Solving Large-Scale Dense Linear Systems on Desktop Computers

Mercedes Marqués
Enrique S. Quintana-Ortí
Gregorio Quintana-Ortí

Universidad Jaime I de Castellón (España)

Robert A. van de Geijn

The University of Texas at Austin

METNUM – June, 2009

Motivation

Linear systems $Ax = b$ with $A \rightarrow n \times n$ dense and $n \rightarrow O(10,000 - 100,000)$:

- Estimation of the Earth gravitational field
- BEM in electromagnetics and acoustics
- Molecular dynamic simulations

- Only two years ago, these problems were considered "large" and a cluster was employed for their solution
- Current multi-core processors provide sufficient *computational power* to tackle them...
  - Computational cost is $O(n^3) \rightarrow$ be patient
  - Storage cost is $O(n^2) \rightarrow$ use disk

Motivation

Conventional wisdom against use of disk

- Disk is too slow to feed the processor
- OS can deal with data on disk, but careful design is required to reduce I/O $\rightarrow$ Out-of-Core (OOC) algorithms
- Programming OOC is cumbersome (e.g., asynchronous I/O)

No longer true!

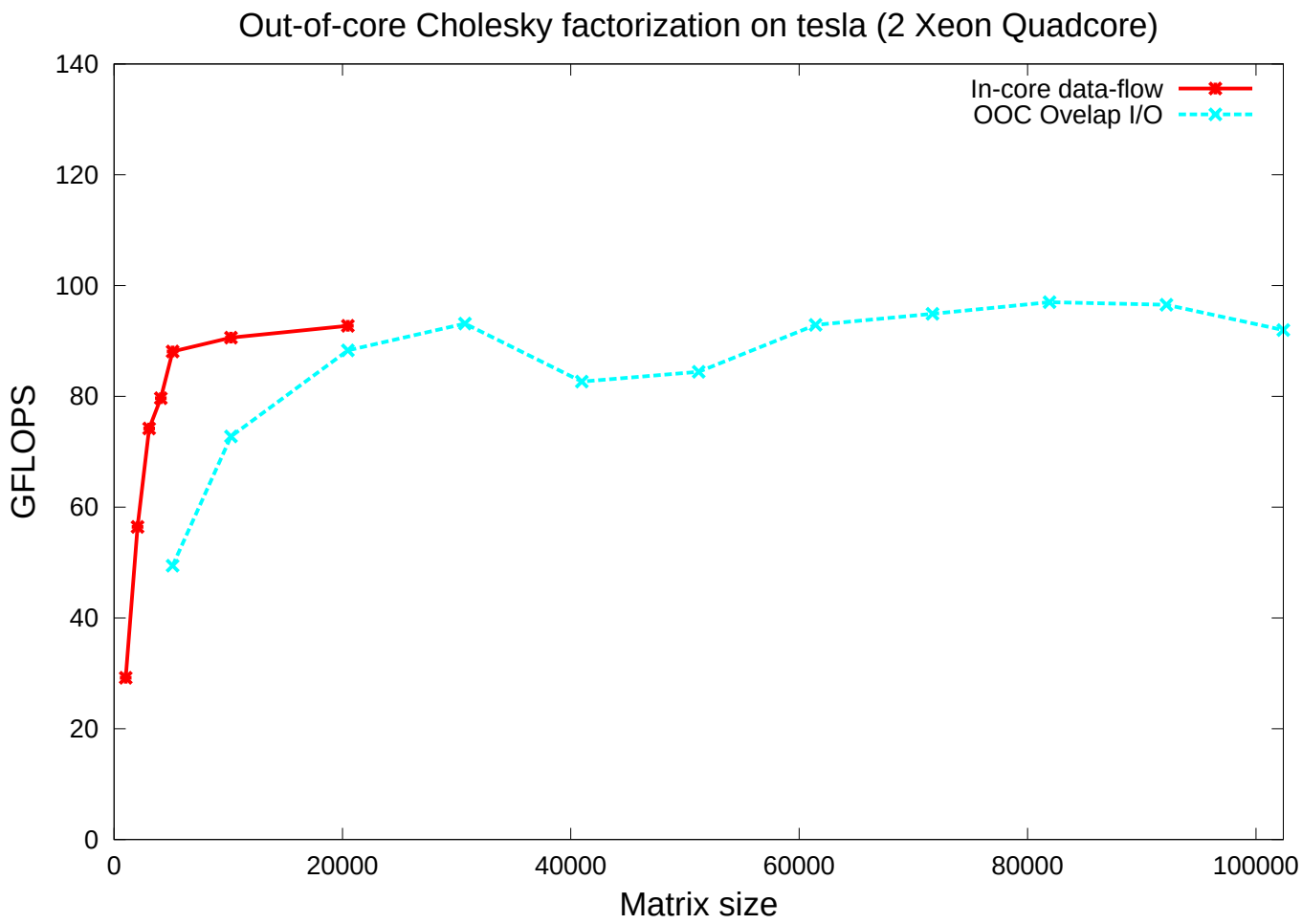
Motivation

Conventional wisdom against use of disk

- Disk is too slow to feed the processor
- OS can deal with data on disk, but careful design is required to reduce I/O $\rightarrow$ Out-of-Core (OOC) algorithms
- Programming OOC is cumbersome (e.g., asynchronous I/O)

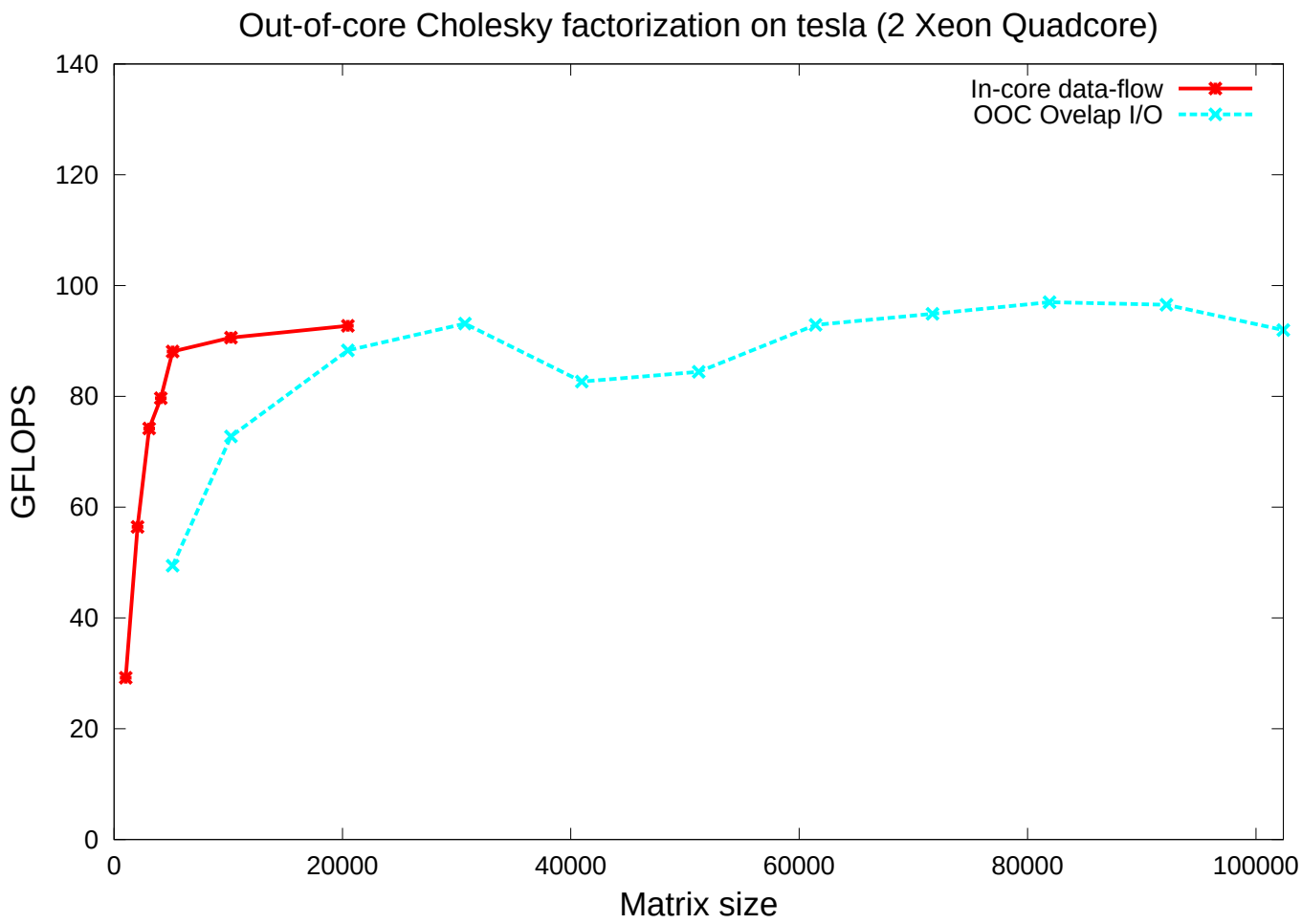
No longer true!

A preview ...



Solution of a 100,000×100,000 s.p.d. linear system in 1 hour and 4 minutes!

A preview ...



Solution of a 100,000×100,000 s.p.d. linear system in 1 hour and 4 minutes!

Outline

1. Motivation
2. Cholesky factorization (Overview of FLAME)
3. OOC
4. Parallelization
5. Experimental results
6. Concluding remarks

The Cholesky Factorization

Definition

Given $A \rightarrow n \times n$ symmetric positive definite, compute

$$A = L \cdot L^T,$$

with $L \rightarrow n \times n$ lower triangular

Algorithms should ideally be represented in a way that captures how we reason about them

The Cholesky Factorization

Definition

Given $A \rightarrow n \times n$ symmetric positive definite, compute

$$A = L \cdot L^T,$$

with $L \rightarrow n \times n$ lower triangular

Algorithms should ideally be represented in a way that captures how we reason about them

FLAME Notation

Algorithm: $[A] := \text{CHOL\_BLK}(A)$

Partition $A \rightarrow \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right)$

where A_{TL} **is** 0×0

while $n(A_{BR}) \neq 0$ **do**

Determine block size b

Repartition

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \rightarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

where A_{11} **is** $b \times b$

$$A_{11} = L_{11} L_{11}^{\mathrm{T}}$$
$$A_{21} := A_{21} L_{11}^{-\mathrm{T}}$$
$$A_{22} := A_{22} - A_{21} A_{21}$$

Continue with

$$\left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline A_{BL} & A_{BR} \end{array}\right) \leftarrow \left(\begin{array}{c|c|c} A_{00} & A_{01} & A_{02} \\ \hline A_{10} & A_{11} & A_{12} \\ \hline A_{20} & A_{21} & A_{22} \end{array}\right)$$

endwhile

FLAME/C Code

```
int FLA_Cholesky_blk( FLA_Obj A, int nb_alg )
{
  /* ... FLA_Part_2x2( ); ... */
  while ( FLA_Obj_width( ATL ) < FLA_Obj_width( A ) ){
    /* ... */
    FLA_Repart_2x2_to_3x3(
            ATL, /**/ ATR,       &A00, /**/ &A01, &A02,
         /* ************ */   /* ******************** */
                                 &A10, /**/ &A11, &A12,
            ABL, /**/ ABR,       &A20, /**/ &A21, &A22,
            b, b, FLA_BR );

    /*--------------------------------------------------------*/
    FLA_Cholesky_unb( A11 );           /* A21 := Cholesky( A11 )  */
    FLA_Trsm( FLA_ONE, A11,
                       A21 );          /* A21 := A21 * inv( A11 )'*/
    FLA_Syrk( FLA_MINUS_ONE, A21,
                             A22 );    /* A22 := A22 - A21 * A21' */
    /*--------------------------------------------------------*/
    /* ... FLA_Cont_with_3x3_to_2x2( ); ... */
  }
}
```

Outline

1. Motivation
2. Cholesky factorization (Overview of FLAME)
3. OOC
4. Parallelization
5. Experimental results
6. Concluding remarks

Algorithms-by-tiles and algorithms-by-blocks

Matrix of tiles/blocks

$$A \rightarrow \begin{pmatrix} \bar{A}_{0,0} & \star & \star & \cdots & \star \\ \bar{A}_{1,0} & \bar{A}_{1,1} & \star & \cdots & \star \\ \bar{A}_{2,0} & \bar{A}_{2,1} & \bar{A}_{2,2} & \cdots & \star \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{A}_{N-1,0} & \bar{A}_{N-1,1} & \bar{A}_{N-1,2} & \cdots & \bar{A}_{N-1,N-1} \end{pmatrix}$$

Hierarchical organization

- Each tile is a matrix of blocks
- The tile is the unit of storage on disk
- The block is the unit of computation in the processor

Conventional implementation of OOC algorithm

```
int FLA_Cholesky_blk( FLA_Obj A, int nb_alg )
{
  /* ... FLA_Part_2x2( ); ... */
  while ( FLA_Obj_width( ATL ) < FLA_Obj_width( A ) ){
    /* ... */
    /* ... FLA_Repart_2x2_to_3x3( ); ... */

    /*-----------------------------------------------------------*/
    FLA_Cholesky_unb( A11 );           /* A21 := Cholesky( A11 )  */
    FLA_Trsm( FLA_ONE, A11,
                       A21 );          /* A21 := A21 * inv( A11 )'*/
    FLA_Syrk( FLA_MINUS_ONE, A21,
                             A22 );    /* A22 := A22 - A21 * A21' */
    /*-----------------------------------------------------------*/
    /* ... FLA_Cont_with_3x3_to_2x2( ); ... */
  }
}
```

- Insert explicit I/O calls to move data between RAM and disk
- Use large tiles to occupy most of RAM
- Asynchronous I/O?

Unconventional OOC: Transparent OOC

First stage: build list of *pending tasks*

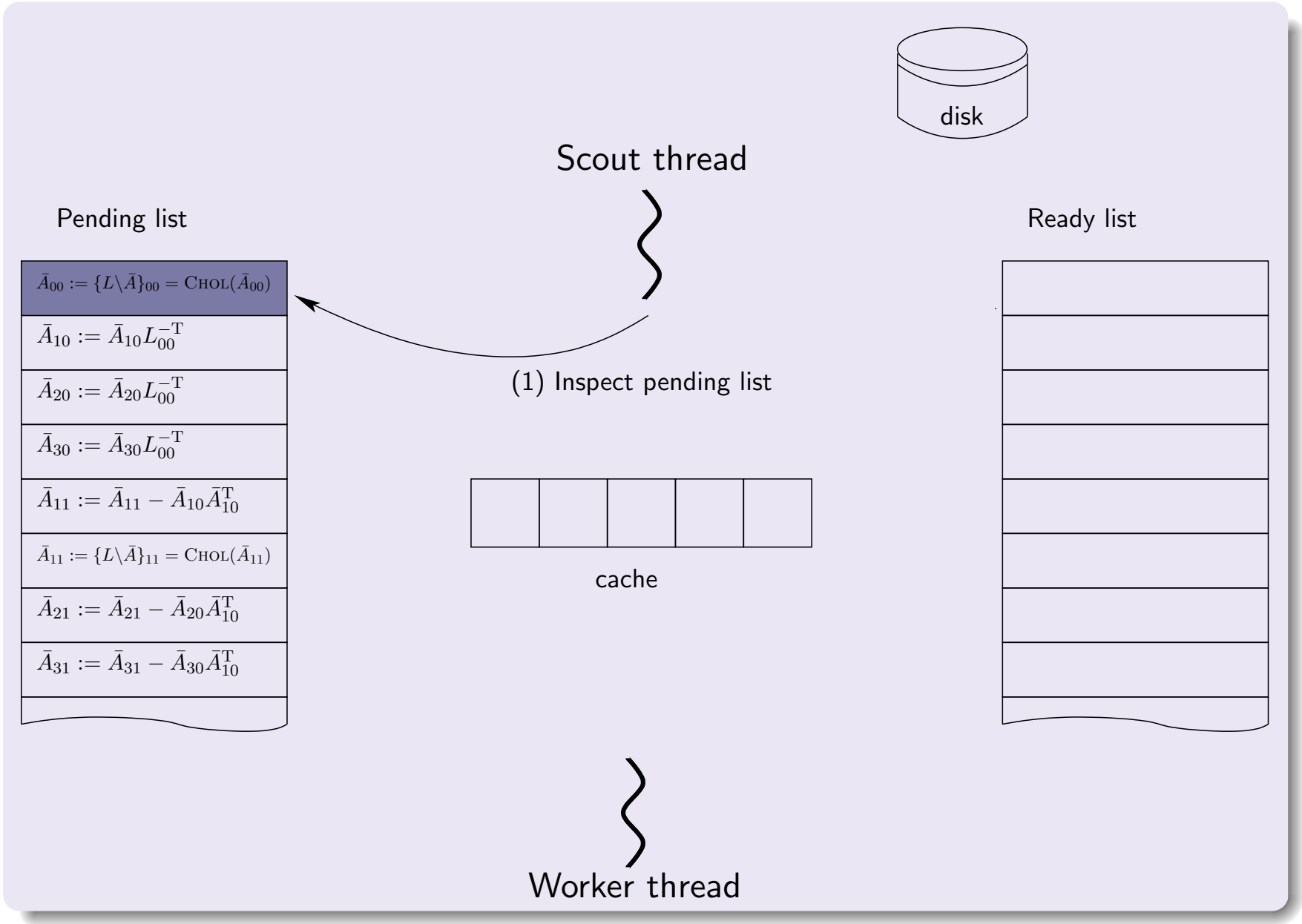
$$\begin{array}{|c|c|c|c|}\hline \bar{A}_{0,0} & \star & \star & \cdots \\ \hline \bar{A}_{1,0} & \bar{A}_{1,1} & \star & \cdots \\ \bar{A}_{2,0} & \bar{A}_{2,1} & \bar{A}_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \hline \end{array}$$

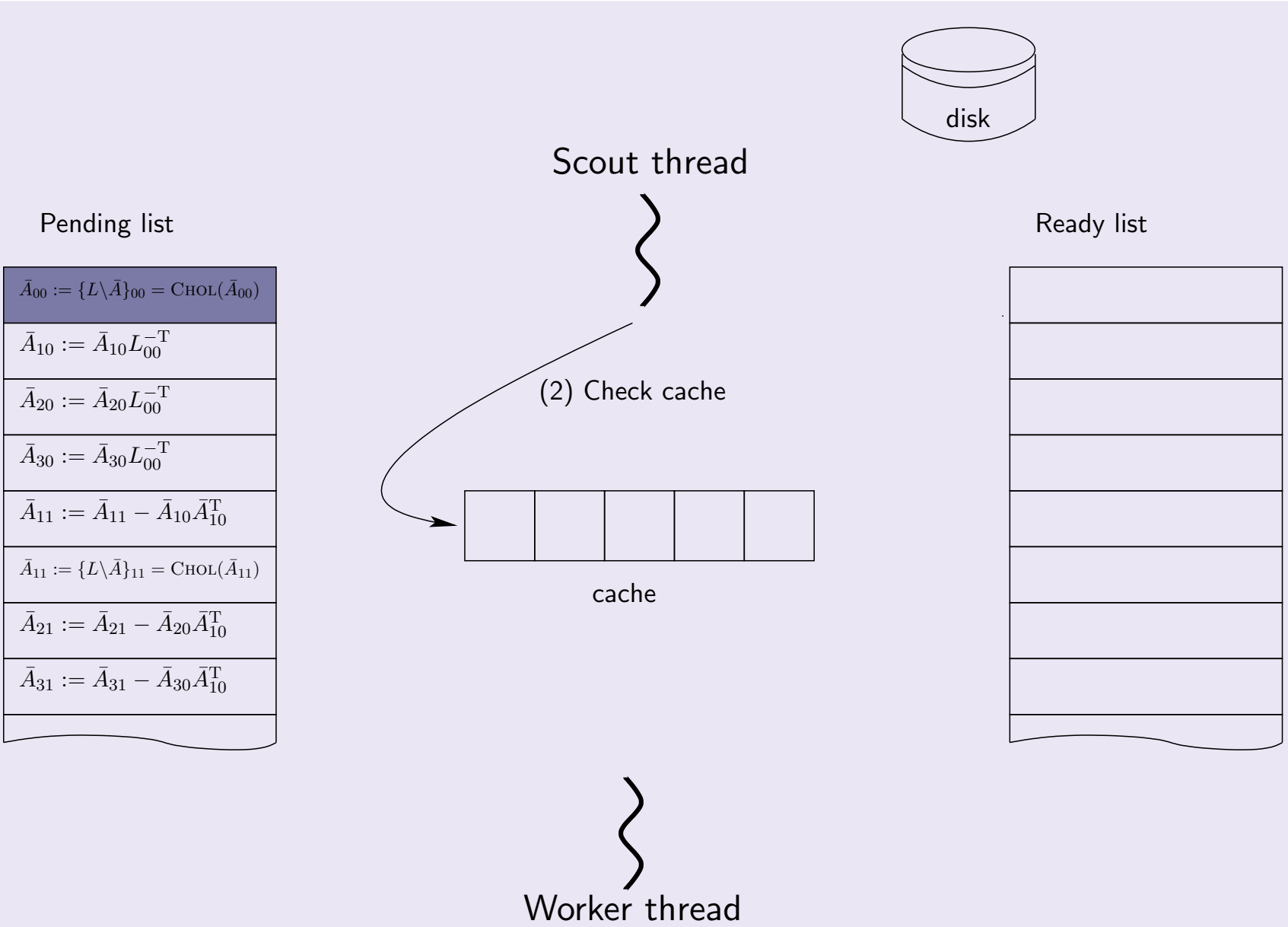
$\longrightarrow$ at runtime

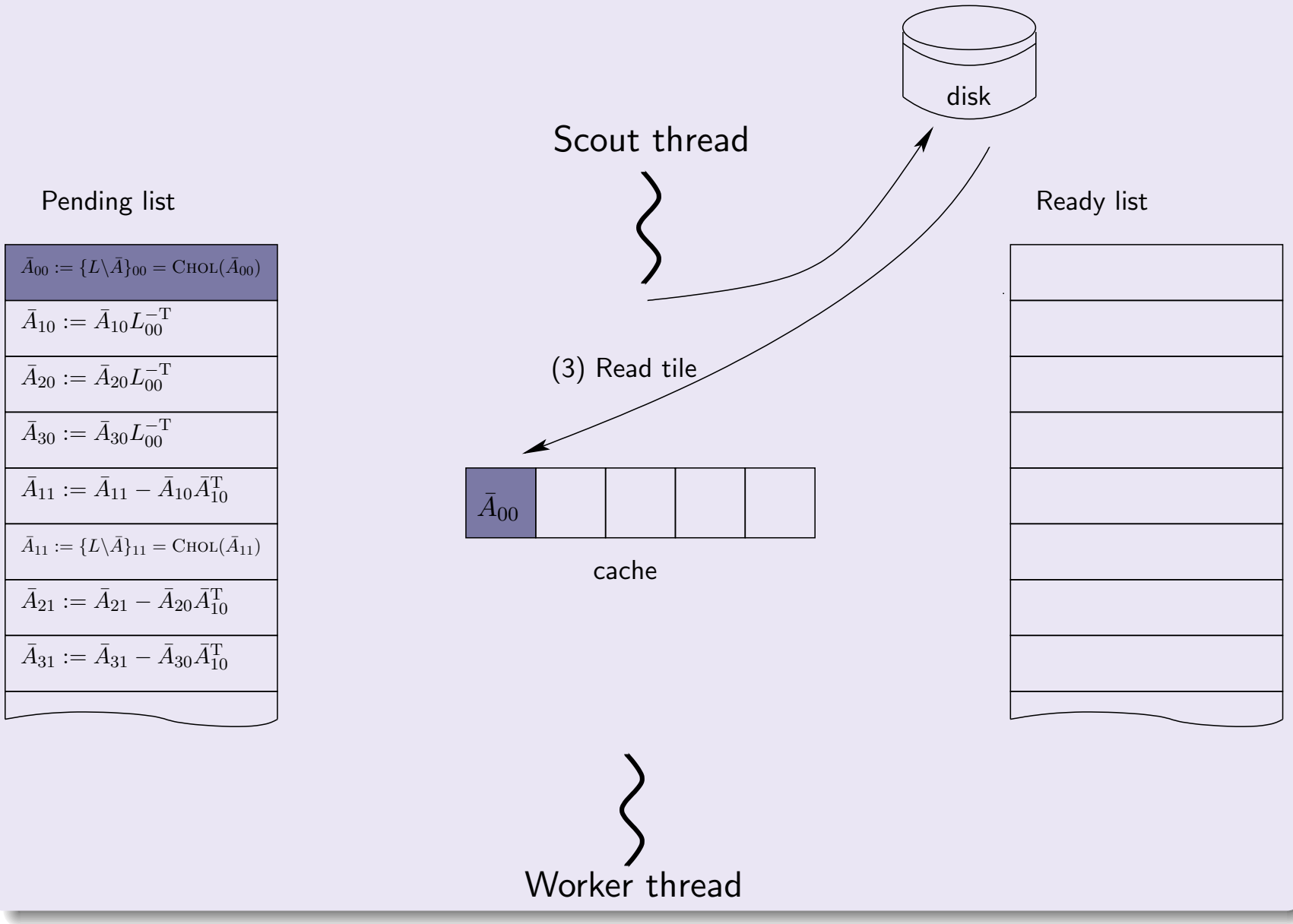
1. $\textsc{Chol}(\bar{A}_{0,0})$
2. $\bar{A}_{1,0} := \bar{A}_{1,0}\ L_{0,0}^{-T}$
3. $\bar{A}_{2,0} := \bar{A}_{2,0}\ L_{0,0}^{-T}$
4. $\vdots$
5. $\bar{A}_{1,1} := \bar{A}_{1,1} - \bar{A}_{1,0}\bar{A}_{1,0}^{T}$
6. $\vdots$

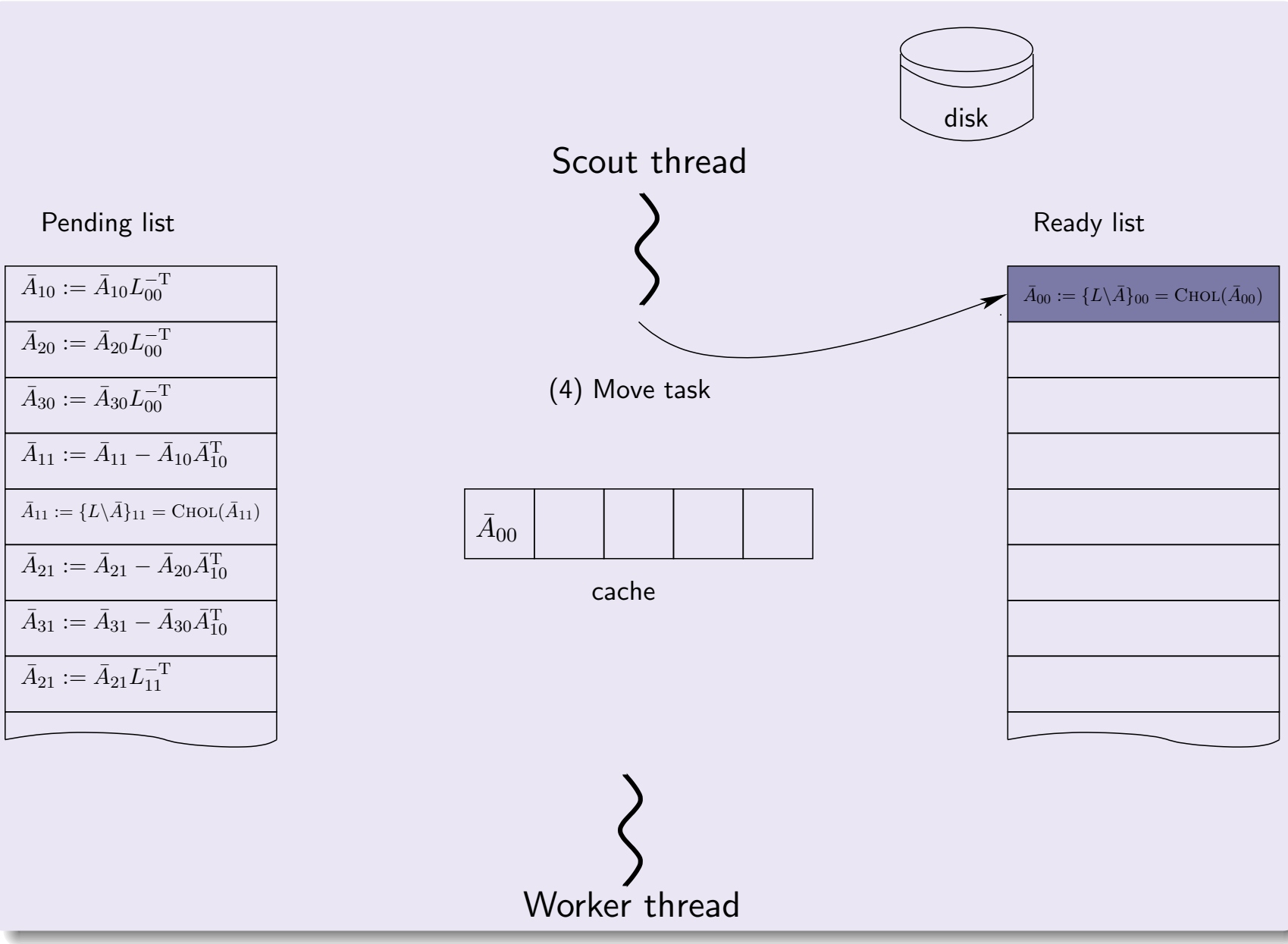
Second stage: runtime dictates the execution

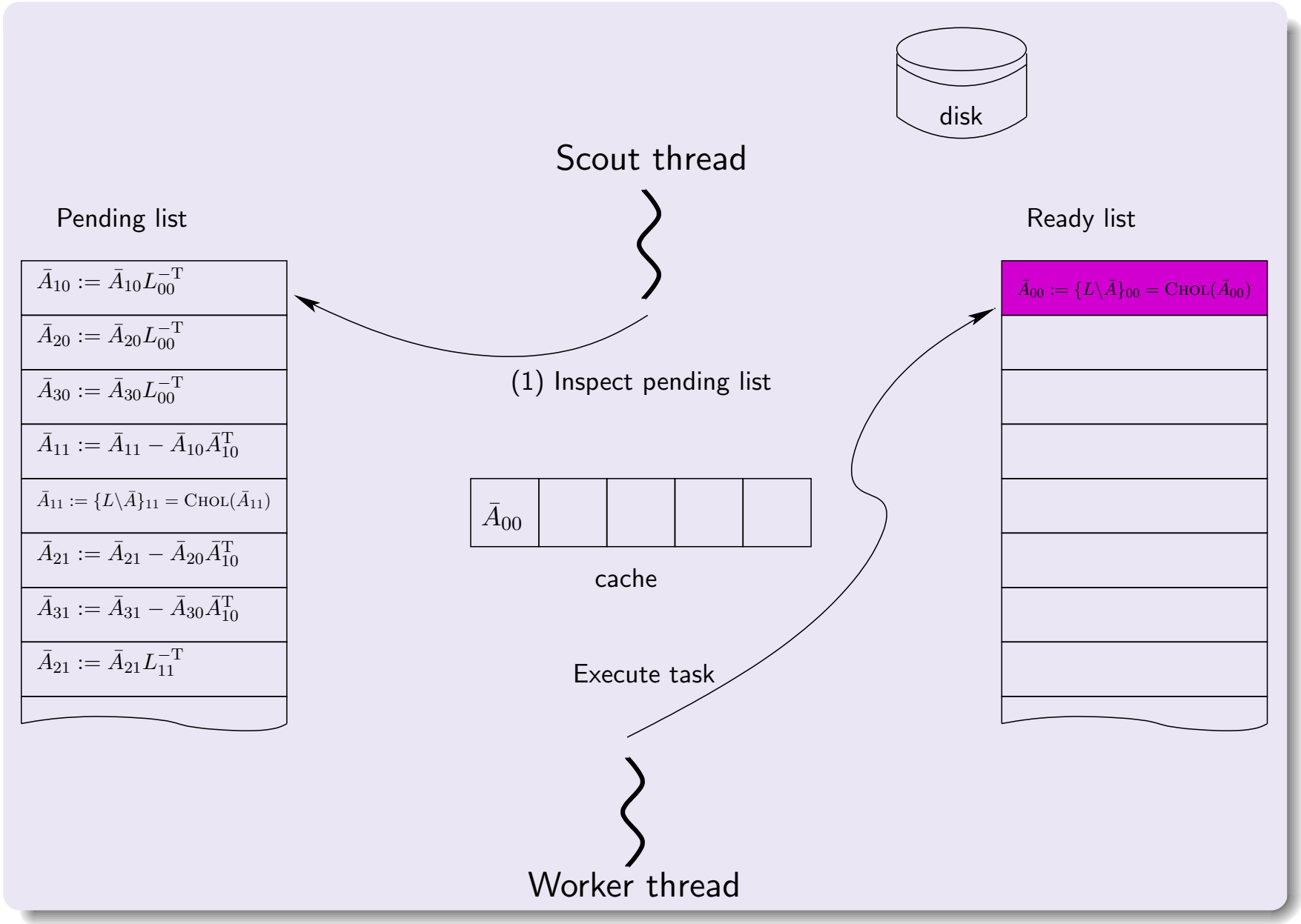
1. RAM is a software cache for data on disk
2. scout and worker threads deliver transparent asynchronous I/O

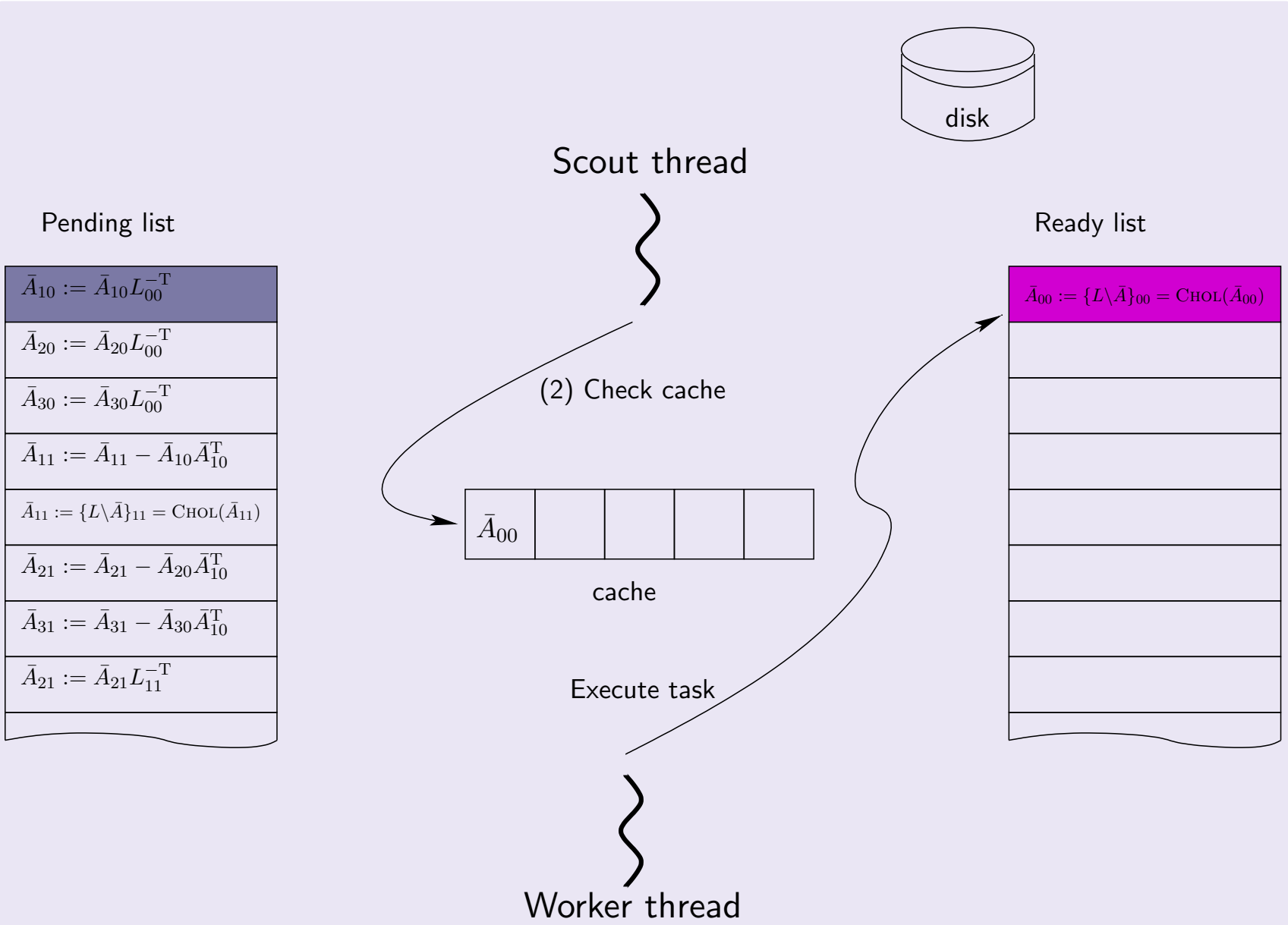

disk
Scout thread
Pending list
Ready list
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \text{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\text{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\text{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\text{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\text{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \text{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\text{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\text{T}}$
(1) Inspect pending list
cache
Worker thread

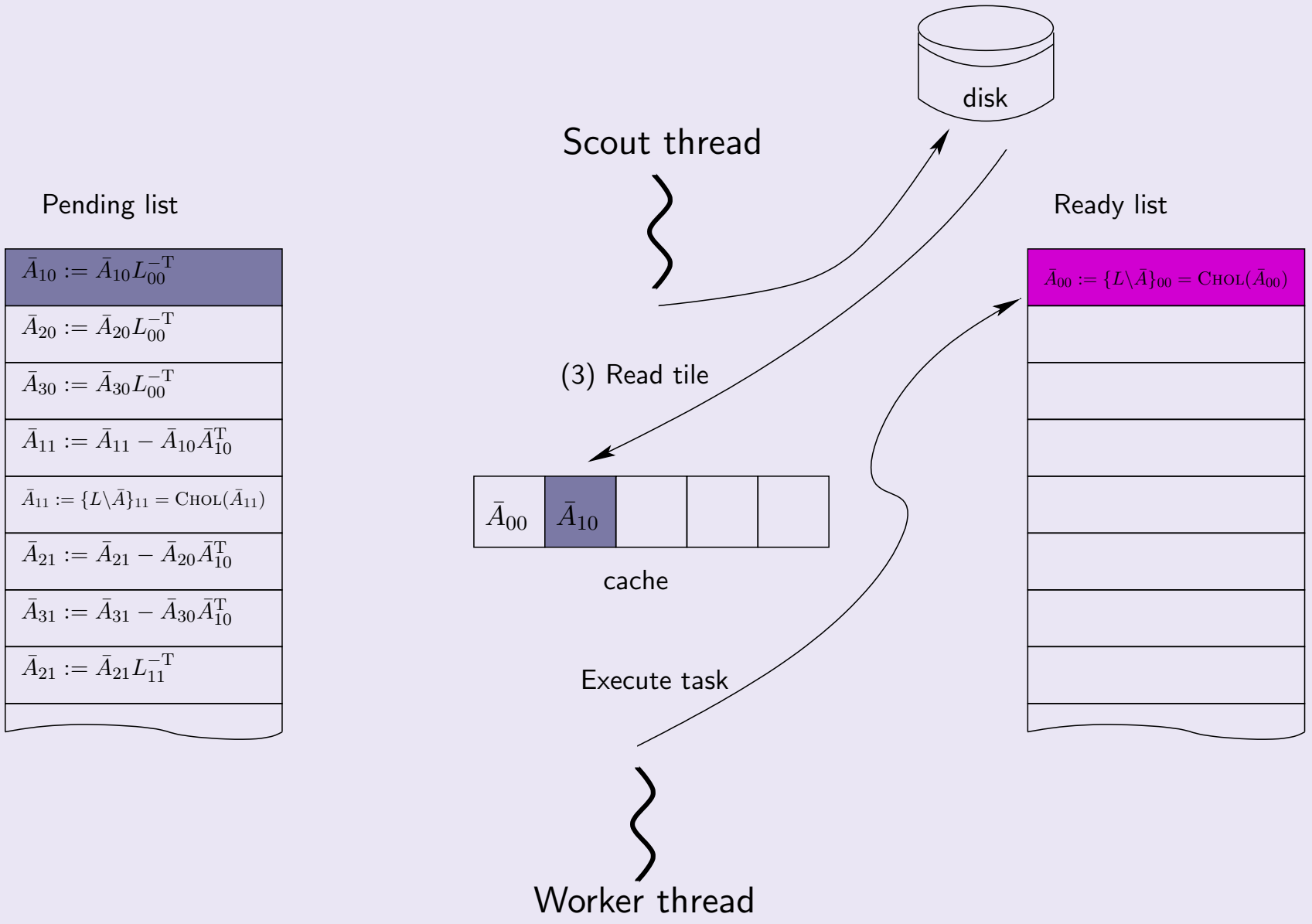

disk
Scout thread
Pending list
Ready list
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \text{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\text{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\text{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\text{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\text{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \text{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\text{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\text{T}}$
(2) Check cache
cache
Worker thread

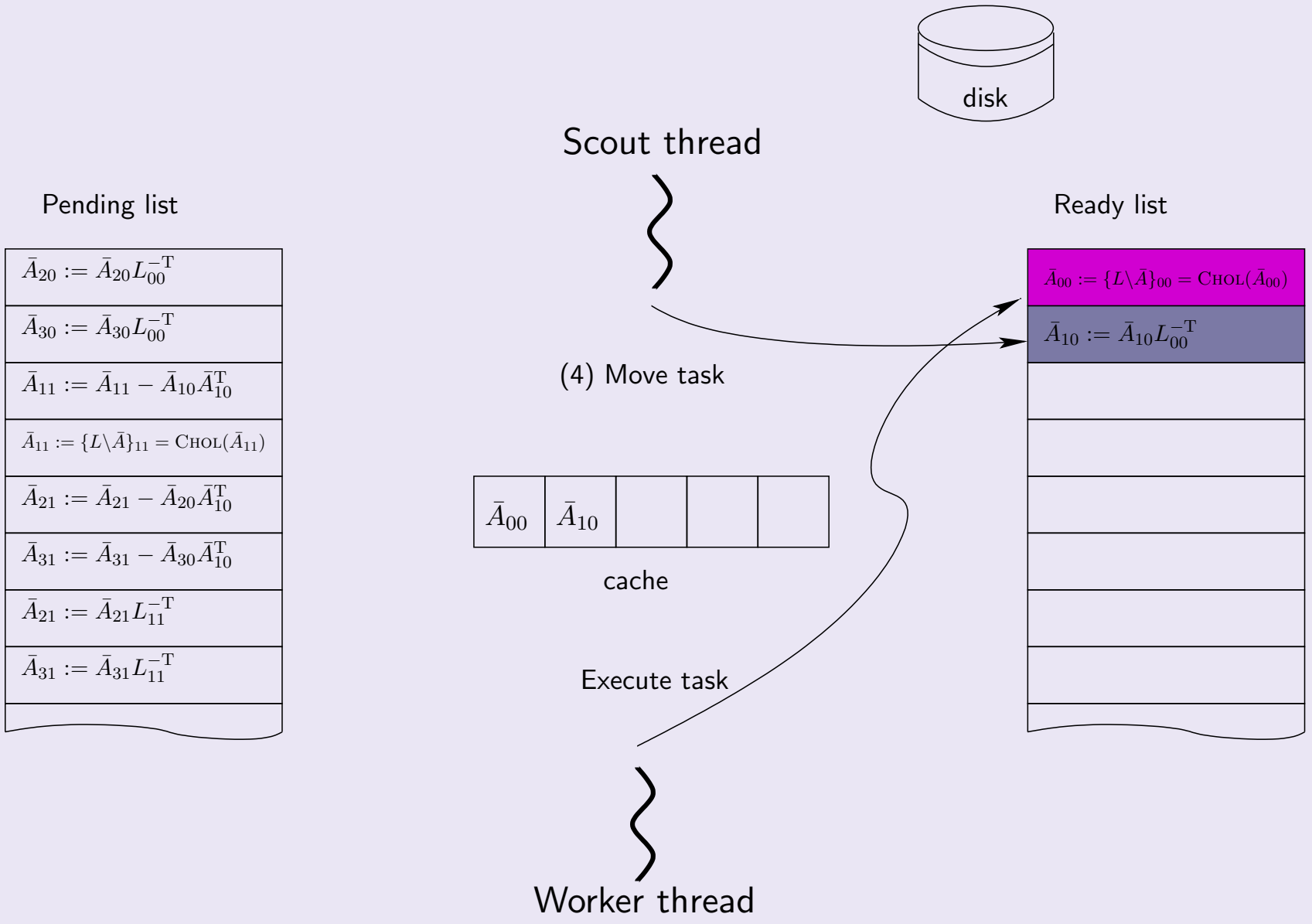

disk
Scout thread
Pending list
Ready list
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \text{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\text{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\text{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\text{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\text{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \text{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\text{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\text{T}}$
(3) Read tile
$\bar{A}_{00}$
cache
Worker thread

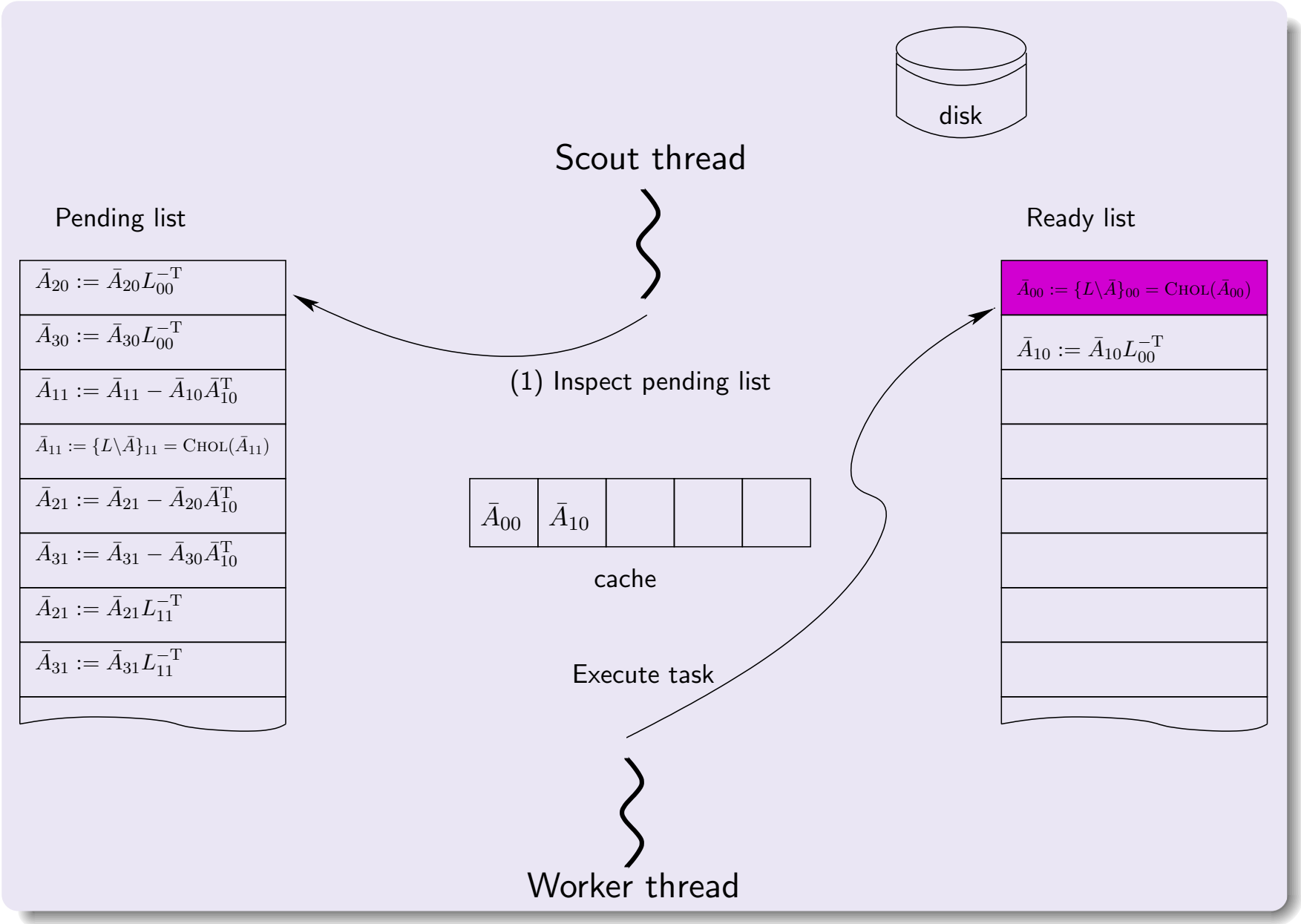

disk
Scout thread
Pending list
Ready list
$\bar{A}_{10} := \bar{A}_{10} L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20} L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30} L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L \backslash \bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21} L_{11}^{-\mathrm{T}}$
$\bar{A}_{00} := \{L \backslash \bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
(4) Move task
$\bar{A}_{00}$
cache
Worker thread

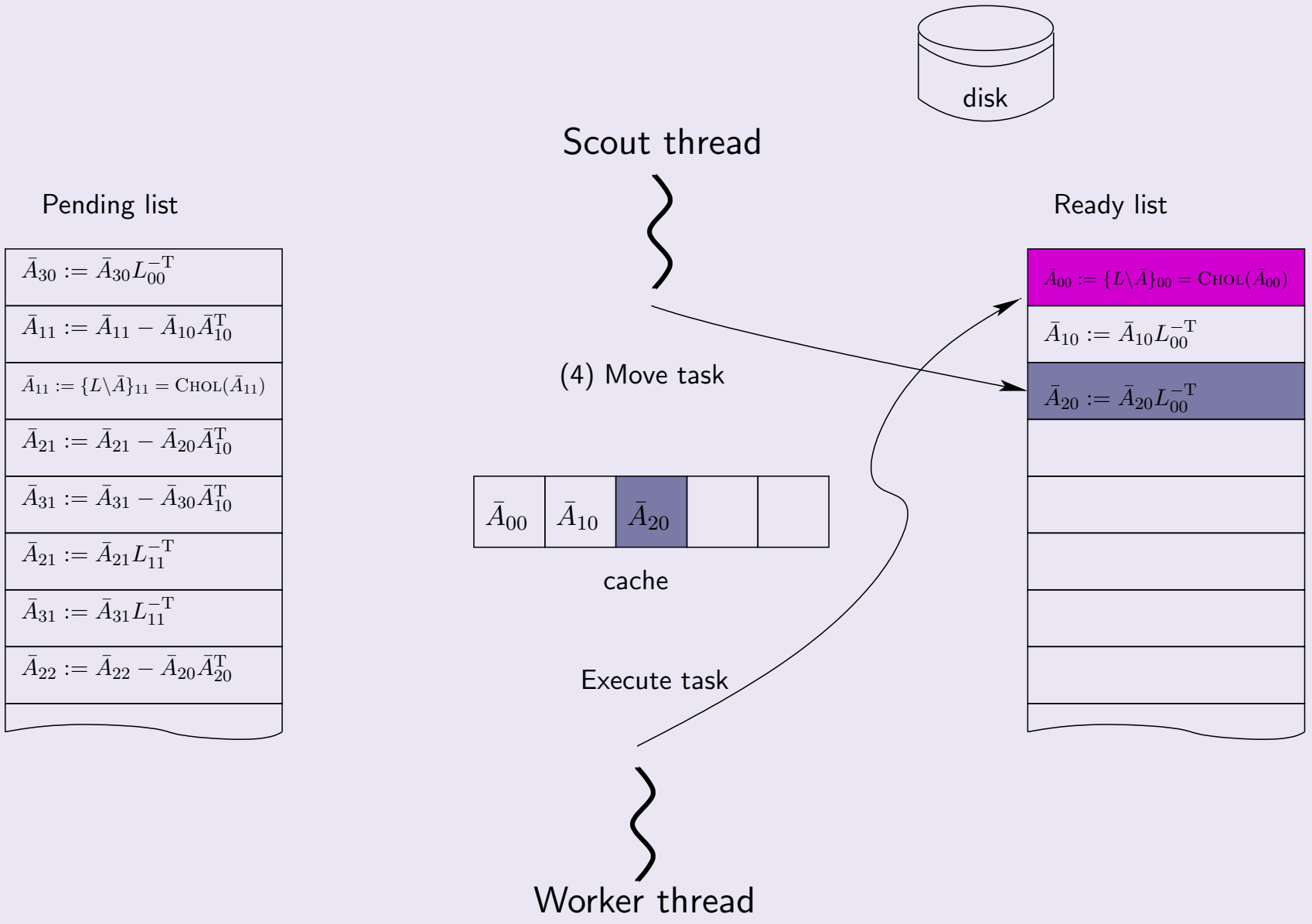

disk
Scout thread
Pending list
Ready list
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
(1) Inspect pending list
$\bar{A}_{00}$
cache
Execute task
Worker thread

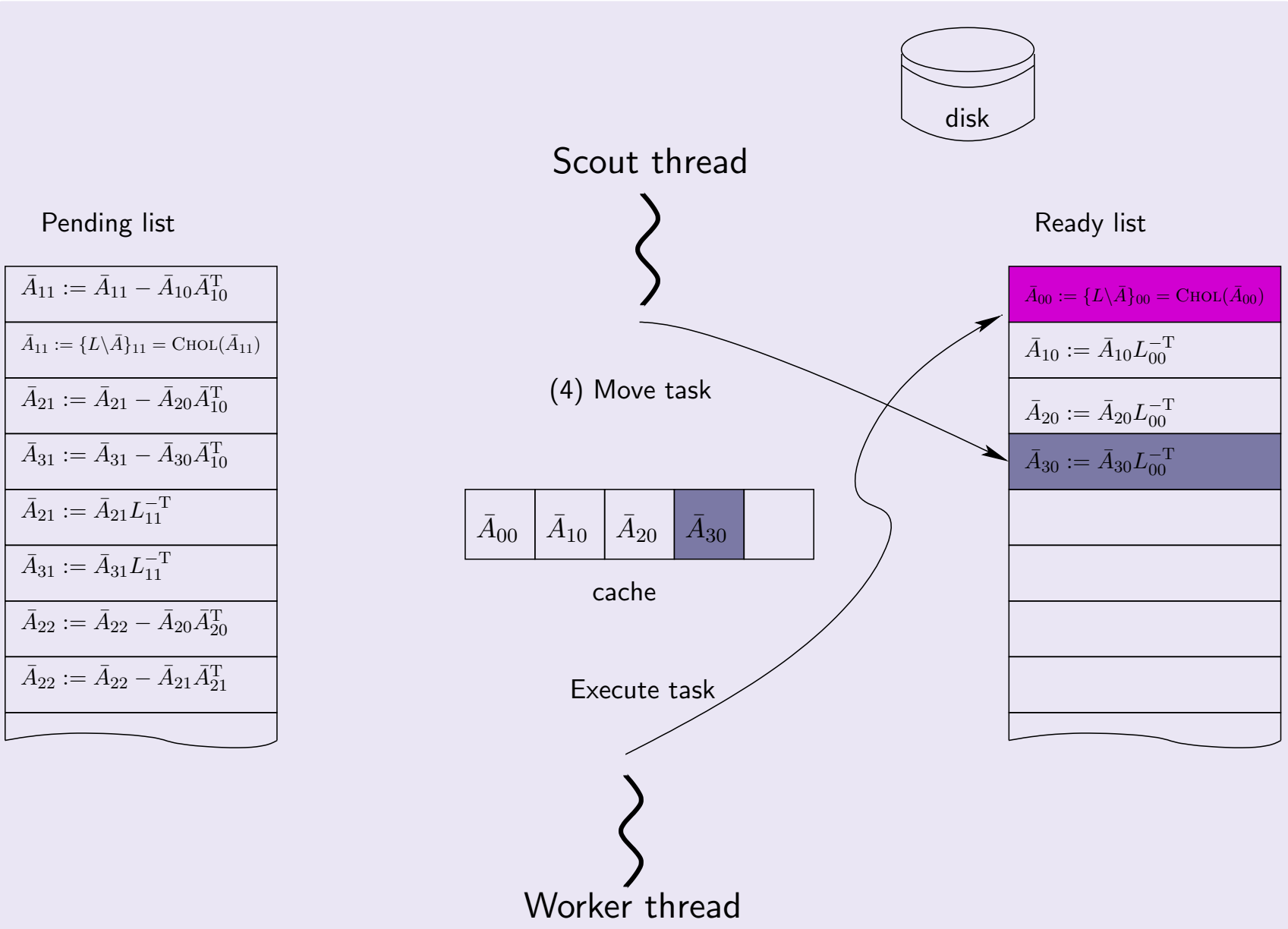

disk
Scout thread
Pending list
Ready list
$\bar{A}_{10} := \bar{A}_{10} L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20} L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30} L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L \backslash \bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21} L_{11}^{-\mathrm{T}}$
$\bar{A}_{00} := \{L \backslash \bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
(2) Check cache
$\bar{A}_{00}$
cache
Execute task
Worker thread

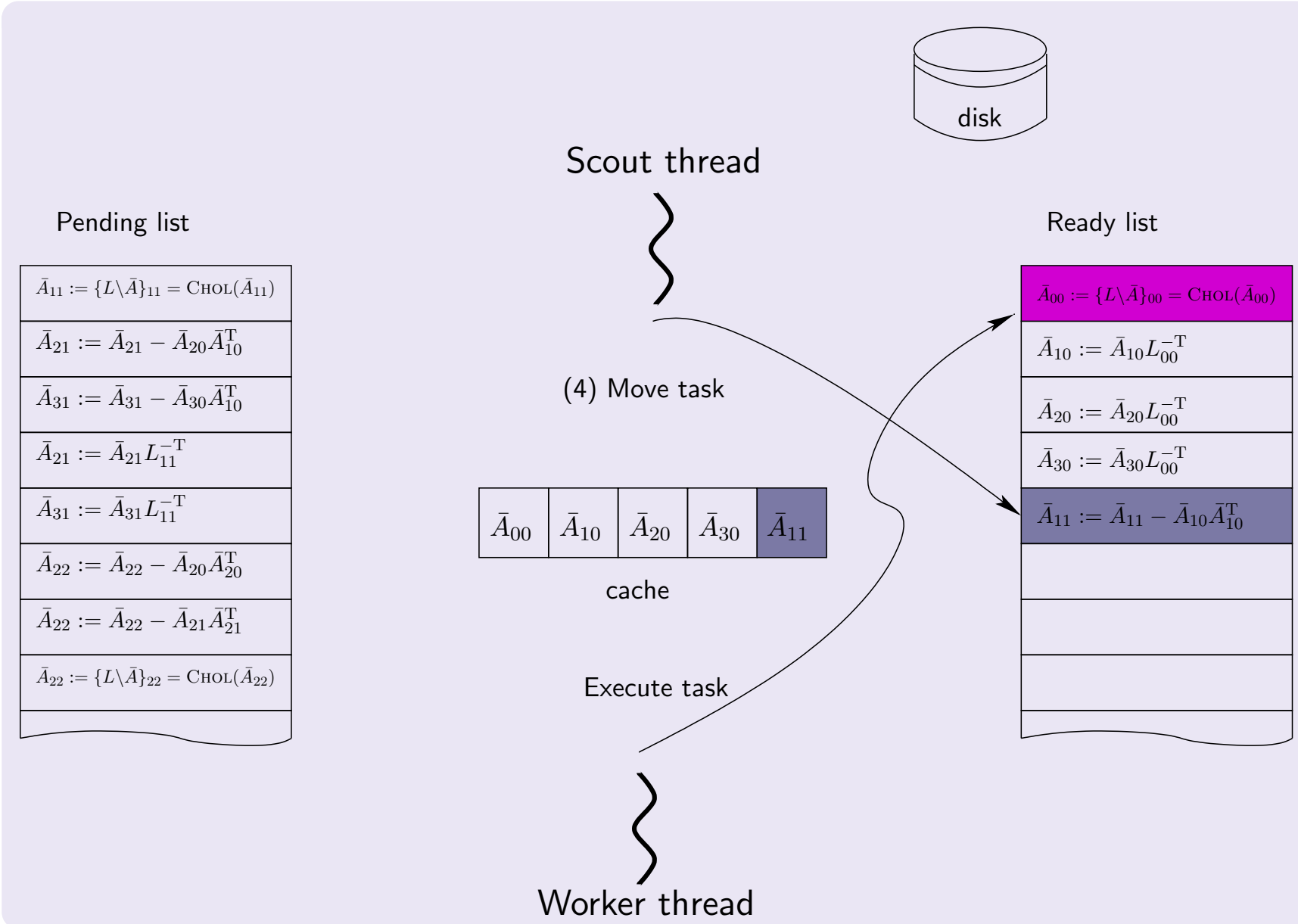

disk
Scout thread
Pending list
Ready list
$\bar{A}_{10} := \bar{A}_{10} L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20} L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30} L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L \backslash \bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30} \bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21} L_{11}^{-\mathrm{T}}$
$\bar{A}_{00} := \{L \backslash \bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
(3) Read tile
$\bar{A}_{00}$
$\bar{A}_{10}$
cache
Execute task
Worker thread

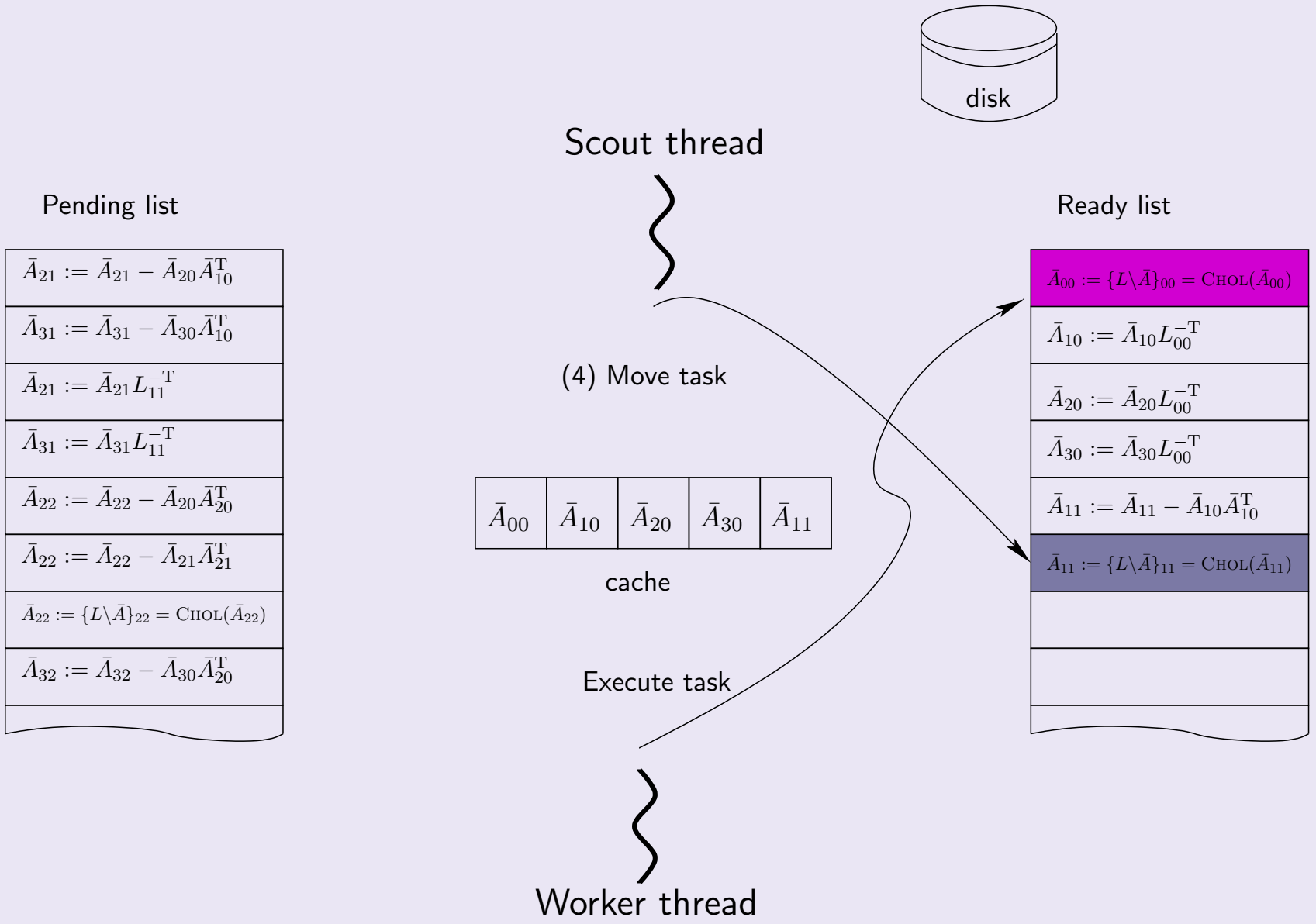

disk
Scout thread
Pending list
Ready list
$\bar{A}_{20} := \bar{A}_{20} L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30} L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L \backslash \bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21} L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} L_{11}^{-\mathrm{T}}$
$\bar{A}_{00} := \{L \backslash \bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10} L_{00}^{-\mathrm{T}}$
(4) Move task
$\bar{A}_{00}$
$\bar{A}_{10}$
cache
Execute task
Worker thread

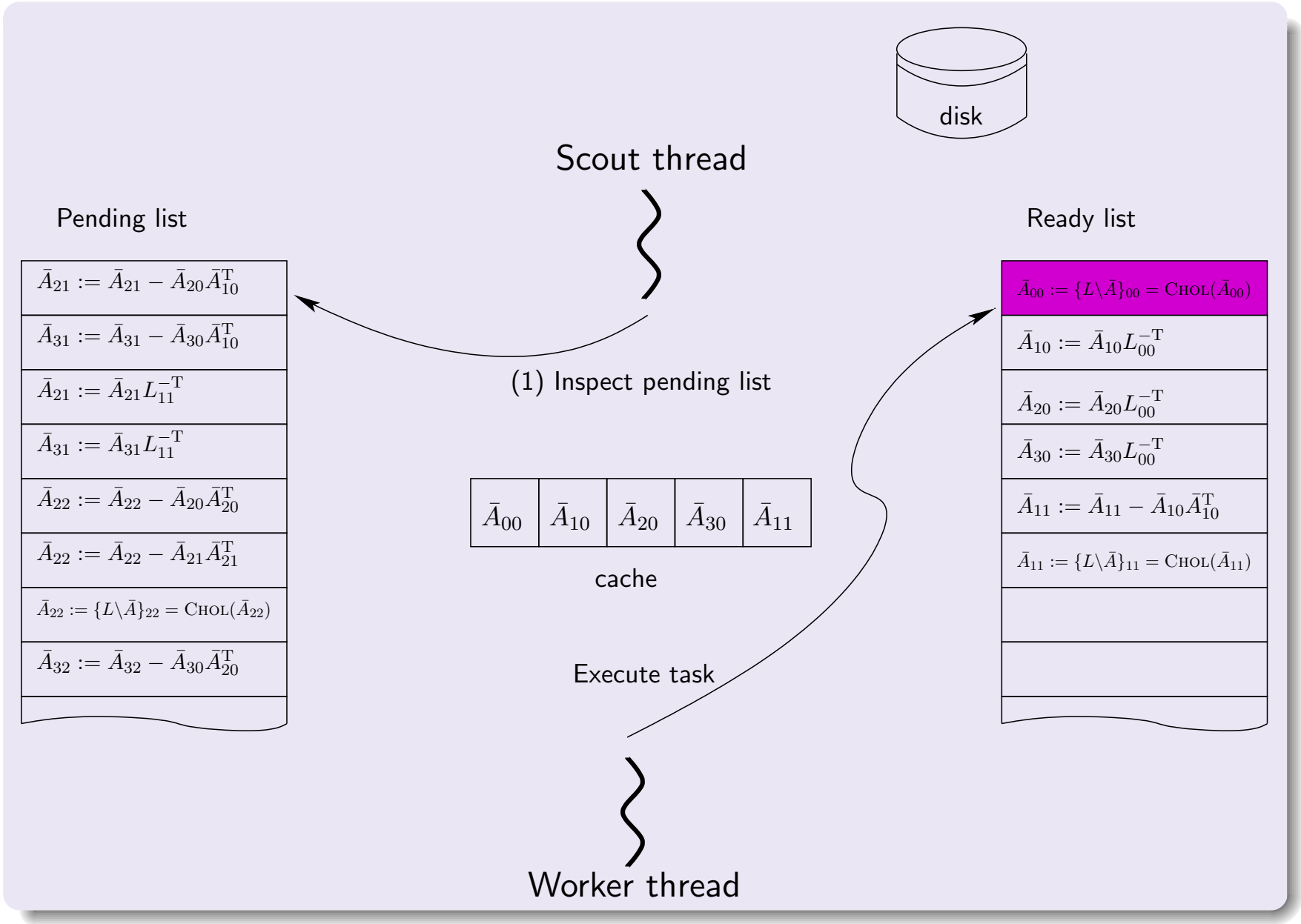

disk
Scout thread
Pending list
Ready list
$\bar{A}_{20} := \bar{A}_{20} L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30} L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21} L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} L_{11}^{-\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10} L_{00}^{-\mathrm{T}}$
(1) Inspect pending list
$\bar{A}_{00}$
$\bar{A}_{10}$
cache
Execute task
Worker thread

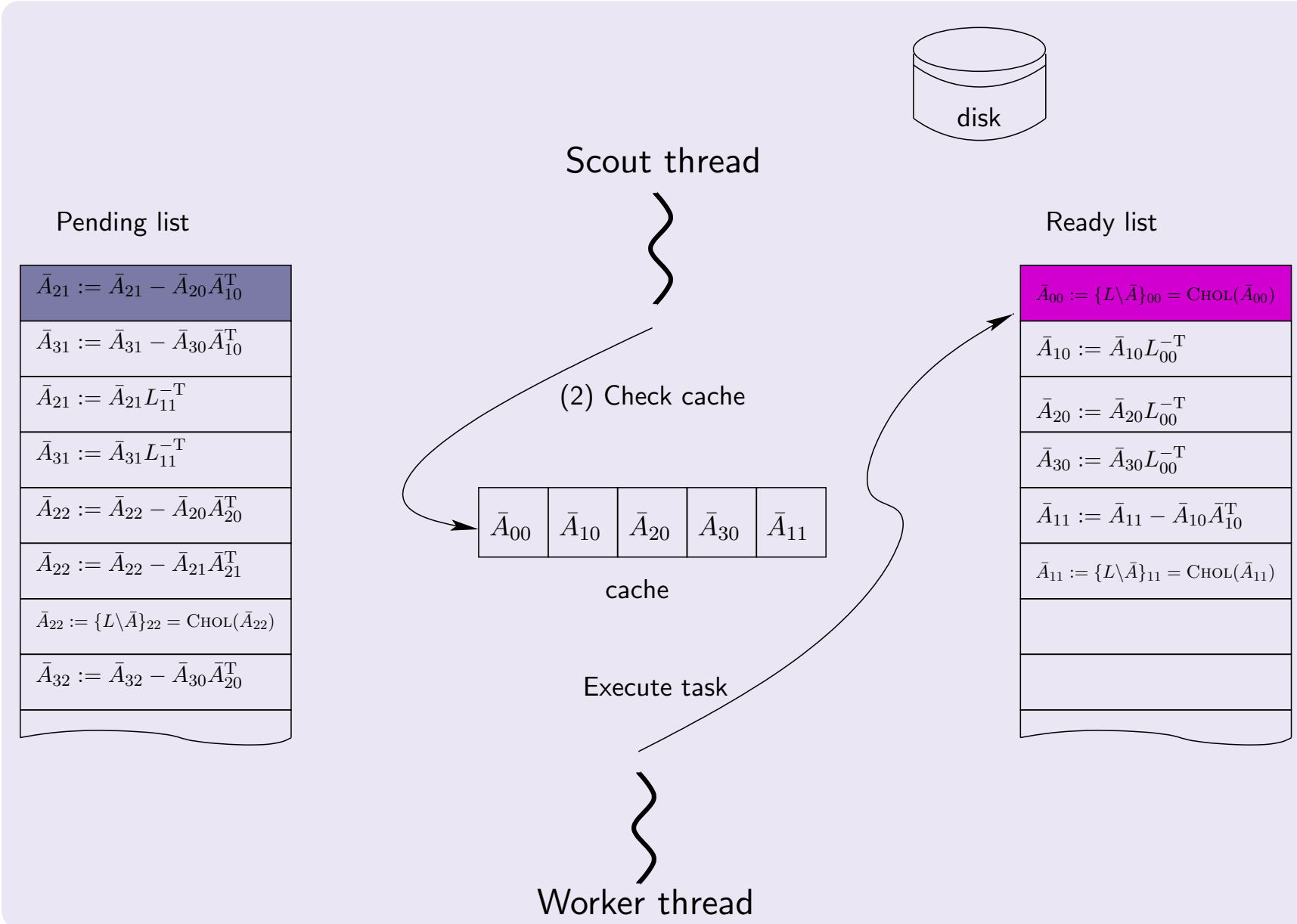

disk
Scout thread
Pending list
Ready list
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
(4) Move task
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$
cache
Execute task
Worker thread

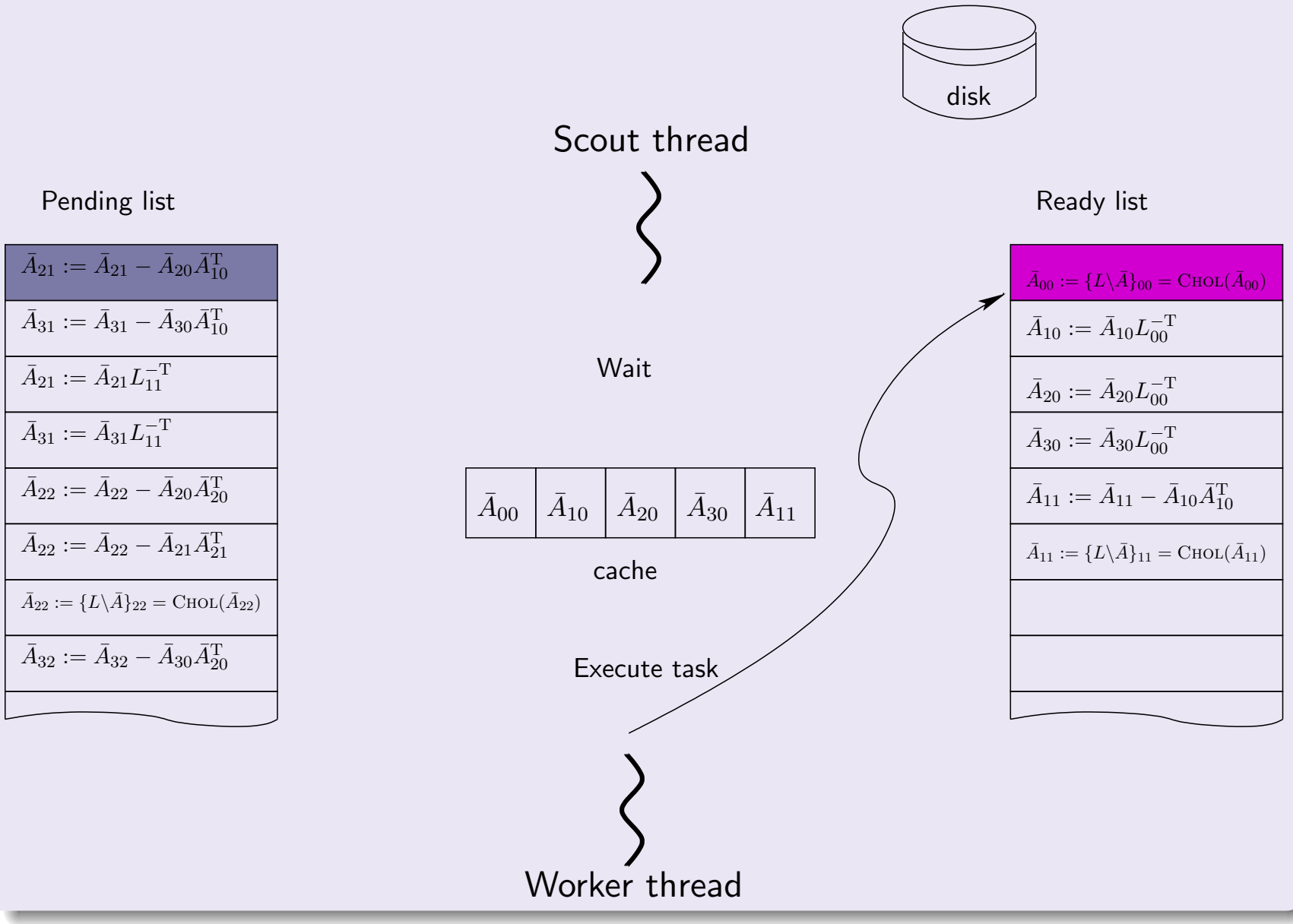

disk
Scout thread
Pending list
Ready list
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
(4) Move task
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$
cache
Execute task
Worker thread

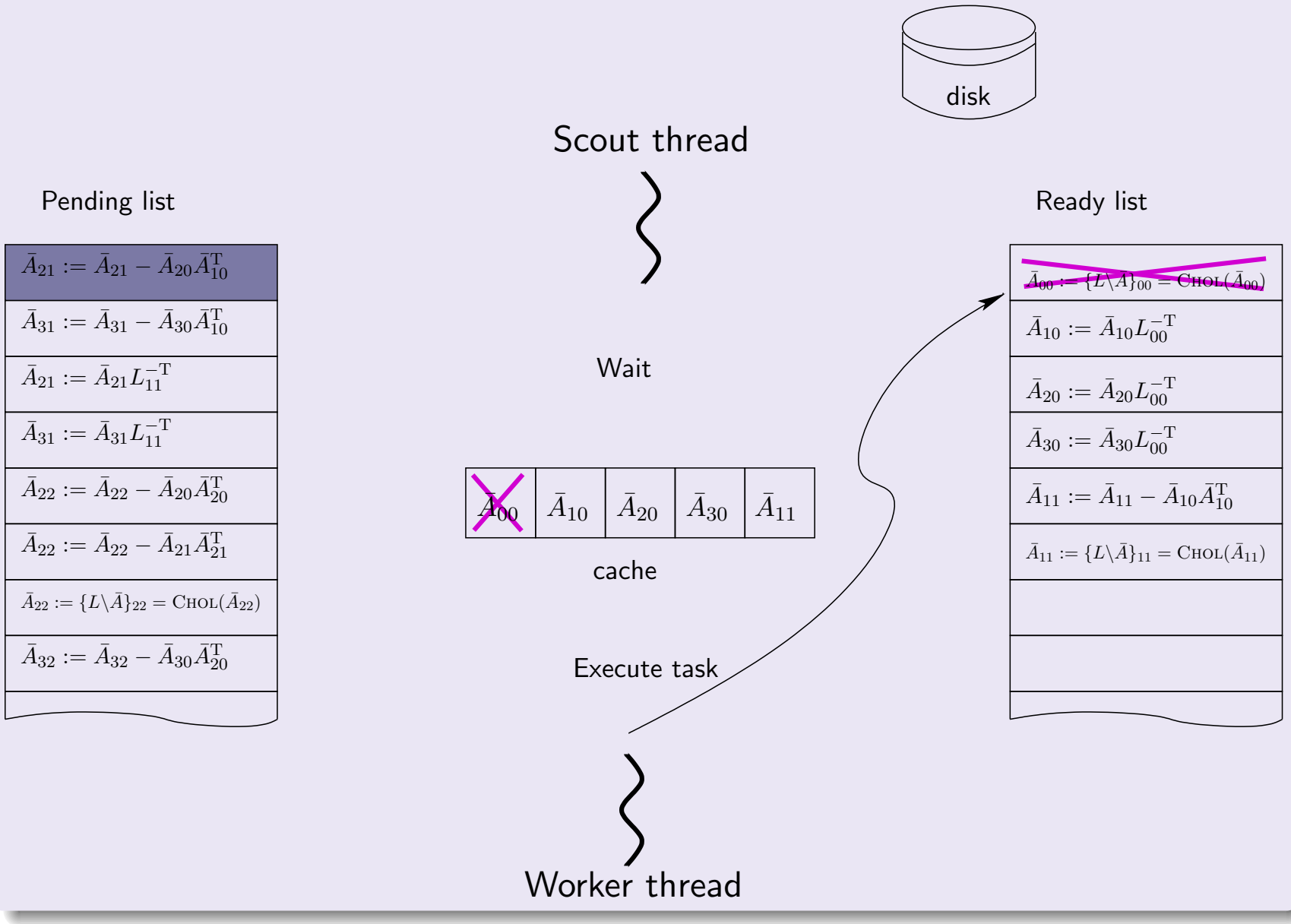

disk
Scout thread
Pending list
Ready list
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \text{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\text{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\text{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\text{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\text{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\text{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\text{T}}$
$\bar{A}_{22} := \{L\backslash\bar{A}\}_{22} = \text{CHOL}(\bar{A}_{22})$
(4) Move task
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$ $\bar{A}_{11}$
cache
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \text{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\text{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\text{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\text{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\text{T}}$
Execute task
Worker thread

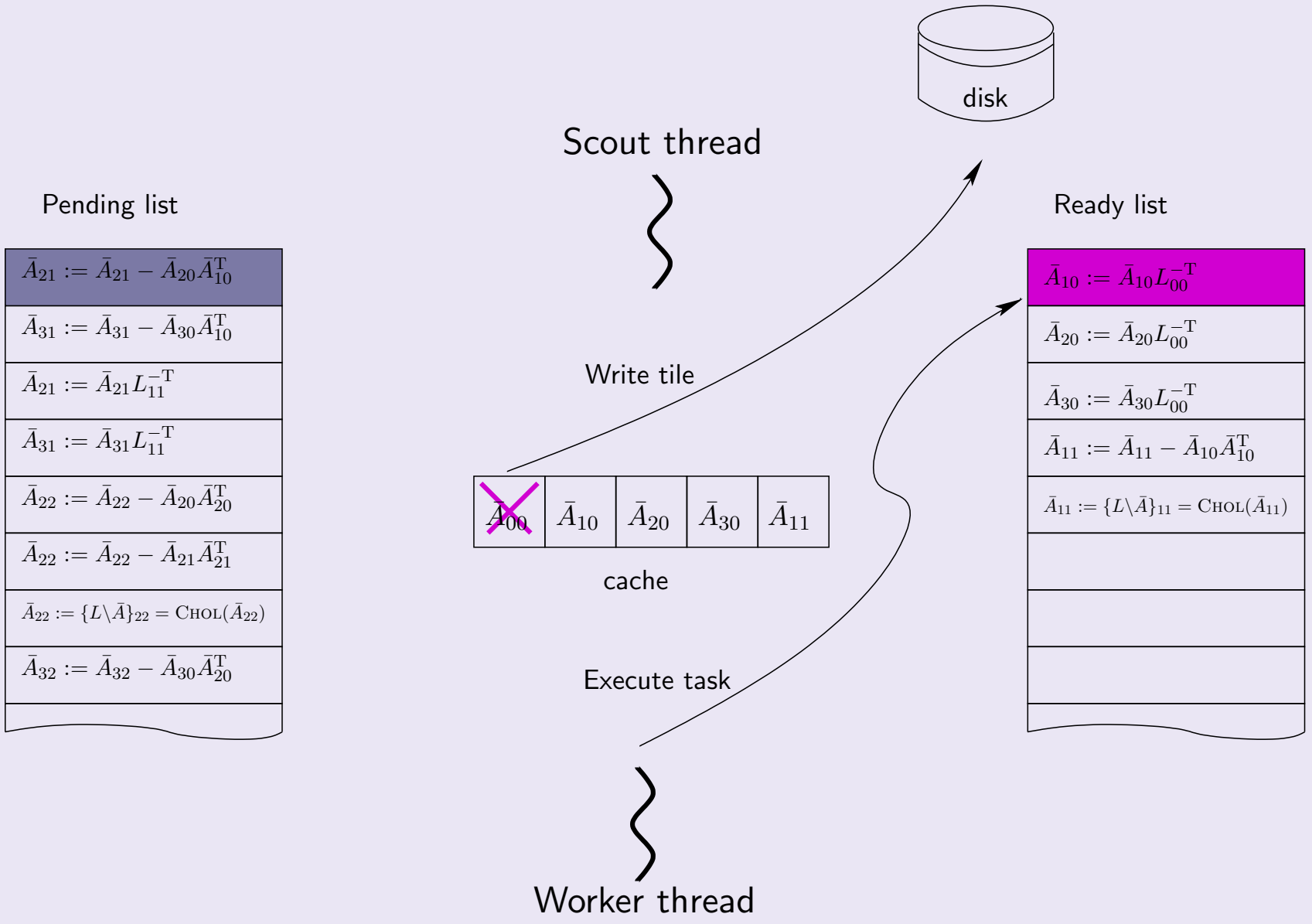

disk
Scout thread
Pending list
Ready list
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{22} := \{L\backslash\bar{A}\}_{22} = \mathrm{CHOL}(\bar{A}_{22})$
$\bar{A}_{32} := \bar{A}_{32} - \bar{A}_{30}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
(4) Move task
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$ $\bar{A}_{11}$
cache
Execute task
Worker thread

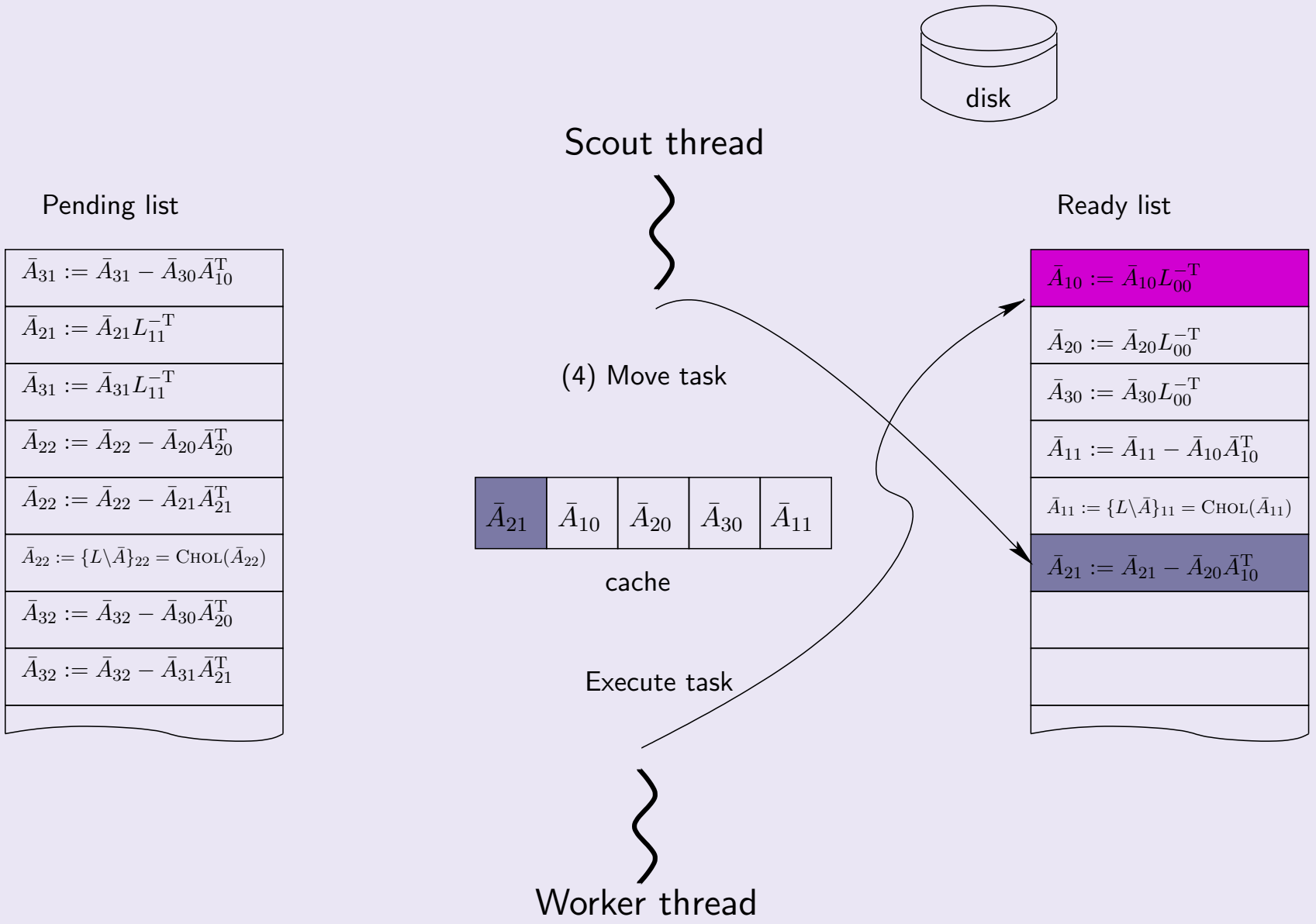

disk
Scout thread
Pending list
Ready list
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{22} := \{L\backslash\bar{A}\}_{22} = \mathrm{CHOL}(\bar{A}_{22})$
$\bar{A}_{32} := \bar{A}_{32} - \bar{A}_{30}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
(1) Inspect pending list
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$ $\bar{A}_{11}$
cache
Execute task
Worker thread


disk
Scout thread
Pending list
Ready list
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{22} := \{L\backslash\bar{A}\}_{22} = \mathrm{CHOL}(\bar{A}_{22})$
$\bar{A}_{32} := \bar{A}_{32} - \bar{A}_{30}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
(2) Check cache
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$ $\bar{A}_{11}$
cache
Execute task
Worker thread


disk
Scout thread
Pending list
Ready list
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{22} := \{L\backslash\bar{A}\}_{22} = \mathrm{CHOL}(\bar{A}_{22})$
$\bar{A}_{32} := \bar{A}_{32} - \bar{A}_{30}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
Wait
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$ $\bar{A}_{11}$
cache
Execute task
Worker thread


disk
Scout thread
Pending list
Ready list
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{22} := \{L\backslash\bar{A}\}_{22} = \mathrm{CHOL}(\bar{A}_{22})$
$\bar{A}_{32} := \bar{A}_{32} - \bar{A}_{30}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{00} := \{L\backslash\bar{A}\}_{00} = \mathrm{CHOL}(\bar{A}_{00})$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
Wait
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$ $\bar{A}_{11}$
cache
Execute task
Worker thread


disk
Scout thread
Pending list
Ready list
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{22} := \{L\backslash\bar{A}\}_{22} = \mathrm{CHOL}(\bar{A}_{22})$
$\bar{A}_{32} := \bar{A}_{32} - \bar{A}_{30}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
Write tile
$\bar{A}_{00}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$ $\bar{A}_{11}$
cache
Execute task
Worker thread


disk
Scout thread
Pending list
Ready list
$\bar{A}_{31} := \bar{A}_{31} - \bar{A}_{30}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{21} := \bar{A}_{21}L_{11}^{-\mathrm{T}}$
$\bar{A}_{31} := \bar{A}_{31}L_{11}^{-\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{20}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{22} := \bar{A}_{22} - \bar{A}_{21}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{22} := \{L\backslash\bar{A}\}_{22} = \mathrm{CHOL}(\bar{A}_{22})$
$\bar{A}_{32} := \bar{A}_{32} - \bar{A}_{30}\bar{A}_{20}^{\mathrm{T}}$
$\bar{A}_{32} := \bar{A}_{32} - \bar{A}_{31}\bar{A}_{21}^{\mathrm{T}}$
$\bar{A}_{10} := \bar{A}_{10}L_{00}^{-\mathrm{T}}$
$\bar{A}_{20} := \bar{A}_{20}L_{00}^{-\mathrm{T}}$
$\bar{A}_{30} := \bar{A}_{30}L_{00}^{-\mathrm{T}}$
$\bar{A}_{11} := \bar{A}_{11} - \bar{A}_{10}\bar{A}_{10}^{\mathrm{T}}$
$\bar{A}_{11} := \{L\backslash\bar{A}\}_{11} = \mathrm{CHOL}(\bar{A}_{11})$
$\bar{A}_{21} := \bar{A}_{21} - \bar{A}_{20}\bar{A}_{10}^{\mathrm{T}}$
(4) Move task
$\bar{A}_{21}$ $\bar{A}_{10}$ $\bar{A}_{20}$ $\bar{A}_{30}$ $\bar{A}_{11}$
cache
Execute task
Worker thread

Outline

Parallelization of In-Core Task

Target: multi-core processor

$$\textsc{Chol}(\bar{A}_{k,k})$$

- Standard parallelization $\rightarrow$ multithreaded BLAS
- Advanced parallelization $\rightarrow$ data-flow algorithm

Outline

Porting `libflame`. Experimental results

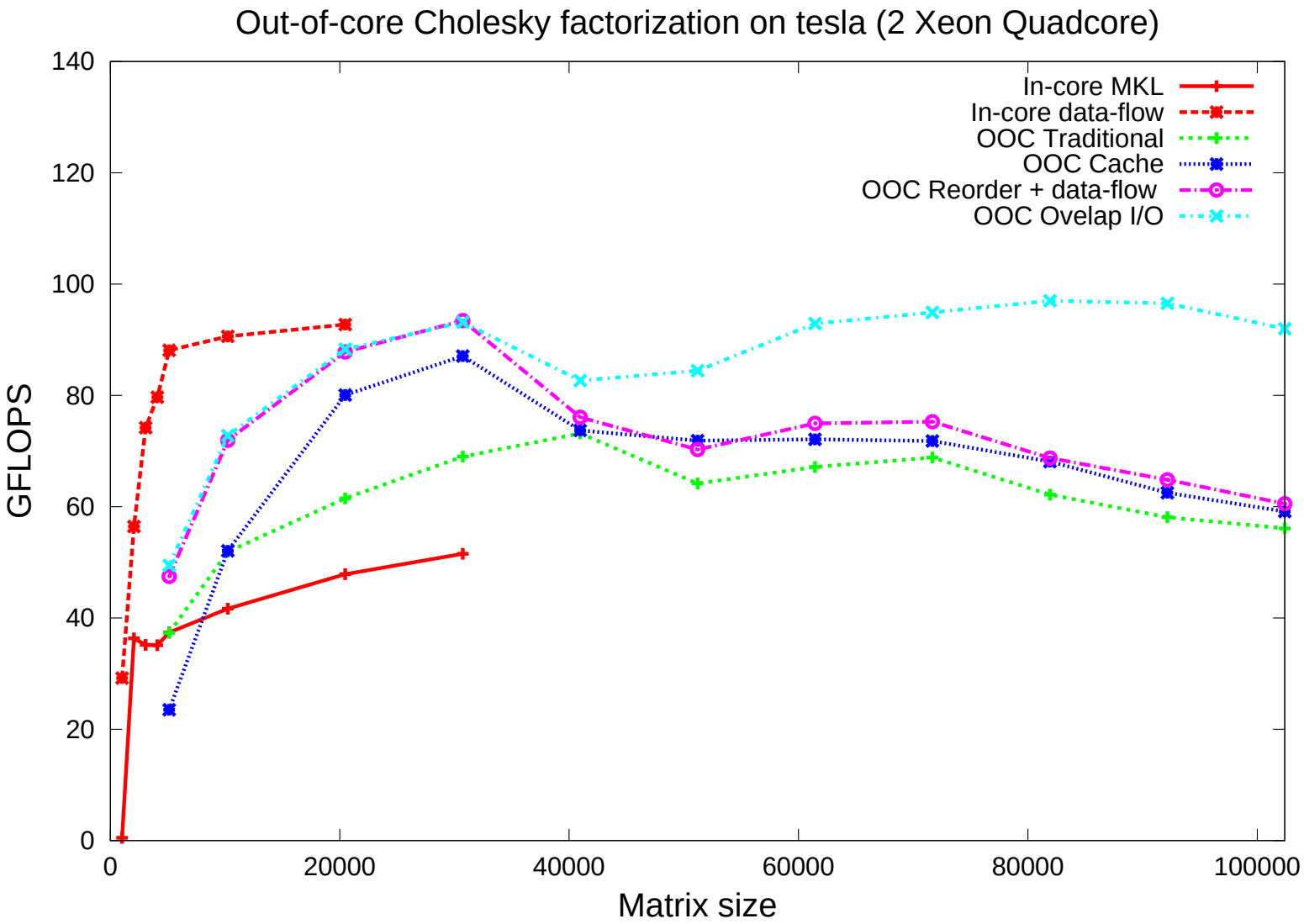
Hardware & Software

- Two Intel Xeon QuadCore E5405, 2.0 GHz, 8 GBytes RAM
- I/O interface of 1.5 Gbites/sec
- SATA-I disk with 160 GBytes
- MKL 10.1, single precision

Algorithms & Implementations

- In-core MKL
- In-core data-flow
- OOC Traditional
- OOC Cache
- OOC Reordered + data-flow
- OOC Overlap I/O

Porting `libflame`. Experimental results



Porting `libflame`. Experimental results

Matrix size	Time	MBytes of required RAM
10,240	4.9sec	400
51,200	8min 49.9sec	10,000
102,400	1h 4min 52.0sec	40,000

Conclusion

Against conventional wisdom

- Disk is fast enough to feed the processor
- Programming OOC is transparent to the library developer
  - Job for the runtime
  - Transparent port of the functionality of `libflame`
- Solving large problems possible using a few multi-core processors

For more information...

Visit `http://www.cs.utexas.edu/users/flame/`