Additional Spatial Dimensions Can Help Speed Up Computations

Luc Longpré, Olga Kosheleva, and Vladik Kreinovich

University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA
longpre@utep.edu, olgak@utep.edu, vladik@utep.edu
1. Many Computational Problems Require Too Much Computation Time

- It is known that many practical computational problems are NP-hard.
- This means, crudely speaking, that:
  - unless $P = NP$ (which most computer scientists do not believe to be possible),
  - any algorithm that always solves the corresponding problem will require,
  - at least for some inputs of reasonably large size,
  - an unrealistically long time to solve,
  - e.g., time larger than the lifetime of the Universe.
2. Parallelization Can Help – At Least to Some Extent

- If for a person, some task takes too much time, this person can (and does) ask for help.
- When two or more people work on some task, they can perform it faster.
- Similarly, when a certain computational task requires too much time on a single computer:
  - a natural way to speed up computations is to divide the original task between several computers,
  - i.e., to parallelize computations.
- Many modern high-performance computers consists of thousands of processors working on the same task.
- For many computational tasks, this indeed leads to a drastic speed-up.
3. Fundamental Limitations of Parallelization Speed-Up

• In general, parallelization is not a panacea: this idea has limitations.

• Some of these limitations are technical.

• These limitations will hopefully be overcome in the future.

• However, there are also fundamental limitations on how much speed-up can be achieved by parallelization.

• Indeed, let us assume that we have a parallel computer that finishes its computations in time $T_{\text{par}}$.

• Let us show how we can simulate its computations sequentially.
4. Limitations (cont-d)

- According to modern physics, the speed of all processes is bounded by the speed of light \(c\).
- During the time \(T_{\text{par}}\), the information from the processors must reach the user.
- This means that the processors that participate in this computation must be located within the distance

\[ R \overset{\text{def}}{=} c \cdot T_{\text{par}}. \]

- In geometric terms, they must be inside the sphere of radius \(R\) centered at the user location.
- The overall volume of this area is equal to

\[ V = \frac{4}{3} \cdot \pi \cdot R^3 = \frac{4}{3} \cdot \pi \cdot c^3 \cdot T_{\text{par}}^3. \]

- Let us denote by \(\Delta V\) the smallest possible volume of a single processor.
5. Limitations (cont-d)

- Then, the number of processor $N_{\text{proc}}$ that can fit inside this sphere cannot exceed the value

$$N_{\text{proc}} \leq N_{\text{max}} \overset{\text{def}}{=} \frac{V}{\Delta V} = \frac{4}{3} \cdot \Delta V \cdot \pi \cdot c^3 \cdot T_{\text{par}}^3.$$ 

- Whatever we can compute in parallel on $N_{\text{proc}}$ processors, we can also compute sequentially, if we:
  - first simulate all the first steps of all the processor,
  - then all the second steps of all the processors, etc.

- This way, each step of the parallel computer requires $N_{\text{proc}}$ steps of the sequential computer; thus:
  - what was computed on a parallel computer in time $T_{\text{par}}$
  - can be computed on a sequential computer in time $T_{\text{seq}} = N_{\text{proc}} \cdot T_{\text{par}}$. 
6. Limitations (cont-d)

• Due to the above inequality, we have

\[ T_{\text{seq}} \leq \frac{4}{3 \cdot \Delta V} \cdot \pi \cdot c^3 \cdot T^3_{\text{par}} \cdot T_{\text{par}} = C \cdot T^4_{\text{par}}, \]

where we denoted \( C \overset{\text{def}}{=} \frac{4}{3 \cdot \Delta V} \cdot \pi \cdot c^3. \)

• So:
  - if the fastest time that it takes for a sequential computer to solve a problem is \( T \),
  - the fastest time \( T_{\text{par}} \) that this problem can be solved on a parallel computer satisfies

\[ T \leq T_{\text{seq}} \leq C \cdot T^4_{\text{par}}. \]

• Thus \( T_{\text{par}} \geq C^{-1/4} \cdot T^{1/4}. \)
7. Limitations (cont-d)

- This implies that by using parallelization, we can speed up, at best, to the 4-th root of the sequential time.

- This is good, but not ideal:
  - if the original sequential time $T$ was exponential – as for NP-hard problems,
  - the parallel time is still exponential.
8. Extra Spatial Dimensions

• The above argument assumes that we live in a 3-dimensional space.

• However:
  – according to modern physics,
  – the requirement that quantum field theory is consistent implies that the dimension of space is at least 10.

• A natural question is: how does the presence of these extra spatial dimensions affect computations?

• This is the question that we study in this talk.
9. A Seemingly Natural Idea

- At first glance, the situation is straightforward.
- Instead of 3 spatial dimensions we have $d > 3$ dimensions.
- So, the volume of the area inside the sphere of radius $R$ is equal to $V = c_d \cdot R^d$ for some constant $c_d$.
- Taking into account that $R = c \cdot T_{\text{par}}$, we conclude that $V = c_d \cdot c^d \cdot T_{\text{par}}^d$.
- Thus, the number $N_{\text{proc}}$ of processors is bounded by the number
  $$N_{\text{proc}} \leq N_{\text{max}} \overset{\text{def}}{=} \frac{V}{\Delta V} = \frac{c_d}{\Delta V} \cdot c^d \cdot T_{\text{par}}^d.$$
10. A Seemingly Natural Idea (cont-d)

- So, this parallel computation can be simulated on a sequential computer in time

\[ T_{\text{seq}} \leq N_{\text{proc}} \cdot T_{\text{par}} = \frac{c_d}{\Delta V} \cdot c^d \cdot T_{\text{par}}^d \cdot T_{\text{par}} = C_d \cdot T_{\text{par}}^{d+1}, \]

where this time \( C_d \overset{\text{def}}{=} \frac{c_d}{\Delta V} \cdot c^d. \)

- So:
  - instead of the previous rather-high lower bound
    \[ T_{\text{par}} \geq \text{const} \cdot T_{\text{seq}}^{1/4}, \]
  - we get a much better lower bound
    \[ T_{\text{par}} \geq \text{const} \cdot T_{\text{seq}}^{1/(d+1)}, \text{ with } d \geq 10. \]
11. Why This Idea Is Naive

• The above result looks good.

• However, it is based on the simplified idea that extra spatial dimensions are similar to the current three ones.

• The fact that we currently observe only three dimensions means that different dimensions are different.

• There are two possible approaches to how to explain that other dimensions are not yet observable.

• The first natural approach is to conclude that:
  – since we cannot observe any changes in other spatial dimensions,
  – this means that these dimensions are very small in size,
  – e.g., that each of these dimensions represents not a line, but a circle of a small radius.
12. Why This Idea Is Naive (cont-d)

- The second natural approach is to assume that:
  - while all our processes are happening in a very small fragment of the additional dimensions,
  - these dimensions actually have larger size.

- Only due to some physical reasons, we cannot leave this small fragment.

- An analogy is when we are in a narrow valley between two mountain ranges.

- In principle, we can get out of this valley, but this requires climbing high mountains.

- And for that, we will need lots of energy and probably special equipment, which few of us have.
13. First Approach: How Actual Compactification Affects Computations

- We used the formula for the volume $V$ of the inside of the sphere of radius $R = c \cdot T_{par}$, where:
  - $c$ is the speed of light and
  - $T_{par}$ is the computation time.

- In the analysis of the 3-D situation, we used the formula for the volume of a sphere in the 3-D space.

- How will the resulting calculations change in the multi-D space?

- To answer this question, we need to find, for this space, what is the corresponding volume $V$.

- The distance between the points $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ in the multi-D space is equal to
  $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2 + \ldots}$. 
14. First Approach (cont-d)

• For reasonable computation time $T_{\text{par}}$, the radius $R = c \cdot T_{\text{par}}$ is large.

• Thus, it is much larger than the size $s_e$ of each extra dimension.

• Remember that this size is so small that we do not notice these extra spatial dimensions; thus:
  
  – the terms $(x_4 - y_4)^2, \ldots$ corresponding to differences in extra dimensions – and which are of order $s_e^2$
  
  – are much much smaller than the terms describing the distance in the 3-D space:

  $$(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2.$$
15. First Approach (cont-d)

• Thus, with high accuracy, we can safely assume that:
  - the distance between the two multi-D points
  - is equal to the distance between their 3-D parts:

\[
d(x, y) \approx d_3(x, y) := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}.
\]

• So, the set of all the points which are at distance \( \leq R \) from the user can be described as follows:
  - we take all the points \((x_1, x_2, x_3)\) from the corresponding 3-D sphere, and
  - for each of these points, we consider all possible combinations \((x_4, \ldots)\) of additional coordinates.
16. First Approach (cont-d)

- The size of each additional coordinate is $s_e$.
- In a $d$-dimensional space, there are $d - 3$ additional spatial coordinates.
- Thus, the overall volume of the additional part of $s_e^{d-3}$.
- So, the overall volume of the sphere in $d$-dimensional space is equal to $\frac{4}{3} \cdot \pi \cdot R^3 \cdot s_e^{d-3}$. 

17. How Many Processors Can We Fit Now?

- The multi-D volume \( \Delta V \) of a processor can be obtained by multiplying:
  - its 3-D volume \( \Delta V_3 \) by
  - its volume \( \Delta V_e \) in the extra dimensions.

- At present:
  - the size of the processor in additional dimensions is \( s_e \),
  - we get the exact same number of processors as in the 3-D case,
  - no gain at all from the existence of additional spatial dimensions.
18. How Many Processors Can We Fit (cont-d)

• However:

  – if we manage to decrease the size of a processor in extra dimensions to less than \( s_e \),
  – so that the volume \( \Delta V_e \) of a processor in the extra dimensions is smaller than \( s_e^{d-3} \),
  – then, by dividing the overall multi-D volume by the volume of a single processor,
  – we get the new value for the number of processors:

\[
N_{\text{proc}} \leq N_{\text{max}} = \frac{V}{\Delta V} = \frac{4}{3 \cdot \Delta V_3} \cdot \pi \cdot R^3 \cdot \frac{s_e^{d-3}}{\Delta V_e}.
\]
19. How Many Processors Can We Fit (cont-d)

- Since we consider the case when \( \Delta V_e < s_e^{d-3} \):
  - this number of processors is larger than the corresponding 3-D number of processors \( \frac{4}{3 \cdot \Delta V_3} \cdot \pi \cdot R^3 \)
  - by a factor of \( C = \frac{s_e^{d-3}}{\Delta V_e} > 1 \).
20. Conclusion for This Approach

- The first approach to multi-D space-time is when all extra dimensions are actually compactified:
  - after an appropriate level of miniaturization,
  - we will be able to get a $C$ times increase in number of processors that we can fit into each area.
- Thus, in principle, we get a constant times computation speed-up.
- This is not as spectacular as we could imagine based on the naive approach, but any speed up is good.
21. Second Approach: How It Affects Computations

• Here:
  – if we limit ourselves to the same small area of extra dimensions where all observable processes occur,
  – then we get the exact same situation as in the first approach – and
  – thus, we can get the same constant times increase,
  – where the constant depends on how successful we are in minituarizing our processors.

• However, we do not have to limit ourselves to the small area that contains all observable processes.
22. Second Approach (cont-d)

- There are other areas as well, it is just that these areas are difficult to reach.
- Since going there requires a lot of energy, thus preventing usual particles from going there.
- What if we apply this considerable amount of energy and reach these additional areas?
- What do we gain with respect to computations?
23. First Gain: All the Promises of the Naive Approach Turn out to Be True

- We are allowed to use a significant area in extra dimensions.
- So, we can have all the advantages promised by the above-described naive approach:
  - instead of being able to fit $\sim T_{\text{par}}^3$ processors into an area of radius $R = c \cdot T_{\text{par}}$,
  - we can fit a much larger amount of $\sim T_{\text{par}}^d$ processors.
- Thus:
  - instead of the possibility to reduce the sequential computation time $T_{\text{seq}}$ to $T_{\text{par}} \sim T_{\text{seq}}^{1/4}$,
  - we can get a much more drastic speed-up $T_{\text{par}} \sim T_{\text{seq}}^{1/(d+1)}$. 
24. Interestingly, There Is an Additional Speed-Up

- *All* the processes are limited to a narrow area of values of extra spatial dimensions.

- This means, in effect, that:
  - this limitation is the property of the underlying space-time,
  - not of any specific physical field.

- In other words, this means that:
  - the space-time is not as flat as the space-time of our usual 3D space – that would have enabled particles to easily go in all possible spatial directions,
  - but rather curved.

- In modern physics, General Relativity describes curved space-time.
25. Additional Speed-Up (cont-d)

• In this theory:
  – free particles move along geodesic lines,
  – i.e., lines in which the resulting proper time $\Delta s$ between the each two locations is the shortest possible.

• In terms of coordinate time $t$:
  – this overall proper time can be computed
  – by adding up proper times $ds = \frac{ds}{dt} \cdot dt$ corresponding to different parts of the trajectory, i.e., as:
    $$\Delta s = \int \frac{ds}{dt} \, dt.$$ 

• According to General Relativity, the ratio $\frac{ds}{dt}$ is, in general, smaller than 1.
26. Additional Speed-Up (cont-d)

- In a gravitational field, all the processes slow down.
- If this field is very strong – e.g., near a black hole – then it can slow down drastically.
- When the outside world measures 10 years, people near the black hole will only count several months.
- In our previous work, we considered possible computational consequences of this effect in the 3D space.
- Interestingly, in the 2nd approach to the multi-D cases, we have an additional possibility to use this effect.
- Indeed, for all the particles, the optimal path is by going via the narrow zone of observable processes.
- This means that in this zone, the ratio $\frac{ds}{dt}$ is much smaller than in the neighboring zones.
27. Additional Speed-Up (cont-d)

• Similarly:
  – the fact that the fastest way to get from two points in the US usually involves taking a freeway
  – is an indication that the allowed speed on the freeway is larger than on all other roads.

• For example, if we are in the vicinity of a gravitating body, where the ratio \( \frac{ds}{dt} \) is smaller than 1.

• This is thus an analogue of a freeway.

• Particles will tend to move close to this vicinity, which we observe as gravitational attraction.
28. Additional Speed-Up (cont-d)

- The stronger the gravitational field:
  - the smaller the ratio \( \frac{ds}{dt} \) and thus,
  - the more probable it is that the particles will bend towards this vicinity,
  - so the larger the observed gravitational attraction.

- In our multi-D case:
  - the fact that in the neighborhood of our zone the value of the ratio is much larger than in the zone,
  - means that during the time \( \Delta t \), the proper time \( \Delta s \) in this neighborhood is larger than in the zone.
29. Additional Speed-Up (cont-d)

• In other words, during the same coordinate time:
  – the processor located in the neighborhood will be able to perform more operations
  – than a processor that stays in our zone.

• Thus, we will get an additional speed-up.
30. Conclusion for This Approach

- In the second approach,
  - we can have more processor working in parallel,
  - by placing additional processors outside the narrow zone where the observable processes occur.

- Also:
  - the processors placed outside this zone will compute much faster than the ones in the zone,
  - which will lead to an additional speedup.
31. Acknowledgments

This work was supported in part by the National Science Foundation grants:

- 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science);
- HRD-1834620 and HRD-2034030 (CAHSI Includes).

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478.