Solving Nonconvex Power Control Problems in Wireless Networks: Low SIR Regime and Distributed Algorithms

Chee Wei Tan, Daniel P. Palomar, Mung Chiang
Department of Electrical Engineering, Princeton University, Princeton, NJ, 08544
Email: {cheetan, danielp, chiangm}@princeton.edu
Phone: (609) 258-5071

Abstract: In wireless cellular networks that are interference-limited, a variety of power control problems can be formulated as nonlinear optimization with a system-wide objective, e.g., maximizing total system throughput under many QoS constraints from individual users. Previous work have been done in the high SIR regime by turning these problems with nonlinear objectives and constraints into convex optimization problems. However, in the medium to low SIR regime, these problems cannot be transformed into tractable convex optimization problems. This paper makes two contributions: (1) In the low SIR regime, we propose a method with centralized computation to obtain the global optimal solution by solving a series of geometric programs. (2) While efficient and robust algorithms have been extensively studied for centralized solutions of geometric programs, distributed algorithms have not been fully explored. We present a systematic method of distributed algorithms for power control based on geometric programs in high SIR regime. These two contributions can be readily combined to solve any nonlinear power control problems in general SIR regime. These techniques are illustrated through a series of examples.

Keywords: Wireless networks transmit power optimization, geometric programming, distributed algorithms
Solving Nonconvex Power Control Problems in Wireless Networks: Low SIR Regime and Distributed Algorithms

Chee Wei Tan, Daniel P. Palomar, Mung Chiang
Department of Electrical Engineering, Princeton University

Abstract—In wireless cellular networks that are interference-limited, a variety of power control problems can be formulated as nonlinear optimization with a system-wide objective, e.g., maximizing total system throughput under many QoS constraints from individual users. Previous work have been done in the high SIR regime by turning these problems with nonlinear objectives and constraints into convex optimization problems. However, in the medium to low SIR regime, these problems cannot be transformed into tractable convex optimization problems. This paper makes two contributions: (1) In the low SIR regime, we propose a method with centralized computation to obtain the global optimal solution by solving a series of geometric programs. (2) While efficient and robust algorithms have been extensively studied for centralized solutions of geometric programs, distributed algorithms have not been fully explored. We present a systematic method of distributed algorithms for power control based on geometric programs in high SIR regime. These two contributions can be readily combined to solve any nonlinear power control problems in general SIR regime. These techniques are illustrated through a series of examples.

I. INTRODUCTION

As wireless networks support an increasingly wide variety of applications, e.g., voice, data, video, Quality of Service (QoS) provisioning has become an important issue. Due to the broadcast nature of radio transmission, data rates and other QoS metrics in wireless networks are affected by signal interference from other users. The Signal to Interference Ratio (SIR) is often used to capture the effect of both co-channel and adjacent channel interference. Power control can be used to control SIR, and in doing so, indirectly control the QoS seen by users in the network. While a system-wide goal must be supported by the network, individual users’ QoS requirements must also be satisfied. Any power allocation must be constrained by a feasible set formed by these minimum requirement from the users. Hence, there exists a tradeoff between user-centric constraints and network centric objective.

Traditionally, power control problem is often formulated as minimizing a linear objective function such as the total power, e.g., [2], and the methods are usually not general enough to support a diverse set of nonlinear system objective and QoS constraints because many QoS metrics are nonlinear functions of SIR, which is in turn a nonlinear function of transmit powers. Hence, it may appear that many power control problems involving QoS metrics are not efficiently solvable. Recently in [1], [3], [4], it is shown that a variety of power control problems can be formulated as nonlinear optimization with a system-wide objective, e.g., maximizing total system throughput or worst case user throughput, under many QoS constraints from individual users e.g., on data rate, delay, and outage probability, through a Geometric Programming (GP) framework. GP can be used to efficiently compute a variety of resource allocation in wireless network, and efficiently determine the feasibility of user requirements by returning either a feasible set of solutions or a certificate of infeasibility [1], [3], [4]. This leads to an effective admission control and admission pricing method. However, there are two main limitations in the GP-based power control methods. The first limit is the high SIR assumption used to render some nonconvex QoS problems tractable by GP, and the second limit is the need for centralized computation. This paper shows how to overcome both of these limitations in GP-based power control.

The paper is organized as follows. In section II, we describe the system model of a cellular network. In section III, we describe GP-based power control. In section IV, we solve the nonconvex problem of power control in low SIR regime using a series of GP. In section V, we show how special structures in GP and its Lagrange dual problem lead to distributed algorithms which can be used to solve any standard GP problem. Finally, we conclude the paper in section VI.

II. SYSTEM MODEL

Due to space limit, we focus only on cellular networks, but the results are readily extendable to ad hoc networks, as in [1], [3]. We consider a single cell cellular network with one base station and $n$ mobile users. The setup has $n$ logical links (or users) with $n$ transmitters and receivers.¹ Transmit powers for each user are denoted by $P_1, \ldots, P_n$. Under Rayleigh fading, the power received from transmitter $j$ at receiver $i$ is given by $G_{ij}F_{ij}P_j$ where $G_{ij} \geq 0$ represents the path gain and is often modeled as proportional to $d_{ij}^{-\gamma}$ where $d_{ij}$ is distance and $\gamma$ is the power fall-off factor. We also let $G_{ij}$ encompass antenna gain and coding gain. The numbers $F_{ij}$ model Rayleigh fading and are assumed to be independent with unit mean. The distribution of the received power from transmitter $j$ at receiver $i$ is exponential with mean value $E[G_{ij}F_{ij}P_j] = G_{ij}P_j$. The SIR for the receiver on logical link $i$ is:

$$
\text{SIR}_i = \frac{P_iG_{ii}F_{ii}}{\sum_{j \neq i} P_j G_{ij}F_{ij} + n_i}
$$

¹Different receivers do not necessarily mean different physical receivers as it may mean the same physical receiver with different frequency channels, codes, or antenna beams in an antenna array.
where $n_i$ is the noise for receiver $i$.

The constellation size $M$ used by a link can be closely approximated for MQAM modulations as follows: $M = 1 + \ln(\phi_B) - \text{BER}$ where BER is the bit error rate and $\phi_1, \phi_2$ are constants that depend on the modulation type. Defining $K = \ln(\phi_B)$ leads to an expression of the data rate $R_i$ on the $i$th link as a function of SIR:

$$R_i = \frac{1}{T} \log_2(1 + K \text{SIR}_i),$$

which can be approximated as $R_i = \frac{1}{T} \log_2(K \text{SIR}_i)$ when $K \text{SIR}$ is much larger than 1. This approximation is reasonable either when the signal level is much higher than the interference level or when the spreading gain is large. For notational simplicity in the rest of this paper, we redefine $G_{ii}$ as $K$ times the original $G_{ii}$, thus absorbing constant $K$ into the definition of SIR.

### III. Geometric Programming-based Power Control

It is shown in [1] that in high SIR regime, many nonlinear QoS constraints can be solved using GP-based power control. GP-based power control is a convex optimization framework for resource allocation. There are two equivalent forms of GP: standard form and convex form. We first define a monomial as a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$f(x) = d_1 x_1^{a_1} \cdots x_n^{a_n},$$

where the multiplicative constant $d \geq 0$ and the exponential constants $a_j \in \mathbb{R}$, $j = 1, 2, \ldots, n$. A sum of monomials, indexed by $k$ below, is called a posynomial:

$$f(x) = \sum_{k=1}^{K} d_k x_1^{a_{1k}} \cdots x_n^{a_{nk}}.$$

where $d_k \geq 0$, $k = 1, 2, \ldots, K$, and $a_{jk} \in \mathbb{R}$, $j = 1, 2, \ldots, n, k = 1, 2, \ldots, K$. The key features about posynomials are its positivity and convexity (in log domain).

Minimizing a posynomial subject to posynomial upper bound inequality constraints and monomial equality constraints is called GP in standard form:

$$\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \leq 1, \quad i = 1, 2, \ldots, m, \\
& \quad h_l(x) = 1, \quad l = 1, 2, \ldots, M
\end{align*}$$

where $f_i$, $i = 0, 1, \ldots, m$, are posynomials: $f_i(x) = \sum_{k=1}^{K} d_{ik} x_1^{a_{1k}} \cdots x_n^{a_{nk}}$ and $h_l(x) = 1, 2, \ldots, M$ are monomials: $h_l(x) = d_{lk} x_1^{a_{1l}} x_2^{a_{2l}} \cdots x_n^{a_{nl}}$.

GP in standard form is not a convex optimization problem, because posynomials are not convex functions. However, with a logarithmic change of the variables and multiplicative constants: $y_i = \log x_i, b_k = \log d_k, b_l = \log d_l$, we can turn it into the following equivalent problem in $y$:

$$\begin{align*}
\text{minimize} & \quad p_0(y) = \log \sum_{k=1}^{K} \exp(a_{1k} y + b_k) \\
\text{subject to} & \quad p_i(y) = \log \sum_{k=1}^{K} \exp(a_{1k} y + b_k) \leq 0, \quad \forall i, \\
& \quad q_l(y) = a_{1l} y + b_l = 0, \quad l = 1, 2, \ldots, M.
\end{align*}$$

This is referred to as GP in convex form, which is a convex optimization problem since it can be verified that the log-sum-exp function is convex, and the equality constraint functions are affine. Fig. 1 shows the approach of GP-based power control for general SIR regime. In high SIR regime, it is known that we solve only one GP. The first contribution of this paper (in the next section) shows that, in medium to low SIR regime, we can solve truly nonconvex power control problems that cannot be turned into convex formulation through a series of GPs.

### IV. Transmit Power Allocation in Low SIR Regime

A limitation in the use of GP is the assumption that SIR is much larger than 0dB, i.e., when SIR is not much larger than 0dB, the approximation of $\log(1 + \text{SIR})$ as log SIR does not hold. Unlike SIR, which is an inverted posynomial, $1 + \text{SIR}$ is not an inverted posynomial. Instead, $1 + \text{SIR}$ is a ratio between two posynomials: To overcome this issue, GP can be extended to Signomial Programming (SP): minimizing a signomial subject to upper bound inequality constraints on signomials, where a signomial is a sum of monomials, possibly with negative multiplicative coefficients:

$$s(x) = \sum_{i=1}^{N} c_i g_i(x)$$

where $c \in \mathbb{R}^N$ and $g_i(x)$ are monomials. SP is more general than GP since it allows the use of signomials in lieu of posynomials.

An SP cannot be directly solved since it is not convex and cannot be rewritten in convex form (as opposed to a GP). We first convert an SP into a Complementary GP, which allows upper bound constraints on the ratio between two posynomials, and then apply a monomial approximation iteratively to obtain a series of GPs. This is called the condensation method (see [1] and references therein). The first step of conversion from an SP into a Complementary GP is simple. An inequality in SP of the following form

$$f_{i1}(x) - f_{i2}(x) \leq 1,$$

where $f_{i1}, f_{i2}$ are posynomials, is clearly equivalent to

$$\frac{f_{i1}(x)}{1 + f_{i2}(x)} \leq 1.$$

The problem is still not convex and cannot be rewritten as a convex problem; therefore, there is no practical algorithm to
obtain the optimal solution. However, in this form, it turns out to be convenient to use monomial approximation. Now we have two choices to make the monomial approximation. One is to approximate the denominator $1 + f_{12}(x)$ with a monomial but leave the numerator $f_{11}(x)$ as a posynomial. This is called the (single) condensation method, and results in a GP approximation of an SP. An iterative procedure can again be carried out: given a feasible $x^k$, from which a monomial approximations using $\alpha(x^k)$ (described in the next subsection) can be made and a GP formed, from which an optimizer can be computed and used as $x^{k+1}$, which becomes the starting point for the next iteration. This sequence of computation of $x$ may converge to $x^*$, an optimizer of the original SP. Another choice is to make the monomial approximation for both the denominator $1 + f_{12}(x)$ and numerator posynomials $f_{11}(x)$. That turns all the constraints into monomials, and after a log transformation, approximate SP as a linear program. This is called the double condensation method, and a similar iterative procedure can be carried out as in the single condensation case.

A. Monomial Approximation

There are many ways to make a monomial approximation of a posynomial. One possibility is based on the following simple inequality: arithmetic mean is greater than or equal to geometric mean, i.e.,

$$\sum_i \alpha_i v_i \geq \prod_i \left(\frac{u_i}{\alpha_i}\right)^{\alpha_i},$$

where $v > 0$ and $\alpha \succeq 0$, $1^T \alpha = 1$. Letting $u_i = \alpha_i v_i$, we can write this basic inequality as

$$\sum_i u_i \geq \prod_i \left(\frac{u_i}{\alpha_i}\right)^{\alpha_i}.$$

Let $\{u_i(x)\}$ be the monomial terms in a posynomial $f(x) = \sum_i u_i(x)$. A lower bound inequality on posynomial $f(x)$ can now be approximated by an upper bound inequality on the following monomial:

$$\prod_i \left(\frac{u_i(x)}{\alpha_i}\right)^{-\alpha_i}.$$

This approximation is in the conservative direction because the original constraint is now tightened. There are many choices of $\alpha$. One possibility is to let

$$\alpha_i(x) = u_i(x)/f(x), \quad \forall i,$$

which obviously satisfies the condition that $\alpha > 0$ and $1^T \alpha = 1$. Given an $\alpha$ for each lower bound posynomial inequality, a standard form GP can be obtained based on the above geometric mean approximation.

B. Applications to power control

GP-based power control problems in medium to small SIR regimes become sigmoidal programs, which can be solved by single or double condensation method. We discuss about single condensation method first. Consider a representative problem formulation of maximizing total system throughput in a cellular wireless network subject to user rate and outage probability constraints:

$$\begin{align*}
\text{maximize} & \quad R_{\text{system}}(P) \\
\text{subject to} & \quad R_i(P) \geq R_{\text{req},i}, \forall i, \\
& \quad P_{\alpha,i}(P) \leq P_{\alpha,i,\max}, \forall i, \\
& \quad P_i \leq P_{i,\max}, \forall i,
\end{align*}$$

which is explicitly written out as:

$$\begin{align*}
\text{minimize} & \quad \prod_{i=1}^N \frac{1}{1 + \text{SIR}_i} \\
\text{subject to} & \quad (2^{R_{\text{req},i}} - 1) \frac{1}{\text{SIR}_i} \leq 1, \forall i, \\
& \quad (\text{SIR}_{\text{th}})^{N-1}(1 - P_{\alpha,i,\max}) \prod_{i \neq k} G_{ik} P_k G_{ik i} P_i \leq 1, \forall i, \\
& \quad P_i (P_{i,\max})^{-1} \leq 1, \forall i,
\end{align*}$$

where $R_{\text{req},i}$ is the minimum required data rate for each user, $\text{SIR}_{\text{th}}$ denotes the threshold for SIR outage, $P_{\alpha,i,\max}$ is the maximum outage probability, and $P_{i,\max}$ is the maximum power constraint. The variables are $P$. The outage probability can be interpreted as the fraction of time the $i$-th logical link experiences an outage due to fading, and a derivation of the above outage probability constraint can be found in [4]. All the constraints are posynomials. However, the objective is not a posynomial, but a ratio between two posynomials. This power control problem is a Complementary GP, and can be solved by condensation method by solving a series of GPs. Specifically, we have the following algorithm:

STEP 0: Choose an initial feasible $P$.

STEP 1: Evaluate the denominator posynomial of the (6) objective function with the given $P$.

STEP 2: Compute for each term $i$ in this posynomial, $\alpha_i = \frac{\text{value of } i\text{-th term in posynomial}}{\text{value of posynomial}}$.

STEP 3: Condense the denominator posynomial of the (6) objective function into a monomial using (5) with weights $\alpha_i$.

STEP 4: Solve the resulting GP using interior point method.

STEP 5: Go to STEP 1 using $P$ of STEP 4.

STEP 6: Terminate the $k$-th loop if $\|D^{(k)} - D^{(k-1)}\| \leq \epsilon$ where $\epsilon$ is the error tolerance for exit condition.

As condensing the objective in the above problem gives us an underestimate of the objective value, each GP in the condensation iteration loop tries to improve the accuracy of the approximation to a particular minimum in the original feasible region.

Example 1. We consider a cellular wireless network with 3 users. Let $T = 10^{-6}$, $G_{ij} = 1.5$, and generate $G_{ij}, i \neq j$, as independent random variables uniformly distributed between 0 and 0.3. Threshold SIR is $\text{SIR}_{\text{th}} = -10$ dB, and minimal data rate requirements are 100 kbps, 600 kbps and 1000 kbps for logical links 1, 2 and 3 respectively. Maximal outage probabilities are 0.01 for all links, and maximal transmit powers are 3mW, 4mW and 5mW for link 1, 2 and 3 respectively.
For each instance of SP in (6), we pick a random initial feasible power vector $P$ uniformly between 0 and $P_{max}$. Fig. 2 compares the maximized total network throughput achieved over five hundred sets of experiments with different initial vectors. With (single) condensation method, SP converges to different optima over the entire set of experiments, achieving (or coming very close to) the global optimum at 5290 bps (96% of the time) and a local optimum at 5060 bps (4% of the time), thus very likely to converge to or very close to the global optimum. The number of GP iterations required by condensation method over the same set of experiments is 15 GPs if an extremely tight exit condition is picked for SP condensation iteration: $\epsilon = 1 \times 10^{-10}$. This average can be substantially reduced by using a larger $\epsilon$, e.g., increasing $\epsilon$ to $1 \times 10^{-2}$ requires on the average 4 GPs.

![Figure 2: Maximized total system throughput achieved by (single) condensation method for 500 different initial feasible vectors (Example 1).](image)

The optimum of power control produced by condensation method may be a local one. The following heuristics of solving a series of SPs (each solved through a series of GPs) can be further applied to help find the global optimum. After the original SP (6) is solved, a slightly modified SP is formulated and solved:

\[
\text{minimize} \quad t \\
\text{subject to} \quad \frac{1}{\prod_{i=1}^{N} \frac{1}{1 + \text{SIR}_i}} \leq t, \\
\qquad \quad t \leq \frac{t_0}{\alpha},
\]

Same set of constraints as problem (6),

where $\alpha$ is a constant slightly larger than 1. At each iteration of a modified SP, the previous computed optimum value is set to constant $t_0$ and the modified problem (7) is solved to yield an objective value that is better than the objective value of the previous SP by at least $\alpha$. The auxiliary variable $t$ is introduced so as to turn the problem formulation into SP in $(P, t)$. If we obtain the global optimal solution in Eqn (6), Eqn (7) will be infeasible, and thus only one SP iteration is needed.

**Example 2.** The above heuristics is applied to the instances of Example 5 where solving SP returns a locally optimal power allocation, and is found to obtain the globally optimal solution within 1 or 2 rounds of solving additional SPs (7).

We have discussed a power control problem (6) where the objective function needs to be condensed. The method is also applicable if some constraint functions are signomials and need to be condensed. For example, consider the case of differentiated services where a user expects to obtain a predicted QoS relatively better than the other users. We may have a proportional delay differentiation model where a user who pays more tariff obtains a delay proportionally lower as compared to users who pay less.

Packet traffic entering the base station at the transmitter of logical link $i$ is assumed to be Poisson with parameter $\lambda_i$ and to have an exponentially distributed length with parameter $\Gamma$. Using the model of an $M/M/1$ queue, the total packet arrival rate at queue $i$ is $\Lambda_i$. The expected delay $\bar{D}_i$ can be written as

\[
\bar{D}_i = \frac{1}{\Gamma (R_i (P) - \Lambda_i)}.
\]

Then for a particular delay ratio between any user $i$ and $j$, $\sigma_{ij}$, we have $\frac{\bar{D}_i}{\bar{D}_j} = \sigma_{ij}$, which, by (8), is equivalent to

\[
\frac{1 + \text{SIR}_j}{(1 + \text{SIR}_i)^{\sigma_{ij}}} = 2^{(\lambda_j - \sigma_{ij} \lambda_i)T/\Gamma}.
\]

The denominator on the left hand side is a posynomial raised to a positive power. Therefore, double condensation method can be readily used to solve the proportional delay differentiation problem because the function on the left hand side can be condensed to a monomial, and a monomial equality constraint is allowed in GP.

**Example 3.** We consider the wireless cellular network in Example 1 with an additional constraint $\frac{D_i}{D_j} = 1$. The arrival rates of each user at base station is measured and input as network parameters into Eqn. (9). Fig. 3 and 4 show the convergence towards satisfying all the QoS constraints including the DiffServ constraint. As shown on the figures, the convergence is extremely fast with the power allocations very close to the optimal power allocation by the 8-th GP iteration.

<table>
<thead>
<tr>
<th>Method</th>
<th>System throughput</th>
<th>$P^1_1$</th>
<th>$P^2_2$</th>
<th>$P^3_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhaustive search</td>
<td>6626 kbps</td>
<td>0.65</td>
<td>0.77</td>
<td>0.79</td>
</tr>
<tr>
<td>Single condensation</td>
<td>6626 kbps by solving 17 GPs</td>
<td>0.65</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td>Added DiffServ constraint</td>
<td>6624 kbps by solving 17 GPs</td>
<td>0.68</td>
<td>0.75</td>
<td>0.68</td>
</tr>
</tbody>
</table>

**TABLE I**

**Example 3.** Row 2 shows Example 1 using single condensation on system throughput only, and Row 1 shows the optimal solutions found by exhaustive search. Row 3 shows Example 3 with an additional DiffServ constraint $D_1/D_3 = 1$.

V. DISTRIBUTED ALGORITHMS FOR GP-BASED POWER CONTROL

Although GP has highly efficient interior point algorithms that compute the global optimal solution in polynomial time,
these algorithms requires centralized computation. Hence, distributed and scalable algorithms are often desired, especially in a practical system like a large network. In this section, we use the dual decomposition method to decompose a geometric program into smaller sub-problem whose solutions are jointly and iteratively coordinated by the use of dual variables. The key step is to introduce auxiliary variables and addition equality constraints, thus transferring the coupling in the objective to coupling in the constraints, which can be solved by introducing ‘consistency pricing’. ‘Consistency prices’ are updated via local communication channels among the variables that are coupled together with each other. We illustrate our idea through an unconstrained minimization problem followed by an application of our technique to wireless network power control.

Suppose we have the following unconstrained standard GP minimization problem in $x > 0$:

$$
\text{minimize } \sum_s f_s(x_s, \{x_i\}_{i \in I(s)})
$$

where $x_s$ denotes the local variable, $\{x_i\}_{i \in I(s)}$ denote the coupled variables from other users, and $f_s$ is either a monomial or posynomial. Make a change of variable $y_s = \ln x_s$ in the original problem. Hence, we have

$$
\text{minimize } \sum_s f_s(e^{y_s}, \{e^{y_i}\}_{i \in I(s)}).
$$

By introducing auxiliary variables $y_s$ for the coupled arguments and additional equality constraints to enforce consistency:

$$
\begin{align*}
\text{minimize} & \quad \sum_s f_s(e^{y_s}, \{e^{y_i}\}_{i \in I(s)}) \\
\text{subject to} & \quad y_{si} = y_i, \forall i \in I(s), \forall s.
\end{align*}
$$

(11)

Each $s$-th sub-problem controls the local variables $(y_s, \{y_{si}\}_{i \in I(s)})$. Next, the Lagrangian of (11) is formed as

$$
L(\{y_s\}, \{y_{si}\}; \{\gamma_{si}\}) = \sum_s f_s(e^{y_s}, \{e^{y_i}\}_{i \in I(s)}) + \sum_{i \in I(s)} \gamma_{si}(y_i - y_{si}).
$$

The minimization of the Lagrangian is done simultaneously at each sub-problem. In addition, the following master dual problem has to be solved:

$$
\text{maximize } g(\{\gamma_{si}\})
$$

(12)

where

$$
g(\{\gamma_{si}\}) = \sum_s \min L_s(y_s, \{y_{si}\})
$$

and each function $L_s(y_s, \{y_{si}\})$ is given by

$$
f_s(e^{y_s}, \{e^{y_i}\}_{i \in I(s)}) + \left( \sum_{i \in I(s)} \gamma_{si}e^{y_i} - \sum_{i \in I(s)} \gamma_{si}y_{si} \right).
$$

In the general case where $f_s$ is constrained, $g(\{\gamma_{si}\})$ contains the additional term $-\theta^2 I$ where $\theta$ includes the Lagrange multipliers for every constraint excluding the auxiliary variable equality constraints. Note that the transformed primal problem is convex. Hence, the Lagrange dual problem solves the original standard GP problem assuming zero duality gap.

At the $s$-th sub-problem, the function $L_s$ involving only local variables is minimized upon receiving the updated dual variables $\{\gamma_{si}, i: s \in I(i)\}$ (recall that $\{\gamma_{si}, i \in I(s)\}$ are local dual variables). Each $s$-th sub-problem in turn updates the consistency prices:

$$
\gamma_{si}(t + 1) = \gamma_{si}(t) - \theta(t)(y_i(t) - y_{si}(t)), \forall i \in I(s).
$$

Appropriate choice of the stepsize $\alpha(t) > 0$ leads to stability and convergence of the dual algorithm [1].

**Example 4.** We maximize total system throughput for three logical links with only maximum power and outage probability constraint. For simplicity of illustration, we assume that all channel gains are fixed at one for all users, and thermal noise is negligible. We have from (6), after some manipulations, the simplified problem as

$$
\begin{align*}
\text{minimize } & \quad \sum_{i=1}^{3} e^{y_i} \left( \sum_{j=1, j \neq i}^{3} e^{-y_j} \right) \\
\text{subject to } & \quad e^{y_i} (P_i^{max})^{-1} \leq 1, i = 1, 2, 3 \\
& \quad \sum_{i=1}^{3} \alpha e^{-2y_i} \sum_{j=1, j \neq i}^{3} e^{y_j} \leq 1
\end{align*}
$$

where $\alpha = \frac{1}{2} \text{SIR}_{th}(1 - P_{o,i,\text{max}}), \forall i$.

In order to decouple the optimizing variables, we introduce auxiliary variables $y_{ij}$ and additional constraints $y_{ij} = y_j, \forall i, j \neq i$. The Lagrangian is formed as:

$$
L(y_i, y_{ij}; \lambda, \nu, \mu) = \sum_{i=1}^{3} L_i(y_i, y_{ij}, \lambda, \nu, \mu) - \nu^T 1 - \lambda
$$
where \( \nu \) and \( \lambda \) are the Lagrange multipliers associated with the maximum power and the probability outage constraints respectively. \( \mu \) is a vector that contains the Lagrange multipliers associated with auxiliary variable constraints. The function \( L_i(y_i, y_j, \lambda, \nu, \mu) \), \( \forall i \) or \( L_i \) for brevity is

\[
L_i = e^{y_i} \sum_{j=1, j \neq i}^{3} e^{-y_{ij}} + \lambda \alpha e^{-2y_i} e^{\sum_{j=1, j \neq i}^{3} y_{ij}} + (P^\text{max}_i)^{-1} \nu_i e^{y_i} + \sum_{j=1, j \neq i}^{3} (\mu_{ij} y_{ij} - \mu_{ji} y_i).
\]

Note that the optimizing variables and constant parameters in each \( L_i \), \( \forall i \) are local to each link \( i \). The only coupling is in the \( \mu \) variables. Each \( L_i \), \( \forall i \) can now be jointly minimized in parallel and the local dual variables can be updated by the following iterative gradient method:

\[
\mu_{ij}(t + 1) = \mu_{ij}(t) + \alpha(t)(y_{ij} - y_j), \quad \forall i, j \neq i
\]

where \( t \) is the iteration number and \( \alpha(t) \) is the step size. The global dual variable can similarly be updated as

\[
\lambda(t + 1) = \lambda(t) + \alpha(t) \sum_{i=1}^{3} \alpha e^{-2y_i} e^{\sum_{j=1, j \neq i}^{3} y_{ij}} - 1).
\]

In our simulations, we let \( \alpha(t) \) be a constant value of 1/10. Fig. 5 shows the number of iterations needed for the Lagrangian at User 1 and 2 to converge to the optimum solution. Fig. 6 shows the convergence of the Lagrange multipliers updated at User 1 to the optimal solution. The average number of iterations or message passing between the three logical links requires on average 20 iterations.

**Interpretation and extensions of message passing algorithm:** Each term \( L_i \) (or monomial in the standard GP problem) above has a corresponding interpretation of a network element (end user or intermediate nodes). In the above example, the Lagrangian is decoupled and minimized at each logical link using local variables only. The sub-problems are usually smaller in size as compared to the original problem with fewer number of constraints. Updating the consistency prices using message passing through local communication channels incurs overhead. The overhead in our example above can be quantified as follows: each link updates one auxiliary variable from other links that are associated with it and receives two units of its updated auxiliary variables from the other links. Hence, at each iteration, four units of message per link are exchanged among the links. The overheads incurred can be interpreted as part of a network protocol that solves a global optimization problem. This overhead can be reduced by limiting the scope of message passing, e.g., ignoring the messages from links that are physically further. An alternate way is to take into account the form of coupling through interference, e.g., a summation of interference terms can be replaced by a single auxiliary variable.

VI. CONCLUSION

GP-based power control has recently been developed to compute globally optimal power allocation in high-SIR regime for a variety of nonlinear objectives and constraints. In this paper, we present two contributions that overcome the current bottlenecks of GP power control. First, in the low SIR regime, sigmoidal programming with centralized computation is presented to obtain the global optimal solution by solving a series of geometric programs. Two, we present a systematic theory of distributed algorithms for GP power control in high SIR regime using dual decomposition methods. Numerical methods show that the convergence to the global optimal solutions using these methods is fast. Lastly, these two contributions can be combined to distributively obtain a globally optimal solution for general SIR regime, which is also applicable to Digital Subscriber Line (DSL) spectrum management.

REFERENCES