDMA. Consider a synchronous direct sequence CDMA communication system with m users. During each symbol interval, the ith user, for $1 \leq i \leq m$, sends a data symbol q_i with zero mean and unit variance to the base station. The symbol q_i is spread by a signature sequence $\mathbf{g}_i$ of length n before transmission. The received signal is, therefore, an n-dimensional vector of the form

$$\mathbf{q} = \sum_{i=1}^m \sqrt{p_i} q_i \mathbf{g}_i + \mathbf{n}_{therm} \tag{1.8}$$

where p_i represents the transmitted power of the signal of the ith user. The vector $\mathbf{n}_{therm}$ represents the contribution of thermal noise.

It can be shown [8] that the minimum mean square error (MMSE) receiver is again given by (1.4) with the resulting STR given by (1.5). Satisfactory performance, determined by target STR's, may be achieved by varying the power output of each of the users. This power control problem has been of much interests in wireless communications.

Previous research on power control in different wireless communication systems have been focused on the effects of power control [9, 10, 11], the development of algorithms [12, 13, 14, 15, 16], and the proofs of convergence of iterative algorithms [17, 18]. However, the problem on the existence of a power control solution is not fully addressed. In this paper, we consider the existence of a power control solution for a synchronous CDMA system with MMSE receivers.

1.3. Existence Problem. We observe that (1.8) and (1.1), as well as the corresponding receivers, are identical. We are faced with the same problem in both cases. We focus on the antenna array setting. In fact, explicit results are easily derived to provide optimal solutions for the cases:

1. When the number of users m, is equal to the number of antennas in the array n plus one.
2. When the number of antennas in the array n is equal to one.

However, for the case when the number of users m is greater than the number of antennas in the array n plus one, only iterative *ad-hoc* algorithms are available, a proof of existence has been left open (the case when the number of users m, is less than the number of antennas in the array n is trivial). In this paper we establish the existence of a solution for the case of $m > n+1$ under certain constraints, using an optimization on a manifold approach [19].

This paper is organized as follows: Section 2 sets up our problem, states our main existence result, and provides lemmas concerning conditions in our existence result. In Section 3 we establish our existence result by construction, using an optimization on a manifold argument. Gradient flows are defined which are guaranteed to flow to an unique solution at a global minimum within the manifold. Section 4 provides analytical results for the special case examples, and an iterative algorithm which is asymptotically Newton and hence quadratically convergent is introduced to provide practical calculation of the desired solution. Also, a decentralized *ad-hoc* algorithm is given which is more practical in some applications. Section 5 gives conclusions.

2. Existence Result.

2.1. Problem Statement. Consider the equation

$$\mathbf{R} = \left[\mathbf{PG}\left(\mathbf{G}^H\mathbf{PG}\right)^{-1}\mathbf{G}^H - \mathbf{X}\right]_{diag} = 0 \tag{2.1}$$

with matrix diagonal constraints $r_j, p_j, x_j \in \mathbb{R}$ and $\mathbf{g}_i \in \mathbb{C}^n$, $n < m$.

$$\mathbf{R} = \begin{bmatrix} r_1 & & \\ 0 & \ddots & 0 \\ & & r_m \end{bmatrix}, \mathbf{P} = \begin{bmatrix} p_1 & & \\ 0 & \ddots & 0 \\ & & p_m \end{bmatrix}, \tag{2.2}$$

$$\mathbf{X} = \begin{bmatrix} x_1 & & \\ 0 & \ddots & 0 \\ & & x_m \end{bmatrix}, \mathbf{G} = \begin{bmatrix} \mathbf{g}_1^H \\ \vdots \\ \mathbf{g}_m^H \end{bmatrix}$$

Here, $A = (B)_{diag}$ denotes a diagonal matrix A where the elements of A are the diagonal elements of B, and $\mathbf{G}^H$ denotes the Hermitian (transpose conjugate) of $\mathbf{G}$. Of course, it is necessary that $\mathbf{G}$ is full rank n and that $\mathbf{P}$ is rank n to ensure the existence of the inverse $\left(\mathbf{G}^H\mathbf{P}\mathbf{G}\right)^{-1}$.

Taking the trace of (2.1), then for all $\mathbf{P}$, satisfying (2.1),

$$\text{tr}\left(\mathbf{R}\right) = \text{tr}\left[\mathbf{P}\mathbf{G}\left(\mathbf{G}^H\mathbf{P}\mathbf{G}\right)^{-1}\mathbf{G}^H - \mathbf{X}\right] = \text{tr}\left(I_n\right) - \text{tr}\left(\mathbf{X}\right) = 0 \tag{2.3}$$

using the trace operator property $\text{tr}\left(AB\right) = \text{tr}\left(BA\right)$, so that necessary conditions for (2.1) to hold are that $\text{tr}\left(\mathbf{X}\right) = n$, $\text{tr}\left(\mathbf{R}\right) = 0$.

Notice that if some $\mathbf{P}^*$ is a solution of (2.1), so also is $\beta\mathbf{P}^*$ for an arbitrary nonzero scaling factor β.

2.2. Main Existence Theorem. THEOREM 2.1. *Suppose the following constraints are satisfied*

1. $\mathbf{G}$ *is generic in that any selection of* n *rows of* $\mathbf{G}$ *are linearly independent.*
2. $0 < x_j < 1$ *for all* j.
3. $\sum_{j=1}^{m} x_j = n$, *or equivalently* $\sum_{j=1}^{m} r_j = 0$.

Then there exists a positive definite solution $\mathbf{P} = \mathbf{P}^*$ *of* (2.1) *(trivially satisfying* $\left(\mathbf{G}^H\mathbf{P}^*\mathbf{G}\right)^{-1} > 0$*).*

Moreover, there are no limiting solutions $\mathbf{P}^{**}$ *with* $\left(\mathbf{G}^H\mathbf{P}^{**}\mathbf{G}\right)$ *singular.*

Also, Condition 3 and $\mathbf{G}$ *full rank are necessary conditions for* (2.1) *to hold.*

The remainder of this section is devoted to developing a proof of this theorem. The necessity result has been trivially established in (2.3).

2.3. Genericity Condition on G. The concept of a generic $\mathbf{G}$, has important implications in the development of our proof for Theorem 2.1. Consequences of generic $\mathbf{G}$ selections are now developed:

If $\mathbf{G}$ is generic, in that Condition 1 of Theorem 2.1 holds, then for an arbitrary set $\mathcal{J}_{n^\dagger} := \{j_1, j_2, \ldots, j_{n^\dagger}\}$ with $n^\dagger \leq m$, an equivalent condition is

$$\text{rank}\left(\sum_{i\in\mathcal{J}_{n^\dagger}} \mathbf{g}_i\mathbf{g}_i^H\right) = \begin{cases} n^\dagger, & n^\dagger < n \\ n, & n^\dagger \geq n \end{cases}. \tag{2.4}$$

Also, for $n^\dagger < n$ an eigenvalue decomposition, gives the further equivalent conditions

$$\sum_{i\in\mathcal{J}_{n^\dagger}} \mathbf{g}_i\mathbf{g}_i^H = \begin{bmatrix} U_{n^\dagger} \\ U_{n-n^\dagger} \end{bmatrix} diag\left(\lambda_1, \lambda_2, \ldots, \lambda_{n^\dagger}, 0, \ldots, 0\right) \begin{bmatrix} U_{n^\dagger} \\ U_{n-n^\dagger} \end{bmatrix}^H, \tag{2.5}$$

where $\lambda_i > 0$ for $i = 1, 2, \ldots, n^\dagger$, and $\left[U_{n^\dagger}^H \ \ U_{n-n^\dagger}^H\right]$ is unitary.

LEMMA 2.2. *With generic* $\mathbf{G}$*, as in Condition 1 of Theorem 2.1, and* $n^\dagger < n$. *Then*

$$\Re\left(\left\{\mathbf{g}_i|_{i\in\mathcal{J}^{n^\dagger}}\right\}\right) \equiv \Re\left(U_{n^\dagger}^H\right), \ \ \mathbf{g}_j|_{j\notin\mathcal{J}_{n^\dagger}} \notin \Re\left(U_{n^\dagger}^H\right), \tag{2.6}$$

and

$$\mathbf{g}_j^H\big|_{j\in\mathcal{J}_{n^\dagger}} U_{n-n^\dagger} = 0,\ \ \mathbf{g}_j^H\big|_{j\notin\mathcal{J}_{n^\dagger}} U_{n-n^\dagger} \neq 0. \tag{2.7}$$

Here $\Re(B)$ *denotes the range space of* B.

Proof. Clearly for generic $\mathbf{G}$, (2.5) holds and

$$\left(\sum_{i\in\mathcal{J}_{n^\dagger}} \mathbf{g}_i\mathbf{g}_i^H\right) U_{n-n^\dagger}^H = 0 \Rightarrow \Re\left(\{\mathbf{g}_{i\in\mathcal{J}_{n^\dagger}}\}\right) \perp \Re\left(U_{n-n^\dagger}^H\right).$$

But $\Re\left(U_{n-n^\dagger}^H\right) \perp \Re\left(U_{n^\dagger}^H\right)$ since $\left[U_{n^\dagger}^H \ \ U_{n-n^\dagger}^H\right]$ is unitary so that the first equation in (2.6) holds, from which the second follows, as then does (2.7). □

Consider the related ϵ-dependent square Hermitian matrix

$$A(\epsilon) = \epsilon I_n + \sum_{i\in\mathcal{J}_{n^\dagger}} \mathbf{g}_i\mathbf{g}_i^H, \tag{2.8}$$

where $\epsilon \geq 0$. Then

$$A(\epsilon) = \begin{bmatrix} U_{n^\dagger} \\ U_{n-n^\dagger} \end{bmatrix} \left[\epsilon I_n + diag\left(\lambda_1, \lambda_2, \ldots, \lambda_{n^\dagger}, 0, \ldots, 0\right)\right] \begin{bmatrix} U_{n^\dagger} \\ U_{n-n^\dagger} \end{bmatrix}^H. \tag{2.9}$$

LEMMA 2.3. *With generic* $\mathbf{G}$*, as in Condition 1 of Theorem 2.1 and* $n^\dagger < n$*. Then*

$$\limsup_{\epsilon\downarrow 0} \mathbf{g}_j^H A(\epsilon)^{-1} \mathbf{g}_j\Big|_{j\in\mathcal{J}_{n^\dagger}} < \infty,\ \ \liminf_{\epsilon\downarrow 0} \mathbf{g}_j^H A(\epsilon)^{-1} \mathbf{g}_j\Big|_{j\notin\mathcal{J}_{n^\dagger}} = \infty. \tag{2.10}$$

Proof. For generic $\mathbf{G}$, (2.5) holds, as then does (2.9), so that

$$\mathbf{g}_j^H A(\epsilon)^{-1} \mathbf{g}_j = \mathbf{g}_j^H \begin{bmatrix} U_{n^\dagger} \\ U_{n-n^\dagger} \end{bmatrix}$$
$$diag\left[(\lambda_1+\epsilon)^{-1}, (\lambda_2+\epsilon)^{-1}, \ldots, (\lambda_{n^\dagger}+\epsilon)^{-1}, \epsilon^{-1}, \ldots, \epsilon^{-1}\right] \begin{bmatrix} U_{n^\dagger} \\ U_{n-n^\dagger} \end{bmatrix}^H \mathbf{g}_j.$$

The desired result follows from application of the Lemma 2.3 result (2.7). □

3. Optimization on Manifold. Our proof of Theorem 2.1 is by construction using an optimization approach. Let us here seek a solution of (2.1) by minimizing the cost function defined as

$$\Phi(\mathbf{P}) := \frac{1}{2}\sum_{j=1}^{m} \frac{1}{\gamma_j} r_j^2, \tag{3.1}$$

A number of our results focus on the two sets

$$J_{min} := \{j : r_j \geq r_i \ \forall \ i\}, \ J^{max} := \{j : r_j \leq r_i \ \forall \ i\}. \tag{3.2}$$

The cost function is defined only over the open smooth manifold

$$M := \{\mathbf{P}|\mathbf{P} > 0, \mathbf{PG}(\mathbf{G}^H\mathbf{PG})^{-1} exists\} \tag{3.3}$$

and is subject to the constraint $\sum_{j=1}^{m} r_j = 0$. Actually this later constraint is satisfied for all $\mathbf{P} \in M$.

Notice that if the cost function is zero on the manifold, then $\mathbf{R} = 0$, and the associated value of $\mathbf{P}$, denoted $\mathbf{P}^*$, is a solution of (2.1), that is $\mathbf{R}(\mathbf{P}^*) = 0$.

The introduction of the γ_j terms is to assist in establishing the following:

1. A global minimum of the cost function belongs to M.
2. There are no local minimum or other critical points in M, other than the global minimum.
3. There are no limiting solutions of (2.1) or limiting critical points of $\Phi(\mathbf{P})$ on the boundary of M, denoted δM, and defined in terms of a matrix condition number $CN(\cdot)$, as

$$\delta M := \left\{\mathbf{P}|\mathbf{P} \geq 0, CN\left(\mathbf{G}^H\mathbf{PG}\right) = \infty\right\} \tag{3.4}$$

Our first task is to show that under the conditions of the theorem, the only possible critical points of $\Phi(\mathbf{P})$, denoted $\mathbf{P_0}$, belonging to M are the global minimum of $\Phi(\mathbf{P})$, which would occur at $\mathbf{P}^*$, a solution of (2.1), that is $\mathbf{R}(\mathbf{P}^*) = 0$.

3.1. Cost Function Gradients. The gradients of $\Phi(\mathbf{P})$ are given from

$$\left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{\gamma=0} = r_j \frac{\partial r_j}{\partial p_j} + \sum_{i \neq j} r_i \frac{\partial r_i}{\partial p_j}. \tag{3.5}$$

Noting that

$$r_j = p_j \mathbf{y}_{j,j} - x_j, \ \mathbf{y}_{i,j} := \mathbf{g}_i^H \left(\mathbf{G}^H\mathbf{PG}\right)^{-1} \mathbf{g}_j,$$

and

$$\frac{\partial\left(\mathbf{G}^H\mathbf{PG}\right)}{\partial p_j} = \mathbf{g}_j \mathbf{g}_j^H, \tag{3.6}$$

then

$$\frac{\partial r_i}{\partial p_j} = \begin{cases} \mathbf{y}_{j,j}\left(1 - p_j \mathbf{y}_{j,j}\right) & for \ i = j \\ -p_i \mathbf{y}_{i,j} \mathbf{y}_{j,i} \leq 0 & for \ i \neq j \end{cases}. \tag{3.7}$$

Substituting for $\frac{\partial r_j}{\partial p_j}$ and $\frac{\partial r_i}{\partial p_j}$ in the second part of (3.5) gives

$$\left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{\gamma=0} = r_j \mathbf{y}_{j,j}\left(1 - p_j \mathbf{y}_{j,j}\right) - \sum_{i \neq j} r_i p_i \mathbf{y}_{i,j}\mathbf{y}_{j,i} = r_j \mathbf{y}_{j,j} - \sum_{i=1}^{m} r_i p_i \mathbf{y}_{i,j}\mathbf{y}_{j,i}. \tag{3.8}$$

It follows that

$$\left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{\gamma=0} = r_j \mathbf{y}_{j,j} - \mathbf{g}_j^H \left(\mathbf{G}^H\mathbf{P}\mathbf{G}\right)^{-1} \mathbf{G}^H\mathbf{P}\mathbf{R}\mathbf{G}\left(\mathbf{G}^H\mathbf{P}\mathbf{G}\right)^{-1}\mathbf{g}_j, = \mathbf{g}_j^H \left(\mathbf{G}^H\mathbf{P}\mathbf{G}\right)^{-1} \mathbf{G}^H\mathbf{P}(r_j I - \mathbf{R})\mathbf{G}\left(\mathbf{G}^H\mathbf{P}\mathbf{G}\right)^{-1}\mathbf{g}_j \tag{3.9}$$

LEMMA 3.1. *Consider the nonempty sets* J^{max}, J_{min} *of* (3.2), *then*

$$\left(r_j I - R\right)|_{j \in J^{max}} \geq 0, \quad \left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{j \in J^{max}} \geq 0. \tag{3.10}$$

$$\left(r_j I - R\right)|_{j \in J_{min}} \leq 0, \quad \left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{j \in J_{min}} \leq 0. \tag{3.11}$$

Moreover, for $\mathbf{R} \neq 0$, *then there is some* r_i *not in* J^{max}, *and some* r_i *not in* J_{min}, *and also*

$$\left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{j \in J^{max}} > 0, \quad \left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{j \in J_{min}} < 0. \tag{3.12}$$

Proof. The first inequalities involving r_j follow from the definitions of J^{max}, J_{min}. The first set of inequalities on the gradients follow immediately from substituting the first inequalities into (3.9).

Now recall that the sum of the r_j are zero, so that in the case $\mathbf{R} \neq 0$, at least for one pair i,j, then $r_j \neq r_i$. In an expansion of the right hand side of (3.9), involving terms $(r_j - r_i)$, then at least one of the summands would be nonzero, so that the inequalities can be strengthened to strict inequalities, as claimed. □

3.2. Cost Function Properties. Any critical points $\mathbf{P_0}$ of $\Phi(\mathbf{P})$ within M, are characterized by

$$\left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{\mathbf{P}=\mathbf{P}} = 0 \ \forall \ j. \tag{3.13}$$

Observe that the constraint $\sum_{j=1}^{m} r_j = 0$ is satisfied for all $\mathbf{P} \in M$, so there is no need for any projection onto this constraint. Useful equivalent conditions to (3.13) are given in the following lemma.

LEMMA 3.2. *Consider the cost function* (3.1), *under the constraint* $\sum_{j=1}^{m} r_j = 0$. *The only critical points of the cost function* $\Phi(\mathbf{P})$ *within* M, *satisfying* (3.13), *if any* *exist, are the global minimum of the cost function, characterized by* $\mathbf{R}(\mathbf{P}^*) = 0$*, being the solution of* (2.1).

Moreover, for all elements in the nonempty sets J^{max}, J_{min} *of* (3.2)*, a necessary and sufficient condition for* $\mathbf{P} \in M$ *to achieve a global minimum for* $\Phi(\mathbf{P})$ *are that*

$$\left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{j \in J^{max}} = 0, \text{ or } \left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{j \in J_{min}} = 0. \tag{3.14}$$

Furthermore, for $\mathbf{P} \neq \mathbf{P}^*$*, then of necessity the strict inequalities* (3.12) *hold.*

Proof. Necessity of (3.14), follows by implication from (3.13). For sufficiency, first note that for $j \in J^{max}$ defined in (3.2), and applying the constraint $\sum_{j=1}^{m} r_j = 0$, then from (3.9)

$$\begin{aligned}
\left.\frac{\partial \Phi(\mathbf{P})}{\partial p_j}\right|_{j \in J^{max}} = 0 &\Rightarrow \left. r_j \right|_{j \in J^{max}} = 0 \\
&\Rightarrow r_i = 0 \;\forall\; i \text{ (under the constraint } \sum_{j=1}^{m} r_j = 0) \\
&\Rightarrow \Phi(\mathbf{P}) = 0 \text{ (from (3.1))} \\
&\Rightarrow \frac{\partial \Phi(\mathbf{P})}{\partial p_i} = 0 \;\forall\; i \text{ (from(3.9))}
\end{aligned} \tag{3.15}$$

Likewise, working with J_{min}, and the first parts of the lemma are established. The last part follows from application of the second part result to the Lemma 3.1. □

From the results so far, we see the following: In seeking to minimize $\Phi(\mathbf{P})$ on M, via a gradient flow, for example, if the gradient flow was always repelled by the boundary, it would converge to a critical point in M. By the above lemma, this critical point $\mathbf{P}^*$ would be the desired solution of (2.1).

Our next step is to further study the cost function on the boundary, and to exclude the case that a solution of (2.1) exists as a limiting solution approaching the boundary δM.

LEMMA 3.3. *Consider the cost function* $\Phi(\mathbf{P})$ *of* (3.1)*, and that the Conditions of Theorem 2.1 apply. Consider also any gradient flow reducing the cost function assumed to converge to a limiting point* $\mathbf{P_1}$ *on the boundary of* M*, namely* δM *Then* $\mathbf{R}(\mathbf{P_1}) \neq 0$*, and* $\mathbf{P_1}$ *is not a solution of* (2.1).

Proof.

First consider the re-organization

$$\begin{aligned}
\mathbf{G}^H \mathbf{P} \mathbf{G} &= \sum_{j=1}^{m} \mathbf{g}_j p_j \mathbf{g}_j^H \\
&= \mathbf{g}_j p_j \mathbf{g}_j^H + \sum_{i=1, i \neq j}^{m} \mathbf{g}_i p_i \mathbf{g}_i^H \\
&= \mathbf{g}_j p_j \mathbf{g}_j^H + \bar{\mathbf{G}}_j^H \bar{\mathbf{P}}_j \bar{\mathbf{G}}_j,
\end{aligned}$$

with obvious definitions of $\bar{\mathbf{P}}, \bar{\mathbf{G}}_j$, (not functions of $p_j, \mathbf{g}_j$). Substitution into (3.6) yields

$$r_j = p_j \mathbf{g}_j^H \left(\mathbf{g}_j p_j \mathbf{g}_j^H + \bar{\mathbf{G}}_j^H \bar{\mathbf{P}}_j \bar{\mathbf{G}}_j\right)^{-1} \mathbf{g}_j - x_j$$
$$= \mathbf{g}_j^H \left(\mathbf{g}_j \mathbf{g}_j^H + \bar{\mathbf{G}}_j^H \frac{\bar{\mathbf{P}}_j}{p_j} \bar{\mathbf{G}}_j\right)^{-1} \mathbf{g}_j - x_j,$$

and applying the Matrix Inversion Lemma, then

$$r_j = 1 - \left[1 + \mathbf{g}_j^H \left(\bar{\mathbf{G}}_j^H \frac{\bar{\mathbf{P}}_j}{p_j} \bar{\mathbf{G}}_j\right)^{-1} \mathbf{g}_j\right]^{-1} - x_j. \tag{3.16}$$

Now consider a sequence of $\mathbf{P}$, as $\{\mathbf{P}(k) : k = 1, 2, \ldots, \infty\}$, which reduces $\Phi(\mathbf{P}(k))$ as k increases, such as can be extracted from gradient flow equations. Consider also that $p_j(k) \to \infty$ as $k \to \infty$. Then, we wish to study the limiting behavior, as $k \to \infty$, of

$$r_j(k) = 1 - \left[1 + \mathbf{g}_j^H \left(\bar{\mathbf{G}}_j^H \frac{\bar{\mathbf{P}}_j(k)}{p_j(k)} \bar{\mathbf{G}}_j\right)^{-1} \mathbf{g}_j\right]^{-1} - x_j. \tag{3.17}$$

Let us assume that there is a subset of n_∞ indices j, denoted J_∞, such that $p_j(k) \to \infty$ at the same maximum rate as $k \to \infty$. Thus for some $c_{min}, c_{max} > 0$,

$$J_\infty := \left\{ j | \lim_{k\to\infty} \frac{p_i(k)}{p_j(k)}\bigg|_{i \notin J_\infty, j \in J_\infty} = 0,\ c_{max} \geq \lim_{k\to\infty} \frac{p_i(k)}{p_j(k)}\bigg|_{i,j \in J_\infty} \geq c_{min} \right\}. \tag{3.18}$$

Let us also consider a dual subset of n_0 indices j, denoted J_0, such that $p_j(k) \to 0$ at the same maximum rate as $k \to \infty$. Thus, again for some $c_{min}, c_{max} > 0$,

$$J_0 := \left\{ j | \lim_{k\to\infty} \frac{p_i(k)}{p_j(k)}\bigg|_{j \notin J_0, i \in J_0} = 0,\ c_{max} \geq \lim_{k\to\infty} \frac{p_i(k)}{p_j(k)}\bigg|_{i,j \in J_0} \geq c_{min} \right\}. \tag{3.19}$$

The dual results are useful since the solution of (2.1) is invariant of the scaling of $\mathbf{P}$, so that with a change of scaling, the role of zero and infinity can be interchanged.

We consider limiting $r_j(k)$ of (3.17), for two cases of n_∞ or n_0.

Case 1: $0 < n_\infty \leq n$, or $m > n_0 \geq m - n$.. Consider first an application of (3.18) as

$$\liminf_{k\to\infty} \left(\bar{\mathbf{G}}_j^H \frac{\bar{\mathbf{P}}_j(k)}{p_j(k)} \bar{\mathbf{G}}_j\right)\bigg|_{j \in J_\infty} = \liminf_{k\to\infty} \sum_{i \neq j} \mathbf{g}_i \frac{p_i(k)}{p_j(k)} \mathbf{g}_i^H \bigg|_{j \in J_\infty}$$
$$\geq c_{min} \sum_{i \neq j} \mathbf{g}_i \mathbf{g}_i^H \bigg|_{i,j \in J_\infty}. \tag{3.20}$$

for some constant $c_{min} > 0$. Clearly, with generic $\mathbf{G}$, as in Condition 1 of Theorem 2.1, and $0 < n_\infty \leq n$, the Property (2.4) tells us that the right hand side of the above inequality has rank $(n_\infty - 1) < n$. Also, the conditions of Lemma 2.2 and Lemma 2.3 are met. Thus let us identifying $n^* < n$ with $n_\infty - 1 < n$, and $\lim_{\epsilon\to 0} A(\epsilon)$ as an upper bound on the left hand side of the above inequality. Then application of the right hand equation of (2.10) yields

$$\lim_{k\to\infty} \mathbf{g}_j^H \left(\bar{\mathbf{G}}_j^H \frac{\bar{\mathbf{P}}_j(k)}{p_j(k)} \bar{\mathbf{G}}_j\right)^{-1} \mathbf{g}_j \bigg|_{j\in J_\infty} = \infty.$$

Taking limits in (3.17), and including the corresponding dual situation result without giving details, then

$$\lim_{k\to\infty} r_j(k) = (1 - x_j) > 0 \ \forall\ j \in J_\infty|_{0<n_\infty\le n},\ or\ j \notin J_0|_{m>n_0\ge m-n}. \tag{3.21}$$

where the inequality follows from the Condition 2 of Theorem 2.1. Equivalently, there is asymptotic ill-conditioning of $\mathbf{PG}\left(\mathbf{G}^H\mathbf{PG}\right)^{-1}$, and $\mathbf{P}(k)$ converges to the manifold boundary δM as follows:

$$\lim_{k\to\infty} \mathbf{P}(k)\,|_{j\in J_\infty|_{0<n_\infty\le n},\ or\ j\in J_0|_{m>n_0\ge m-n}} \in \delta M. \tag{3.22}$$

The nonzero limits on $r_j(k)$ established for $0 < n_\infty \le n$ and for $n_0 \ge m - n$ in (3.21), give the result that $R(\mathbf{P_1}) \ne 0$, or equivalently, the solution of (2.1) does not exist as a limit on the boundary of M, δM, and the lemma is established. □

For completeness we add the further limiting results.

Case 2: $m > n_\infty \ge n$, or $0 < n_0 \le m - n$. Consider the inequality, which follows by neglecting terms involving $p_i(k)|_{i\notin J_\infty}$, as

$$\left(\bar{\mathbf{G}}_j^H \frac{\bar{\mathbf{P}}_j(k)}{p_j(k)} \bar{\mathbf{G}}_j\right)\bigg|_{j\notin J_\infty} \ge \sum_{i\in J_\infty} \mathbf{g}_i \frac{p_i(k)}{p_j(k)} \mathbf{g}_i^H \bigg|_{j\notin J_\infty}.$$

From the generic assumption property (2.4), recall that $m > n_\infty \ge n$, and thus both terms above are full rank. Now taking limits in the above inequality, then application of the first of the limits (3.18) in inverse form, leads to the limit

$$\lim_{k\to\infty} \left(\bar{\mathbf{G}}_j^H \frac{\bar{\mathbf{P}}_j(k)}{p_j(k)} \bar{\mathbf{G}}_j\right)\bigg|_{j\notin J_\infty} \ge \lim_{k\to\infty} \sum_{i\in J_1} \mathbf{g}_i \frac{p_i(k)}{p_j(k)} \mathbf{g}_i^H \bigg|_{j\notin J_\infty} = \infty I. \tag{3.23}$$

It follows that

$$\lim_{k\to\infty} \mathbf{g}_j^H \left(\bar{\mathbf{G}}_j^H \frac{\bar{\mathbf{P}}_j(k)}{p_j(k)} \bar{\mathbf{G}}_j\right)^{-1} \mathbf{g}_j \bigg|_{j\notin J_\infty} = 0.$$

Application in (3.17) gives the following, which includes the corresponding dual result,

$$\lim_{k\to\infty} r_j(k) = -x_j < 0 \ \forall\ j \notin J_\infty|_{m>n_\infty\ge n},\ or\ j \in J_0|_{0<n_0\le m-n}. \tag{3.24}$$

where the inequality follows from the Condition 2 of Theorem 2.1.

REMARK 3.1. *Clearly, generic* $\mathbf{G}$*, as in Condition 1 of Theorem 2.1, is sufficient for Lemma 3.3 to hold. However, if* $\mathbf{g}_j|_{j\in J_\infty}$ *of* (3.21) *is linearly independent of the* $\mathbf{g}_i|_{j\in J_\infty}$*, then Lemma 3.3 will still hold. Thus generic* $\mathbf{G}$*, as in Condition 2.1 of Theorem 2.1, is not a necessary condition.*

However, in the case of non-generic $\mathbf{G}$ *since we can not guarantee that* $\mathbf{g}_j|_{j\in J_\infty}$ *of* (3.21) *is linearly independent of the* $\mathbf{g}_j|_{j\in J_\infty}$*. Then to establish our existence result we must restrict the selection of* $\mathbf{G}$ *to generic* $\mathbf{G}$ *as in Condition 1 of Theorem 2.1.*

The above derivations are summarized in the first part of the proof of the following lemma.

LEMMA 3.4. *Consider the cost function* $\Phi(\mathbf{P})$ *of* (3.1)*, on the manifold* M *with boundary* δM*, and that the Conditions of Theorem 2.1 apply.*

Then $\mathbf{GP}(\mathbf{G}^H\mathbf{PG})^{-1}$ *becomes ill-conditioned asymptotically, and* $\mathbf{P}\to\delta M$*, and* (3.21), (3.22) *apply, if and only if, in a minimization of the cost function, the sets* J_∞, J_0 *of* (3.18), (3.19)*, have* $0 < n_\infty \le n$*, or* $n_0 \ge m-n$ *elements, respectively. .*

Moreover, a lower bound on the cost function is $\frac{1}{2}(1-x_{max})^2$*, where* x_{max} *is the maximum of the set of* x_j*.*

Proof. The first part of the proof follows from the derivation of (3.21), (3.22). The second part follows since at the boundary δM, then (3.21) holds for at least one value of j, so we take the minimum r_j value to give a minimum cost as in the lemma. □

We proceed by first showing that any gradient flow reducing the cost function and converging to a limiting point $\mathbf{P_1}$ on the boundary of M, namely δM, does not achieve a limiting solution $\mathbf{P}^{**}$ of (2.1).

3.3. Gradient Flows. et us first consider two "downhill" gradient flows for $\Phi(\mathbf{P}) > 0$, or equivalently $\mathbf{R}\neq 0$, on the manifold M, with boundary δM. In particular,

$$\dot{p}_j = -\gamma_j \frac{\partial\Phi(\mathbf{P})}{\partial p_j}\bigg|_{j\in J^{max}} < 0, \text{ and } \dot{p}_j = -\gamma_j \frac{\partial\Phi(\mathbf{P})}{\partial p_j}\bigg|_{j\in J_{min}} > 0. \tag{3.25}$$

Here the sets J^{max}, J_{min} are defined in (3.2), and consist at time t of the j such that, respectively, $r_j(t)\ge r_i(t)\ \forall\ i\ne j$, and $r_j(t)\le r_i(t)\ \forall\ i\ne j$, and the inequality properties of Lemma 3.1 are made explicit. The gain factors γ_j are chosen to be positive, but so as to upper bound the rate of change of p_j or p_j^{-1} to less than some exponential rate, preferably differing for each j to avoid any tendency towards ill conditioning .

LEMMA 3.5. *Consider either of the two flows, or both acting together, of* (3.25) *associated with arbitrary positive definite initial conditions* $\mathbf{P}>0$*, and that the conditions of Theorem 2.1 apply.*

Then sufficiently close to the boundary δM*, there is no critical point* $\mathbf{P_0}$ *for either of the flows.*

Moreover for both flows, there is no $p_j|j\in J^{max}$ *which approaches infinity, or* $p_j|_{j\in J_{min}}$ *which approaches, and there is no convergence of* $\mathbf{P}$ *to the boundary* δM*.*

Also, for after some sufficiently large time, the cost $\Phi(\mathbf{P})$ *becomes smaller than any minimum finite cost in the neighborhood of the boundary* δM*: Such a bound in given in Lemma 3.4.*

Furthermore, each of the flows converges to a critical point in M, being the solution of $\mathbf{R}(\mathbf{P}^*) = 0$*. Equivalently* $\mathbf{P}^*$ *is a solution of* (2.1)*, unique to within a scaling factor* β*. That is, the existence and uniqueness result of Theorem 2.1 holds.*

Proof. The first result follows immediately from Lemma 3.3. The second results, in part, follow by the inequalities in the gradient flows, being properties of Lemma 3.1, preventing p_j increasing for the first flow and decreasing for the second flow.

Even should some set of p_j approach zero in the first flow, or infinity in the second flow, then the gain selections γ_j ensure that there is a limited rate of exponential growth of p_j, p_j^{-1} for $j \in J^{max}, J_{min}$, respectively. Consequently, there is no finite escape time for $p_j|_{j \in J^{max}}$ or $p_j^{-1}|_{j \in J_{min}}$. Eventually the cost function is lower than any prior threshold, and in particular the lower bound of the cost function on the boundary δM.

Subsequently, there is no attraction to the boundary for either of the two flows, and both flows converge to the same critical point within M. Application of Lemma 3.2 completes the proof. □

REMARK 3.2. *When the γ_j are selected to achieve different exponential rate bounds for different j, then this imposes the constraints $n_\infty, n_0 \leq 1$. Thus asymptotic ill-conditioning can occur only for $n_\infty = 1 \leq n$, or $n_0 = 1 \geq m - n$. Actually, the constraints $n \geq 1, m - n \geq 1$, are assumed to be satisfied here, otherwise* (2.1) *has trivial solutions.*

Considering the first flow where we have established that there are no p_j approaching infinity, let us assume that $n_0 = 1 \geq m - n$, then from (3.24)*, for the one $j \in J_0$, the corresponding $r_j(k)$ is asymptotically negative, so then asymptotically $j \notin J^{max}$. There is a contradiction. Consequently, we see that the first flow converges to a critical point in M, without asymptotic ill-conditioning, that is, without going to the boundary δM, as claimed.*

Likewise, considering the second flow where we have established that there are no p_j approaching zero, let us assume $n_\infty = 1$, then from (3.21)*, the corresponding r_j approaches a positive value, giving a contradiction, so ensuring convergence to a critical point in M, as claimed.*

THEOREM 3.6. *Consider that the conditions of Theorem 2.1 apply. Consider also a gradient flow for $\Phi(\mathbf{P}) > 0$ or equivalently $\mathbf{R} \neq 0$, on the manifold M, with boundary δM. In particular, denoting $\bar{p} := [p_1\ p_2\ \ldots p_{m-1}]$ and $\nabla\Phi(\mathbf{P}) := \frac{\partial \Phi(\mathbf{P})}{\partial \bar{\mathbf{p}}}$, then the flow*

$$\dot{\bar{p}} = -Q\nabla\Phi(\mathbf{P}) \tag{3.26}$$

for arbitrary $Q > 0$. Then this flow will converge to a critical point $\mathbf{P}^* \in M$ satisfying $\nabla\Phi(\mathbf{P}^*) = 0$, and $\mathbf{R}(\mathbf{P}^*) = 0$. Equivalently $\mathbf{P}^*$ is a solution of (2.1)*, and the existence result of Theorem 2.1 holds.*

Proof. The gradient flow (3.26) on M, avoids the boundary δM, since as established in the above lemma, there is a "downhill" path from arbitrary initial conditions which is repelled by the boundary δM. The flow then converges to a critical point satisfying $\nabla\Phi(\mathbf{P}) = 0$, which by Lemma 3.2 is a global minimum and yields $\mathbf{P}^*$ the solution of $\mathbf{R}(\mathbf{P}^*) = 0$, and of (2.1). □

The existence and uniqueness result of Theorem 2.1 are now established by construction.

4. Construction Results. Although the existence result of the previous section is derived by a construction procedure, this is by no means the best method for construction. In this section, given the existence result of the previous section we develop efficient construction methods.

We first focus on two special cases for which explicit solutions can be calculated before proceeding to a general algorithm.

4.1. Special Case Solution: $n = 1$. Consider that $n = 1$, and rewrite (2.1) as

$$\left[\mathbf{P}^H \mathbf{g} \left(\mathbf{g}^H \mathbf{P} \mathbf{g}\right)^{-1} \mathbf{g}^H\right]_{diag} = \mathbf{X},$$

or equivalently

$$\frac{p_j g_j^H g_j}{\sum_{l=1}^m p_l g_l^H g_l} = x_j, \; p_j = \frac{x_j}{g_j^H g_j} \sum_{l=1}^m p_l g_l^H g_l,$$

where $j, l = 1, 2, \ldots m$. Let $\beta = \sum_{l=1}^m p_l g_l^H g_l$, then

$$\mathbf{P} = \beta \times diag\left\{\frac{x_1}{g_1^H g_1}, \frac{x_2}{g_2^H g_2}, \ldots, \frac{x_m}{g_m^H g_m}\right\}, \tag{4.1}$$

for arbitrary $\beta \neq 0$, so that without loss of generality set $\beta = 1$. We summarize the above result as the following theorem

THEOREM 4.1. *Given that the conditions of Theorem 2.1 apply, then an explicit solution* $\mathbf{P}^*$ *of* (2.1), *for the case* $n = 1$ *is given by* (4.1).

4.2. Special Case Solution: $m = n+1$. Consider that $m = n+1$, and rewrite (2.1) as

$$\left[\mathbf{W}^H \mathbf{G} \left(\mathbf{G}^H \mathbf{W} \mathbf{W}^H \mathbf{G}\right)^{-1} \mathbf{G}^H \mathbf{W}\right]_{diag} = \mathbf{X}, \tag{4.2}$$

where $\mathbf{W}\mathbf{W}^H = \mathbf{P}$. Let us denote the square augmented matrix $\tilde{\mathbf{G}} = [\mathbf{G} \;\; \mathbf{g}]$, where $\tilde{\mathbf{G}}$ is full rank, and $\mathbf{W}^H \mathbf{g}$ is orthogonal to all columns of $\mathbf{W}^H \mathbf{G}$. Then

$$\begin{aligned}\tilde{\mathbf{G}}^H \mathbf{W} \mathbf{W}^H \tilde{\mathbf{G}} &= \begin{bmatrix} \mathbf{G}^H \\ \mathbf{g}^H \end{bmatrix} \mathbf{W} \mathbf{W}^H \begin{bmatrix} \mathbf{G}^H & \mathbf{g}^H \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{G}^H \mathbf{W} \mathbf{W}^H \mathbf{G} & 0 \\ 0 & \mathbf{g}^H \mathbf{W} \mathbf{W}^H \mathbf{g} \end{bmatrix},\end{aligned}$$

and

$$\left(\tilde{\mathbf{G}}^H \mathbf{W} \mathbf{W}^H \tilde{\mathbf{G}}\right)^{-1} = \begin{bmatrix} \left(\mathbf{G}^H \mathbf{W} \mathbf{W}^H \mathbf{G}\right)^{-1} & 0 \\ 0 & \left(\mathbf{g}^H \mathbf{W} \mathbf{W}^H \mathbf{g}\right)^{-1} \end{bmatrix}.$$

Then from the full rank property of $W^H \tilde{\mathbf{G}}$,

$$
\begin{aligned}
I_{n+1} &= \left[\mathbf{W}^H\tilde{\mathbf{G}}\left(\tilde{\mathbf{G}}^H\mathbf{W}\mathbf{W}^H\tilde{\mathbf{G}}\right)^{-1}\tilde{\mathbf{G}}^H\mathbf{W}\right]_{diag} \\
&= \left\{\mathbf{W}^H\begin{bmatrix}\mathbf{G} & \mathbf{g}\end{bmatrix}\begin{bmatrix}\left(\mathbf{G}^H\mathbf{W}\mathbf{W}^H\mathbf{G}\right)^{-1} & 0 \\ 0 & \left(\mathbf{g}^H\mathbf{W}\mathbf{W}^H\mathbf{g}\right)^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{G}^H \\ \mathbf{g}^H\end{bmatrix}\mathbf{W}\right\}_{diag} \\
&= \left[\mathbf{W}^H\mathbf{G}\left(\mathbf{G}^H\mathbf{W}\mathbf{W}^H\mathbf{G}\right)^{-1}\mathbf{G}^H\mathbf{W}\right]_{diag} + \left[\mathbf{W}^H\mathbf{g}\left(\mathbf{g}^H\mathbf{W}\mathbf{W}^H\mathbf{g}\right)^{-1}\mathbf{g}^H\mathbf{W}\right]_{diag}.
\end{aligned}
\tag{4.3}
$$

Now (4.2) can be written as

$$
\left[\mathbf{W}^H\mathbf{g}\left(\mathbf{g}^H\mathbf{W}\mathbf{W}^H\mathbf{g}\right)^{-1}\mathbf{g}^H\mathbf{W}\right]_{diag} = I_{n+1} - \mathbf{X},
$$

or equivalently,

$$
\frac{p_j g_j^H g_j}{\sum_{l=1}^{n+1} p_l g_l^H g_l} = 1 - x_j,\ p_j = \frac{1 - x_j}{g_j^H g_j}\sum_{l=1}^{n+1} p_l g_l^H g_l,
$$

for $j, l = 1, 2, \ldots, n+1$. Let us denote $\tilde{\mathbf{g}}$, the null vector of $\mathbf{G}^H$, then

$$
\mathbf{G}^H\tilde{\mathbf{g}} = 0,\ \mathbf{G}^H\mathbf{W}\mathbf{W}^H\mathbf{P}^{-1}\tilde{\mathbf{g}} = 0,\ \mathbf{g} = \mathbf{P}^{-1}\tilde{\mathbf{g}}. \tag{4.4}
$$

It follows that

$$
p_j = \beta\frac{(1 - x_j)p_j^2}{\tilde{g}_j^H\tilde{g}_j}, \tag{4.5}
$$

where $\beta = \sum_{l=1}^{n+1} p_l g_l^H g_l$. Then

$$
\mathbf{P} = \beta \times diag\left\{\frac{\tilde{g}_1^H\tilde{g}_1}{1 - x_1}, \frac{\tilde{g}_2^H\tilde{g}_2}{1 - x_2}, \ldots, \frac{\tilde{g}_m^H\tilde{g}_m}{1 - x_m}\right\}, \tag{4.6}
$$

for arbitrary $\beta \neq 0$, so that without loss of generality set $\beta = 1$. We summarize the above results as the following theorem

THEOREM 4.2. *Given that the conditions of Theorem 2.1 apply, then an explicit solution* $\mathbf{P}^*$ *of* (2.1), *for the case* $m = n + 1$ *is given by* (4.6).

4.3. General Solution. To achieve a general solution of (2.1), our approach is to design a practical gradient descent algorithm to minimize the cost function $\Phi(\mathbf{P})$ of (3.1) to achieve a critical point $\mathbf{P}^*$ satisfying $\nabla\Phi(\mathbf{P}^*)\,|_{\mathbf{P}=\mathbf{P}^*} = 0$, where ∇ denotes the gradient. A first approach to achieve a practical algorithm is to work with a gradient algorithm of the form

$$
\mathbf{p}(k+1) = \mathbf{p}(k) - \alpha(k)\,Q(k)\,\nabla\Phi(\mathbf{P}(k)), \tag{4.7}
$$

where

$$\nabla\Phi(\mathbf{P}) := \frac{\partial\Phi(\mathbf{P})}{\partial\mathbf{p}} = S\mathbf{r}. \tag{4.8}$$

Here $\mathbf{p} = [p_1, p_2, \ldots, p_m]^T$, $\mathbf{r} = [r_1, r_2, \ldots, r_m]^T$, and from (3.7) $S := (S)_{ij}\Big| S_{ij} = \frac{\partial r_j}{\partial p_i}$ for $i, j = 1, 2, \ldots, m$.

A suitable selection of $Q(k)$ would be some form of positive definite matrix to ensure a "downhill" direction, yet being an approximation to the inverse of the Hessian matrix to achieve quadratic convergence rates. A suitable selection of $\alpha(k)$ could be the result of a line search so that $\Phi(\mathbf{P}(k+1))$ is minimized yet maintains $\mathbf{G}^H\mathbf{P}\mathbf{G} > 0$. A maximum step size would be when $|\mathbf{G}^H\mathbf{P}\mathbf{G}| = 0$, which is given as the minimum real positive root of a polynomial equation.

The Hessian matrix at the critical point when $\mathbf{R} = 0$ is derived as follows. The ij-th element of $\mathcal{H}_\Phi(\mathbf{P})|_{\mathbf{R}=0}$ is

$$\left.\frac{\partial^2\Phi(\mathbf{P})}{\partial p_i \partial p_j}\right|_{\mathbf{R}=0} = \left.\left[e_j^T\frac{\partial S}{\partial p_i}\mathbf{r} + e_j^T S\frac{\partial \mathbf{r}}{\partial p_i}\right]\right|_{\mathbf{R}=0} = e_j^T S S^H e_i,$$

where $e_i = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ with 1 in the i-th position, and

$$\mathcal{H}_\Phi(\mathbf{P})|_{\mathbf{R}=0} = SS^H \tag{4.9}$$

Now it is readily shown that S is singular since

$$S\underline{1} = 0,$$

where $\underline{1}$ is a vector with all entries unity. Thus $(SS^H)^{-1}$ does not exist. Indeed this property reflects the fact that the solutions $\mathbf{P}^*$ of (2.1), when these exist under the conditions of Theorem 2.1, are unique only up to a scaling factor.

The approach we propose here to side-step the singularity problem in the manifold M and achieve a unique $\mathbf{P}^*$ is to work with a cost function related to $\phi(\mathbf{P})$ as

$$\bar{\phi}(\mathbf{P}) := \frac{1}{2}\sum_{i=1}^{m-1} r_i^2. \tag{4.10}$$

Observe that (4.10) leads to

$$\nabla\bar{\Phi}(\mathbf{P}) := \frac{\partial\bar{\Phi}(\mathbf{P})}{\partial\bar{\mathbf{p}}} = \bar{S}\bar{\mathbf{r}}, \tag{4.11}$$

where $\bar{\mathbf{p}}_i = \mathbf{p}_i$, $\bar{\mathbf{r}}_i = \mathbf{r}_i$, $\bar{S}_{ij} = S_{ij}$ for $i, j = 1, 2, \ldots, m-1$. Recalling the constraint $\sum_{i=1}^m r_i = 0$, it is immediate that $\bar{\mathbf{R}} = 0 \Leftrightarrow \mathbf{R} = 0$, and consequently,

$$\nabla\bar{\Phi}(\mathbf{P}^*) = 0 \Leftrightarrow \nabla\Phi(\mathbf{P}^*) = 0.$$

Under the conditions of Theorem 2.1, any critical point $\mathbf{P}^*$ of $\bar{\Phi}(\mathbf{P})$ is also a critical point of $\Phi(\mathbf{P})$, this being unique to within a scaling factor . Moreover, since the gradient is zero only when $\mathbf{R} = 0$, or equivalently $\mathbf{r} = 0$, then $\bar{S}$ is always full rank.

THEOREM 4.3. *Consider the gradient algorithm, assuming existence of the inverse,*

$$\begin{aligned} \bar{\mathbf{p}}(k+1) &= \bar{\mathbf{p}}(k) - \alpha(k)\left(\bar{S}(k)\bar{S}^H(k)\right)^{-1} \nabla\bar{\phi}(\mathbf{P}(k)) \\ &= \bar{\mathbf{p}}(k) - \alpha(k)\left(\bar{S}^H(k)\right)^{-1}\bar{\mathbf{r}}(k), \ \bar{\mathbf{p}}(0) > 0. \end{aligned} \tag{4.12}$$

Here, the step size $\alpha(k) \in (0,1]$ is chosen to preserve $\mathbf{G}^H\mathbf{P}\mathbf{G} > 0$ and to minimize (or merely reduce) $\bar{\Phi}(\mathbf{P}(k+1))$ in a line search.

Then under the conditions of Theorem 2.1:

i) The algorithm is well defined for all k, or equivalently the inverse exists for all k.

ii) The algorithm (4.12) *has a limiting solution*

$$\lim_{k\to\infty} \bar{\mathbf{p}}(k) = \bar{\mathbf{p}}^* \tag{4.13}$$

where $\mathbf{P}^* = diag\{\bar{\mathbf{p}}^*, p_m\}$ is the unique critical point of $\Phi(\mathbf{P})$. $\mathbf{P}^*$ is unique given by $\nabla\Phi(\mathbf{P}^*) = 0$, and satisfying $\Phi(\mathbf{P}^*) = 0$.

iii) The algorithm (4.12) *is equivalent to*

$$\bar{\mathbf{p}}(k+1) = \bar{\mathbf{p}}(k) - \alpha(k)\,\mathcal{H}_{\bar{\Phi}}^{-1}(\mathbf{P}(k))\big|_{\mathbf{R}=0}\nabla\bar{\Phi}(\mathbf{P}(k)). \tag{4.14}$$

Here $\mathcal{H}_{\bar{\Phi}}(\mathbf{P})$ is the Hessian matrix of $\bar{\Phi}(\mathbf{P})$ at $\mathbf{P}$. Also, as k approaches ∞, $\mathbf{P}(k)$ approaches $\mathbf{P}^*$. Also, $\mathcal{H}_{\bar{\Phi}}(\mathbf{P}(k))|_{\mathbf{R}=0}$ approaches asymptotically the Hessian $\mathcal{H}_{\bar{\Phi}}(\mathbf{P}^*)$ at the unique critical point $\mathbf{P}^*$.

iv) A line search on $\alpha(k)$ achieves the property

$$\lim_{k\to\infty} \alpha(k) = 1. \tag{4.15}$$

v) The algorithm (4.12) *with* (4.15) *holding is asymptotically a Newton algorithm, and converges quadratically.*

Proof. Consider the gradient (4.11). As noted above the theorem, S is full rank in M, so that $S(k)^{-1}$ exists for all k. Thus the second equality in (4.12) holds. Also, the ij-th element of $\mathcal{H}_{\bar{\Phi}}(\mathbf{P})|_{\mathbf{R}=0}$ is

$$\left.\frac{\partial^2\Phi(\bar{\mathbf{P}})}{\partial p_i \partial p_j}\right|_{\mathbf{R}=0} = e_j^T \frac{\partial\bar{S}}{\partial p_i}\bar{r}\Big|_{\mathbf{R}=0} + e_j^T\bar{S}\frac{\partial\bar{r}}{\partial p_i} = e_j^T\bar{S}\bar{S}^H e_i^T.$$

Thus

$$\mathcal{H}_{\bar{\Phi}}(\mathbf{P})|_{\mathbf{R}=0} = \bar{S}\bar{S}^H. \tag{4.16}$$

and the equivalence of (4.12) and (4.14) is established.

Now, since we are dealing with a gradient algorithm, clearly the selection of $\alpha(k)$ by a line search, achieves the property that $\bar{\Phi}(\mathbf{P}(k+1)) < \bar{\Phi}(\mathbf{P}(k))$, except at a critical point. This in turn, from the properties of a line search, is a continuous mapping. Applying standard results yields convergence to a critical point of $\bar{\Phi}(\mathbf{P})$. As established above, such a critical point is the desired solution of (2.1).

The remaining results of the theorem are standard for asymptotic Newton algorithms. □

REMARK 4.1.

If there exists a quadratic upper bound $\bar{\Phi}(\alpha)$, *of* $\bar{\Phi}(\mathbf{P})$ *then it would be possible to find an explicit solution for* $\alpha(k)$, *making the line search unnecessary. At this stage, we have no simple to calculate quadratic upper bound. We then require a line search for* $\alpha(k)$.

A line search iteration requires a computational effort which is linear in $m-1$. *This is a negligible cost compared to an iteration which requires the inverse of* $\bar{S}$ *is cubic in* $m-1$.

We know asymptotically the optimal step size is unity, so that in practice a step size of unity is first used and reduced only if this gives a reduced cost without ill conditioning. Any line search is for a step size over the range $(0,1]$.

We also note that it is sometimes useful to work with a simpler algorithm in the first few iterations, since there is no convergence rate advantage to a Newton type algorithm except asymptotically. The simpler algorithm we propose replaces (4.12) *with one which approximates* $\bar{S}$ *by a diagonal matrix consisting of the diagonal elements of* $\bar{S}$, *namely* $[\bar{S}]_{diag}$, *as follows.*

$$\bar{\mathbf{p}}(k+1) = \bar{\mathbf{p}}(k) - \alpha(k)\left([\bar{S}]_{diag}(k)\right)^{-1} F(k),\ \bar{\mathbf{p}}(0) > 0. \tag{4.17}$$

The computational effort for this iteration is now only quadratic in $m-1$.

On many randomly generated examples, our algorithms found $\mathbf{P}^*$ *to within the accuracy of the computer in seven iterations.*

4.3.1. Decentralized Algorithm. There is motivation in some applications to work with a decentralized algorithm as in (4.17). It turns out that asymptotically this algorithm is equivalent to one where $s_{ii} = y_{i,i}(1-p_i y_{i,i})$ is replaced by $s_{ii} = y_{i,i}(1-x_i)$, and $s_{ij} = -p_j y_{j,i} y_{i,j}$ is replaced by $s_{ij} = 0$.

Indeed, simulation studies suggest that such a substitution gives the preferred decentralized algorithm

$$p_i(k+1) = p_i(k) - \frac{r_i}{y_{i,i}(1-x_i)} \ for\ i = 1,2,\ldots,m, \tag{4.18}$$

with a step-size adjustment $\alpha(k)$ not needed in our simulations. For the simplest randomly generated examples m=4 and n=2, the decentralized algorithm required between 60-80 iterations to converge within the accuracy of the computer. The number of iterations required for the decentralized algorithm to converge within the accuracy for the computer increases with the number of users in the system e.g. for m=20 and n=2 at least 200 iterations before satisfactory convergence is achieved.

5. Conclusion. In this paper, we have established an existence result for an antenna array problem. The existence result is established by construction, and by then employing an optimization on a manifold argument.

Finally, in Section 4, we provide explicit results for the cases when $m = n + 1$ and $m = 1$, and a general iterative solution for the case of $m > n + 1$ (where m is the number of users, and n is the number of antennas in the array). The resultant general solution is a centralized algorithm and asymptotically Newton, and supersedes the existing *ad-hoc* decentralized algorithms, providing improvements in convergence, to the accuracy of the computer, of at least one order of magnitude.

REFERENCES

[1] S.Haykin, *Adaptive Filter Theory*, Third Edition, Prentice-Hall, 1996.
[2] G.J.Foschini, *Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas*, Bell Labs Tech. J., pp. 41–59, 1996.
[3] J.H.Winters, *Smart antennas for wireless systems*, IEEE Pers. Commun., pp. 23–27, Feb., 1998.
[4] A.J.Paulraj and B.C.Ng, *Space-time modems for wireless personal communications*,
[5] T.F.Wong, T.M.Lok et al, *A linear receiver for direct-sequence spread-spectrum multiple-access systems with antenna arrays and blind adaptation*, IEEE Trans. on Inform. Theory, 44:2(1998), pp. 659–676.
[6] V.Tarokh, N.Seshadri and A.Calderbank, *Space time codes for high data rates wireless communications, performance criterion and code construction*, IEEE Trans. on Inform. Theory, 44:2(1998), pp. 744–765.
[7] , K.Sheikh, D.Gesbert et al, *Smart antennas for broadband wireless access networks*, IEEE Commun. Mag., pp. 100–105, Nov., 1999.
[8] U.Madhow and M.L.Honig, *MMSE interference suppression for direct-sequence spread-spectrum CDMA*, IEEE Trans. on Commun., 42:12(1994), pp. 3178–3188.
[9] K.S.Gilhousen, I.M.Jacobs et al, *On the capacity of a cellular CDMA system*, IEEE Trans. on Veh. Technol., 40:2(1991), pp. 303–312.
[10] J.Zander, *Performance of optimum transmitter power control in cellular radio systems*, IEEE Trans. on Veh. Technol., 41:2(1992), pp. 57–62.
[11] A.M.Viterbi and A.J.Viterbi, *Erlang capacity of a power-controlled CDMA system*, IEEE Journal on Selected Areas in Commun., pp. 892–900, Aug., 1993.
[12] G.J.Foschini and Z.Miljanic, *A simple distributed autonomous power control algorithm and its convergence*, IEEE Trans. on Veh. Technol., 42:4(1993), pp. 641-646.
[13] J.C.I.Chuang and N.R.Sollenberger, *Uplink power control for TDMA portable radio channels*, IEEE Trans. Veh. Technol., 43:1(1994), pp. 33–39.
[14] S.A.Grandhi, R.Vijayan and D.J.Goodman, *Distributed power control in cellular radio systems*, IEEE Trans. on Commun., 42:2(1994), pp. 226–228.
[15] R.D.Yates and C.Y.Huang, *Integrated power control and base station assignment*, IEEE Trans. on Veh. Technol., 44:3(1995), pp. 638–644.
[16] C.W.Sung and W.S.Wong, *A distributed fixed-step power control algorithm with quantization and active link protection*, IEEE Trans. on Veh. Technol., 48:2(1999), pp. 553–562.
[17] R.D.Yates, *A framework for uplink power control in cellular radio systems*, IEEE Journal on Selected Areas in Commun., pp. 1341–1347, Sep. 1995.
[18] C.W.Sung and W.S.Wong, *The convergence of an asynchronous cooperative algorithm for distributed power control in cellular systems*, IEEE Trans. on Veh. Technol., 48:2(1999), pp. 563–570.
[19] U.Helmke and J.B.Moore, *Optimization and Dynamical Systems*, Springer-Verlag, 1994.