Conference Paper

On Strong and Weak Sustainability, with an Application to Self-Suspending Real-Time Tasks

Felipe Cerqueira
Geoffrey Nelissen
Björn B. Brandenburg

CISTER-TR-180502
On Strong and Weak Sustainability, with an Application to Self-Suspending Real-Time Tasks

Felipe Cerqueira, Geoffrey Nelissen, Björn B. Brandenburg

*CISTER Research Centre
Polytechnic Institute of Porto (ISEP-IPP)
Rua Dr. António Bernardino de Almeida, 431
4200-072 Porto
Portugal
Tel.: +351.22.8340509, Fax: +351.22.8321159
E-mail: http://www.cister.isep.ipp.pt

Abstract

Motivated by an apparent contradiction regarding whether certain scheduling policies are sustainable, we revisit the topic of sustainability in real-time scheduling and argue that the existing definitions of sustainability should be further clarified and generalized. After proposing a formal, generic sustainability theory, we relax the existing notion of (strongly) sustainable scheduling policy to provide a new classification called weak sustainability. Proving weak sustainability properties allows reducing the number of variables that must be considered in the search of a worst-case schedule, and hence enables more efficient schedulability analyses and testing regimes even for policies that are not (strongly) sustainable. As a proof of concept, and to better understand a model for which many mistakes were found in the literature, we study weak sustainability in the context of dynamic self-suspending tasks, where we formalize a generic suspension model using the Coq proof assistant and provide a machine-checked proof that any JLFP scheduling policy is weakly sustainable with respect to job costs and variable suspension times.
On Strong and Weak Sustainability, with an Application to Self-Suspending Real-Time Tasks

Felipe Cerqueira
Max Planck Institute for Software Systems (MPI-SWS), Kaiserslautern, Germany
felipe@mpi-sws.org

Geoffrey Nelissen
CISTER Research Centre, ISEP, Polytechnic Institute of Porto (IPP), Porto, Portugal
grrpn@isep.ipp.pt

Björn B. Brandenburg
Max Planck Institute for Software Systems (MPI-SWS), Kaiserslautern, Germany
bbb@mpi-sws.org

Abstract
Motivated by an apparent contradiction regarding whether certain scheduling policies are sustainable, we revisit the topic of sustainability in real-time scheduling and argue that the existing definitions of sustainability should be further clarified and generalized. After proposing a formal, generic sustainability theory, we relax the existing notion of (strongly) sustainable scheduling policy to provide a new classification called weak sustainability. Proving weak sustainability properties allows reducing the number of variables that must be considered in the search of a worst-case schedule, and hence enables more efficient schedulability analyses and testing regimes even for policies that are not (strongly) sustainable. As a proof of concept, and to better understand a model for which many mistakes were found in the literature, we study weak sustainability in the context of dynamic self-suspending tasks, where we formalize a generic suspension model using the Coq proof assistant and provide a machine-checked proof that any JLFP scheduling policy is weakly sustainable with respect to job costs and variable suspension times.

2012 ACM Subject Classification Software and its engineering → Real-time schedulability

Keywords and phrases real-time scheduling, sustainability, self-suspending tasks, machine-checked proofs

Digital Object Identifier 10.4230/LIPIcs.ECRTS.2018.26

Funding This work was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — 391919384, and partially supported by National Funds through FCT (Portuguese Foundation for Science and Technology) within the CISTER Research Unit (CEC/04234).

1 What Really is Sustainability?

Since the seminal paper by Liu and Layland [14], the analysis and certification of real-time systems has often relied on the fundamental notion of sustainability [6], which at a high level expresses the idea that “if a system is proven to be safe under extreme conditions, then it will remain safe if the conditions improve at runtime.” By allowing system designers to focus on such extreme scenarios (rather than the entire state space of the system), sustainability plays a fundamental role in the design, prototyping, analysis, and validation of real-time systems.

One common application of this principle is to determine the schedulability of the system by identifying worst-case scheduling scenarios. For example, any schedulability analysis for uniprocessor fixed-priority (FP) scheduling of sporadic tasks [15] that assumes that jobs
execute for their worst-case execution time (WCET) or arrive at maximum rate exploits the fact that the FP scheduling policy for sporadic tasks is sustainable, i.e., the occurrence of “better” job parameters (namely, larger inter-arrival times or lower execution times) at runtime does not cause any deadline miss.

While precursors to this concept were already identified and proven in earlier papers [11, 12], the general concept of sustainability was first formalized by Baruah and Burns [6, 7] and later refined by Baker and Baruah [3]. Although the definition by Baker and Baruah is more rigorous than the original definition, we argue in this paper that there is still a need for improvement in terms of clarity and precision.

To support our claim, in §1.1 and §1.2 we present an example in the context of uniprocessor scheduling with self-suspending tasks [16], where we show a scheduling policy that can be interpreted as both sustainable and not sustainable with respect to job execution times (also called job costs hereafter). Both claims are correct according to the existing definitions of sustainability and only depend on varying interpretations by the reader. This example shows that, despite being a well-established concept, the theory of sustainability needs further clarification and formalization.

1.1 Uniprocessor EDF Scheduling with Self-Suspicions is not Sustainable w.r.t. Job Costs

Consider uniprocessor earliest-deadline-first (EDF) scheduling of self-suspending tasks under the segmented suspension model. Self-suspending tasks are used to model workloads that may have their execution suspended at given times, for example, to perform remote operations on co-processors, acquire locks, wait for data, or synchronize with other tasks. The segmented self-suspending task model can be formalized as follows.

Definition 1 (Sporadic Task Model with Segmented Self-Suspicions). Let \( \tau \) be a task set and let \( J \) be a job set generated by \( \tau \). Each task \( T_i \in \tau \) is defined by a period \( p_i \), deadline \( d_i \) and a sequence of execution and suspension segments \( S_i = [e_i^1, s_i^1, e_i^2, s_i^2, \ldots, e_i^n] \). These
task parameters encode the constraints that any two jobs generated by $T_i$ must be separated by a minimum inter-arrival time $p_i$. Each job released by $T_i$ must finish its execution by a relative deadline $d_i$, and alternates between execution and suspension segments as defined by the sequence $S_i$. The execution time of the $k$-th execution segment of job $j$ is upper-bounded by $e_k^i$, and the suspension time of its $k$-th suspension segment is upper-bounded by $s_k^i$.

Next, let us recall the definition of sustainable policy as proposed by Burns and Baruah [7].

**Definition 2 (Sustainable Policy – original definition from [7]).** A scheduling policy and/or a schedulability test for a scheduling policy is sustainable if any system deemed schedulable by the schedulability test remains schedulable when the parameters of one or more individual tasks are changed in any, some, or all of the following ways: (i) decreased execution requirements; (ii) larger periods; (iii) smaller jitter; and (iv) larger relative deadlines.

As explained by Burns and Baruah [7], the interpretation of Definition 2 for scheduling policies concerns the values of job parameters at runtime: “[…] a scheduling policy that guarantees to retain schedulability if actual execution requirements during run-time are smaller than specified WCET’s, and if actual jitter is smaller than the specified maximum jitters, would be said to be sustainable with respect to WCET’s and jitter”.

Thus, in order to show that a scheduling policy is not sustainable with respect to execution requirements (i.e., job costs), we must find a counterexample that shows a job set $J$ that is schedulable under that policy, along with a job set $J_{\text{better}}$ with lower or equal job execution times that is not schedulable under the same policy.

Fig. 1 depicts such a counterexample for uniprocessor EDF scheduling on the segmented self-suspending task model, adapted from prior work by Abdeddaïm and Masson [2]. Fig. 1-(a) shows the original EDF schedule of three tasks $T_1$, $T_2$ and $T_3$, which contains no deadline misses. Next, by reducing the cost of $T_3$’s job by 1 time unit as shown in Fig. 1-(b), the different interleaving of suspension times during the time interval [13, 16) increases the interference incurred by task $T_1$, causing a deadline miss at time 18.

This counterexample, which is simple enough to make the claim non-disputable, proves that, according to Definition 2, EDF scheduling under the segmented suspension model is not sustainable with respect to job costs.

### 1.2 Uniprocessor EDF Scheduling with Self-Suspensions is Sustainable w.r.t. Job Costs

Consider the same platform, task model and scheduling policy as in §1.1, and recall the definition of sustainable policy proposed by Baker and Baruah [3].

**Definition 3 (Sustainable Policy – original definition from [3]).** Let $A$ denote a scheduling policy. Let $\tau$ denote any sporadic task system that is $A$-schedulable. Let $J$ denote a collection of jobs generated by $\tau$. Scheduling policy $A$ is said to be sustainable if and only if $A$ meets all deadlines when scheduling any collection of jobs obtained from $J$ by changing the parameters of one or more individual jobs in any, some, or all of the following ways: (i) decreased execution requirements; (ii) larger relative deadlines; and (iii) later arrival times with the restriction that successive jobs of any task $T_i \in \tau$ arrive at least $p_i$ time units apart.

Definition 3 is similar to Definition 2, except that it explicitly makes the difference between the notion of jobs and tasks. It requires the job set $J$ with original parameters to be generated by a task set $\tau$ that is $A$-schedulable, i.e., all job sets generated by $\tau$ exhibit no deadline misses when scheduled by $A$. However, note that the modified job set obtained from $J$ (which we call $J_{\text{better}}$) does not have to be generated by $\tau$. 
Now, we must check whether the counterexample in Fig. 1 is still valid. At a first glance, the job sets \( J \) and \( J_{\text{better}} \) depicted in Fig. 1-(a) and Fig. 1-(b) seem to prove that uniprocessor EDF scheduling with segmented self-suspending tasks is not sustainable with respect to job costs, according to Definition 3. After all, we can assume that job set \( J \) is generated for instance by some task set \( \tau = \{(p_1 = 12, d_1 = 6, S_1 = [2, 2, 2]), (p_2 = 9, d_2 = 7, S_2 = [2, 2, 2]), (p_3 = 10, d_3 = 10, S_3 = [2])\} \).

However, let us consider the alternative job set \( J_{\text{susp}} \) in Fig. 1-(c), in which the job of task \( T_3 \) has the original cost of 2 time units, and the suspension time of the second job of task \( T_2 \) is reduced by 1 time unit. Clearly, \( J_{\text{susp}} \) can be generated by task set \( \tau \), since the job costs are the same as in \( J \) and the suspension segments are no larger than those in \( J \), which is allowed by the segmented suspension model. Moreover, we can observe that in the schedule of \( J_{\text{susp}} \), task \( T_1 \) again misses a deadline at time 18.

Since job set \( J_{\text{susp}} \) generated by \( \tau \) is not schedulable, it is clear that \( \tau \) does not satisfy the assumption of being \( A \)-schedulable (i.e., EDF-schedulable) required by Definition 3. Therefore, job sets \( J \) and \( J_{\text{better}} \) in Fig. 1 are not a valid counterexample for establishing that the policy is not sustainable. Since the counterexample is not valid, what can we really say about the sustainability of this policy? Why do the two definitions disagree?

One aspect that is implicit but unclear in both definitions is whether all job parameters other than the sustainable parameter (i.e., job costs) must remain constant. In fact, as shown in \( J_{\text{susp}} \) from Fig. 1-(c), in some cases we can vary the other parameters (i.e., job suspension times) to compensate the increase in interference that would otherwise cause the scheduling anomaly. Since this parameter variation is allowed by the task constraints, this suggests that a task set that is schedulable for any possible job suspension times may in effect be resilient to scheduling anomalies on job costs, even though individual schedulable job sets are not.

In fact, by constructing job sets similar to \( J_{\text{susp}} \) in the example above, we provide a mechanized proof (i.e., a proof that is verified by the Coq proof assistant) in §4 that establishes that uniprocessor job-level fixed priority (JLFP) scheduling of sporadic tasks under the dynamic suspension model is, what we later define as, weakly sustainable with respect to job costs and variable suspension times.

Note that this result does not make the counterexample of Abdeddaïm and Masson incorrect. Their result is simply based on a different interpretation of sustainability where nothing but the job parameter under consideration for the sustainability property can vary between the compared schedules; thus, the results stated in §1.1 and §4 are both correct. In §3, we will complement the existing sustainability theory with the notions of strong and weak sustainability to distinguish those contradictory but correct interpretations of sustainability.

1.3 This Paper

The seemingly contradictory observations in §1.1 and §1.2 suggest the need for clarification in the definitions of sustainability, which are currently restricted to the standard sporadic task model and are not precise with respect to how parameters can vary across the original and modified job sets \( J \) and \( J_{\text{better}} \).

We believe that the solution to this problem lies in formalizing the abstract concepts of real-time scheduling meta-theory such as “job and task parameters” in a rigorous way, so that the different notions of sustainability can be stated precisely. Additionally, this approach allows transcribing those concepts into a proof assistant such as Coq to formalize and mechanically prove key results [8]. With that in mind, we propose a formal sustainability theory for real-time scheduling, which we present in §2.
Our goal in this paper is not only to clarify what sustainability means, but also to provide a foundation for more efficient schedulability analyses for policies that are sustainable with varying parameters (such as the suspension times in the example from §1.2), a new concept that we call weakly sustainable policy. The exact definition and implications of weak sustainability will be discussed in §3.

Finally, we apply this newly defined notion of weak sustainability in §4, where we formalize self-suspending tasks in Coq and mechanically prove that uniprocessor, job-level fixed priority (JLFP) scheduling of self-suspending tasks under the dynamic suspension model is weakly sustainable with respect to job costs and varying suspension times.

To summarize, this paper makes the following contributions:

1. a formal theory of sustainability in real-time scheduling, with definitions of sustainable policy [3, 7], sustainable analysis [3, 6, 7] and self-sustainable analysis [3] generalized to any scheduling policy and any task and platform models (§2);
2. the definition of the new notions of strongly and weakly sustainable policies (§3), and the corresponding composition rules (§3.2);
3. the first formalization of sustainability theory and real-time scheduling with self-suspending tasks in a proof assistant (§4.1 and online appendix [1]); and
4. a mechanized proof of weak sustainability of uniprocessor JLFP scheduling of dynamic self-suspending tasks with respect to job costs and varying suspension times (§4.2–§4.4 and online appendix [1]).

2 Formalization of Sustainability Theory

In this section, we formalize the theory of sustainability in real-time scheduling and characterize the basic notions of sustainability proposed in the literature, namely sustainable policy [3, 7], sustainable analysis [6, 7] and self-sustainable analysis [3].

Our motivation for developing this theory is twofold: we aim to (a) clarify and generalize the existing notions of sustainability so that they become compatible with any scheduling policy and any task and platform models, and (b) provide the theoretical support for defining the new concept of weak sustainability, which will be covered in §3 and mechanically proven in §4 for uniprocessor JLFP scheduling of dynamic self-suspending tasks.

Note that this section does not introduce fundamentally new concepts; rather, it spells out precisely common implicit assumptions about the task and platform models and gives a more formal presentation of the underlying real-time scheduling meta-theory, which will be used to mechanically prove the results (see §4).

In order to distinguish the different nuances of sustainability, one must be able to correlate the variation of job and task parameters with schedulability. Hence, we must formalize the system model and present the basic definitions related to jobs and tasks.

2.1 Platform Model

We begin by stating assumptions about the platform model, in particular the notions of time and platform parameters, which specify part of the scheduling problem to be solved. Note that all definitions in this paper are compatible with both discrete and dense time.

Definition 4 (Processor Platform). Let platform II be the system on which jobs are scheduled.

Definition 5 (Platform Parameter). Each platform II has a finite set of parameters $\mathcal{P}_{\text{plat}}$. 
Example 1 (Common Platforms). Examples of platforms include uniprocessor systems, identical multiprocessors [10], and uniform multiprocessors [4]. Multiprocessor platforms usually have an associated parameter $m \in P_{\text{plat}}$ that indicates the number of processors.

Note that Definition 5 does not limit the set of parameters defining a platform to its number of processors; in fact, the set of parameters $P_{\text{plat}}$ could also express the heterogeneity of the platform [5], its power consumption, or execution speed profiles [17]. We keep the set of parameters unspecified in order to retain maximal generality and not limit our definitions to a fixed subset of system models.

This approach is uncommon. Most works tend to limit their results to a specific system model (e.g., task-level fixed priority scheduling of sequential tasks on single or multi-core processors). Instead, we prefer generality to specificity, so that the concepts and properties presented hereafter can be instantiated for any scheduling problem.

2.1.1 Jobs

After discussing the general aspects of the system model, we now define a job set.

Definition 6 (Job Set). A job set $\mathcal{J}$ is a (potentially infinite) collection of jobs.

Next, in order to define sustainability without being restricted to a particular task model, we generalize the notion of a job parameter.

Definition 7 (Job Parameter). We denote as job parameters any finite set $P_{\text{job}}$, where each parameter $p \in P_{\text{job}}$ is a function over jobs.

Example 2. Common job parameters include $\text{cost}(j)$, the actual job execution time, $\text{arrival}(j)$, the absolute job arrival time, and $\text{deadline}(j)$, the relative job deadline. They may for instance also include the job suspension time in the case of self-suspending jobs, its level of parallelism and/or its energy consumption if such properties are of interest.

Next, we define the notion of scheduling policy, which specifies the strategy for selecting jobs to be scheduled, i.e., allocated to a processor at a given time.

Definition 8 (Scheduling Policy). Given a platform $\Pi$ and a job set $\mathcal{J}$ with job parameters $P_{\text{job}}$, we define a scheduling policy $\sigma$ as any algorithm that determines whether a job $j \in \mathcal{J}$ is scheduled at a given time $t$ on a processor $\pi \in \Pi$.

For job sets that have associated deadlines, we can also define whether they are schedulable.

Definition 9 (Schedulable Job Set). Assume that jobs have a deadline as one of their parameters. Then we say that a job set $\mathcal{J}$ is schedulable on platform $\Pi$ under policy $\sigma$ iff none of its jobs misses a deadline when scheduled on $\Pi$ under policy $\sigma$.

To compare different job sets, we must also be able to express how job parameters can vary across job sets (e.g., job costs may increase while their arrival times remain constant). For that, we define whether two job sets differ only by a given set of parameters.

Definition 10 (Varying Job Parameters in $V$). Consider any subset of job parameters $V \subseteq P_{\text{job}}$, which we call variable parameters, and consider two enumerated job sets $\mathcal{J} = \{j_1, j_2, \ldots\}$ and $\mathcal{J}' = \{j'_1, j'_2, \ldots\}$. We say that $\mathcal{J}$ and $\mathcal{J}'$ differ only by $V$ iff $|\mathcal{J}| = |\mathcal{J}'|$ and $\forall i, \forall p \in (P_{\text{job}} \setminus V)$, $p(j_i) = p(j'_i)$, where $|\mathcal{J}|$ denotes the cardinality of job set $\mathcal{J}$.

Example 3. By stating that $\{j_1, j_2\}$ and $\{j'_1, j'_2\}$ differ only by $V = \{\text{cost}\}$, we claim that jobs $j_1$ and $j'_1$ (respectively, $j_2$ and $j'_2$) are identical in all parameters other than cost. This is useful to formalize, for example, the idea that “schedulability is maintained when reducing only the cost of a job.”
2.1.2 Tasks

While some notions of sustainability apply exclusively to job sets, one can also describe how
the variation of task parameters affects schedulability analysis results. To be able to reason
at the task level, we begin by defining task set and task parameters.

► Definition 11 (Task Set). A task set $\tau$ is as a finite collection of tasks $\{T_1, \ldots, T_n\}$.

From a mathematical point of view, tasks are opaque objects, elements of an enumerated
set. Their utility comes from defining task parameters and using them to constrain the sets
of jobs that can possibly be generated at runtime (see Definition 14 below).

► Definition 12 (Task Parameters). We call task parameters any finite set $P_{task}$, where each
parameter $p \in P_{task}$ is a function over tasks.

► Example 4. Similar to the job parameters in Example 2, common task parameters include,
but are not limited to, $WCET(T_i)$, the worst-case execution time of task $T_i$, and $period(T_i)$,
the period or minimum inter-arrival time of task $T_i$.

Next, we define a task model, which determines how job sets are related to task sets.

► Definition 13 (Task Model). A task model $M$ is the collection of all task sets that can
be defined with given task parameters $P_{task}$, along with a set of constraints relating job
parameters with task parameters.

► Definition 14 (Generated Job Sets). Every task set $\tau \in M$ generates a (potentially infinite)
collection of job sets, denoted $jobsets(\tau) = \{J_1, J_2, \ldots\}$, with the condition that, for every
job set $J \in jobsets(\tau)$ and every job $j \in J$, (a) $j$ belongs to an associated task in $\tau$, denoted
task($j$), and (b) the job parameters of $j$ are constrained by the task parameters of task($j$),
as determined by $M$.

One example of such a task model constraint is the upper bound on job execution times.

► Example 5 (Constraint on Job Execution Time). Let $M$ be the sporadic task model. Let the
job parameter cost($j$) denote the actual execution time of job $j$ and let the task parameter
$WCET(T_i)$ denote the WCET of task $T_i$. For every job set $J$ generated by $M$, the cost of
each job $j \in J$ is upper-bounded by the cost of its task, i.e.,

$$\forall \tau \in M, \forall J \in jobsets(\tau), \forall j \in J, cost(j) \leq WCET(task(j)).$$

Using the notion of generated job sets, we can now define whether a task set is schedulable.

► Definition 15 (Schedulable Task Set). A task set $\tau \in M$ is schedulable on platform II under
scheduling policy $\sigma$ iff every generated job set $J \in jobsets(\tau)$ is schedulable on II under $\sigma$.

Similarly to Definition 10, in order to relate parameters across task sets, we define whether
two task sets differ only by a given set of parameters.

► Definition 16 (Varying Task Parameters in $V$). Consider any subset of task parameters
$V \subseteq P_{task}$, which we call variable parameters, and consider two task sets $\tau = \{T_1, T_2, \ldots\}$
and $\tau' = \{T'_1, T'_2, \ldots\}$. We say that $\tau$ and $\tau'$ differ only by $V$ iff $|\tau| = |\tau'|$ and $\forall i, \forall p \in (P_{task} \setminus V), p(T_i) = p(T'_i)$, where $|\tau|$ denotes the cardinality of task set $\tau$.

2.2 Generalized Sustainability Definitions

In this section, we use the basic concepts of jobs and tasks to formalize the notions of
sustainability found in the literature, namely sustainable policy (§2.2.1), sustainable analysis
(§2.2.2) and self-sustainable analysis (§2.2.3). Note that, differently from prior work [3, 6, 7],
our definitions are generic and compatible with different task and platform models.
2.2.1 Sustainable Scheduling Policy

We begin by generalizing the concept of a sustainable scheduling policy [3, 7], which was briefly discussed in §1. The definition captures the idea that, if a policy is sustainable with respect to a set of job parameters, having “better” values for those parameters (e.g., lower job execution costs, larger periods, less jitter, etc.) at runtime does not cause any deadline miss. We call this notion “strong sustainability,” for reasons that will be made clear in §3.

**Definition 17 (Strongly Sustainable Policy).** Assume any scheduling policy $\sigma$ and platform $\Pi$, and consider any subset of job parameters $S \subseteq \mathcal{P}_{\text{job}}$, which we call sustainable parameters. For each parameter $p \in S$, let $\preceq_p$ be any partial order over job sets, such that $J \preceq_p J'$ holds iff every job in $J$ has no worse parameter $p$ than its corresponding job in $J'$. Then we say that the scheduling policy $\sigma$ is strongly sustainable with respect to the job parameters in $S$ iff

$$\forall J \text{ s.t. } J \text{ is schedulable on platform } \Pi \text{ under policy } \sigma, \quad \forall J_{\text{better}} \text{ s.t. } J \text{ and } J_{\text{better}} \text{ differ only by } S \text{ and } \forall p \in S, J_{\text{better}} \preceq_p J,$$

$J_{\text{better}}$ is schedulable on platform $\Pi$ under policy $\sigma$.

Definition 17 states that, under a strongly sustainable scheduling policy $\sigma$, whenever we compare two job sets and show that the job set with “worse parameters” does not miss any deadline, then the job set with “better parameters” must also not miss any deadline. Note that the relation $\preceq_p$ is a crucial part of the specification and should be clearly indicated in the sustainability claim, as shown in the next examples.

**Example 6 (Sustainability with Decreasing Job Costs).** Let $\sigma$ denote any uniprocessor work-conserving, fixed-priority scheduling policy and let $\text{cost}(j)$ denote the actual execution time of job $j$. Given any job sets $J = \{j_1, j_2, \ldots\}$ and $J' = \{j'_1, j'_2, \ldots\}$, we define the relation $J \preceq_{\text{cost}} J'$ as $\forall i, \text{cost}(j_i) \leq \text{cost}(j'_i)$.

Using the relation $\preceq_{\text{cost}}$, we can instantiate Definition 17. This property expresses the notion that, under policy $\sigma$, decreasing job execution times does not render the system unschedulable. This property was proven by Ha and Liu [12] for various job models.

Similarly, one can also define sustainability with respect to job inter-arrival times. It just requires a more nuanced partial order definition, as shown in the following example.

**Example 7 (Sustainability with Increasing Job Inter-Arrival Times).** Let $\sigma$ denote any work-conserving, fixed-priority scheduling policy and let $\text{arrival}(j)$ denote the absolute arrival time of job $j$. Next, given any job sets $J = \{j_1, j_2, \ldots\}$ and $J' = \{j'_1, j'_2, \ldots\}$, we define the relation $\preceq_{\text{interarrival}}$ as

$$\forall i, \forall j_{\text{prev}}, \forall j'_{\text{prev}} \text{ s.t. }$$

$$\text{task}(j_i) = \text{task}(j'_{\text{prev}}) = \text{task}(j'_{\text{prev}}) = \text{task}(j'_{\text{prev}}) \text{ and }$$

$$\text{arrival}(j_{\text{prev}}) < \text{arrival}(j_i) \text{ and } \text{arrival}(j'_{\text{prev}}) < \text{arrival}(j'_i),$$

$$\text{arrival}(j_i) - \text{arrival}(j_{\text{prev}}) \geq \text{arrival}(j'_i) - \text{arrival}(j'_{\text{prev}}).$$

This relation expresses that, if $J \preceq_{\text{interarrival}} J'$, then the distance between two jobs of the same task in $J$ is no worse (i.e., no smaller) than in $J'$.

Finally, note that Definition 17 differs from Definition 3 (in §1.2) due to Baker and Baruah [3], as it does not require the original job set $J$ to belong to some schedulable task set $\tau$. Thus, according to our definition, Figs. 1-(a) and 1-(b) are a valid counterexample for establishing the non-sustainability (in the strong sense) of JLFP schedulers w.r.t. job costs in the presence of self-suspensions, which agrees with Definition 2 in §1.1.
2.2.2 Sustainable Schedulability Analysis

Having discussed how sustainability applies to scheduling policies, we now present the corresponding definitions for schedulability analyses, starting with the notion of sustainable schedulability analysis [3, 6, 7]. Before we proceed, we must define schedulability analysis.

◮ Definition 18 (Schedulability Analysis). Let a schedulability analysis \( A \) for task model \( M \), platform \( \Pi \), and scheduling policy \( \sigma \) denote any algorithm that assesses whether a task set \( \tau \in M \) is schedulable on \( \Pi \) under policy \( \sigma \).

Now we state whether a given schedulability analysis \( A \) is sustainable. The intuition is that, if analysis \( A \) is sustainable with respect to certain job parameters, then, if a task set \( \tau \) is deemed schedulable by \( A \), any job set with “better” parameters than those of a job set generated by \( \tau \) does not miss any deadlines.

◮ Definition 19 (Sustainable Analysis). Consider any schedulability analysis \( A \) for task model \( M \), platform \( \Pi \), and scheduling policy \( \sigma \), and consider any subset of job parameters \( S \subseteq P_{\text{job}} \), which we call sustainable parameters. For each parameter \( p \in S \), let \( \preceq_p \) be any partial order over job sets, such that \( J \preceq_p J' \) holds iff every job in \( J \) has no worse parameter \( p \) than its corresponding job in \( J' \). Then we say that analysis \( A \) is sustainable with respect to \( S \) iff

\[
\forall \tau \in M \text{ s.t. } \tau \text{ is deemed schedulable by } A,
\forall J \in \text{jobsets}(\tau), \forall J_{\text{better}} \text{ s.t. } \\
J \text{ and } J_{\text{better}} \text{ differ only by } S \text{ and } \forall p \in S, J_{\text{better}} \preceq_p J,
J_{\text{better}} \text{ is schedulable on } \Pi \text{ under policy } \sigma.
\]

Although the definitions of strongly sustainable policy (Definition 17) and sustainable analysis (Definition 19) both refer to the runtime behavior of the policy, the two notions are different. If the analyzed policy \( \sigma \) is strongly sustainable w.r.t. some parameters \( S \), then any sufficient or exact schedulability analysis for \( \sigma \) is also sustainable w.r.t. \( S \). However, even if \( \sigma \) is not strongly sustainable, it is possible to find sufficient schedulability analyses that are sustainable. In fact, we argue this is exactly the case that an intuitive notion of a “safe analysis” is trying to address: the underlying policy \( \sigma \) may exhibit various kinds of scheduling anomalies, but if a specific task set is deemed schedulable by a sustainable analysis, then no deadlines will be missed in the actual system even if some parameters turn out to be “better in the real system than assumed during analysis.”

2.2.3 Self-Sustainable Analysis

Another type of sustainability that can be found in the literature, also related to schedulability analysis, is the notion of self-sustainable analysis [3]. The intuition is that, if analysis \( A \) is self-sustainable with respect to a set of task parameters, then, if a task set \( \tau \) is deemed schedulable by analysis \( A \), every task set with “better” parameters than \( \tau \) will also be deemed schedulable by \( A \).

◮ Definition 20 (Self-Sustainable Analysis). Let \( A \) be any schedulability analysis for task model \( M \), platform \( \Pi \), and scheduling policy \( \sigma \), and consider any subset of task parameters \( S \subseteq P_{\text{task}} \). For each parameter \( p \in S \), let \( \preceq_p \) be any partial order over task sets, such that \( \tau \preceq_p \tau' \) holds iff every task in \( \tau \) has no worse parameter \( p \) than its corresponding task in \( \tau' \).
Then we say that schedulability analysis \( A \) is self-sustainable with respect to \( S \) iff
\[
\forall \tau \in \mathcal{M} \text{ s.t. } \tau \text{ is deemed schedulable by } A,
\forall \tau_{\text{better}} \text{ s.t. } \tau_{\text{better}} \text{ and } \tau \text{ differ only by } S \text{ and } \forall p \in S, \tau_{\text{better}} \preceq_p \tau,
\tau_{\text{better}} \text{ is deemed schedulable by } A.
\] (1)

To clarify the definition, we provide an example.

**Example 8 (RTA is Self-Sustainable with respect to Decreasing Task Costs).** Let \( A \) be some response-time analysis (RTA) for the sporadic task model and let WCET\((T_i)\) denote the worst-case execution time of task \( T_i \). Given any task sets \( \tau = \{T_1, T_2, \ldots\} \) and \( \tau' = \{T'_1, T'_2, \ldots\} \) with the same number of tasks, we define the relation \( \tau \preceq_{\text{WCET}} \tau' \) as
\[
\forall i, \text{WCET}(T_i) \leq \text{WCET}(T'_i).
\]

Based on the task parameter WCET and the relation \( \preceq_{\text{WCET}} \), we can instantiate the self-sustainability property as in Definition 20, which then expresses that, if the RTA claims \( \tau \) to be schedulable, then it must also claim task sets with lower WCETs to be schedulable.

Note that, despite their similarity, the notions of sustainable and self-sustainable analysis are fundamentally different. While sustainability refers to job parameters, self-sustainability concerns task parameters. Moreover, to prove that an analysis \( A \) is sustainable, one must show that the job sets generated by a task set \( \tau \) deemed schedulable by \( A \) do not have any anomalies. On the other hand, proving that analysis \( A \) is self-sustainable is a purely algorithmic property, akin to a notion of monotonicity, of the analysis procedure itself and has nothing to do with the safety at runtime of a system under analysis. For example, to prove the property in Example 8, one must show that, if the RTA computes a fixed point \( R \) for given task costs, then it will compute a fixed point \( R' \leq R \) if lower task costs are provided.

### 3 Weakly Sustainable Scheduling Policies

Recall from §1.1 that uniprocessor EDF scheduling of self-suspending tasks was proven to be not sustainable with respect to job costs [2], and as mentioned at the end of §2.2.1, this result agrees with our notion of a strongly sustainable policy (Definition 17).

However, in §1.2, we also hinted (but did not prove) that this scheduling policy is still sustainable to some extent with respect to job costs. As shown in Fig. 1-(c), by reducing suspension times (i.e., a transformation that is compliant with the task model and its constraints), we were able to construct a job set \( J_{\text{susp}} \in \text{jobsets}(\tau) \) that is as hard to schedule as job set \( J_{\text{better}} \). This suggests that any schedulability analysis \( A \) applied to task set \( \tau \) would deem it “not schedulable” anyway because of job set \( J_{\text{susp}} \).

Thus, the fact that \( J_{\text{better}} \) itself is not schedulable does not straightforwardly prove that the uniprocessor EDF scheduling policy applied to self-suspending tasks is not sustainable in some sense w.r.t. job costs, at least if self-suspension times may vary at runtime. In fact, whether or not any parameters other than the sustainable parameters should be allowed to vary at runtime is the cause of most confusion in the various interpretations of sustainability found in the state of the art [3, 6, 7], and our motivation for formalizing the notion of varying job and task parameters in Definitions 10 and 16.

While the notion of strongly sustainable policy (Definition 17) expresses that the system remains schedulable if we decrease job costs while maintaining all other parameters constant, we believe that this is too strong an assumption in many settings, since most useful schedulability analyses will consider that job parameters can vary freely and concurrently across a range of possible values. The sustainability property that we are going to define
thus allows other parameters to vary, subject to the constraints defined by the given task set. The rationale for this is that it allows for more efficient schedulability analyses if certain job parameters can be assumed to have maximal values while others are considered variable. The current sustainability theory does not allow such fine-grained categorization.

To develop a supporting theory for schedulability analyses based on this idea, in this section we propose a new classification of sustainable scheduling policies that differentiates between strong sustainability and weak sustainability.

### 3.1 Definition of Weakly Sustainable Policy

As suggested in the previous section, in order to define weak sustainability, we must be able to infer that a collection of job sets remains schedulable when certain parameters are allowed to vary. This idea is captured by the following definition.

**Definition 21 (Schedulable with Varying Job Parameters V).** Given a task set \( \tau \) and a subset of job parameters \( V \subseteq P_{job} \), we say that a job set \( J \) is schedulable with varying parameters \( V \) subject to task set \( \tau \) on platform \( \Pi \) under policy \( \sigma \) iff any job set \( J_{other} \in \text{jobsets}(\tau) \) that differs from \( J \) only by \( V \) is also schedulable on \( \Pi \) under policy \( \sigma \).

To illustrate the definition, we provide an example.

**Example 9 (Schedulable with Varying Costs).** Assume any scheduling policy \( \sigma \) and consider the set of variable parameters \( V = \{ \text{cost} \} \). Given a job set \( J = \{ j_1, j_2 \} \) generated by task set \( \tau \), we say that \( J \) is schedulable with varying costs subject to task set \( \tau \) iff every job set \( J_{other} \) generated by \( \tau \) that has two jobs and the same parameters as \( J \) except for their costs is schedulable. That is, any job set constructed by changing only the job costs of \( J \) (to higher or lower values), without violating the constraints set forth by the parameters of task set \( \tau \), must be schedulable.

In other words, one may understand this notion to mean that job set \( J \) is not only schedulable itself, but also a “schedulability witness” for a whole family of related job sets that are identical in all parameters except for those in \( V \). Based on this concept, we can now define precisely under which conditions a policy is weakly sustainable.

**Definition 22 (Weakly Sustainable Policy).** Assume any platform \( \Pi \), task model \( M \), and scheduling policy \( \sigma \), and consider any disjoint subsets of job parameters \( S \subseteq P_{job} \) and \( V \subseteq P_{job} \), which we call sustainable and variable parameters, respectively. For each sustainable parameter \( p \in S \), let \( \preceq_p \) be any partial order over job sets, such that \( J \preceq_p J' \) holds iff every job in \( J \) has no worse parameter \( p \) than its corresponding job in \( J' \). Then we say that scheduling policy \( \sigma \) is weakly sustainable with sustainable parameters \( S \) and variable parameters \( V \) iff

\[
\forall \tau \in M, \forall J \in \text{jobsets}(\tau) \text{ s.t. } J \text{ is schedulable with varying } V \text{ subject to } \tau \text{ on platform } \Pi \text{ under policy } \sigma, \\
\forall J_{better} \text{ s.t. } J \text{ and } J_{better} \text{ differ only by } S \text{ and } \forall p \in S, J_{better} \preceq_p J, \\
J_{better} \text{ is schedulable on platform } \Pi \text{ under policy } \sigma.
\]

The idea of weak sustainability is that, if we can determine that a job set is schedulable for all variations of parameters in \( V \) (subject to the constraints imposed by its associated task set), then all job sets with better parameters \( S \) must be schedulable. For clarity, we provide the following example.
Example 10 (Weak Sustainability w.r.t. Job Costs and Varying Suspension Times). Consider a uniprocessor JLF scheduling policy $\sigma$ and the dynamic suspension model, i.e., jobs can suspend at any time but the total suspension duration of each job is bounded by its task’s maximum suspension time. Let $\text{susp}(j)$ denote the total suspension time of job $j$ and $\text{cost}(j)$ the execution time of job $j$.

By defining the sets of job parameters $S = \{\text{cost}\}$ and $V = \{\text{susp}\}$, and the relation $\preceq_{\text{cost}}$ as in Example 6, one can instantiate Definition 22 and prove (as shown in §4) that, for any task set $\tau \in \mathcal{M}$, if job set $J$ generated by $\tau$ is schedulable for all possible suspension times (subject to the upper limit imposed by $\tau$), then all job sets with lower or equal job costs are also schedulable.

In the specific case where the set of varying parameters $V$ is empty, we call the scheduling policy strongly sustainable.

Definition 23 (Strongly Sustainable Policy). We say that a policy is strongly sustainable with respect to the job parameters in $S$ iff it is weakly sustainable with respect to the sustainable parameters in $S$ and an empty set of variable parameters $V = \emptyset$.

Note that if $V = \emptyset$, proving that job set $J$ is schedulable with varying parameters $V$ is the same as establishing that $J$ itself is schedulable. That implies the following equivalence, which connects the definitions of sustainable policy in §2 and §3.

Corollary 1 (Equivalence of Strong Sustainability). The notion of strongly sustainable policy as defined in Definition 23 is equivalent to Definition 17.

The weak sustainability property is useful for constraining the search space when developing schedulability analyses. As is already known, if some policy $\sigma$ is strongly sustainable with respect to the parameters in $S$, maximizing/minimizing such parameters enables constructing worst-case scenarios (e.g., the critical instant for uniprocessor FP scheduling of sporadic tasks [14]), so that only a single worst-case scenario must be analyzed (rather than the entire space of all possible parameter combinations).

However, recall that policy $\sigma$ might not be strongly sustainable with respect to $S$. But if we are still able to prove that $\sigma$ is weakly sustainable with respect to $S$ and variable parameters $V$, we can still maximize/minimize the parameters in $S$, as long as the schedulability analysis covers all values of the parameters in $V$. In other words, establishing a weak sustainability property can be thought of as a dimensionality reduction of the search space that must be considered by a safe schedulability analysis.

For instance, having proven in Theorem 4 in §4.4 that uniprocessor JLF scheduling of self-suspending tasks is weakly sustainable with respect to job costs and variable suspension times, we know that any schedulability analysis for that model may assume that all jobs generated by the tasks execute for their maximum execution time, and must search only for the worst-case assignments of job suspension times.

3.2 Composing Weak and Strong Sustainability Results

Although the definition of strong sustainability refers to a set $S$ of multiple parameters, one can still establish the sustainability of each parameter in isolation. In fact, the critical instant for the sporadic task model is obtained by composing worst-case assumptions about individual job parameters: maximizing job costs, minimizing inter-arrival time, etc.

As will be shown in Theorem 1, this composition rule applies not only for strong sustainability (as discussed in prior work [3]), but can also be extended to weak sustainability. Before presenting the theorem, we first provide an alternative (but equivalent) definition of
weak sustainability based on the contrapositive of Definition 22, which simplifies the proof of Theorem 1 below.

**Definition 24 (Weakly Sustainable Policy – alternative definition).** Assume any platform II, task model \( M \) and scheduling policy \( \sigma \), and consider any disjoint subsets of job parameters \( S \subseteq P_{\text{job}} \) and \( V \subseteq P_{\text{job}} \), which we call sustainable and variable parameters, respectively. For each sustainable parameter \( p \in S \), let \( \preceq_p \) be any partial order over job sets, such that \( J \preceq_p J' \) holds iff every job in \( J \) has no worse parameter \( p \) than its corresponding job in \( J' \).

Then we say that the scheduling policy is weakly sustainable with sustainable parameters \( S \) and variable parameters \( V \) iff

\[
\forall J \text{ s.t. } J \text{ is not schedulable on platform II under policy } \sigma,
\forall \tau \in M, \forall J_{\text{worse}} \in \text{jobsets}(\tau) \text{ s.t.}
\ J \text{ and } J_{\text{worse}} \text{ differ only by } S \text{ and } \forall p \in S, J \preceq_p J_{\text{worse}},
\exists J'_{\text{worse}} \in \text{jobsets}(\tau) \text{ s.t.}
\ J_{\text{worse}} \text{ and } J'_{\text{worse}} \text{ differ only by } V \text{ and}
\ J'_{\text{worse}} \text{ is not schedulable on platform II under policy } \sigma.
\]

Put differently, for any job set \( J \) that is not schedulable, if we can find another job set \( J_{\text{worse}} \) that is generated by some task set \( \tau \) and \( J \) is “better” than \( J_{\text{worse}} \), then there exists a member in \( J_{\text{worse}} \)’s “family” of related job sets that is also not schedulable.

For instance, recall that this is the same reasoning underlying the counterexample in §1.2: given a job set \( J \) that is not schedulable (Fig. 1-(b)) and a job set \( J_{\text{worse}} \) with higher job costs (Fig. 1-(a)), we were able to show that there exists a job set \( J'_{\text{worse}} \) that only differs from \( J_{\text{worse}} \) in its suspension times and that also misses a deadline (Fig. 1-(c)).

In addition, we must introduce the notion of independent sets of job parameters.

**Definition 25 (Independent Sets of Job Parameters).** We say that subsets of job parameters \( A \subseteq P_{\text{job}} \) and \( B \subseteq P_{\text{job}} \) are independent with respect to task model \( M \) iff for each task parameter \( p_{\text{task}} \) defined by \( M \), and for every \( p_A \in A \) and \( p_B \in B \), if \( p_A \) is constrained by \( p_{\text{task}} \) according to model \( M \), then \( p_B \) is not constrained by \( p_{\text{task}} \) according to model \( M \).

In most task models commonly considered in the real-time literature, job parameters are usually independent of each other.

**Example 11 (Parameters Are Usually Independent).** In the sporadic task model with self-suspending tasks, the sets of job parameters \( A = \{ \text{cost, arrival} \} \) and \( B = \{ \text{susp} \} \) have independent task constraints, since these job parameters are each constrained by a different task parameter, namely, the task WCET, minimum inter-arrival time and maximum suspension time. In contrast, in a hypothetical task model where every job \( j \) is split into two execution sections of length \( \text{cost}_1(j) \) and \( \text{cost}_2(j) \) such that \( \text{cost}_1(j) + \text{cost}_2(j) \leq \text{WCET}(\text{task}(j)) \), the parameters \( \{ \text{cost}_1 \} \) and \( \{ \text{cost}_2 \} \) are clearly not independent.

Using the definition of weak sustainability above (Definition 24) and the notion of independent sets of job parameters (Definition 25), we establish the composition rule for weakly sustainable policies.

**Theorem 1 (Composition Rule: Weak – Weak).** Consider any task model \( M \), scheduling policy \( \sigma \) and processor platform II. Let \( S_a \), \( V_a \), \( S_b \), \( V_b \) denote subsets of the job parameters \( P_{\text{job}} \) such that \( S_a \cap V_b = \emptyset \) and \( S_b \cap V_a = \emptyset \), and such that either \( S_b \) is independent of \( P_{\text{job}} \setminus S_a \), or \( S_a \) is independent of \( P_{\text{job}} \setminus S_a \), with respect to task model \( M \). Assume that
(a) \( \sigma \) is weakly sustainable with respect to \( S_a \) and variable parameters \( V_a \), and that (b) \( \sigma \) is weakly sustainable with respect to \( S_b \) and variable parameters \( V_b \). Then (c) \( \sigma \) is weakly sustainable with respect to \( S_a \cup S_b \) and variable parameters \( V_a \cup V_b \).

**Proof.** Consider a job set \( J \) that is not schedulable on platform \( \Pi \) under policy \( \sigma \). Let \( \tau \) be any task set, and let \( J_{\text{worse}} \in \text{jobs}(\tau) \) be a job set that only differs from \( J \) by the parameters in \( S_a \cup S_b \) and that has no better parameters than \( J \) w.r.t. \( S_a \cup S_b \). Then, according to Definition 24, we must prove that there exists a job set \( J'_{\text{worse}} \in \text{jobs}(\tau) \) that only differs from \( J_{\text{worse}} \) with respect to \( V_a \cup V_b \) and that is also not schedulable.

Using the independent parameters assumption, assume without loss of generality that it is \( S_a \) that is independent of all other job parameters \( \mathcal{P}_{\text{job}} \setminus S_a \), with respect to model \( \mathcal{M} \). If this is not the case, then by assumption we have that \( S_a \) is independent of other parameters \( \mathcal{P}_{\text{job}} \setminus S_a \) and we can exchange the indices \( a \) and \( b \) in the remainder of the proof.

1. **Step 1 – Construction of \( J'_a \) from \( J \):** Let \( J_a \) be the same job set as \( J \), but with the same job parameters in \( S_a \) as \( J_{\text{worse}} \). That is, let \( J = \{j_1, j_2, \ldots\} \) and \( J_{\text{worse}} = \{j_1^w, j_2^w, \ldots\} \) and recall that they have the same number of jobs. Then we define \( J_a = \{j_1^a, j_2^a, \ldots\} \) with the same cardinality such that, for any index \( i \), we have \( \forall p \in S_a, p(j_i^a) = p(j_i^w) \) and \( \forall p \notin S_a, p(j_i^a) = p(j_i) \).

Next, we construct a task set \( \tau_a \in \mathcal{M} \) such that, for every task parameter \( p_{\text{task}} \) that constrains job parameters in \( \mathcal{P}_{\text{job}} \setminus S_b \), the value of \( p_{\text{task}} \) in \( \tau_a \) is the same as in \( \tau \), and for every task parameter \( p_{\text{task}} \) that constrains job parameters in \( S_b \), the value of \( p_{\text{task}} \) in \( \tau_a \) is the same as the task set that generated job set \( J \). Since \( J_a \) only differs from \( J_{\text{worse}} \in \text{jobs}(\tau) \) with respect to \( S_b \), and \( S_b \) is independent of the other job parameters, it follows that \( J_a \in \text{jobs}(\tau_a) \).

Since \( J \) is not schedulable, and \( J \) and \( J_a \) differ only by \( S_a \), we can exploit the fact that \( \sigma \) is weakly sustainable with \( S_a \) and varying \( V_a \). Thus, it follows that there exists a job set \( J'_a \in \text{jobs}(\tau_a) \) that differs from \( J_a \) only by the parameters in \( V_a \) and that is not schedulable on platform \( \Pi \) under policy \( \sigma \).

2. **Step 2 – Construction of \( J'_{ab} \) from \( J'_a \):** Let \( J_{ab} \) be the same job set as \( J'_a \) except that the job parameters in \( S_b \) are the same as in \( J_{\text{worse}} \). That is, let \( J'_a = \{j_1^a, j_2^a, \ldots\} \) and \( J_{\text{worse}} = \{j_1^w, j_2^w, \ldots\} \) and recall that they have the same number of jobs. Then we define \( J_{ab} = \{j_1^{ab}, j_2^{ab}, \ldots\} \) with same cardinality such that, for any index \( i \), we have \( \forall p \in S_b, p(j_i^{ab}) = p(j_i^w) \) and \( \forall p \notin S_b, p(j_i^{ab}) = p(j_i^a) \).

Note that by construction, \( J_{ab} \) has the same job parameters as \( J_{\text{worse}} \in \text{jobs}(\tau) \), except for those in \( V_a \), which were obtained when generating \( J'_a \) via weak sustainability. However, note that \( J'_a \) is generated by task set \( \tau_a \), which has the same constraints for \( V_a \) as \( \tau \), since \( V_a \cap S_b = \emptyset \). Thus, every job parameter of \( J_{ab} \) is compatible with \( \tau \), i.e., \( J_{ab} \in \text{jobs}(\tau) \).

Since \( J'_a \) is not schedulable, and \( J'_a \) and \( J_{ab} \) differ only by \( S_b \), we can exploit the fact that \( \sigma \) is weakly sustainable with \( S_b \) and varying \( V_b \). Thus, there must exist a job set \( J'_{ab} \in \text{jobs}(\tau) \) that differs from \( J_{ab} \) only by the parameters in \( V_b \) and that is not schedulable on platform \( \Pi \) under policy \( \sigma \).

Since \( J \) has the same parameters as \( J_{\text{worse}} \) except for those in \( S_a \cup S_b \), and because \( J_a \) and \( J_{ab} \) were constructed from \( J \) by copying the parameters \( S_a \) and \( S_b \) from \( J_{\text{worse}} \) and varying the parameters in \( V_a \cup V_b \), it follows that \( J'_{ab} \) has the same parameters as \( J_{\text{worse}} \), except for the variable parameters \( V_a \) and \( V_b \). Moreover, since \( S_a \cap V_b = \emptyset \) and \( S_b \cap V_a = \emptyset \), this guarantees that \( S_a \) and \( S_b \) do not vary during the construction of \( J'_a \) and \( J'_{ab} \), so for every \( p \in S_a \cup S_b \), the order \( \preceq_p \) is preserved across the successive job set transformations.
Thus, there exists a job set $J'_{worse} = J'_{ab}$ that belongs to $\text{jobsets}(\tau)$, that only differs from $J_{worse}$ with respect to $V_a \cup V_b$ and is also not schedulable on platform $\Pi$ under policy $\sigma$. ◭

Assuming $V_b = \emptyset$ yields a rule for combining strong and weak sustainability results.

**Corollary 2 (Composition Rule: Weak – Strong).** Consider any scheduling policy $\sigma$ and processor platform $\Pi$. Let $S_a$, $V_a$ and $S_b$ denote subsets of the job parameters $P_{job}$ such that $S_b \cap V_a = \emptyset$, and such that either $S_b$ is independent of $P_{job} \setminus S_b$, or $S_a$ is independent of $P_{job} \setminus S_a$, with respect to task model $M$. Assume that $\sigma$ is weakly sustainable with respect to $S_a$ and variable $V_a$ and also strongly sustainable with respect to $S_b$. Then $\sigma$ is weakly sustainable with respect to $S_a \cup S_b$ and variable $V_a$.

Finally, assuming $V_a = V_b = \emptyset$ yields the composition rule for strong sustainability, which was already proven by Baker and Baruah [3].

**Corollary 3 (Composition Rule: Strong – Strong).** Consider any scheduling policy $\sigma$ and processor platform $\Pi$. Let $S_a$ and $S_b$ denote subsets of the job parameters $P_{job}$, and such that either $S_b$ is independent of $P_{job} \setminus S_b$, or $S_a$ is independent of $P_{job} \setminus S_a$, with respect to task model $M$. Assume that $\sigma$ is strongly sustainable with respect to $S_a$ and also strongly sustainable with respect to $S_b$. Then $\sigma$ is strongly sustainable with respect to $S_a \cup S_b$.

Note that, although necessary in the general case, the parameter independence constraint in Corollary 3 was not explicitly stated in the original definition [3], since Baker and Baruah only considered the sporadic task model, in which all parameters are independent.

## 4 Uniprocessor Scheduling of Dynamic Self-Suspending Tasks is Weakly Sustainable w.r.t. Job Costs and Variable Suspensions

In this section, we prove that uniprocessor JLFP scheduling with dynamic self-suspending tasks is weakly sustainable with respect to job costs and variable suspension times. Although we could have focused on other real-time task models, we chose to study the sustainability of self-suspending tasks for the following reasons.

1. **Recent errors:** This topic has faced many misconceptions in the past, with a considerable number of unsound results being published [9]. We hope that our work on sustainability introduces helpful formalism and a better understanding of the task model.

2. **Future work on schedulability analysis:** Proving weak sustainability of uniprocessor JLFP scheduling of dynamic self-suspending tasks can provide directions for future work. It enables more efficient schedulability analyses to be developed, by reducing the search space to only the parameters that must be kept variable (i.e., suspension times), while the others (i.e., execution times) can remain constant.

To address the issue of recent errors and increase the degree of confidence in the results, our proof has been mechanized in PROSA [8], a library for the COQ proof assistant for formal specification and machine-checked proofs of real-time scheduling theory. The specification and proofs are available online [1] and can be checked independently with the COQCHK tool. Simple step-by-step instructions are provided on the website.

Note that, despite being phrased in terms of sporadic tasks for the sake of simplicity, this proof is conceptually also compatible with other job arrival models (periodic, bursty, etc.).

The rest of this section is structured as follows. First, we present our formalization of the dynamic suspension model, which is required for stating the theorems in PROSA. Next, we provide an overview of our proof strategy based on schedule reductions, which can be reused in other sustainability proofs. In the remaining subsections, we discuss the high-level steps...
of the proof, which despite being specific for scheduling with self-suspending, highlight key steps necessary in a rigorous proof of sustainability.

4.1 A Generic Suspension Model

In order to instantiate the sustainability claim for real-time scheduling of self-suspending tasks, we must formally define the concept of self-suspension.

Definition 26 (Job Suspension Time). We define job suspension time as a function $\text{sus}(j, s)$ such that, for any job $j$ and any value $s \in \mathbb{N}$, $\text{sus}(j, s)$ expresses the duration for which $j$ must suspend immediately after receiving $s$ units of service.

The job suspension parameter is explained more clearly in the following example.

Example 12 (Table of Suspension Durations). Job suspension times $\text{sus}(j, s)$ can be understood as a table containing the duration of the suspension intervals associated with job $j$. For example, for a job $j$ such that $\text{cost}(j) = 5$, we can define $\text{sus}(j, s)$ to equal 0 except for $\text{sus}(j, 3) = 2$ and $\text{sus}(j, 4) = 3$.

This suspension table indicates that job $j$ executes for 3 time units, then suspends for 2 time units, then executes for 1 more time unit, then suspends for 3 more time units and finally completes its last time unit of execution. Note that this assignment is both an instance of the dynamic suspension model (with total suspension time equal to 5) and of the segmented suspension model (with execution segments $[e_1 = 3, s_1 = 2, e_2 = 1, s_2 = 3, e_3 = 1]$).

By allowing arbitrary suspension durations between each unit of service, this model is generic enough to represent any suspension pattern under discrete time. Thus, it supports both segmented [16] and dynamic [13] suspension models, as shown in Example 12.

Next, by accumulating suspension durations, we define the total suspension time of a job.

Definition 27 (Total Suspension Time). We define the total suspension time $\text{sus}(j)$ of job $j$ as the cumulative suspension time up to completion, i.e., $\text{sus}(j) = \sum_{s < \text{cost}(j)} \text{sus}(j, s)$.

After clarifying job suspension times, we now define task suspension times and show how both are related under the dynamic suspension model.

Definition 28 (Task Suspension Time). For any task $T_i$, we define the task suspension time $\text{sus}(T_i)$ as an upper-bound on the total suspension time of any job of $T_i$.

Definition 29 (Suspension Time Constraints). The dynamic suspension model requires that the total suspension time of any job is upper-bounded by the suspension time of its task, i.e.,

$$\forall \tau \in \mathcal{M}, \forall \mathcal{J} \in \text{jobsets}(\tau), \forall j \in \mathcal{J}, \text{sus}(j) \leq \text{sus}(\text{task}(j)).$$

Beside its suspension time, every task $T_i$ is defined by a WCET, a minimum inter-arrival time or period, and a deadline, as stated in Definition 1.

4.2 Overview of the Proof Strategy

Having presented the main characteristics of the dynamic self-suspending task model, we now explain our proof strategy for establishing weak sustainability of uniprocessor JLFP scheduling of dynamic self-suspending tasks w.r.t. job costs and variable suspension times. For simplicity, the proof is based on the alternative definition of weakly sustainable policy.
Figure 2 Proof strategy for establishing weak sustainability with respect to job costs and variable suspension times. Given a job $j_i$ that misses a deadline in schedule $S$ of the original job set $\mathcal{J}$, we construct a new job set $\mathcal{J}_{\text{worse}}$ and a new schedule $S'$ where the corresponding job $j'_i$ misses a deadline. Note that in schedule $S'$, job costs are no smaller than in $S$, suspension times can be defined arbitrarily (within the bounds of the task set), and all other job parameters (i.e., arrival time, deadline) remain unchanged.

(Definition 24). According to Definition 24, we must prove that

$$\forall J \text{ s.t. } J \text{ is not schedulable under a uniprocessor JLFP scheduling policy } \sigma,$$

$$\forall \tau \in \mathcal{M}, \forall J_{\text{worse}} \in \text{jobsets}(\tau) \text{ s.t. }$$

$$J \text{ and } J_{\text{worse}} \text{ differ only by } S = \{\text{cost}\} \text{ and } J \preceq_{\text{cost}} J_{\text{worse}},$$

$$\exists J'_{\text{worse}} \in \text{jobsets}(\tau) \text{ s.t. }$$

$$J_{\text{worse}} \text{ and } J'_{\text{worse}} \text{ differ only by } V = \{\text{susp}\} \text{ and }$$

$$J'_{\text{worse}} \text{ is not schedulable under policy } \sigma.$$

That is, first we consider any job set $J$ that is not schedulable and any job set $J_{\text{worse}}$ that has “no better job costs” than $J$ (and that is otherwise identical). Then we must show that there exists a job set $J'_{\text{worse}}$ generated by the same task set as $J_{\text{worse}}$ only by its job suspension times and that it is not schedulable. In particular $J$ and $J_{\text{worse}}$ have equal suspension times (but not necessarily equal execution costs), whereas $J_{\text{worse}}$ and $J'_{\text{worse}}$ have equal execution costs (but not necessarily equal suspension times).

Our proof begins by considering any job set $J$ and its associated schedule $S$ where some job misses a deadline. Then we construct a job set $J'_{\text{worse}}$ together with its schedule $S'$ where some job also misses a deadline. This strategy is illustrated in Fig. 2.

In the next section, we present an algorithm for iteratively constructing schedule $S'$ (and hence the associated job set $J'_{\text{worse}}$) based on $S$. It is followed by the two main proof obligations: (a) proving that some job misses a deadline in $S'$ (§4.3.1) and (b) proving that $S'$ is a valid schedule of $J'_{\text{worse}}$ (§4.3.2). This proves that $J'_{\text{worse}}$ is not schedulable.

4.3 Constructing $J'_{\text{worse}}$ and Schedule $S'$

Based on the strategy proposed in §4.2, we now present the algorithm to construct schedule $S'$ and the associated job set $J'_{\text{worse}}$, based on the original schedule $S$. In the remainder of this paper, whenever we want to refer to the same job before and after the parameter transformation (i.e., from $J$ to $J'_{\text{worse}}$), we refer to them as corresponding jobs $j_i$ and $j'_i$.

Before proceeding, we must first introduce the concept of job service.

Definition 30 (Job Service). Given a schedule $S$, we define service($j, t$) as the cumulative amount of time during which job $j$ executes in the interval $[0, t]$.

Next, recall that jobs in $J'_{\text{worse}}$ have no better job costs than in $J$, i.e., for any corresponding jobs $j_i$ and $j'_i$, cost($j_i$) ≤ cost($j'_i$). Since they might have to execute for different durations, we begin by defining the notion of added cost.
Definition 31 (Added Cost). We define the added cost $\Delta_{\text{cost}}(j'_i)$ of job $j'_i$ in $J'_{\text{worse}}$ as the difference between its original and inflated costs, i.e., $\Delta_{\text{cost}}(j'_i) = \text{cost}(j'_i) - \text{cost}(j_i) \geq 0$.  

In order to guarantee that schedule $S'$ becomes as hard as schedule $S$ (i.e., so that jobs still miss their deadlines) and at the same time easy to compare in terms of received job service (i.e., the time for which a job executed since its release), we construct $S'$ based on the idea of “picking jobs that are late with respect to $S$,” where late is defined as follows.

Definition 32 (Late job). We say that job $j'_i$ is late in schedule $S'$ at time $t$ iff the service received by $j'_i$ in $S'$ up to time $t$ is less than the service received by the corresponding job $j_i$ in schedule $S$ (compensated by the added cost), i.e., $\text{service}(j'_i, t) < \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i)$.  

We now present the algorithm used to iteratively build schedule $S'$ and job set $J'_{\text{worse}}$. Algorithm 1 ensures that (i) every job $j'_i \in J'_{\text{worse}}$ executes for its total execution cost $\text{cost}(j'_i)$ ($\geq \text{cost}(j_i)$), (ii) every job $j'_i \in J'_{\text{worse}}$ has a total suspension time $\text{susp}(j'_i)$ upper-bounded by the suspension time $\text{susp}(j_i)$ of its corresponding job in schedule $S$, and (iii) at least one job of $J'_{\text{worse}}$ misses its deadline in $S'$ (as proven in Sec 4.3.1).

Algorithm 1 (Construction of Job Set $J'_{\text{worse}}$ and Schedule $S'$). Consider any time $t$ and let $J(t)$ denote the set that contains every job $j'_i$ that is ready (i.e., released, not completed, and not suspended) in schedule $S'$ at time $t$ and such that either (a) $j'_i$ is late at time $t$ or (b) the corresponding job $j_i$ is scheduled in $S$ at time $t$.

1. **Schedule:** We schedule in $S'$ at time $t$ the highest-priority job in $J(t)$, or idle the processor if $J(t)$ is empty.
2. **Suspensions:** Any job $j'_i \in J'_{\text{worse}}$ suspends in $S'$ at time $t$ iff the corresponding job $j_i$ is suspended in $S$ and $j'_i$ is not late.

Note that Algorithm 1 not only picks late jobs, but also favors higher-priority jobs and tries to copy schedule $S$ if possible. While rule (a) ensures that the schedule respects the JLFP policy, rule (b) provides a tie-breaking rule if there are multiple jobs that can be picked, in which case we choose the same job as the job scheduled in $S$.

It only remains to be shown that schedule $S'$ results in a deadline miss (Theorem 2) and schedule $S'$ does not violate any property of the scheduling policy, platform, and task model, such as work conservation, priority enforcement, etc. (Theorem 3).

4.3.1 Proving that $S'$ Misses a Deadline

In order to prove that some job $j'_i \in J'_{\text{worse}}$ misses a deadline in $S'$, we establish the following key invariant that relates the service in the two schedules $S$ and $S'$.

Lemma 1 (Service Invariant). For any corresponding jobs $j_i \in J$ and $j'_i \in J'_{\text{worse}}$, at any time $t$, $\text{service}(j'_i, t) \leq \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i)$.  

**Proof.** Proven in PROSA [1]. Consider any pair of corresponding jobs $j_i$ and $j'_i$. The proof follows by induction on time $t$.

1. **Base Case:** At time $t = 0$, jobs have received no service, thus $\text{service}(j'_i, 0) = 0 = \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i)$.
2. **Inductive Step:** Assume as the induction hypothesis that, for some $t$, $\text{service}(j'_i, t) \leq \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i)$. Then we must prove $\text{service}(j'_i, t+1) \leq \text{service}(j_i, t+1) + \Delta_{\text{cost}}(j'_i)$.  


First, consider the simple case where job \( j'_i \) is not scheduled in \( S' \) at time \( t \). Then,

\[
\text{service}(j'_i, t + 1) = \text{service}(j'_i, t) \quad (j'_i \text{ is not scheduled in } S' \text{ at } t)
\]

\[
\leq \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i) \quad (\text{by induction hypothesis})
\]

\[
\leq \text{service}(j_i, t + 1) + \Delta_{\text{cost}}(j'_i). \quad (\text{by monotonicity of service})
\]

Otherwise, assume that \( j'_i \) is scheduled in \( S' \) at time \( t \). From the schedule construction (Algorithm 1), it follows that either (a) \( S \) and \( S' \) schedule corresponding jobs at time \( t \), or (b) \( S' \) schedules a late job at time \( t \). We analyze both cases.

a. **Corresponding Jobs are Scheduled:** The corresponding jobs scheduled in \( S \) and \( S' \) at time \( t \) must be \( j_i \) and \( j'_i \), so

\[
\text{service}(j'_i, t + 1) = \text{service}(j'_i, t) + 1 \quad (j'_i \text{ is scheduled in } S' \text{ at time } t)
\]

\[
\leq \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i) + 1 \quad (\text{by induction hypothesis})
\]

\[
= \text{service}(j_i, t + 1) + \Delta_{\text{cost}}(j'_i) \quad (j_i \text{ is scheduled in } S \text{ at time } t).
\]

b. **Late Job:** Job \( j'_i \) must be the highest-priority late job in \( S' \) at time \( t \). By the definition of late job (Definition 32), it follows that \( \text{service}(j'_i, t) < \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i) \), so

\[
\text{service}(j'_i, t + 1) = \text{service}(j'_i, t) + 1 \quad (j'_i \text{ is scheduled in } S' \text{ at time } t)
\]

\[
< \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i) + 1 \quad (\text{by assumption})
\]

\[
\leq \text{service}(j_i, t) + \Delta_{\text{cost}}(j'_i). \quad (\text{by converting } < \text{ to } \leq)
\]

The claim holds in all cases, which concludes the proof by induction.

Since we must prove that schedule \( S' \) results in a deadline miss, we use the service invariant above to conclude that jobs complete earlier in \( S \) than in \( S' \).

**Corollary 4 (Jobs Complete Earlier in \( S \)).** For any corresponding jobs \( j_i \in J \) and \( j'_i \in J'_{\text{worse}} \), if \( j'_i \) has completed in schedule \( S' \) by time \( t \), then \( j_i \) has completed in \( S \) by time \( t \).

**Proof.** Proven in PROSA [1]. Follows from Lemma 1, since \( j_i \) receives enough service in \( S \) to complete before the corresponding \( j'_i \) in \( S' \).

Recall that we initially assumed that some job misses a deadline in \( S \). We can thus conclude that the corresponding job also misses a deadline in \( S' \).

**Theorem 2 (Deadline Miss).** There exists a job \( j'_i \in J'_{\text{worse}} \) that misses a deadline in \( S' \).

**Proof.** Proven in PROSA [1]. Recall from Corollary 4 that corresponding jobs complete earlier in \( S \) than in \( S' \). Since by assumption there exists a job \( j_i \) that misses a deadline in \( S \), the corresponding job \( j'_i \) must also miss a deadline in \( S' \).

**4.3.2 Proving that \( S' \) is a Valid Schedule**

Although we have already established the non-schedulability of the generated schedule \( S' \), it remains to be shown that schedule \( S' \) is valid and compatible with the task model.

**Theorem 3 (Valid Schedule).** Schedule \( S' \) is a valid uniprocessor schedule of job set \( J'_{\text{worse}} \) assuming JLFP scheduling of sporadic, dynamic self-suspending tasks.

**Proof.** Proven in PROSA [1]. Follows from Algorithm 1, since suspension intervals in schedule \( S' \) are no longer than those in \( S \) and the fact that the dynamic self-suspension model imposes only an upper bound on total job suspension time, and since by construction the derived schedule \( S' \) is work-conserving, respects self-suspensions, and respects job priorities.
4.4 Main Claim

Based on the strategy explained in §4.2, by combining Theorems 2 and 3, we prove that the scheduling policy is weakly sustainable.

**Theorem 4 (Weak Sustainability).** Uniprocessor JLFP scheduling of sporadic self-suspending tasks under the dynamic suspension model is weakly sustainable with respect to job costs and variable suspension times.

**Proof.** Proven in Prosa [1]. Instantiate Definition 24 with uniprocessor JLFP scheduling of sporadic self-suspending tasks under the dynamic suspension model for $S = \{\text{cost}\}$ and $V = \{\text{susp}\}$. Theorems 2 and 3 imply that, for any schedule $S$ of job set $J$ that contains a deadline miss, there exists a schedule $S'$ of the transformed set $J_{\text{worse}}$ that also contains a deadline miss.

We emphasize that Algorithm 1 builds a schedule $S'$ and hence a job set $J'_{\text{worse}}$ that has different suspension times than the original job set $J$. Therefore, the presented argument indeed proves the weak (but not strong) sustainability of uniprocessor JLFP scheduling under the dynamic self-suspending task model w.r.t. job costs and variable suspension times.

5 Conclusion and Future Work

Sustainability is a central aspect of real-time theory with many applications in the development of real-time systems. By allowing system designers to target only extreme scheduling scenarios, it simplifies the design, prototyping, and analysis of real-time systems. In addition, the use of sustainable scheduling policies and analyses greatly aids the validation and certification process, by ensuring that only a subset of execution scenarios must be checked, and that any variation within the system’s specified bounds does not compromise safety.

In this paper, we have identified that the existing notions of sustainability in real-time scheduling allow for multiple interpretations with regard to whether real-time scheduling of self-suspending tasks is sustainable with respect to job costs. To resolve the issue, we developed a precise sustainability theory for real-time scheduling that is compatible with any task and platform model (§2), and also proposed the new notions of strongly and weakly sustainable policies (§3), which can be used to derive more efficient schedulability analyses for policies that were shown to not be strongly sustainable.

To better understand a model for which many mistakes were found in the literature [9], we chose to study weak sustainability in the context of self-suspending tasks. For that, we developed a generic model for self-suspending tasks (§4.1) that was formalized in the Coq proof assistant and integrated into Prosa [1, 8]. Finally, we mechanically proved in Prosa that uniprocessor JLFP scheduling of self-suspending tasks is weakly sustainable with respect to job costs and variable suspension times (§4.2–§4.4).

In ongoing work, we are working towards leveraging the obtained weak sustainability result to derive new, machine-checked schedulability tests for the dynamic suspension model. In future work, it will be interesting to identify instances of weakly sustainable parameters in other task models, platforms, and scheduling algorithms to improve the complexity of their associated timing analyses and to lessen test coverage requirements.
References


