Max-Plus Algebraic Throughput Analysis of Synchronous Dataflow Graphs

***Abstract*—In this paper we present a novel approach to throughput analysis of synchronous dataflow (SDF) graphs. Our approach is based on describing the evolution of actor firing times as a linear time-invariant system in max-plus algebra. Experimental results indicate that our approach is faster than state-of-the-art approaches to throughput analysis of SDF graphs. The efficiency of our approach is due to the exploitation of the regular structure of the max-plus system's graphical representation, the properties of which we thoroughly prove.**

I. Introduction

Synchronous dataflow (SDF) graphs [10] are well-known models of computation that are widely used to model real-time embedded streaming applications. Timing analysis of SDF graphs aims at finding performance characteristics such as throughput and latency, which is crucial information when exploring the design-space of real-time critical systems.

There are two main approaches to the timing analysis of SDF graphs. The first approach is based on the transformation of an SDF graph into an equivalent homogeneous SDF (HSDF) graph, which is then analysed for its critical cycle. A disadvantage of this approach is that the HSDF graph may become quite large: in the worst case, its size is exponential in the size of the corresponding SDF graph. The second, state-of-the-art approach to timing analysis of SDF graphs is by exploring the state-space of a simulated self-timed execution until a periodic phase is found. Such a simulation-based method avoids the transformation from SDF into HSDF.

In this paper, we present an alternative, analytical approach to timing analysis of SDF graphs. Our approach consists of a novel way of constructing a max-plus algebraic description of the evolution of actor firing times in a self-timed execution of an SDF graph. As a result, we obtain HSDF-like graphs that contain significantly fewer edges than the HSDF graph obtained by the commonly followed transformation from SDF into HSDF. Furthermore, the graphs obtained by our transformation may be efficiently analysed for its maximum cycle ratio. This is due to the regular structure of these graphs, the properties of which are formally proven.

The main contribution of our work is a sound and new basis for the formal analysis of SDF graphs using max-plus algebra, which allows for an efficient method to calculate the throughput of an SDF graph. We confirm the efficiency of our method by comparing it with the state-of-the art simulation-based approach on testsets used in an earlier study [6].

The remainder of this paper is outlined as follows: in section III, we give a brief introduction to SDF graphs, equivalent HSDF graphs, max-plus algebra and the graphical representation of max-plus systems. In sections IV - V we describe how a linear, time-invariant max-plus system may be derived from an SDF graph and graphically represented. Section VI formally proves properties of the structure of these linear max-plus systems and Section VII describes the experimental comparison between our approach and the state-of-the-art simulation-based approach to throughput analysis. Finally, Section VIII concludes the paper and gives directions for future work.

II. Related work

In timing analysis of SDF graphs, the transformation of the graph into an equivalent HSDF graph is a common step that is described by various authors, e.g. [9], [10] or [11]. In these papers, the potentially huge size of the HSDF graph is often given as a main reason to resort to simulation-based methods [6]. In fact, in [6] a comparison between a simulation-based approach in which the state-space of a self-timed execution of an SDF graph is explored and methods based on analysing the equivalent HSDF graph has concluded that simulation is a few orders of magnitude faster. In our approach, we derive an opposite result.

The potentially large size of an SDF graph's equivalent HSDF graphs has been recognised as a problem in [9], where the authors describe an approach to reduce the size of an SDF graph's equivalent HSDF graph. The main drawback of their approach is that they require the full HSDF graph to be constructed first, which is avoided in our approach.

In [5] it is described how reduced HSDF graphs are obtained from SDF graphs by representing each token in the SDF graph by a single linear max-plus expression. Although the size of the reduced HSDF graph may be small for graphs with only very few tokens, constructing the system involves simulation of the SDF graph and the symbolic manipulation of max-plus expressions, which is complicated and requires the administration of all tokens that are produced and consumed during the execution of the SDF graph. Our approach is simpler and does not depend on the number of tokens in the graph.

III. Preliminaries

In this section we will discuss some specification formalisms and their relationships.

A. SDF graphs

Synchronous dataflow (SDF) graphs are often used to model streaming applications. We will assume that the reader is familiar with standard SDF terminology (such as actor, channel,

firing, production/consumption rates, etc), we only define a few SDF-notions that are relevant for this paper.

An SDF graph is *consistent* if a shortest non-empty sequence of actor firings exists, which as a whole will effectively leave the token distribution unchanged. Such a sequence of firings is called a *graph iteration*. The repetition vector q of a consistent SDF graph associates with each actor a the number of times q_a that actor fires within a single graph iteration.

The time between the start and completion of a single firing of an actor a is called the *execution time* of actor a and is denoted by τ_a. The *throughput* of an SDF graph is the average number of graph iterations that are executed per unit of time, measured over a sufficiently large amount of time. The maximum throughput is attained by a *self-timed* execution, which means that each actor fires (possibly several times simultaneously) as soon as it is enabled.

An example SDF graph is depicted in figure 1(a). Each actor is annotated with its execution time: 2 time units for actors a and b, 3 time units for actor c. The graph is consistent: a graph iteration consists of 4 firings of actor a, 2 firings of actor b and 3 firings of actor c. Hence, the repetition vector of the graph is $\langle 4, 2, 3 \rangle$.

The schedule of a self-timed execution of the example graph is shown in Figure 1(b). It starts with two (parallel) firings of actor a to consume the two initial tokens on channel ba. After an initial settling phase of 2 time units, every 9 time units a single iteration is completed. The throughput achieved in a self-timed execution is therefore $\frac{1}{9}$.

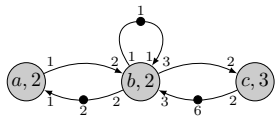
B. Equivalent HSDF graphs

The timing behaviour of a consistent SDF graph can be analysed by transforming the SDF graph into an equivalent *homogeneous* SDF (HSDF) graph, i.e., an SDF graph in which all production and consumption rates are *one*, using the well-known procedures found in e.g., [11] or [10]. Given an SDF graph, the equivalent HSDF graph is constructed by creating an *actor* for each firing in a single iteration of the SDF graph, and a *channel* for each produced/consumed token in the original SDF graph. Figure 1(c) shows the HSDF graph corresponding to the SDF graph in Figure 1(a).

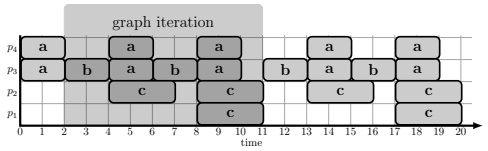
We remark that the complexity of an equivalent HSDF graph is increased (even exponentially) in comparison with the underlying SDF graph. This increase in complexity is the primary reason HSDF graphs are not used in most analysis methods for SDF graphs and gave rise to simulation-based methods instead (cf. e.g., [6]).

C. Max-Plus Algebra

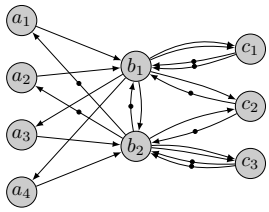
Timed synchronous systems may be mathematically described using *max-plus algebra* [2], [3], [8]. In max-plus algebra, times at which events take place are related to times at which preceding events take place using the operators $\max$ to express synchronisation and $+$ to express duration. The additive *zero element* $\varepsilon = -\infty$ is used to indicate absence of a precedence relation. To emphasise the resemblance between conventional linear algebra and max-plus algebra, it is common



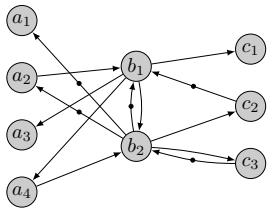
(a) SDF Graph



(b) Self-timed schedule



(c) Equivalent HSDF Graph



(d) Timed Event Graph

$$t_{a_1}(k) = t_{b_2}(k-1) \otimes 2 \qquad t_{c_1}(k) = t_{b_1}(k) \otimes 3$$
$$t_{a_2}(k) = t_{b_2}(k-1) \otimes 2 \qquad t_{c_2}(k) = t_{b_2}(k) \otimes 3$$
$$t_{a_3}(k) = t_{b_1}(k) \otimes 2 \qquad t_{c_3}(k) = t_{b_2}(k) \otimes 3$$
$$t_{a_4}(k) = t_{b_1}(k) \otimes 2$$
$$t_{b_1}(k) = (t_{a_2}(k) \oplus t_{b_2}(k-1) \oplus t_{c_2}(k-1)) \otimes 2$$
$$t_{b_2}(k) = (t_{a_4}(k) \oplus t_{b_1}(k) \oplus t_{c_3}(k-1)) \otimes 2$$

(e) Max-Plus equations

Fig. 1. Example SDF-graph with several derived representations.

to use operators $\otimes$ and $\oplus$ to denote $+$ and $\max$, respectively. We remark that $\otimes$ is distributive over $\oplus$.

For example, the following expression states that the k^{th} occurance of event h happens at least 5 time units after the $(k-1)^{\text{th}}$ occurance of event i and at least 3 time units after the k^{th} occurance of event j:

$$t_h(k) \geq t_i(k-1) \otimes 5 \oplus t_j(k) \otimes 3. \tag{1}$$

Similar to linear system descriptions in conventional algebra, the behaviour of a timed synchronous system can be expressed as an m^{th} order linear max plus system:

$$x(k) = \bigoplus_{i=0}^{m} A_i \otimes x(k-i). \tag{2}$$

Many efficient algorithms are available to analyse such a linear max-plus description [8], [1] or its graphical representation [4].

D. Timed Event Graphs

A linear time-invariant max-plus system may be depicted graphically by a *timed event graph* (TEG, also known as *timed marked graph*). In a TEG, each non-zero element a_{jm} in the

matrix A_i in (2) is represented by an edge from m to j, with i tokens (called *delays*) and a weight a_{jm}. Figure 1(d) shows a TEG for the max-plus system that is derived from the SDF graph of Figure 1(a) (see Section IV for details). An HSDF graph may in fact be considered as a specialisation of a TEG, where the weights of edges are replaced by execution times of actors. In the remainder of this paper, we will omit edge weights in a TEG or an HSDF graph's execution times when they may be inferred from the context.

IV. Linear Max-Plus descriptions of SDF graphs

In this section we will use max-plus algebra to describe the evolution of actor firing times during the self-timed execution of an SDF graph. The events that we will relate in max-plus expressions are the *completion* of firings. These events are related through *precedence constraints*, which are imposed by the channels in an SDF graph: the times at which an actor may fire depends on the times at which sufficient tokens become available on its incoming channels.

For each SDF channel ab we will relate times at which firings of actor b complete to times at which firings of actor a complete. In order to define these relations, we will write m_{ab} for the production rate of actor a on channel ab, n_{ab} for the consumption rate of actor b on channel ab and d_{ab} for the initial number of tokens on channel ab.

Per SDF channel ab, the time at which actor b may start its j^{th} firing is constrained by the time at which the *last required token* for that firing is produced onto channel ab by actor a. The completion of the j^{th} firing of actor b requires the production of *at least* $N = j \cdot n_{ab} - d_{ab}$ tokens by actor a.

To find the firing of actor a that must have completed such that the j^{th} firing of actor b may start, we must thus divide by m_{ab} and round the result towards the nearest higher integer. Let i be this firing of a. We call i the *predecessor* of j on channel ab, denoted $i = \pi_{ab}(j)$, with $\pi_{ab}(j)$ defined as:

$$\pi_{ab}(j) = \left\lceil \frac{j \cdot n_{ab} - d_{ab}}{m_{ab}} \right\rceil. \tag{3}$$

We can use this predecessor function to relate the completion times of an actor's firings to the times at which firings of other actors complete. Let $t_b(j)$ denote the time at which actor b *completes* its j^{th} firing and $\mathcal{E}$ the set of channels in the SDF graph. The following max-plus expression then captures the precedence constraint for firings of actor b, due to the actor's incoming channels:

$$t_b(j) = \bigoplus_{ab \in \mathcal{E}} t_a\left(\pi_{ab}(j)\right) \otimes \tau_b. \tag{4}$$

These constraints may not generally be expressed as the linear time-invariant max-plus system (2). In parlance of system theory, the system expressed by (4) is a so-called *linear time-variant* system, since $\pi_{ab}(j)$ may not generally be replaced by $j - k$ (for some $k \in \mathbb{N}$).

For *consistent* SDF graphs however, equation (4) is *periodically* time-variant and may be expressed by a linear time-invariant system by a change of variables: We let $b_j(k)$ denote the j^{th} firing of actor b in the $(k+1)^{\text{th}}$ iteration, thus $t_{b_j}(k) = t_b(j + k \cdot q_b)$. Note that $j \in \{1, \ldots, q_b\}$. By changing variable b into b_j, equation (4) may be rewritten as follows:

$$t_{b_j}(k) = \bigoplus_{ab \in \mathcal{E}} t_a\left(\left\lceil \frac{j \cdot n_{ab} + k \cdot q_b \cdot n_{ab} - d_{ab}}{m_{ab}} \right\rceil\right) \otimes \tau_b. \tag{5}$$

Since in a single iteration of a consistent SDF graph, with repetition vector q, the number of tokens produced onto each channel is equal to the number of tokens consumed from that channel, we have $q_b \cdot n_{ab} = q_a \cdot m_{ab}$. We use this to simplify (5) into:

$$t_{b_j}(k) = \bigoplus_{ab \in \mathcal{E}} t_a\left(\pi_{ab}(j) + k \cdot q_a\right) \otimes \tau_b. \tag{6}$$

To complete the change of variables, we must rewrite $t_a(\pi_{ab}(j) + k \cdot q_a)$ as $t_a(i + m \cdot q_a)$, which we then write as $t_{a_i}(m)$, with $i \in \{1, \ldots, q_a\}$. Terms i and m are obtained by applying basic modular arithmetic. Note that since we number an actor's firings starting with *one*, decrements and increments by one are required. Let $\tilde{\pi}_{ab}(j)$ be the *firing index* of $\pi_{ab}(j)$ within a graph iteration, defined as follows:

$$\tilde{\pi}_{ab}(j) = (\pi_{ab}(j) - 1) \bmod q_a + 1, \tag{7}$$

and $(\delta_{ab} + 1)$ the *iteration index*: the index of the iteration in which the firing takes place, given by:

$$\delta_{ab}(j) = \left\lfloor \frac{\pi_{ab}(j) - 1}{q_a} \right\rfloor. \tag{8}$$

The following expression then completes the change of variables and gives a linear time-invariant system:

$$t_{b_j}(k) = \bigoplus_{ab \in \mathcal{E}} t_{a_{\tilde{\pi}_{ab}(j)}}\left(k + \delta_{ab}(j)\right) \otimes \tau_b. \tag{9}$$

As an example, consider the SDF graph depicted in Figure 1(a). The time-variant precedence constraint for actor b is:

$$t_b(j) = \left(t_b(j-1) \oplus t_a(2j) \oplus t_c\left(\left\lceil \frac{3j-6}{2} \right\rceil\right)\right) \otimes 2.$$

The linear, time-invariant equations for actor b and the other actors in the SDF graph are shown in Figure 1(e).

V. Linear Constraint Graph

The linear time-invariant max-plus system expressed by equation (9) may be graphically represented by a TEG, where each vertex represents an actor firing in a single graph iteration (the weight on an edge is then equal to the execution time of the actor corresponding to the edge's sink). To distinguish these graphs from arbitrary HSDF graphs, we choose to refer to these TEGs as *Linear Constraint Graph* (LCG) and refer to the tokens in this graph with the term *delay*. Note that the indegree of each vertex is equal to the indegree of the corresponding actor in the SDF graph. The construction of the LCG from a consistent SDF graph is outlined in Algorithm 1. Figure 1(d) shows the TEG obtained by applying Algorithm 1 to the SDF graph of Figure 1(a).

Algorithm 1 Transforms a consistent SDF graph into an LCG

Let $\mathcal{G}$ be a simple, consistent SDF graph with repetition vector q and $\mathcal{H}$ be an empty LCG.
for each actor a in $\mathcal{G}$ **do**
  Add vertices $a_1 \dots a_{q_a}$ to $\mathcal{H}$
end for
for each channel ab in $\mathcal{G}$ **do**
  for $j = 1 \dots q_b$ **do**
    $i \leftarrow \tilde{\pi}_{ab}(j)$
    add edge $a_i b_j$ with $-\delta_{ab}(j)$ delays to $\mathcal{H}$
  end for
end for

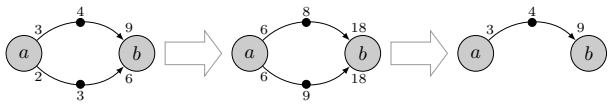


Fig. 2. An SDF multigraph may be transformed into a simple directed graph by equalising the rates of channels between two actors. Only a single channel that has the minimum number of initial tokens needs to be retained.

A. Reducing consistent SDF multigraphs

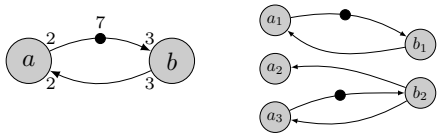
In an SDF multigraph, multiple channels may exist between two actors, in which case the channels are said to be *parallel*. Each of these parallel channels results in a different set of max-plus equations. However, in a consistent SDF multigraph, parallel channels may be *sorted* by the strength of the precedence constraints they imply. A set of parallel channels may then be replaced by the channel that imposes the strongest constraint.

In order to sort channels by the strength of their imposed precedence constraints, their rates first need to be equalised: Since multiplying a channel's rates and initial tokens with the same constant does not alter the channel's imposed precedence constraint, we may choose suitable integers and multiply each channel's production (or consumption) rate such that each channel has the same production (or consumption) rate. In case the SDF multigraph is consistent, each of the parallel channels will then have the same consumption (or production) rate as well (this follows directly from the fact that in a consistent graph we have $m_{ab} \cdot q_a = n_{ab} \cdot q_b$ for any channel ab).

If parallel channels have equal production rates and equal consumption rates, the strongest precedence constraints are imposed by the channel with the fewest tokens. Hence, for a pair of parallel channels, we may remove the SDF channel that, after equalising the channels' rates, has the most initial tokens (see Figure 2). Note that this is a straightforward generalisation of the transformation of an HSDF multigraph to a simple graph found in [11].

B. Linear Constraint Graphs and equivalent HSDF graphs

Although the construction of an LCG from an SDF graph and the transformation of an SDF graph into an equivalent HSDF graph are similar, there is an important, fundamental difference. When transforming an SDF graph into an equivalent HSDF graph, an SDF channel is interpreted as data *communication*. A max-plus algebraic description of the timing behaviour of an SDF graph however, treats channels as data *dependencies* and only considers the strongest of these dependencies. As a result, an LCG contains (much) fewer edges than an equivalent HSDF graph, which is illustrated in Figure 1. Fewer edges also means fewer cycles, which severely impacts the efficiency of maximum-cycle ratio algorithms [4]. Furthermore, as we will demonstrate in the following section, the structure of an LCG may be exploited to allow for a much more efficient analysis.



(a) SDF graph (b) Constraint graph

Fig. 3. An SDF graph and its corresponding constraint graph. The constraint graph contains two cycles, each of which has the same cycle ratio.

VI. Throughput analysis of SDF graphs

The throughput of an SDF graph is the average number of iterations that are completed per unit of time. Since the LCG of an SDF graph has exactly one vertex for each firing of a single graph iteration, the SDF graph's throughput is equal to the minimum of the average number of firings per unit of time over all vertices in the constraint graph. It is well known (see for example [4] or [12]) that this minimum average firing time is determined by the *maximum cycle ratio* of the LCG, which is the maximum of the cycle ratios of all simple cycles in the graph, where the cycle ratio λ of a cycle C is defined as:

$$\lambda(C) = \frac{\sum_{a_i b_j \in C} \tau_b}{\sum_{a_i b_j \in C} -\delta_{ab}(j)} \tag{10}$$

A cycle that has the maximum cycle ratio is said to be a *critical cycle*. Note that a constraint graph may contain multiple critical cycles, see for example the constraint graph shown in Figure 3, which contains 2 critical cycles.

Since the Linear Constraint Graph of an SDF graph may be quite large, we shall first investigate its structure for regularity and redundancy that may be exploited. This structure becomes especially apparent when constraint graphs are depicted in the column-wise representation of Figure 4(b): we group vertices that represent firings of the same SDF actor into columns, and (vertically) order the vertices by the index of the firing they represent. The following sections describe the structural properties of Linear Constraint Graphs, starting with the simplest graph (the LCG that represents a single SDF channel), followed by more complex graphs that represent SDF paths, cycles and, finally, full SDF graphs.

A. Defining the structure of the Linear Constraint Graph: parallel and crossing edges

The structure of the Linear Constraint Graph that represents a single SDF channel emerges from the in-order token con-

sumption (tokens are consumed in the same order they are produced) and the SDF graph's balance equations.

The graph's balance equations state that on SDF channel ab, tokens produced by q_a firings of actor a are consumed by q_b firings of actor b. Due to the presence of initial tokens on the channel, these q_b firings may span at most two consecutive graph iterations. In other words, the number of delays on any two edges in the LCG of an SDF channel can not differ by more than one.

The in-order token consumption orders the number of delays on edges leaving vertices that represent consecutive firings of actor a: if a_i and a_j are two vertices in the LCG with $j > i$, then the number of delays on any edge leaving a_j can not be lower than the number of delays on any edge leaving a_i.

A direct result of these two basic rules is that for disjoint edges (two edges are disjoint if they share neither source nor sink) that represent the same SDF channel, the number of delays may be inferred simply by looking at the firing indices of the edges' sources and sinks. We introduce the following terminology to formalise the structure of a linear constraint graph of a single SDF channel:

Definition 1 (parallel and crossing edges)**.** *Let $e_1 = a_{i_1}b_{j_1}$ and $e_2 = a_{i_2}b_{j_2}$ be two edges (with $i_1 = \tilde{\pi}_{ab}(j_1)$ and $i_2 = \tilde{\pi}_{ab}(j_2)$) in the linear constraint graph that represents* SDF *channel ab. The relations* parallel *and* crossing *are defined as follows:*

- *e_1 is crossing with e_2, denoted $e_1 \nparallel e_2$, if:*

$$(i_2 > i_1 \wedge j_2 < j_1) \vee (i_2 < i_1 \wedge j_2 > j_1)$$

- *e_1 is parallel with e_2, denoted $e_1 \parallel e_2$, if:*

$$(i_1 > i_2 \wedge j_1 > j_2) \vee (i_1 < i_2 \wedge j_1 < j_2)$$

Two crossing edges can not have the same number of delays, since in that case tokens would be consumed out of order (tokens produced by a firing are consumed before tokens produced by an earlier firing are consumed). Therefore, one edge must have precisely one delay more than the other. This is formalised in the following proposition:

Proposition 1 (different delays on crossing edges)**.** *Let $a_{i_1}b_{j_1}$ and $a_{i_2}b_{j_2}$ be two* crossing *edges in the constraint graph that represents* SDF *channel ab, with delays k_1 and k_2, respectively, and with $i_2 > i_1$ (and thus $j_2 < j_1$). Then $k_2 = k_1 + 1$.*

Proof: First of all, since $i_1 = \tilde{\pi}_{ab}(j_1)$ and $i_2 = \tilde{\pi}_{ab}(j_2)$, we have $\tilde{\pi}_{ab}(j_2) > \tilde{\pi}_{ab}(j_1)$. Furthermore, since $j_2 < j_1$, we have $\pi_{ab}(j_2) < \pi_{ab}(j_1)$. It then follows that $\delta_{ab}(j_2) < \delta_{ab}(j_1)$. Edge $a_{i_2}b_{j_2}$ thus has more delays (recall that the number of delays on an edge is $-\delta_{ab}(j)$) than edge $a_{i_1}b_{j_1}$. The fact that the number of delays on the two edges can not differ by more than one completes the proof. ∎

Following a similar reasoning we may infer that two parallel edges in the constraint graph of SDF channel ab carry the same number of delays:

Proposition 2 (same delays on parallel edges)**.** *Let $a_{i_1}b_{j_1}$ and $a_{i_2}b_{j_2}$ be two* parallel *edges in the constraint graph that represents* SDF *channel ab, with delays k_1 and k_2, respectively, and with $i_2 > i_1$ (and thus $j_2 > j_1$). Then $k_2 = k_1$.*

Proof: $j_2 > j_1$ gives $\pi_{ab}(j_2) > \pi_{ab}(j_1)$ and $i_2 > i_1$ gives $\tilde{\pi}_{ab}(j_2) > \tilde{\pi}_{ab}(j_1)$. It then follows that $\delta_{ab}(j_2) = \delta_{ab}(j_1)$. ∎

B. Properties of paths and cycles in linear constraint graphs

The relationship between parallel (and crossing) edges and their delays can be extended to disjoint paths (two paths are disjoint if no vertex is shared between the paths) in an LCG. Instead of two actors a and b and one channel ab, we now consider the situation in which we have n actors $a_1, a_2, \ldots, a_n$, and (at least the) SDF channels a_1a_2, a_2a_3, …, $a_{n-1}a_n$. In the LCG such a sequence of channels is represented by several paths, the indices of which are now denoted as superscripts, so $(a_1^{i_1}a_2^{i_2}\ldots a_{n-1}^{i_{n-1}}a_n^{i_n})$ denotes such a path in which channel a_ka_{k+1} is represented by the edge $a_k^{i_k}a_{k+1}^{i_{k+1}}$. We refer to a path by the sequence of its vertices and denote the delay of a path P (i.e., the sum of the delays of its edges) by $|P|_d$. Paths are assumed to be *simple*, i.e., no vertex is repeated in a path. Similar to the definitions for edges in an LCG, we introduce the following terminology:

Definition 2 (parallel and crossing paths)**.** *Let $\mathcal{G}$ be the* LCG *representing a path $P = (a_1a_2\ldots a_{n-1}a_n)$ in a consistent SDF graph. Furthermore, let $P_i = (a_1^{i_1}a_2^{i_2}\ldots a_{n-1}^{i_{n-1}}a_n^{i_n})$ and $P_j = (a_1^{j_1}a_2^{j_2}\ldots a_{n-1}^{j_{n-1}}a_n^{j_n})$ be two disjoint paths in $\mathcal{G}$. Then P_i and P_j are:*

- parallel*, denoted $P_i \parallel_p P_j$, if $(j_1 > i_1 \wedge j_n > i_n) \vee (j_1 < i_1 \wedge j_n < i_n)$.*
- crossing*, denoted $P_i \nparallel_p P_j$, if $(j_1 > i_1 \wedge j_n < i_n) \vee (j_1 < i_1 \wedge j_n > i_n)$.*

Analogous to the case of disjoint edges, the relative delays on parallel and crossing paths depend only on the first and last vertices of the paths. This property can be derived from Propositions 1 and 2 in a straightforward way, using induction on the number of actors represented by the paths, and is formally stated in the following lemma.

Lemma 1 (relative delays on disjoint paths)**.** *Let $P_i = (a_1^{i_1}\ldots a_n^{i_n})$ and $P_j = (a_1^{j_1}\ldots a_n^{j_n})$ be two paths representing the same n actors and $n-1$ channels, with $i_1 > j_1$. Then:*

(1) *$|P_i|_d = |P_j|_d$ if P_i and P_j are parallel;*
(2) *$|P_i|_d = |P_j|_d + 1$ if P_i and P_j are crossing.*

Proof: We prove this by induction on the number of actors n of the paths. First of all, note that $a_k^{i_k} \neq a_k^{j_k}$ for all $k = 1, 2, \ldots, n$ since two distinct edges can not have the same sink (and by assumption the inequality holds for $k = 1$). In case $n = 2$ both paths are edges and we obtain the result from Propositions 1 and 2. Next, suppose the result holds for all paths with at most n actors, and consider two paths P_i and P_j with $n + 1 \geq 3$ actors and with final edges $e_i = a_n^{i_n}a_{n+1}^{i_{n+1}}$ and $e_j = a_n^{j_n}a_{n+1}^{j_{n+1}}$, respectively. We assume again that $i_1 > j_1$. There are four cases to consider, depending on the relative orders of i_n, j_n and i_{n+1}, j_{n+1}. If $j_n < i_n$ ($j_n > i_n$) and $j_{n+1} < i_{n+1}$, then the paths $P_i - a_{n+1}^{i_{n+1}}$ and

$P_j - a_{n+1}^{j_{n+1}}$ are parallel (crossing) and $a_n^{i_n} a_{n+1}^{i_{n+1}}$ and $a_n^{j_n} a_{n+1}^{j_{n+1}}$ are parallel (crossing), and the claims follow by induction and by Propositions 1 and 2. The other two cases are similar. ■

An important consequence of the above lemma is the following: If we have three pairwise disjoint paths and each crosses at least one of the other two paths, then two of these three paths must be parallel. Furthermore, if a path crosses two other, disjoint paths, these two paths must be parallel. These two implications are captured in the following corollary, which follows directly from the above lemma.

Corollary 1 (restrictions on disjoint paths)**.** *Let* $P_i = (a_1^{i_1} \dots a_n^{i_n})$*,* $P_j = (a_1^{j_1} \dots a_n^{j_n})$ *and* $P_k = (a_1^{k_1} \dots a_n^{k_n})$ *be three* disjoint *paths, with* $k_1 > j_1 > i_1$*. Furthermore let* P_i *cross* P_j*. Then:*

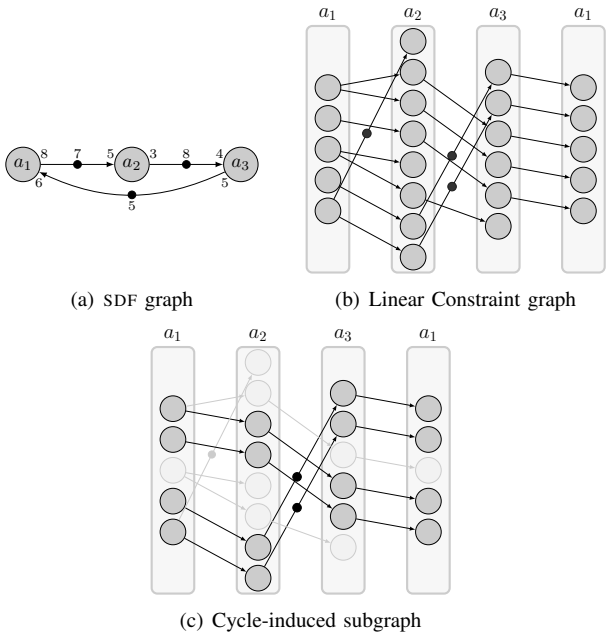
(1) *if* P_j *and* P_k *cross, then* P_i *and* P_k *do not cross;*
(2) *if* P_j *and* P_k *are parallel, then* P_i *and* P_k *cross.*

Proof: Both claims may be easily proven by contradiction using Lemma 1. As the proofs for both claims is similar, it suffices to prove claim (1). Assume paths P_i and P_k do cross. Then by Lemma 1, P_i, P_j and P_k must all have different delays, in particular $|P_i|_d \neq |P_j|_d$. But since $k_1 > j_1$ and $k_1 > i_1$, by the same lemma we have $|P_k|_d = |P_i|_d + 1$ and $|P_k|_d = |P_j|_d + 1$, which implies the contradiction $|P_i|_d = |P_j|_d$. ■

Note that the results of the above lemma and its corollary also hold if we consider walks instead of paths (so nodes may be repeated) in the SDF graph, as long as the sequence of actors of the two walks is the same and between every pair of successive actors there is a channel represented by an edge in the walks. In particular, the result also holds for *closed walks*, i.e. walks for which the first and last actor are identical. We show below how this can help us to analyse the behaviour of an SDF graph that consists of a single (simple) directed cycle.

Let the consistent SDF graph with repetition vector q be a directed cycle consisting of actors $a_1, a_2, \dots, a_n$ and channels $a_1a_2, a_2a_3, \dots, a_{n-1}a_n, a_na_1$. Then the LCG has a sequence of $q_k = q_{a_k}$ nodes $a_k^1, a_k^2, \dots, a_k^{q_k}$ for every actor a_k, and edges $a_k^i a_{k+1}^j$ (and $a_n^i a_1^j$) representing the firings, as defined before. For convenience, we repeat the sequence of q_1 nodes for actor a_1 at the end and think of the LCG as an array of $n+1$ columns, where the sequences of nodes representing the actors are ordered from left to right, and where the leftmost and rightmost sequences are identical, representing the actor $a = a_1$ (see Figure 4(b)). To distinguish these two sequences, we denote the leftmost sequence by $L = L(a)$ and the rightmost by $R = R(a)$.

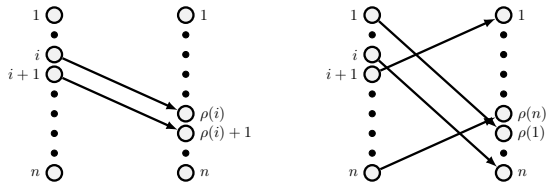
Now consider the Linear Constraint Graph's *cycle-induced* subgraph. This graph is obtained by removing all edges and vertices from the LCG graph that do not lie on a cycle (see Figure 4(c)) and consists of a number of disjoint paths (since each vertex has an indegree of one), each of which starts in L and ends in R. Furthermore, if there are n paths in the subgraph, each column contains precisely n vertices. Because each path has the same length and the (relative) delay of a path is (by Lemma 1), fully determined by its start and end



(a) SDF graph (b) Linear Constraint graph

(c) Cycle-induced subgraph

Fig. 4. An SDF graph and its corresponding LCG in a column representation, with actor a_1 duplicated. The rightmost figure depicts the cycle-induced subgraph of the LCG, which is obtained by retaining all nodes and edges that lie on a cycle. There are two cycles of length 6 in the LCG, and both cycles have a delay of one.



(a) parallel paths (b) crossing paths

Fig. 5. Structure of the permutation ρ

vertices, we choose to compactly represent the cycle-induced subgraph by a *permutation* $\rho : \{1, \dots, n\} \to \{1, \dots, n\}$. This permutation maps the index of a vertex in L to the index of a vertex in R, where the index of a vertex is based on the natural ordering of vertices representing the same actor (i.e., $a_k^i < a_k^j$ if $i < j$). In the remainder of this section, we shall refer to vertices (in L and R) by their index; paths in the cycle-induced subgraph then start in a node $i \in \{1, \dots, n\}$ and terminate in $\rho(i) \in \{1, \dots, n\}$.

Representing the cycle-induced subgraph as a permutation on a set of integers reveals a clear structure in Linear Constraint Graphs that represent SDF cycles. Consider the case where two parallel paths start in subsequent vertices, indexed i and $i+1$. Using the lemma stated above and its corollary we may derive that these paths also terminate in subsequent vertices, or $\rho(i+1) = \rho(i) + 1$ (see Figure 5(a)), which leads to the following proposition:

Proposition 3. *Let* P_i *and* P_{i+k} *be two parallel paths that start in vertices indexed* i *and* $i+k$*, respectively, with* $k > 0$*. Then* $\rho(i+k) = \rho(i) + k$*.*

Proof: Let $k = 1$. Note that since P_i and P_{i+1} are parallel, we have $\rho(i+1) > \rho(i)$. Assume $\rho(i+1) > \rho(i) + 1$. There must exist j such that $\rho(j) = \rho(i) + 1$. Let P_j be the path that connects j with $\rho(i) + 1$. In case $j < i$, we have $P_j \nparallel_p P_i$ and $P_j \parallel_p P_{i+1}$. By Corollary 1 however, we have $P_j \nparallel_p P_{i+1}$, which is a contradiction. The assumption that $j > i + 1$ leads to a contradiction in a similar way. We thus have $\rho(i+1) = \rho(i) + 1$, and by straightforward induction on k it follows that $\rho(i+k) = \rho(i) + k$. ∎

For crossing paths, a similar relation in terms of ρ exists. This is illustrated in Figure 5(b) and may be understood by considering two crossing paths P_i and P_{i+1} that start in subsequent vertices i and $i+1$, respectively. We may divide the set of paths in two subsets: the first subset contains all paths starting in vertices $1, 2, \ldots, i$, and the second contains all paths starting in vertices $i+2, \ldots, n$. By Corollary 1, both subsets contain pairwise parallel paths. Furthermore, each path in one subset crosses all other paths in the other subset. As a consequence, we must have $\rho(i) = n$ and $\rho(i+1) = 1$. The following proposition formally generalises this conclusion:

Proposition 4. *Let P_i and P_{i+1} be two crossing paths that start in subsequent vertices indexed i and $i+1$, respectively. Then $\rho(i+k) = k$ for $k > 0$ and $\rho(i-k) = n - k$ for $k \geq 0$.*

Proof: Assume $\rho(i+1) > 1$. There must exist j such that $\rho(j) = 1$. Let P_j be the path that connects j with $\rho(j)$. In case $j < i$, we have $P_j \parallel_p P_{i+1}$ and $P_j \parallel_p P_i$. By Corollary 1 however, we have $P_i \nparallel_p P_j$, which is a contradiction. In case $j > i + 1$, we have $P_j \nparallel_p P_i$ and $P_j \nparallel_p P_{i+1}$, which again contradicts Corollary 1, thus $\rho(i+1) = 1$. Following a similar reasoning it follows that $\rho(i) = n$. By straightforward induction on k it follows that $\rho(i+k) = k$ and $\rho(i-k) = n - k$. ∎

We are now ready to move from paths in the LCG to cycles. A simple cycle in the LCG may be constructed by repeatedly applying the permutation ρ until the start vertex is reached again. For this, let $\rho^{k+1}(i) = \rho(\rho^k(i))$ and $\rho^1(i) = \rho(i)$. Due to the structure of the LCG, any two (simple) cycles in the graph must have the same cycle ratio. This is formally stated in the following theorem and proven by exploiting the definition of ρ. Note that the theorem is not restricted to *simple* SDF cycles, but applies to any closed walk in the SDF graph.

Theorem 1 (Two cycles have the same length and delay)**.** *Let $\mathcal{G}$ be the cycle-induced subgraph of the Linear Constraint Graph corresponding to an* SDF *cycle, and let $C_i = \{i, \rho(i), \ldots, \rho^{n_i}(i) = i\}$ and $C_j = \{j, \rho(j), \ldots, \rho^{n_j}(j) = j\}$ be two disjoint simple cycles in $\mathcal{G}$. Then $n_i = n_j$ and $|C_i|_d = |C_j|_d$.*

Proof: Let $\mathcal{G}$ consist of N paths. We may define ρ as follows: $\rho(i+k) = (\rho(i) + k - 1) \bmod N + 1$, for some $k \in \{1, \ldots, n\}$. Using the definition of ρ, the length l of a cycle starting at vertex a may be calculated by finding the minimum positive value of l that satisfies the linear congruence: $a + l \cdot k \equiv a \pmod N$. This solution l is independent of a, which implies $n_i = n_j$. Since a cycle is a path that starts and ends in the same vertex, C_1 and C_2 are parallel paths. Then by Lemma 1, C_1 and C_2 must have the same delay. ∎

C. Throughput analysis of arbitrary SDF graphs

Theorem 1 provides an efficient approach to throughput analysis of SDF cycles. Rather than constructing the full LCG of an SDF cycle, it suffices to pick a random vertex and follow edges in reverse direction until a cycle is found. By Theorem 1, this cycle must then be (one of) the graph's critical cycle(s). A straightforward question is whether the same approach works for arbitrary SDF graphs: Can we choose a random vertex and restrict the search for a critical cycle to the subgraph reachable (by following edges in reverse direction) from the initial vertex? In this section we show that this is indeed the case for strongly connected SDF graphs (note that the throughput of an SDF graph that is not strongly connected may be calculated from the throughputs of its strongly connected components, as is described in [7]). An important property in understanding why this approach works, concerns the reachability of vertices in an LCG, which we formalise in the following proposition:

Proposition 5 (reachability)**.** *Let a and b be actors in a consistent and strongly connected* SDF *graph with repetition vector q. Then for each vertex b_j that represents the j^{th} firing of actor b there exists an i such that the* LCG *contains a path from a_i to b_j.*

Proof: In a strongly connected SDF graph, each actor has at least one incoming channel. As a result, in the LCG, each vertex has a nonzero indegree and the claim trivially follows. ∎

In words, Proposition 5 states that if, by following edges in reverse, actor a is reachable from actor b, then from any vertex that represents a firing of actor b we may reach a vertex that represents a firing of actor a. We use this fact together with Theorem 1 to prove that only a subgraph of the LCG needs to be explored for its critical cycle:

Theorem 2 (Subgraph analysis)**.** *Let $\mathcal{G}$ be the* LCG *that represents (consistent)* SDF *graph $\mathcal{G}_{sdf}$, and s an arbitrary vertex in $\mathcal{G}$. Furthermore, let $\mathcal{H}$ be the induced subgraph of $\mathcal{G}$ that consists of those vertices from which a path to s exists. The maximum cycle ratio of $\mathcal{H}$ is the maximum cycle ratio of $\mathcal{G}$.*

Proof: Let the critical cycle in $\mathcal{G}$ be $C = (a_1^{i_1} \ldots a_n^{i_n} = a_1^{i_1})$. Cycle C is contained in the LCG that corresponds to a cycle W in the SDF graph, with $W = (a_1, a_2, \ldots, a_n = a_1)$ (Note that this cycle may be a walk in the SDF graph, i.e., vertices may be repeated).

We prove the theorem by contradiction. Let $\mathcal{H}$ not contain C. Then s obviously does not lie on C. Furthermore, there is no path from a vertex on C to s, since if this were the case, C would be in $\mathcal{H}$.

Now choose a vertex $v = a_1^{i_m}$ that is not reachable from C, but from which there is a path to s (i.e., v is in $\mathcal{H}$). By Proposition 5, such a vertex can always be found. If in the LCG that represents SDF cycle W we follow edges in reverse direction from v, then eventually a cycle C' will be found. By

	Mimic DSP	Large HSDFG	Long transient
State-space exploration			
avg $[s]$	1.1×10^{-3}	6.6×10^{-2}	4.2×10^{-1}
var $[s^2]$	2.1×10^{-2}	2.3×10^{2}	1.7×10^{1}
Linear Constraint Graph			
avg $[s]$	5.6×10^{-5}	1.1×10^{-3}	1.7×10^{-4}
var $[s^2]$	5.3×10^{-5}	2.7×10^{-2}	1.4×10^{-4}

TABLE I
EXPERIMENTAL RESULTS OF THE THREE METHODS

	Mimic DSP	Large HSDFG	Long transient
SDF Graphs			
vertices	20.0	13.4	284
edges	24.4	21.7	359
Equivalent HSDF Graphs			
vertices	1008	8166	284
edges	3151	95321	359
Linear Constraint Graphs			
vertices	119 (11.8%)	754 (9.2%)	284 (100%)
edges	151 (4.8%)	1202 (1.3%)	359 (100%)

TABLE II
GRAPH SIZES

Theorem 1, C' has both the same length and the same delay as C. We may thus restrict our search to $\mathcal{H}$. ∎

Theorem 2 implies that it is not necessary to explore the entire LCG for its critical cycle. More specifically, it does not matter whether the LCG is strongly connected or not. We may thus, in a similar way to the approach proposed in the previous section, choose an arbitrary root vertex in the LCG and search the induced subgraph that consists of vertices from which the root vertex is reachable, for its critical cycle.

VII. RESULTS

To evaluate our approach, we have applied the algorithm described in [12], to find a timed event graph's maximum cycle ratio to the LCG's obtained from SDF graphs. For the sake of comparison, we have used the same three testsets that were used in [6] to compare the performance of state-space exploration with HSDF-based approaches. We remark that in [6] it was found that some of the SDF graphs that lead to large HSDF graphs could not be analysed within 30 minutes.

For each SDF graph in the three testsets, the number of vertices and edges in the LCG as well as the number of actors and channels in equivalent HSDF graphs was recorded. Results were obtained on an Intel Xeon CPU core running at 2.40GHz within a 24-core machine with 64GB of RAM. Table I shows the average and variance of the measured runtimes for the two approaches on the three different test sets. The results clearly indicate that our approach based on the analysis of an LCG outperforms the simulation-based approach by several orders of magnitude.

Table II shows the average reduction achieved by an LCG when compared to an SDF graph's equivalent HSDF graph. Note that the "Long transient" testset contains HSDF graphs, which can not be represented more compactly by an LCG. The percentages included in the last two rows of the table indicate the amount of vertices and edges relative to the equivalent HSDF graph.

VIII. CONCLUSION AND FUTURE WORK

In this paper we have presented a novel approach to throughput analysis of SDF graphs. At the basis of this approach is a max-plus algebraic representation of a consistent SDF graph as a linear, periodically time-variant system, followed by a transformation into a linear time-invariant system by a change of variables. The Linear Constraint Graphs we derive are smaller (i.e., have fewer vertices and edges) than equivalent HSDF graph usually derived from an SDF graph. Furthermore, the regular structure of an LCG (the properties of which we thoroughly prove) may be exploited, which leads to an approach to throughput analysis that is faster than the state-of-the-art state-space exploration method.

We are convinced that the proven regular structure of the constraint graph provides the basis for new and efficient timing analysis techniques for SDF graphs.

REFERENCES

[1] Jean Cochet-terrasson, Guy Cohen, Stéphane Gaubert, Michael Mc Gettrick, and Jean-pierre Quadrat. Numerical Computation of Spectral Elements in Max-Plus Algebra, 1998.

[2] Guy Cohen, Stéphane Gaubert, and Jean-Pierre Quadrat. Max-plus algebra and system theory: Where we are and where to go now. *Annual Reviews in Control*, 23:207–219, January 1999.

[3] Guy Cohen, Geert Jan Olsder, and Jean-pierre Quadrat. *Synchronization and linearity*. Wiley New York, 1992.

[4] Ali Dasdan. Experimental analysis of the fastest optimum cycle ratio and mean algorithms. *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, 9(4):385, 2004.

[5] Marc Geilen. Reduction techniques for synchronous dataflow graphs. *Annual ACM IEEE Design Automation Conference*, pages 911–916, 2009.

[6] A. H. Ghamarian, M. C. W. Geilen, S. Stuijk, T. Basten, B. D. Theelen, M. R. Mousavi, A. J. M. Moonen, and M. J. G. Bekooij. Throughput Analysis of Synchronous Data Flow Graphs. *ACSD*, 2006.

[7] A.H. Ghamarian, M. Geilen, T. Basten, B. Theelen, M.R. Mousavi, and S. Stuijk. Liveness and boundedness of synchronous data flow graphs. *FMCAD06*, (August):68–75, 2006.

[8] B. Heidergott, Geert Jan Olsder, and Jacob van der Woude. *Max Plus at Work: modeling and analysis of synchronized systems*. Princeton University Press, 2006.

[9] Kazuhito Ito and Keshab K. Parhi. Determining the minimum iteration period of an algorithm. *Journal of VLSI Signal Processing*, 11(3):229–244, December 1995.

[10] E.A. Lee and D.G. Messerschmitt. Synchronous data flow. *Proceedings of the IEEE*, 75(9):1235–1245.

[11] Sundararajan Sriram and Shuvra S. Bhattacharyya. Embedded Multiprocessors: Scheduling and Synchronization. February 2009.

[12] N Young, R Tarjan, and J Orlin. Faster Parametric Shortest Path and Minimum Balance Algorithms. *ArXiv Computer Science e-prints*, May 2002.