Asynchronous Systems for Constraint Satisfaction: Filtering and Stability

Tarek A. Lahlou and Thomas A. Baran
Digital Signal Processing Group
Massachusetts Institute of Technology

Abstract—We discuss the formulation of, and stability conditions for, a set of generally asynchronous signal processing systems for solving a class of constraint satisfaction problems. Problems within this class are specifically those where a quadratic conservation principle is known to exist, as commonly occurs in stationarity conditions associated with a variety of convex and nonconvex optimization problems. With the intent of addressing a wide range of system architectures, the presented stability results are formulated for use with both a generally asynchronous update protocol, modeled as sample-and-hold subsystems triggered by independent Bernoulli processes, and also first-order filtering of the asynchronous updates. Numerical examples are provided by illustration and reference, indicating system behavior consistent with the presented stability results.

Index Terms—asynchronous signal processing systems, constraint satisfaction problems, filtering, stability, conservation

I. INTRODUCTION

In a variety of large-scale data processing applications, a key goal is to determine a solution satisfying a particular set of constraints. The PageRank algorithm [1] is one notable example consistent with this. The authors have also recently described a class of asynchronous optimization methods that essentially reduce to algorithms for satisfying constraints corresponding to associated stationarity conditions [2]. In these and other applications, e.g. [3]–[5], the use of an *asynchronous* algorithm in determining a solution to the specific problem at hand is especially important, in particular as the storage, retrieval and processing of data within these applications becomes increasingly distributed across large-scale networks. This is critical in dynamically-changing or heterogeneous networks, where synchronized communication between processing nodes can be impractical.

As asynchronous algorithms become increasingly used for distributed and generally uncoordinated data processing, it will in particular be increasingly important to identify sufficient conditions under which these systems are guaranteed to converge. This paper outlines a set of theorems describing such conditions, with associated proofs being available in [6], pertaining to a class of constraint satisfaction problems where a quadratic conservation principle is known to exist. For this class of constraint satisfaction problems, which we specifically refer to as "conservative constraint satisfaction problems" (CCSPs), the associated algorithm is readily described using an asynchronous signal-flow structure, containing linear and nonlinear elements interconnected through randomly-triggered sample-and-hold subsystems, possibly including first-order asynchronous filtering of the updates. The conditions for stability discussed in this paper are in turn stated in terms of system properties associated with various aggregate properties of the overall system, which are straightforward to identify in a variety of cases including many cases commonly encountered in practice.

We begin in Section II by formally defining the class of CCSPs considered herein and, in Section III, we define a companion class of signal processing systems that can be used to solve them. We establish the necessary machinery related to convergence of asynchronous systems in Section IV in anticipation of the stability analysis in Section V. We close by making several key connections with optimization theory in Section VI and present some numerical experiments in Section VII.

II. CONSERVATIVE CONSTRAINT SATISFACTION PROBLEMS

Constraint satisfaction problems (CSPs) are traditionally defined as a 3-tuple $\langle \mathcal{V}, \mathcal{D}, \mathcal{C} \rangle$ consisting of a set of variables denoted $\mathcal{V}$, a set of corresponding domains $\mathcal{D}$ over which the variables are defined, and a set of constraints between the variables denoted $\mathcal{C}$. These may be written formally as

$$\mathcal{V} = \{v_1, \ldots, v_n\}, \quad \mathcal{D} = \{\mathcal{D}_1, \ldots, \mathcal{D}_n\}, \quad \mathcal{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_r\}, \quad (1)$$

with each variable v_j satisfying $v_j \in \mathcal{D}_j$, $j = 1, \ldots n$, and with each $\mathcal{C}_j$, $j = 1, \ldots, r$, being representable as a set constraint imposed on a particular subset $\{v_j\}$ of the variables in $\mathcal{V}$. We define a conservative CSP (CCSP) as being reducible to a CSP described by a 3-tuple $\langle \mathcal{V}, \mathcal{D}, \mathcal{C} \rangle$ having elements that take the following form:

$$\mathcal{V} = \{\underline{c}, \underline{d}\}, \quad \mathcal{D} = \{\mathbb{R}^k, \mathbb{R}^k\}, \quad \mathcal{C} = \{W, \mathcal{M}\}, \quad (2)$$

where W and $\mathcal{M}$ are constraints imposed on the entire set of variables $\{\underline{c}, \underline{d}\}$, and where in particular W is a linear subspace of $\mathbb{R}^{2k}$ that satisfies the following property:

$$\|\underline{c}\|^2 - \|\underline{d}\|^2 = 0, \quad \begin{bmatrix} \underline{c}^T & \underline{d}^T \end{bmatrix}^T \in W. \quad (3)$$

In (3) and the sequel, $\| \cdot \|$ denotes the 2-norm.

There are a variety of techniques that can be used to verify that a particular CSP is a CCSP, i.e. that can be used to reduce a CSP to a form that satisfies (2)-(3). Among these are algebraic manipulation and reduction techniques, a key ingredient of which would be the identification and transformation of conservation principles representable as a quadratic form that is isomorphic to the left-hand side of (3), as is discussed in detail in [7]. Moreover, the references [7]–[9] discuss a variety of practical engineering problems that are reducible to solving a CCSP. These include solving a broad class of linear and nonlinear optimization problems, as discussed in [2]. Additionally, determining the steady-state voltage and current distributions in a linear or nonlinear electrical network is reducible to a CCSP.

There are generally a variety of algebraic expressions satisfying (2)-(3) that can be used to describe a particular CCSP. With this in mind we focus the scope of discussion in this paper to those CCSPs for which the respective set constraints W and $\mathcal{M}$ are specifically generated using functional relationships between the vectors $\underline{c}$ and $\underline{d}$:

$$W = \left\{ \begin{bmatrix} \underline{c} \\ G\underline{c} \end{bmatrix} \; : \; \underline{c} \in \mathbb{R}^k \right\}, \quad \mathcal{M} = \left\{ \begin{bmatrix} m(\underline{d}) \\ \underline{d} \end{bmatrix} \; : \; \underline{d} \in \mathbb{R}^k \right\} \quad (4)$$

where $G \colon \mathbb{R}^k \to \mathbb{R}^k$ is an orthogonal matrix and $m \colon \mathbb{R}^k \to \mathbb{R}^k$ is a generally nonlinear map. We proceed by casting CCSPs as fixed-point problems using the maps in (4), in particular being concerned with identifying any vectors $(\underline{c}^\star, \underline{d}^\star)$ that solve the implicit equations

$$\underline{d}^\star = G\underline{c}^\star \quad \text{and} \quad \underline{c}^\star = m(\underline{d}^\star). \quad (5)$$
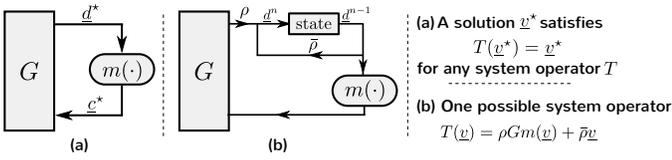
Fig. 1. Example signal processing systems associated with (a) a solution to the algebraic system (5) and (b) a filtered system that can be replaced by (a) in steady-state. An informal definition of a system operator is also provided.

III. ASYNCHRONOUS, FILTERED SIGNAL PROCESSING SYSTEMS

The class of signal processing systems that we address in this paper is specifically composed of those that are consistent with the description in Fig. 1(b). The general strategy for solving CCSPs using these systems is specifically to design a system as in Fig. 1(b) for which, in steady-state, it may be replaced with the associated system depicted in Fig. 1(a), which would graphically represent a solution to the CCSP as indicated in (5).

IV. DEFINITIONS AND CONVENTIONS

We now provide a sequence of definitions to establish a notational convention and to clearly state the notion of stability considered in this paper. To this end, we define a *system operator* as any map $T \colon \mathbb{R}^k \to \mathbb{R}^k$ consistent with (5) and the discussion in Section III, e.g. $T(\underline{v}) = Gm(\underline{v})$ or $T(\underline{v}) = m(G\underline{v})$.

First, we restrict our attention to the class of system operators addressed in the following definition.

Definition 1 (α-conic). *We define a system operator T as being α-conic at $\underline{v}$ provided there exists a constant $\alpha \geq 0$ for which*

$$\sup_{\underline{u} \neq \underline{0}} \frac{\|T(\underline{v} + \underline{u}) - T(\underline{v})\|}{\|\underline{u}\|} \leq \alpha. \tag{6}$$

If T satisfies (6) for all $\underline{v} \in \mathbb{R}^k$ then we call T α-conic everywhere.

Observe that being α-conic everywhere is equivalent to Lipschitz continuity with constant α. For a given operator T, we define its filtered realization next and focus on its implementation thereafter.

Definition 2 (Filtered system operator). *For a given system operator T, we associate a filtered system operator T_f whose action is to take an affine combination of $\underline{v}$ and $T(\underline{v})$ of the form*

$$T_f(\underline{v}) \triangleq \rho T(\underline{v}) + \bar{\rho}\underline{v} \tag{7}$$

where $\rho > 0$ is the filtering parameter and $\bar{\rho} = 1 - \rho$.

Any fixed point of T_f is clearly a fixed point of T and vice versa.

An asynchronous system is one in which the state elements act as coordinatewise, randomly triggered sample-and-hold elements. We define an implementation protocol consistent with this behavior next.

Definition 3 (Asynchronous protocol). *An implementation of T with initial state $\underline{v}^0$ is given by the state sequence $\{\underline{v}^n\}_{n=0}^{\infty}$ generated by*

$$\underline{v}^n = D^{(p)} T_f\left(\underline{v}^{n-1}\right) + \left(I_k - D^{(p)}\right)\underline{v}^{n-1}, \qquad n \geq 1, \tag{8}$$

where I_k is an identity matrix and $D^{(p)}$ is a $k \times k$ stochastic, binary, diagonal matrix whose k diagonal elements are i.i.d. Bernoulli and independent of $\underline{v}^{n-1}$, taking values $D_{ii}^{(p)} = 1$ with probability p and $D_{ii}^{(p)} = 0$ with probability $1 - p$.

Observe that implementing T utilizes T_f rather than T itself.

Referring to (8), an asynchronous signal processing system is able to reduce its overall communication bandwidth by performing only the subset of computation dictated by the delays which trigger for each iteration n. This property is often desirable in distributed computing environments, e.g. in a heterogeneous network with un-coordinated processing nodes [10].

We now formally define stability in the context of this paper.

Definition 4 (Stability in r-th mean [11]). *A system operator T is stable in r-th mean if, using the update (8), the condition*

$$\lim_{n \to \infty} \mathbb{E}\left[\|\underline{v}^n - \underline{v}^\star\|^r\right] = 0 \tag{9}$$

holds for some state $\underline{v}^\star$ where $\mathbb{E}[\cdot]$ denotes the expectation operator. When (9) holds, we write it succinctly as $\underline{v}^n \xrightarrow{r} \underline{v}^\star$.

Finally, we denote by $B(\underline{c}, r)$ a non-empty, closed basin with center c and radius $r > 0$, i.e. $B(\underline{c}, r) = \{\underline{v} \in \mathbb{R}^k \colon \|\underline{v} - \underline{c}\| \leq r\}$.

V. STABILITY ANALYSIS

The purpose of this section is to provide sufficient conditions under which a system operator T is stable. We reiterate that stability is a property of the system operator and not a particular realization of the state sequence. Additionally, many of the coming results reduce to known conditions in special cases, typically by setting $\rho = p = 1$. Proofs of the following theorems are available in [6].

A. Dissipative system operators

We first address α-*dissipative* system operators, i.e. those that are α-conic with $\alpha < 1$. To begin, we bound the domain of a state sequence when the system operator is α-dissipative about an arbitrary state. This restriction does not preclude a system operator from being α-conic about any other state with $\alpha \geq 1$.

Theorem 1 (Finite-time entrapment; $\alpha < 1$). *Let T be α-dissipative about $\underline{c}$. Then, for every $\epsilon > 0$ and initial state $\underline{v}^0$, the state sequence $\{\underline{v}^n\}_{n=n_0}^{\infty}$ is contained within $B(\underline{c}, \frac{\rho\|T(\underline{c}) - \underline{c}\|}{1 - \rho\alpha - |\bar{\rho}|} + \epsilon)$ in mean, i.e.*

$$\mathbb{E}\left[\|\underline{v}^n - \underline{c}\|\right] \leq \frac{\rho\|T(\underline{c}) - \underline{c}\|}{1 - \rho\alpha - |\bar{\rho}|} + \epsilon, \quad n \geq n_0, \tag{10}$$

for some finite integer n_0 provided $\rho \in (0, \frac{2}{1+\alpha})$.

If $\underline{c}$ is a fixed point, then it is unique and Theorem 1 implies that the state sequence becomes arbitrarily close to $\underline{c}$ in finite time. Next, we consider system operators that are α-dissipative over a basin with center $\underline{c}$.

Theorem 2 (Convergence in a basin; $\alpha < 1$). *Let T be α-dissipative about all $\underline{v} \in B(\underline{c}, r)$. Then T has a unique fixed point $\underline{v}^\star$ in $B(\underline{c}, r)$ and is stable in mean, i.e. $\underline{v}^n \xrightarrow{1} \underline{v}^\star$, provided $\underline{v}^0 \in B(\underline{c}, r)$, $\rho \in (0, 1]$, and*

$$\|\underline{c} - T(\underline{c})\| \leq (1 - \alpha)r. \tag{11}$$

We now take T to be α-dissipative everywhere or equivalently to be a contraction. The Banach fixed-point theorem [12] states that T has a unique fixed point and that synchronously ($p = 1$) iterating T ($\rho = 1$) from any initial state produces it at a linear rate. More generally, by using a filtered system operator we obtain

$$\|T_f(\underline{v}) - \underline{v}^\star\| \leq (\rho\alpha + |\bar{\rho}|)\|\underline{v} - \underline{v}^\star\| \tag{12}$$
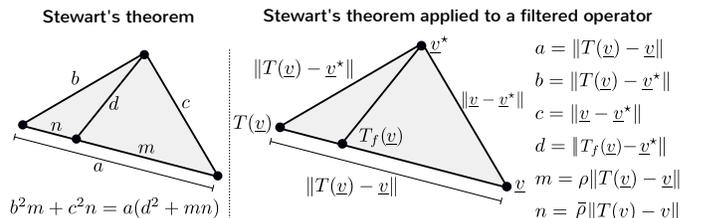
Fig. 2. The application of Stewart's theorem (left) to a filtered system operator T_f (right) illustrating the geometry behind the identity in (13).

If $\underline{v}^{n-1} = \underline{v}$ then $\underline{v}^n$ takes the state:

(i) with probability $(1-p)^2$
(iii) with probability p^2
(iv) with probability $p(1-p)$
(v) with probability $p(1-p)$
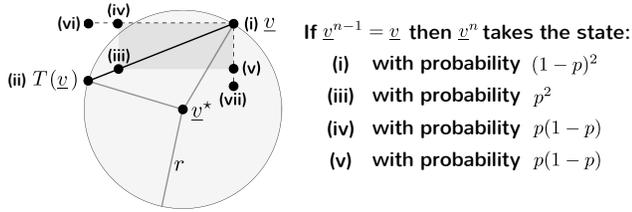
Fig. 3. An illustration of the possible outcomes for a single asynchronous update according to (8) for a passive everywhere system operator T in $\mathbb{R}^2$.

Fig. 4. An example of the possible outcomes for the system operator described in Section V-C in $\mathbb{R}^2$ illustrating the utility of Theorems 1 and 2 in assembling stability properties for α-expansive system operators.

from which we conclude $\rho\alpha + |\bar{\rho}| \in [0,1)$ provided $\rho \in (0, \frac{2}{1+\alpha})$. We extend this to include asynchronous updates in the next theorem.

Theorem 3 (Convergence in $\mathbb{R}^k$; $\alpha < 1$). *Let T be α-dissipative everywhere. Then T has a unique fixed point $\underline{v}^\star$ and is stable in mean square, i.e. $\underline{v}^n \xrightarrow{2} \underline{v}^\star$, provided $\rho \in (0, \frac{2}{1+\alpha^2})$.*

B. Passive system operators

We now consider *passive everywhere* or non-expansive system operators, i.e. those which are α-conic for $\alpha = 1$ and all $\underline{v} \in \mathbb{R}^k$. The fixed-point set $\mathcal{F}_T$ of such an operator T is necessarily convex; this fact is a straightforward consequence of the Cauchy-Schwarz inequality. This does not preclude $\mathcal{F}_T$ from being empty and so we assume hereon that $\mathcal{F}_T$ is non-empty, i.e. that the associated CCSP is well-defined. We proceed in this subsection to focus on stability in mean square for this class of system operators in two stages: (i) characterizing the application of the filtered operator T_f on $\underline{v}^{n-1}$, and (ii) enumerating the possible states for $\underline{v}^n$ and their likelihoods based upon the stochastic combination of $\underline{v}^{n-1}$ and $T_f(\underline{v}^{n-1})$ in (8).

Denote $\underline{v}^{n-1}$ by $\underline{v}$ for clarity. The continuum of states achievable by application of T_f to $\underline{v}$ as a function of $\rho > 0$ is given by the open ray $\rho T(\underline{v}) + \bar{\rho}\underline{v}$. The squared distance of any state $T_f(\underline{v})$ on this continuum to a fixed point $\underline{v}^\star \in \mathcal{F}_T$ is

$$\|T_f(\underline{v}) - \underline{v}^\star\|^2 = \rho \|T(\underline{v}) - \underline{v}^\star\|^2 + \bar{\rho}\|\underline{v} - \underline{v}^\star\|^2 - \rho\bar{\rho}\|T(\underline{v}) - \underline{v}\|^2 \tag{13}$$

Utilizing the passivity of T, (13) reduces further to the inequality

$$\|T_f(\underline{v}) - \underline{v}^\star\|^2 \leq \|\underline{v} - \underline{v}^\star\|^2 - \rho\bar{\rho}\|T(\underline{v}) - \underline{v}\|^2. \tag{14}$$

Restricting $\rho\bar{\rho}$ to be positive, i.e. restricting $\rho \in (0,1)$, observe that the squared distance of $T_f(\underline{v})$ to $\underline{v}^\star$ is no further than the squared distance of $\underline{v}$ to $\underline{v}^\star$ and is strictly closer so long as $\underline{v}$ is not itself a fixed point. This observation suggests that mean square stability of T will not coincide with a linear rate of convergence without further restricting the class of passive everywhere system operators since $T(\underline{v})$ and $\underline{v}$ may be arbitrarily close together in (14).

We briefly justify the identity (13). A proof relying on Euclidean geometry is established by application of Stewart's theorem [13] to the filtered system operator T_f. Specifically, consider a triangle with sides of length a, b, and c with a cevian to side a of length d where the cevian divides a into two pieces of lengths n and m where m is adjacent to c and n is adjacent to b. Stewart's theorem, illustrated in Fig. 2 on the left, ensures that these distances satisfy $b^2 m + c^2 n = a(d^2 + mn)$. By proper assignment of the vertices of such a triangle (listed on the right in Fig. 2), the identity in (13) follows immediately. Referring to Fig. 3, the triangle with corners labeled (i), (ii) and $\underline{v}^\star$ corresponds to the triangle on the right of Fig. 2 and the label (iii) corresponds to the state $T_f(\underline{v})$ for some fixed value of $\rho \in (0,1)$.

We next enumerate the possible outcomes of $\underline{v}^n$ according to the stochastic combination of $\underline{v}$ and $T_f(\underline{v})$ in (8) and reference Fig. 3 as an example in $\mathbb{R}^2$. In particular, $\underline{v}^n$ can take any of the 2^k corners of the k-orthotope (hyperrectangle) defined with opposite corners $\underline{v}$ and $T_f(\underline{v})$. Intuitively, each corner represents one of the 2^k ways i
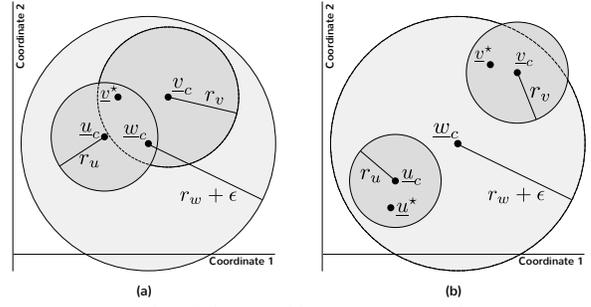
of k state elements can trigger for $0 \leq i \leq k$ and is exhaustive since

$$\sum_{i=0}^{k} C(k,i)p^i(1-p)^{k-i} = 1 \tag{15}$$

where $C(k,i)$ is the binomial coefficient corresponding to the total number of ways i of the k state elements can trigger. Referring to Fig. 3, the labels (i), (iii), (iv), and (v) correspond to these four outcomes for a fixed-value of ρ. The line segments with endpoints labeled (i)-(ii), (i)-(vi), and (i)-(vii) describe the continuum of possible outcomes as ρ varies between 0 and 1.

From the geometry of the dynamics in (8), it is straightforward to conclude that a subset of the possible states taken by $\underline{v}^n$ are further from the fixed point $\underline{v}^\star$ than either $\underline{v}^{n-1}$ or $T_f(\underline{v}^{n-1})$. Referring to Fig. 3, the state labeled (iv) is an example of such an outcome. In general, such outcomes cannot be avoided by further restricting the filtering parameter ρ. The following theorem, however, ensures that the system operator T is indeed stable in mean square so long as the filtering parameter is selected in the open unit interval.

Theorem 4 (Convergence in $\mathbb{R}^k$; $\alpha = 1$). *Let T be passive everywhere with a non-empty set of fixed points $\mathcal{F}_T$. Then T is stable in mean square, i.e. $\underline{v}^n \xrightarrow{2} \underline{v}^\star$ for some $\underline{v}^\star \in \mathcal{F}_T$, provided $\rho \in (0,1)$.*

C. Expansive system operators

The final class of system operators we address are α-*expansive*, i.e. those which are α-conic for $\alpha > 1$. Consistent with the remark preceding Theorem 1, an α-conic system operator may satisfy (6) about generally many states with a different parameter $\alpha \in [0, \infty)$ for each. The stability of such a system operator may be analyzed by aggregating the appropriate stability results for each state, for example by using local results including Theorems 1 and 2.

Consider a system operator T which is α-expansive about some state and additionally satisfies the following conditions:

1) T is α_w-dissipative about $\underline{w}_c$ where $\underline{w}_c \neq T(\underline{w}_c)$ (Theorem 1);
2) T is α_v-dissipative about all $\underline{v} \in B(\underline{c}_v, r_v)$ for a radius r_v satisfying $\|\underline{c}_v - T(\underline{c}_v)\| \leq (1 - \alpha_v)r_v$ (Theorem 2);
3) T is α_u-dissipative about all $\underline{u} \in B(\underline{c}_u, r_u)$ for a radius r_u satisfying $\|\underline{c}_u - T(\underline{c}_u)\| \leq (1 - \alpha_u)r_u$ (Theorem 2).

Figure 4 depicts the domain of T in $\mathbb{R}^2$ for two scenarios consistent with the assumptions listed above. Referring to the figure, we conclude from the first condition that within a finite number of iterations the state sequence will enter the basin $B(\underline{w}_c, r_w + \epsilon)$. Scenario (a) depicts the case where the intersection of $B(\underline{u}^c, r_u)$ and $B(\underline{v}^c, r_v)$ is non-empty and so if the state sequence enters either basin it will converge to the unique fixed point $\underline{v}^\star = \underline{u}^\star$ in mean (provided ρ is appropriately selected). Scenario (b) depicts the complementary case where the intersection of these basins is empty and so T has at least
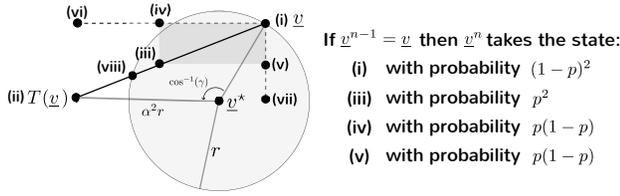
Fig. 5. An illustration of the possible outcomes for a single asynchronous update according to (8) for an α-expansive everywhere system operator T in $\mathbb{R}^2$ which satisfies the conic mixing property in (16).

two distinct fixed points. Furthermore, if the state sequence enters either basin then it will converge in mean to the unique fixed point inside the basin it entered.

Consider now those system operators that are α-expansive everywhere with non-empty, convex fixed-point sets. If a system operator in this class also satisfies a conic mixing property about the fixed-point set, then the filtering coefficient may be judiciously chosen so that the filtered operator is effectively dissipative everywhere in turn behaving stably. This mixing condition and its consequences in terms of stability are stated formally next.

Theorem 5 (Conic mixing; $\alpha > 1$)**.** *Let T be α-expansive everywhere with a non-empty, convex set of fixed points $\mathcal{F}_T$. If T additionally satisfies the conic mixing property*

$$\sup_{\underline{v} \notin \mathcal{F}_T} \frac{\langle T(\underline{v}) - \underline{v}^\star, \underline{v} - \underline{v}^\star \rangle}{\|T(\underline{v}) - \underline{v}^\star\| \, \|\underline{v} - \underline{v}^\star\|} \leq \gamma \qquad (16)$$

for all $\underline{v}^\star \in \mathcal{F}_T$ and some $\gamma \in [-1, 1)$ such that $\alpha\gamma < 1$, then T is stable in mean square, i.e. $\underline{v}^n \xrightarrow{2} \underline{v}^\star$ for some $\underline{v}^\star \in \mathcal{F}_T$, provided

$$\rho \in \left(0, \frac{2(1 - \alpha\gamma)}{1 + \alpha^2 - 2\alpha\gamma}\right). \qquad (17)$$

A connection between Theorem 5 and Fig. 5 is similar to the previously established connection between Theorem 4 and Fig. 3 with two modifications. First, the restriction in (17) of the filtering parameter is analgous to selecting $\rho \in (0, 1)$ in Theorem 4 since it enforces that the squared distance of $T_f(\underline{v})$ to $\underline{v}^\star$ is strictly less than the squared distance of $\underline{v}$ to $\underline{v}^\star$. Referring to Fig. 5, the equidistant state $T_f(\underline{v})$ corresponds to $\rho = \frac{2(1-\alpha\gamma)}{1+\alpha^2-2\alpha\gamma}$ and is labeled (viii). Second, the conic mixing parameter $\gamma < 1$ in (16) implies that $T_f(\underline{v})$ and $\underline{v}$ cannot be arbitrarily close together on the perimeter of the basin $B(\underline{v}^\star, r)$ in Fig. 5. As a consequence, the state sequence converges linearly to a fixed point and the optimal filtering parameter in the sense of provable convergence rates is $\rho^\star = \frac{1-\alpha\gamma}{1+\alpha^2-2\alpha\gamma}$. The analogous choice of filtering paramter in Fig. 3 is $\rho^\star = \frac{1}{2}$. In the respective contexts of Figs. 3 and 5, the optimal filter parameters correspond to selecting the state $T_f(\underline{v})$ which minimizes the squared distance $\|T_f(\underline{v}) - \underline{v}^\star\|^2$.

VI. CONNECTIONS TO OPTIMIZATION THEORY

A connection between the class of CCSPs defined in Section II and a class of generally convex and non-convex optimization problems is essentially established in [2]. The key link to this is that the quadratic conservation principle inherent to the stationarity conditions described in [2] is isomorphic to the quadratic conservation principle inherent to (3), thereby providing direction in the following general strategy: (i) transform the stationarity conditions into a CCSP, (ii) solve the CCSP using an asynchronous processing system informed by the sufficient conditions presented in this paper, and (iii) put the solution obtained through the inverse transform resulting in a solution to the optimization problem. Drawing upon this connection, the asynchronous signal processing systems presented in [8]–[10] for solving optimization problems can readily be viewed as special cases
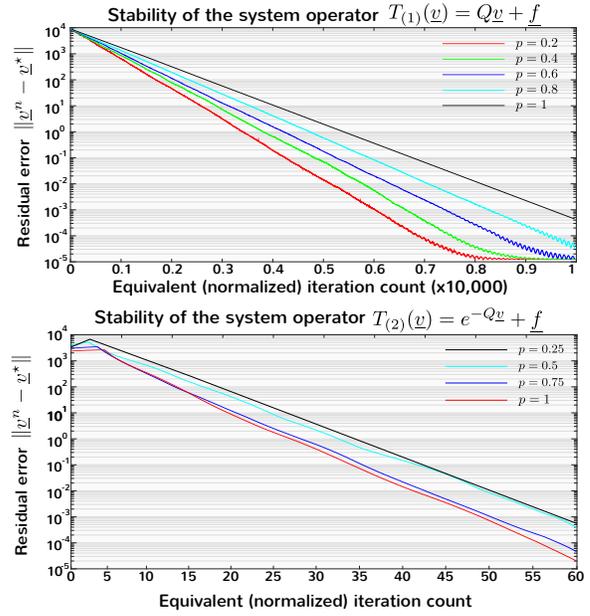
Fig. 6. Numerical stability results obtained by asynchronously implementing the system operators in (18) as discussed in Section VII.

of the systems presented in this paper and in this sense serve as a variety of numerical examples.

VII. NUMERICAL EXPERIMENTS

In this section, we perform a series of numerical experiments to investigate the numerical stability of asynchronously implementing system operators for two CCSPs. Our procedure is as follows:

1) for a given dimensionality k, generate a system operator $T: \mathbb{R}^k \to \mathbb{R}^k$ associated with the CCSP at hand;
2) identify a fixed point $\underline{v}^\star \in \mathcal{F}_T$ of the system operator;
3) generate a state sequence $\{\underline{v}^n\}_{n=0}^\infty$ for various probability values p using (8) with $\rho = \frac{1}{2}$ and track the size of $\underline{v}^n - \underline{v}^\star$;
4) repeat step 3 for many trials with a new initial state $\underline{v}^0$ on each trial and average the results.

Let $T_{(1)}$ and $T_{(2)}$ respectively denote passive everywhere and transcendental system operators of the form:

$$T_{(1)}(\underline{v}) = Q|\underline{v}| + \underline{f}, \quad \text{and} \quad T_{(2)}(\underline{v}) = e^{-Q\underline{v}} + \underline{f}, \quad (18)$$

where Q is an orthogonal matrix with eigenvalues bounded away from -1, $\underline{f}$ is a Gaussian random vector, and the absolute value and exponential are coordinatewise. Q is obtained as follows: a candidate matrix is drawn from a Gaussian ensemble and then projected to the nearest orthogonal matrix in the Frobenius sense and candidates are regenerated until one satisfies the desired spectral property.

The experimental results are depicted in Fig. 6 as a function of "equivalent (normalized) iteration count" by which we mean the total amount of computation performed, i.e. each coordinatewise state update is counted as $\frac{1}{k}$th of an iteration. The performance depicted is averaged over 1000 trials for various values of p and dimensionality $k = 50$. The initial state $\underline{v}^0$ is randomly chosen from the boundary of a scaled unit sphere in $\mathbb{R}^k$ on each trial. Observe for $T_{(1)}$ that the average state sequence initially converges at a linear rate and slows to sublinear performance once it is within a neighborhood sufficiently close to $\underline{v}^\star$. Although $T_{(2)}$ is not α-conic for any finite value of α, the operator defined by $T_{(2)}$ composed with itself m times is for some finite integer m. This manifests itself in Fig. 6 in the first few iterations, after which linear convergence in mean square consistent with Theorem 3 is observed.

REFERENCES

[1] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," 1999.

[2] T. A. Baran and T. A. Lahlou, "Conservative signal processing architectures for asynchronous, distributed optimization part I: General framework," in *IEEE Global Conference on Signal and Information Processing*, Dec 2014, pp. 35–39.

[3] C. E. Lee, A. E. Ozdaglar, and D. Shah, "Solving systems of linear equations: locally and asynchronously," *Computing Research Repository*, vol. abs/1411.2647, 2014.

[4] P. Tseng, D. P. Bertsekas, and J. N. Tsitsiklis, "Partially asynchronous, parallel algorithms for network flow and other problems," *SIAM Journal on Control and Optimization*, vol. 28, no. 3, pp. 678–710, May 1990.

[5] D. Mitra, "Asynchronous relaxations for the numerical solution of differential equations by parallel processors," *SIAM Journal on Scientific and Statistical Computing*, 1987.

[6] T. A. Lahlou and T. A. Baran, "Signal processing structures for solving conservative constraint satisfaction problems," in *arXiv*, Aug 2015.

[7] T. A. Baran, *Conservation in signal processing systems*, Ph.D. thesis, Massachusetts Institute of Technology, 2012.

[8] T. A. Baran and T. A. Lahlou, "Conservative signal processing architectures for asynchronous, distributed optimization part II: Example systems," in *IEEE Global Conference on Signal and Information Processing*, Dec 2014, pp. 40–44.

[9] T. A. Lahlou and T. A. Baran, "Asynchronous algorithms for solving linear programs," *ArXiv e-prints*, Feb. 2015.

[10] T. A. Lahlou and T. A. Baran, "Implementing conservative signal processing systems for asynchronous, distributed optimization as a web service," *ArXiv e-prints*, Sept. 2015.

[11] R. G. Gallager, *Stochastic processes, theory for applications*, Cambridge University Press, 2013.

[12] R. M. Brooks and K. Schmitt, *The contraction mapping principle and some applications*, Electronic journal of differential equations: Monograph. Department of Mathematics, Texas State University, 2009.

[13] J. W. Russell, *An elementary treatise on pure geometry: With numerous examples*, Clarendon Press series. Clarendon Press, 1893.