

# Cost-constrained Viterbi Algorithm for Resource Allocation in Solar Base Stations

Viet Hung Tran and Marceau Coupechoux

**Abstract**—Solar energy is currently a popular renewable resource, yet limited daily. In green cellular networks, multiple constraints optimization (MCO) problems arise naturally. For example, a typical objective is to control the power transmission of hybrid base stations (connected to both solar panels and electrical grid) in order to maximize user’s average throughput, under the constraints of consumed grid energy and user’s blocking rate. However, such problems have been generally proved to be NP-hard. In this paper, we formulate this generic MCO problem as a quantized Markovian cost-reward model, with no assumption on input data. We then propose a novel algorithm, namely Cost-constrained Viterbi Algorithm, which recursively returns the optimal policy with linear computational complexity for this model. As an application, we provide engineering rules for the design of hybrid base stations through extensive simulations. In comparison with brute force method for a simple scenario, we find that our algorithm does achieve the constrained optimal policy.

**Index Terms**—Constrained optimization, Viterbi algorithm, energy harvesting, cellular networks, hybrid base station.

## I. INTRODUCTION

Renewable energy has been widely considered as next-generation technology for mobile networks [1]. Solar energy in particular offers very good advantages as it is predictable and solar panels are now available at affordable price, with low impact on the environment. The harvested solar energy is however sharply varying during the day time and is daily constrained. Moreover, there is an hourly mismatch in general between the mobile user’s traffic and the incoming solar energy. This mismatch might drain out solar panel’s battery at the time of high traffic, while preserving too much energy at the time of low traffic.

In this paper, we consider a hybrid base station (BS), connected to solar panels and the national electrical grid. For countering the mismatch between harvesting energy and user’s traffic processes, we allow the BS to control its transmit power  $P_{TX}$  along the day such that the Quality of Service (QoS) is maximized. We impose a constraint on the amount of energy to be taken from the grid, and possibly on the maximum user blocking rate. To solve this constrained optimization problem, we propose a novel algorithm, namely Cost-constrained Viterbi Algorithm (CVA), which recursively returns the constrained optimal policy for transmit power’s control with linear computational complexity. We illustrate

the CVA’s performance in realistic scenarios of solar BSs and European user’s traffic. Note that, because we impose no assumption on data model at the input, CVA also works for arbitrary renewable energy process (for example wind energy [2]).

### A. Related Work

Because BSs take more than 50% of operating power of a cellular network [3], several constrained optimization problems involving energy consumption have been studied in recent literature. For example in [4], the optimal criterion is a weighted combination of the consumed grid energy and of the number of dropped packets, while the constraint is the maximum transmitted power. In [5], the optimal criterion is energy efficiency, while the constraints are consumed grid energy and user’s data throughput. In [6], the constraints are transmitted data rate and battery size while consumed grid energy is minimized. In [7], the constraints are defined as user’s blocking rate and user’s data throughput while optimal criterion is to minimize consumed grid energy. In [8], [9], the problem is formulated as sum-rate maximization problem, in which QoS is defined as user’s average throughput and the constraint is consumed energy.

Unfortunately, the sum-rate maximization was proved to be an NP-hard problem in general [3], [4], [10]. To overcome the complexity issue, several papers propose heuristics, e.g. in [3], [5], [11]. They however do not provide any guarantee on the optimality of the proposed solution. Another approach is to transform the original problem into a convex optimization problem. However, this idea is only feasible in simplistic system models. For example in [5], [6], [8], [9], the convexity for constrained sum-rate was achieved, but without considering data traffic process and user’s admission control. Lastly, dynamic programming via discrete Markov Decision Process (MDP) were proposed in [4], [7]. With MDP, the curse of dimensionality remains and requires Markovian models, for example, for traffic or harvested energy [6].

Our MCO problem can also be seen as a special case of the Constrained Shortest Path (CSP) problem in graph theory. CSP problem was proved to be NP-hard [12] and, hence, currently has to rely on heuristics or discretization forms of dynamic programming [12], [13]. Our graph is however a tree because we assumed continuous cost-reward values and thus there are in general no two paths reaching the same node. As a consequence, the current recursive CSP algorithms cannot be efficiently applied to our model. In contrast, we propose a quantization step in CVA, which merely merges

V. H. Tran and M. Coupechoux are with LTCI, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France. V. H. Tran is also with Hochiminh city University of Technology, Vietnam (Email: tranviet Hung@hcmut.edu.vn, marceau.coupechoux@telecom-paristech.fr).

V. H. Tran is supported by french competitiveness cluster project SYSTEM-ATIC LCI4D and the french research cluster Digiteo.

close vertices and extracts a representative path among the paths reaching these vertices. In other words, CVA groups at each step close-by trajectories and extracts one such that the difference with others can be arbitrarily made small as the number of quantization levels increases. This step is not a rounding as done in recursive CSP algorithms so that the exact value of the path can be returned. This path extraction is the key step to extend CVA from one constraint to multiple constraints with the same computational complexity.

In broader context, our MCO problem is also a special case of constrained MDP problem [14]–[17]. Unlike our CVA, however, the state-of-the-art techniques for solving constrained MDP are based on Lagrangian linear programming, whose computational complexity in this case is inefficient in practice [16], [17]. For faster computation, a Lagrangian dynamic programming method was recently proposed (c.f. [14], [15]), in which linear programming was replaced by dynamic programming in solving Lagrangian unconstrained MDP equations. Nevertheless, the computational complexity of this method is at least a polynomial function dependent on number of constraints and magnitude of cost-reward values [17] and, hence, is very high in practice.

Our CVA can also be considered as a special case of constrained Viterbi algorithms, which were scarcely proposed in inference and decoding context (e.g. in [18], [19]). These algorithms in literature merely differ from traditional Viterbi algorithm by one extra step of validating constraints upon edges of a given Markov chain. In contrast, our CVA does not truncate violating edges on a given Markov chain, but makes truncation (together with quantization and constraint’s validation) a part of establishing Markov chain process.

## B. Contributions and Organization

Compared to the literature, we consider a generic multiple constraints optimization (MCO) problem with no assumption on input data. We propose a novel formulation of this problem, namely quantized Markovian cost-reward. This allows us to formulate any deterministic or probabilistic problem, where both optimal reward and constraint costs are independently accumulated over time. For example, all accumulated reward and cost based on  $\mathcal{L}_p$ -norm fall in this framework. This generalization makes all models proposed in above cited works on hybrid BS become its special cases.

In our approach, we rely on quantization, which yields the optimal solution when quantized intervals tends to zero [10]. The drawback is that, however, the number of quantized levels for MCO problem will grow exponentially in this case, which yields impractical computational complexity. In order to avoid this complexity, we adapt the famous Viterbi algorithm in decoding literature to this quantized Markovian model and return the globally constrained optimal trajectory. Owing to this novel algorithm, namely Cost-constrained Viterbi Algorithm, it is sufficient to only quantize one out of multiple constraints in order to asymptotically achieve the optimal policy for original model. Furthermore, the computational complexity of CVA only grows linearly with the number of quantization levels and time points, which makes it completely tractable. This

approach is in contrast to the state-of-the-art dynamic programming techniques, which consider the Cartesian product of quantized values of constrained variables and thus see their complexity grow exponentially with the number of variables. To the best of our knowledge, this is the first time Viterbi algorithm is applied to MCO problem.

As an application of CVA in this paper, our aim is to maximize user’s average throughput, subject to constraints of both consumed grid energy and user’s average blocking rate in hybrid BS. We show that our aim is equivalent to maximizing the  $\mathcal{L}_1$ -norm of user’s throughputs over time, subject to two constraints. The first constraint is  $\mathcal{L}_\infty$ -norm between accumulated renewable energy and accumulated station’s operating energy. The second constraint is  $\mathcal{L}_1$ -norm of user’s blocking rates over time.

In summary, this paper strives for both theoretical and practical solutions to constrained optimization of resource allocation. The contributions are as follows:

- A novel formulation for constrained optimization problem, namely quantized Markovian cost-reward is provided. This formulation is flexible and applicable to both continuous and discrete Markovian cost-rewards, whose special case is based on  $\mathcal{L}_p$ -norm.
- We then design a novel algorithm CVA, whose contribution is twofold. On one hand, it recursively returns an approximated constrained optimization solution, whose error is arbitrarily small up to quantization level. On the other hand, its computational complexity only grows linearly with coming data, owing to its recursive quantization scheme.
- For the first time, an  $\mathcal{L}_p$ -norm formulation is provided for energy process in a hybrid BS. Owing to this formulation, our constrained resource allocation problem can be recognized as special case of Markovian cost-reward system, for which CVA can be efficiently used.
- Our simulation is carefully calibrated to fit practical scenarios of a hybrid macro BS in LTE system. The key factors in designing a practical solar BS, e.g. solar panel size, necessary maximum power transmission, optimal trade-off between QoS and consumed power, etc., are illustrated and discussed. Hence, simulation results are of practical use in solar BS design. For example, in reference scenario for solar macro BS in European urban area, 40% of BS’s consumed energy can be typically saved by controlling  $P_{TX}$  according to CVA, in comparison with traditional fixed  $P_{TX}$  scheme.
- A novel iterative CVA scheme is also proposed for a solar BS network. This iterative scheme, which essentially applies CVA to each BS in a network one-by-one, always converges to local maximum of total reward for the whole network, under the same set of constraints of each BS. Owing to this decentralized approach, iterative CVA is able to optimally reduce (possibly switch-off) BSs’ interference power, achieving both lower grid consumed energy and higher QoS in the network. For example, our simulation shows that, iterative CVA can save up to 85% of consumed grid energy in typical scenario of hexagonal BSs, given the same QoS constraint of fixed  $P_{TX}$  scheme.

As CVA can be applied to a broader context than hybrid BS, we first present the theoretical part before focusing on the application. As a consequence, the paper is organized as follows: Section II presents the Markovian approach and CVA for MCO problem in generic form. Section III introduces the mathematical model for solar mobile station. Section IV illustrates the optimality of CVA via simulation. The paper is concluded in section V.

## II. MULTIPLE CONSTRAINTS OPTIMIZATION PROBLEM

In this section, we propose a Markovian formulation for solving an MCO problem and present our algorithm.

### A. Constrained optimization objective

Let us consider a policy  $\pi_i \triangleq \{a_1, a_2, \dots, a_i\} \in \mathcal{A}^i$ , in which each action  $a_i$  belongs to the same finite action set  $\mathcal{A}$  and the set  $\Pi_i \triangleq \mathcal{A}^i$  is called the policy space up to time point  $i = 1, 2, \dots, N$ . Let us denote  $\text{card}(\cdot)$  as cardinal number operator and define  $K \triangleq \text{card}(\mathcal{A})$ .

We assume that actions generate a sequence of independent rewards and costs. In this typical memoryless system, let us associate with each action  $a_i$  a pair  $\{b_i, c_i\}$  of scalar reward  $b_i \triangleq b_i(a_i) : \mathcal{A} \rightarrow \mathbb{R}$  and an  $L$ -dimensional vector of cost  $c_i = [c_{1,i}, \dots, c_{L,i}]' \triangleq c_i(a_i) : \mathcal{A} \rightarrow \mathbb{R}^L$ . Let us denote  $\{v_i, u_i\} \in \mathbb{R}^{L+1}$  the pair of scalar accumulated reward  $v_i$  and an  $L$ -dimensional vector of accumulated cost  $u_i$ , up to time point  $i$ , as follows:

$$\begin{aligned} v_i &\triangleq v(\pi_i) = \sum_{j=1}^i b_j(a_j), \\ u_i &\triangleq u(\pi_i) = \sum_{j=1}^i c_j(a_j), \end{aligned} \quad (1)$$

where  $v(\pi_i) : \mathcal{A}^i \rightarrow \mathbb{R}$  and  $u(\pi_i) : \mathcal{A}^i \rightarrow \mathbb{R}^L$ . Note that the notations  $u_i = [u_{1,i}, \dots, u_{L,i}]' \triangleq u(\pi_i) = [u_1(\pi_i), \dots, u_L(\pi_i)]'$  in (1) are element-wise, as follows:  $u_{l,i} \triangleq u_l(\pi_i) = \sum_{j=1}^i c_{l,j}(a_j)$ ,  $l = 1, 2, \dots, L$ .

Let us first define  $\pi_N^*$  as the optimal policy for non-constraint case, i.e.,  $\pi_N^* = \text{argmax}_{\pi_N \in \Pi_N} v(\pi_N)$ . Owing to independence and additivity in (1), the solution of the non-constraint case can be easily found, i.e.,  $\pi_N^* = \{a_1^*, \dots, a_N^*\}$  and  $\forall i, a_i^* = \text{argmax}_{a_i \in \mathcal{A}} b(a_i)$ .

In an MCO problem with threshold  $\tau$ , the solution is much harder to find. The objective is to find the constrained optimal policy  $\pi_N^{*(\tau)}$ , which maximizes the accumulated reward over a finite horizon, subject to  $L$  constrained conditions:

$$\begin{aligned} \pi_N^{*(\tau)} &= \text{argmax}_{\pi_N \in \Pi_N^{(\tau)} : \tau_1 \leq u(\pi_N) \leq \tau_2} v(\pi_N), \\ \text{with } \pi_N \in \Pi_N^{(\tau)} : &\begin{cases} \tau_{1,1} \leq u_1(\pi_N) \leq \tau_{1,2} \\ \vdots \\ \tau_{L,1} \leq u_L(\pi_N) \leq \tau_{L,2} \end{cases}, \quad (2) \end{aligned}$$

where  $\Pi_N^{(\tau)}$  denotes the constrained subspace in policy space  $\Pi_N$ , the two  $L$ -dimensional vectors of thresholds  $\tau =$

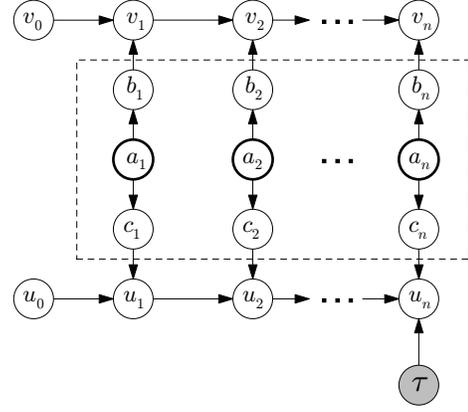


Figure 1. Markovian recursions for independent cost-reward scheme. Parameter  $\tau$  denotes constrained threshold.

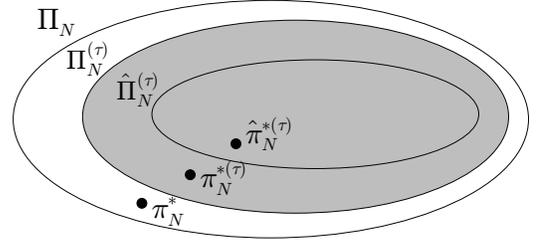


Figure 2. Venn diagram for constrained optimization space.  $\Pi_N$  is the set of unconstrained policies,  $\Pi_N^{(\tau)}$  is the set of constrained policies and  $\hat{\Pi}_N^{(\tau)}$  is the set of policies explored by CVA. In every set, there is an optimal policy represented by the dots.

$\{\tau_1, \tau_2\} \in \mathbb{R}^L \times \mathbb{R}^L$  are lower and upper bound of constraints, respectively, and inequalities are element-wise. This problem can be illustrated via Venn diagram in Fig. 2.

### B. Curse of dimensionality

In this constrained case, we can see that the action  $a_i^{(\tau)}$  has to belong to a constrained action set  $\mathcal{A}_i^{(\tau)} \subseteq \mathcal{A}$  at time  $i$ , such that the policy  $\pi_N^{(\tau)} = \{a_1^{(\tau)}, \dots, a_N^{(\tau)}\} \in \Pi_N^{(\tau)}$  satisfies the constraint (2). Hence, in general, the choice of  $a_i^{(\tau)}$  depends on both constrained policies in the past,  $\pi_{i-1}^{(\tau)} \in \Pi_{i-1}^{(\tau)}$ , and future,  $\pi_{i+1:N}^{(\tau)} \in \Pi_{i+1:N}^{(\tau)}$ . Owing to this lack of independence property, the constrained optimal policy (2) can only be found brute-forcedly at time  $N$ , after computing and verifying all  $K^N$  values of  $\pi_N$  with the constraint in (2). This computational complexity  $O(K^N)$  grows exponentially with time  $N$  and, hence, yields the curse of dimensionality for this constrained optimization problem.

In literature, the exact optimal solution for this constrained sum-rate problem was also proved to be NP-hard [3], [10].

### C. Markov chain formulation

From (1), we can recognize the natural Markovian property of  $\{v_i, u_i\}$  as follows:

$$\begin{aligned} v(\pi_i) &= b_i(a_i) + v(\pi_{i-1}), \\ u(\pi_i) &= c_i(a_i) + u(\pi_{i-1}), \end{aligned} \quad (3)$$

where by convention  $\pi_0 = \emptyset$ ,  $u_0 = v_0 = 0$ . The Markovian model (3) can be illustrated via a directed acyclic graph (DAG) [20] in Fig. 1.

As explained in subsection II-B, the number of action's trajectories<sup>1</sup>  $\pi_i$  grows exponentially with time-point  $i$ . Our idea is, at each time point, to categorize them into  $M$  representative groups, whose boundaries are defined via quantization method. As we will see, these  $M$  groups will construct an  $M$ -states Markov chain form over time. Note that, this categorization is not an approximate procedure, but the goal is merely to reformulate the original Markovian model (3) into a Markov chain form, which facilitates the reduction step for computational load in next subsection.

1) *Scalar quantization*: At each time point  $i = 1, 2, \dots, N$ , there are  $L + 1$  scalar variables in our MCO problem (2): the accumulated reward  $v_i$  and accumulated costs  $u_{l,i}$ , with  $l = 1, 2, \dots, L$ . The interesting point is, as we will show in section II-D, quantizing any one of these  $L + 1$  scalar quantization is enough to establish a tractable Markovian recursive model. Hence, let us define a quantization variable  $q_i \in \{v_i, u_{1,i}, u_{2,i}, \dots, u_{L,i}\}$  to be one of those  $L + 1$  scalar variables. Note that  $q_i$  is a function of the chosen policy  $\pi_i$  and can also be written  $q_i \triangleq q(\pi_i)$ . At each time point  $i$ , let us quantize scalar value  $q_i$  into  $M$  intervals  $[\tilde{q}_{k-1,i}, \tilde{q}_{k,i})$ , divided by  $M + 1$  levels  $\tilde{q}_i \triangleq [\tilde{q}_{0,i}, \tilde{q}_{1,i}, \dots, \tilde{q}_{M,i}]' \in \mathbb{R}^{M+1}$ , where  $\tilde{q}_{0,i} < \tilde{q}_{1,i} < \dots < \tilde{q}_{M,i}$ .

For fast computation, the quantized values might be set uniformly at fixed values, or via iterative method like the well-known Lloyd-Max algorithm [21], [22], which converges to locally mean-square-error (MSE). However, in our simulations, we found no significant gain in applying Lloyd-Max algorithm. For simplicity, we present in this paper the uniform setting, although our simulation results are almost the same with both Lloyd-Max (non-uniform) and uniform quantization.

At time point  $i$ , if  $\text{card}(\Pi_i) = K^i \leq M$ , we do not really need quantization step because the exact values  $q(\pi_i)$  for all values of  $\pi_i$ , under sorting technique, can be regarded as levels themselves. Otherwise, if  $\text{card}(\Pi_i) = K^i > M$ , at any  $i$ , we can assign:

$$\tilde{q}_{k,i} = \begin{cases} \min_{\pi_i \in \Pi_i} q(\pi_i) & , k = 0 \\ \max_{\pi_i \in \Pi_i} q(\pi_i) & , k = M \\ \tilde{q}_{0,i} + \frac{k}{M}(\tilde{q}_{M,i} - \tilde{q}_{0,i}) & , 0 < k < M \end{cases}. \quad (4)$$

Note that, in the case of quantizing scalar constraint value  $u_{l,i} = u_l(\pi_i)$ ,  $l = 1, 2, \dots, L$ , the extremum levels  $\tilde{q}_{0,i}$  and  $\tilde{q}_{M,i}$  might be set as thresholds  $\tau_1$  and  $\tau_2$  in constraint-space, instead of minimal and maximal values in (4), respectively. By this way, the accumulated quantization error is bounded by  $\tau_1$  and  $\tau_2$ . However, this online thresholding approach can only be applied if functional form of  $u_l(\pi_i)$  is monotonic, as explained in section II-E.

In practice, choosing which variable should be used for quantization step generally depends on two factors: its potential reduction on computational load and its potential gain in the approximation step. From our experience, we noticed that the quantization of  $v_i$  may not be preferable with respect

to the computational reduction because maximizing operator in (2) already reduced all values of  $v_i$  to a single optimal one.

2) *Trajectory segmentation*: We say that  $q(\pi_i)$  belongs to state  $k$  if  $q(\pi_i)$  belongs to  $k$ th interval  $[\tilde{q}_{k-1,i}, \tilde{q}_{k,i})$ ,  $k = 1, 2, \dots, M$ . By this way, we can divide the policy space  $\Pi_i$  into  $M$  separated subspaces  $\pi_i \in \Pi_i = \{\Pi_{1,i}, \dots, \Pi_{M,i}\}$ , in which  $\Pi_{k,i}$  is the set of all trajectories  $\pi_i = \pi_i^k \in \Pi_{k,i}$  making  $q(\pi_i^k)$  hit state  $k$  at time  $i$ . If there is no such trajectory, we assign  $\Pi_{k,i} = \emptyset$ . This is similar to the case of Markov chain with no visited state  $k$  at time  $i$ . We also divided each policy subspace  $\Pi_{k,i} = \{\Pi_{1,k,i}, \dots, \Pi_{M,k,i}\}$  into further  $M$  separated subspaces, where  $\Pi_{j,k,i}$  is the set of all trajectories  $\pi_i = \pi_i^{j,k}$  hitting state  $j$  at time  $i - 1$  and state  $k$  at  $i$ .

3) *Markovian trajectory model*: The original Markovian model (3) can be re-formulated via segmented trajectory model, as follows:

$$\begin{aligned} v(\pi_i^{j,k}) &= b_i(a_i^{j,k}) + v(\pi_{i-1}^j), \\ u(\pi_i^{j,k}) &= c_i(a_i^{j,k}) + u(\pi_{i-1}^j), \end{aligned} \quad (5)$$

$\forall j, k \in \{1, \dots, M\}$ ,  $i = 1, 2, \dots, N$ . Each action  $a_i^{j,k}$  represents a transition of policy from state  $j$  to state  $k$  (i.e. from  $\pi_{i-1}^j$  to  $\pi_i^k$ ), as illustrated in Fig. 3. Let us substitute the objective function (2) to Markovian trajectory model (5). For clarity, let us divide this substitution step into three consecutive sub-steps, as follows:

$$\begin{aligned} \max_{\pi_i} v(\pi_i) &= \max_k \left( \max_{\pi_i^k} v(\pi_i^k) \right), \\ \max_{\pi_i^k} v(\pi_i^k) &= \max_j \left( \max_{\pi_i^{j,k}} v(\pi_i^{j,k}) \right), \\ \max_{\pi_i^{j,k}} v(\pi_i^{j,k}) &= \max_{\pi_{i-1}^j} \left( \max_{a_i^{j,k}} b_i(a_i^{j,k}) + v(\pi_{i-1}^j) \right), \end{aligned} \quad (6)$$

in which  $j, k \in \{1, \dots, M\}$ ,  $i = 1, 2, \dots, N$  and max-operators are taken over constrained policy spaces:  $\pi_i \in \Pi_i^{(\tau)}$ ,  $\pi_i^k \in \Pi_{k,i}^{(\tau)}$ ,  $\pi_i^{j,k} \in \Pi_{j,k,i}^{(\tau)}$  and  $\pi_{i-1}^j \in \Pi_{j,i-1}^{(\tau)}$ . Because our trajectory segmentation via quantization is not an approximate step, the curse of dimensionality discussed in subsection II-B remains the same. However, we will use right below this Markovian formulation to reduce the computational load from  $K^i$  original trajectories up to time  $i$  to  $M$  by replacing all trajectories in  $\Pi_{k,i}$  with the best policy at time  $i$  in state  $k$ .

#### D. Forward-Backward recursion

In order to achieve a tractable recursion for (6), our idea is to approximate (6) by bringing operators  $\max_{\pi_{i-1}^j}$  into parentheses via plug-in substitution technique. This means we reject all policies  $\pi_i$ , except a single representative policy  $\hat{\pi}_i$  (see Fig. 2), corresponding to the highest accumulated reward among remaining policies up to time  $i$ . In the same way, we define  $\hat{\pi}_i^k$  as the representative policy at state  $k$  at time  $i$  and

<sup>1</sup>The words *policy* and *trajectory* are interchangeably used in this paper.

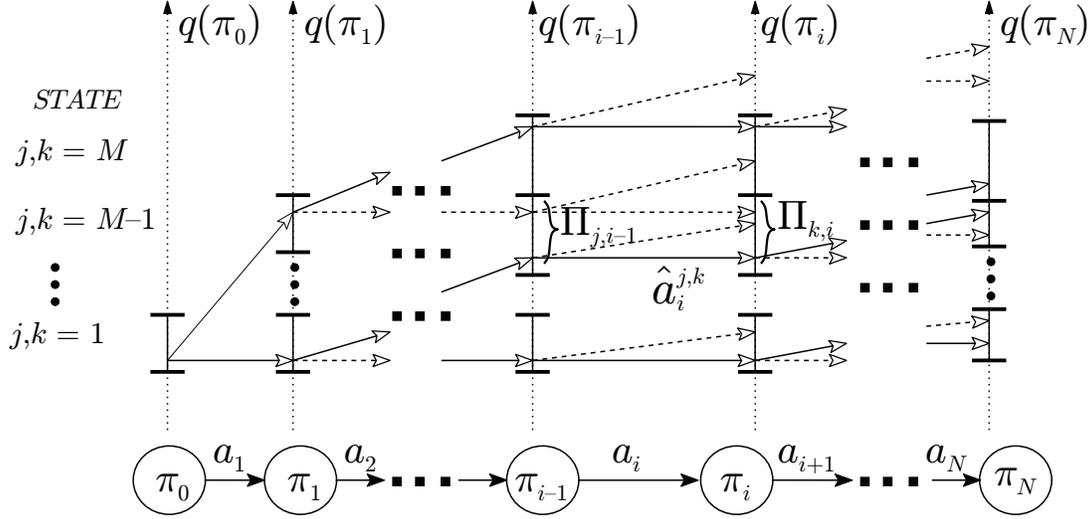


Figure 3. Quantized trellis diagram for CVA's Markovian recursions. Quantization variable  $q$  can be either accumulated reward  $v$  or cost  $u$ . Each vertex is a specific value of  $q$ . Each edge is a specific action. The dashed arrows are rejected trajectories. Focusing on the transition between  $i-1$  and  $i$ : all remaining trajectories reaching state  $j$ , i.e. in  $\Pi_{j,i-1}$ , are rejected except the one with highest accumulated reward  $\hat{\pi}_{i-1}^j$ . From this vertex, the action leading  $\hat{\pi}_{i-1}^j$  to state  $k$  at time  $i$  is  $\hat{a}_i^{j,k}$ , such that  $\hat{\pi}_i^k = \{\hat{a}_i^{j,k}, \hat{\pi}_{i-1}^j\}$ . All trajectories that violate the constraints are rejected.

$\hat{\pi}_i^{j,k}$  as the representative policy between states  $j$  at  $i-1$  and  $k$  at  $i$ . By this way, we achieve the following approximation:

$$\begin{aligned} v(\hat{\pi}_i) &\approx \max_{\pi_i} v(\pi_i), \\ v(\hat{\pi}_i^k) &\approx \max_{\pi_i^k} v(\pi_i^k), \\ v(\hat{\pi}_i^{j,k}) &\approx \max_{\pi_i^{j,k}} v(\pi_i^{j,k}). \end{aligned} \quad (7)$$

1) *Forward recursion for optimization function:* We now show that the representative policies can be computed recursively. Substituting (7) into (6), we can deduce consecutive relationships for (7), as follows:

$$\begin{aligned} v(\hat{\pi}_i) &= \max_k v(\hat{\pi}_i^k), \\ v(\hat{\pi}_i^k) &= \max_j v(\hat{\pi}_i^{j,k}), \\ v(\hat{\pi}_i^{j,k}) &= b_i(\hat{a}_i^{j,k}) + v(\hat{\pi}_{i-1}^j), \end{aligned} \quad (8)$$

in which  $\hat{\pi}_0^j = \pi_0^j$  and then the value  $v(\hat{\pi}_i^k)$ ,  $k = 1, 2, \dots, M$ , can be evaluated recursively from  $v(\hat{\pi}_{i-1}^j)$ ,  $j = 1, 2, \dots, M$ . For later use, we provide the following definitions:

$$\begin{aligned} \hat{\pi}_i &\triangleq \hat{\pi}_i^{k=\hat{k}_i}, \\ \hat{\pi}_i^k &\triangleq \hat{\pi}_i^{j=\hat{j}_i(k),k}, \\ \hat{\pi}_i^{j,k} &\triangleq \{\hat{a}_i^{j,k}, \hat{\pi}_{i-1}^j\}, \end{aligned} \quad (9)$$

together with:

$$\begin{aligned} \hat{k}_i &\triangleq \operatorname{argmax}_k v(\hat{\pi}_i^k), \\ \hat{j}_i(k) &\triangleq \operatorname{argmax}_j v(\hat{\pi}_i^{j,k}), \\ \hat{a}_i^{j,k} &\triangleq \operatorname{argmax}_{a_i^{j,k}} b_i(a_i^{j,k}). \end{aligned} \quad (10)$$

2) *Forward recursion for constraint space:* Given the representative policies  $\hat{\pi}_i^k$  found in (9), the forward recursion for accumulated cost can be feasibly computed, as follows:

$$u(\hat{\pi}_i^k) = c_i(\hat{a}_i^{j=\hat{j}_i(k),k}) + u(\hat{\pi}_{i-1}^{j=\hat{j}_i(k)}), \quad (11)$$

$\forall k = 1, 2, \dots, M$ ,  $i = 1, 2, \dots, N$ . At time  $N$ , the values  $u(\hat{\pi}_N^k)$ , computed via (11), will be used for verifying constraint condition (2), as shown below.

3) *Backward recursion:* At time  $N$ , the forward recursion step provides  $M$  representative policies  $\hat{\pi}_N^k$ ,  $k = 1, \dots, M$ , from which we return the highest reward policy  $\hat{\pi}_N^{(\tau)}$  not violating the constraints  $\tau$  in (2). Policy  $\hat{\pi}_N^{(\tau)}$  can be found via a fast back-tracking step from stored values of reverse jumps  $\hat{j}(k)$ , already computed and memorized in forward step (8-10). The optimal state jumps will be traced back recursively, as follows:

$$\begin{aligned} \hat{k}_N^* &= \operatorname{argmax}_{k : u(\hat{\pi}_N^k) \in [\tau_1, \tau_2]} v(\hat{\pi}_N^k), \\ \hat{k}_{i-1}^* &= \hat{j}_i(\hat{k}_i^*), \quad i = N, \dots, 2. \end{aligned} \quad (12)$$

Finally, from (10), the Viterbi's constrained optimal policy  $\hat{\pi}_N^{(\tau)} \triangleq \{\hat{a}_1^{(\tau)}, \dots, \hat{a}_N^{(\tau)}\}$  in (2) is:

$$\hat{a}_i^{(\tau)} = \hat{a}_i^{j=\hat{j}_i(\hat{k}_i^*), k=\hat{k}_i^*}. \quad (13)$$

If there is no  $k$  satisfying the condition on  $u$  in (12), it means that CVA could not find any policy satisfying the constraint condition (2).

4) *Complexity:* The complexity of the linear quantization step is  $O(MN)$  for  $N$  time points (see section II-C1). The complexity of the forward recursion step (8-10) is  $O(KMN)$ , since total number of forward transitions that need updating at each time is  $O(KM)$ . Owing to fast memory-retrieving step, the complexity of the backward recursion is  $O(N)$ . The complexity of CVA is hence  $O(KMN)$ , which is dominated by that of forward recursion.

### E. Monotonic constraint class

There is a special class of constraints, in which the accuracy of the CVA solution increases by exponentially reducing the number of explored policies. Indeed, if constraints are monotonic, at any time point  $i$ , we can reject trajectories that violate the constraints without being forced to wait until the end of the horizon for rejecting such trajectories. More formally, by definition, the accumulated cost  $u$  is called monotonic if it either never decreases nor increases:

$$u(\pi_1) \leq \dots \leq u(\pi_N) \quad \text{or} \quad u(\pi_1) \geq \dots \geq u(\pi_N), \quad (14)$$

in which, by our convention, the inequalities are element-wise for all  $L$ -dimensional vectors  $u(\pi_i)$ ,  $\forall i$ . Assuming without loss of generality the non-decreasing property (14), the constraint (2) can be found violated immediately if, at any time  $i = 1, 2, \dots, N$ , we find that  $u(\pi_i) > \tau_2$  for some specific  $\pi_i$ .

From this remark, we can verify the constraint (2) in tandem with forward recursion of optimization function (8). For later use, let us consider three examples of monotonic constraints. Firstly, if the cost  $c_i$  is always non-negative or non-positive, the accumulated cost satisfies the monotonic property (14). Secondly, a special case of the first is when the accumulated cost can be written as the  $\mathcal{L}_p$ -norm of the cost vector:  $u_i = \sum_{j=1}^i |c_j|^p$ . The third example is when  $u_i$  takes the form of a maximum of the sequence of possibly negative costs:

$$\begin{aligned} u_i &= \max(c_i, \dots, c_1) \\ &= \max(c_i, \max(c_{i-1}, \dots, \max(c_2, c_1))) \\ &= c_i \oplus c_{i-1} \oplus \dots \oplus c_1, \end{aligned} \quad (15)$$

where  $x \oplus y \triangleq \max(x, y)$  and max operator is element-wise for vectors. Written in this way,  $u_i$  still appears as an accumulated cost because the max is a special case of ring-sum in ring theory [23], [24]. Note that this form includes the classical  $\mathcal{L}_\infty$ -norm.

### F. Pseudo-code of Cost-constrained Viterbi Algorithm

In literature, the traditional Viterbi algorithm [24], [25] involves two steps: the first step is to define a max-recursion similar to (8-10), the second step is the back-tracking step similar to (12). The key differences in CVA are the plug-in substitution (7) and the constraint verification step (11,14). We provide the pseudo-code of CVA in Algorithm 1.

## III. CONSTRAINED RESOURCE ALLOCATION FOR SOLAR BASE STATION

### A. System model

In this section, we consider a solar BS as an input-output system (see Fig. 4a). The consumed energy of the BS at the input comes from either solar battery system or electricity grid. The BS consumes in priority the solar energy stored in the battery. At the output, the BS's transmitted power will affect the user's QoS. Because cellular operators have to pay for grid energy, its amount of usage is desired to be kept under a daily threshold, while QoS needs to be maximized.

---

### Algorithm 1 Cost-constrained Viterbi Algorithm

---

#### Inputs:

- a set of actions  $a \in \mathcal{A}$  over  $K$  choices.
- the reward-cost  $b, c$  upon action (Fig. 1)
- constraint thresholds  $\{\tau_1, \tau_2\}$

#### Precompute:

- pick quantization variable  $q$  (section II-C1)
- set number  $M$  of quantization levels

#### Forward recursion:

- For  $i = 1, 2, \dots, N$  {
- evaluate linear quantization (4)
  - evaluate plug-in substitution (7)
  - verify constraint condition (11,14)
  - evaluate max-recursion (8-10)}

#### Backward recursion:

- For  $i = N, \dots, 1$  {
- evaluate back-tracking from memory (12)}

#### Output:

- return the approximate (13) for constrained optimal policy (2)

**Computational complexity:**  $O(KMN)$

---

1) *Energy process:* For studying energy process in BS, let us define the following notations:

- $t$ : time (in hours)
- $P_V$ : harvested solar power, i.e., output power of photovoltaic (PV) panel
- $P_{BS}$ : BS's operational power, an affine function of the transmitted power  $P_{TX}$
- $E_G$ : accumulated energy taken from grid  $E_G(t) \geq 0$
- $E_B$ : accumulated energy stored in battery,  $E_B(t) \geq 0$
- $E_0$ : initial energy stored in battery, i.e.  $E_0 = E_B(0) \geq 0$
- $E_V$ : accumulated harvested solar energy, i.e.  $E_V(t) = \int_0^t P_V(t) dt$
- $E_{BS}$ : accumulated BS's operational energy, i.e.  $E_{BS}(t) = \int_0^t P_{BS}(t) dt$
- $E_U$ : accumulated saving energy level,  $E_U(t) \triangleq E_0 + E_V(t) - E_{BS}(t)$ , i.e., the difference between input and output energy

A typical process is illustrated in Fig. 4b. We have the following relationship:

$$\frac{dE_U(t)}{dt} = \begin{cases} \frac{dE_B(t)}{dt}, & \text{if } E_U(t) > E_{\min}(t), \\ -\frac{dE_G(t)}{dt}, & \text{if } E_U(t) = E_{\min}(t), \end{cases} \quad (16)$$

in which  $\frac{dE_U(t)}{dt} = P_V(t) - P_{BS}(t)$ , the reference point is defined as  $E_{\min}(t) \triangleq [\min_{0 \leq \omega \leq t} E_U(\omega)]^-$  with operator  $[x]^- \triangleq \min(0, x)$ , and we start with initial energy  $E_0$  stored in battery at  $t = 0$ , i.e.,  $E_U(0) = E_0 \geq 0$ . The explanation of (16) is given as follows:

- The case  $E_U(t) > 0$ : Because from 0 to  $t$  the system harvested more energy than it consumed, the grid is not active at time  $t$ . Hence, at that time the BS's energy comes only from the battery, i.e.  $\frac{dE_U(t)}{dt} = \frac{dE_B(t)}{dt}$ . Also, battery is charged if  $\frac{dE_U(t)}{dt} \geq 0$  and discharged if  $\frac{dE_U(t)}{dt} < 0$ . By definition,  $E_U(t) > 0 \geq E_{\min}(t)$  in this case.
- The case  $E_U(t) \leq 0$ : Because from 0 to  $t$  the energy process is deficient, the BS has consumed some energy

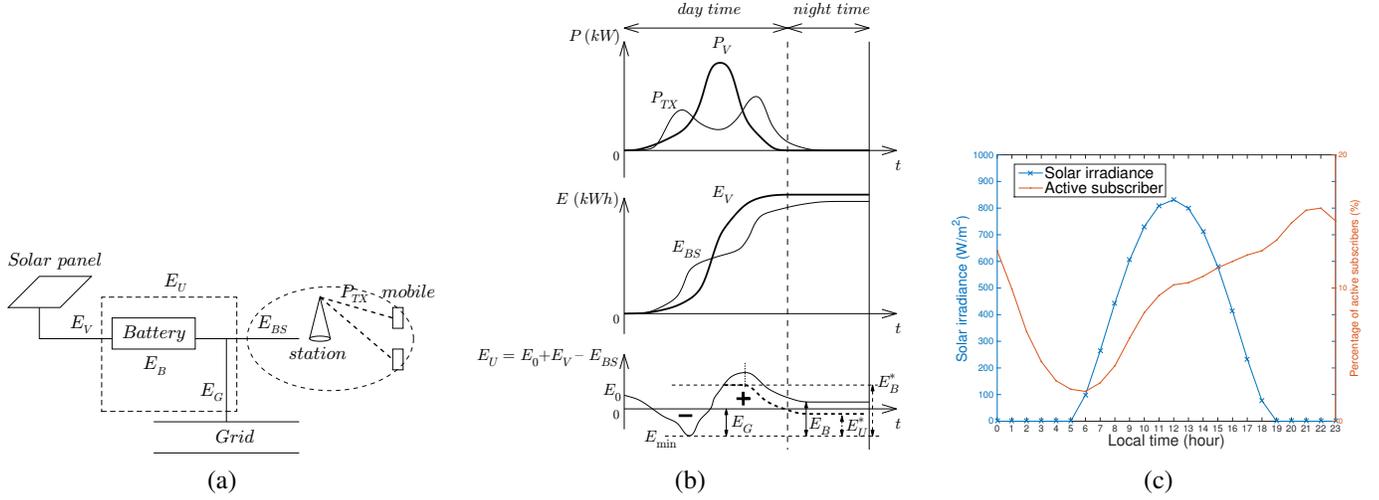


Figure 4. (a) Solar base station model. (b) Illustrative example for power and energy processes. Harvested energy is maximum at noon. There are peaks of user's traffic corresponding to peaks of  $P_{TX}$ . Both  $E_G$ , the total energy taken from the grid, and  $E_B$ , the battery energy level, are computed with respect to  $E_{min}$ , the minimum of the non-positive part of  $E_U$ .  $E_0$  is battery's initial energy. (c) Average solar irradiance estimated for Rome and average percentage  $\chi$ (%) of active subscribers in Europe.

from the grid up to the time  $t_0 \in [0, t]$  where  $E_U(t_0) = E_{min}(t_0)$ . If  $t_0 < t$  and therefore  $E_U(t) > E_{min}(t)$ , more energy has been harvested than consumed in the interval  $[t_0, t]$ . Thus, the BS consumes energy from the battery and this case is similar to the first one. If  $t_0 = t$ , then  $E_U(t)$  is decreasing such that  $E_U(t) = E_{min}(t)$ , the system keeps setting up new reference point and, hence, the grid is active, i.e.,  $\frac{dE_U(t)}{dt} = -\frac{dE_G(t)}{dt}$ .

By integration of (16), this formulation is equivalent to:

$$\begin{aligned} E_G(t) &= |E_{min}(t)|, \\ E_B(t) &= E_U(t) + |E_{min}(t)|. \end{aligned} \quad (17)$$

In literature, the above energy process was studied for the case of  $E_U(t) > 0$  in [11], in which the grid is not available. To the best of our knowledge, this is the first time reference point concept  $E_{min}$  is applied to evaluation of both  $E_G$  and  $E_B$ . In Fig. 4b, we also illustrate the case of limited battery with capacity  $E_B^*$  by dotted version of  $E_U$  process, called  $E_U^*$ .

In the case of unlimited battery, the  $E_U$  process keeps going up at the time of charging battery  $\frac{dE_U(t)}{dt} \geq 0$ . In contrast,  $E_U^*$  stays constant after battery energy level  $E_B$  reaches its limit at  $E_B^*$ , and all harvested solar energy during this period will be lost. During the time of discharging battery, however, both  $E_U$  and  $E_U^*$  processes decrease with the same rate, hence the same decreasing slope during this period. Limited battery may or may not increase the consumed energy from the grid. If  $E_B^*$  is too low, the  $E_U^*$  process may go below the reference point  $E_{min}$  of  $E_U$  and, hence, extra grid energy is required. On the contrary, if  $E_B^*$  is high enough,  $E_U^*$  process will not set up a new reference point  $E_{min}$  and, hence, no extra grid energy is consumed.

2) *Harvested energy model:* In this section, we compute  $P_V$ . Let us assume that photovoltaic (PV) panel is set up parallel with earth's surface. We then employ the well-known Hottel's [26] and Liu-Jordan's [27] models in order to estimate solar irradiance  $G$  (in W/m<sup>2</sup>) perpendicular to earth's surface,

at any time during a day and at any place on earth, as follows [28]–[30]:

$$\begin{aligned} P_V &= \zeta P_{solar} = \zeta G \times A_V, \\ G &= G_{ET} \times (0.271 + 0.706\beta) [\cos(\theta_z)]^+, \end{aligned} \quad (18)$$

in which the generated electric power  $P_V$  of PV is proportional to absorbed solar power  $P_{solar}$ ,  $A_V$  is PV panel's area (in m<sup>2</sup>),  $\theta_z$  is the solar zenith angle and  $[x]^+ \triangleq \max(0, x)$ .

In state-of-the-art PV modules, solar efficiency  $\zeta$  varies from about 10% to about 30% at standard condition [31]. For detail calculation of  $\zeta = \frac{I_P \times V_P}{G}$ , in which  $I_P$  and  $V_P$  are respectively current and voltage in operation of PV module, we will employ the characteristics  $I_P$ - $V_P$  of simplified single-diode circuit of semiconductor PV module [32], [33]. The reason is that calculation for this simplified circuit is tractable and only requires PV module's parameters at standard condition, which are available in most of manufacturer's datasheet. Owing to availability of maximum power point tracking (MPPT) algorithms in commercial products [34], let us assume that  $V_P = V_{MPP}$ , where  $V_{MPP}$  is standard maximum power point voltage, given by manufacturer's datasheet. Then, from Photowatt PW1650-24V datasheet and formula  $I_P$  of solar single-diode circuit, both given in [32], the theoretical solar coefficient for PW1650-24V is  $\zeta = 17.52\%$ . We use this value of  $\zeta$  in our simulations.

The Hottel's atmospheric transmittance  $\beta$  is a function of  $\theta_z$  and local altitude  $H$  (in km). For calculation of  $\beta$ , we will apply the empirical Hottel's 23 km visibility haze model, given in [26], [30], which is reasonable for PV panel set up on top of buildings in a city. The extra-terrestrial solar irradiance,  $G_{ET} \triangleq G_{ET}(\gamma)$ , is a function of day number of the year  $\gamma \in \{1, \dots, 365\}$ , as given in [30], [35]. The value of  $\cos(\theta_z)$  can be geometrically calculated via day number  $\gamma$ , latitude  $\phi$ , longitude  $\varphi$  and time-zone  $T_{GMT}$ , as given in [28], [35]. Let us consider the case of Rome, Italy on 15th August. Substituting  $\{\phi, \varphi, H, T_{GMT}, \gamma\} = \{42^\circ, 12.5^\circ, 0.024, 1, 227\}$  into

formulae in above cited references, we can feasibly calculate hourly values of  $G$  via (18), as plotted in Fig. 4c. Note that, this hourly curve of solar irradiance is consistent with experimental measures at University of Rome Tor Vergata in mid-summer [36], [37].

3) *Consumed energy model for mobile station:* According to a recent study of FP7 program EARTH [38], BS can operate in two separating modes: operating and sleep-mode. The necessary power  $P_{BS}$  for operating BS can be approximately regarded as a linear function of transmitted power  $P_{TX}$ , as follows [38], [39]:

$$P_{BS} = \begin{cases} N_{TRX}(P_0 + slope \times P_{TX}) & , 0 < P_{TX} \leq P_{TX}^{\max} \\ N_{TRX}P_{sleep} & , P_{TX} = 0 \end{cases}. \quad (19)$$

The notation is defined as follows:  $N_{TRX}$  is total number of transceiver chains, which is a product of number of BS sectors, antennas and carriers.  $P_0$  is constant power for equipment's operation.  $slope$  is linear coefficient of load-dependent power transmission.  $P_{sleep}$  is sleep-mode power consumption,  $P_{sleep} < P_0$ .  $P_{TX}^{\max}$  is maximum transmitted power per antenna. Note that, sleep-mode is regarded as the key for energy-saving schemes. In state-of-the-art BS, the operating power is set independent of load, which leaves the potential of saving energy to future BS [38], [40]. For macro BS, the estimated values of these parameters are [38]:  $P_0 = 130$  W,  $slope = 4.7$ ,  $P_{sleep} = 75$  W and  $P_{TX}^{\max} = 20$  W.  $N_{TRX} = 1$  and  $N_{TRX} = 2$  for SISO and MIMO  $2 \times 2$ , respectively. In our simulations, the range of  $P_{TX}$  value is equally divided into  $K$  levels, i.e.  $P_{TX,k} \triangleq \frac{k}{K}P_{TX}^{\max}$ ,  $k = 1, \dots, K$ . Time period  $T$  for switching  $P_{TX}$  level is one hour.

4) *Physical data rates:* By Shannon's theorem, the user's maximum data rate depends on channel bandwidth  $B$  and signal-to-noise ratio, as follows:

$$C(x) = kB \log_2(1 + SNR(x)),$$

$$SNR(x) = N_{TRX} \frac{P_{TX}/PL(x)}{N_0B}, \quad (20)$$

in which  $B$  is channel bandwidth,  $k$  is coefficient gap between practical and optimal coding schemes,  $SNR$  is signal-to-noise ratio dependent on transmitted power  $P_{TX}$  and user's path-loss  $PL(x)$  to BS, and  $N_0$  is standard thermal noise power spectral density. Coding coefficient in (20) was estimated as  $k \approx 0.5$ ,  $k \approx 0.58$  and  $k \approx 0.54$  for SISO, SIMO  $1 \times 2$  and MIMO  $2 \times 2$  systems, respectively [41]. These figures have been obtained by fitting (20) with the outputs of an operator simulator compliant with 3GPP recommendations [42] and in particular take into account fast fading and MIMO multiplexing gain. The path-loss  $PL(x)$  for non-line-of-sight (NLOS) macro BS in urban area and carrier frequency  $f_c = 2.6$  GHz can be evaluated from 3GPP technical report [42]:  $PL(x) = 39.1 \log_{10}(x) + 21.8$  (dB), where  $10 < x < 5000$  in meter. The channel bandwidth  $B = 20$  MHz is also assumed in our simulations. We first consider a single-cell scenario without shadowing and then show results with 6 interferers and shadowing with standard deviation of 6 dB.

5) *User's traffic model:* In this paper, the coverage area  $\mathcal{S}$  of BS is considered as a ring, whose coverage radius is

assumed to be between  $r_a = 10$  m and  $r_b = 1000$  m. Mobile user's location is assumed to be fixed during the call and users appear uniformly inside  $\mathcal{S}$ . In queueing theory, our user's traffic can be modeled as  $M/G/1/\kappa$  processor-sharing queue (i.e. with admission control), in which  $\kappa$  is maximum number of users being served [43], [44]. The utilization factor  $\rho$  in this case can be calculated as follows [45], [46]:

$$\rho = \int_{x \in \mathcal{S}} \rho(x) dx, \quad \text{with} \quad \rho(x) \triangleq \frac{\lambda(x)}{\mu(x)} = \frac{\xi}{C(x)/\sigma}, \quad (21)$$

where  $\lambda$  (user/s) is mean rate of arrival users,  $\xi$  (user/s/km<sup>2</sup>) is mean arrival rate per surface unit,  $\mu$  (user/s) is mean value of serving rate of BS and  $\sigma$  (Mbit/user) is mean size of transmitted data per user. From Shannon's rate expression (20), it is difficult to provide solution for integral (21) in closed form. We calculate (21) via discretization of area  $\mathcal{S}$  in our simulations. The blocking probability  $\text{Prob}_b$  and the mean throughput per user in this case are [45], [47]:

$$\text{Prob}_b = \rho^\kappa \frac{1 - \rho}{1 - \rho^{\kappa+1}},$$

$$R = \psi_S \frac{(1 - \rho)}{\rho} \frac{1 - \rho^\kappa}{1 - (\kappa + 1)\rho^\kappa + \kappa\rho^{\kappa+1}}, \quad (22)$$

where  $\psi_S$  (Mbps) is total data traffic demand inside  $\mathcal{S}$  and  $\psi_S \triangleq \sigma \int_{x \in \mathcal{S}} \lambda(x) dx = \sigma \xi \pi (r_b^2 - r_a^2)$ . In our simulations, the admission control is based on the maximum number  $\kappa$  of active users, such that data rate for any active user is at least  $C_{\min} = 0.1$  Mbps, i.e., we set  $\kappa = \left\lfloor \frac{C(r_b)}{C_{\min}} \right\rfloor$ , with floor function  $\lfloor \cdot \rfloor$ , as done in [47, section 4.1]. Then, we adopt the scenario #1 (ordinary traffic) and scenario #2 (heavy traffic) in EARTH project [38] for average mobile traffic demand in European urban area. All mobile users equally subscribe to 3 operators. However, only active subscribers require data traffic from operator. As extracted from [38, Fig. 4], the population density for urban deployment is assumed 1000 people per km<sup>2</sup> and the hourly percentage  $\chi$  of active subscribers in average is shown in Fig. 4c, whose maximum is 16% at peak hour. From those numbers in [38], we can calculate the hourly user's traffic demand  $\sigma \xi$ . The maximum value of user's traffic demand  $\sigma \xi$  for scenario #1 and #2 is 9.2 Mbps/km<sup>2</sup> and 17.25 Mbps/km<sup>2</sup> at peak hour, respectively. For simplicity, we only consider scenario #2 in subsection IV-C8.

## B. Markovian model for base station

1) *Action and policy definitions:* Let us assume that transmitted power  $P_{TX}$  cannot be changed during the fixed time period  $T$ . If we consider the decision on transmitted power value  $P_{TX}^{[i]}$ , at  $i$ th time period, as an action  $a_i = P_{TX}^{[i]}$ , within a finite set of  $K$  possible values  $\mathcal{A} = \{P_{TX,1}, P_{TX,2}, \dots, P_{TX,K}\}$ , the policy in this case is a sequence of  $P_{TX}$ , as follows:  $\pi_i = \{a_1, \dots, a_i\} = \{P_{TX}^{[1]}, P_{TX}^{[2]}, \dots, P_{TX}^{[i]}\} \in \Pi_i$ .

2) *Quality of service as a reward:* We assume that the reward for our system is a measure of the user's Quality of Service (QoS). In this paper, for a given policy  $\pi_N$ , we consider the user's average throughput  $\bar{R}(\pi_N)$  (in Mbps/user) as the QoS. For this purpose, let us assume that, during  $i$ th time period, BS has served totally  $z_i$  mobile users, whose

average throughput per user is  $R_i(a_i)$  (in Mbps/user) given by (22). The number  $z_i$  of served users is then a function of the blocking rate  $\text{Prob}_b^{[i]}(a_i)$  given by (22), i.e. we have  $z_i = (1 - \text{Prob}_b^{[i]}(a_i))\aleph_i$ , where  $\aleph_i$  denotes number of arrival mobile users during  $i$ th time period. In practice, blocking probabilities  $\text{Prob}_b^{[i]}$  are kept sufficiently small. Then, in this case, we can approximate  $z_i \approx \aleph_i$ . Moreover,  $\aleph_i$  is proportional to  $\chi_i$  in Fig. 4c, so that:

$$\bar{R}(\pi_N) = \frac{\sum_{i=1}^N R_i(a_i) z_i}{\sum_{i=1}^N z_i} \approx \frac{\sum_{i=1}^N R_i(a_i) \chi_i}{\sum_{i=1}^N \chi_i}. \quad (23)$$

In our Markovian model, the reward  $b_i$  can be defined as  $b_i(a_i) = R_i(a_i) \frac{\chi_i}{\sum_{i=1}^N \chi_i}$ , so that the accumulated reward  $v_N$ , as defined in (1), is equal to the overall average throughput  $\bar{R}(\pi_N)$ :

$$v(\pi_N) = \sum_{i=1}^N b_i(a_i) = \bar{R}(\pi_N). \quad (24)$$

Note that, in admission control scenario, the approximation for average throughput (23) was proposed in order to achieve the recursive property for  $\bar{R}(\pi_N)$  (24). With this approximation, the reward can be written as a sum of independent rewards at every time period, so that the reward definition has the desired Markov property.

3) *Blocking rate and grid energy as costs*: In our model, we assume that blocking rate and consumed grid energy are two costs that must be constrained. We will thus consider cost vectors  $c_i = [c_{1,i}, c_{2,i}]$  of  $L = 2$  dimensions,  $i = 1, 2, \dots, N$ . The first cost series  $c_1 = [c_{1,1}, \dots, c_{1,N}]'$  is then defined as a sequence of saving energy amount during  $i$ th time period:  $c_{1,i}(a_i) = E_U^{[i]}(a_i)$  and  $E_U^{[i]}(a_i) \triangleq (P_V^{[i]} - P_{BS}^{[i]}(a_i)) \times T$ , in which  $P_V^{[i]}$  and  $P_{BS}^{[i]}$  via (18-19) are the average harvested solar power and BS power during period  $i$  respectively. The accumulated cost  $u_1 = [u_{1,1}, \dots, u_{1,N}]'$  in this case is equal to accumulated saving energy level:

$$u_1(\pi_N) = E_0 + \sum_{i=1}^N c_{1,i}(a_i) = E_0 + E_U(\pi_N). \quad (25)$$

From (17), the total consumed grid energy is:

$$E_G(\pi_N) = -\min(0, u_1(\pi_1), \dots, u_1(\pi_N)) \quad (26)$$

The second cost value  $c_2 = [c_{2,1}, \dots, c_{2,N}]'$  is defined as  $c_{2,i}(a_i) = \text{Prob}_b^{[i]}(a_i) \frac{\chi_i}{\sum_{i=1}^N \chi_i}$ . The accumulated cost  $u_2$  in this case is equal to the total average blocking rate:

$$u_2(\pi_N) = \sum_{i=1}^N c_{2,i}(a_i) = \bar{P}_b(\pi_N). \quad (27)$$

From equations (25), (26) and (27), we see that costs can be written as the sums of independent costs at every time period and have thus the desired Markov property.

4) *Constrained optimization formulation for BS*: From the Markovian model for BS (24-27), let us consider following optimization problem:

$$\begin{aligned} \pi_N^* &= \underset{\pi_N \in \Pi_N^{(\tau)}}{\text{argmax}} \bar{R}(\pi_N), \\ \text{under two constraints } \pi_N &\in \Pi_N^{(\tau)} : \begin{cases} E_G(\pi_N) \leq \tau_E \\ \bar{P}_b(\pi_N) \leq \tau_b \end{cases}, \end{aligned} \quad (28)$$

where  $\tau_E$  and  $\tau_b$  are desired thresholds of total consumed grid energy and average blocking rate, respectively. If grid energy is not available in practice, we can simply set  $\tau_E = 0$ . This problem can be feasibly solved via Forward-Backward recursion of CVA in section II-D. Also, we recognize that the cost sequences  $E_G(\pi_i)$  and  $\bar{P}_b(\pi_i)$  are monotonic. Hence, we can verify these constraints in tandem with CVA's forward recursion, as explained in section II-E. Note that, we can regard the min operators in (26) as special cases of ring-sum operator, as illustrated in (15), when applying (11).

## IV. SIMULATIONS

In this section, we will consider a macro BS in a European urban environment. The BS consumes energy in hybrid scheme, which includes solar battery and grid energy. The adopted simulation parameters (e.g. channel model, carrier frequency, etc.) are given in section III-A. Our purpose is then to control transmitted power  $P_{TX}$  in order to maximize user's average throughput under two constraints: total consumed grid energy and total blocking rate. Two approaches will be compared to each other: the traditional one keeping  $P_{TX}$  fixed all the time versus our novel CVA presented above.

### A. Allocation methods for transmitted power level

Two approaches (fixed  $P_{TX}$  and CVA) for allocating  $P_{TX}$  level over time will be considered. In fixed  $P_{TX}$  scheme, the  $P_{TX}$  is kept at chosen level for the whole time horizon, regardless of whether there is any active user in the cell. The fixed  $P_{TX}$  level is chosen so as to meet the energy constraint. In CVA scheme, we study three rejection strategies, namely free-blocking-rate, blocking-rate-constraint and minimum-blocking-rate.

- In free-blocking-rate strategy, there is no constraint on the blocking rate<sup>2</sup>. CVA will maximize user's average throughput under constraint of consumed grid energy  $E_G$ . For quantization step in CVA, the energy process  $E_U$  in section III-A1 is uniformly quantized up to  $M = 100$  levels.
- In blocking-rate-constraint, CVA will maximize user's average throughput under two constraints, one for  $E_G$  and the other is 5% for total blocking rate. We quantize the energy process in 100 levels and the blocking rate in 10 levels, i.e.,  $M = 1000$ .
- In minimum-blocking-rate, CVA will minimize blocking rate under constraint of consumed grid energy  $E_G$ . The

<sup>2</sup>Nevertheless, the blocking rate is still computed in order to serve as a tie-breaking rule for policies leading the same user's average throughput.

energy process is quantized with  $M = 1000$  levels. User's throughput is ignored<sup>3</sup>.

For illustration, in CVA scheme, there are  $K = 10$  levels of  $P_{TX}$ , while in fixed  $P_{TX}$  there are  $2K$  levels.

### B. Reference scenario

Reference scenario is a practical and typical case where optimization technique, i.e. CVA, is performing much better than fixed  $P_{TX}$  scheme. In this scenario, SISO system is adopted; the observation's time-length constraint is one day; the size of solar panel is assumed to be  $4 \text{ m}^2$ . The effect of these parameters will be investigated in subsequent scenarios. The fixed  $P_{TX}$  scheme and CVA with free-blocking-rate strategy are then applied and results are shown in Fig. 5.

In Fig. 5a and throughout the paper, the dotted lines are of CVA scheme, the marks without line are of fixed  $P_{TX}$  scheme. Rectangles correspond to the average throughput of accepted users while the crosses correspond to all users (assuming zero throughput for blocked users). The difference between these two curves show the impact of the blocking rate. We can see that CVA yields significant improvement, compared to fixed  $P_{TX}$  when  $E_G$  constraint is tight. CVA is able to meet the constraint with respect to  $E_G$ .

Fig. 5b shows the evolution of accumulated saving energy across the day for CVA and two fixed  $P_{TX}$  schemes (namely with minimum and maximum power). The five curves of CVA correspond to five energy constraints taken from Fig. 5a. CVA curve's slopes vary between the fixed minimum and maximum  $P_{TX}$  curve's slopes according to the chosen  $P_{TX}$  level. At first hours of the day (from 0:00 to 07:00), energy is mostly consumed from the grid and hence  $P_{TX}$  is adjusted according to the constraint. During daylight (from 07:00 to 18:00), energy is harvested from the sun and is sufficient to transmit at higher power levels or even full power. At night (from 18:00 to 24:00), the energy harvested during daylight is sufficient to transmit at full power. Note that the system has an empty battery at 00:00 but has stored some harvested energy at 24:00 that can be used for the next day.

Fig. 5c is a histogram of  $P_{TX}$  levels for the five cases (from left to right, respectively) of CVA scheme in Fig. 5a. Most of the time CVA uses full power. When the energy constraint is tighter, CVA switches to lower  $P_{TX}$  level in order to save up energy.

### C. Case-studies

1) *Effect of time horizon*: The setting in this scenario is the same as reference scenario, except that the time horizon is set to 1 and 30 days, corresponding to right and left set of curves, respectively, in Fig. 6a. Also, we assume a monthly energy constraint for 30 days curves. For comparison with 1 day, we however scale the x-axis to a daily constraint in Fig. 6a. Fig. 6b shows the energy process for the case of 30 days. For clarity, Fig. 6c is the zoomed version of Fig. 6b by excluding the fixed  $P_{TX}$  curves.

In Fig. 6a, all curves are moved to the left when horizon increases, i.e., the average consumed grid energy per day is reduced when time horizon increases. The reason is that the system stores some harvested solar energy in the battery from one day to the next one (see Fig. 5b). With fixed  $P_{TX}$  scheme, the curve for 30 days has two parts: a steep increase and a lower slope. When  $E_U$  is higher than zero at the end of the first day (as for  $P_{TX}^{(\min)}$  in Fig. 5b), no new grid energy is required in the following days because the average input energy is higher than the average output one. Therefore, the grid energy constraint is used only for the first hours of the first day when the battery is empty. This corresponds to the steep increase part. When  $E_U$  is less than zero at the end of the first day, we require new grid energy everyday. Therefore consumed grid energy scales with the time horizon. This corresponds to the lower slope. With a 30 day horizon, CVA still outperforms the fixed  $P_{TX}$  scheme. When the constraint is tighter and tighter, fixed  $P_{TX}$  sees performance dropping while CVA maintain a high level of throughput.

In Fig. 6b, we see that there are two modes for the fixed  $P_{TX}$  schemes. Either there is enough harvested energy compared to the consumed energy and the saved energy  $E_U$  is growing indefinitely, or there is not enough and more and more grid energy is consumed (see minimum and maximum fixed  $P_{TX}$  curves respectively). There is thus a best fixed  $P_{TX}$  that balances input and output energy processes on the long term.

In Fig. 6c, we observe two behaviors. If the constraint is loose, the system takes energy from the grid every day and we observe a decreasing trend towards the constraint. If the constraint is tight (almost zero), during the first half of the month, the system save energy every day in the battery and releases it in the second half of the month to meet the constraint after 30 days. In contrast fixed  $P_{TX}$  policies either save or release energy every day without being able to modulate between the days. This flexibility in tight constraint helps CVA increase the user's average throughput by about 40%, from 9.3 Mbps in fixed  $P_{TX}$  scheme to 12.8 Mbps in CVA.

2) *Effect of panel dimensioning*: In this scenario, we study the effect of the solar panel size. For illustration, we assume the reference scenario with time-horizon of one day and the minimum-blocking-rate strategy. Solar panel's size is set to 0, 2 and  $4 \text{ m}^2$  (i.e from right to left set of curves in Fig. 7a, respectively). We first observe that reducing the panel size is equivalent to requiring more grid energy as curves are moving to the right. If the panel is small the gain of CVA is also small with respect to fixed  $P_{TX}$  policy because there is less room for regulation between days. From 0 to  $4 \text{ m}^2$ , we observe important gains of CVA with respect to fixed  $P_{TX}$ . Above  $4 \text{ m}^2$ , harvested energy is more than sufficient and there is no further gains. The blocking rate for CVA is also very small in this case, in which the largest one is 5.9% (vs. 50% for fixed  $P_{TX}$ ) while the rest is almost zero. As conclusions for this subsection, the panel size has to be chosen according to the grid energy constraint, CVA outperforms fixed  $P_{TX}$  scheme in all cases, and even if there is no solar panel, there is still room for optimization.

<sup>3</sup>It is however still computed to serve as a tie-breaking rules for policies leading the same blocking rate.

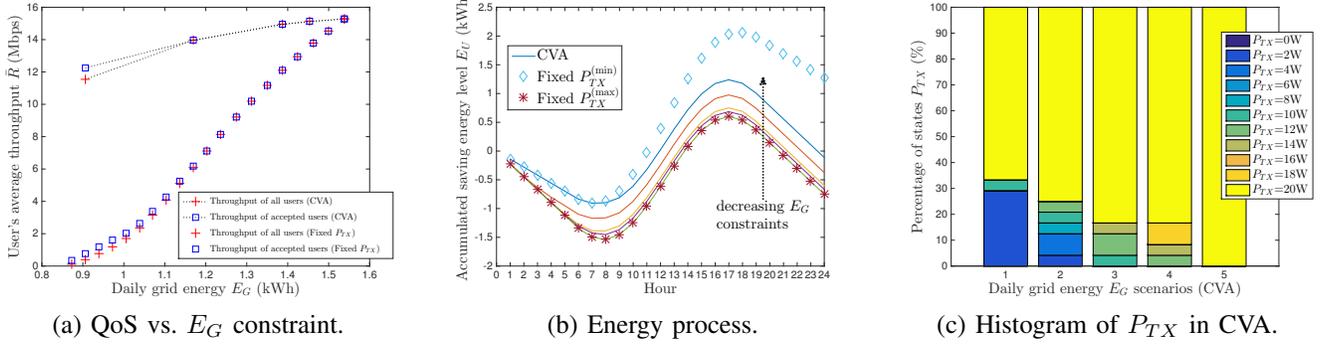
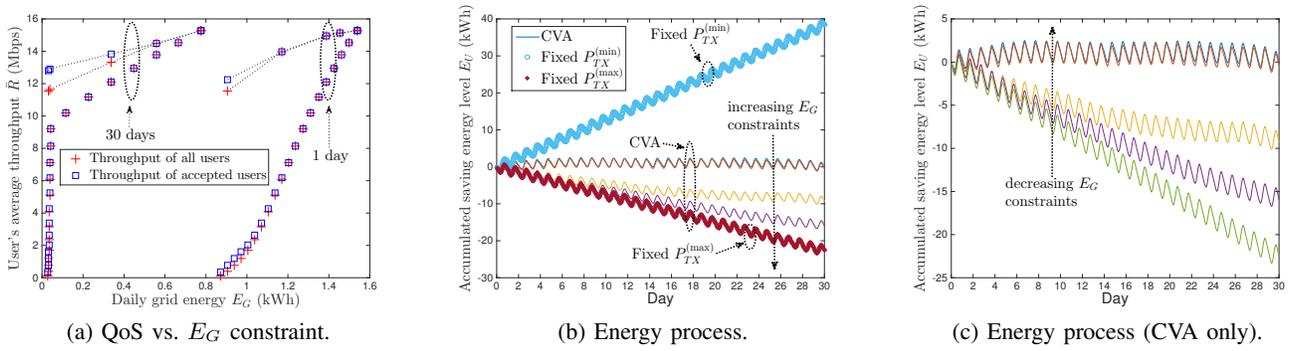
Figure 5. CVA vs. fixed  $P_{TX}$  scheme (reference scenario, free-blocking-rate strategy).

Figure 6. Effect of different time horizons (reference scenario, free-blocking-rate strategy).

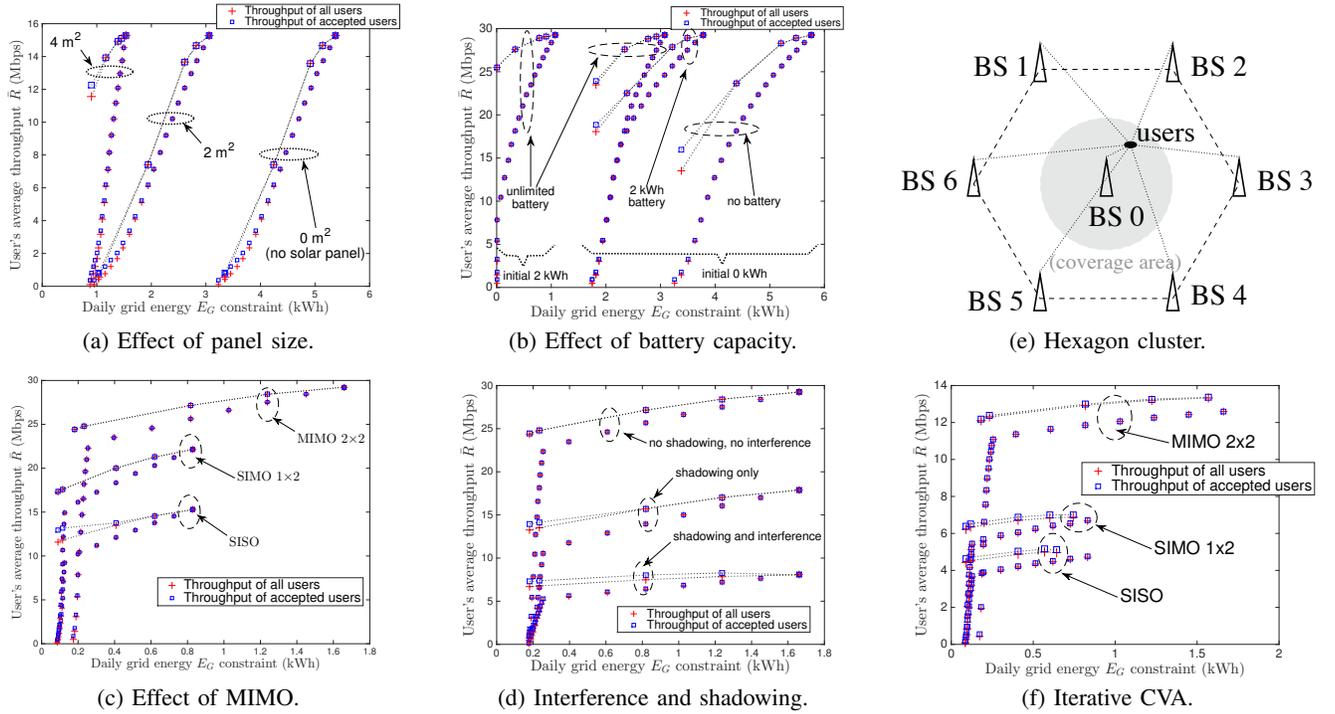


Figure 7. (a-d) Effect of panel size, battery capacity, MIMO, interference and shadowing. (e-f) Iterative CVA scheme for multi BSs

3) *Effect of battery size:* Similar to previous scenario, we now assume MIMO  $2 \times 2$  and a panel size of  $8 \text{ m}^2$ . The effect of varying battery size and non-empty initial battery ( $E_0 > 0$ ), as explained in section III-A1, is shown in Fig. 7b. We first observe that, similar to panel dimensioning, varying battery size moves curves horizontally: obviously, when battery size increases, less grid energy is required. CVA clearly outperforms fixed  $P_{TX}$ . Also, CVA works better with large batteries because there is more freedom for power regulation. Having an initial non-empty battery has a similar effect because this stored energy is used during first hours of the process when daily light is not available. The case of unlimited battery allows us to see that a battery level of about 4.3 kWh is sufficient. By cutting off battery capacity down to almost a half at 2 kWh or even by removing completely the battery, a significant amount of harvested energy is lost and more energy is required from the grid.

4) *Effect of MIMO system:* In this subsection, we study the influence of SIMO  $1 \times 2$  and MIMO  $2 \times 2$ . A 10 days time-horizon, no sleep-mode, free-blocking-rate strategy are assumed. Panel sizes of  $4 \text{ m}^2$ ,  $4 \text{ m}^2$  and  $8 \text{ m}^2$  are assumed for SISO, SIMO and MIMO cases, respectively. Results are shown in Fig. 7c. We first see how MIMO increases the average throughput but requires more energy. Moreover, the blocking rate is reduced with MIMO (almost zero). The reason is that the traffic demand is the same while the serving rate is increased. CVA in MIMO case outperforms fixed  $P_{TX}$  scheme and the relative gain with respect to fixed  $P_{TX}$  is similar to the SISO case. The interesting case is SIMO, which compromises between SISO and MIMO. Given the same amount of harvested energy, the average throughput in SIMO is significantly higher than that in SISO, while the blocking rate in SIMO case is dramatically lower (almost zero). This is owing to the increase in serving rate in SIMO, without the need of increase in energy like MIMO. Although the average throughput of SIMO is not as high as that of MIMO, we suggest that SIMO should be preferred in application of CVA, owing to its efficiency in energy per bit.

5) *Effect of shadowing and interference:* Given the same MIMO  $2 \times 2$  case in previous subsection, we wish to study the effect of shadowing and interference. At equal angles around our central BS, we set up six extra macro BSs as interferers, whose inter-distance to central BS is two times the cell radius  $r_b$ , as shown in Fig. 7e. For simplicity, we assume that interferers transmit at 10 W, i.e., half of the maximum power. For every received power, we consider an independent shadowing of  $\sigma_s = 6 \text{ dB}$  [42] in the path-loss formula in (20). The best server policy is employed, i.e., a user is served by the BS providing the highest received power. Results in this case are shown in Fig. 7d. As expected, shadowing effect reduces QoS significantly, down to about 50% of QoS in no shadowing case. The interference effect then further reduces QoS down to about 50% of QoS in shadowing case. Nevertheless, we again observe that, for all cases, the CVA always yields better QoS than fixed  $P_{TX}$  for our central BS under the same constraint of grid energy. The rejection rate in CVA is also sufficiently low.

6) *Effect of iterative CVA in BS network:* Let us now study a network of BSs as shown in Fig. 7e. The hexagon area is served by seven solar BSs (the coverage radius is nearly doubled with  $r_b = 1990 \text{ m}$  compared to previous cases). In such a single frequency network, every BS tries to optimize its QoS, while generating interference on other BSs and thus influencing their own QoS. To solve this issue, we propose an iterative CVA, which iteratively applies CVA to each BS one-by-one while keeping the  $P_{TX}$  trajectories of all other BSs unchanged. Along the iteration steps, the network average QoS never decreases. Indeed if we fix  $P_{TX}$  policies of all BSs but one and tune network QoS via the remaining  $P_{TX}$  policy, the QoS can only improve because CVA provides a constrained near-optimal policy. Hence, iterative CVA is guaranteed to converge to a local maximum for the QoS of the whole network.

In simulations, we observed that not more than four rounds of iteration are required for convergence, given that our initialization is fixed  $P_{TX}$  scheme of the same grid energy constraint and each round consists of one iteration for each BS. In Fig. 7f, we assume that the harvested solar energy process and daily grid consumed energy constraint are the same for all BSs. Users are served by their best server, which explains that the MIMO's average throughput in this case is better than that of shadowing and interference case in Fig. 7d. The SISO and SIMO, however, have much lower throughput than those in Fig 7c, owing to shadowing and interference in hexagon area. In all cases, iterative CVA outperforms the fixed  $P_{TX}$  scheme. This is due to the fact that the iterative CVA is now able to regulate the  $P_{TX}$  of all interferers in order to balance the other-cell interference. As shown in Fig. 7f, given the same average user's throughput, iterative CVA can save up to 85% of grid consumed energy, in comparison with fixed  $P_{TX}$  scheme. Also, given the same grid energy constraint, all user's throughput in iterative CVA are higher (up to 25%) than that of fixed  $P_{TX}$  scheme.

7) *Effect of sleep-mode:* We now allow the BS to go in sleep-mode, i.e., whatever the policy (CVA or fixed  $P_{TX}$ ), the BS consumes less energy when there is no user to serve. We assume a 10 days time-horizon, a panel size of  $2.8 \text{ m}^2$  and a minimum-blocking-rate strategy. Results are shown in Fig. 8a,b. As with sleep-mode we need less energy, the grid energy constraint can be much tighter to achieve the same throughput. As a consequence, we can reduce the panel size with respect to the reference scenario: about  $3 \text{ m}^2$  are sufficient to run a BS with sleeping-mode and without grid energy, which explains why we picked a slightly smaller panel for this scenario. An interesting observation is that CVA achieves much higher gains in presence of sleep-mode when the constraint is tight and almost no gain when the constraint is loose. The high gain can be explained as follows. CVA offers higher throughputs than fixed  $P_{TX}$ , so that users leave the system quicker. As a consequence the proportion of time during which there is no user in the system is higher (see Fig. 8b) and the BS switches to sleep-mode more often (from 10% for fixed  $P_{TX}$  to 40% for CVA in the tightest constraint scenario).

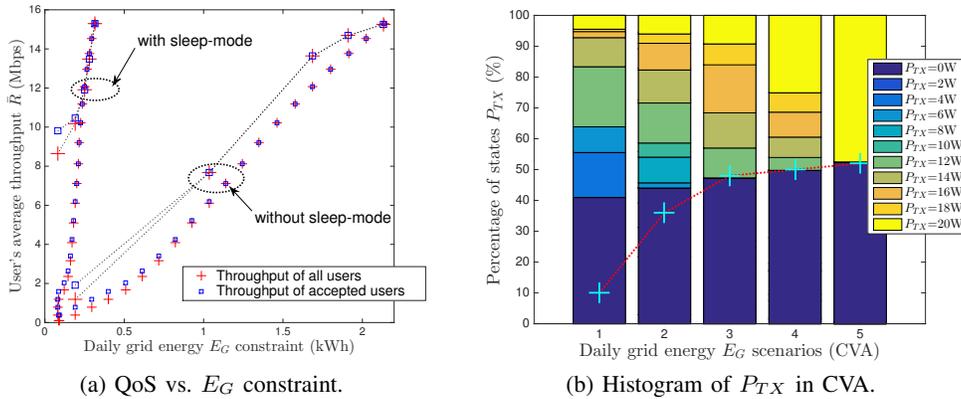


Figure 8. Effect of sleep-mode (“+” correspond to proportions of sleep-mode in fixed  $P_{TX}$ ).

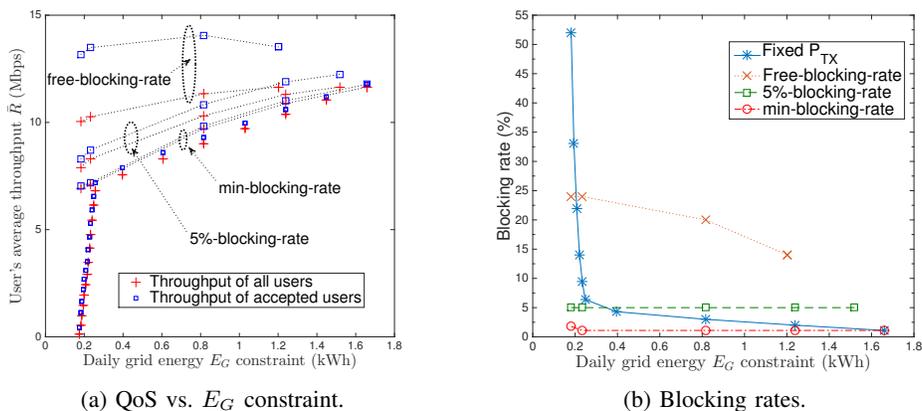


Figure 9. Effect of rejection policy.

	$P_{TX}$ levels (W)	Grid energy (kWh)	Throughput of all users (Mbps)	Throughput of accepted users (Mbps)	Blocking rate (%)	Running time (s)
Fixed $P_{TX}$	[2, 20]	[0.9, 1.5]	[0.5, 15.4]	[0.7, 15.4]	[49, 0]	0.0001
	[2, 10, 20]	[0.9, 1.1, 1.5]	[0.5, 6.2, 15.4]	[0.7, 6.3, 15.4]	[49, 1.3, 0]	0.0001
	[2, 6, 12, 20]	[0.9, 1.0, 1.2, 1.5]	[0.5, 2.3, 8.2, 15.4]	[0.7, 2.6, 8.2, 15.4]	[49, 11, 0, 0]	0.0001
Brute-force	[2, 20]	[0.9, 1.5]	[11.3, 15.4]	[12, 15.4]	[5.9, 0]	3879
	[2, 10, 20]	[0.9, 1.1, 1.5]	[11.4, 13.7, 15.4]	[12, 13.7, 15.4]	[5.6, 0, 0]	110
	[2, 6, 12, 20]	[0.9, 1.0, 1.2, 1.5]	[11.4, 12.6, 14.4, 15.4]	[12, 12.6, 14.4, 15.4]	[5.6, 0, 0, 0]	3471
CVA	[2, 20]	[0.9, 1.5]	[11.3, 15.4]	[12, 15.4]	[5.9, 0]	0.007
	[2, 10, 20]	[0.9, 1.1, 1.5]	[11.4, 13.7, 15.4]	[12, 13.7, 15.4]	[5.6, 0, 0]	0.005
	[2, 6, 12, 20]	[0.9, 1.0, 1.2, 1.5]	[11.4, 12.6, 14.4, 15.4]	[12, 12.6, 14.4, 15.4]	[5.6, 0, 0, 0]	0.011

Table I  
SIMULATION RESULTS FOR REFERENCE SCENARIO WITH DIFFERENT POSSIBLE STATES OF  $P_{TX}$  (W).

8) *Effect of rejection strategy*: We now study the effect of different rejection strategies. We use the same assumption as in the previous subsection with MIMO  $2 \times 2$ . In order to stress the system, we increase the traffic and assume the scenario #2, in which 50% of active users are heavy users. With this high traffic demand, we compare the three rejection strategies for CVA. The result is shown in Fig. 9a,b. We first observe that CVA always outperforms fixed  $P_{TX}$  in terms of user throughput. There is however a trade-off between the gain and the blocking rate: as the blocking rate constraint is tighter (from free to 5% to minimum strategies on the figures), the throughput gain is lower (see Fig. 9a). The free-blocking-rate strategy yields the most throughput gain for CVA, but with high blocking rate (from 24% to 14% for top set of curves in

Fig. 9b). The blocking-rate-constraint meets the constraint at 5% blocking rate, but yields lower gain in throughput. Lastly, the minimum-blocking-rate strategy yields the least blocking-rate, around 1.1% in Fig. 9a,b, but has little throughput gain over traditional fixed  $P_{TX}$  scheme. When the constraint is very tight, this strategy is very interesting as both throughput and blocking rate gains are huge (see the first left points of the two figures). As a conclusion, CVA offers the flexibility to control and balance user throughput and blocking rate.

9) *Comparison with brute-force method*: We now consider a very simple scenario and compare the output of CVA with fixed  $P_{TX}$  scheme and the optimal policy obtained with the brute-force approach. In order to make it computable, we only consider the case of two, three and four possible  $P_{TX}$  values,

as shown in Table I, over a single day. The number of policies is  $2^{24}$ ,  $3^{12}$  and  $4^{12}$  for the case of two possible  $P_{TX}$  values over 24 slots of one hour period, three and four possible  $P_{TX}$  values over 12 slots of two hours period, respectively. The grid energy constraint  $E_G$  kWh is set so as fixed  $P_{TX}$  meets the constraint by always using only one  $P_{TX}$  state all the time. In all cases, brute-force and CVA provide the same performance and clearly outperform fixed  $P_{TX}$  scheme in terms of throughput and blocking rate. From numerical results we observe that indeed the optimal policy is returned by CVA in all cases. Running time is computed using MATLAB and a commercial PC (Intel i7 CPU, 2.67 GHz) and shows clear advantage of CVA.

## V. CONCLUSION

In this paper, we tackled generic optimization problems under multiple constraints, whose objective and constraint functions are based on  $\mathcal{L}_p$ -norm. We formulated this kind of problems as a quantized Markovian cost-reward process. Next, we designed a low complexity near-optimal optimization algorithm, called Cost-constrained Viterbi Algorithm, to solve these problems.

We applied this framework to resource allocation in hybrid mobile base stations. For the first time, we formulate the hybrid energy process using  $\mathcal{L}_p$ -norm. Our algorithm was shown to be robust, flexible and near-optimal in practical settings of solar mobile base station.

We showed in which practical scenarios optimization is particularly interesting compared to fixed transmit power policies: when the constraints are tight (grid energy, solar panel size, short time horizon, blocking rate, battery size) or the system is stressed (high traffic demand), our Cost-constrained Viterbi Algorithm achieves huge gains in user's average throughput and blocking rate. In a more typical scenario, about 40% gain in user throughput can be achieved under the same constraint of grid energy for single BS and about 85% reduction in grid energy can be achieved under the same constraint of user throughput for hexagon network of BSs.

## REFERENCES

- [1] G. Piro, M. Miozzo, G. Forte, N. Baldo, L. A. Grieco, G. Boggia, and P. Dini, "HetNets powered by renewable energy sources: Sustainable next-generation cellular networks." *IEEE Internet Computing*, vol. 17, no. 1, pp. 32–39, Feb. 2013.
- [2] C. Mcguire, M. R. Brew, F. Darbari, S. Weiss, and R. Stewart, "WindFi - a renewable powered base station for rural broadband," *IEEE 19th Int. conference on systems, signals and image processing (IWSSIP)*, pp. 265–268, Apr. 2012.
- [3] T. Han and N. Ansari, "On optimizing green energy utilization for cellular networks with hybrid energy supplies," *IEEE Trans. on Wireless Communications*, vol. 12, no. 8, pp. 3872–3882, Aug. 2013.
- [4] Y. Mao, J. Zhang, and K. B. Letaief, "Grid energy consumption and QoS tradeoff in hybrid energy supply wireless networks," *IEEE Trans. on Wireless Communications*, vol. 15, no. 5, pp. 3573–3586, May 2016.
- [5] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with hybrid energy harvesting base station," *IEEE Trans. on Wireless Communications*, vol. 12, no. 7, pp. 3412–3427, Jun. 2013.
- [6] I. Ahmed, A. Ikhlef, D. W. K. Ng, and R. Schober, "Power allocation for an energy harvesting transmitter with hybrid energy sources," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 6255–6267, Dec. 2013.
- [7] J. Gong, J. S. Thompson, S. Zhou, and Z. Niu, "Base station sleeping and resource allocation in renewable energy powered cellular networks," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 3801–3813, Nov. 2014.
- [8] J. Xu, Y. Guo, and R. Zhang, "CoMP meets energy harvesting: A new communication and energy cooperation paradigm," *IEEE GLOBECOM*, pp. 2508–2513, Dec. 2013.
- [9] S. Hu, Y. Zhang, X. Wang, and G. B. Giannakis, "Weighted sum-rate maximization for MIMO downlink systems powered by renewables," *IEEE Trans. on Wireless Communications*, vol. 15, no. 8, pp. 5615–5625, Aug. 2016.
- [10] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
- [11] A. Huszak, "Bandwidth and consumption controller algorithms for solar-powered base stations," *International Conference on Communications and Information Technology (ICCIT)*, pp. 336–340, Jun. 2012.
- [12] Y. Xiao, K. Thulasiraman, G. Xue, and A. Juttner, "The constrained shortest path problem: algorithmic approaches and an algebraic study with generalization," *AKCE Int. J. of Graphs and Combinatorics*, vol. 2, no. 2, pp. 63–86, Nov. 2005.
- [13] S. Chen, M. Song, and S. Sahni, "Two techniques for fast computation of constrained shortest paths," *IEEE/ACM Trans. on networking*, vol. 16, no. 1, pp. 105–115, Feb. 2008.
- [14] A. Swami, Q. Zhao, Y.-W. Hong, and L. Tong, *Wireless Sensor Networks: Signal Processing and Communications*. Wiley, 2007.
- [15] D. V. Djonin and V. Krishnamurthy, "MIMO transmission control in fading channels - a constrained Markov Decision Process formulation with monotone randomized policies," *IEEE Trans. on Signal Processing*, vol. 55, no. 10, pp. 5069–5083, Oct. 2007.
- [16] A. Zadorojnyi, G. Even, and A. Shwartz, "A strongly polynomial algorithm for controlled queues," *Journal Mathematics of Operations Research*, vol. 34, no. 4, pp. 992–1007, Nov. 2009.
- [17] C. Caramanis, N. B. Dimitrov, and D. P. Morton, "Efficient algorithms for budget-constrained Markov Decision Processes," *IEEE Trans. on Automatic Control*, vol. 59, no. 10, pp. 2813–2817, Oct. 2014.
- [18] J. Li, A. Najmi, and R. M. Gray, "Image classification by a two-dimensional hidden Markov model," *IEEE Trans. on Signal Processing*, vol. 48, no. 2, pp. 517–533, Feb. 2000.
- [19] L. Cao and C. W. Chen, "A novel product coding and recurrent alternate decoding scheme for image transmission over noisy channels," *IEEE Trans. on Communications*, vol. 51, no. 9, pp. 1426–1431, Sep. 2003.
- [20] K. Thulasiraman and M. N. S. Swamy, *Graphs: theory and algorithms*. John Wiley and Sons, 1992.
- [21] J. Max, "Quantizing for minimum distortion," *IRE Trans. on Information Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.
- [22] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. on Information Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.
- [23] S. M. Aji and R. J. McEliece, "The generalized distributive law," *IEEE Trans. on Information Theory*, vol. 46, no. 2, pp. 325–343, Mar. 2000.
- [24] V. H. Tran, "Variational inference in digital receivers," Ph.D. dissertation, Trinity College Dublin, 2014.
- [25] G. Forney, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.
- [26] H. C. Hottel, "A simple model for estimating the transmittance of direct solar radiation through clear atmospheres," *Solar Energy*, vol. 18, no. 2, pp. 129–134, Dec. 1976.
- [27] B. Y. Liu and R. C. Jordan, "The interrelationship and characteristic distribution of direct, diffuse and total solar radiation," *Solar Energy*, vol. 4, no. 3, pp. 1–19, Jul. 1960.
- [28] Y. Bao, X. Wang, X. Liu, S. Zhou, and Z. Niu, "Solar radiation prediction and energy allocation for energy harvesting base stations," *IEEE International Conference on Communications (ICC)*, Jun. 2014.
- [29] J. A. Duffie and W. A. Beckman, *Solar engineering of thermal processes*, 4th ed. Wiley, 2013.
- [30] M. M. Khan and M. J. Ahmad, "Estimation of global solar radiation using clear sky radiation in Yemen," *Journal of Engineering Science and Technology Review*, vol. 5, no. 2, pp. 12–19, Aug. 2012.
- [31] M. A. Green, K. Emery, Y. Hishikawa, W. Warta, and E. D. Dunlop, "Solar cell efficiency tables (version 45)," *Progress in Photovoltaics: Research and applications*, no. 23, pp. 1–9, Jan. 2015.
- [32] A. Bellini, S. Bifaretti, V. Iacovone, and C. Cornaro, "Simplified model of a photovoltaic module," *Applied Electronics Conference*, pp. 47–51, Sep. 2009.
- [33] D. Rekioua and E. Matagne, *Optimization of photovoltaic power systems - modelization, simulation and control*. Springer, 2012.

- [34] A. Bellini, S. Bifaretti, and V. Iacovone, "MPPT algorithm for current balancing of partially shaded photovoltaic modules," *IEEE International Symposium on Industrial Electronics (ISIE)*, pp. 933–938, Jul. 2010.
- [35] S. A. R. Zaidi, M. Ghogho, D. C. McLernon, and A. Swami, "Energy harvesting empowered cognitive metro-cellular networks," *1st International Workshop on Cognitive Cellular Systems (CCS)*, pp. 1–5, Sep. 2014.
- [36] A. Spina and C. Cornaro, "Global, direct and diffuse radiation measurements at ground by the new environmental station of the University of Rome Tor Vergata," in *International congress Energy, environment and technological innovation*, vol. 5, Jan. 2004.
- [37] A. Spina, C. Cornaro, and S. Serafini, "Outdoor ESTER test facility for advanced technologies PV modules," *33rd IEEE Photovoltaic Specialists Conference*, pp. 1–5, May 2008.
- [38] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermarck, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Wireless Communications Magazine*, vol. 18, no. 5, pp. 40–49, Oct. 2011.
- [39] G. Auer, V. Giannini, I. Godor, P. Skillermarck, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, C. Desset, and O. Blume, "Cellular energy efficiency evaluation framework," *73rd IEEE Vehicular Technology Conference*, pp. 1–6, May 2011.
- [40] J. Wu, Y. Zhang, M. Zukerman, and E. K.-N. Yung, "Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 2, pp. 803–826, Feb. 2015.
- [41] M. K. Karray and M. Jovanovic, "Theoretically feasible QoS in a MIMO cellular network compared to the practical LTE performance," *Proc. of ICWMC*, pp. 49–55, Jun. 2012.
- [42] *3GPP TR 36.814 V9.0.0 (2010-03), Further advancements for E-UTRA physical layer aspects (Release 9)*, Std. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/36814.htm>
- [43] F. P. Kelly, *Reversibility and stochastic networks*. John Wiley and Sons, 1979.
- [44] D. Denisov and A. Sapozhnikov, "On the distribution of the number of customers in the symmetric M/G/1 queue," *Queueing Systems*, vol. 54, no. 4, pp. 237–241, Dec. 2006.
- [45] Y. Wu, C. Williamson, and J. Luo, "On processor sharing and its applications to cellular data network provisioning," *Performance Evaluation*, vol. 64, no. 9-12, pp. 892–908, Jun. 2007.
- [46] M. K. Karray and M. Jovanovic, "A queueing theoretic approach to the dimensioning of wireless cellular networks serving variable-bit-rate calls," *IEEE Trans. on Vehicular Technology*, vol. 62, no. 6, pp. 2713–2723, Jul. 2013.
- [47] T. Bonald and A. Proutiere, "Wireless downlink data channels: User performance and cell dimensioning," *Proceedings of the 9th annual international conference on Mobile computing and networking (MobiCom)*, pp. 339–352, Sep. 2003.



**Marceau Coupechoux** is Associate Professor at Telecom ParisTech since 2005 and Professeur Chargé de Cours at Ecole Polytechnique since 2016. He obtained his Master from Telecom ParisTech (1999) and from University of Stuttgart (2000), his Ph.D. from Institut Eurecom (2004), his Habilitation from University Pierre et Marie Curie (2015). From 2000 to 2005, he was with Alcatel-Lucent (in Bell Labs former Research & Innovation and then in the Network Design department). He was Visiting Scientist at the Indian Institute of Science, Bangalore, India, in 2011-2012. He has been General Co-Chair of WiOpt 2017. In the Computer and Network Science department of Telecom ParisTech, he is working on cellular networks, wireless networks, ad hoc networks, cognitive networks, internet of things, focusing mainly on performance evaluation, optimization and resource management.



**Viet Hung Tran** has been a lecturer at Hochiminh city University of Technology, Vietnam, since 2008. He received his Master degree from ENS Cachan in Paris, France, in 2009 and his Ph.D. degree from Trinity College Dublin, Ireland, in 2014. From 2014 to 2016, he was a postdoc at Telecom ParisTech. His research interest is optimal algorithms for Bayesian learning network and information theory.