# Optimal Jacobian Accumulation

## TTI/UChicago Open Problem Session

Andrew Lyons

Computation Institute, University of Chicago
(but I work at Argonne)

June 3, 2008

# Outline


(Structural) Optimal Jacobian Accumulation
- SOJA
- Complexity on restricted DAG classes

Generalized OJA
- Where does this problem come from?
- OJA
- Structural OJA
- Structural OJA with Unit Edges

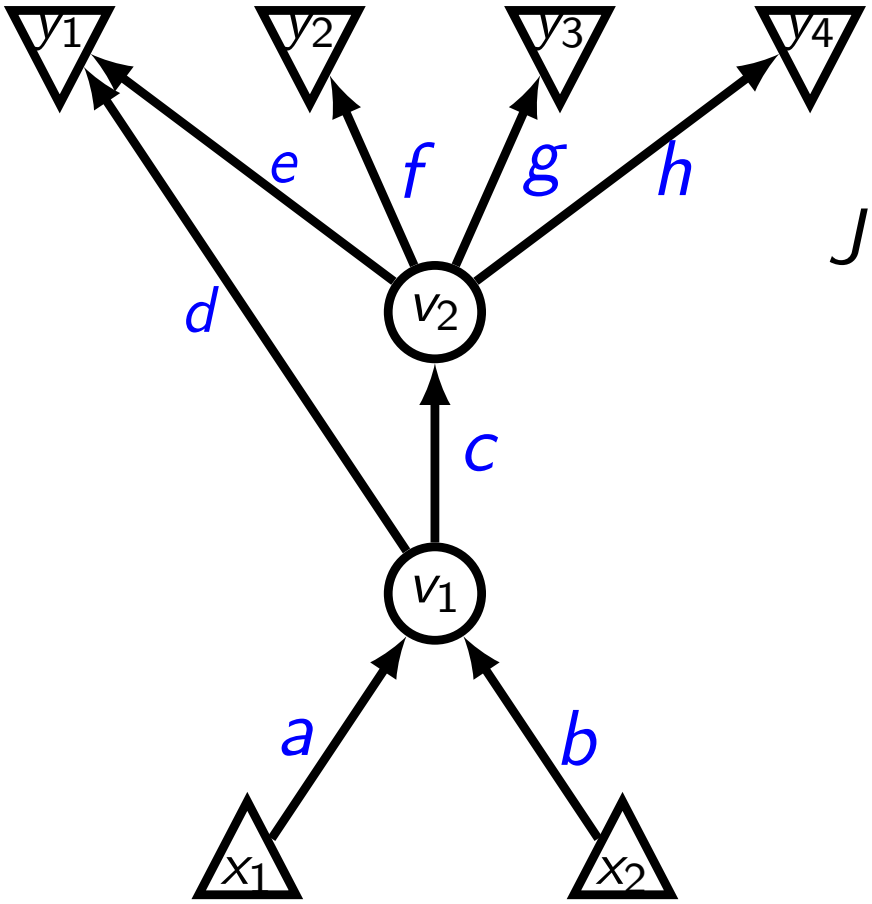
Multiplicative Complexity of Structural OJA with Unit Edges
- Biclique Partition
- Scalar, all paths length $\leq 3$
- Scalar, indegree $\leq 2$

Further Complexity Issues
- Minimizing Additions with Multiplications Fixed
- The Power of Subtraction
- When ISN'T OJA with unit edges harder?

Open Questions

# Baur's Formula

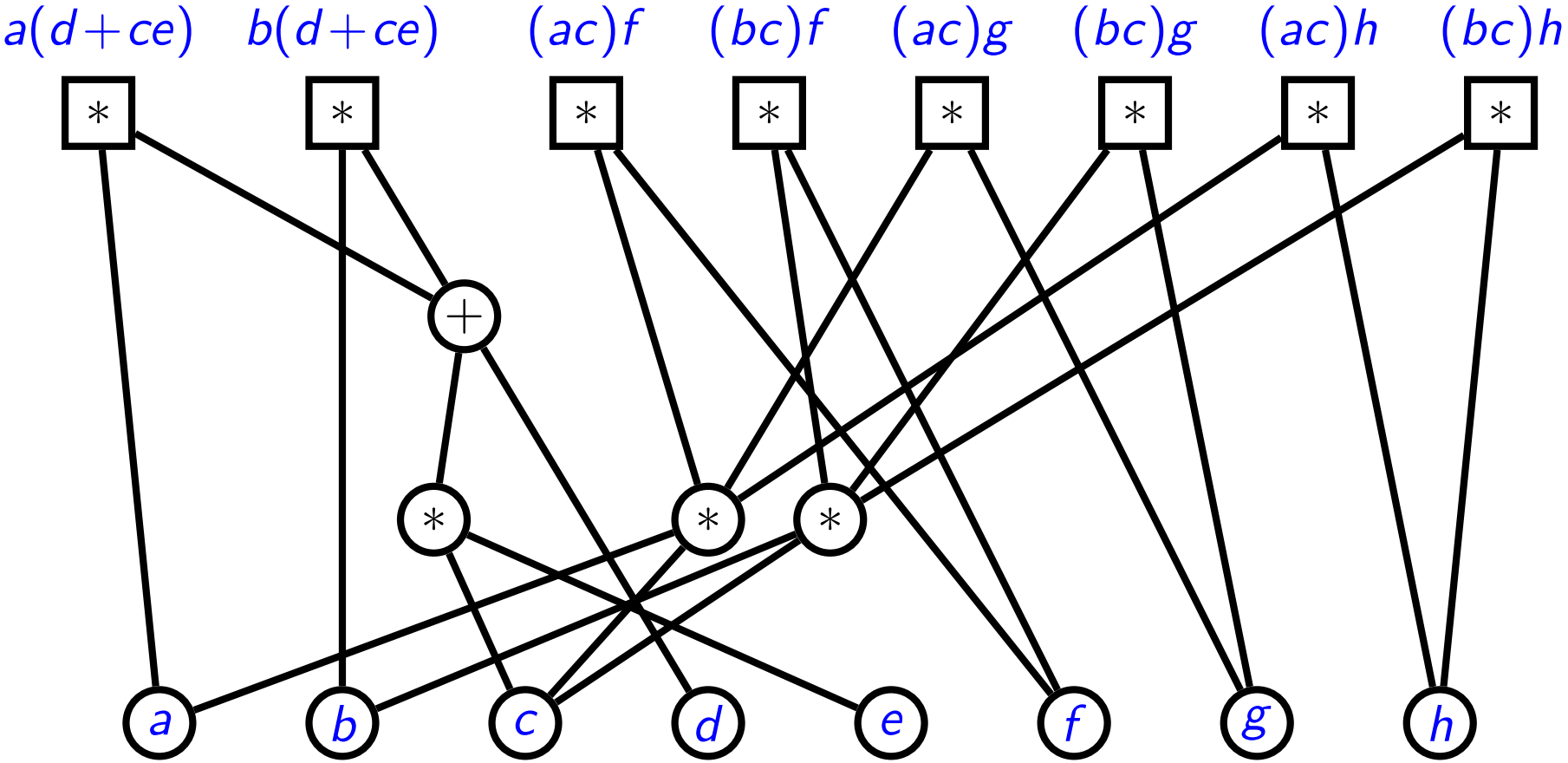
$$J_{j,i} \equiv \frac{\partial y_j}{\partial x_i} = \sum_{\mathcal{P}_{x_i}^{y_j} \in G} \prod_{e \in \mathcal{P}_{x_i}^{y_j}} e$$




$$J = \begin{bmatrix} ad + ace & bd + bce \\ acf & bcf \\ acg & bcg \\ ach & bch \end{bmatrix}$$

$$= \begin{bmatrix} a(d + ce) & b(d + ce) \\ (ac)f & (bc)f \\ (ac)g & (bc)g \\ (ac)h & (bc)h \end{bmatrix}$$

# Algebraic Circuit



# Straight-Line Program

$\mathtt{dv}_1$ = c*e

$\mathtt{dv}_2$ = d+$\mathtt{dv}_1$

$\mathtt{dv}_3$ = a*$\mathtt{dv}_2$ $= a(c+de)$

$\mathtt{dv}_4$ = b*$\mathtt{dv}_2$ $= b(c+de)$

$\mathtt{dv}_5$ = a*c

$\mathtt{dv}_6$ = b*c

$\mathtt{dv}_7$ = $\mathtt{dv}_5$*f $= (ac)f$

$\mathtt{dv}_8$ = $\mathtt{dv}_6$*f $= (bc)f$

$\mathtt{dv}_9$ = $\mathtt{dv}_5$*g $= (ac)g$

$\mathtt{dv}_{10}$ = $\mathtt{dv}_6$*g $= (bc)g$

$\mathtt{dv}_{11}$ = $\mathtt{dv}_5$*h $= (ac)h$

$\mathtt{dv}_{12}$ = $\mathtt{dv}_6$*h $= (bc)h$

# (Structural) Optimal Jacobian Accumulation

As a decision problem:

**Instance:** Dag $G = (V, E)$, where each $e \in E$ is labeled with a *unique* real variable $c_e$ such that all $c_e$ are algebraically independent, positive integer $K$.

**Question:** Is there a straight-line program (algebraic circuit) using operations in $\{+, *\}$ of length $K$ or less that computes every entry in $J$ such that every operand is some $c_e$ or the result of a previous operation?

# Some Open Problems

What is the complexity of minimizing multiplications or total operations?
(We suspect it's **NP**-complete; details if time permits)

# Some Open Problems

What is the complexity of minimizing multiplications or total operations?
(We suspect it's **NP**-complete; details if time permits)

Are divisions or subtractions useful?
(More on subtractions later, if time permits)

# Complexity on restricted DAG classes

Classes for which $\mathrm{SOJA}$ is known to be in **P**:

- Trees
- Single expression-use graphs
- Two-terminal series-parallel graphs

# Complexity on restricted DAG classes

Classes for which $\mathrm{SOJA}$ is known to be in **P**:

- Trees
- Single expression-use graphs
- Two-terminal series-parallel graphs

Open classes:

- Polytree (underlying undirected graph is a tree)
- Addition-free (at most one directed path between any two vertices)

# Complexity on restricted DAG classes

Classes for which $\mathrm{SOJA}$ is known to be in **P**:

- Trees
- Single expression-use graphs
- Two-terminal series-parallel graphs

Open classes:

- Polytree (underlying undirected graph is a tree)
- Addition-free (at most one directed path between any two vertices)
- All nodes have indegree $\leq 2$
- Two-terminal (scalar Jacobian)
- Two-terminal planar

# Outline


(Structural) Optimal Jacobian Accumulation
- SOJA
- Complexity on restricted DAG classes

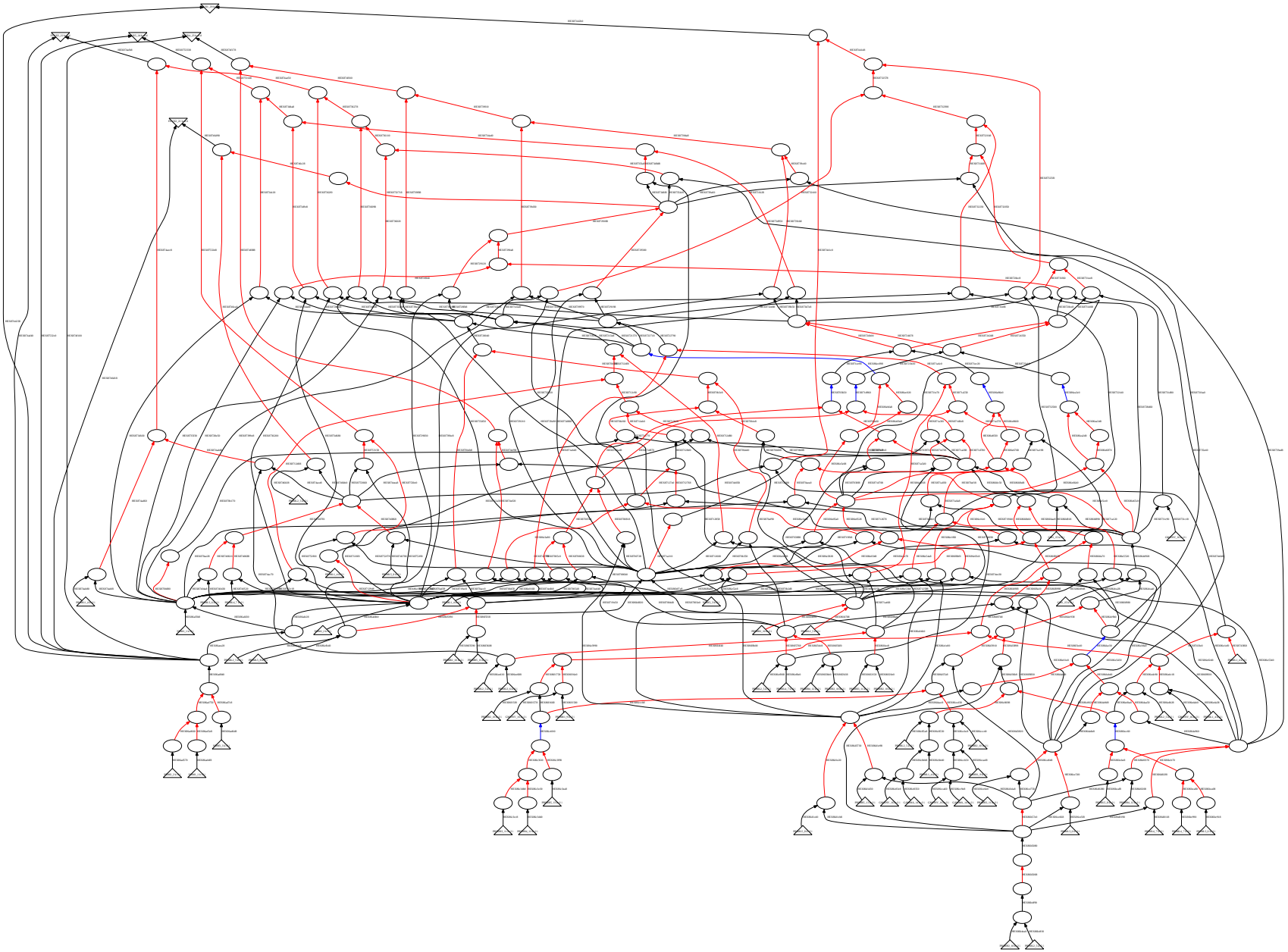
Generalized OJA
- Where does this problem come from?
- OJA
- Structural OJA
- Structural OJA with Unit Edges

Multiplicative Complexity of Structural OJA with Unit Edges
- Biclique Partition
- Scalar, all paths length $\leq 3$
- Scalar, indegree $\leq 2$

Further Complexity Issues
- Minimizing Additions with Multiplications Fixed
- The Power of Subtraction
- When ISN'T OJA with unit edges harder?

Open Questions

# Linearized Computational Graphs



# Complexity of OJA

**Theorem (Naumann, 2008)**

*Minimizing total operations for* OJA *is NP-complete.*

**Proof.**

Reduction from Ensemble Computation... □

# Ensemble Computation

**Instance**: Finite set $S$, collection $C = \{C_1, \ldots, C_r\}$ of subsets of $S$, positive integer $K$.

**Question**: Can the elements of $C$ be built up from the elements of $S$ using $K$ or fewer union operations?

# Ensemble Computation

**Instance**: Finite set $S$, collection $C = \{C_1, \ldots, C_r\}$ of subsets of $S$, positive integer $K$.

**Question**: Can the elements of $C$ be built up from the elements of $S$ using $K$ or fewer union operations?

Example: $S = \{a, b, c, d\}$, $C = \{\{a, b\}, \{b, c, d\}, \{a, c, d\}\}$, $K = 4$.

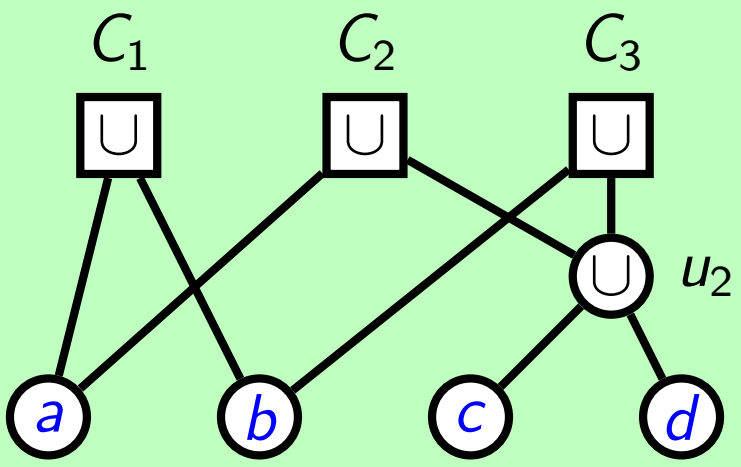
# Ensemble Computation

**Instance**: Finite set $S$, collection $C = \{C_1, \ldots, C_r\}$ of subsets of $S$, positive integer $K$.

**Question**: Can the elements of $C$ be built up from the elements of $S$ using $K$ or fewer union operations?

Example: $S = \{a, b, c, d\}$, $C = \{\{a, b\}, \{b, c, d\}, \{a, c, d\}\}$, $K = 4$.

$$
\begin{aligned}
u_1 &= \{a\} \cup \{b\} &&= C_1 \\
u_2 &= \{c\} \cup \{d\} \\
u_3 &= \{b\} \cup u_2 &&= C_2 \\
u_4 &= \{a\} \cup u_2 &&= C_3
\end{aligned}
$$



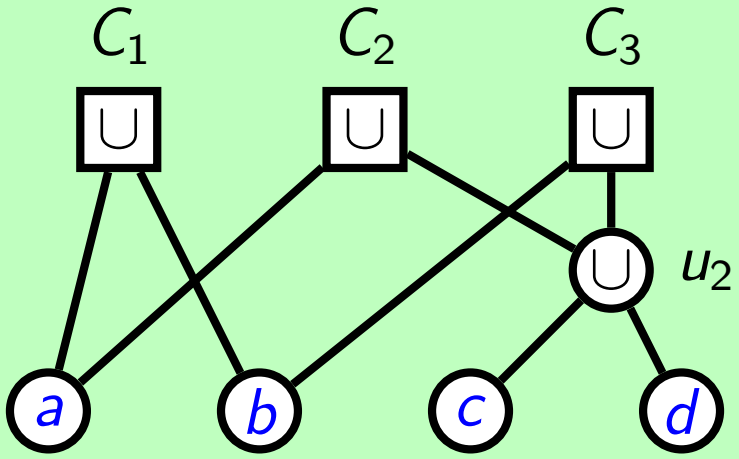
# Ensemble Computation

**Instance**: Finite set $S$, collection $C = \{C_1, \ldots, C_r\}$ of subsets of $S$, positive integer $K$.

**Question**: Can the elements of $C$ be built up from the elements of $S$ using $K$ or fewer union operations?

Example: $S = \{a, b, c, d\}$, $C = \{\{a, b\}, \{b, c, d\}, \{a, c, d\}\}$, $K = 4$.

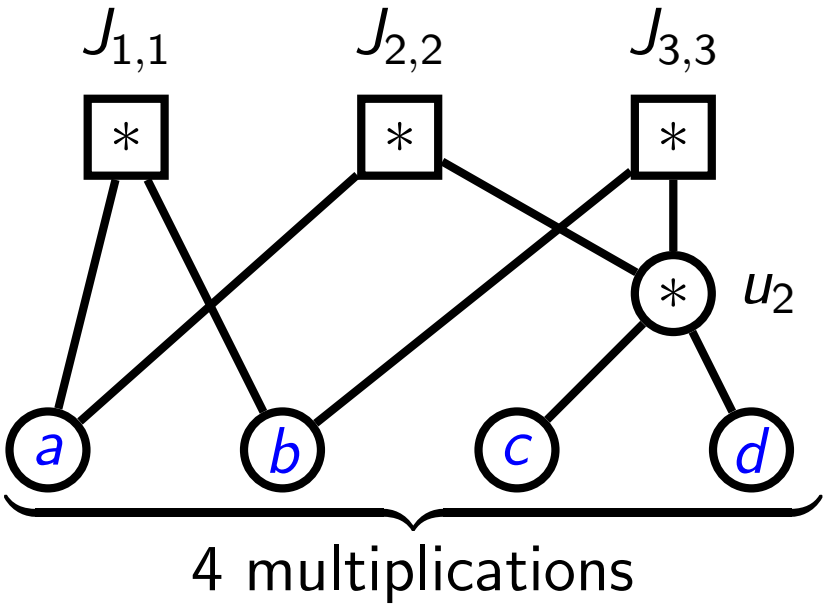
$$
\begin{aligned}
u_1 &= \{a\} \cup \{b\} &&= C_1 \\
u_2 &= \{c\} \cup \{d\} \\
u_3 &= \{b\} \cup u_2 &&= C_2 \\
u_4 &= \{a\} \cup u_2 &&= C_3
\end{aligned}
$$




**Answer: YES**

# Reducing Ensemble Computation to OJA

$$J = \begin{bmatrix} J_{1,1} & 0 & 0 \\ 0 & J_{2,2} & 0 \\ 0 & 0 & J_{3,3} \end{bmatrix}$$

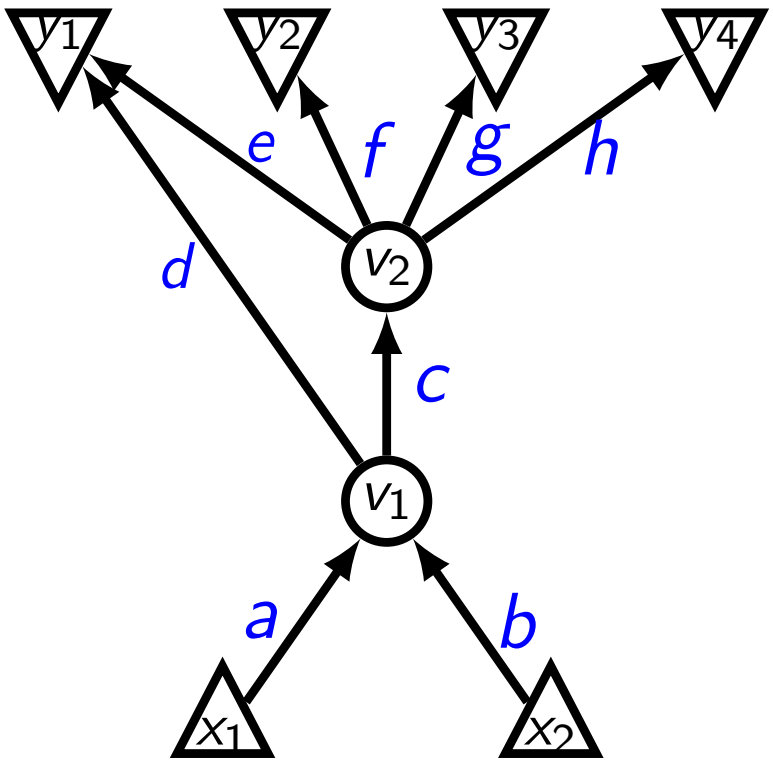


# Reducing Ensemble Computation to OJA

### Corollary

*Minimizing total ops and minimizing multiplications are NP-complete for* OJA *even when $J$ is scalar (single input and single output) and all paths in $G$ have length $\leq 3$.*

# Structural OJA



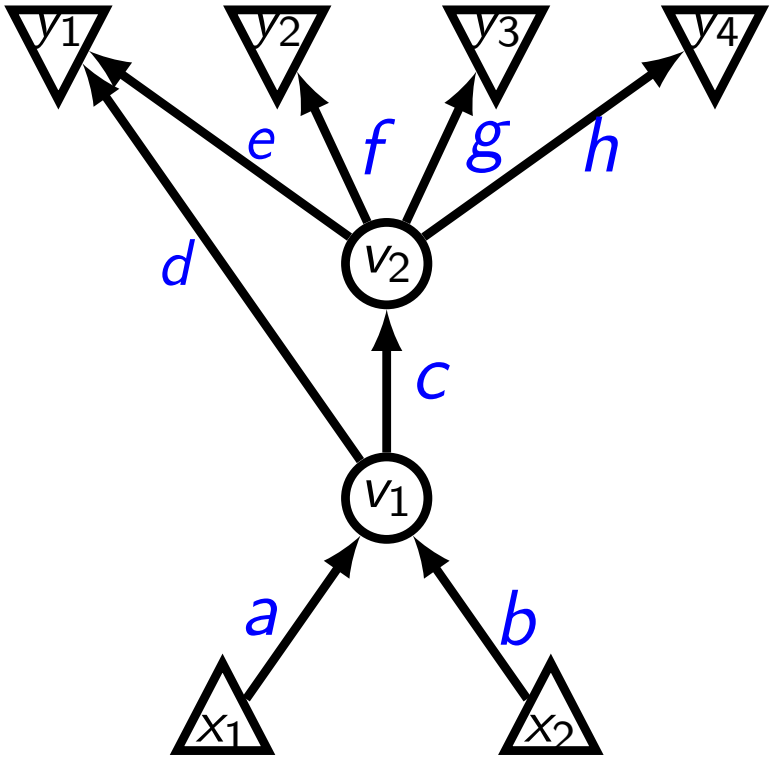

$$J_{j,i} \equiv \frac{\partial y_j}{\partial x_i} = \sum_{\mathcal{P}_{x_i}^{y_j} \in G} \prod_{e \in \mathcal{P}_{x_i}^{y_j}} e$$

$$J = \begin{bmatrix} ad + ace & bd + bce \\ acf & bcf \\ acg & bcg \\ ach & bch \end{bmatrix}$$

$$= \begin{bmatrix} a(d + ce) & b(d + ce) \\ (ac)f & (bc)f \\ (ac)g & (bc)g \\ (ac)h & (bc)h \end{bmatrix}$$

# Structural OJA

$$J_{j,i} \equiv \frac{\partial y_j}{\partial x_i} = \sum_{\mathcal{P}_{x_i}^{y_j} \in G} \prod_{e \in \mathcal{P}_{x_i}^{y_j}} e$$



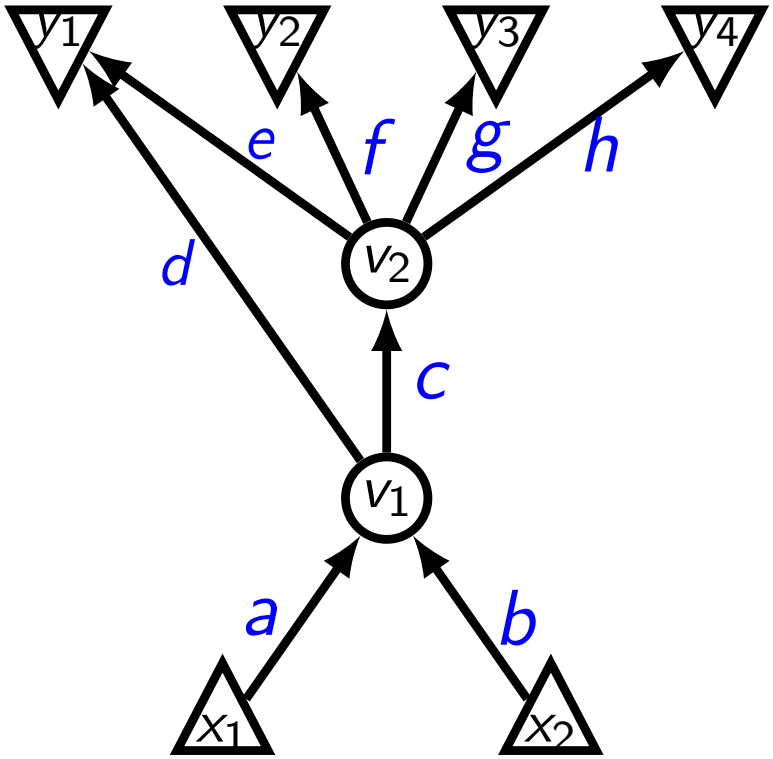

$$J = \begin{bmatrix} ad + ace & bd + bce \\ acf & bcf \\ acg & bcg \\ ach & bch \end{bmatrix}$$

$$= \begin{bmatrix} a(d + ce) & b(d + ce) \\ (ac)f & (bc)f \\ (ac)g & (bc)g \\ (ac)h & (bc)h \end{bmatrix}$$

- Vertex, Edge, Face elimination methods.

# Structural OJA

$$J_{j,i} \equiv \frac{\partial y_j}{\partial x_i} = \sum_{\mathcal{P}_{x_i}^{y_j} \in G} \prod_{e \in \mathcal{P}_{x_i}^{y_j}} e$$



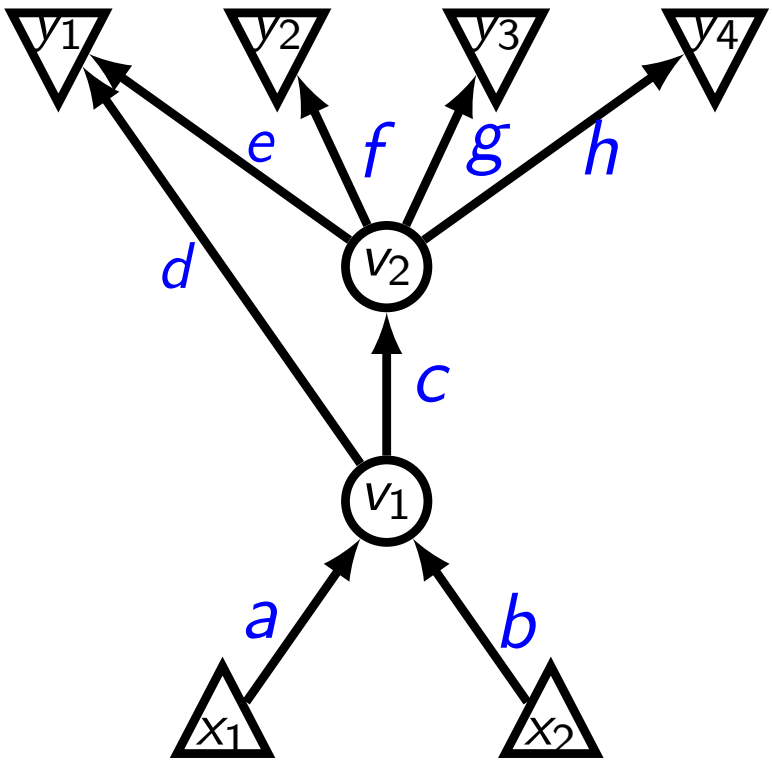

$$J = \begin{bmatrix} ad + ace & bd + bce \\ acf & bcf \\ acg & bcg \\ ach & bch \end{bmatrix}$$

$$= \begin{bmatrix} a(d + ce) & b(d + ce) \\ (ac)f & (bc)f \\ (ac)g & (bc)g \\ (ac)h & (bc)h \end{bmatrix}$$

- Vertex, Edge, Face elimination methods.
- Solved for SEU graphs [Naumann], TTSP graphs [Lyons].

# Structural OJA

$$J_{j,i} \equiv \frac{\partial y_j}{\partial x_i} = \sum_{\mathcal{P}_{x_i}^{y_j} \in G} \prod_{e \in \mathcal{P}_{x_i}^{y_j}} e$$

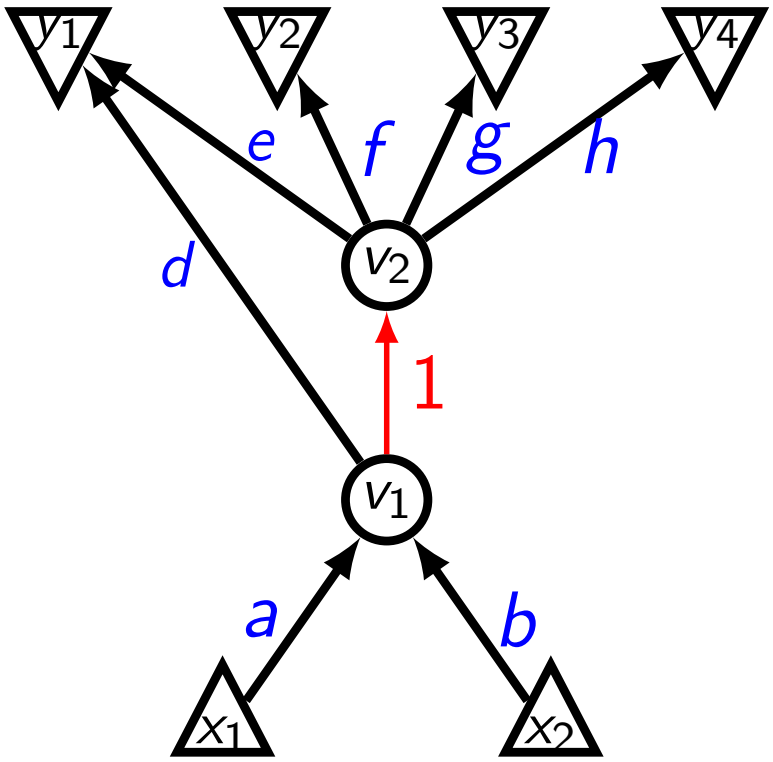



$$J = \begin{bmatrix} ad + ace & bd + bce \\ acf & bcf \\ acg & bcg \\ ach & bch \end{bmatrix}$$

$$= \begin{bmatrix} a(d + ce) & b(d + ce) \\ (ac)f & (bc)f \\ (ac)g & (bc)g \\ (ac)h & (bc)h \end{bmatrix}$$

- Vertex, Edge, Face elimination methods.
- Solved for SEU graphs [Naumann], TTSP graphs [Lyons].
- Complexity still open.

# Structural OJA with Unit Edges



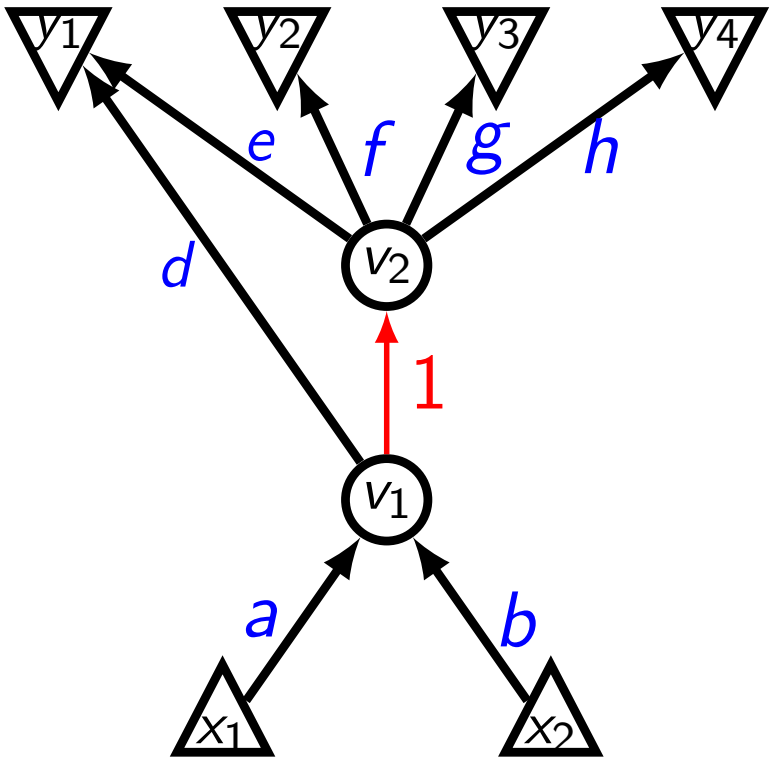

$$J_{j,i} \equiv \frac{\partial y_j}{\partial x_i} = \sum_{\mathcal{P}_{x_i}^{y_j} \in G} \prod_{e \in \mathcal{P}_{x_i}^{y_j}} e$$

$$J = \begin{bmatrix} ad + ae & bd + be \\ af & bf \\ ag & bg \\ ah & bh \end{bmatrix}$$

$$= \begin{bmatrix} a(d+e) & b(d+e) \\ af & bf \\ ag & bg \\ ah & bh \end{bmatrix}$$

- Recognize that some edges are unit labeled.

# Structural OJA with Unit Edges




$$J_{j,i} \equiv \frac{\partial y_j}{\partial x_i} = \sum_{\mathcal{P}_{x_i}^{y_j} \in G} \prod_{e \in \mathcal{P}_{x_i}^{y_j}} e$$

$$J = \begin{bmatrix} ad + ae & bd + be \\ af & bf \\ ag & bg \\ ah & bh \end{bmatrix}$$

$$= \begin{bmatrix} a(d+e) & b(d+e) \\ af & bf \\ ag & bg \\ ah & bh \end{bmatrix}$$

- Recognize that some edges are unit labeled.
- Vertex, Edge, Face eliminations don't apply.

# Outline


(Structural) Optimal Jacobian Accumulation
- SOJA
- Complexity on restricted DAG classes

Generalized OJA
- Where does this problem come from?
- OJA
- Structural OJA
- Structural OJA with Unit Edges

Multiplicative Complexity of Structural OJA with Unit Edges
- Biclique Partition
- Scalar, all paths length $\leq 3$
- Scalar, indegree $\leq 2$

Further Complexity Issues
- Minimizing Additions with Multiplications Fixed
- The Power of Subtraction
- When ISN'T OJA with unit edges harder?

Open Questions

# Multiplicative Complexity of Structural OJA with Unit Edges

**Theorem**

*Minimizing multiplications for* Structural OJA with Unit Edges *is NP-complete.*

**Proof.**

Reduction from Biclique Partition... □

# Biclique Partition

**Instance**: Bipartite graph $G = (A, B, E)$, positive integer $K$.

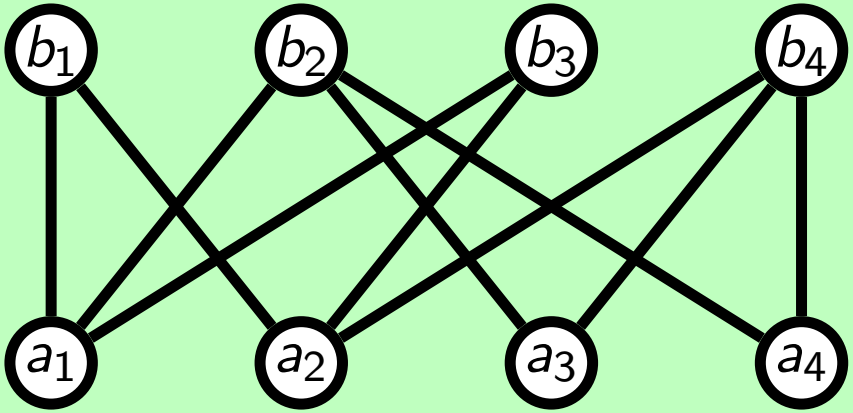
**Question**: Can the edge set $E(G)$ be partitioned into $K$ or fewer complete bipartite graphs?

# Biclique Partition

**Instance**: Bipartite graph $G = (A, B, E)$, positive integer $K$.

**Question**: Can the edge set $E(G)$ be partitioned into $K$ or fewer complete bipartite graphs?

Example

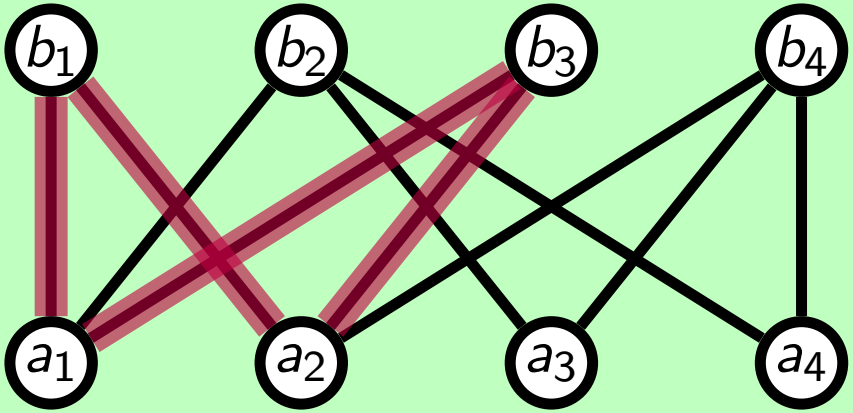


# Biclique Partition

**Instance**: Bipartite graph $G = (A, B, E)$, positive integer $K$.

**Question**: Can the edge set $E(G)$ be partitioned into $K$ or fewer complete bipartite graphs?

Example



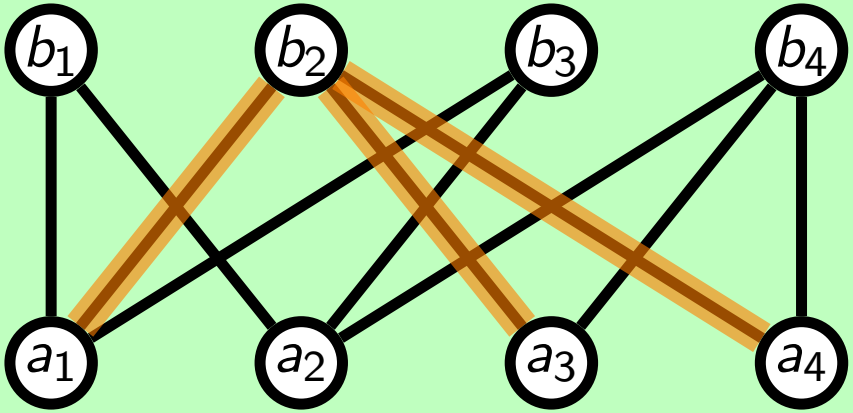

$\{\{a_1, a_2\}, \{b_1, b_3\}\}$

# Biclique Partition

**Instance**: Bipartite graph $G = (A, B, E)$, positive integer $K$.

**Question**: Can the edge set $E(G)$ be partitioned into $K$ or fewer complete bipartite graphs?

Example



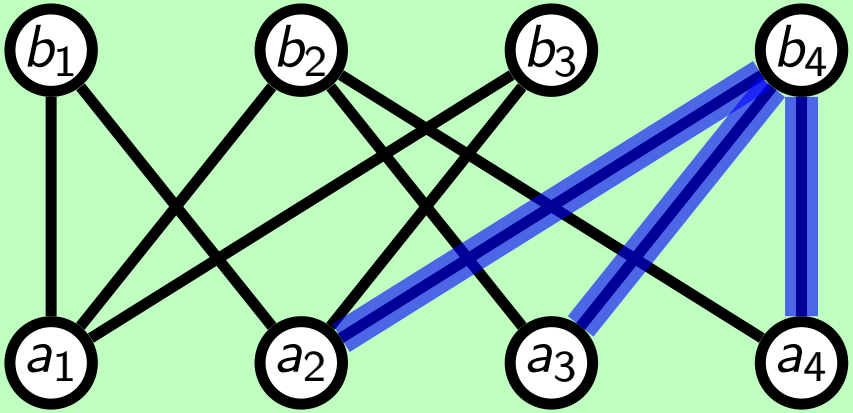

$\{\{a_1, a_3, a_4\}, \{b_2\}\}$

# Biclique Partition

**Instance**: Bipartite graph $G = (A, B, E)$, positive integer $K$.

**Question**: Can the edge set $E(G)$ be partitioned into $K$ or fewer complete bipartite graphs?

Example



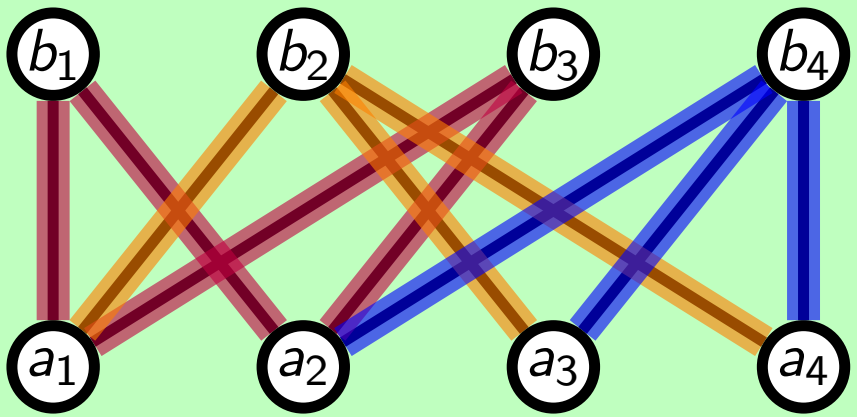

$\{\{a_2, a_3, a_4\}, \{b_4\}\}$

# Biclique Partition

**Instance**: Bipartite graph $G = (A, B, E)$, positive integer $K$.

**Question**: Can the edge set $E(G)$ be partitioned into $K$ or fewer complete bipartite graphs?

Example

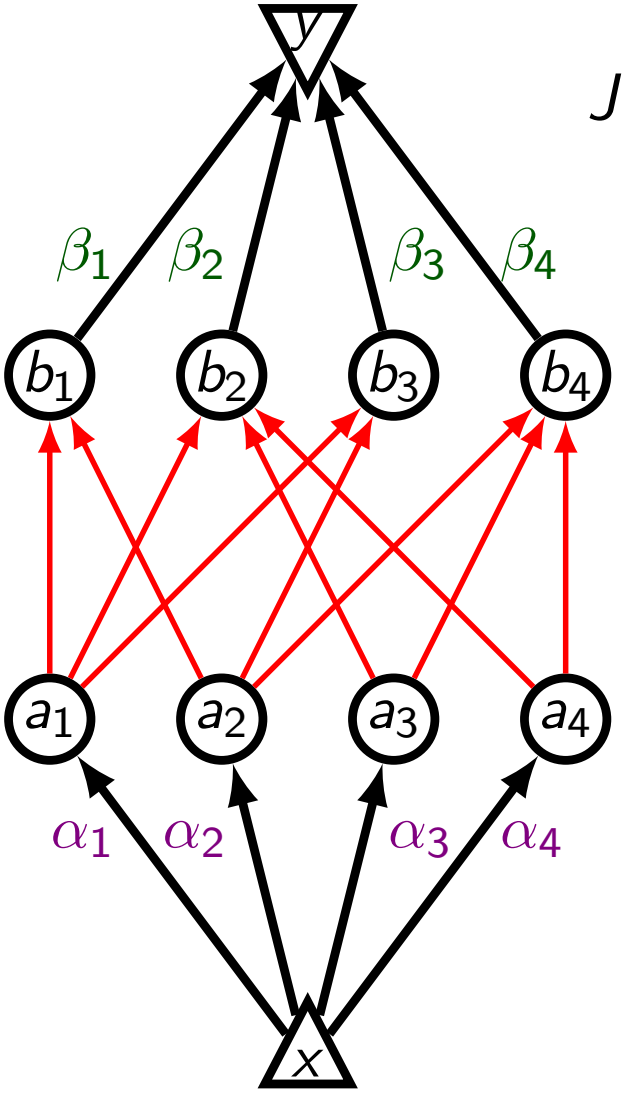



$\{\{a_1, a_2\}, \{b_1, b_3\}\}$

$\{\{a_1, a_3, a_4\}, \{b_2\}\}$

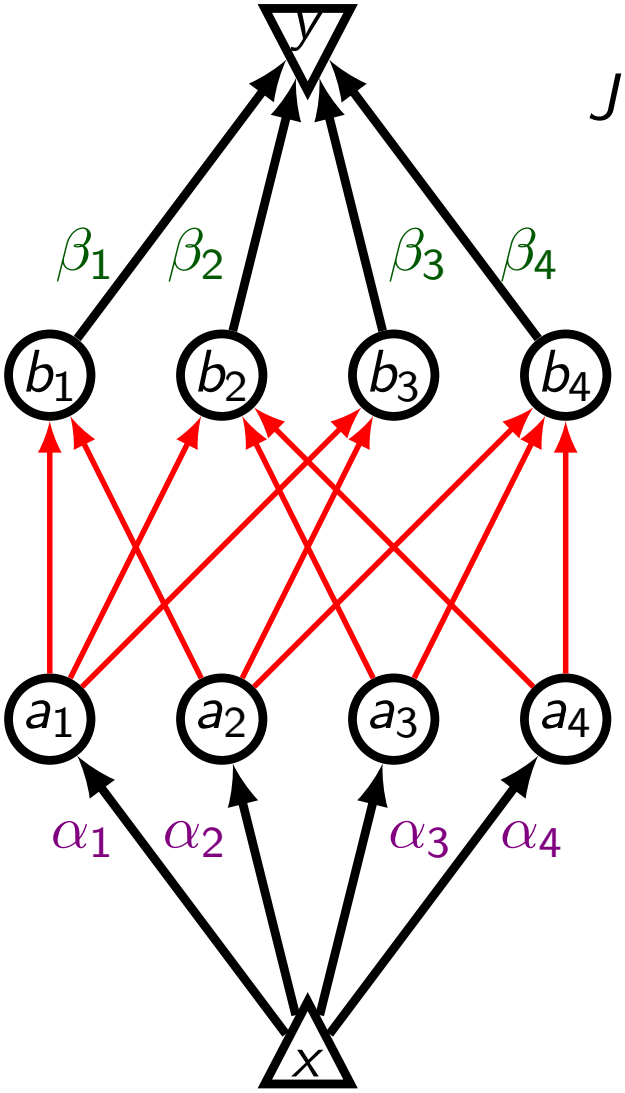
$\{\{a_2, a_3, a_4\}, \{b_4\}\}$

# Example: scalar, all paths length $\leq 3$




$$J = \sum_{\mathcal{P}_x^y \in G} \prod_{e \in \mathcal{P}_x^y} e$$

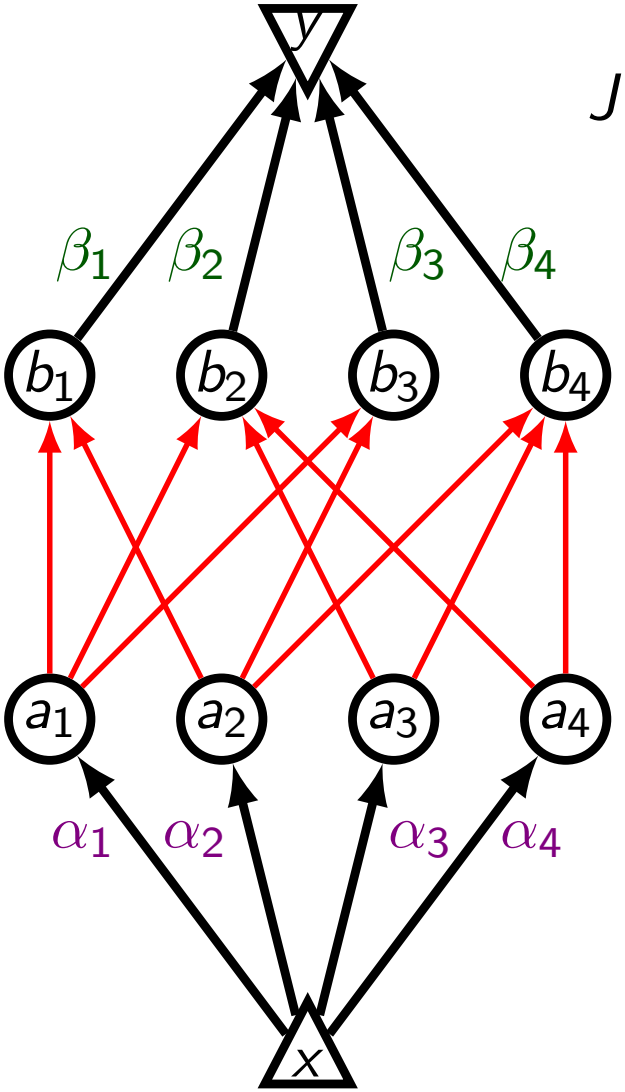
# Example: scalar, all paths length $\leq 3$




$$J = \sum_{\mathcal{P}_x^y \in G} \prod_{e \in \mathcal{P}_x^y} e$$

$$= \underbrace{\begin{array}{c}\alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_1\beta_3 + \alpha_2\beta_1 + \alpha_2\beta_3 \\ + \alpha_2\beta_4 + \alpha_3\beta_2 + \alpha_3\beta_4 + \alpha_4\beta_2 + \alpha_4\beta_4\end{array}}_{\text{10 multiplications}}$$

# Example: scalar, all paths length $\leq 3$

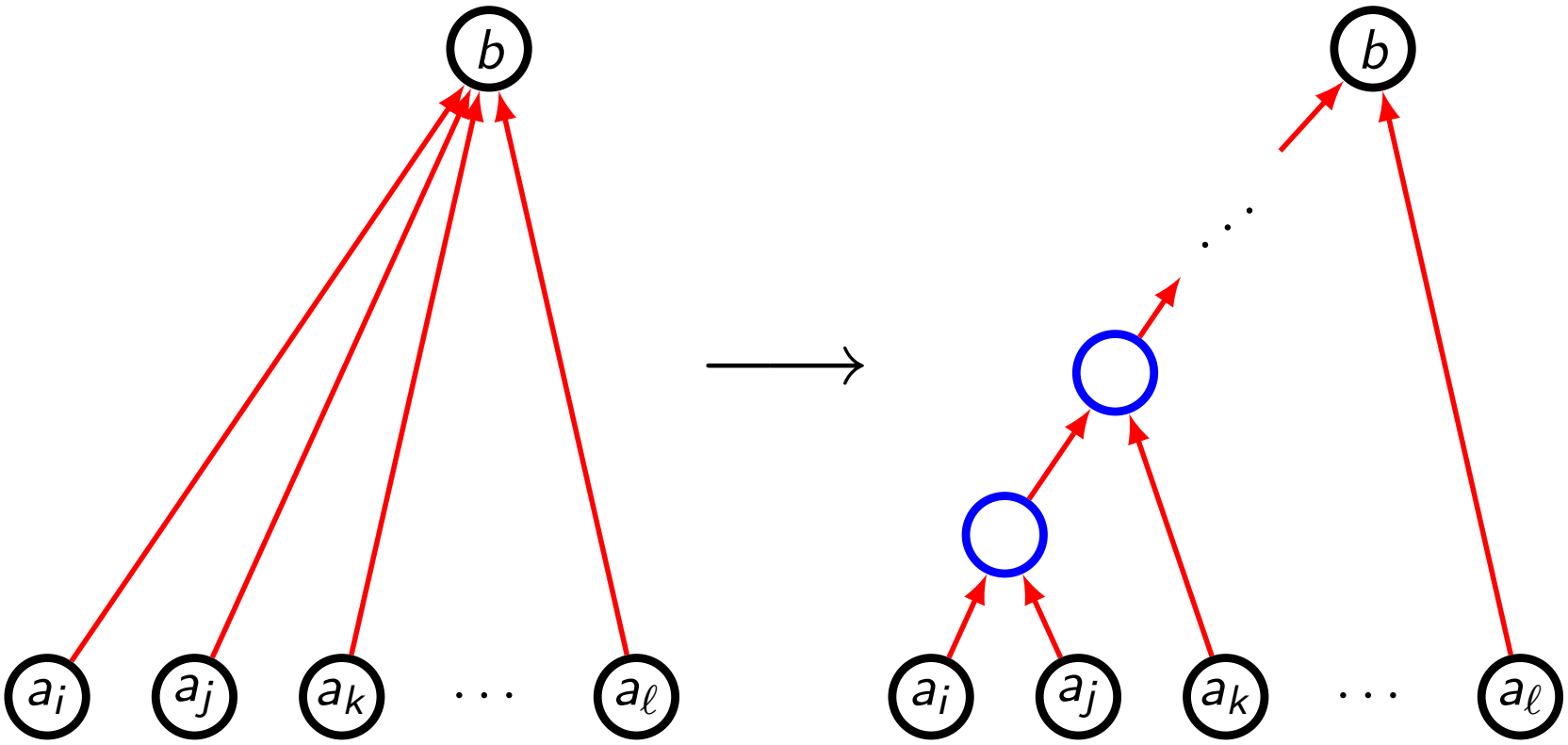



$$J = \sum_{\mathcal{P}_x^y \in G} \prod_{e \in \mathcal{P}_x^y} e$$

$$= \underbrace{\begin{aligned}&\alpha_1\beta_1 + \alpha_1\beta_2 + \alpha_1\beta_3 + \alpha_2\beta_1 + \alpha_2\beta_3 \\ &+ \alpha_2\beta_4 + \alpha_3\beta_2 + \alpha_3\beta_4 + \alpha_4\beta_2 + \alpha_4\beta_4\end{aligned}}_{\text{10 multiplications}}$$

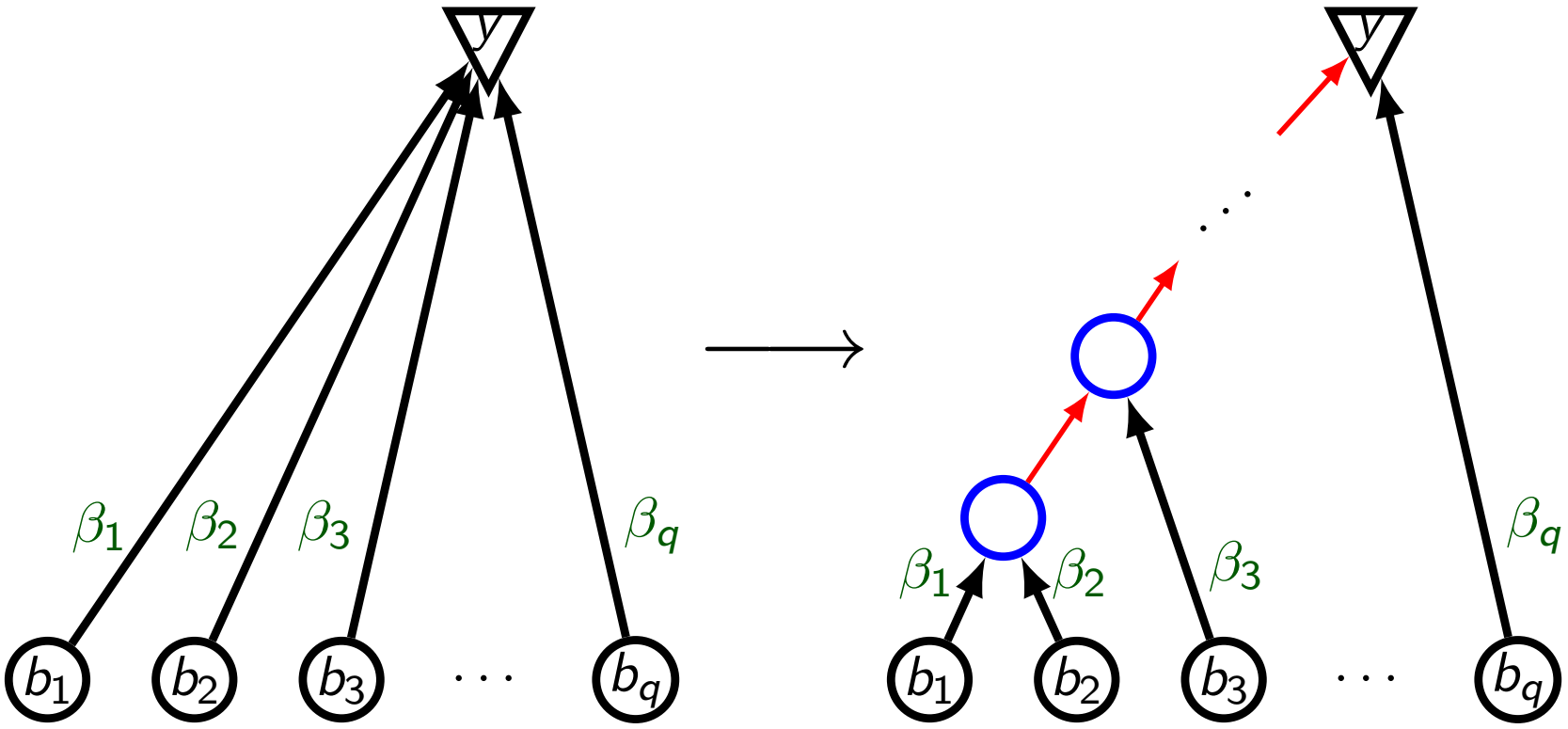
$$= \underbrace{\begin{aligned}&(\alpha_1 + \alpha_2)(\beta_1 + \beta_3) \\ &+ (\alpha_1 + \alpha_3 + \alpha_4)\beta_2 \\ &+ (\alpha_2 + \alpha_3 + \alpha_4)\beta_4\end{aligned}}_{\text{3 multiplications}}$$

# Scalar, indegree $\leq 2$



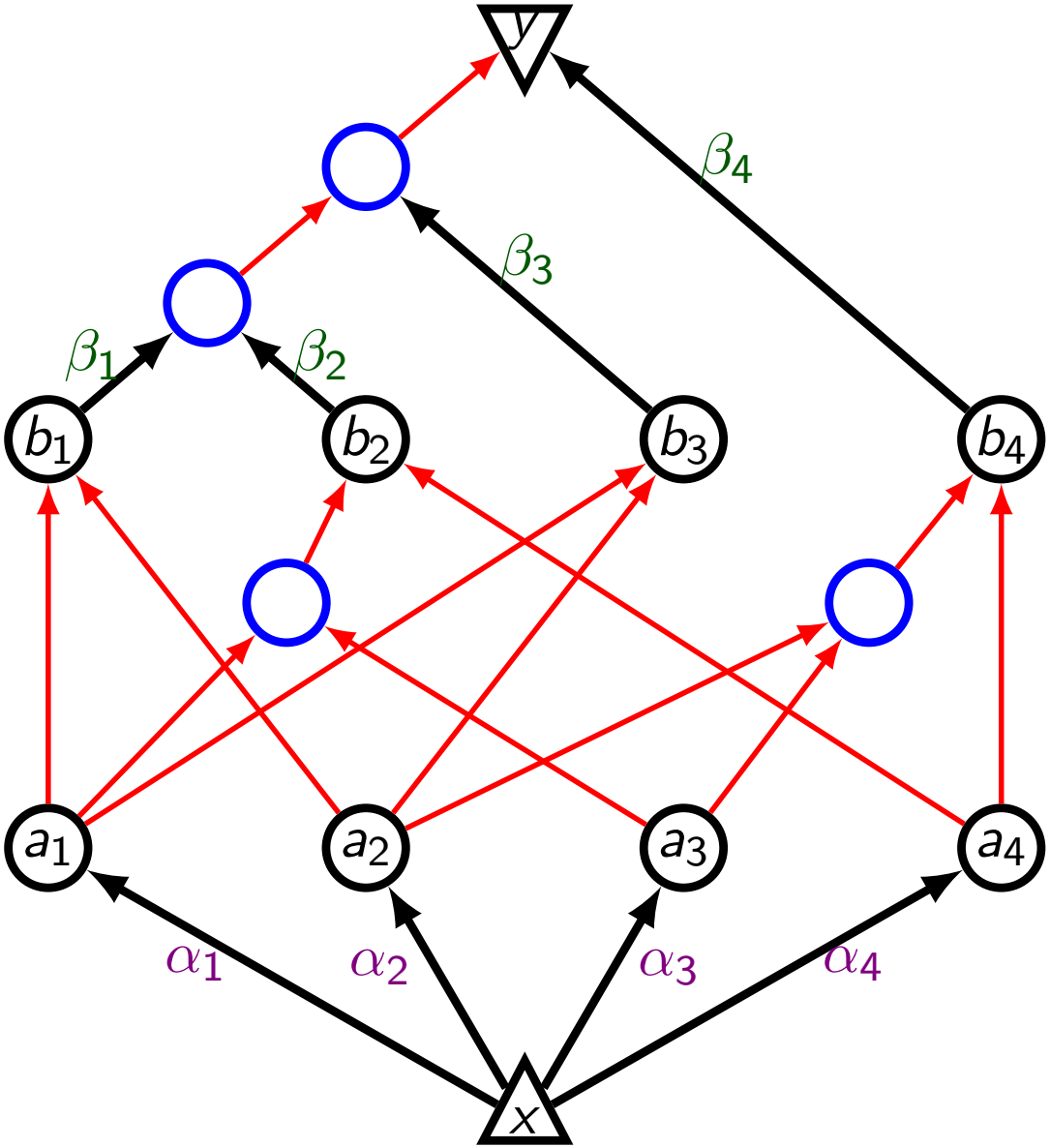

Creating $O(|B||A|)$ new vertices.

# Scalar, indegree $\leq 2$




Creating $|B| - 1$ new vertices.

# Scalar, indegree $\leq 2$



# Multiplicative Complexity of Structural OJA with Unit Edges

### Corollary

*Minimizing multiplications is NP-complete for* Structural OJA with Unit Edges *under the following restrictions:*

- *$J$ is scalar and all paths in $G$ have length $\leq 3$.*

# Multiplicative Complexity of Structural OJA with Unit Edges

### Corollary

*Minimizing multiplications is NP-complete for* Structural OJA with Unit Edges *under the following restrictions:*

- *$J$ is scalar and all paths in $G$ have length $\leq 3$.*
- *$J$ is scalar and all vertices in $G$ have indegree $\leq 2$.*

# Outline


(Structural) Optimal Jacobian Accumulation
- SOJA
- Complexity on restricted DAG classes

Generalized OJA
- Where does this problem come from?
- OJA
- Structural OJA
- Structural OJA with Unit Edges

Multiplicative Complexity of Structural OJA with Unit Edges
- Biclique Partition
- Scalar, all paths length $\leq 3$
- Scalar, indegree $\leq 2$

Further Complexity Issues
- Minimizing Additions with Multiplications Fixed
- The Power of Subtraction
- When ISN'T OJA with unit edges harder?

Open Questions

# Minimizing Additions with Multiplications Fixed

Optimal accumulation for example:

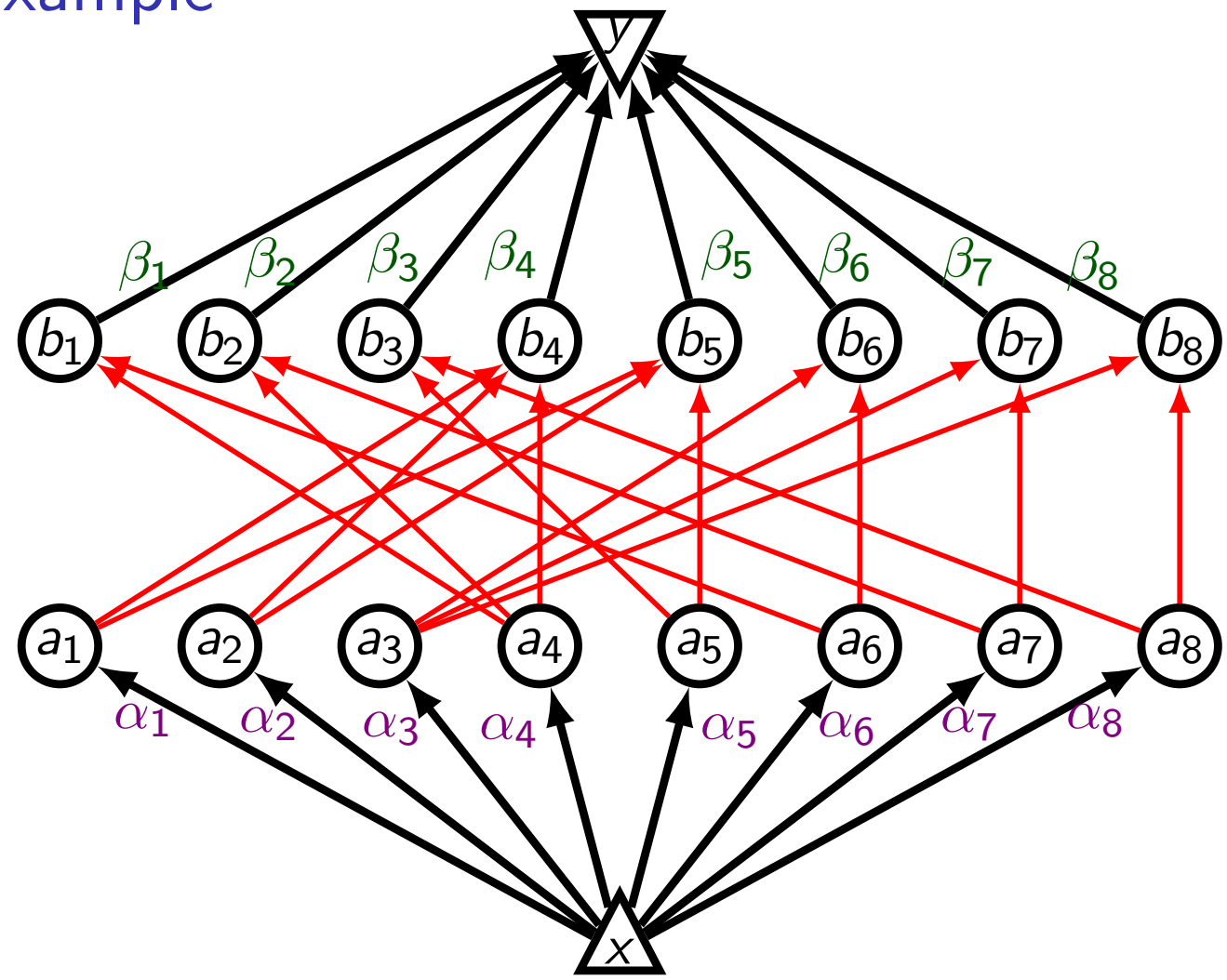
$$(\alpha_1 + \alpha_2)(\beta_1 + \beta_3) + (\alpha_1 + \alpha_3 + \alpha_4)\beta_2 + (\alpha_2 + \alpha_3 + \alpha_4)\beta_4$$

Results in two instances of Ensemble Computation:

- $S = \{\alpha_1, \ldots, \alpha_{|A|}\}$,
  $C = \{\{\alpha_1, \alpha_2\}, \{\alpha_1, \alpha_3, \alpha_4\}, \{\alpha_2, \alpha_3, \alpha_4\}\}$

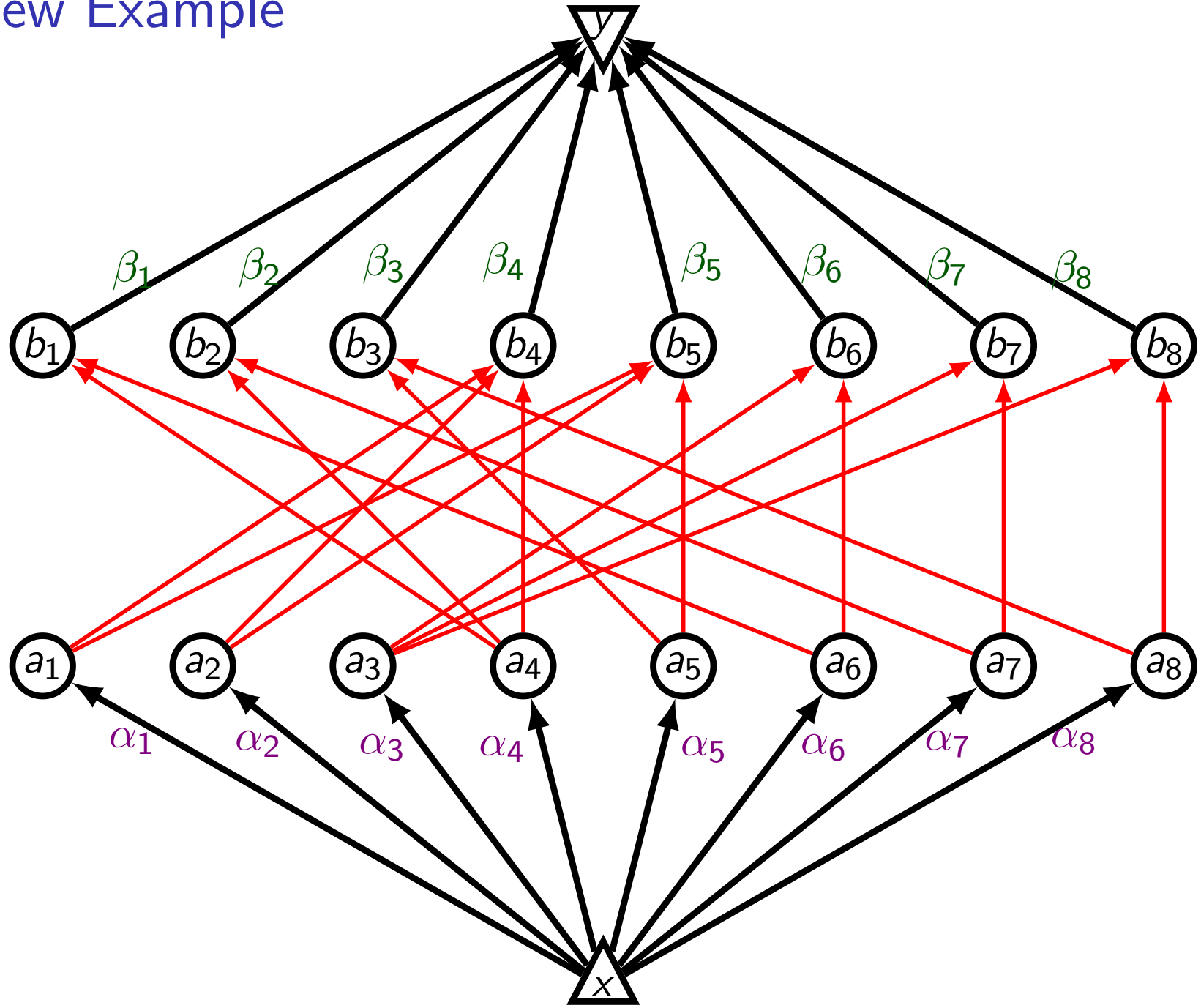
- $S = \{\beta_1, \ldots, \beta_{|B|}\}$,
  $C = \{\{\beta_1, \beta_3\}, \{\beta_2\}, \{\beta_4\}\}$

# A New Example

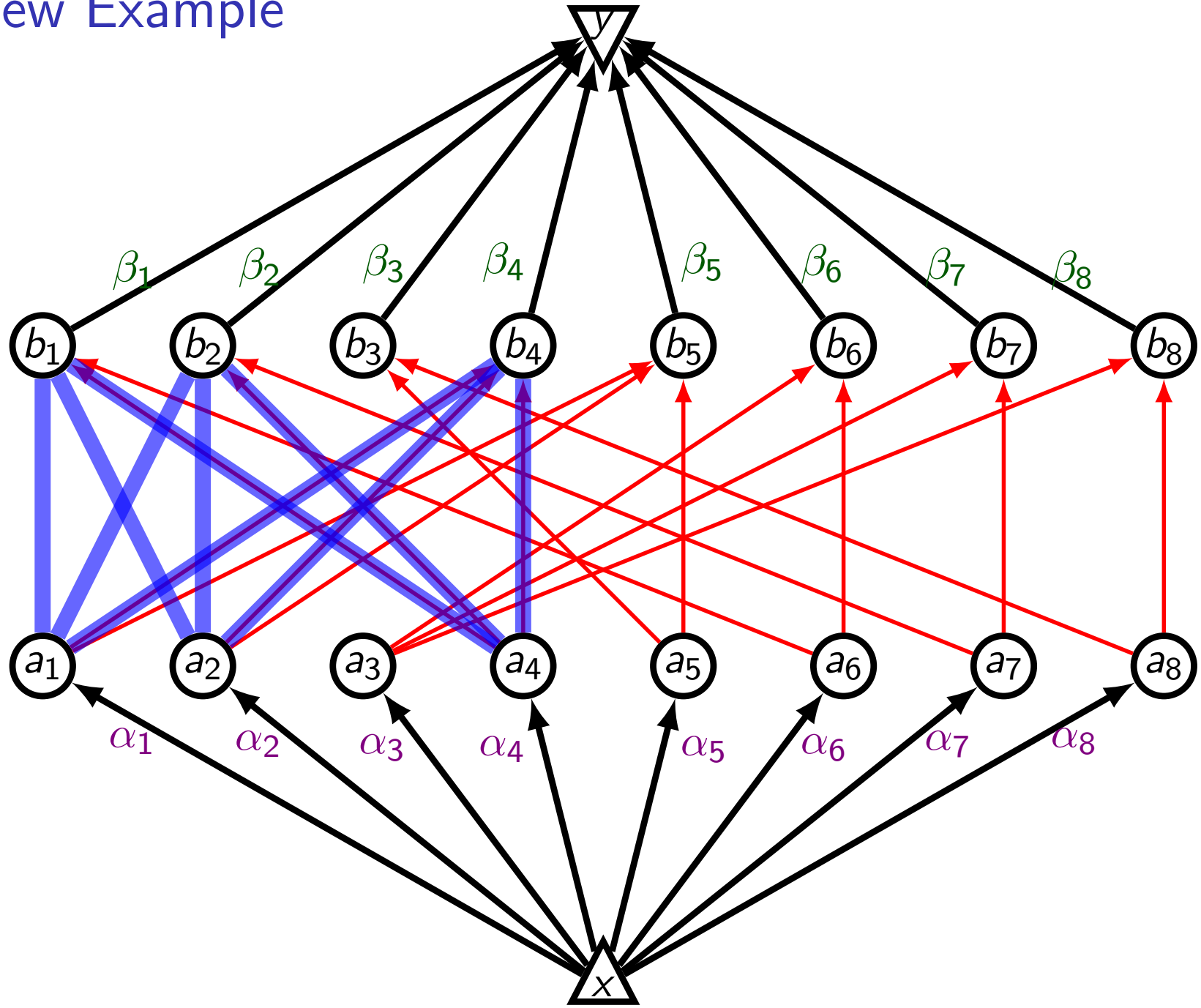



$$
\begin{aligned}
J \;=\; & \alpha_1\beta_4 + \alpha_1\beta_5 + \alpha_2\beta_4 + \alpha_2\beta_5 + \alpha_3\beta_6 + \alpha_3\beta_7 + \alpha_3\beta_8 \\
& + \alpha_4\beta_1 + \alpha_4\beta_2 + \alpha_4\beta_4 + \alpha_5\beta_3 + \alpha_5\beta_5 + \alpha_6\beta_1 + \alpha_6\beta_6 \\
& + \alpha_7\beta_2 + \alpha_7\beta_7 + \alpha_8\beta_3 + \alpha_8\beta_8
\end{aligned}
$$

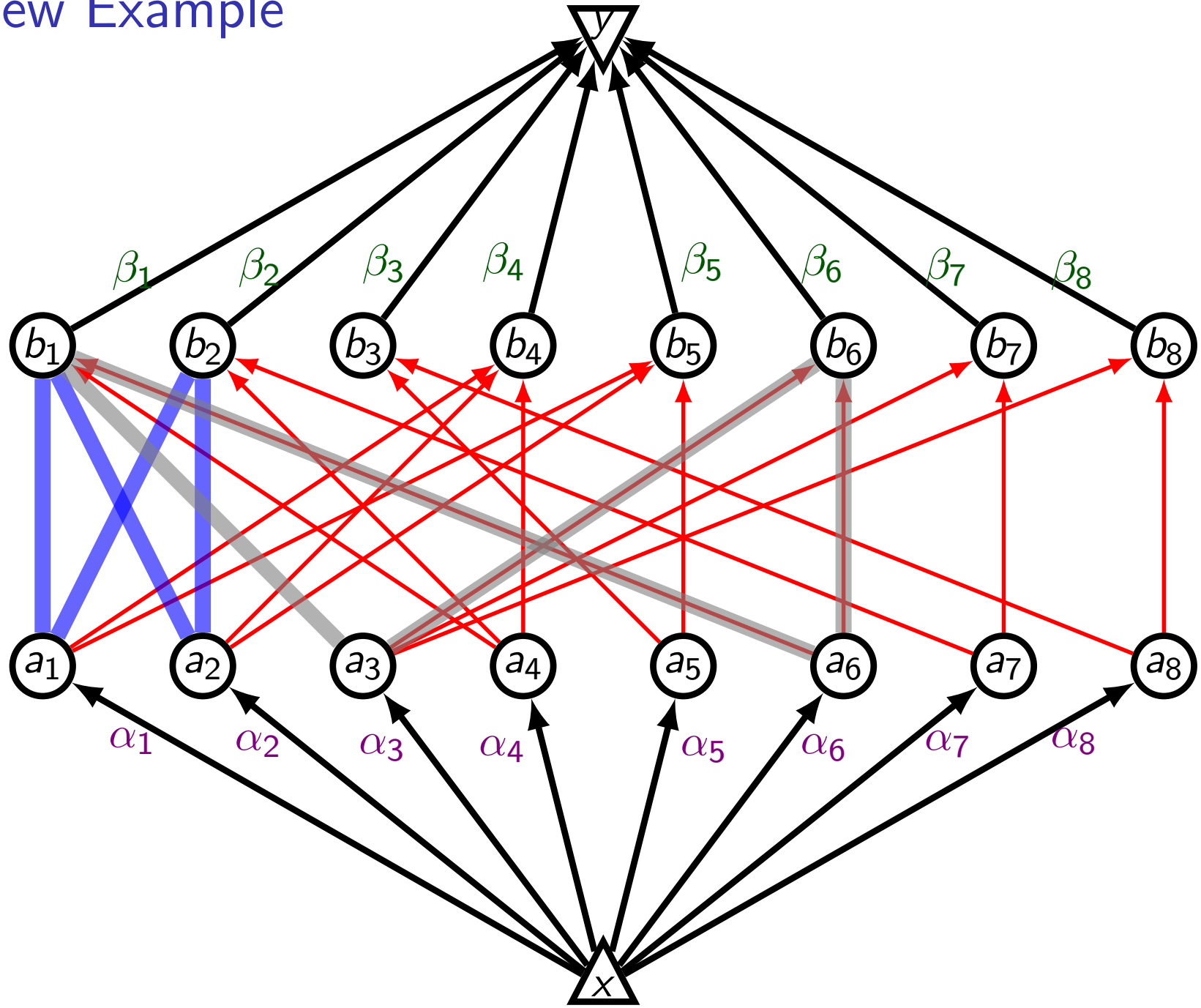
# A New Example



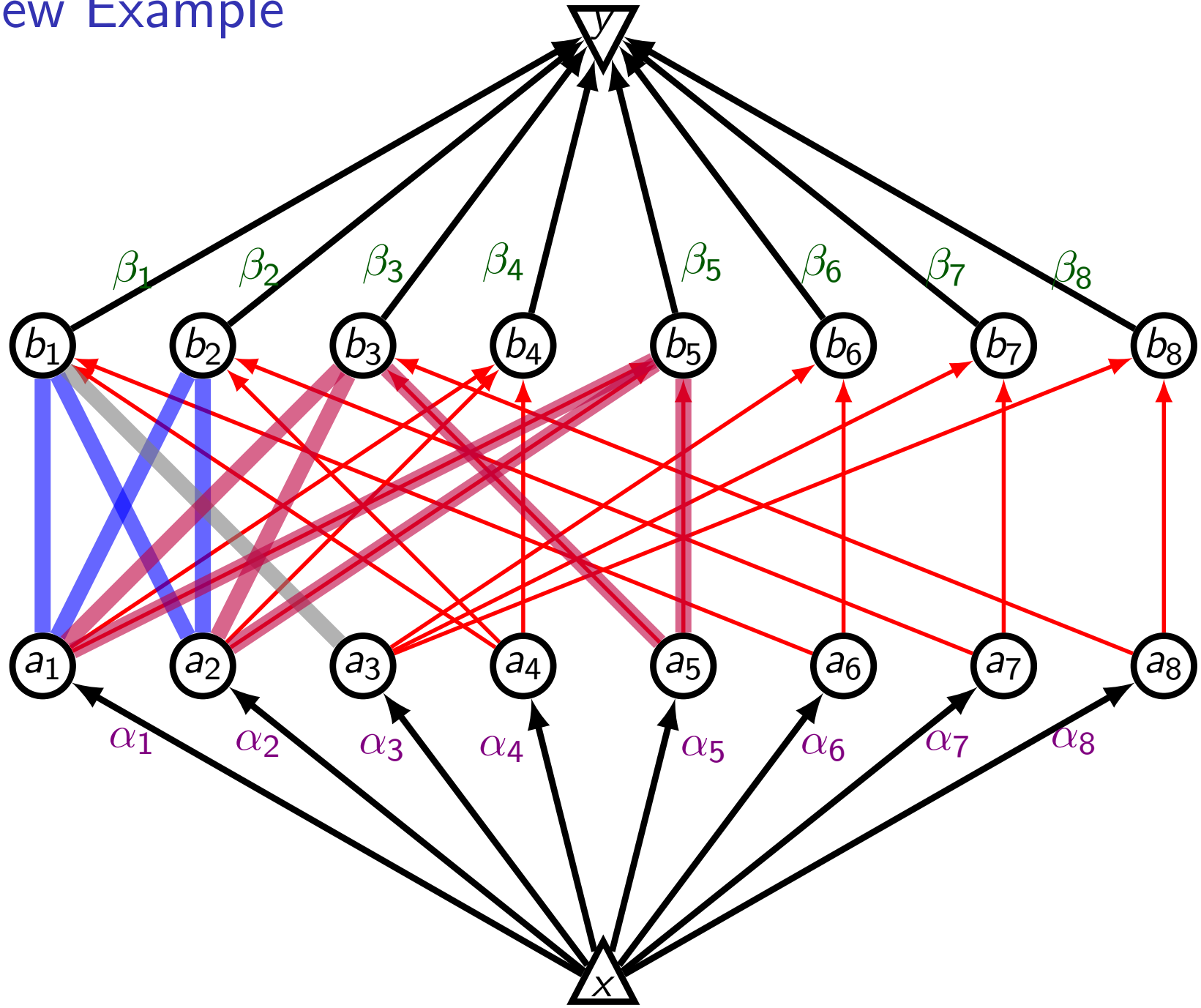
# A New Example



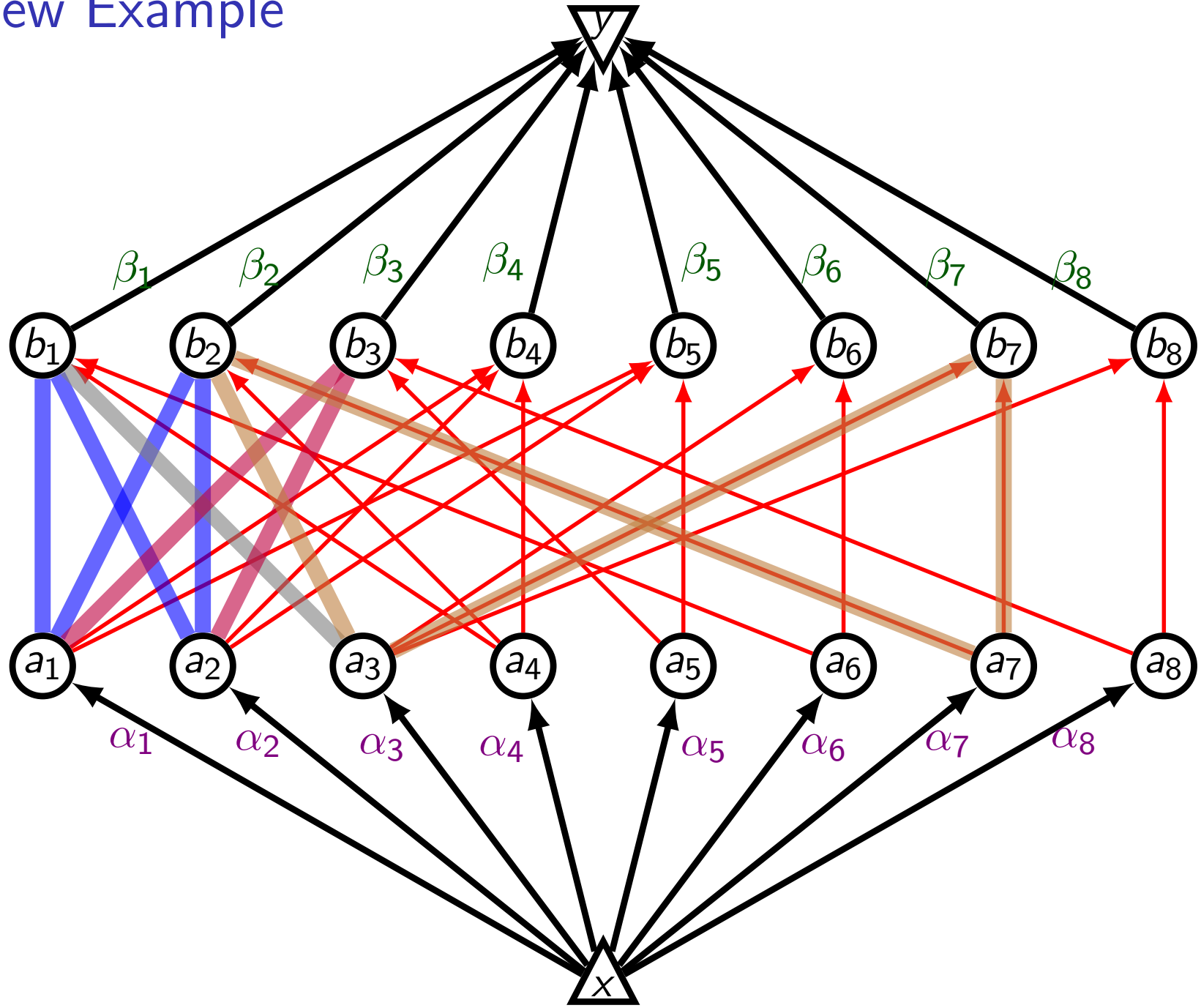
# A New Example



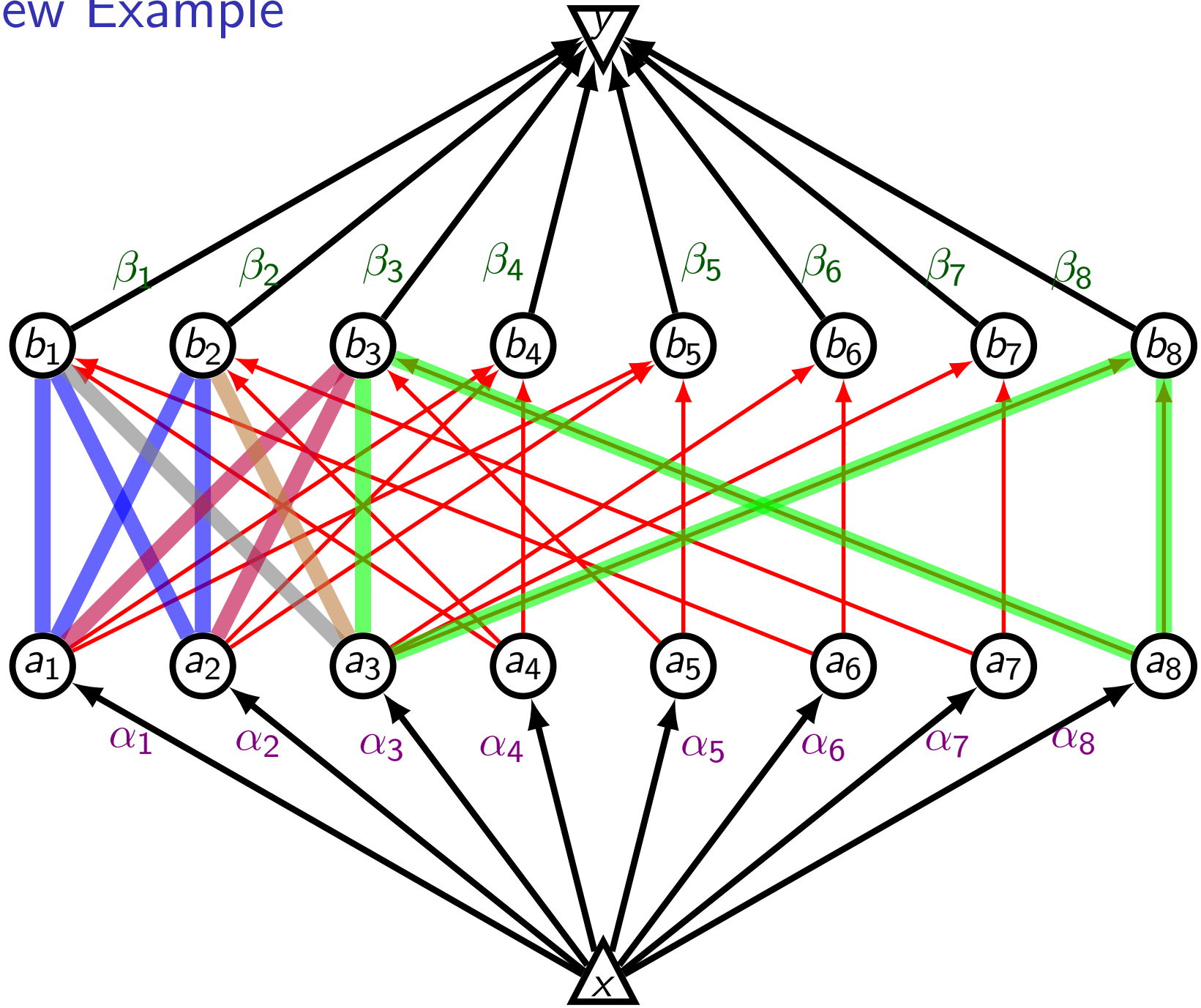
# A New Example



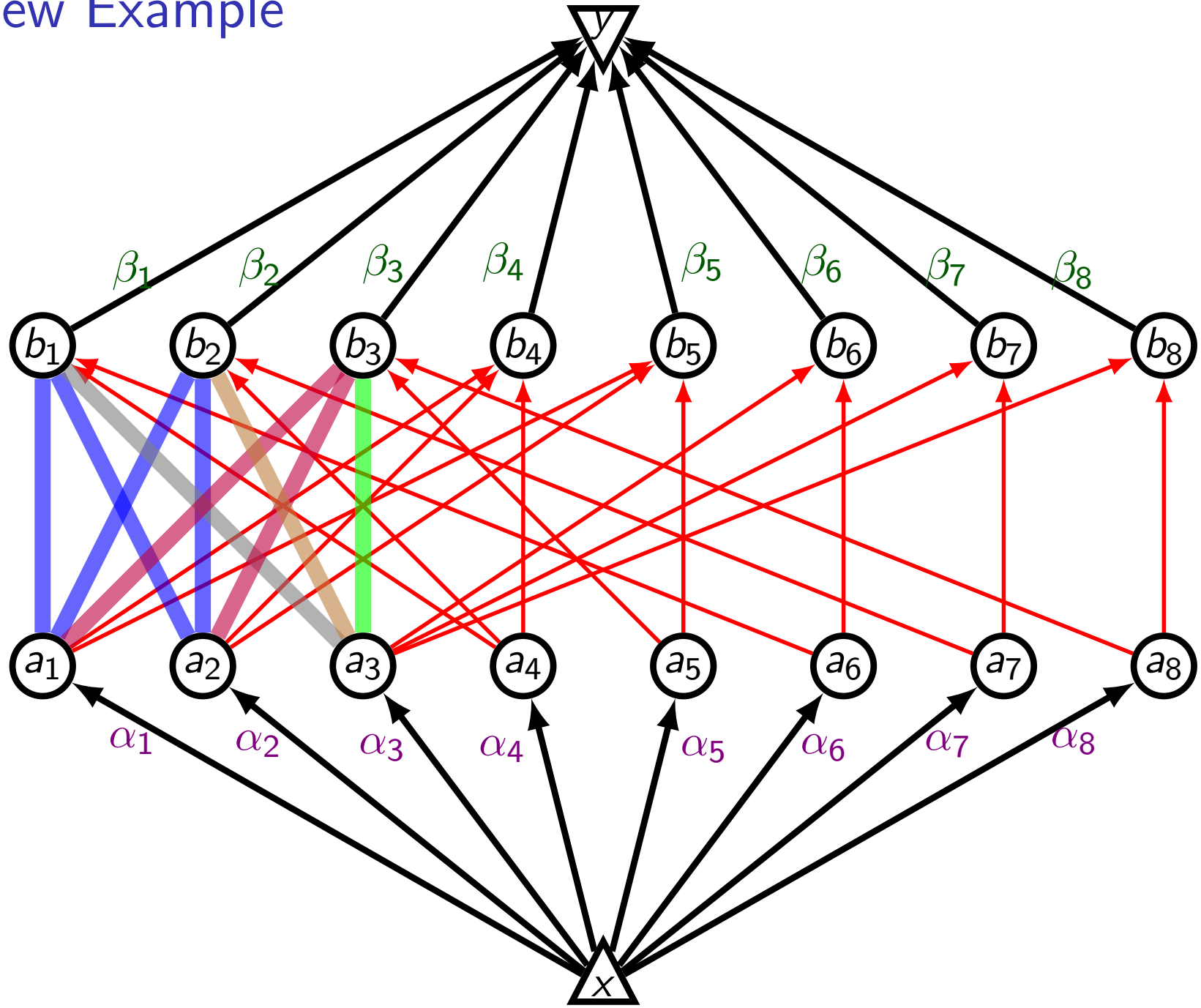
# A New Example



# A New Example



# A New Example



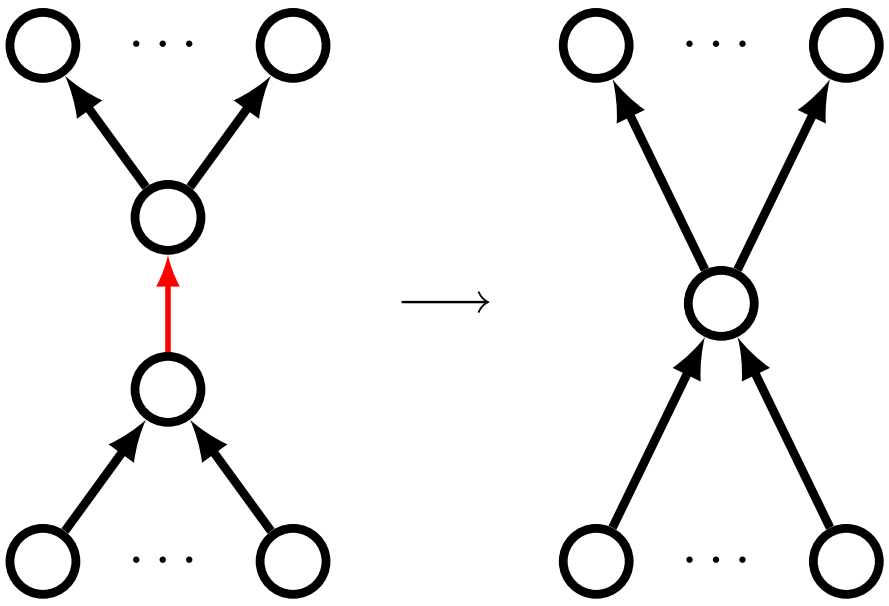
# Bilinear Forms (Gonzalez and JáJá, 1980)

$$
\begin{aligned}
J &= \alpha_1\beta_4 + \alpha_1\beta_5 + \alpha_2\beta_4 + \alpha_2\beta_5 + \alpha_3\beta_6 + \alpha_3\beta_7 + \alpha_3\beta_8 \\
&\quad + \alpha_4\beta_1 + \alpha_4\beta_2 + \alpha_4\beta_4 + \alpha_5\beta_3 + \alpha_5\beta_5 + \alpha_6\beta_1 + \alpha_6\beta_6 \\
&\quad + \alpha_7\beta_2 + \alpha_7\beta_7 + \alpha_8\beta_3 + \alpha_8\beta_8 \\
&= (\alpha_1\ \alpha_2\ \alpha_3\ \alpha_4\ \alpha_5\ \alpha_6\ \alpha_7\ \alpha_8)
\begin{pmatrix}
\square & \square & \square & \blacksquare & \blacksquare & \square & \square & \square \\
\square & \square & \square & \blacksquare & \blacksquare & \square & \square & \square \\
\square & \square & \square & \square & \square & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \square & \blacksquare & \square & \square & \square & \square \\
\square & \square & \blacksquare & \square & \blacksquare & \square & \square & \square \\
\blacksquare & \square & \square & \square & \square & \blacksquare & \square & \square \\
\square & \blacksquare & \square & \square & \square & \square & \blacksquare & \square \\
\square & \square & \blacksquare & \square & \square & \square & \square & \blacksquare
\end{pmatrix}
\begin{pmatrix}
\beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8
\end{pmatrix}
\end{aligned}
$$

# Bilinear Forms (Gonzalez and JáJá, 1980)

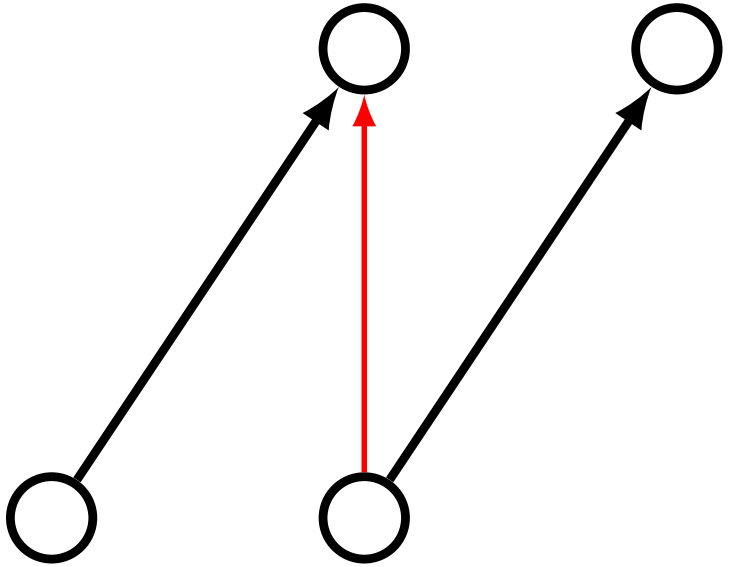
$$
\begin{aligned}
J &= \alpha_1\beta_4 + \alpha_1\beta_5 + \alpha_2\beta_4 + \alpha_2\beta_5 + \alpha_3\beta_6 + \alpha_3\beta_7 + \alpha_3\beta_8 \\
&\quad + \alpha_4\beta_1 + \alpha_4\beta_2 + \alpha_4\beta_4 + \alpha_5\beta_3 + \alpha_5\beta_5 + \alpha_6\beta_1 + \alpha_6\beta_6 \\
&\quad + \alpha_7\beta_2 + \alpha_7\beta_7 + \alpha_8\beta_3 + \alpha_8\beta_8 \\
&= (\alpha_1\ \alpha_2\ \alpha_3\ \alpha_4\ \alpha_5\ \alpha_6\ \alpha_7\ \alpha_8)
\begin{pmatrix}
\square & \square & \square & \blacksquare & \blacksquare & \square & \square & \square \\
\square & \square & \square & \blacksquare & \blacksquare & \square & \square & \square \\
\square & \square & \square & \square & \square & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \square & \blacksquare & \square & \square & \square & \square \\
\square & \square & \blacksquare & \square & \blacksquare & \square & \square & \square \\
\blacksquare & \square & \square & \square & \square & \blacksquare & \square & \square \\
\square & \blacksquare & \square & \square & \square & \square & \blacksquare & \square \\
\square & \square & \blacksquare & \square & \square & \square & \square & \blacksquare
\end{pmatrix}
\begin{pmatrix}
\beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8
\end{pmatrix} \\
&= (\alpha_1 + \alpha_2 + \alpha_3)(\beta_1 + \beta_2 + \beta_4) \\
&\quad + (\alpha_3 + \alpha_6)(\beta_1 + \beta_6) + (\alpha_1 + \alpha_2 + \alpha_5)(\beta_3 + \beta_5) \\
&\quad + (\alpha_3 + \alpha_7)(\beta_2 + \beta_7) + (\alpha_3 + \alpha_8)(\beta_3 + \beta_8) \\
&\quad \boxed{- (\alpha_1 + \alpha_2 + \alpha_3)(\beta_1 + \beta_2 + \beta_3)}
\end{aligned}
$$

# When are the unit edges harmless?



When all the unit edges can be *contracted*.

# When are the unit edges harmless?



This type of unit edge cannot be contracted.

# Outline


(Structural) Optimal Jacobian Accumulation
- SOJA
- Complexity on restricted DAG classes

Generalized OJA
- Where does this problem come from?
- OJA
- Structural OJA
- Structural OJA with Unit Edges

Multiplicative Complexity of Structural OJA with Unit Edges
- Biclique Partition
- Scalar, all paths length $\leq 3$
- Scalar, indegree $\leq 2$

Further Complexity Issues
- Minimizing Additions with Multiplications Fixed
- The Power of Subtraction
- When ISN'T OJA with unit edges harder?

Open Questions

# Open Questions

Algorithms:

- Alternative to vertex, edge, face elimination heuristics.

Complexity:

- Relationship between general OJA and structural OJA with unit edges.
- Could subtraction help general OJA or structural OJA?
- Complexity of minimizing total ops for OJA with unit edges.