Schlegel Diagram and Optimizable Immediate Snapshot Protocol

Susumu Nishimura

Dept. Math, Kyoto University

susumu@math.kyoto-u.ac.jp

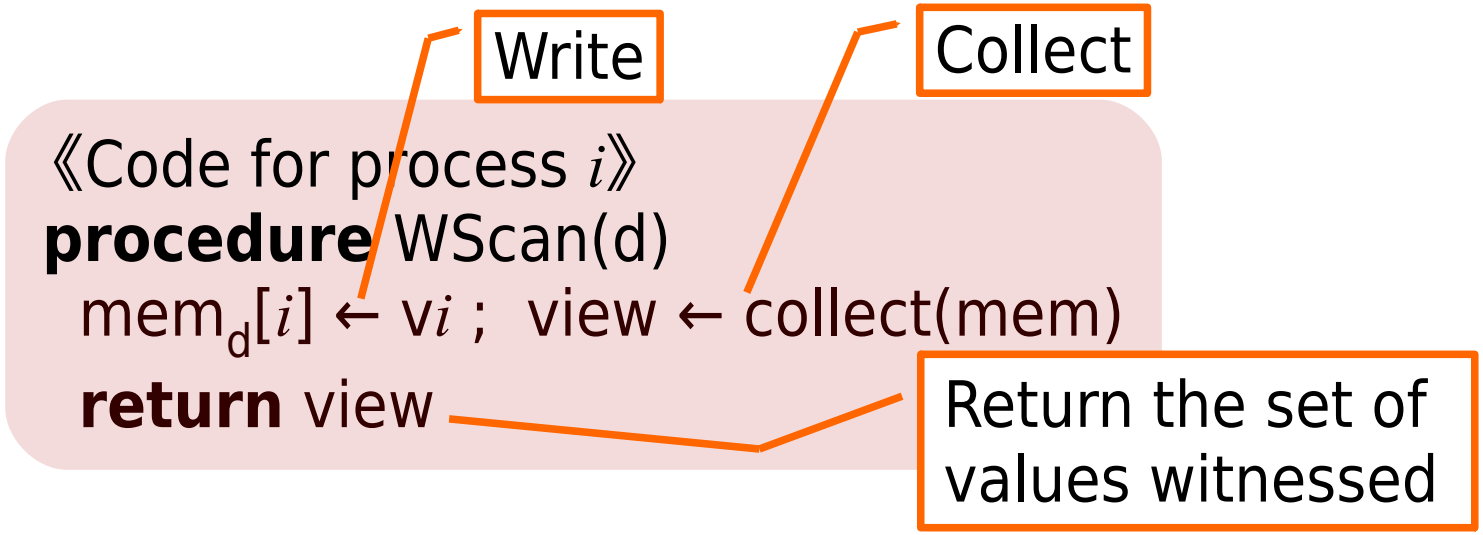
GRADUATE SCHOOL OF FACULTY OF SCIENCE
KYOTO UNIVERSITY

Immediate snapshot

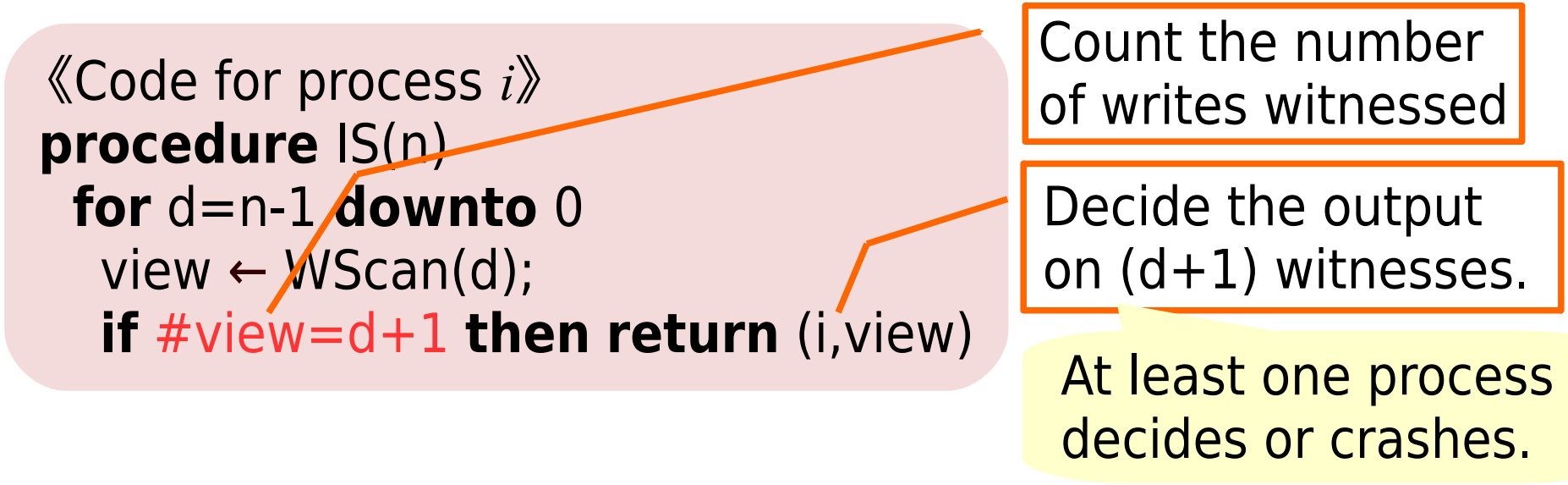
- A memory snapshot operation in the asynchronous read-write shared memory model
  - Central to the study of wait-free computability, in the topological theory of distributed computing
  - No native support by real computing devices
  - Wait-free protocol by Borowsky&Gafni[1993]

Immediate snapshot protocol [BorowskyGafni93]

- *Write&Scan* – a single protocol round

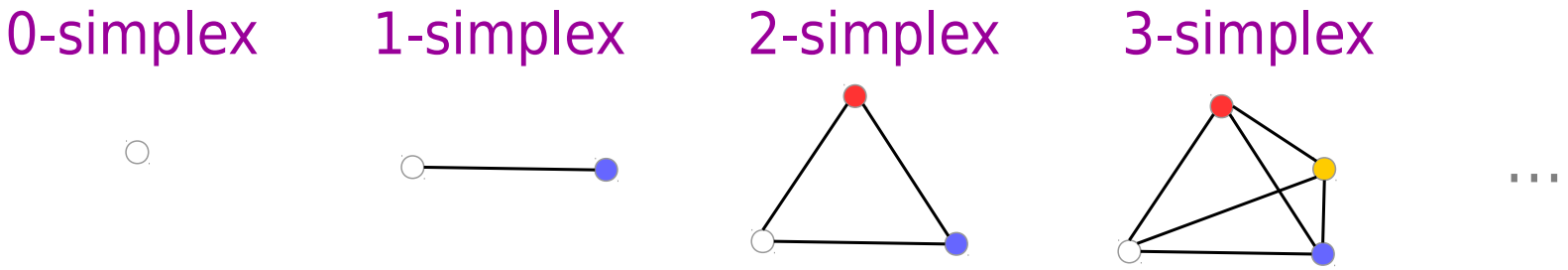


- Multi-round immediate snapshot protocol for n processes

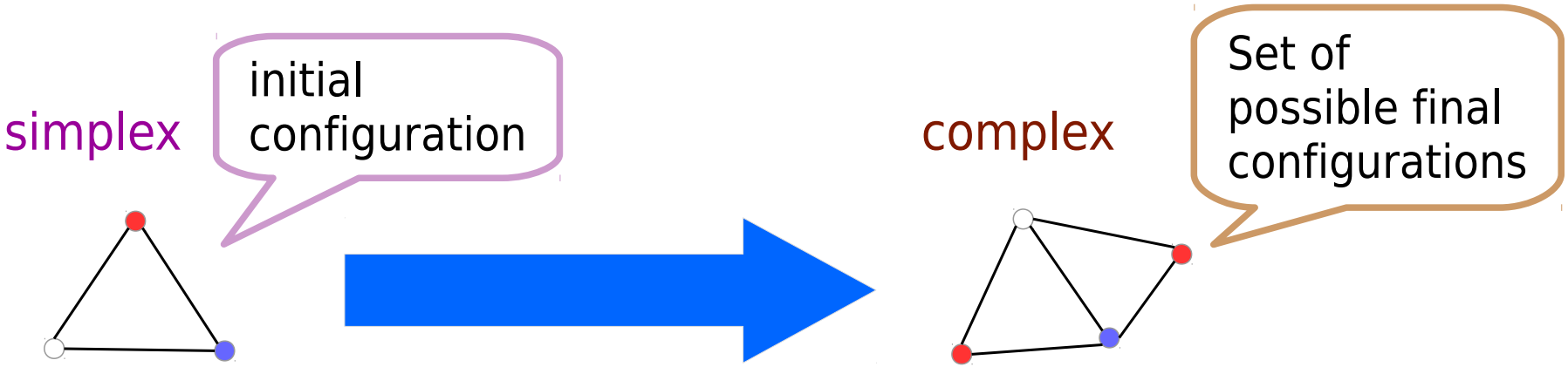


Combinatorial topological model

- *Process* = colored *vertex.* (color=process id)
- *Configuration* of a system of n processes = colored (n-1)-*simplex*.

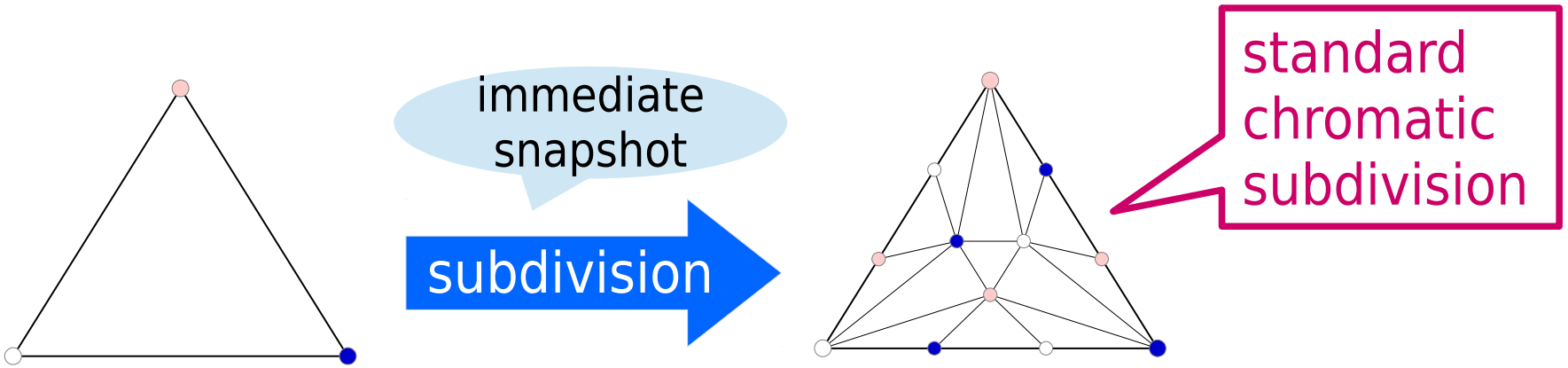


- Distributed system = topological transformation



Immediate snapshot as a subdivision

- Immediate snapshot is a *subdivision*.

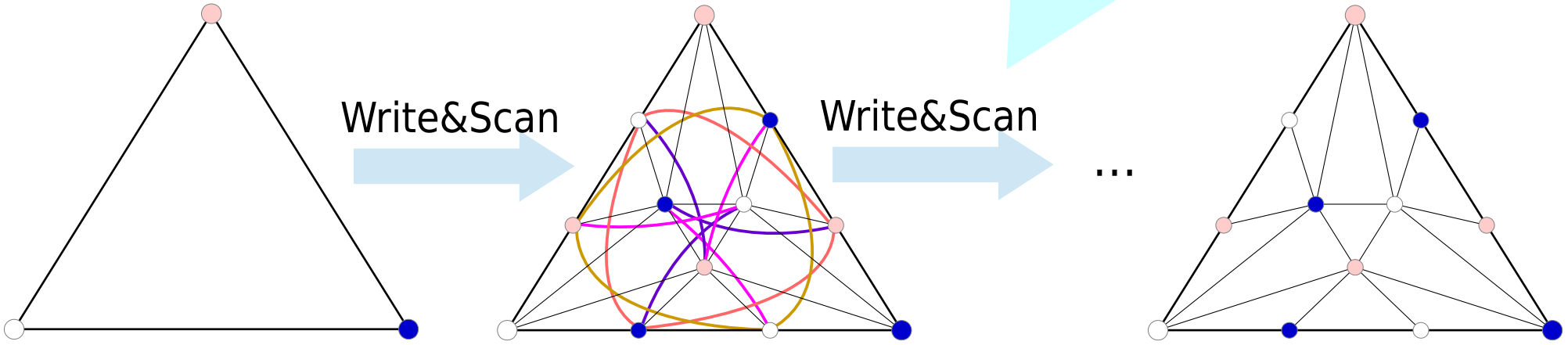


  - **Asynchronous Computability Theorem** **[Herlihy&Shavit99]** gives a topological characterization of wait-free solvability by means of standard chromatic subdivision.

Topological structure of Borowsky-Gafni protocol

- The first round (Write&Scan) of Borowsky-Gafni protocol produces an intricate topological structure.

Superfluous simplexes are *collapsed* in the subsequent protocol rounds. [Benavides&Rajsbaum16]



Contributions

I. Reformulation of Borowsky-Gafni immediate snapshot protocol

- Direct correspondence to *Schlegel diagram*
- Simpler topological structure

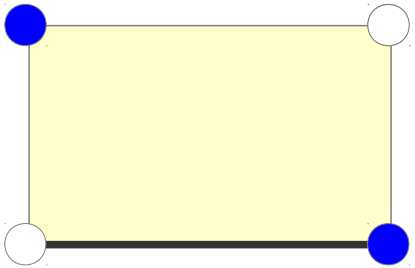
II. Protocol optimization

- Reduced shared memory access
- Mechanizable optimization w/ program specialization

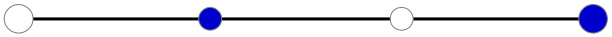
I. Reformulation of Borowsky-Gafni protocol

Schlegel diagram

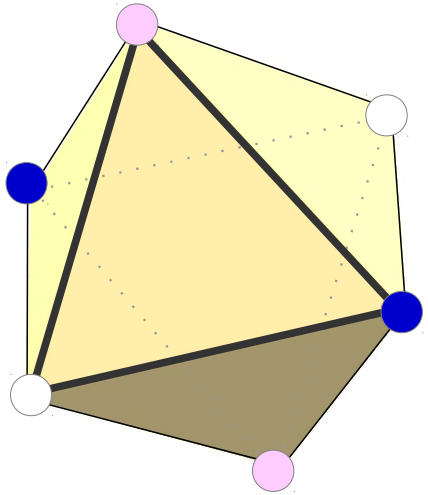
- A Schlegel diagram is a subdivision of a simplex derived from a crosspolytope.
  - Quadrilateral



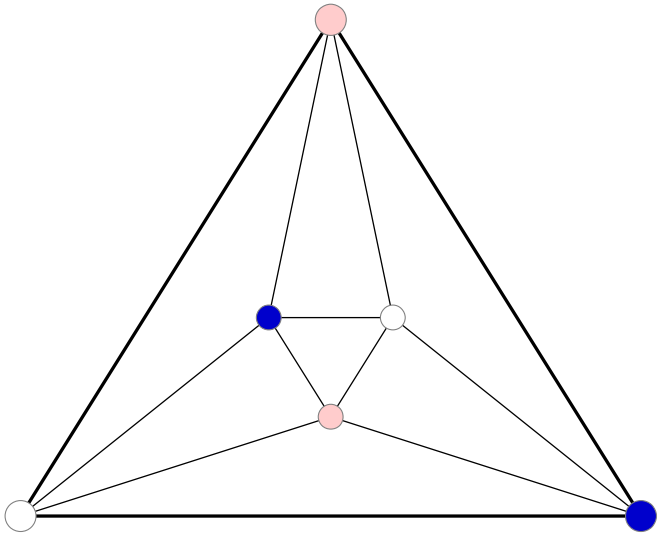
Schlegel diagram (1-simplex)



  - Octahedron

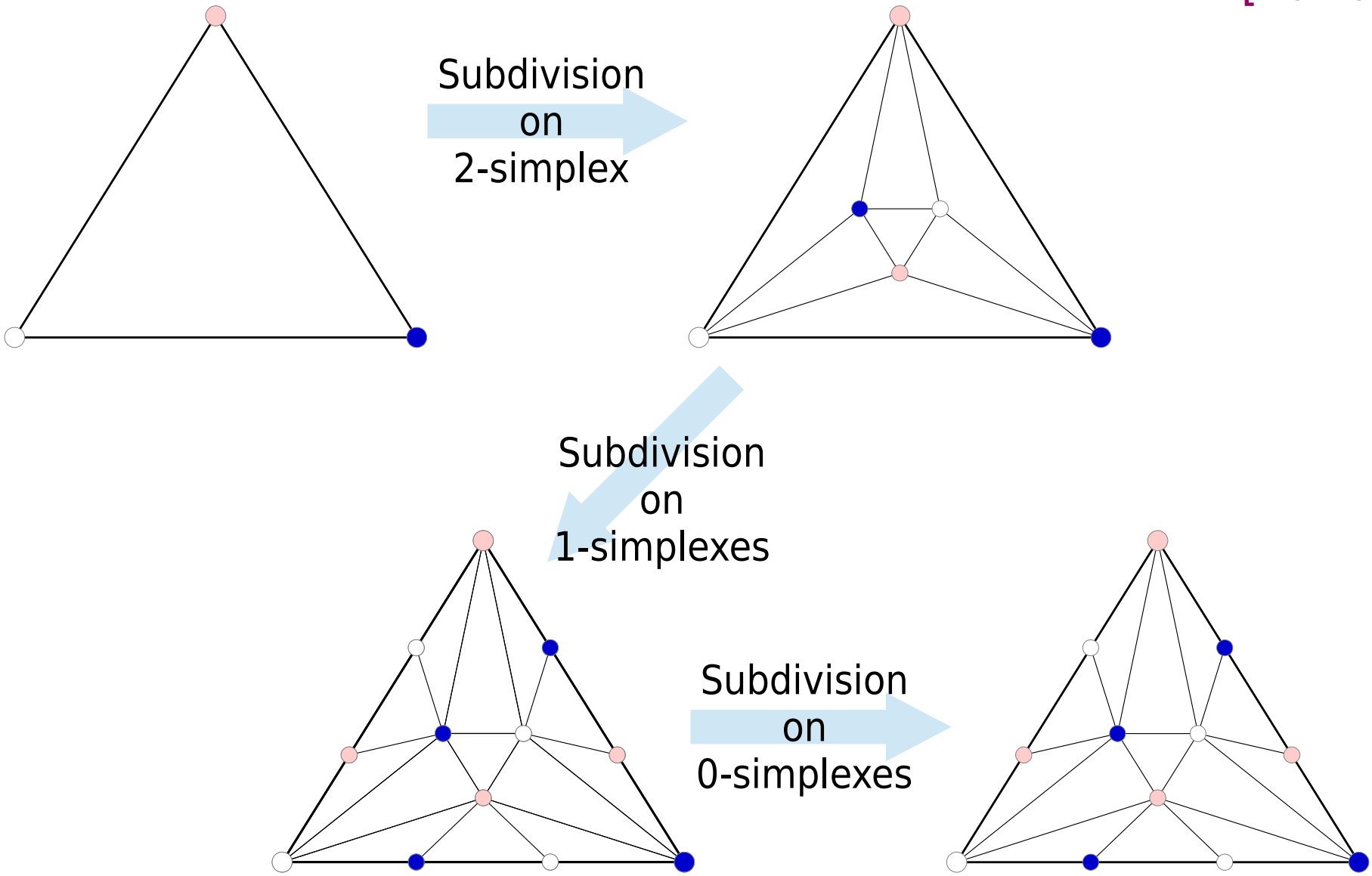


Schlegel diagram (2-simplex)

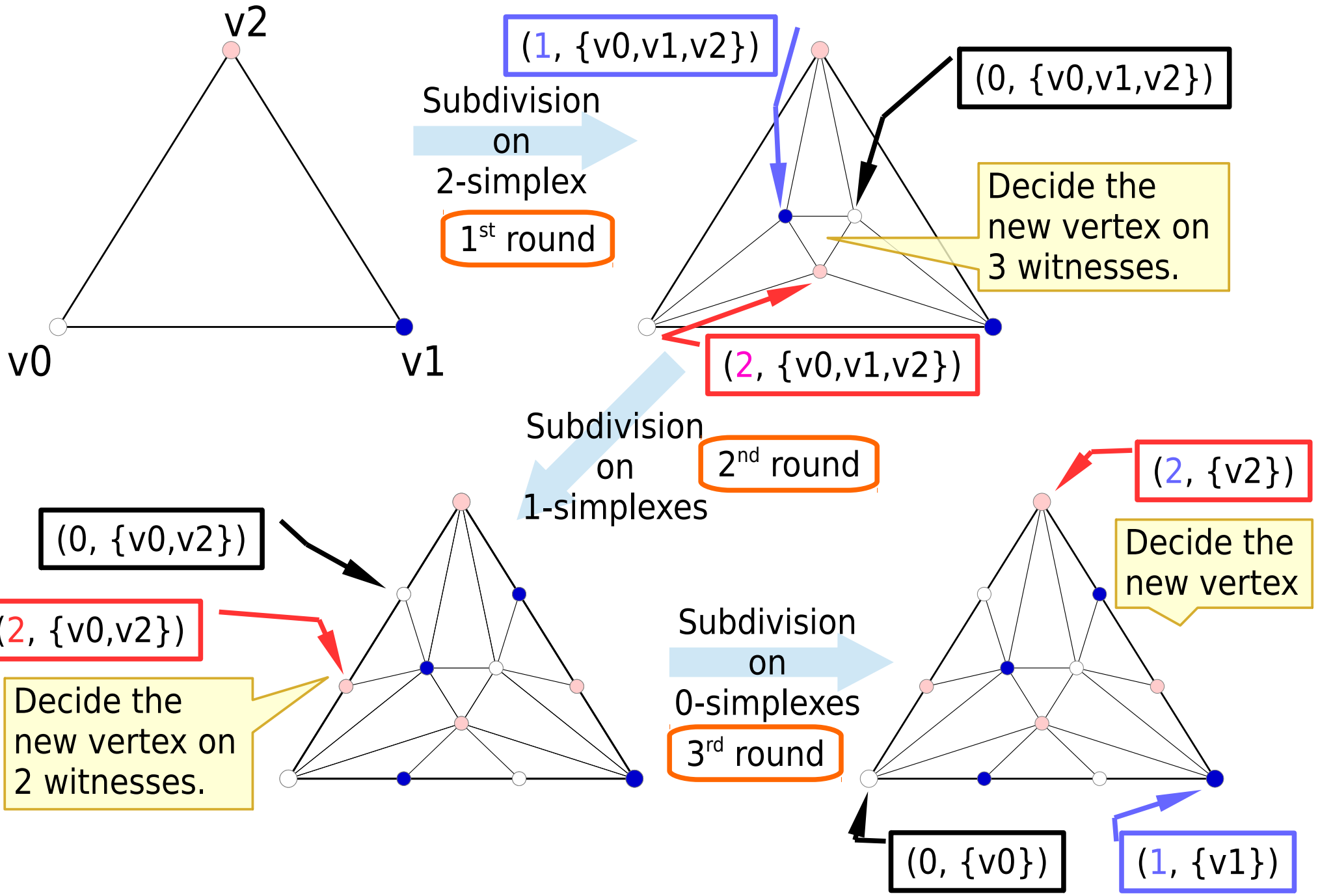


Standard chromatic subdivision via Schlegel diagram

[Kozlov12]

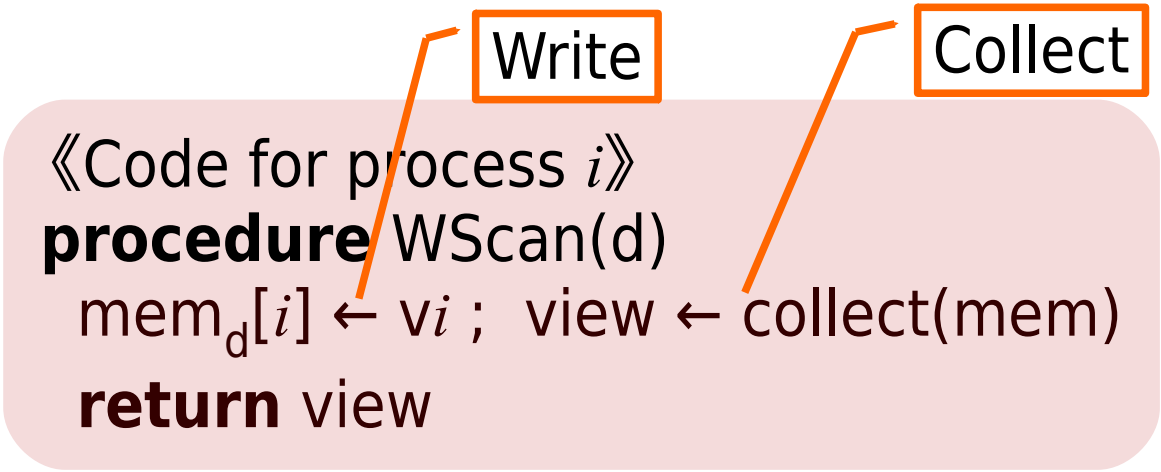


Immediate snapshot as iterated subdivision

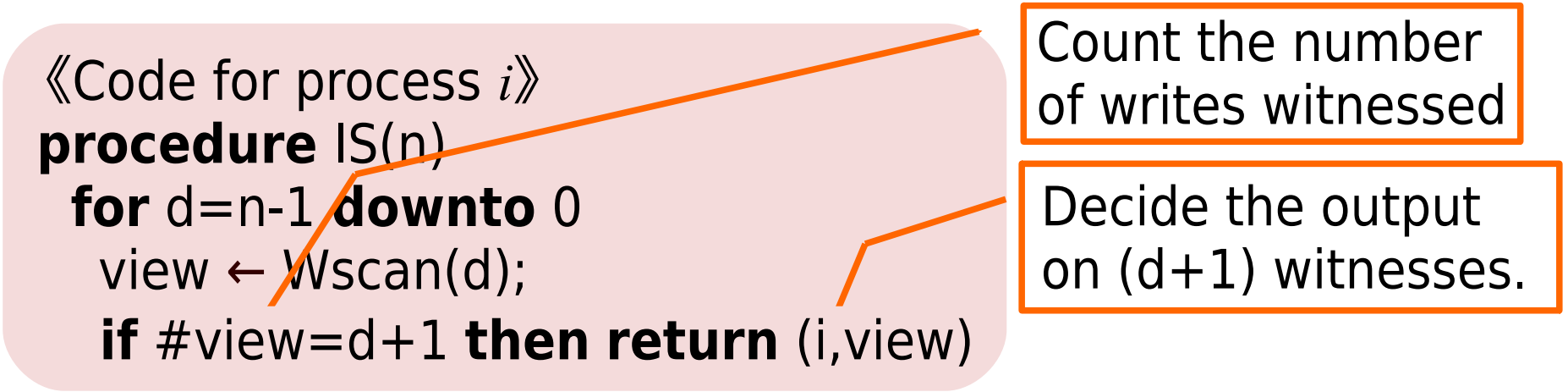


Immediate snapshot protocol [BorowskyGafni93]

- *Write&Scan* – a single protocol round

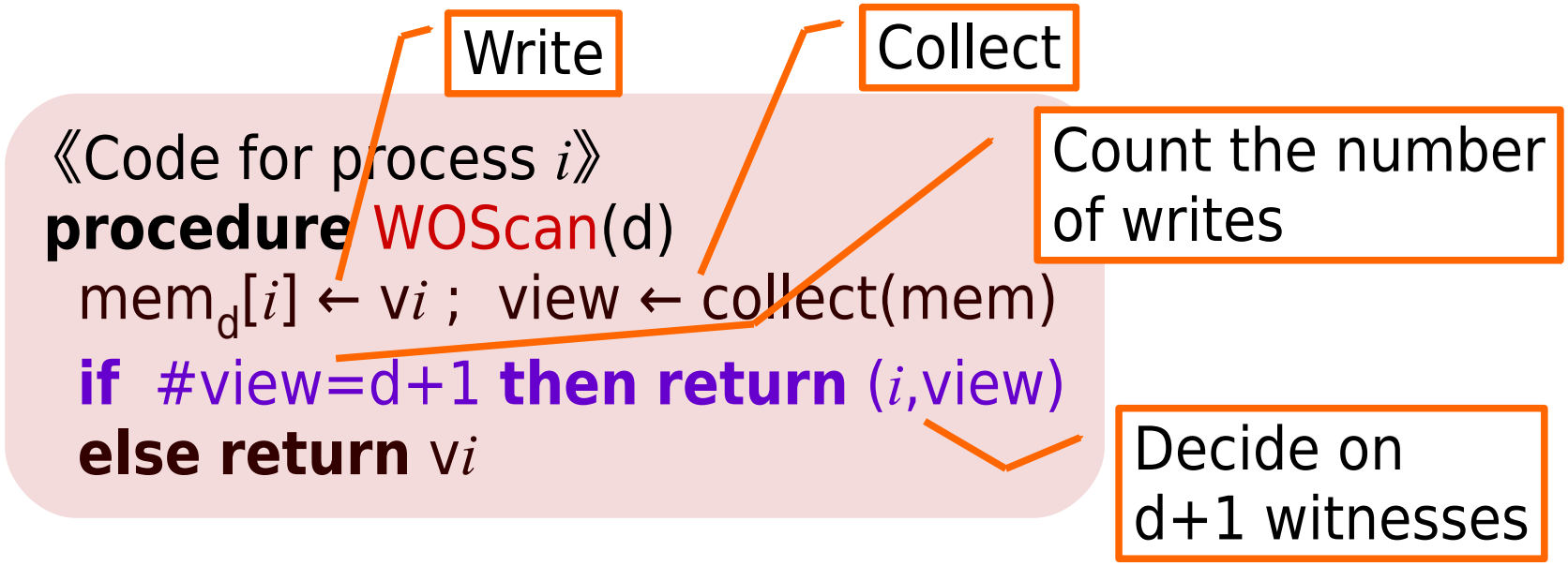


- Multi-round immediate snapshot protocol for n processes

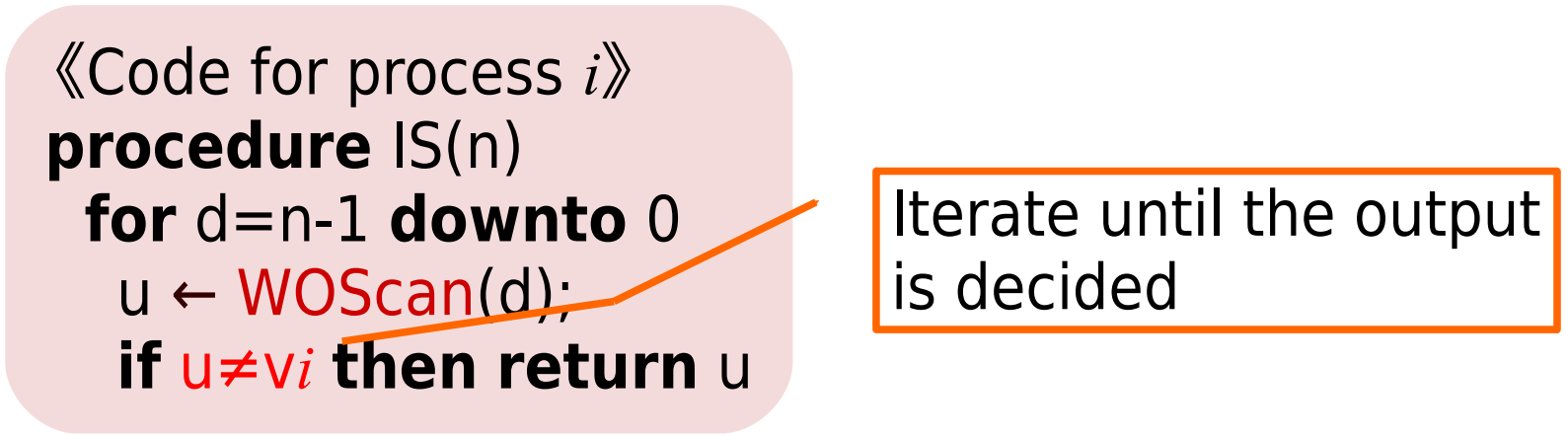


Immediate snapshot protocol, reformulated

- *Write&Oblivious Scan* – subdivision using Schlegel diagram



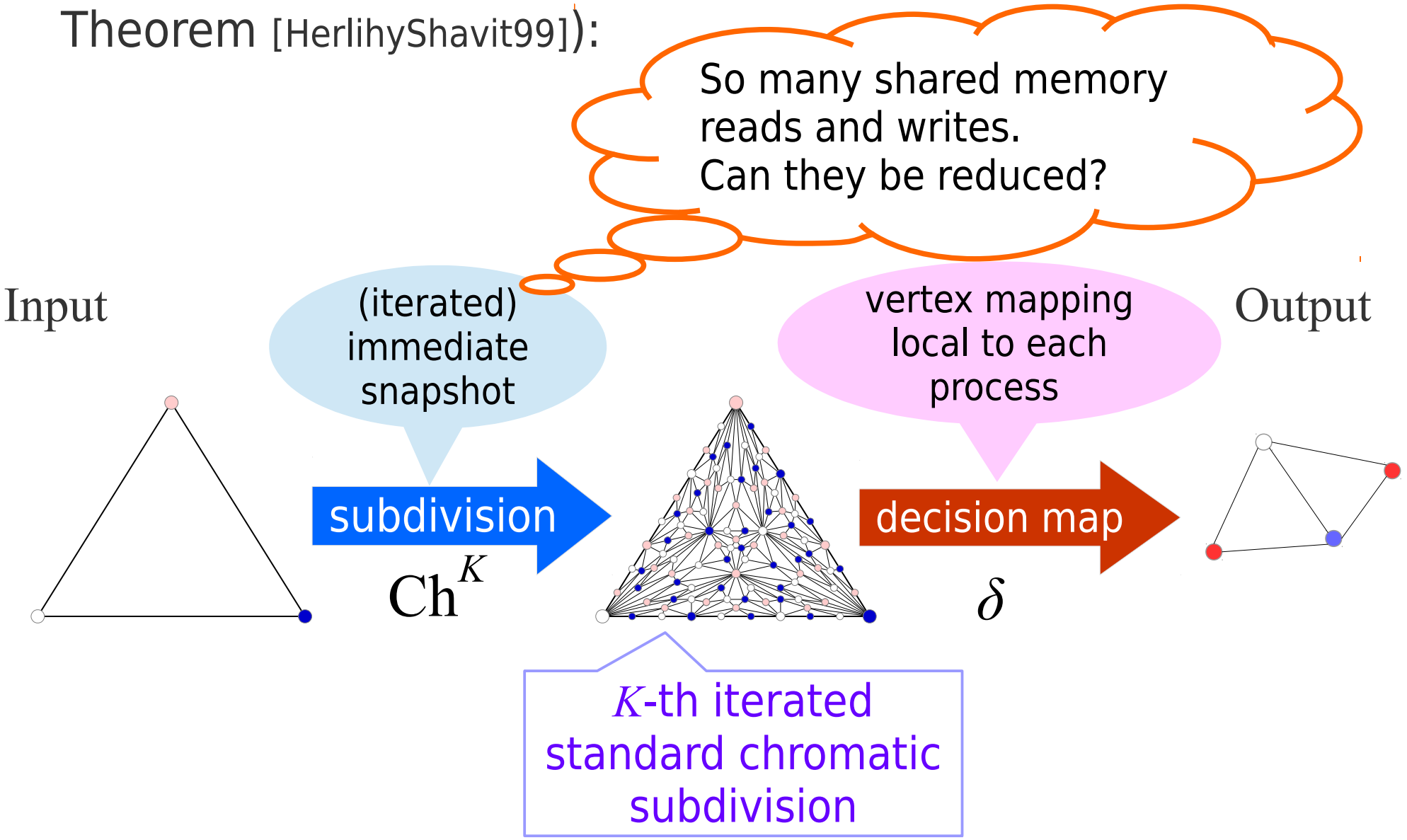
- Multi-round immediate snapshot protocol for n processes



II. Protocol Optimization

Universal implementation of wait-free protocol

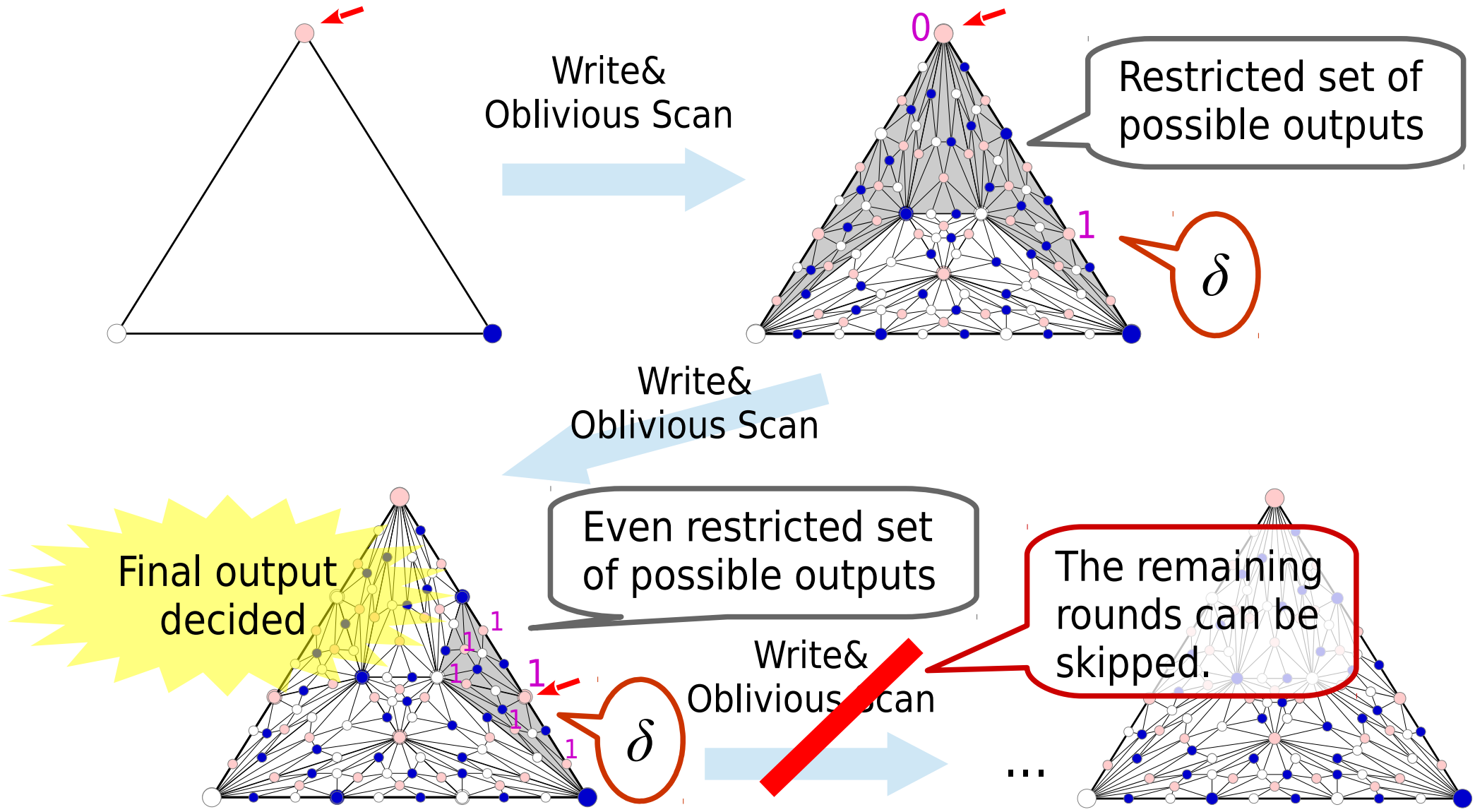
- Every wait-free computable task is implementable in a universal way (by way of Asynchronous Computability Theorem [HerlihyShavit99]):



Optimizing universal implementation

- Each Write&Oblivious Scan restricts the set of possible final outputs to a smaller subset.
- Once restricted to a singleton set, the final output can be decided even at an earlier intermediate protocol round.

Earlier decision making for an instance $\delta \circ \mathrm{Ch}^2$



Optimized iterated immediate snapshot

- K-iterated immediate snapshot for n processes

《Code for process i》
procedure IIS(K,n)
  for k=1 **to** K
    for d=n-1 **downto** 0
      if $\delta(\pi^{-1}(v_i))=\{u\}$ for some u **then return** u
      u ← WOScan(d);
      if u≠v_i **then** v_i ← u

The set $\delta(\pi^{-1}(v_i))$ of possible outputs has converged to a singleton set.

N.B. The degree of optimization depends on each particular instance of protocol.

- No general theorem on speedup ratio or complexity improvement.
- Reduced rounds relieve the shared memory contention.

Further optimization by program specialization

- Optimization by partial evaluation [Jones et al. 1993]
  - For each different vi, the set of possible outputs $\delta(\pi^{-1}(\mathrm{v}i))$ is *statically* computed at compile-time (i.e., in advance of run-time).
    - Every convergence codition is evaluated to a constant truth value, true or false;
    - The unreachable conditional branch is eliminated.

Conclusion and future work

- Topologically simpler reformulation of Borowsky-Gafni immediate snapshot protocol
  - Direct correspondence of each protocol round with Schlegel diagram
  - Mechanizable optimization for reduced shared memory access

- Future direction of research
  - Richer memory model beyond read-write shared memory
    - k-concurrency [Gafni+14], t-resilience [Saraph+16], etc.
    - Protocol implementation with direct topological correspondence.