7.3 The Converse

The Dependency Graph

The Dependency Graph

- The communication system consists of the r.v.'s

$$W, X_1, Y_1, X_2, Y_2, \cdots, X_n, Y_n, \hat{W}$$

  generated in this order.

The Dependency Graph

- The communication system consists of the r.v.'s

$$W, X_1, Y_1, X_2, Y_2, \cdots, X_n, Y_n, \hat{W}$$

  generated in this order.

- The memorylessness of the DMC imposes the following Markov constraint for each i:

$$\underbrace{(W, X_1, Y_1, \cdots, X_{i-1}, Y_{i-1})}_{T_{i-}} \rightarrow X_i \rightarrow Y_i$$

The Dependency Graph

- The communication system consists of the r.v.'s

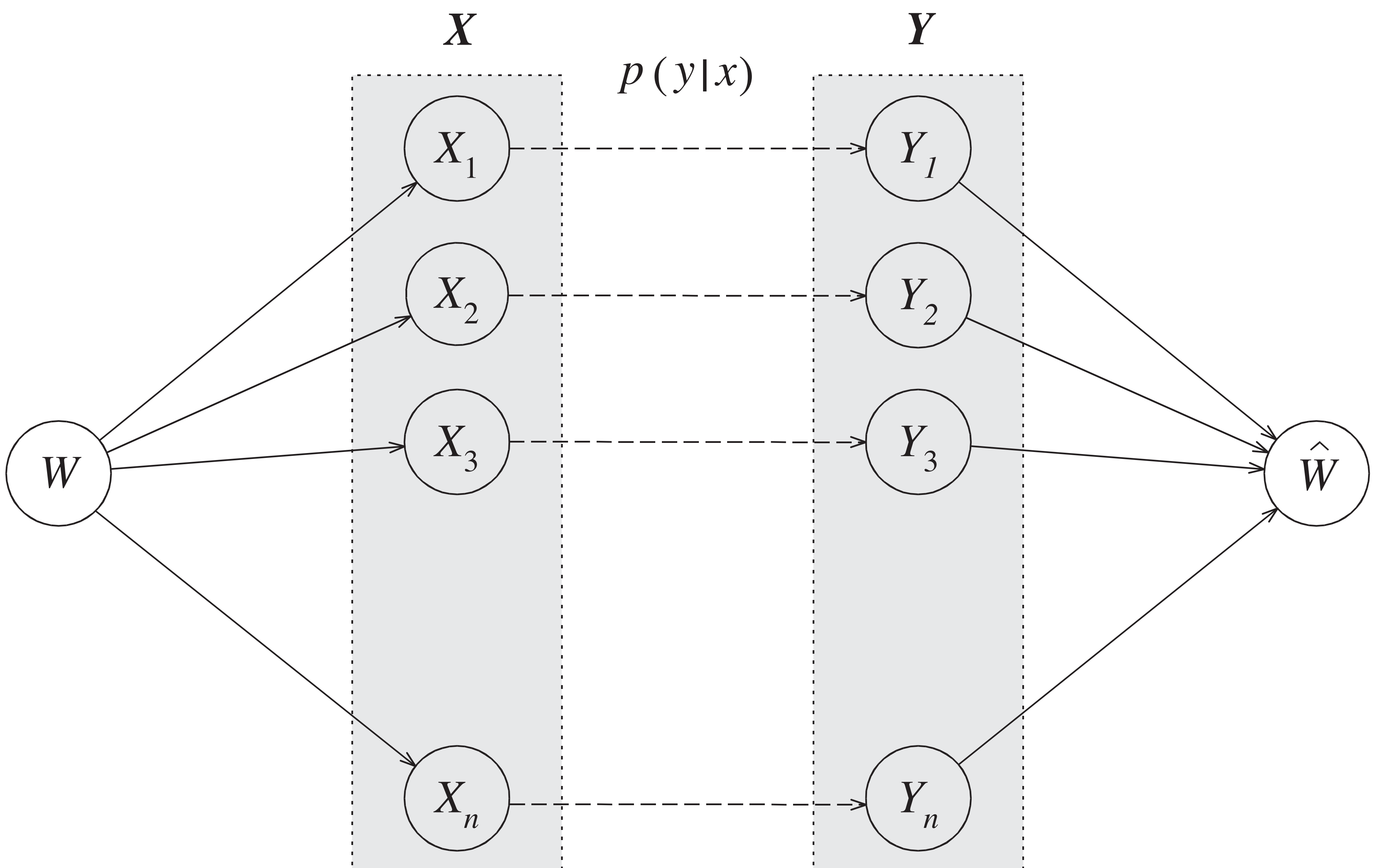
$$W, X_1, Y_1, X_2, Y_2, \cdots, X_n, Y_n, \hat{W}$$

  generated in this order.

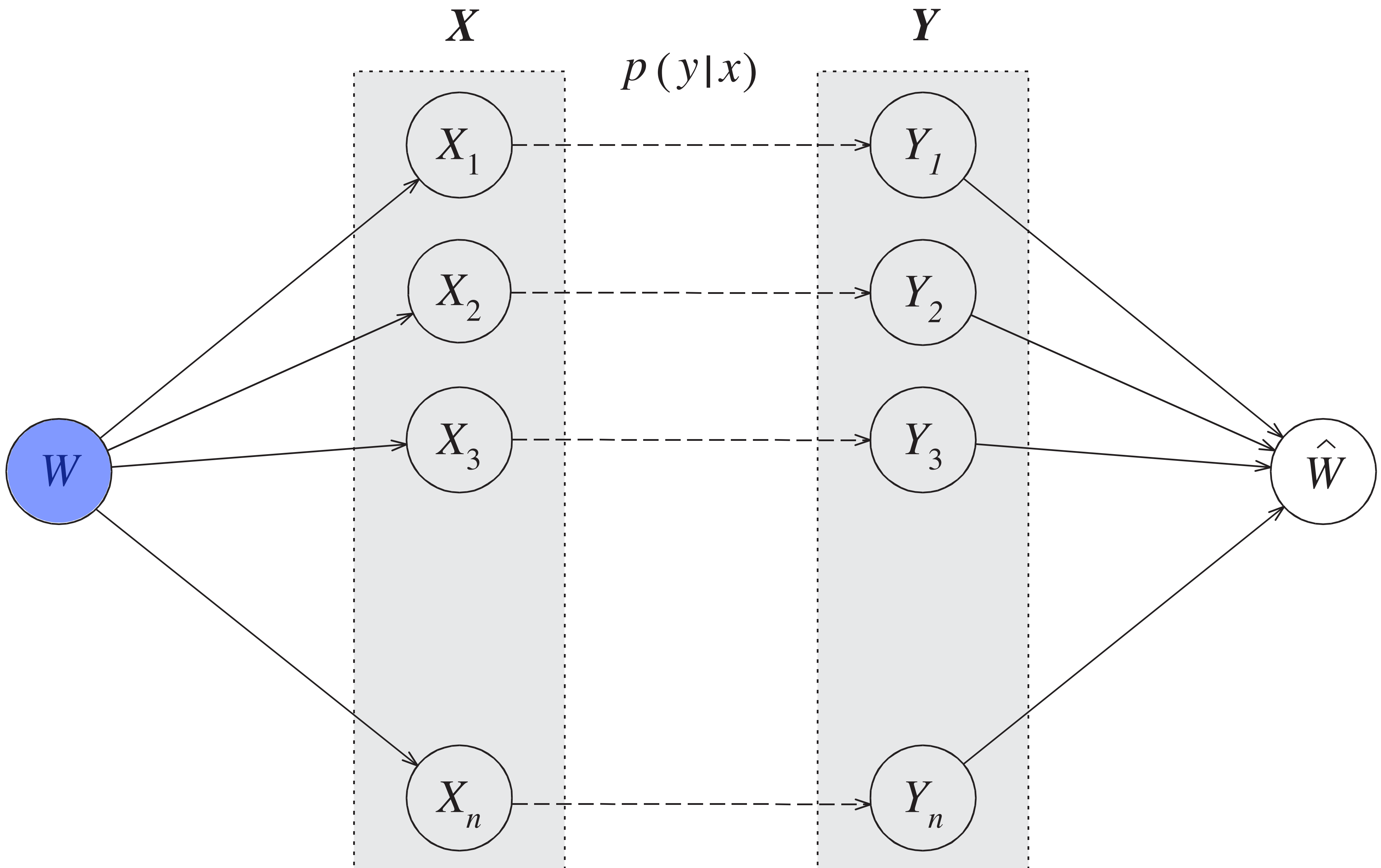
- The memorylessness of the DMC imposes the following Markov constraint for each i:

$$\underbrace{(W, X_1, Y_1, \cdots, X_{i-1}, Y_{i-1})}_{T_{i-}} \rightarrow X_i \rightarrow Y_i$$

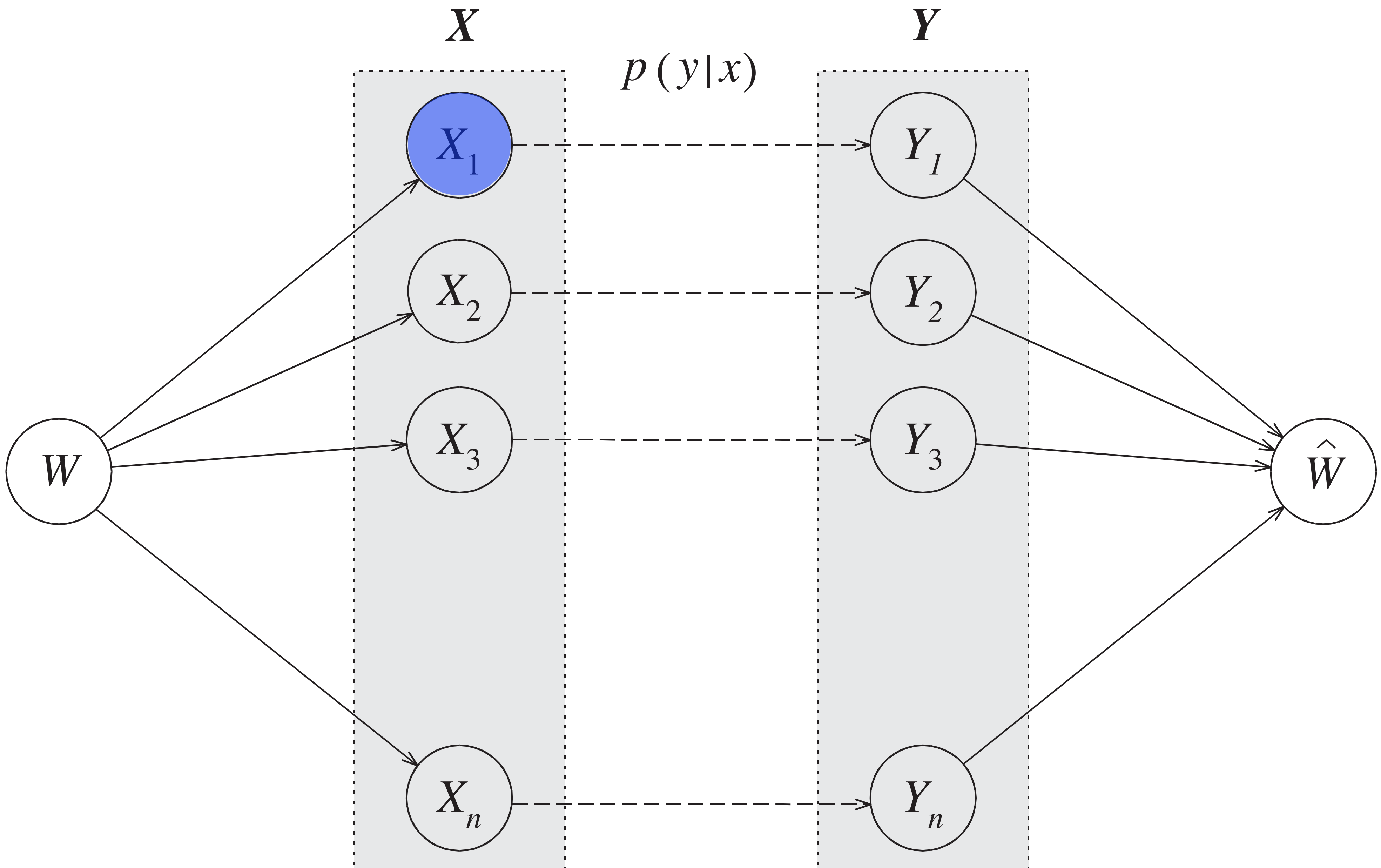
- The dependency graph can be composed accordingly.



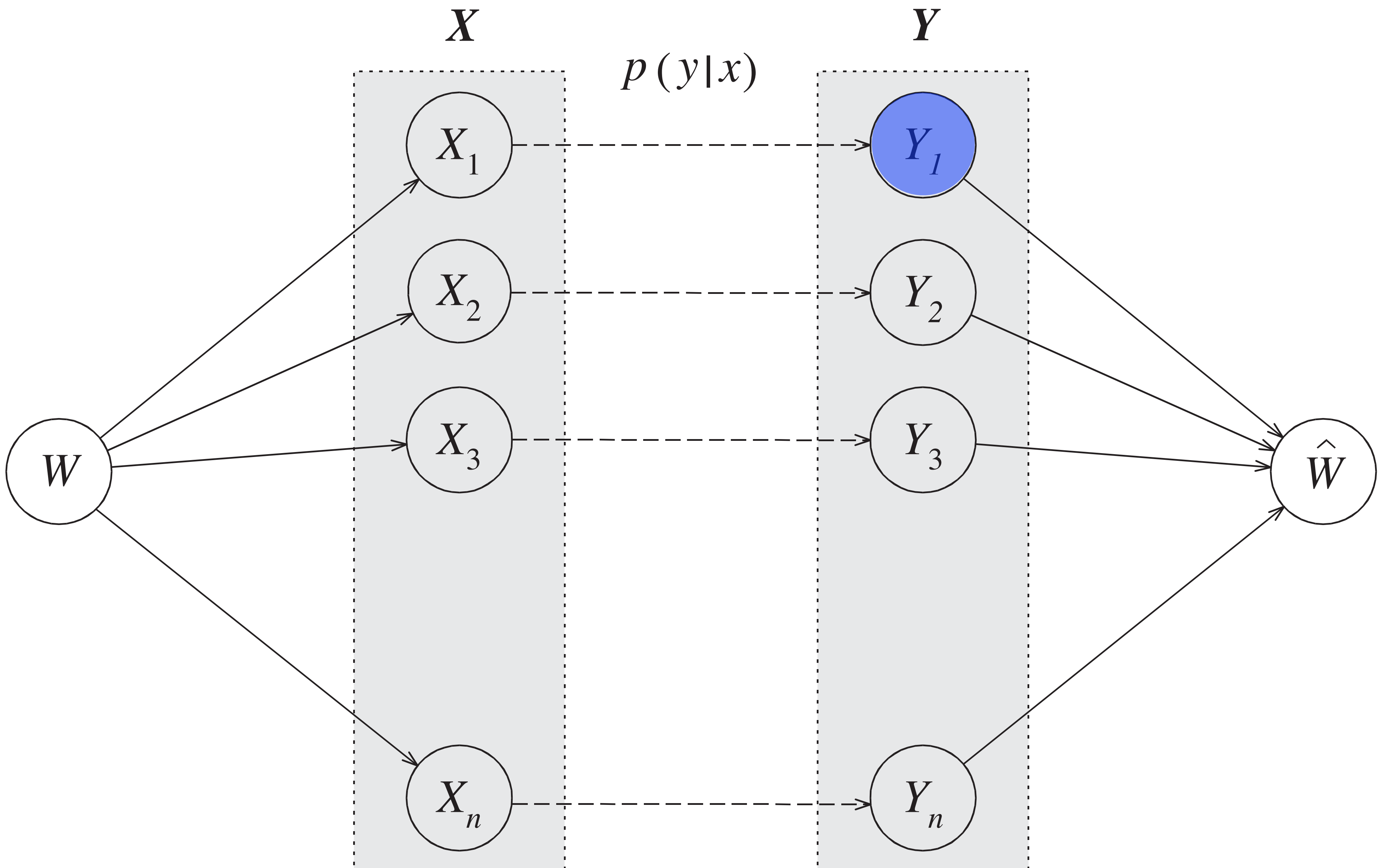
X
Y
p (y | x)
W
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
Ŵ



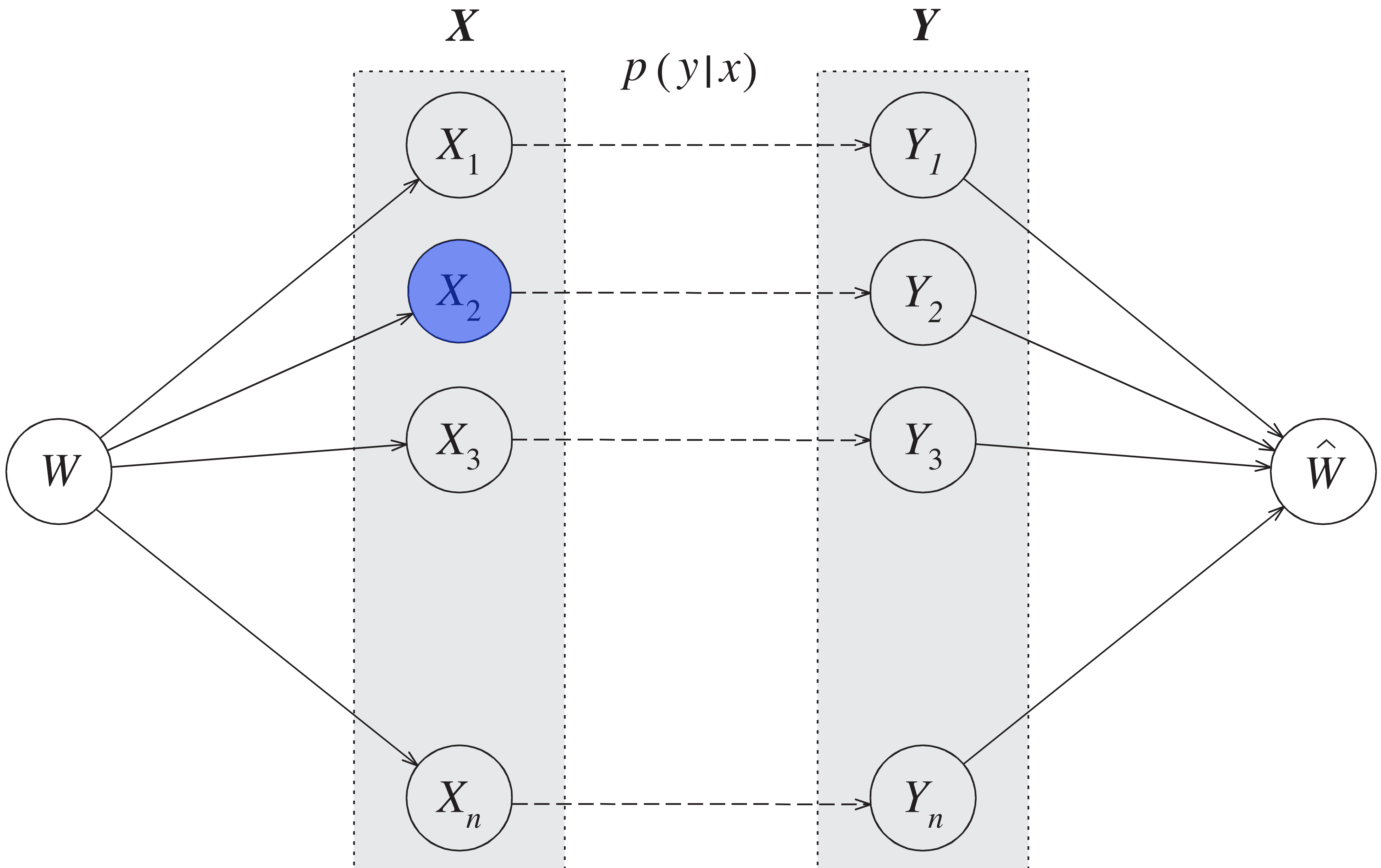
X
Y
p (y | x)
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
W
Ŵ

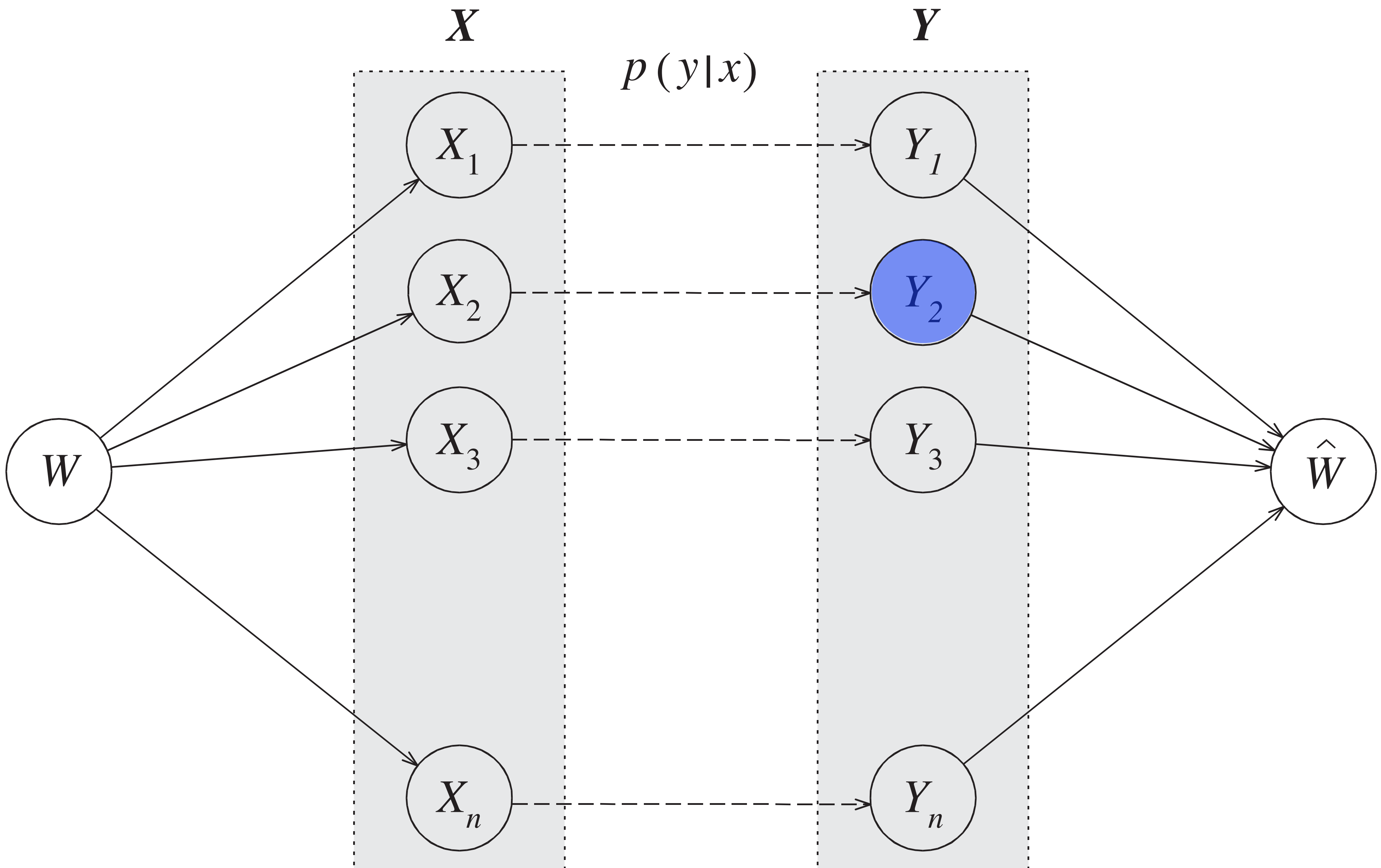


X
Y
p (y | x)
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
W
Ŵ

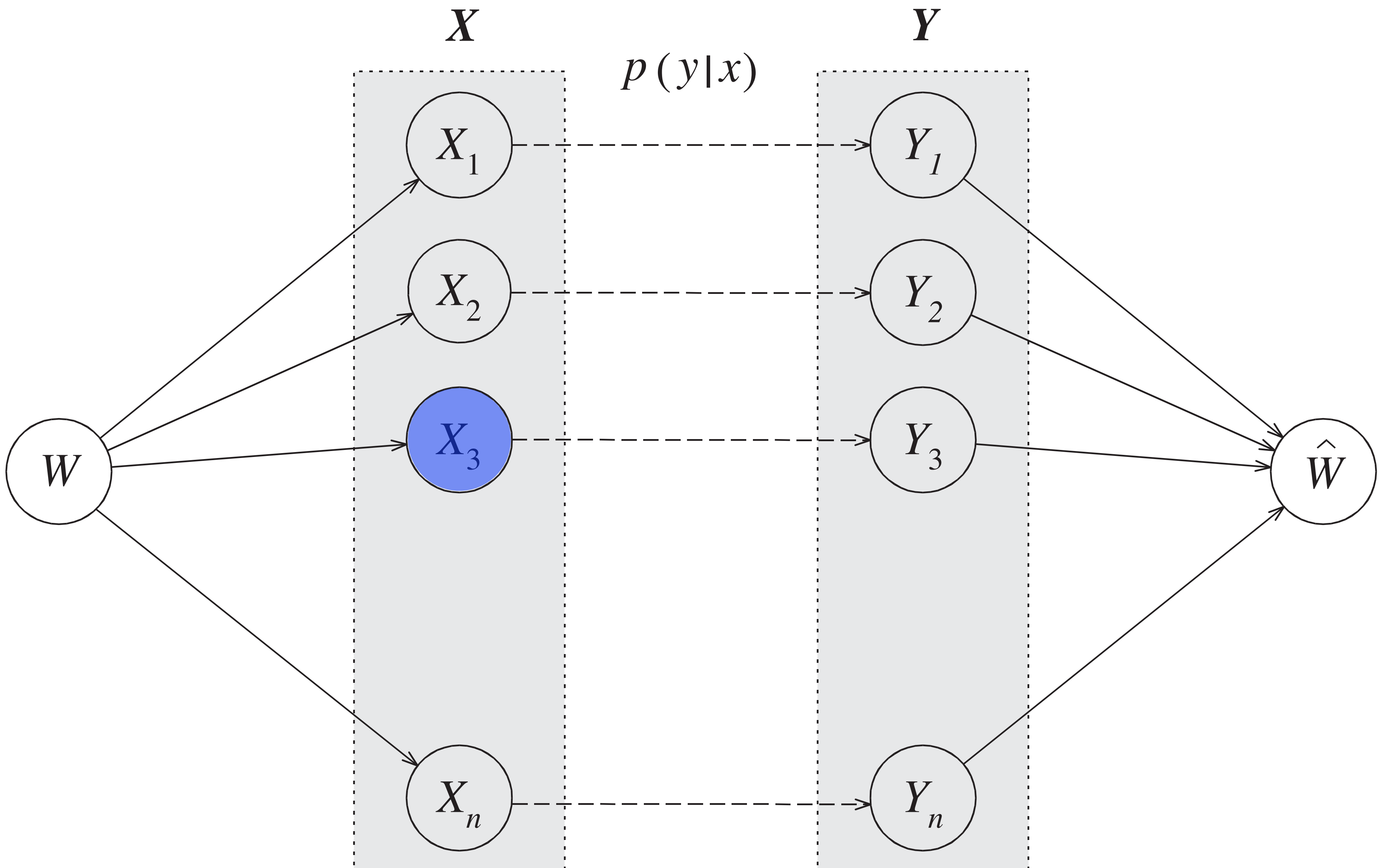


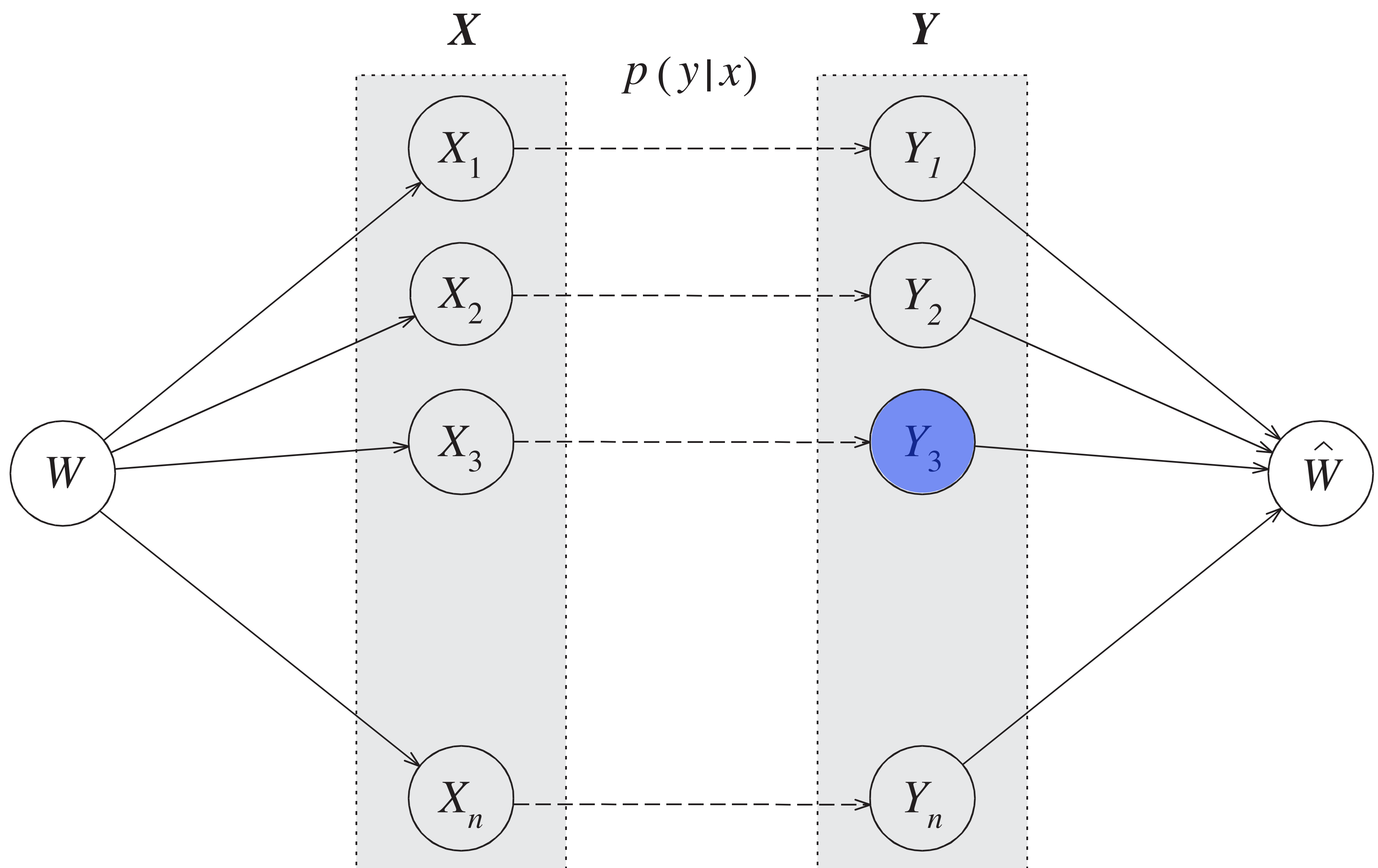
X
Y
p (y | x)
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
W
Ŵ

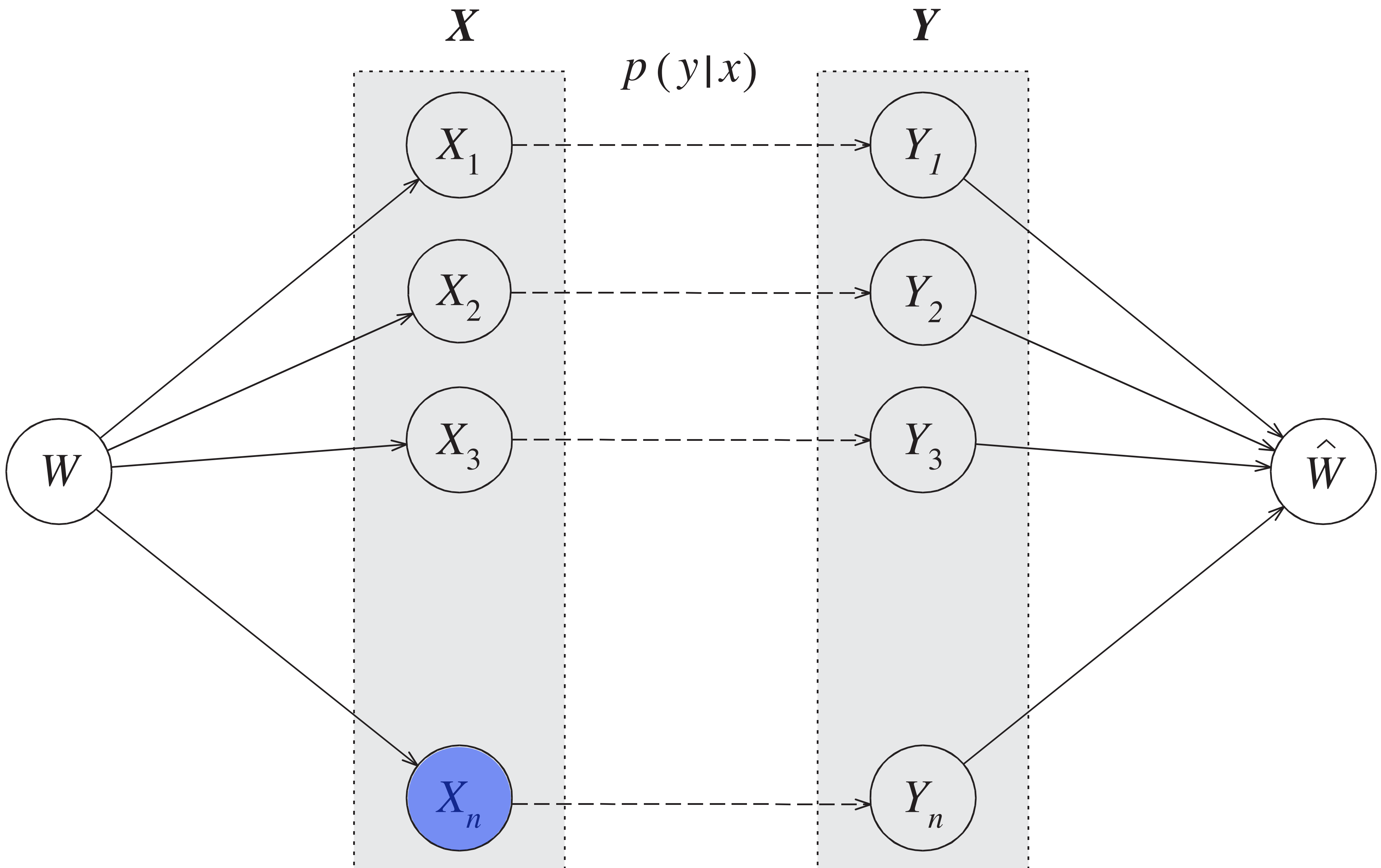

X
Y
p(y|x)
W
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
Ŵ

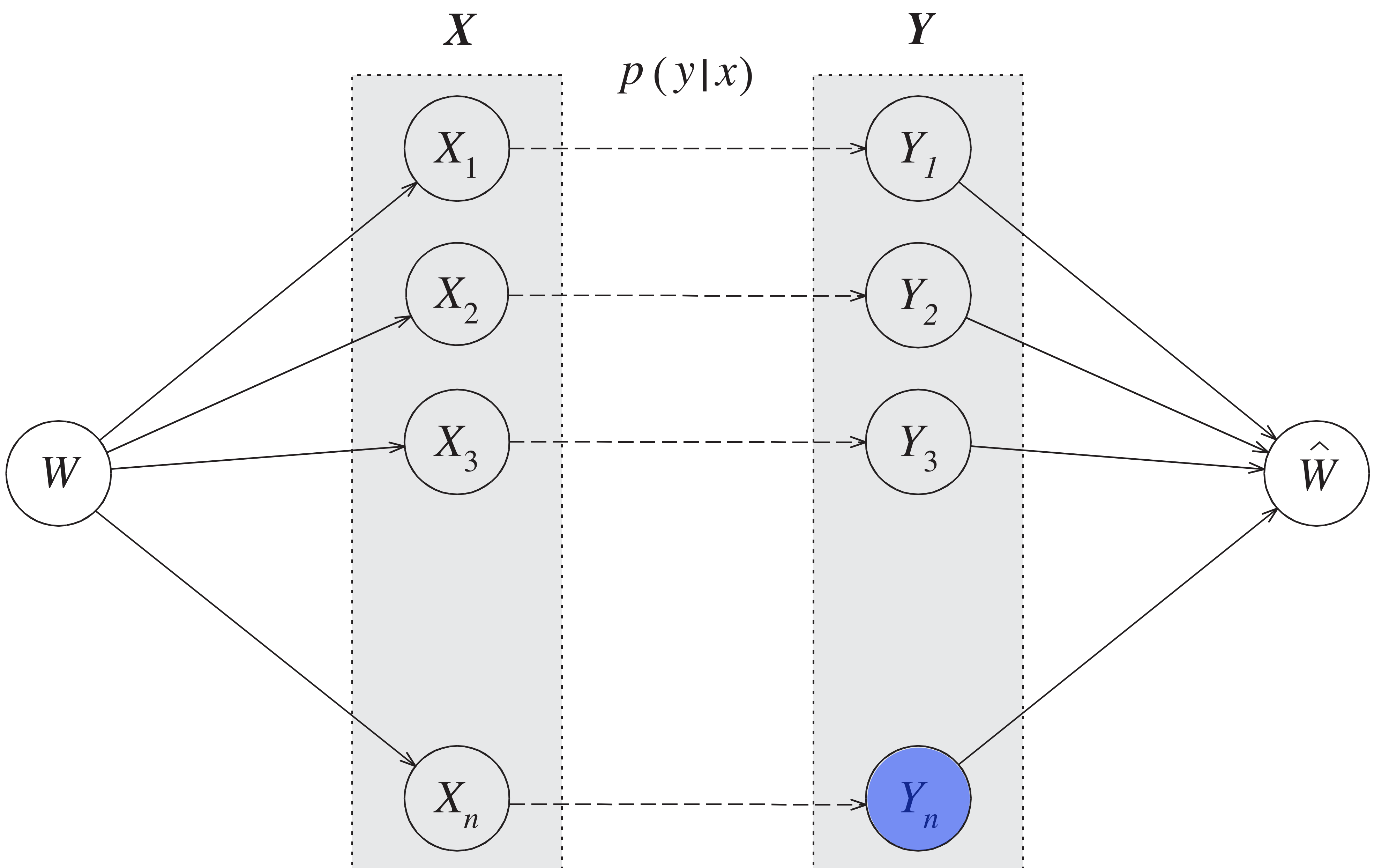


X
Y
p (y | x)
W
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
Ŵ

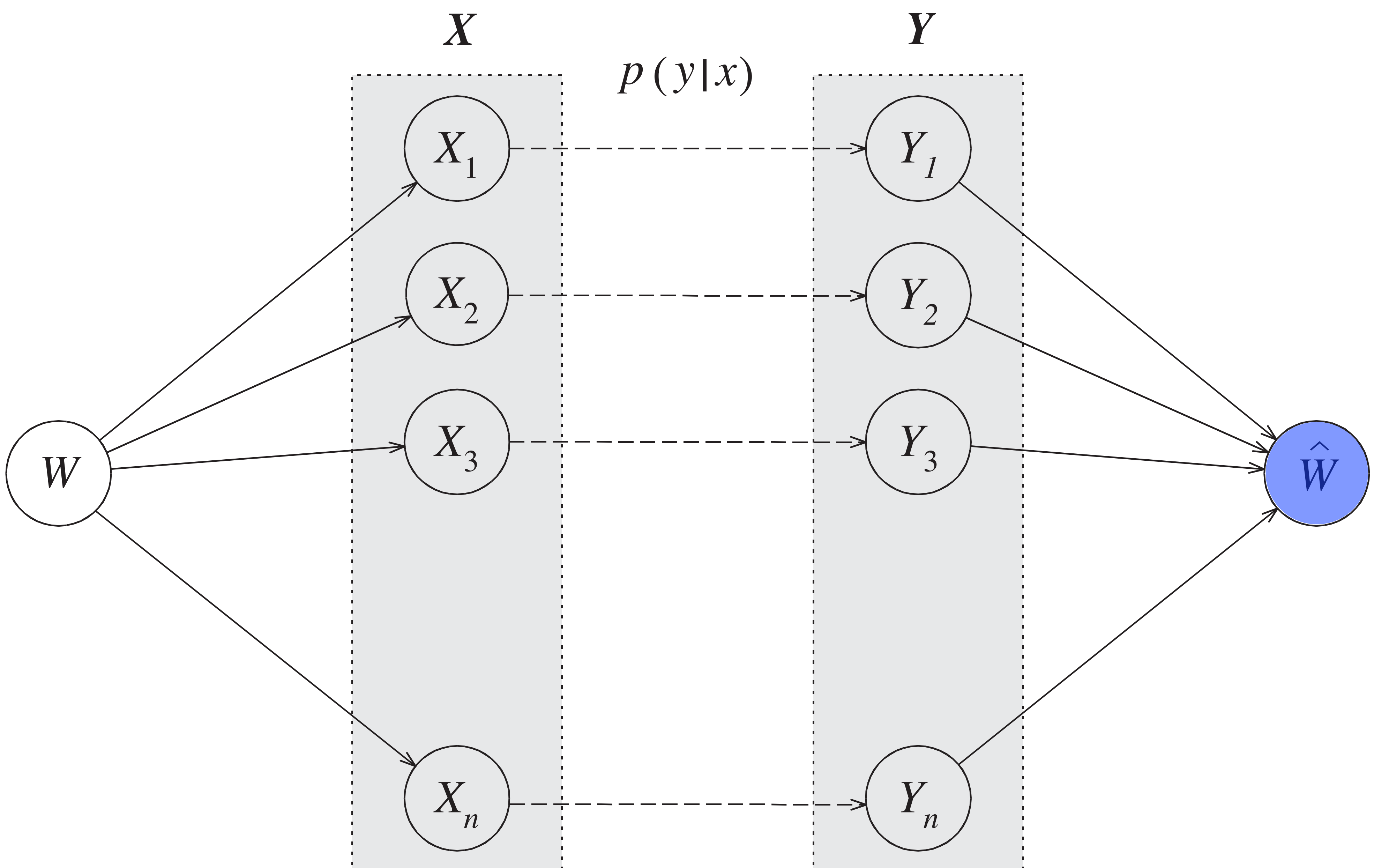

X
Y
p (y | x)
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
W
Ŵ


X
Y
p (y | x)
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
W
Ŵ

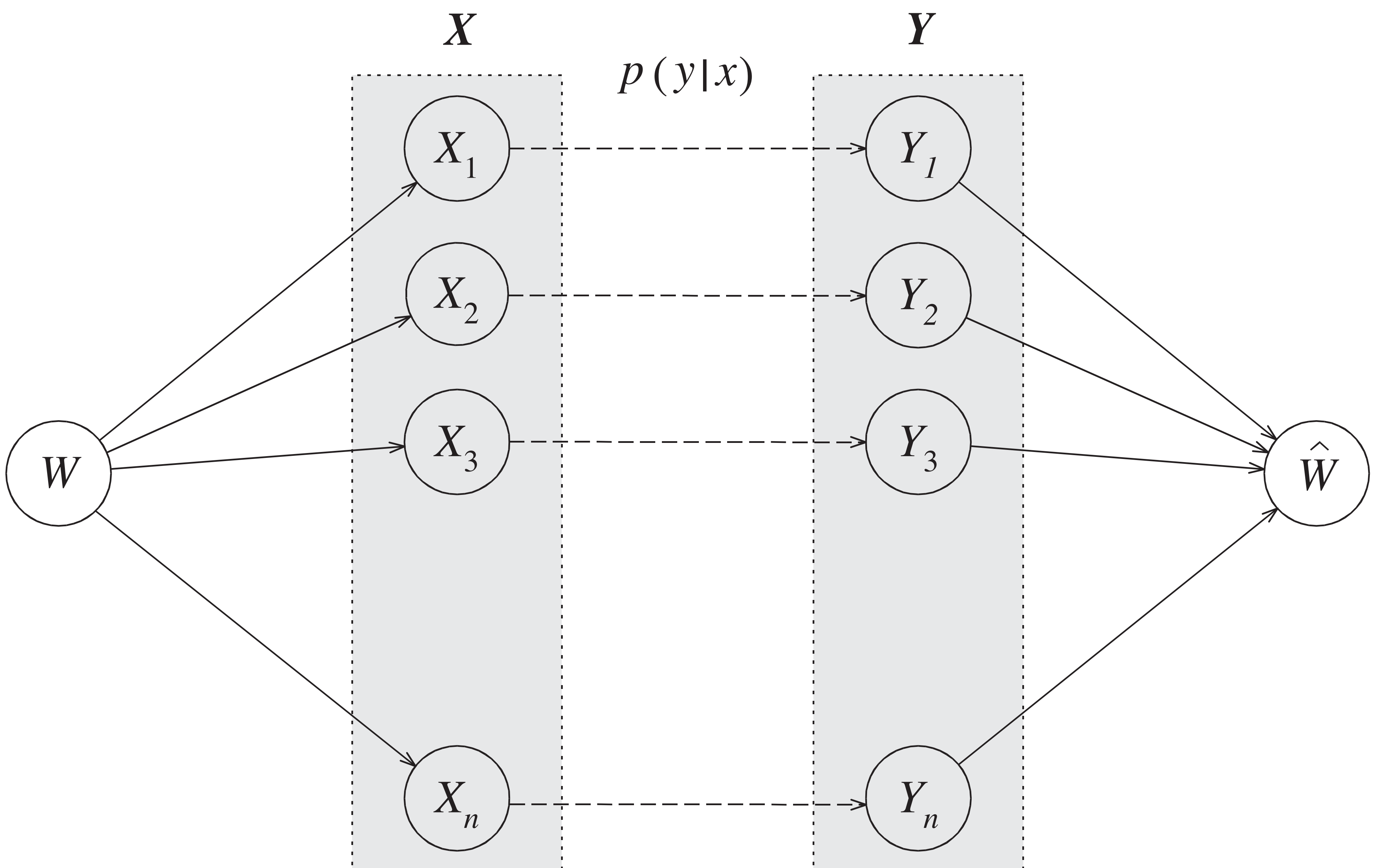

X
Y
p(y|x)
X_1
X_2
X_3
X_n
Y_1
Y_2
Y_3
Y_n
W
$\hat{W}$



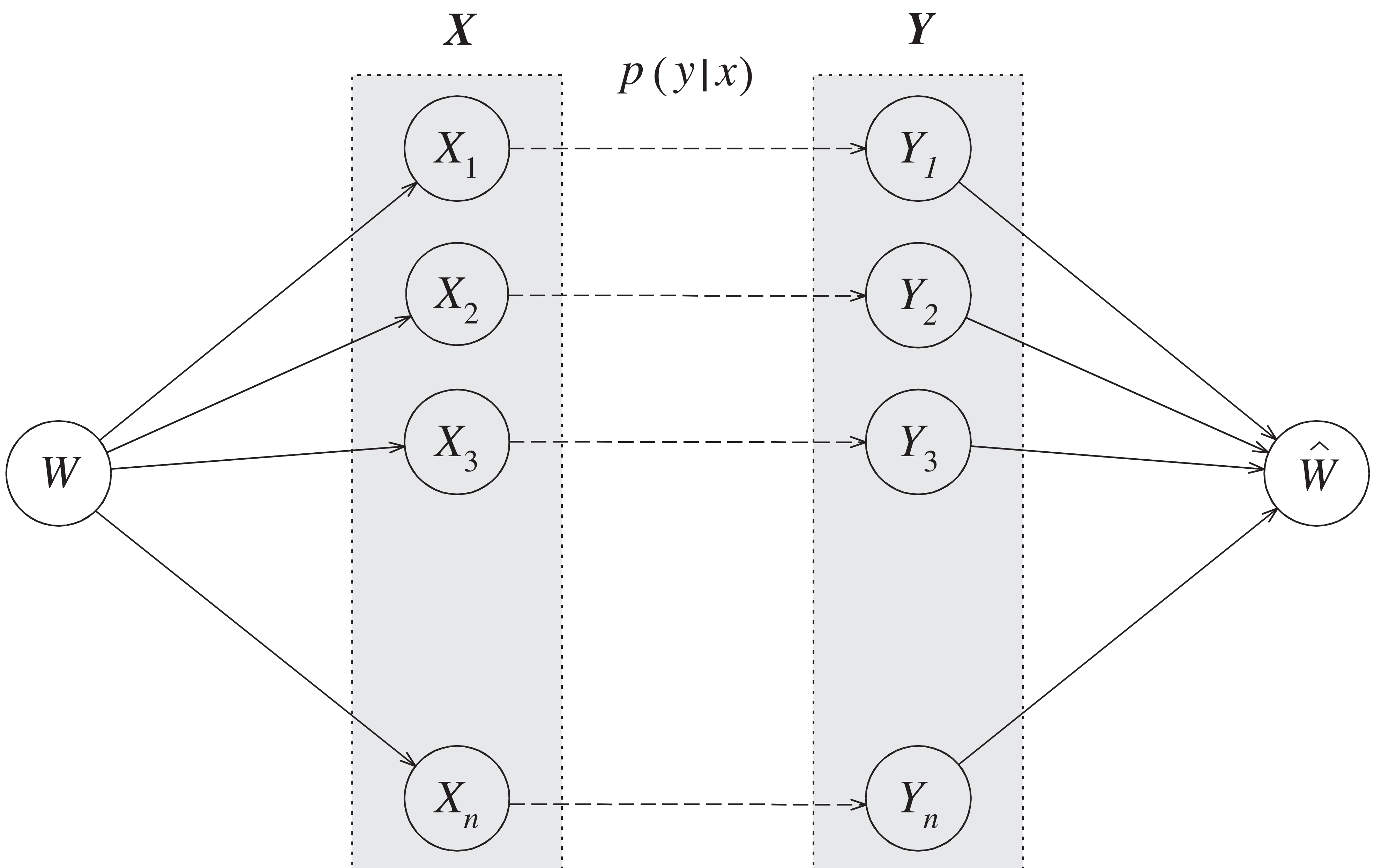
X
Y
p (y | x)
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
W
Ŵ



X
Y
p (y | x)
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
W
Ŵ

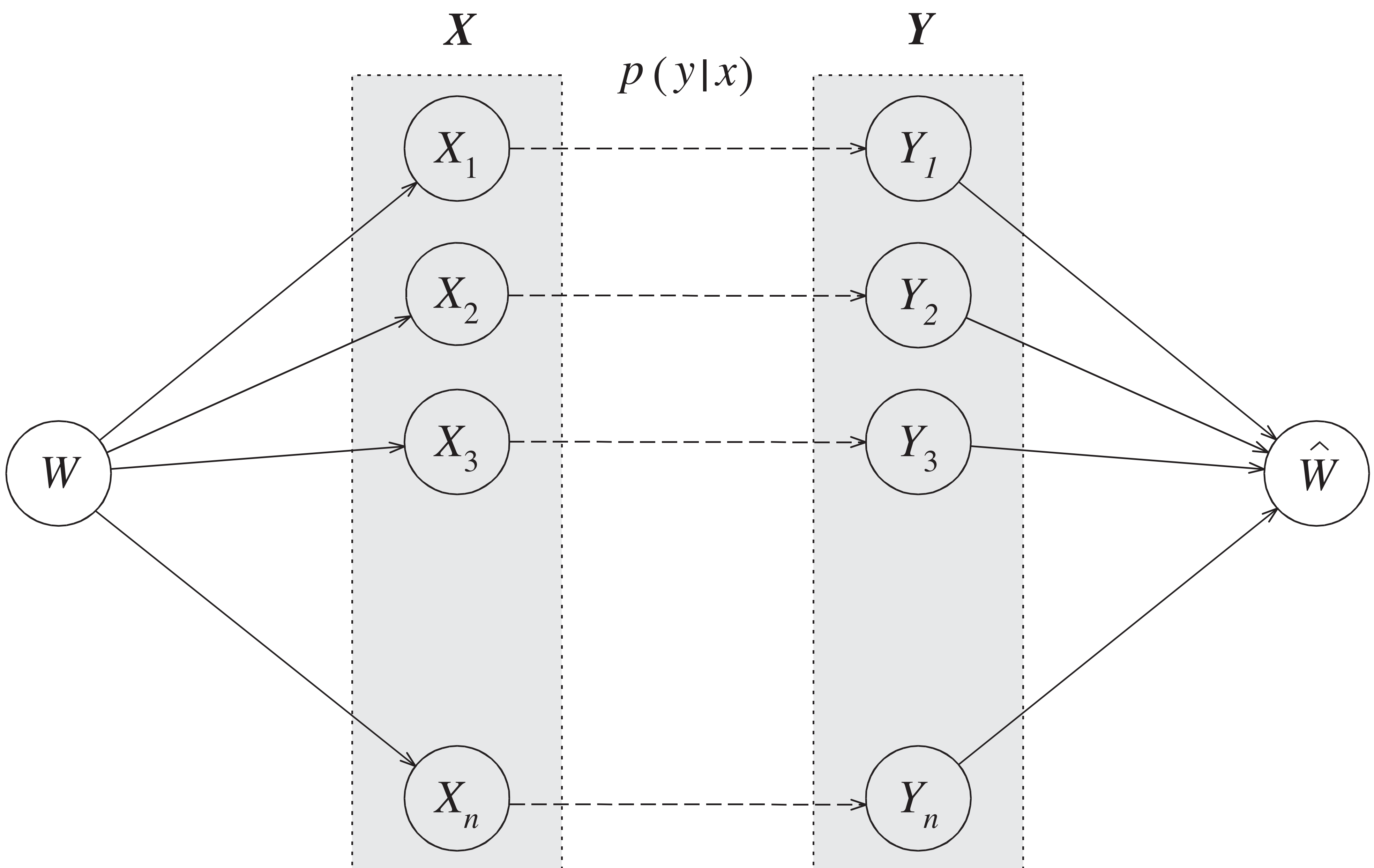


Let q denote the joint distribution for all the r.v.'s:



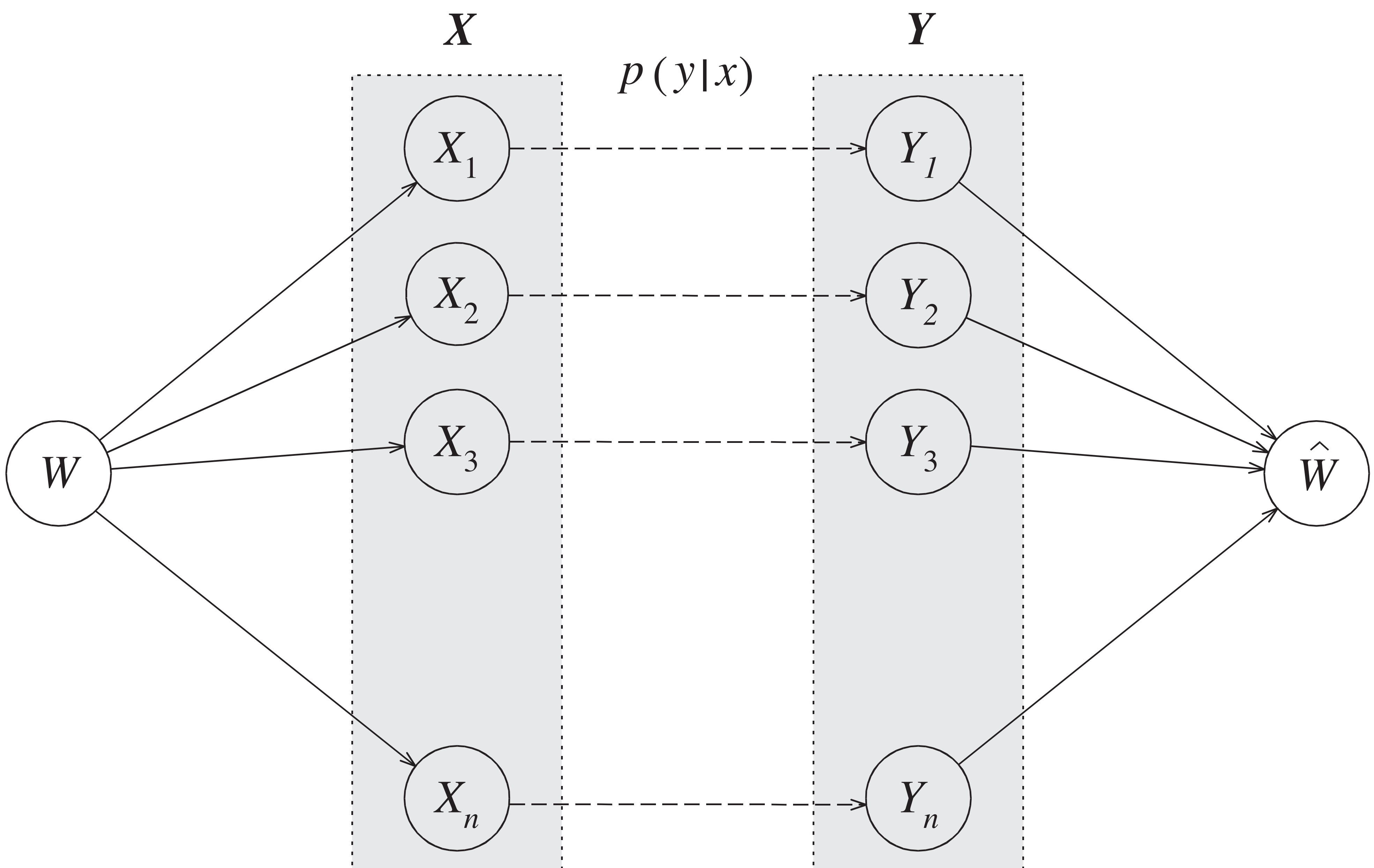
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|w, x_1)q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



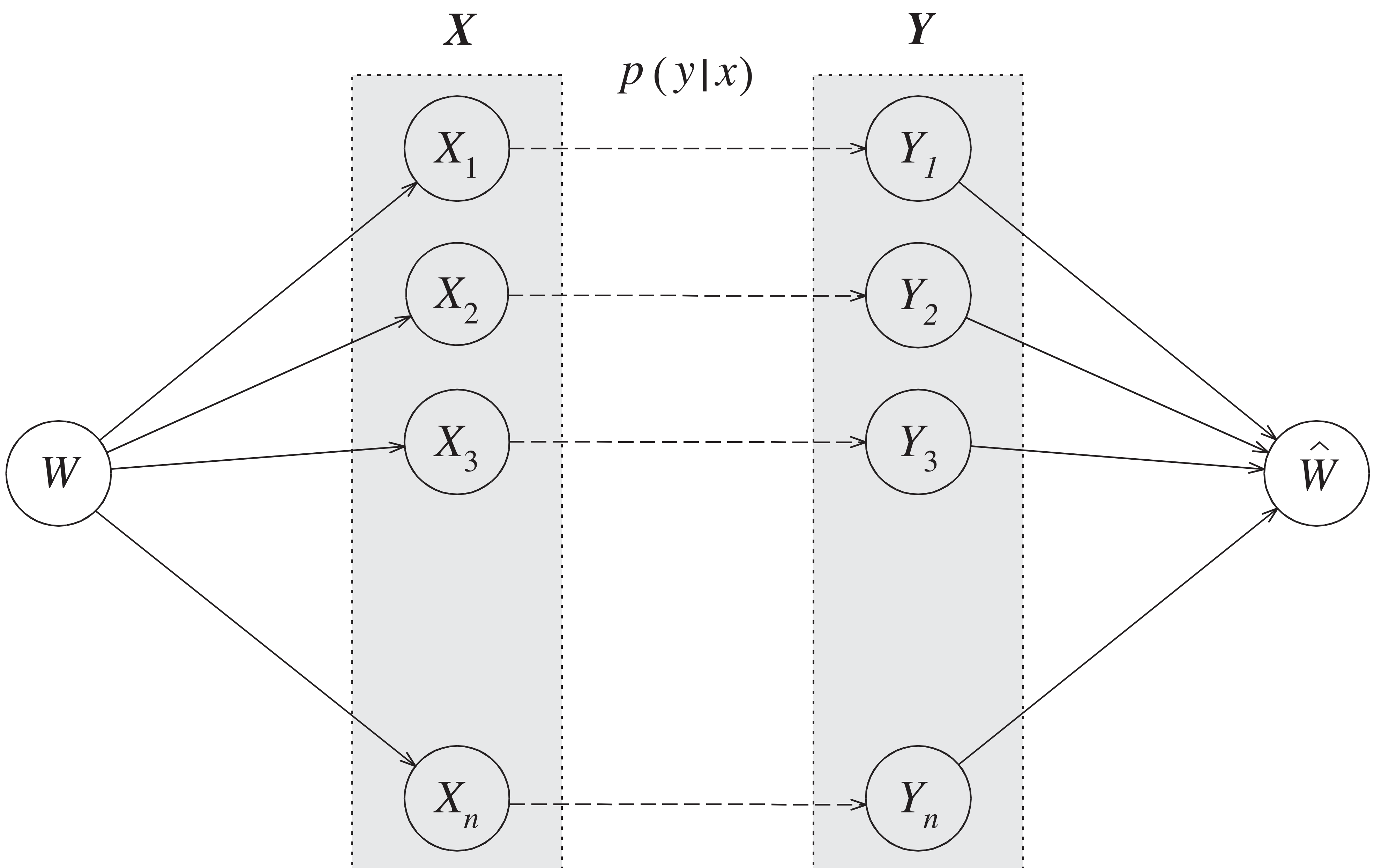
Let q denote the joint distribution for all the r.v.'s:

$$q(\underline{w,} x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(\underline{w})q(x_1|w)q(y_1|w, x_1)q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



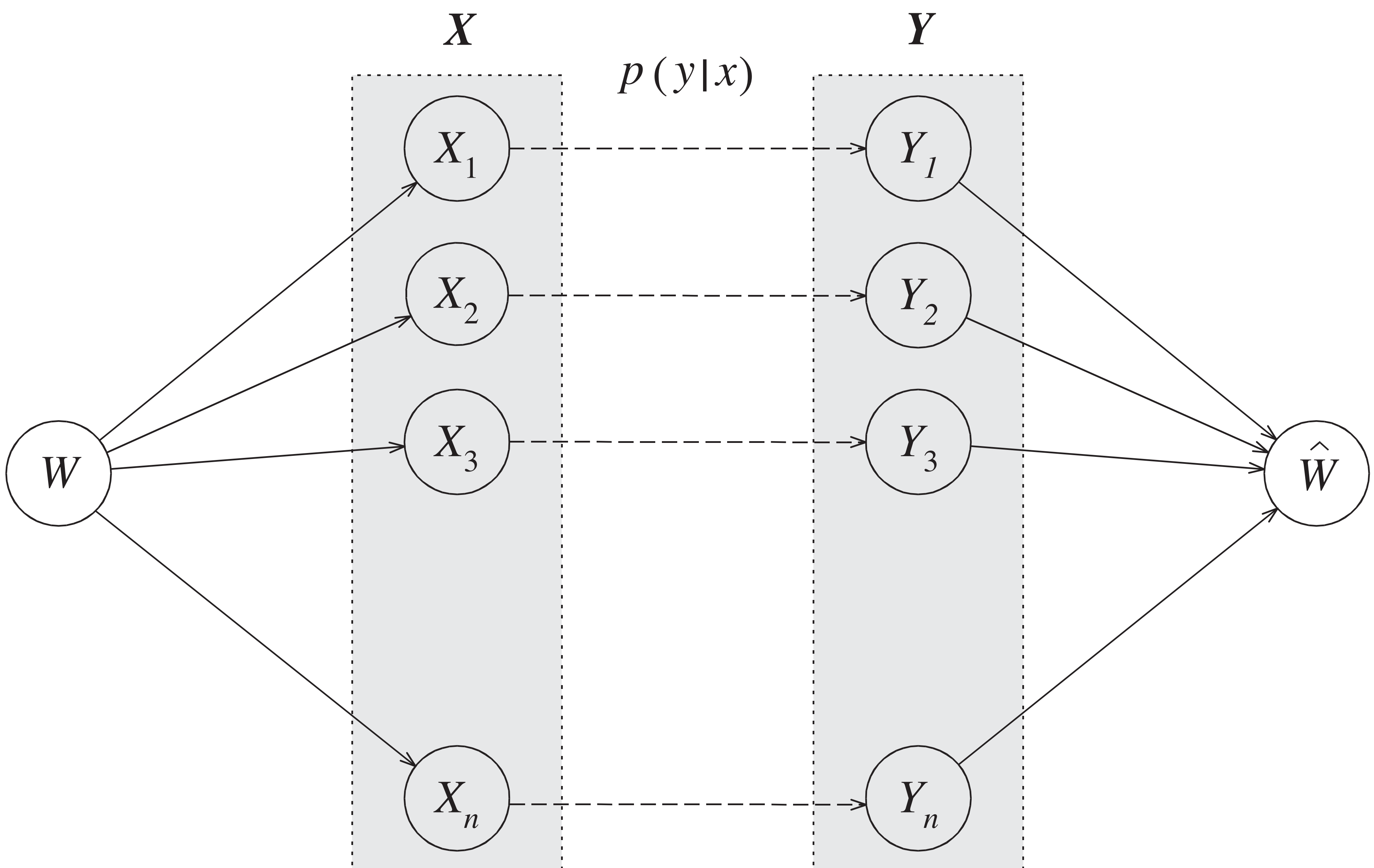
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|w, x_1)q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



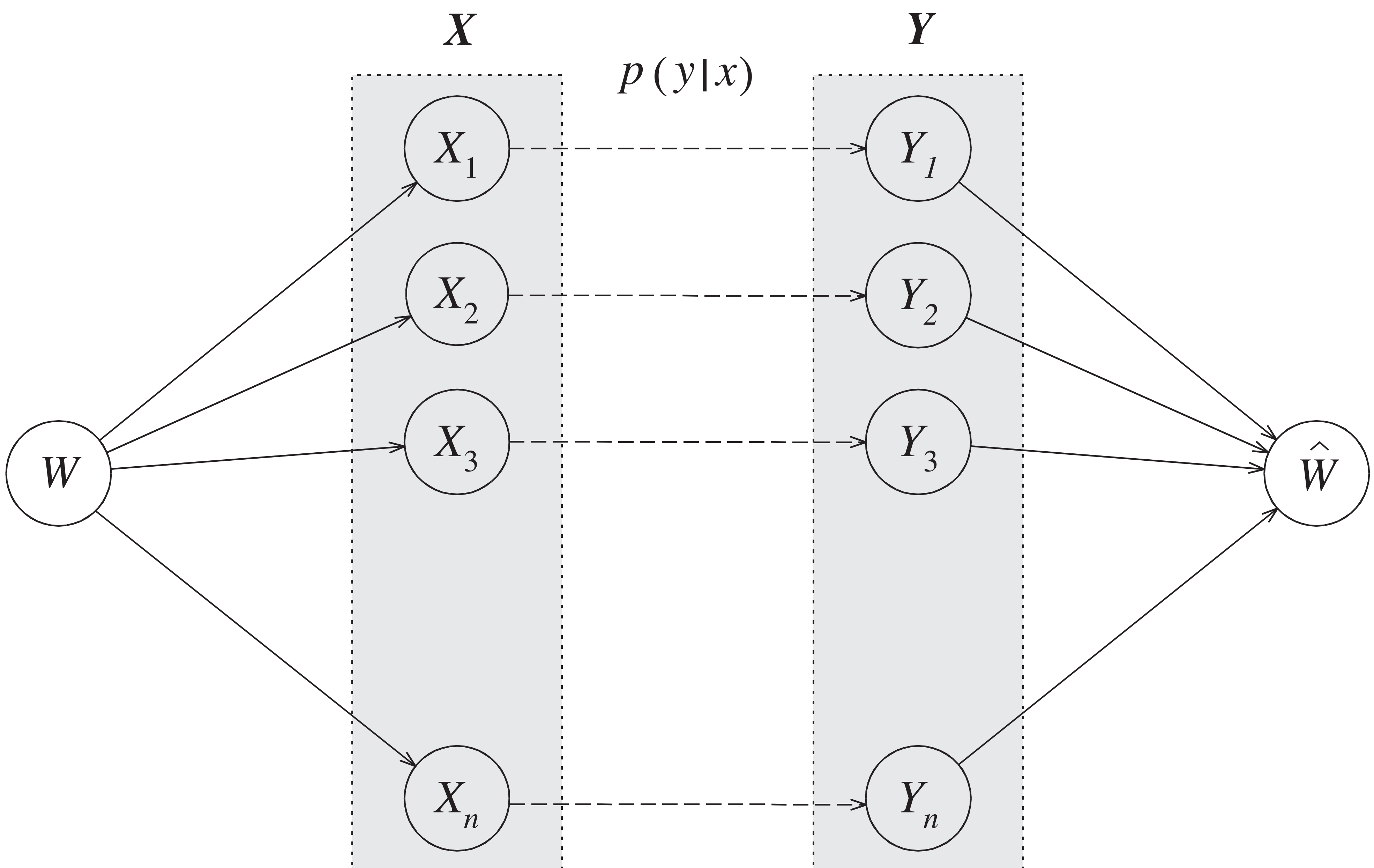
Let q denote the joint distribution for all the r.v.'s:

$$q(\underline{w,}\,\underline{x_1},y_1,x_2,y_2,\cdots,x_n,y_n,\hat{w}) \quad = \quad q(w)q(\underline{x_1}|\underline{w})q(y_1|w,x_1)q(x_2|w,x_1,y_1)q(y_2|w,x_1,y_1,x_2)\cdots$$



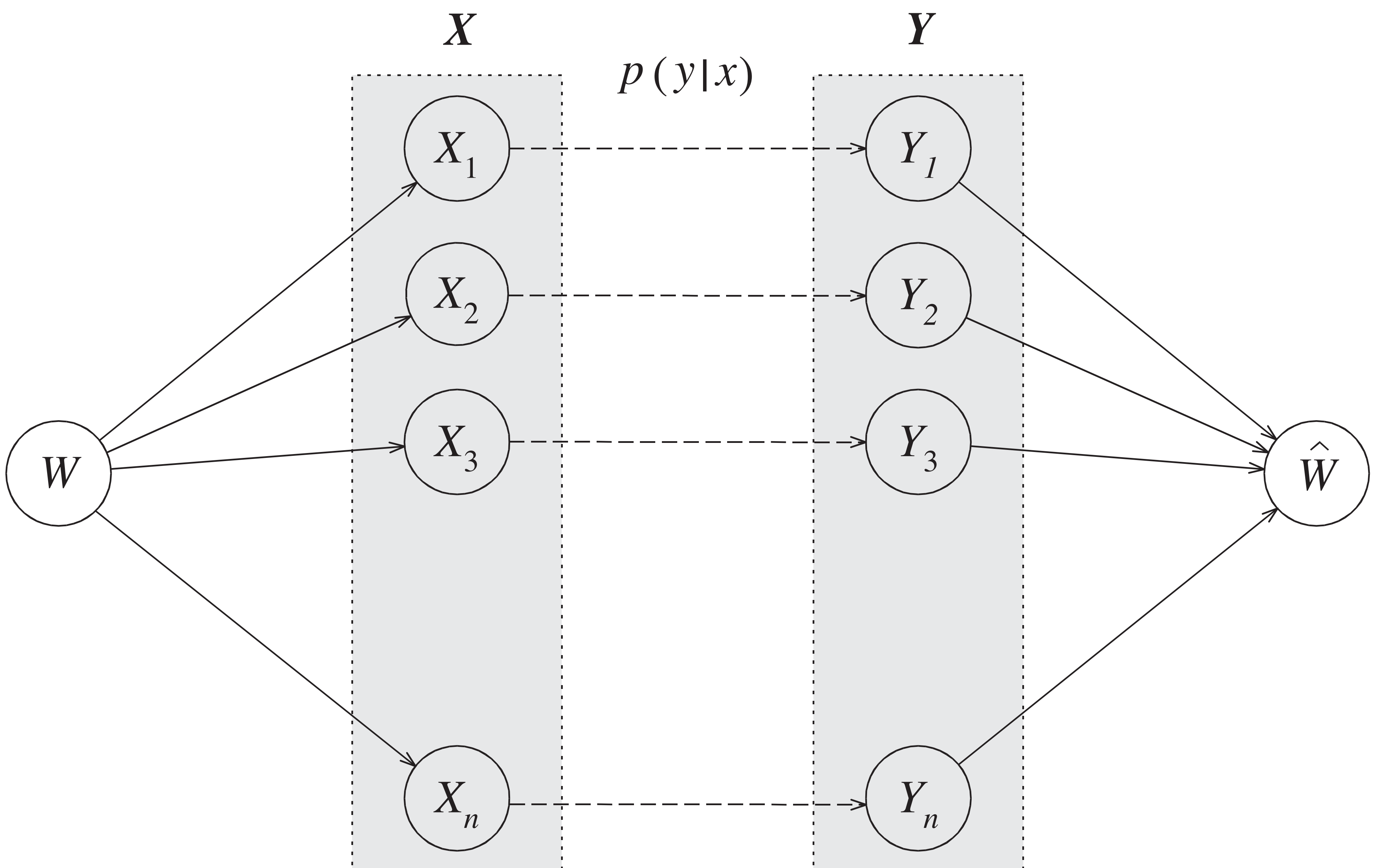
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, \underline{y_1}, x_2, y_2, \cdots, x_n, y_n, \hat{w}) = q(w)q(x_1|w)q(\underline{y_1}|w, x_1)q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



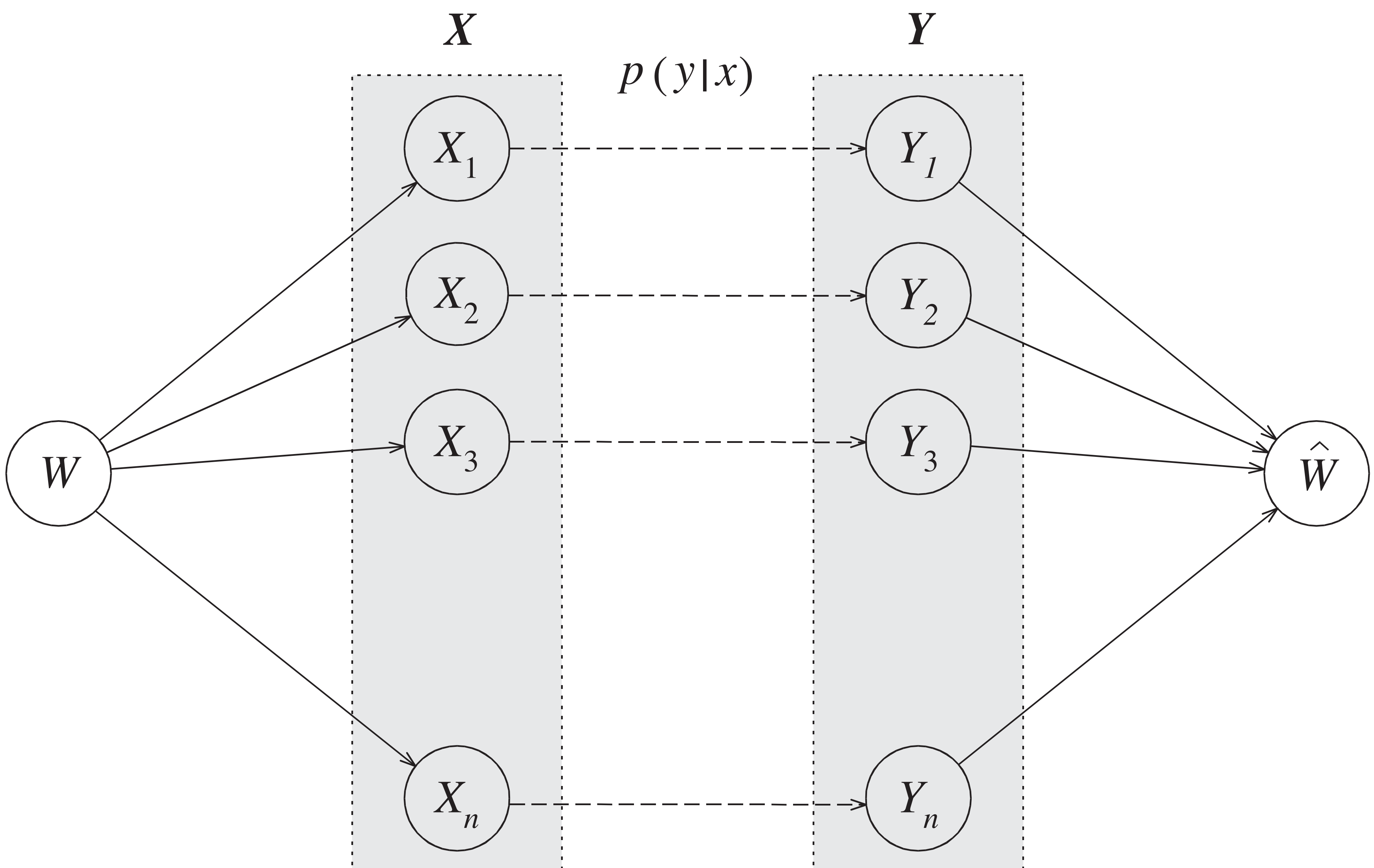
Let q denote the joint distribution for all the r.v.'s:

$$q(\underline{w, x_1,} \underline{y_1}, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(\underline{y_1}|\underline{w, x_1})q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



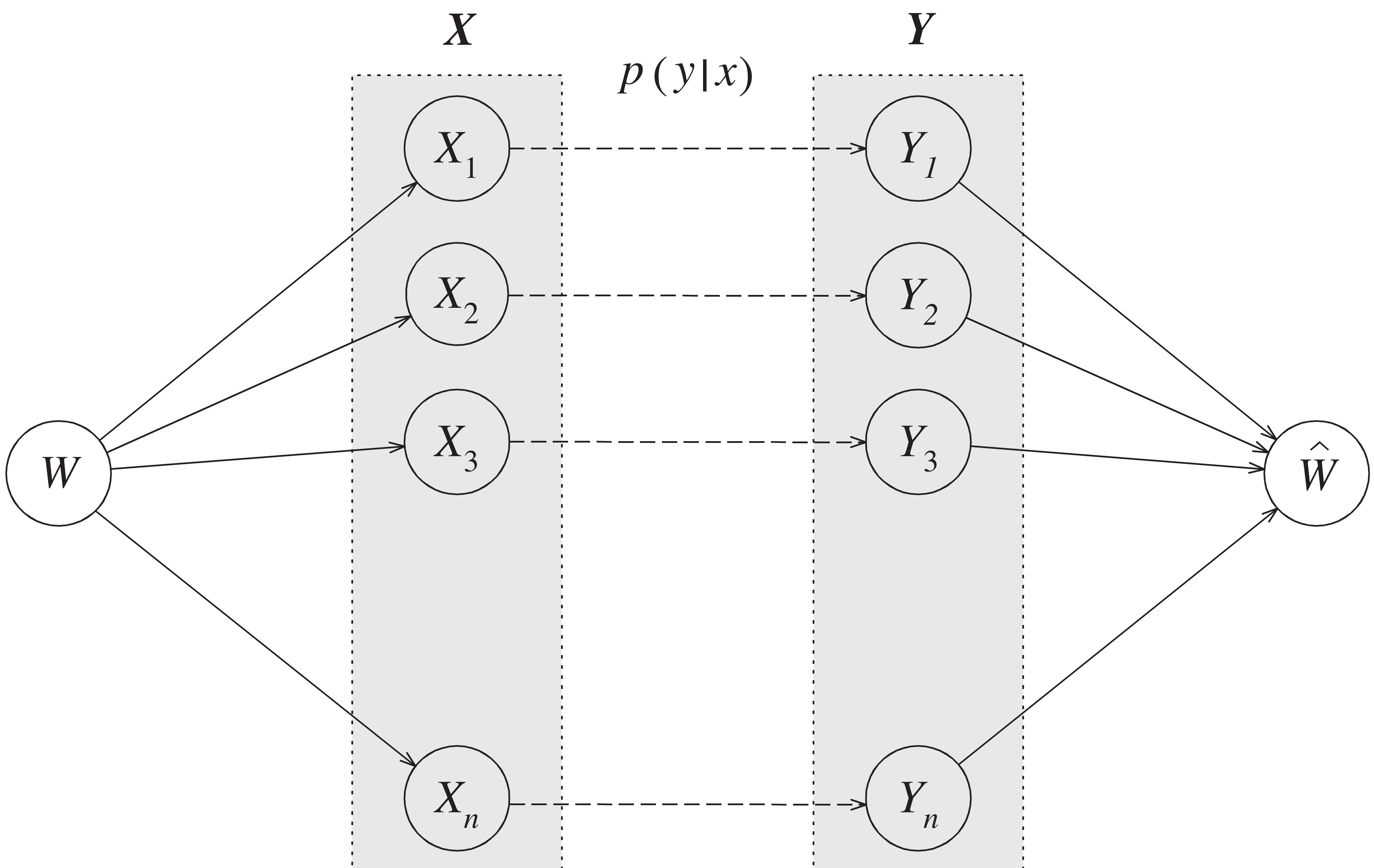
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, \underline{x_2}, y_2, \cdots, x_n, y_n, \hat{w}) = q(w)q(x_1|w)q(y_1|w, x_1)q(\underline{x_2}|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



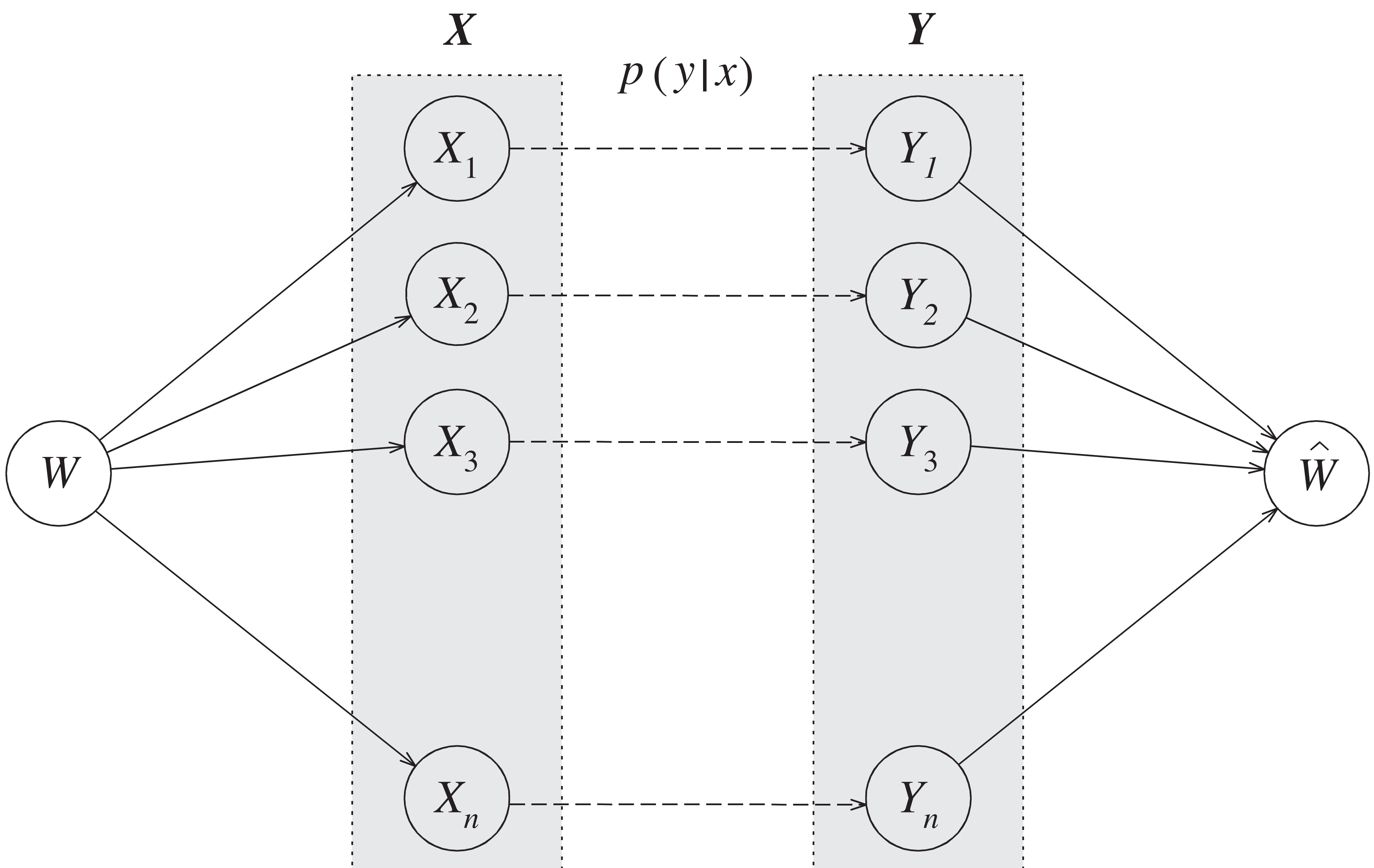
Let q denote the joint distribution for all the r.v.'s:

$$q(\underline{w, x_1, y_1}, \underline{x_2}, y_2, \cdots, x_n, y_n, \hat{w}) = q(w)q(x_1|w)q(y_1|w, x_1)q(\underline{x_2}|\underline{w, x_1, y_1})q(y_2|w, x_1, y_1, x_2)\cdots$$



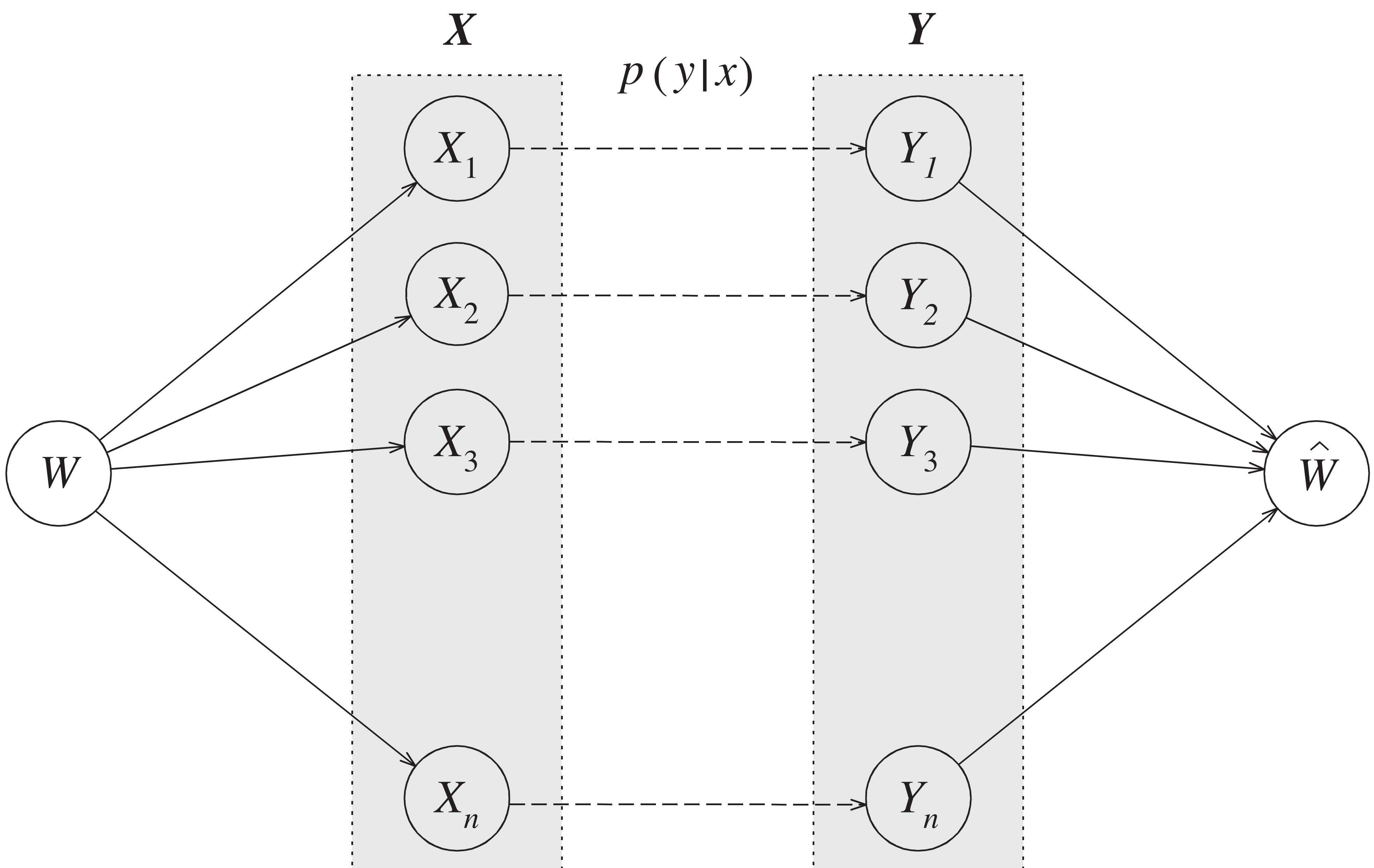
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|w, x_1)q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



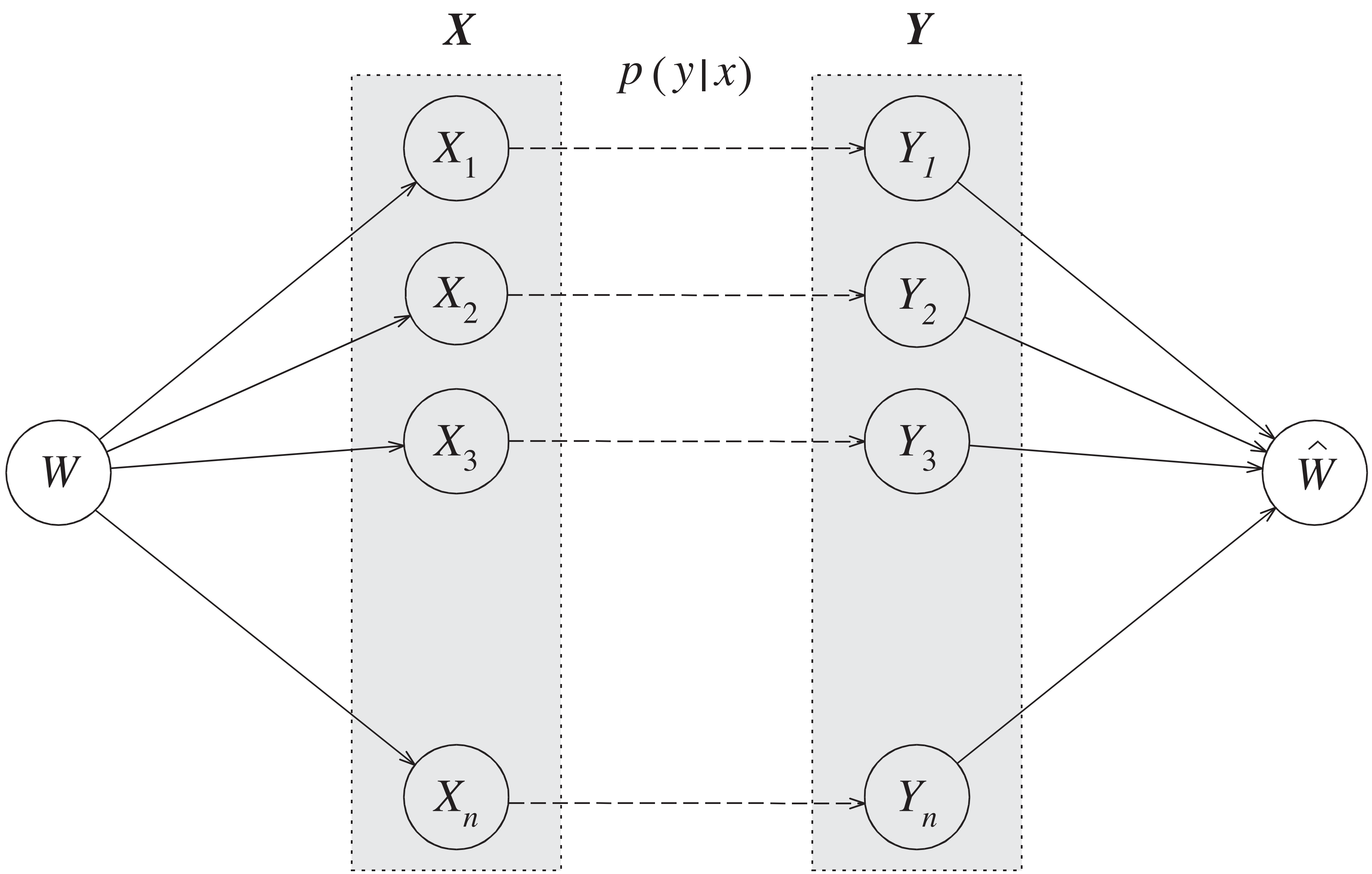
Let q denote the joint distribution for all the r.v.'s:

$$q(\underline{w, x_1, y_1, x_2}, \underline{y_2}, \cdots, x_n, y_n, \hat{w}) = q(w)q(x_1|w)q(y_1|w, x_1)q(x_2|w, x_1, y_1)q(\underline{y_2}|\underline{w, x_1, y_1, x_2}) \cdots$$



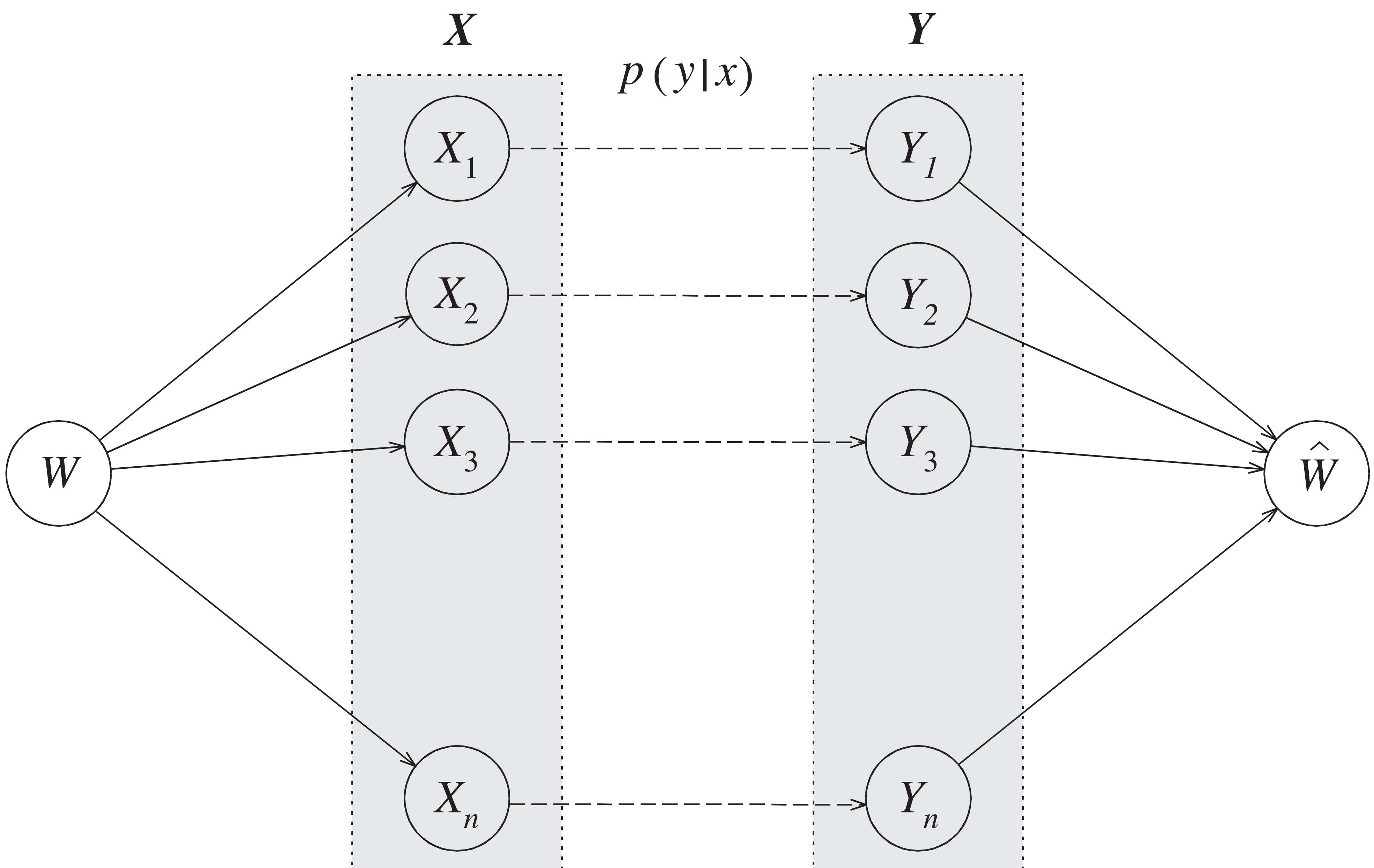
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) = q(w)q(x_1|w)q(y_1|w, x_1)q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



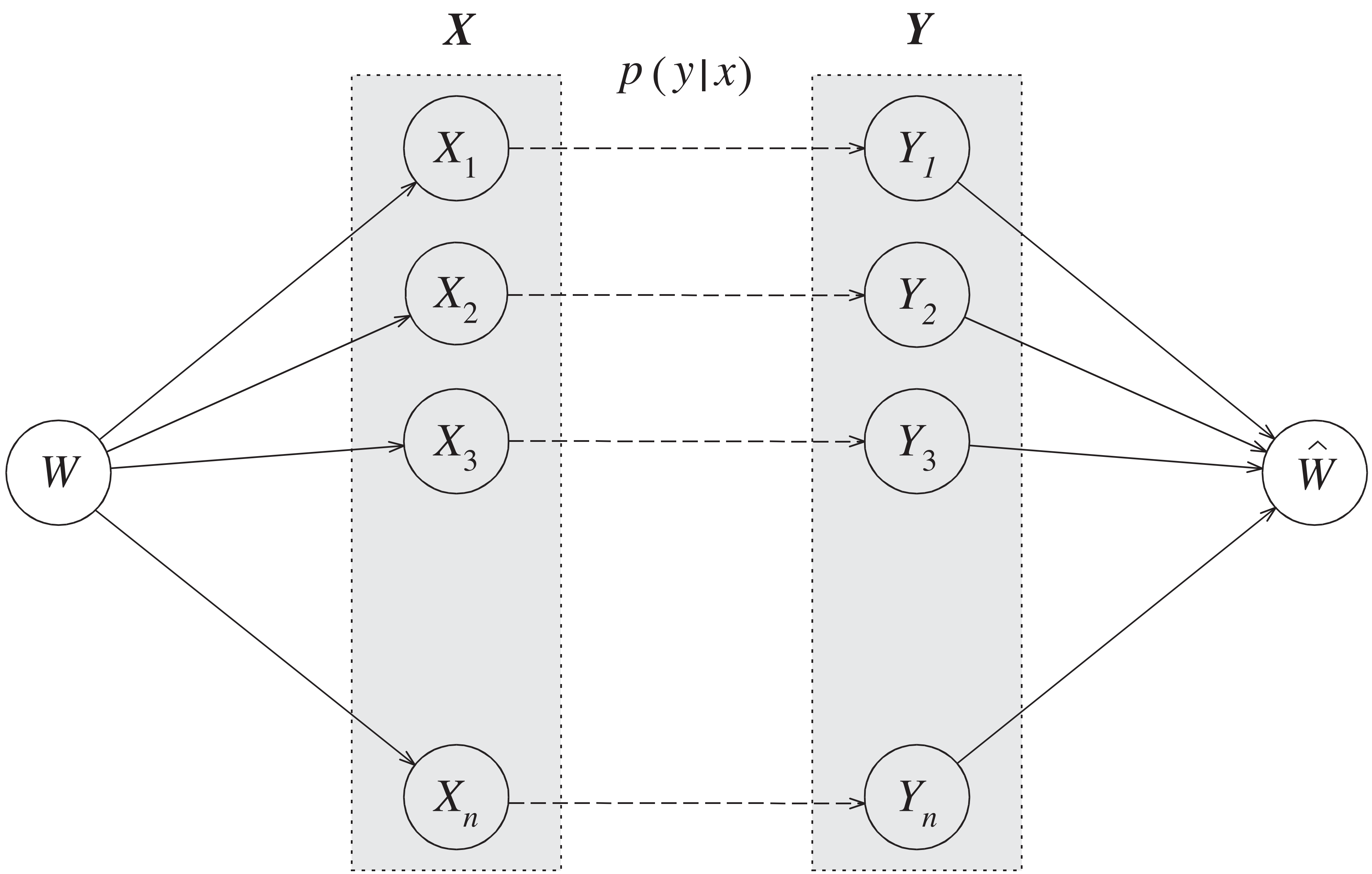
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|w, x_1)q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



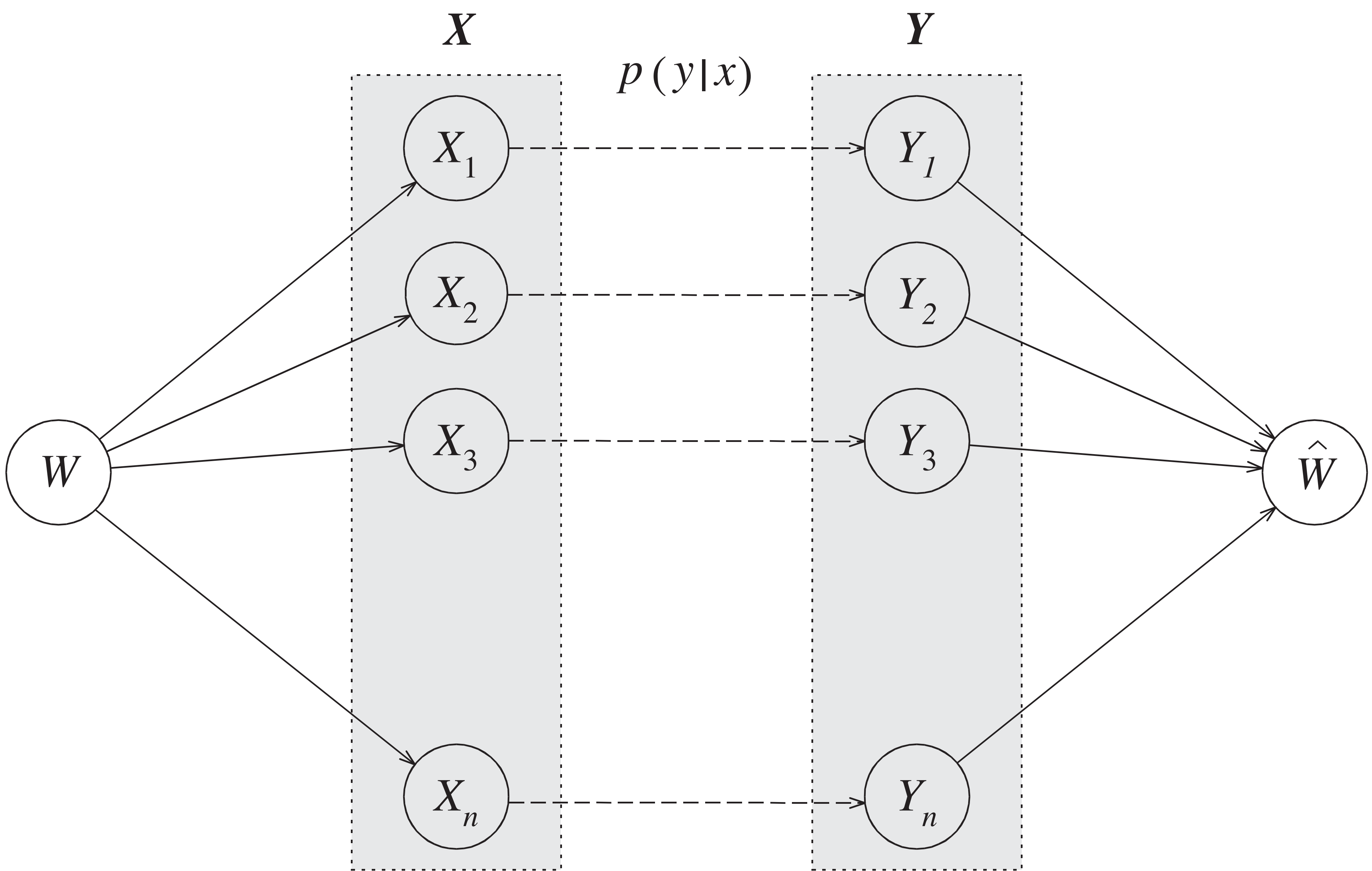
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, x_1, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



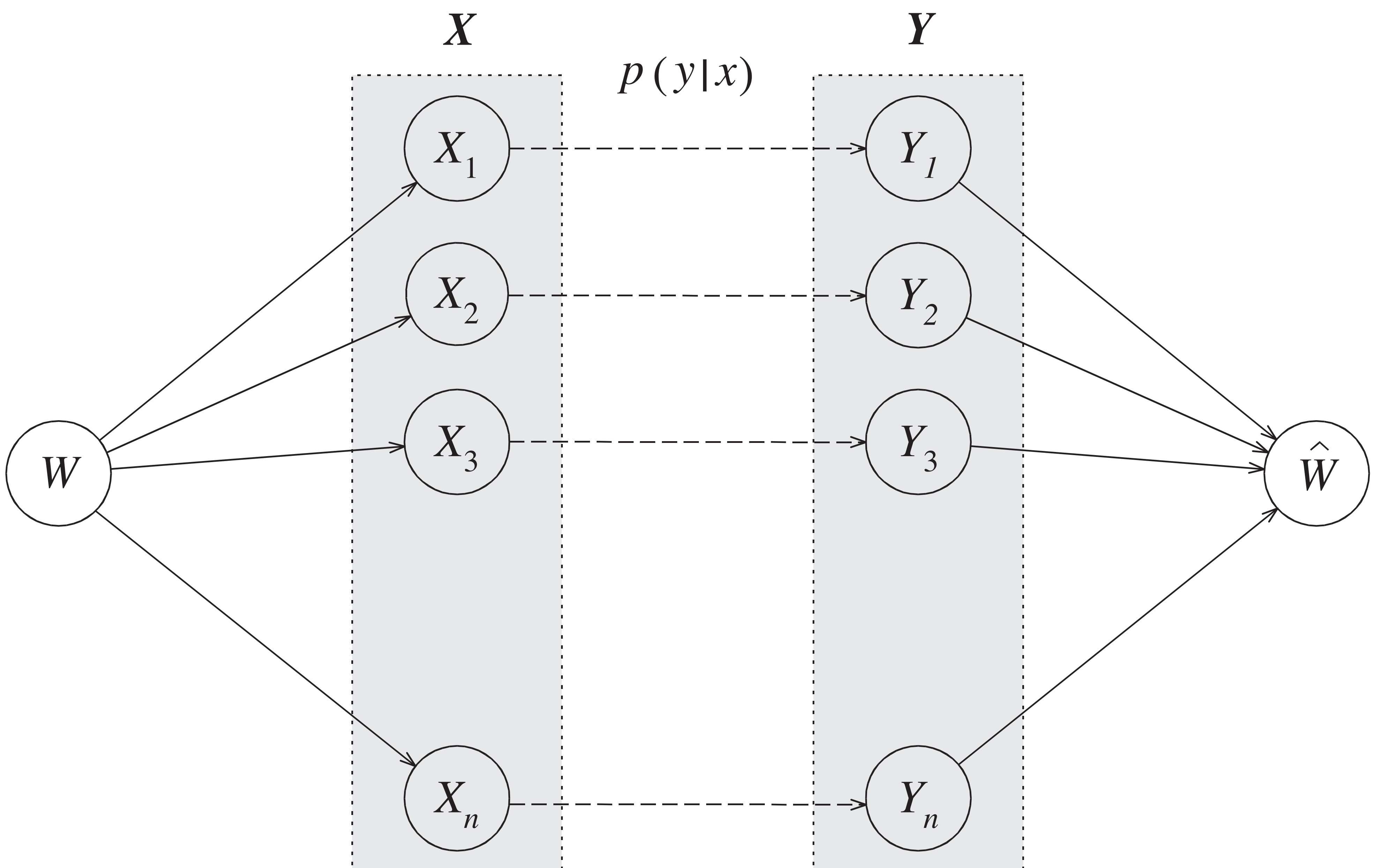
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, y_1)q(y_2|w, x_1, y_1, x_2)\cdots$$



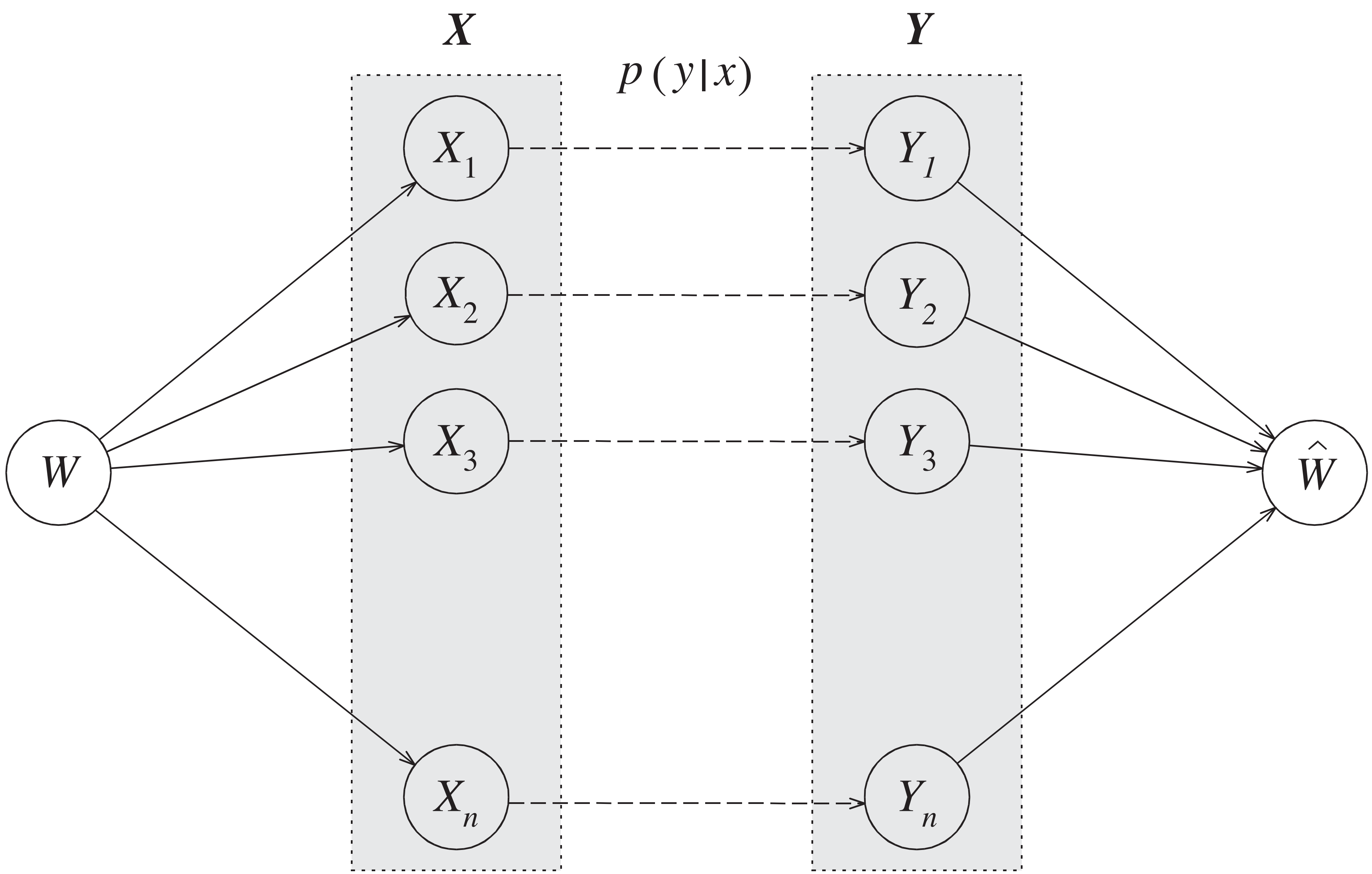
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|w, x_1, y_1, x_2)\cdots$$



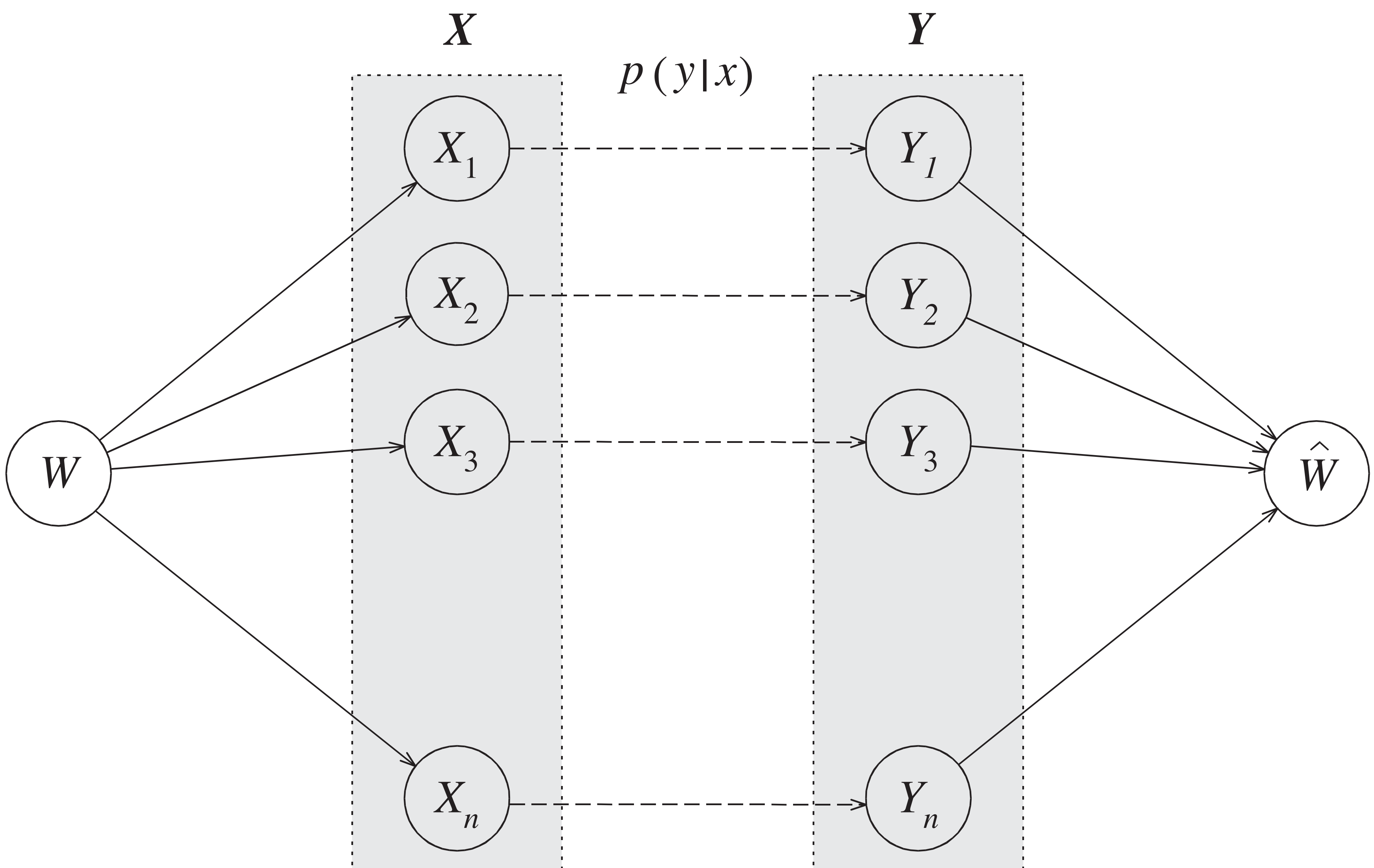
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|w, x_1, y_1, x_2)\cdots$$



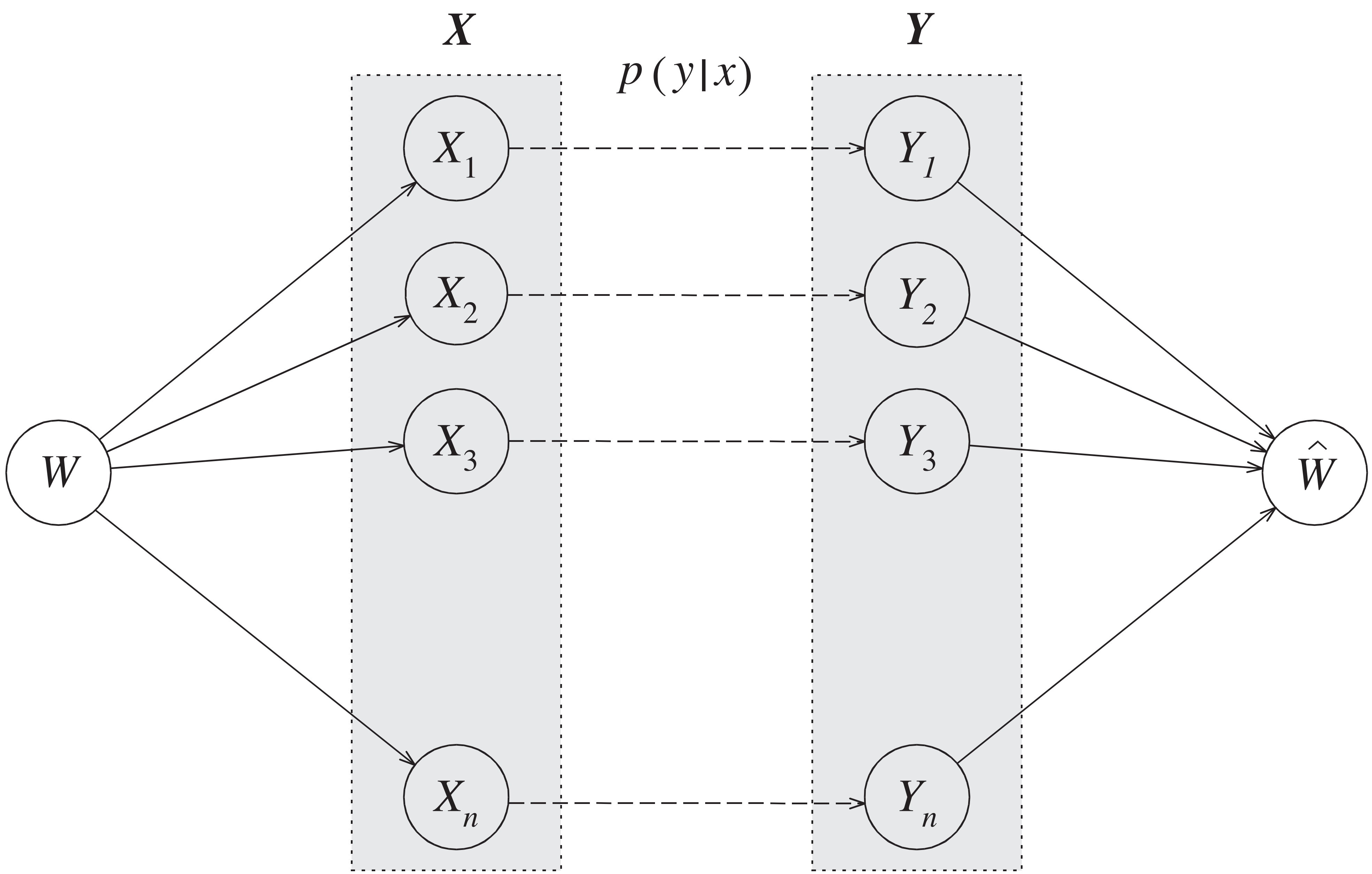
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, x_1, y_1, x_2)\cdots$$



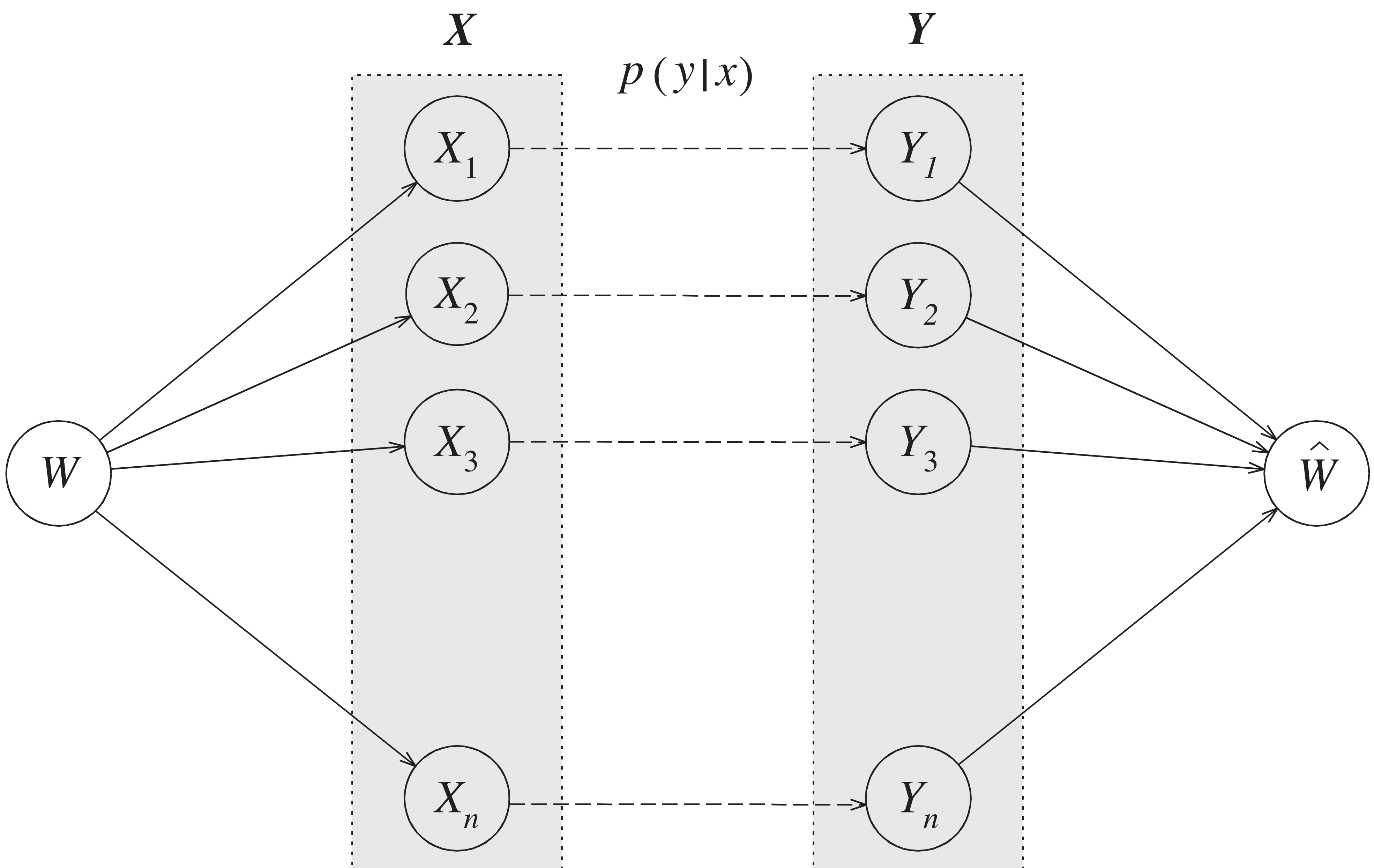
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) \quad = \quad q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, y_1, x_2)\cdots$$



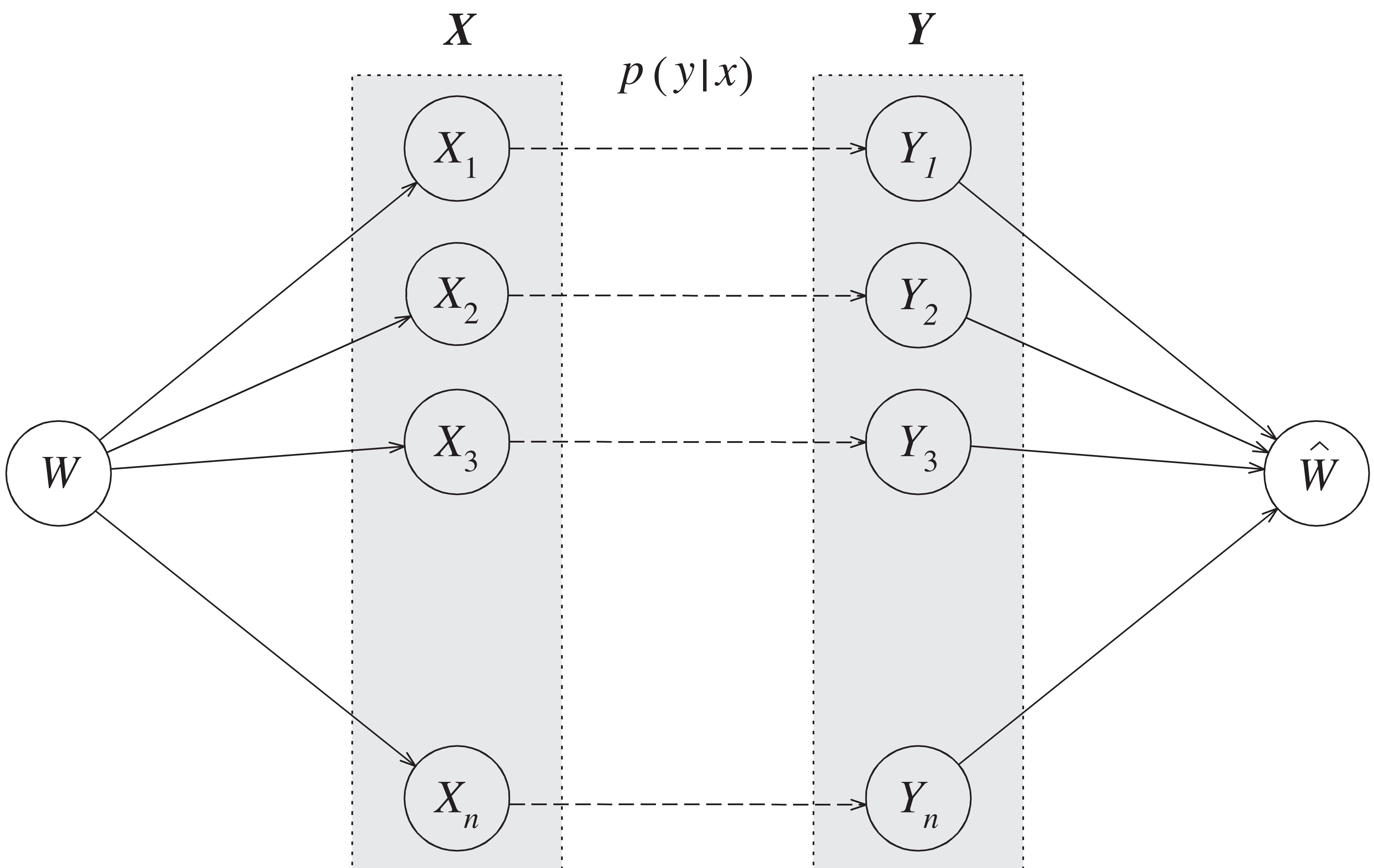
Let q denote the joint distribution for all the r.v.'s:

$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) = q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots$$



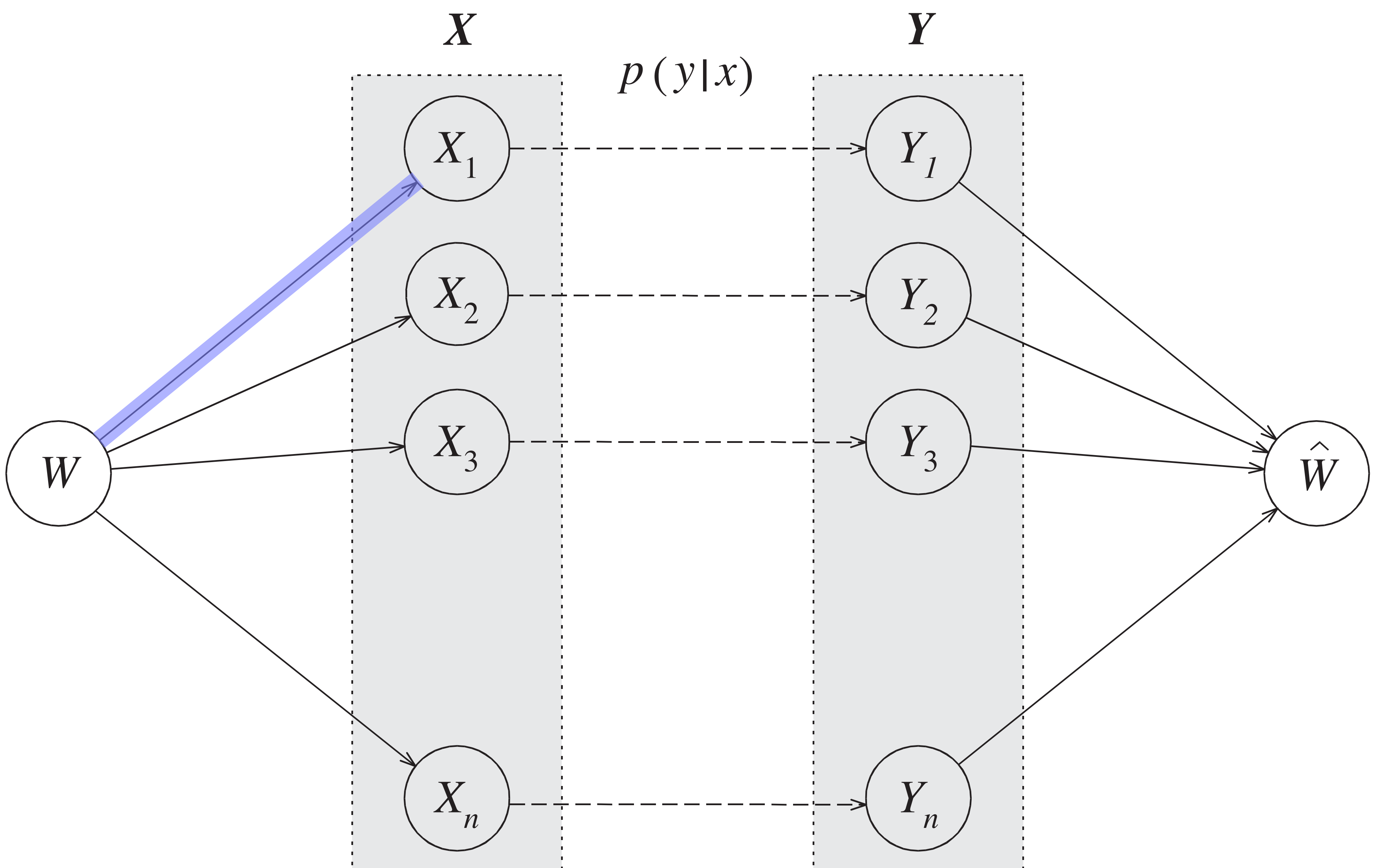
Let q denote the joint distribution for all the r.v.'s:

$$\begin{aligned} q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\ &= q(w)q(x_1|w)p(y_1|x_1)q(x_2|w)p(y_2|x_2)\cdots \end{aligned}$$



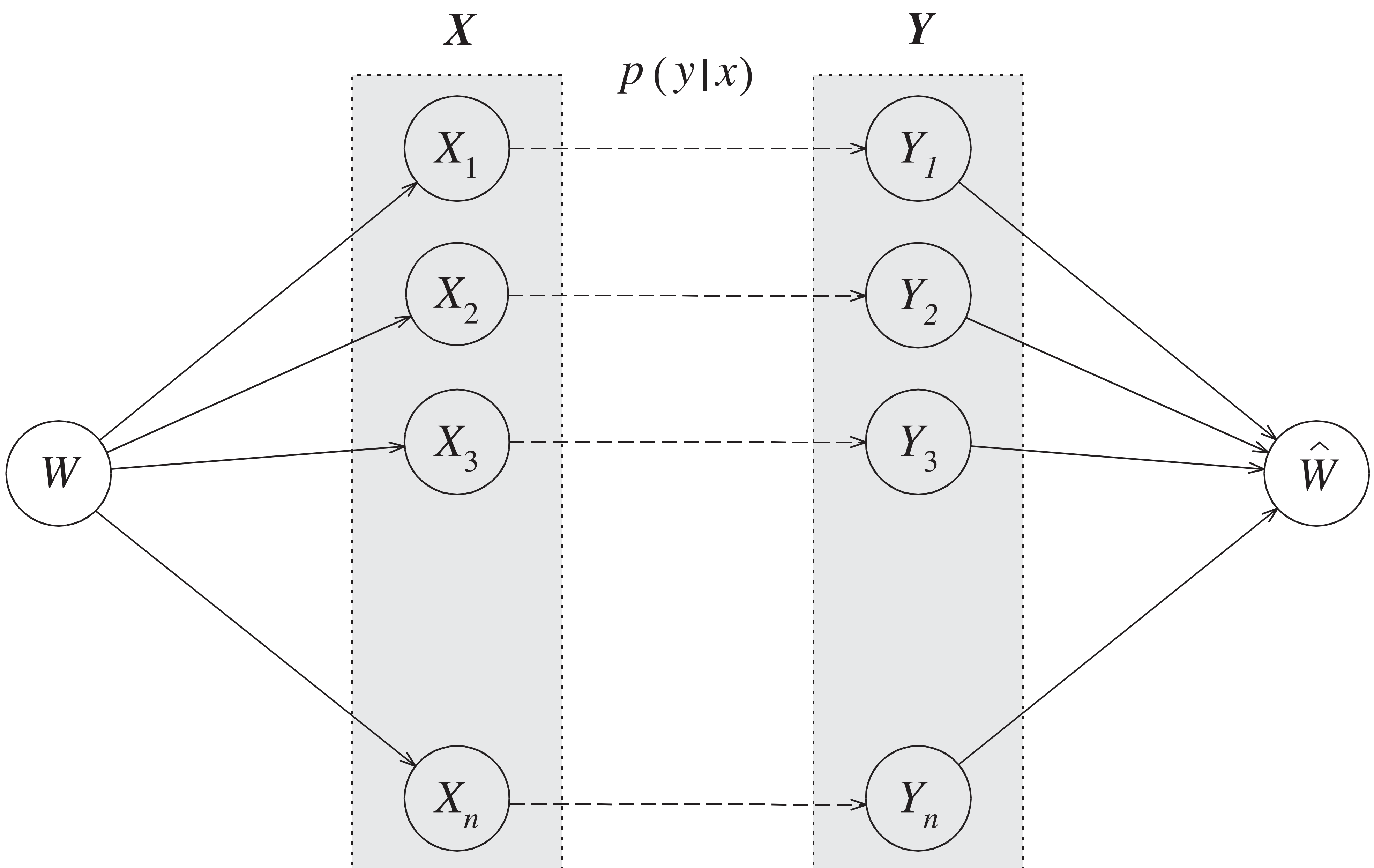
Let q denote the joint distribution for all the r.v.'s:

$$\begin{aligned} q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\ &= q(w)\underline{q(x_1|w)}p(y_1|x_1)q(x_2|w)p(y_2|x_2)\cdots \end{aligned}$$



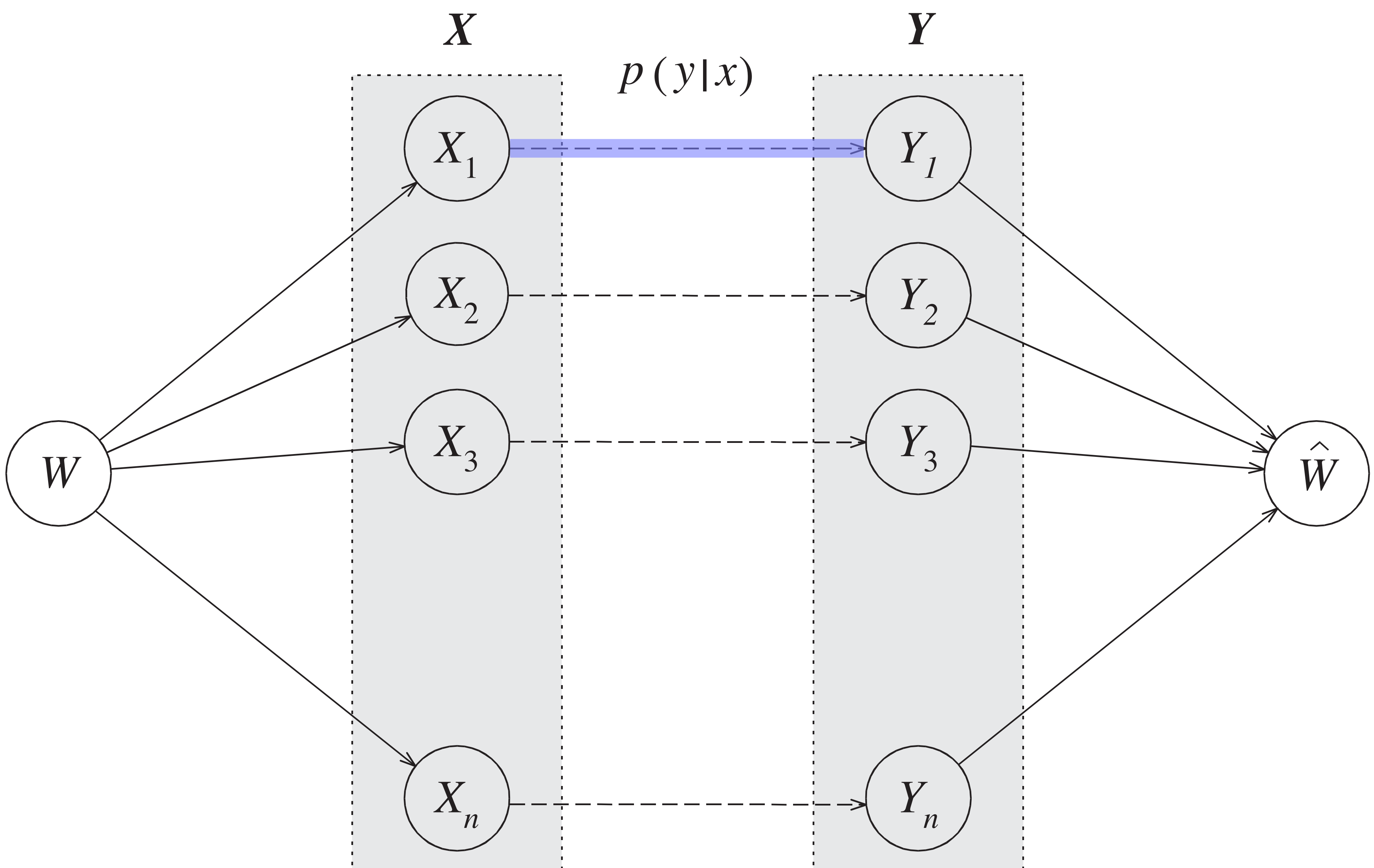
Let q denote the joint distribution for all the r.v.'s:

$$\begin{aligned} q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\ &= q(w)\underline{q(x_1|w)}p(y_1|x_1)q(x_2|w)p(y_2|x_2)\cdots \end{aligned}$$



Let q denote the joint distribution for all the r.v.'s:

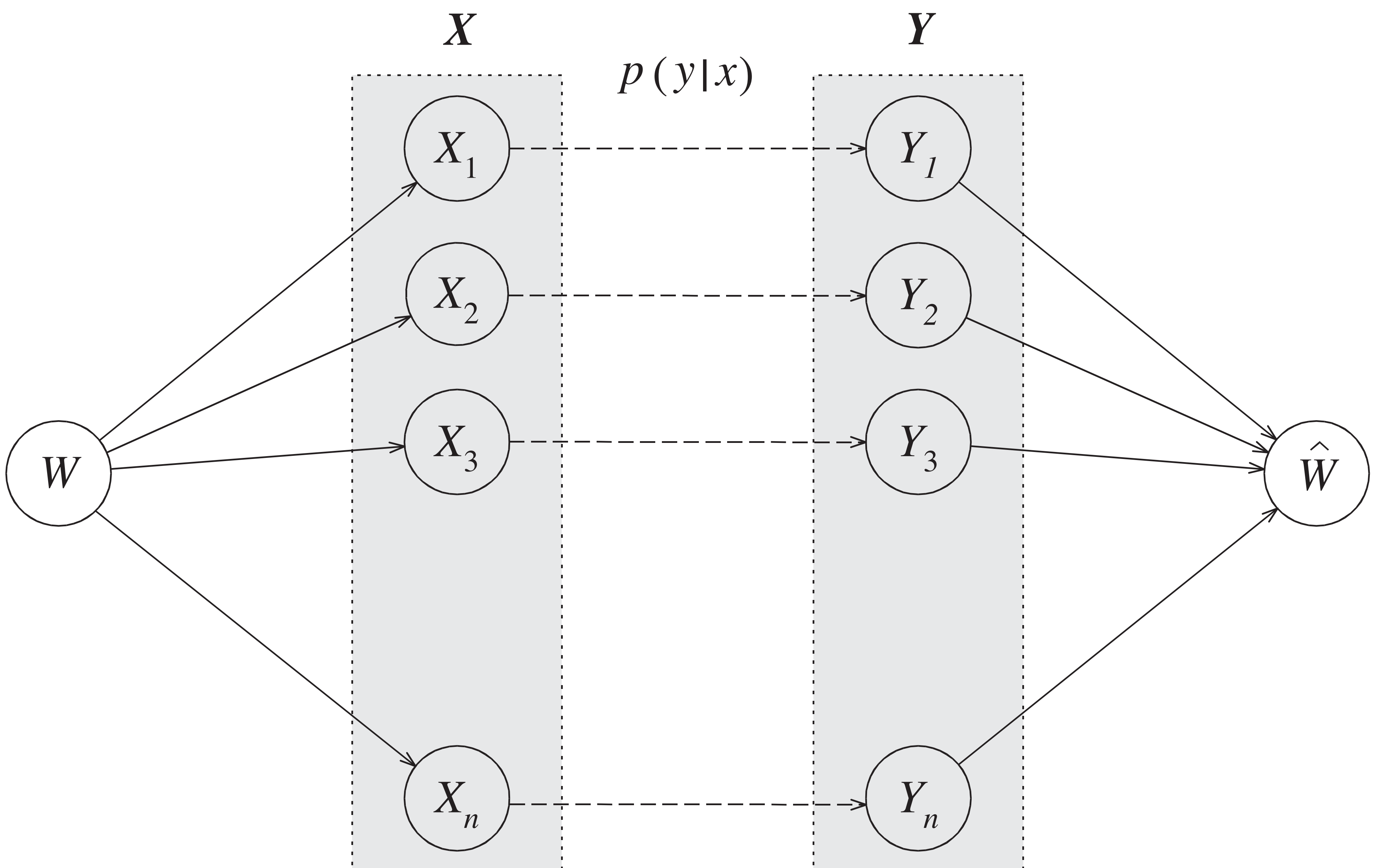
$$
\begin{aligned}
q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\
&= q(w)q(x_1|w)\underline{p(y_1|x_1)}q(x_2|w)p(y_2|x_2)\cdots
\end{aligned}
$$



Let q denote the joint distribution for all the r.v.'s:

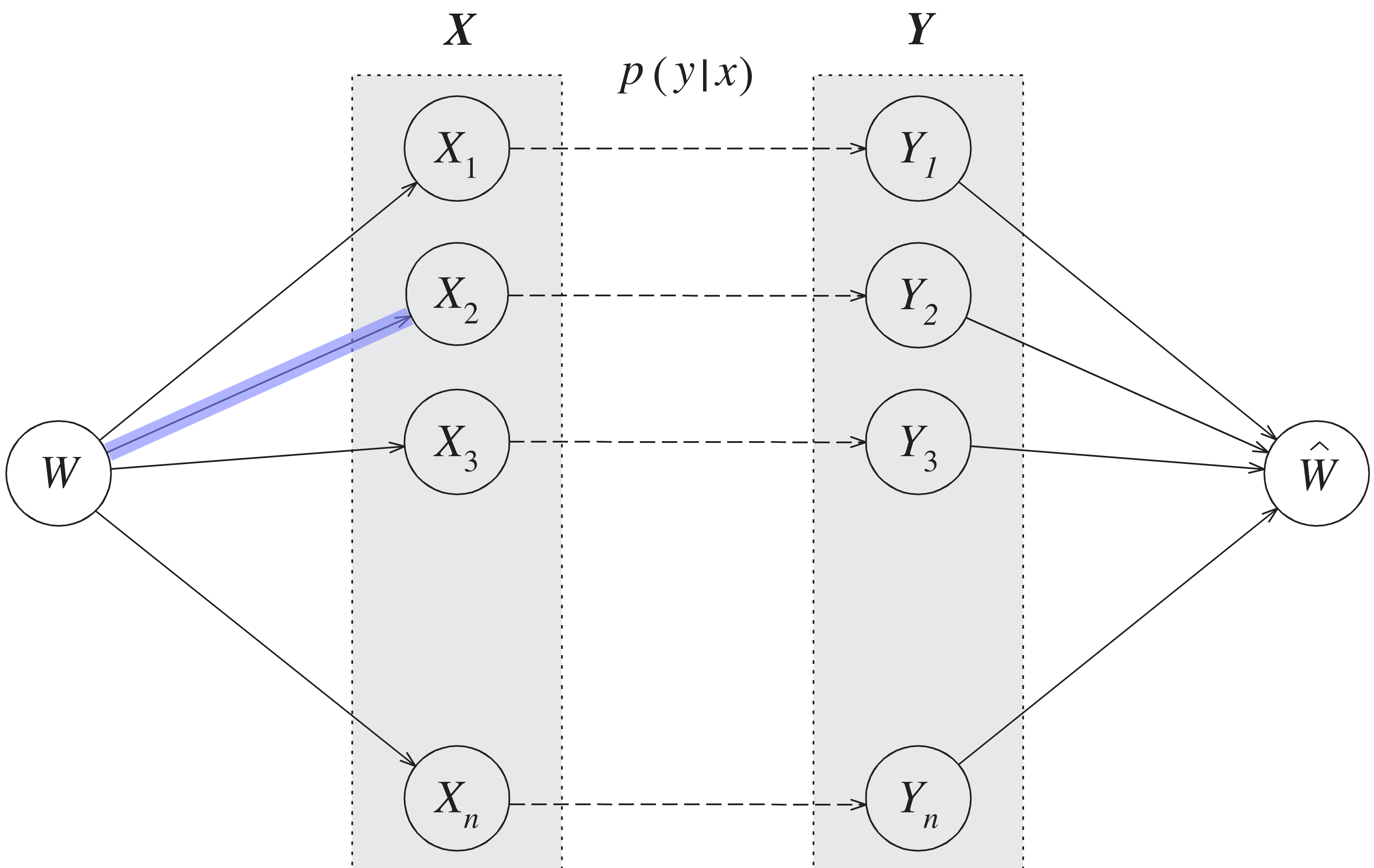
$$q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) = q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots$$

$$= q(w)q(x_1|w)\underline{p(y_1|x_1)}q(x_2|w)p(y_2|x_2)\cdots$$



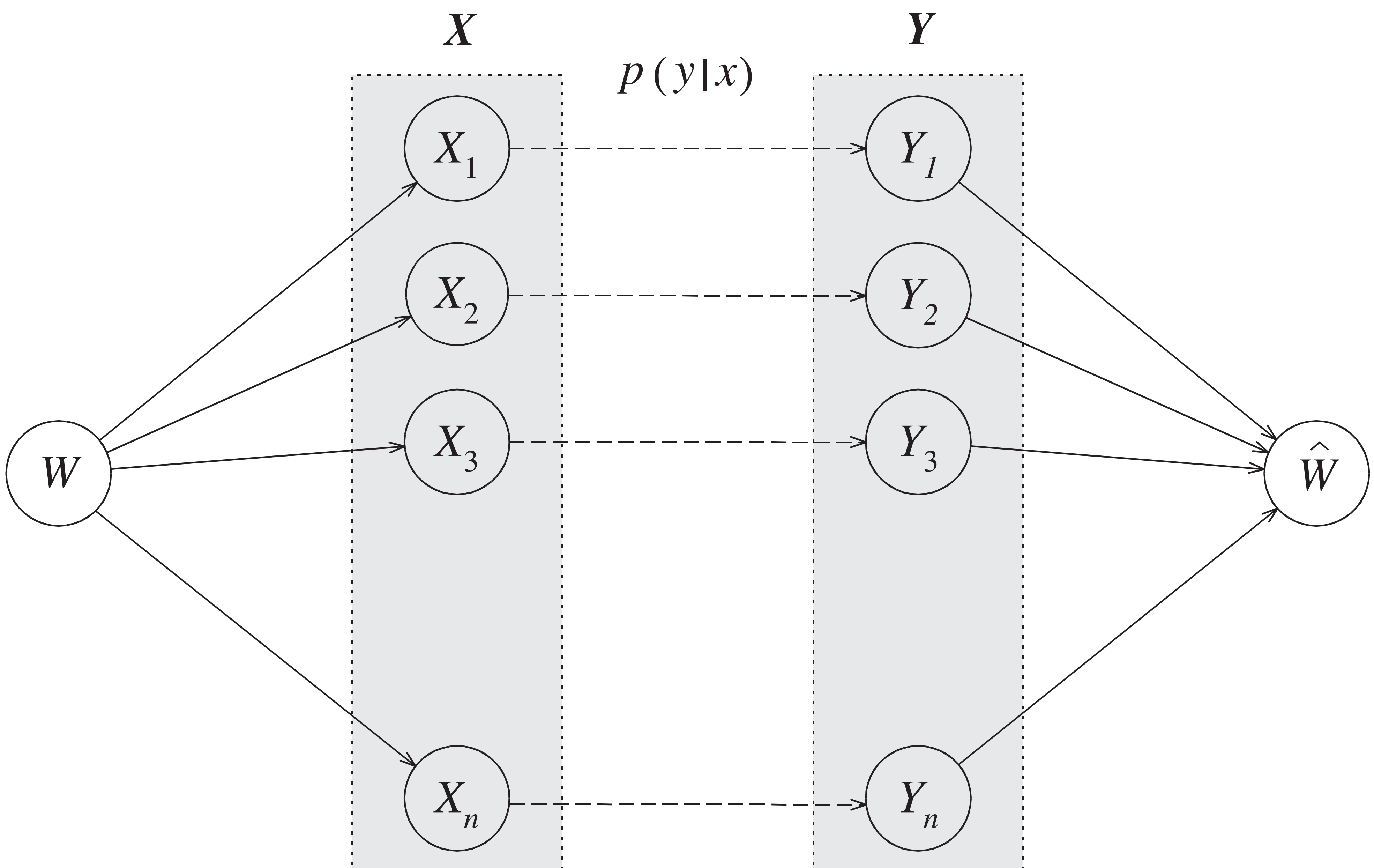
Let q denote the joint distribution for all the r.v.'s:

$$
\begin{aligned}
q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\
&= q(w)q(x_1|w)p(y_1|x_1)\underline{q(x_2|w)}p(y_2|x_2)\cdots
\end{aligned}
$$



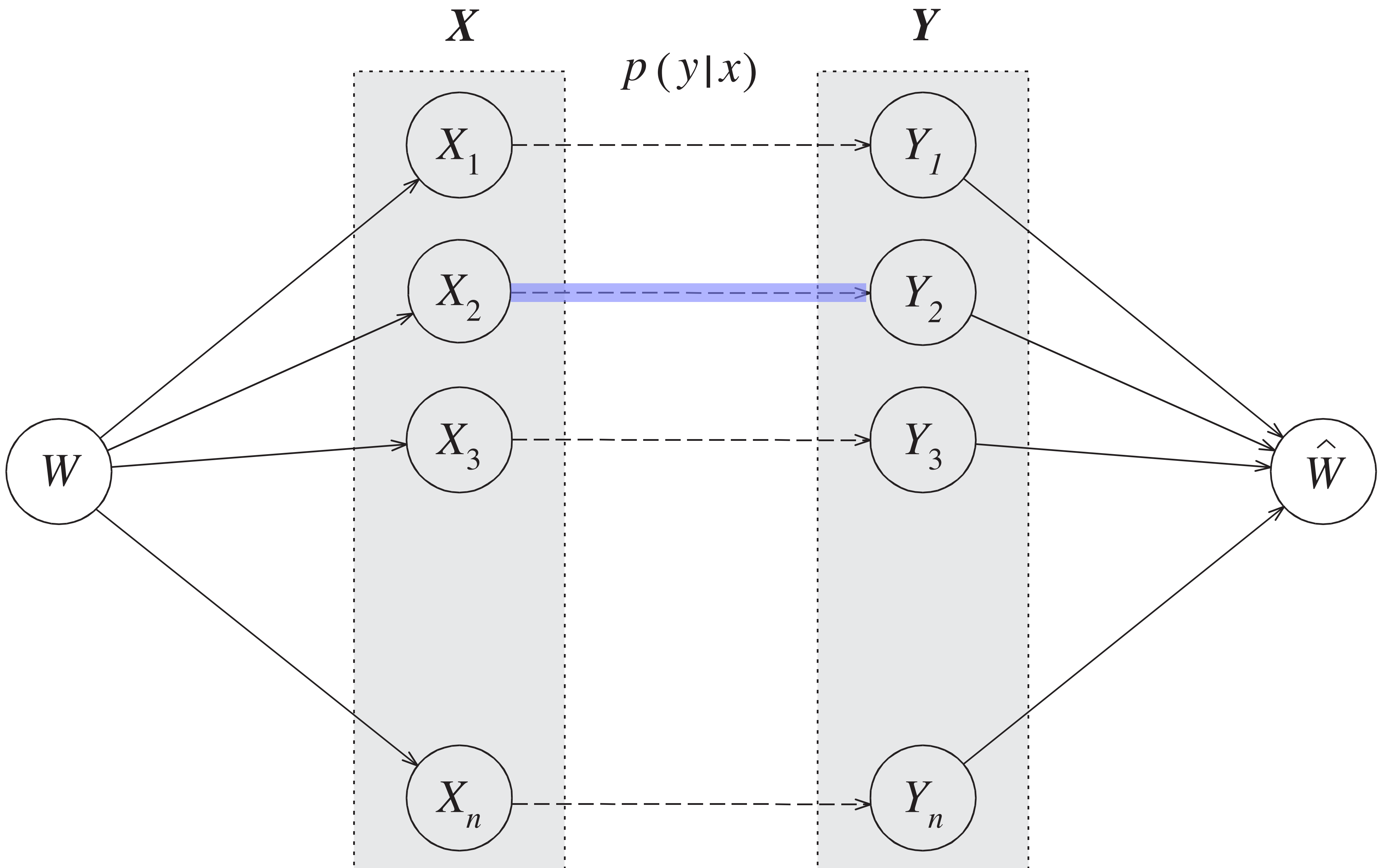
Let q denote the joint distribution for all the r.v.'s:

$$
\begin{aligned}
q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\
&= q(w)q(x_1|w)p(y_1|x_1)\underline{q(x_2|w)}p(y_2|x_2)\cdots
\end{aligned}
$$



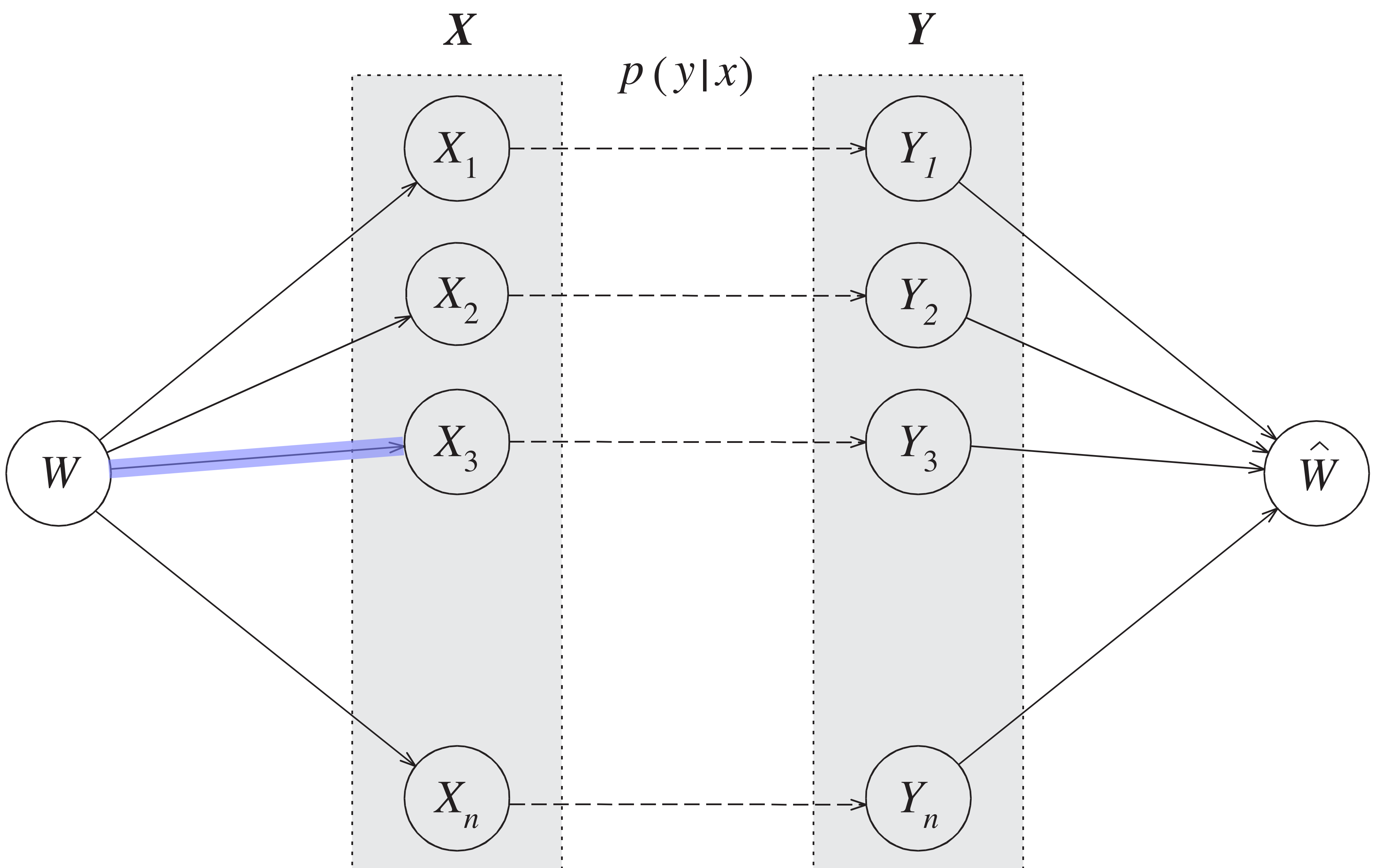
Let q denote the joint distribution for all the r.v.'s:

$$
\begin{aligned}
q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\
&= q(w)q(x_1|w)p(y_1|x_1)q(x_2|w)\underline{p(y_2|x_2)}\cdots
\end{aligned}
$$



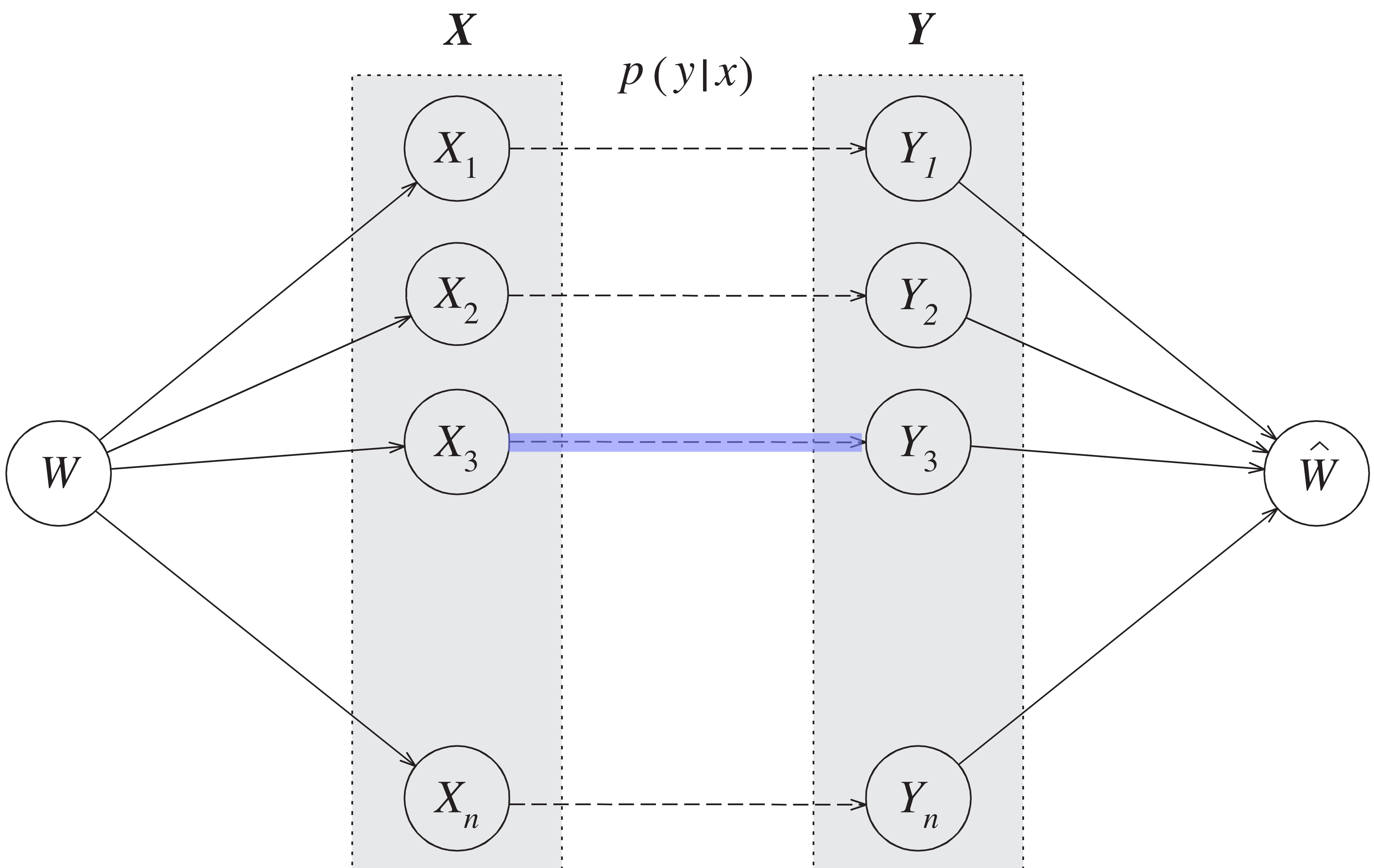
Let q denote the joint distribution for all the r.v.'s:

$$
\begin{aligned}
q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\
&= q(w)q(x_1|w)p(y_1|x_1)q(x_2|w)\underline{p(y_2|x_2)}\cdots
\end{aligned}
$$



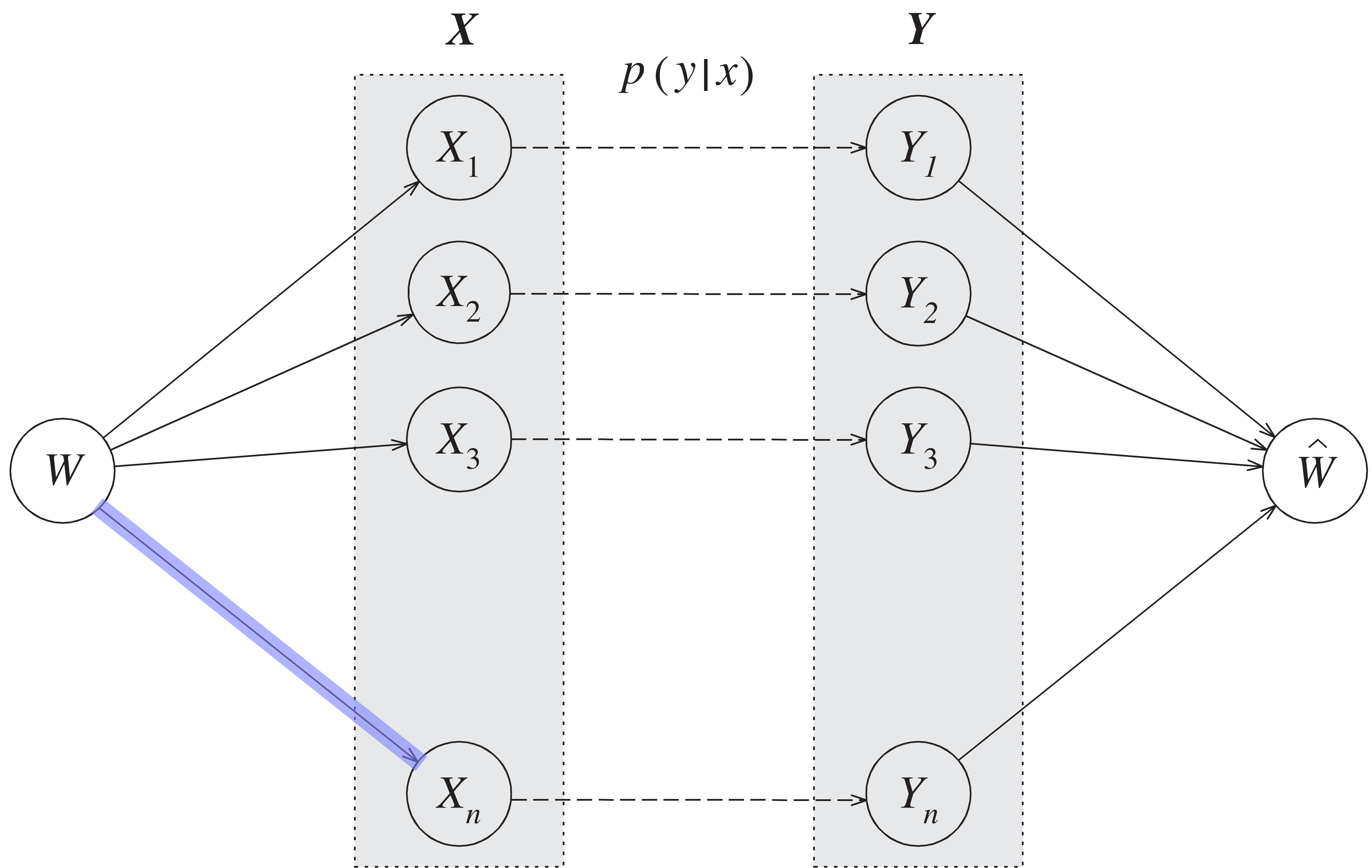
Let q denote the joint distribution for all the r.v.'s:

$$\begin{aligned} q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\ &= q(w)q(x_1|w)p(y_1|x_1)q(x_2|w)p(y_2|x_2)\cdots \end{aligned}$$



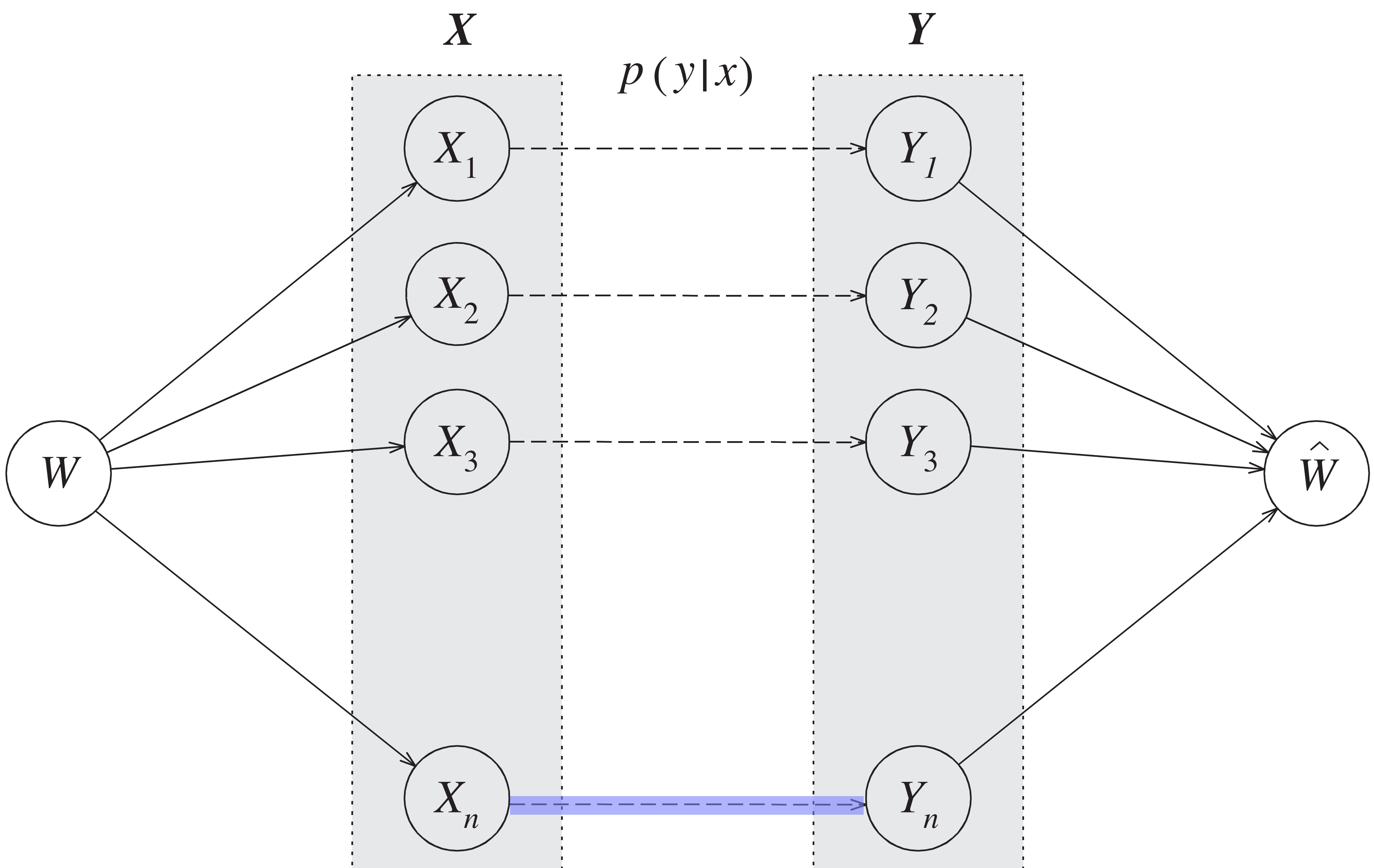
Let q denote the joint distribution for all the r.v.'s:

$$\begin{aligned} q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\ &= q(w)q(x_1|w)p(y_1|x_1)q(x_2|w)p(y_2|x_2)\cdots \end{aligned}$$



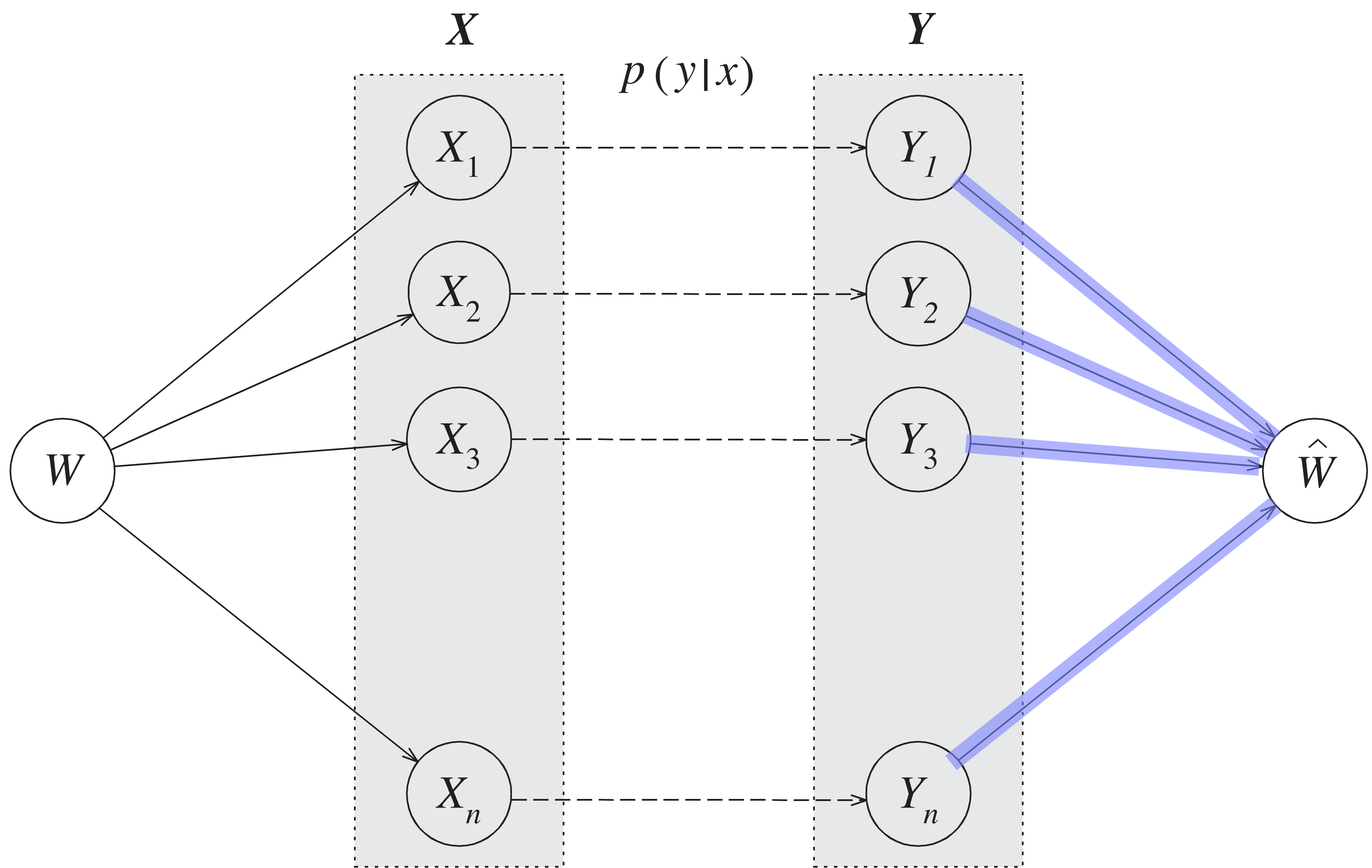
Let q denote the joint distribution for all the r.v.'s:

$$
\begin{aligned}
q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\
&= q(w)q(x_1|w)p(y_1|x_1)q(x_2|w)p(y_2|x_2)\cdots
\end{aligned}
$$



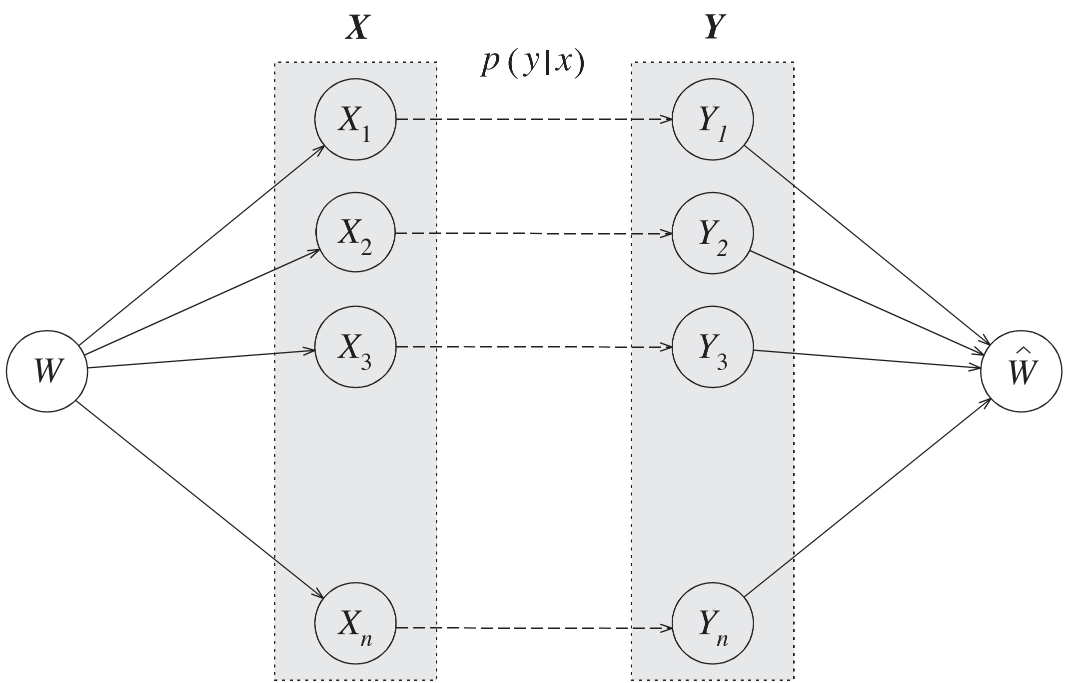
Let q denote the joint distribution for all the r.v.'s:

$$
\begin{aligned}
q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\
&= q(w)q(x_1|w)p(y_1|x_1)q(x_2|w)p(y_2|x_2)\cdots
\end{aligned}
$$

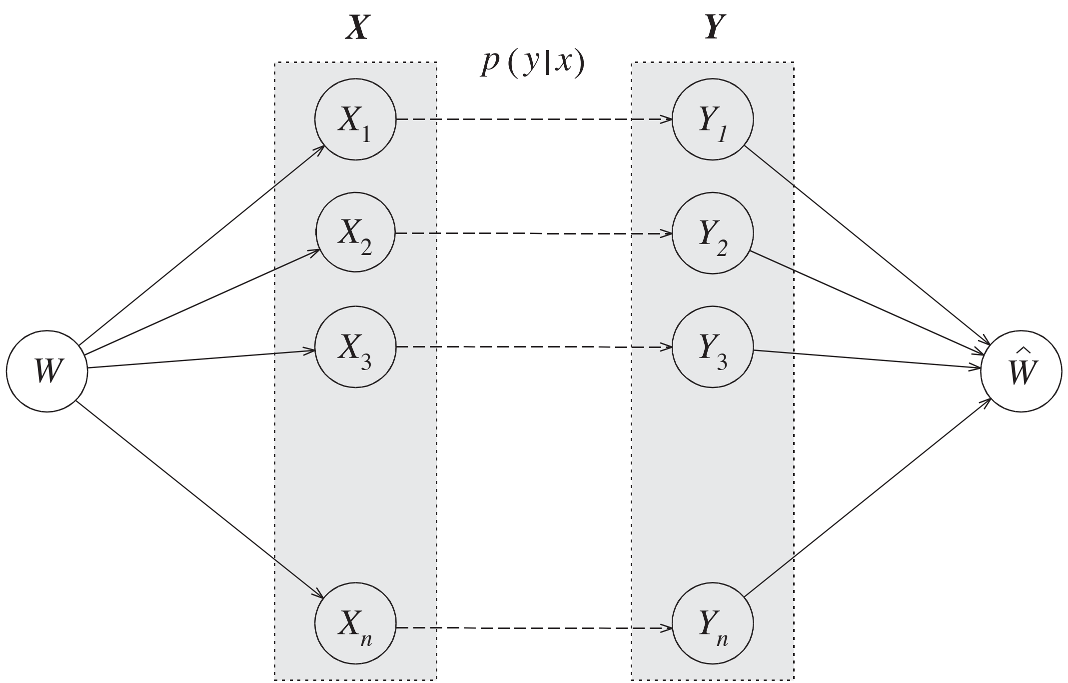


Let q denote the joint distribution for all the r.v.'s:

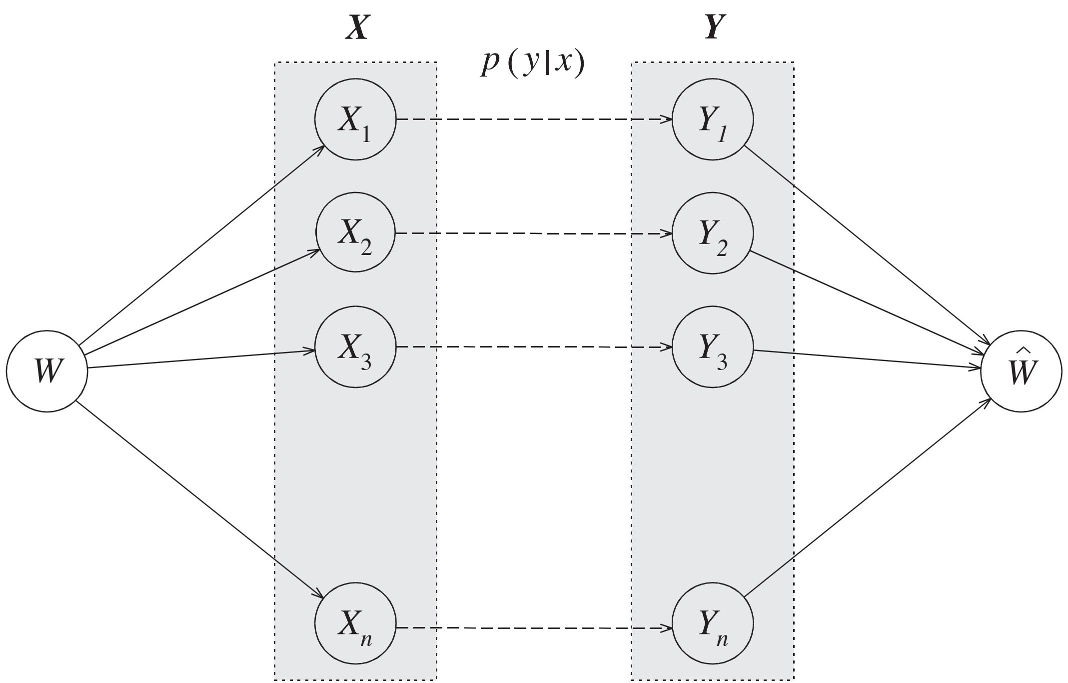
$$\begin{aligned} q(w, x_1, y_1, x_2, y_2, \cdots, x_n, y_n, \hat{w}) &= q(w)q(x_1|w)q(y_1|\cancel{w}, x_1)q(x_2|w, \cancel{x_1}, \cancel{y_1})q(y_2|\cancel{w}, \cancel{x_1}, \cancel{y_1}, x_2)\cdots \\ &= q(w)q(x_1|w)p(y_1|x_1)q(x_2|w)p(y_2|x_2)\cdots \end{aligned}$$



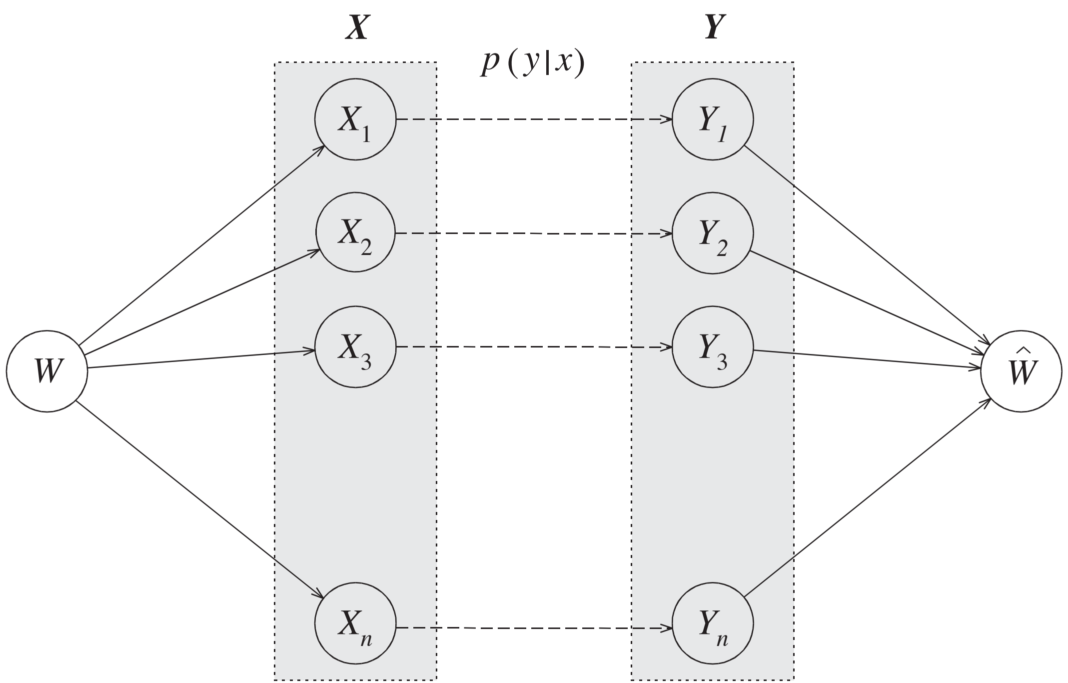
X
p(y|x)
Y
W
X1
X2
X3
Xn
Y1
Y2
Y3
Yn
Ŵ



- Use q to denote the joint distribution and marginal distributions of all r.v.'s.



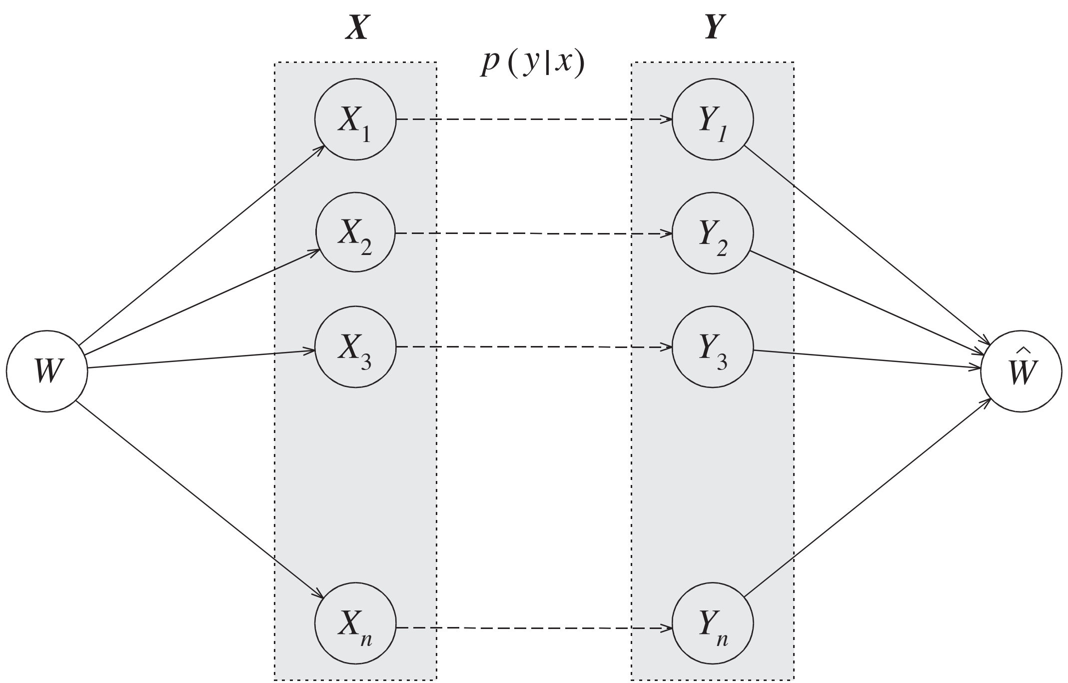
- Use q to denote the joint distribution and marginal distributions of all r.v.'s.
- For all $(w, \mathbf{x}, \mathbf{y}, \hat{w}) \in \mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \hat{\mathcal{W}}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,



- Use q to denote the joint distribution and marginal distributions of all r.v.'s.

- For all $(w, \mathbf{x}, \mathbf{y}, \hat{w}) \in \mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \hat{\mathcal{W}}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

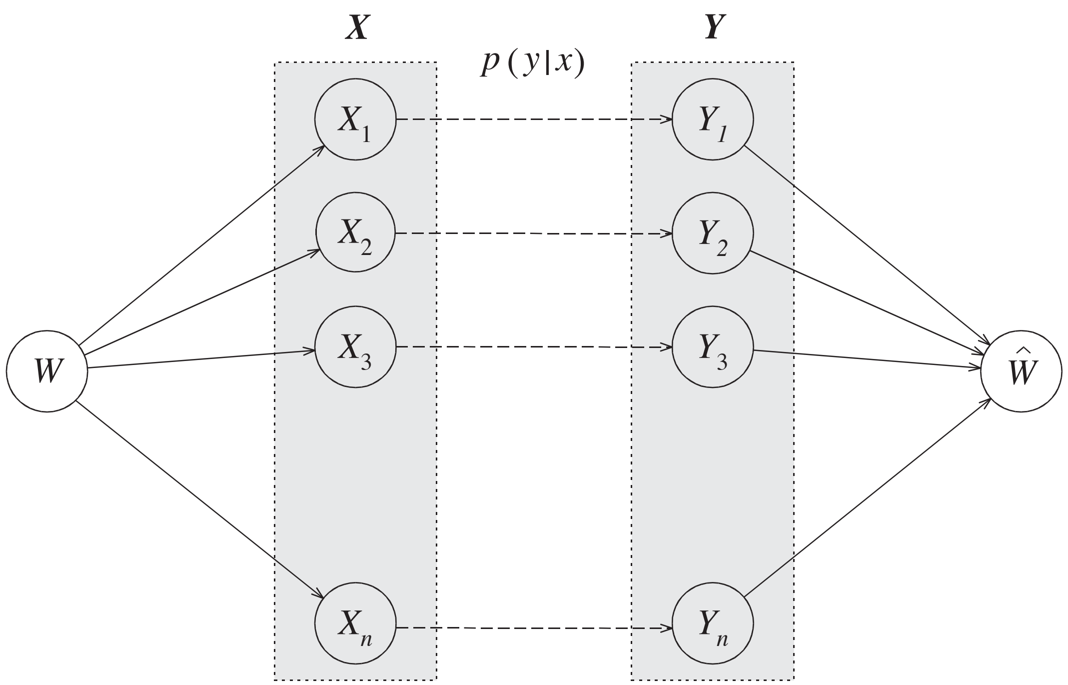
$$q(w, \mathbf{x}, \mathbf{y}\ \hat{w}) = q(w) \left(\prod_{i=1}^{n} q(x_i|w) \right) \left(\prod_{i=1}^{n} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}).$$



- Use q to denote the joint distribution and marginal distributions of all r.v.'s.
- For all $(w, \mathbf{x}, \mathbf{y}, \hat{w}) \in \mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \hat{\mathcal{W}}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$q(w, \mathbf{x}, \mathbf{y}\ \hat{w}) = q(w) \left(\prod_{i=1}^{n} q(x_i|w) \right) \left(\prod_{i=1}^{n} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}).$$

- $q(w) > 0$ for all w so that $q(x_i|w)$ are well-defined.



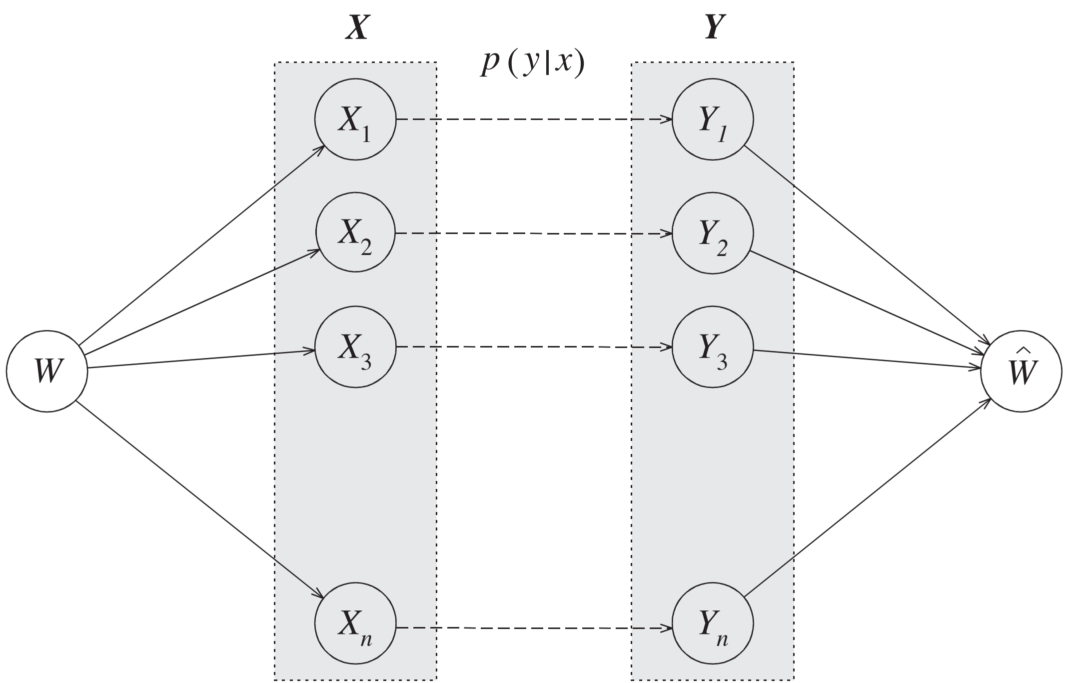
- Use q to denote the joint distribution and marginal distributions of all r.v.'s.

- For all $(w, \mathbf{x}, \mathbf{y}, \hat{w}) \in \mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \hat{\mathcal{W}}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$q(w, \mathbf{x}, \mathbf{y}\ \hat{w}) = q(w)\left(\prod_{i=1}^{n} q(x_i|w)\right)\left(\prod_{i=1}^{n} p(y_i|x_i)\right)q(\hat{w}|\mathbf{y}).$$

- $q(w) > 0$ for all w so that $q(x_i|w)$ are well-defined.

- $q(x_i|w)$ and $q(\hat{w}|\mathbf{y})$ are deterministic.



- Use q to denote the joint distribution and marginal distributions of all r.v.'s.

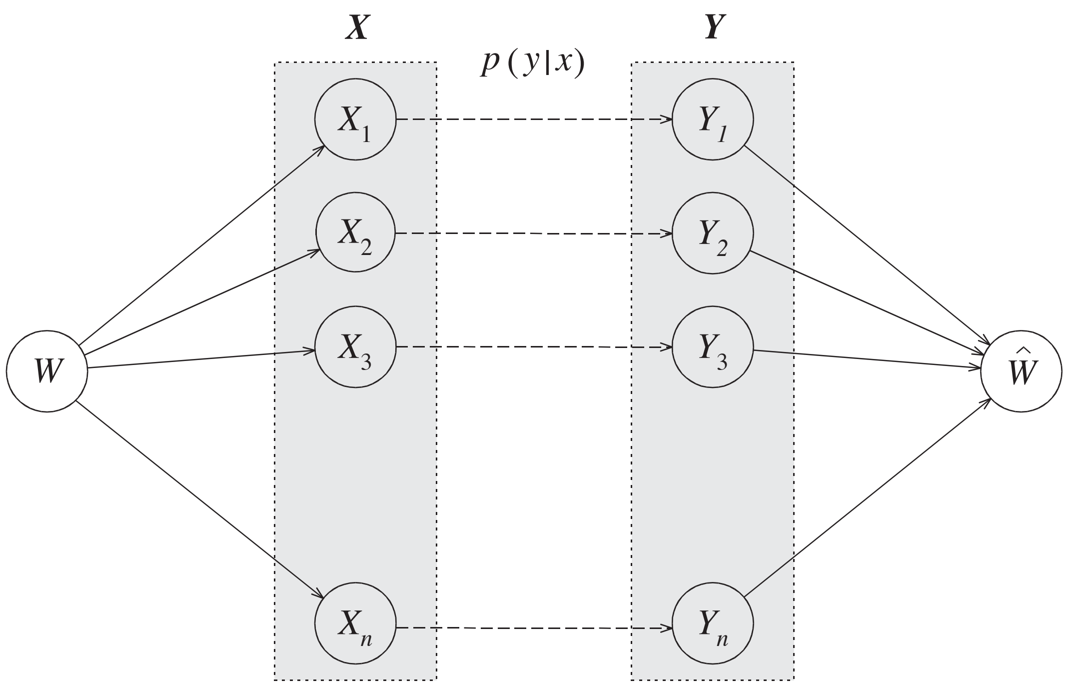
- For all $(w, \mathbf{x}, \mathbf{y}, \hat{w}) \in \mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \hat{\mathcal{W}}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$q(w, \mathbf{x}, \mathbf{y}\ \hat{w}) = q(w) \left(\prod_{i=1}^{n} q(x_i|w) \right) \left(\prod_{i=1}^{n} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}).$$

- $q(w) > 0$ for all w so that $q(x_i|w)$ are well-defined.

- $q(x_i|w)$ and $q(\hat{w}|\mathbf{y})$ are deterministic.

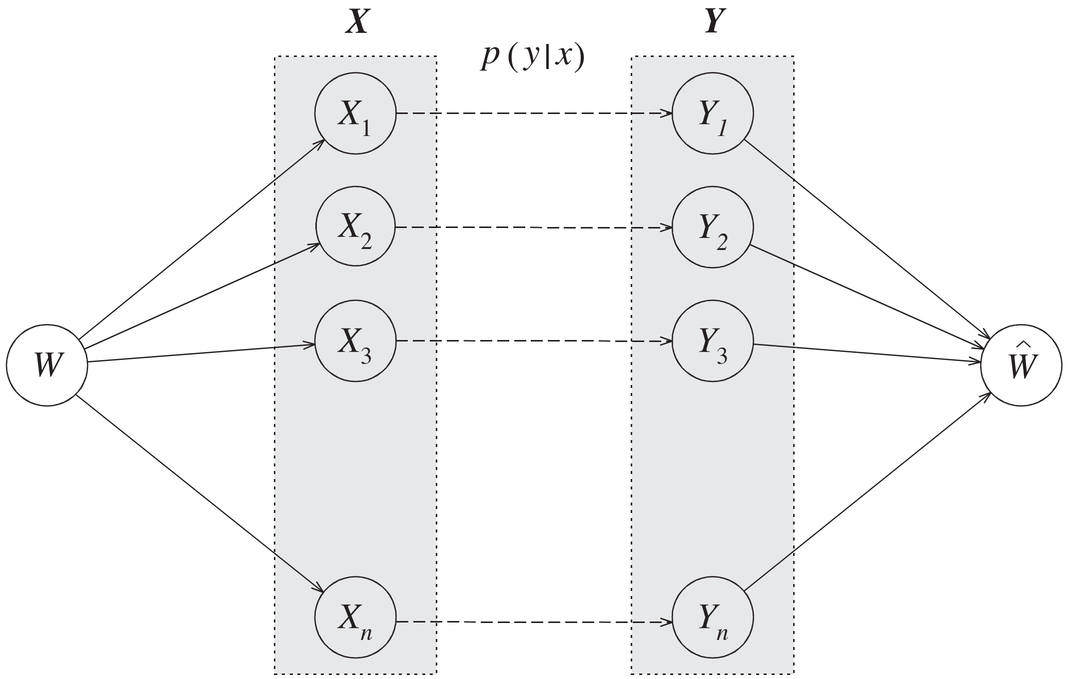
- The dependency graph suggests the Markov chain $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$.



- Use q to denote the joint distribution and marginal distributions of all r.v.'s.
- For all $(w, \mathbf{x}, \mathbf{y}, \hat{w}) \in \mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \hat{\mathcal{W}}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$q(w, \mathbf{x}, \mathbf{y}\ \hat{w}) = q(w) \left(\prod_{i=1}^{n} q(x_i|w) \right) \left(\prod_{i=1}^{n} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}).$$

- $q(w) > 0$ for all w so that $q(x_i|w)$ are well-defined.
- $q(x_i|w)$ and $q(\hat{w}|\mathbf{y})$ are deterministic.
- The dependency graph suggests the Markov chain $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$.
- This can be formally justified by invoking Proposition 2.9.



- Use q to denote the joint distribution and marginal distributions of all r.v.'s.
- For all $(w, \mathbf{x}, \mathbf{y}, \hat{w}) \in \mathcal{W} \times \mathcal{X}^n \times \mathcal{Y}^n \times \hat{\mathcal{W}}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$q(w, \mathbf{x}, \mathbf{y}\ \hat{w}) = \overbrace{q(w)}^{a(w)} \overbrace{\left(\prod_{i=1}^{n} q(x_i|w)\right)}^{b(w,\mathbf{x})} \overbrace{\left(\prod_{i=1}^{n} p(y_i|x_i)\right)}^{c(\mathbf{x},\mathbf{y})} \overbrace{q(\hat{w}|\mathbf{y})}^{d(\mathbf{y},w)}.$$

- $q(w) > 0$ for all w so that $q(x_i|w)$ are well-defined.
- $q(x_i|w)$ and $q(\hat{w}|\mathbf{y})$ are deterministic.
- The dependency graph suggests the Markov chain $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$.
- This can be formally justified by invoking Proposition 2.9.

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$q(\mathbf{x}, \mathbf{y}) = \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w})$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\ &= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\ &= \sum_w \sum_{\hat{w}} \underline{q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right)} q(\hat{w}|\mathbf{y}) \end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} \underline{q(w) \left(\prod_{i} q(x_i|w)\right) \left(\prod_{i} p(y_i|x_i)\right)} q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} \underline{q(w) \left(\prod_{i} q(x_i|w)\right) \left(\prod_{i} p(y_i|x_i)\right)} \sum_{\hat{w}} q(\hat{w}|\mathbf{y})
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y})
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y})
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \underline{\left(\prod_{i} p(y_i|x_i) \right)} \cancel{\sum_{\hat{w}} q(\hat{w}|\mathbf{y})}
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \qquad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$q(\mathbf{x}) = \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y})$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$q(\mathbf{x}) = \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y})$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right]
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \qquad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right]
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \quad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\underline{\sum_w q(w) \prod_i q(x_i|w)}\right] \left[\prod_i p(y_i|x_i)\right] \\
&= \left[\underline{\sum_w q(w) \prod_i q(x_i|w)}\right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)\right]
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \qquad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right]
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \underline{\sum_{\mathbf{y}}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\underline{\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_{i}} p(y_i|x_i) \right]
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_{i} p(y_i|x_i) \right]
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\underline{\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)} \right]
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\underline{\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)} \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \underline{\prod_i \left(\sum_{y_i} p(y_i|x_i) \right)}
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \qquad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\underline{\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)} \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \underline{\prod_i \left(\sum_{y_i} p(y_i|x_i) \right)}
\end{aligned}$$

For $n = 2$,

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w}\sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w}\sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w)\right] \left[\underline{\sum_{y_1}\sum_{y_2}\cdots\sum_{y_n}\prod_i p(y_i|x_i)}\right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w)\right] \underline{\prod_i \left(\sum_{y_i} p(y_i|x_i)\right)}
\end{aligned}$$

For $n = 2$,

$$\sum_{y_1}\sum_{y_2}\prod_{i=1}^{2} p(y_i|x_i)$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \quad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\underline{\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)} \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \underline{\prod_i \left(\sum_{y_i} p(y_i|x_i) \right)}
\end{aligned}$$

For $n = 2$,

$$\begin{aligned}
&\sum_{y_1} \sum_{y_2} \prod_{i=1}^{2} p(y_i|x_i) \\
&\quad = \sum_{y_1} \sum_{y_2} p(y_1|x_1) p(y_2|x_2)
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\underline{\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)}\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \underline{\prod_i \left(\sum_{y_i} p(y_i|x_i)\right)}
\end{aligned}$$

For $n = 2$,

$$\begin{aligned}
&\sum_{y_1} \sum_{y_2} \prod_{i=1}^{2} p(y_i|x_i) \\
&\quad = \sum_{y_1} \sum_{y_2} p(y_1|x_1) p(y_2|x_2) \\
&\quad = \left(\sum_{y_1} p(y_1|x_1)\right) \left(\sum_{y_2} p(y_2|x_2)\right)
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \quad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\underline{\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)} \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \underline{\prod_i \left(\sum_{y_i} p(y_i|x_i) \right)}
\end{aligned}$$

For $n = 2$,

$$\begin{aligned}
&\sum_{y_1} \sum_{y_2} \prod_{i=1}^{2} p(y_i|x_i) \\
&\quad = \sum_{y_1} \sum_{y_2} p(y_1|x_1) p(y_2|x_2) \\
&\quad = \left(\sum_{y_1} p(y_1|x_1) \right) \left(\sum_{y_2} p(y_2|x_2) \right) \\
&\quad = \prod_{i=1}^{2} \left(\sum_{y_i} p(y_i|x_i) \right)
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right)
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right)
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \prod_{i} \left(\cancel{\sum_{y_i} p(y_i|x_i)} \right) \\
&= \sum_{w} q(w) \prod_{i} q(x_i|w).
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \prod_{i} \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_{w} q(w) \prod_{i} q(x_i|w).
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \quad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \prod_{i} \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_{w} q(w) \prod_{i} q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_{i} p(y_i|x_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \prod_i \left(\sum_{y_i} p(y_i|x_i)\right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_{w} q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \prod_i \left(\sum_{y_i} p(y_i|x_i)\right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \qquad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\ &= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\ &= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right]. \end{aligned}$$

2. Furthermore,

$$\begin{aligned} q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\ &= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\ &= \sum_w q(w) \prod_i q(x_i|w). \end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \qquad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \prod_i \left(\sum_{y_i} p(y_i|x_i)\right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\ &= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\ &= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right]. \end{aligned}$$

2. Furthermore,

$$\begin{aligned} q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\ &= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\ &= \sum_w q(w) \prod_i q(x_i|w). \end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \qquad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)\right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \prod_i \left(\sum_{y_i} p(y_i|x_i)\right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\ &= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\ &= \sum_w q(w) \left(\prod_i q(x_i|w)\right) \left(\prod_i p(y_i|x_i)\right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\ &= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right]. \end{aligned}$$

2. Furthermore,

$$\begin{aligned} q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\ &= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\prod_i p(y_i|x_i)\right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i)\right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w)\right] \prod_i \left(\sum_{y_i} p(y_i|x_i)\right) \\ &= \sum_w q(w) \prod_i q(x_i|w). \end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_w q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_{w} q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

$$-E \log q(\mathbf{Y}|\mathbf{X}) = -E \log \prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n} \left[-E \log p(Y_i|X_i) \right],$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x},\mathbf{y}) &= \sum_{w}\sum_{\hat{w}} q(w,\mathbf{x},\mathbf{y},\hat{w}) \\ &= \sum_{w}\sum_{\hat{w}} q(w)\left(\prod_i q(x_i|w)\right)\left(\prod_i p(y_i|x_i)\right) q(\hat{w}|\mathbf{y}) \\ &= \sum_{w} q(w)\left(\prod_i q(x_i|w)\right)\left(\prod_i p(y_i|x_i)\right)\sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\ &= \left[\sum_{w} q(w)\prod_i q(x_i|w)\right]\left[\prod_i p(y_i|x_i)\right]. \end{aligned}$$

2. Furthermore,

$$\begin{aligned} q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x},\mathbf{y}) \\ &= \sum_{\mathbf{y}}\left[\sum_{w} q(w)\prod_i q(x_i|w)\right]\left[\prod_i p(y_i|x_i)\right] \\ &= \left[\sum_{w} q(w)\prod_i q(x_i|w)\right]\left[\sum_{y_1}\sum_{y_2}\cdots\sum_{y_n}\prod_i p(y_i|x_i)\right] \\ &= \left[\sum_{w} q(w)\prod_i q(x_i|w)\right]\prod_i\left(\sum_{y_i} p(y_i|x_i)\right) \\ &= \sum_{w} q(w)\prod_i q(x_i|w). \end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x},\mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x},\mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x},\mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

$$-E\log q(\mathbf{Y}|\mathbf{X}) = -E\log\prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n}\left[-E\log p(Y_i|X_i)\right],$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\ &= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\ &= \sum_{w} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\ &= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right]. \end{aligned}$$

2. Furthermore,

$$\begin{aligned} q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\ &= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\ &= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\ &= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\ &= \sum_{w} q(w) \prod_i q(x_i|w). \end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

$$-E \log q(\mathbf{Y}|\mathbf{X}) = -E \log \prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n} \left[-E \log p(Y_i|X_i) \right],$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \prod_{i} \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_{w} q(w) \prod_{i} q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_{i} p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

$$-E \log q(\mathbf{Y}|\mathbf{X}) = -E \log \prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n} \left[-E \log p(Y_i|X_i) \right],$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\ &= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\ &= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right]. \end{aligned}$$

2. Furthermore,

$$\begin{aligned} q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\ &= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\ &= \sum_w q(w) \prod_i q(x_i|w). \end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

$$-E \log q(\mathbf{Y}|\mathbf{X}) = -E \log \prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n} \left[-E \log p(Y_i|X_i) \right],$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \prod_{i} \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_{w} q(w) \prod_{i} q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_{i} p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

$$-E \log q(\mathbf{Y}|\mathbf{X}) = -E \log \prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n} \left[-E \log p(Y_i|X_i) \right],$$

or

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \tag{1}$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_{w} q(w) \prod_i q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

$$-E \log q(\mathbf{Y}|\mathbf{X}) = -E \log \prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n} [-E \log p(Y_i|X_i)],$$

or

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \quad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned} q(\mathbf{x}, \mathbf{y}) &= \sum_w \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\ &= \sum_w \sum_{\hat{w}} q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\ &= \sum_w q(w) \left(\prod_i q(x_i|w) \right) \left(\prod_i p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right]. \end{aligned}$$

2. Furthermore,

$$\begin{aligned} q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\ &= \sum_{\mathbf{y}} \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\prod_i p(y_i|x_i) \right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_i p(y_i|x_i) \right] \\ &= \left[\sum_w q(w) \prod_i q(x_i|w) \right] \prod_i \left(\sum_{y_i} p(y_i|x_i) \right) \\ &= \sum_w q(w) \prod_i q(x_i|w). \end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_i p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

$$-E \log q(\mathbf{Y}|\mathbf{X}) = -E \log \prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n} [-E \log p(Y_i|X_i)],$$

or

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i).$$

Proposition Show that for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i). \qquad (1)$$

Proof

1. First, for $\mathbf{x}$ and $\mathbf{y}$ such that $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) > 0$,

$$\begin{aligned}
q(\mathbf{x}, \mathbf{y}) &= \sum_{w} \sum_{\hat{w}} q(w, \mathbf{x}, \mathbf{y}, \hat{w}) \\
&= \sum_{w} \sum_{\hat{w}} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) q(\hat{w}|\mathbf{y}) \\
&= \sum_{w} q(w) \left(\prod_{i} q(x_i|w) \right) \left(\prod_{i} p(y_i|x_i) \right) \sum_{\hat{w}} q(\hat{w}|\mathbf{y}) \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right].
\end{aligned}$$

2. Furthermore,

$$\begin{aligned}
q(\mathbf{x}) &= \sum_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}) \\
&= \sum_{\mathbf{y}} \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \left[\sum_{y_1} \sum_{y_2} \cdots \sum_{y_n} \prod_{i} p(y_i|x_i) \right] \\
&= \left[\sum_{w} q(w) \prod_{i} q(x_i|w) \right] \prod_{i} \left(\sum_{y_i} p(y_i|x_i) \right) \\
&= \sum_{w} q(w) \prod_{i} q(x_i|w).
\end{aligned}$$

3. Therefore, for $\mathbf{x}$ such that $q(\mathbf{x}) > 0$,

$$q(\mathbf{y}|\mathbf{x}) = \frac{q(\mathbf{x}, \mathbf{y})}{q(\mathbf{x})} = \prod_{i} p(y_i|x_i).$$

Proposition $H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$.

Proof

1. For any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, if $q(\mathbf{x}, \mathbf{y}) > 0$, then $q(\mathbf{x}) > 0$. Thus by the above proposition, (1) holds.

2. Therefore by (1),

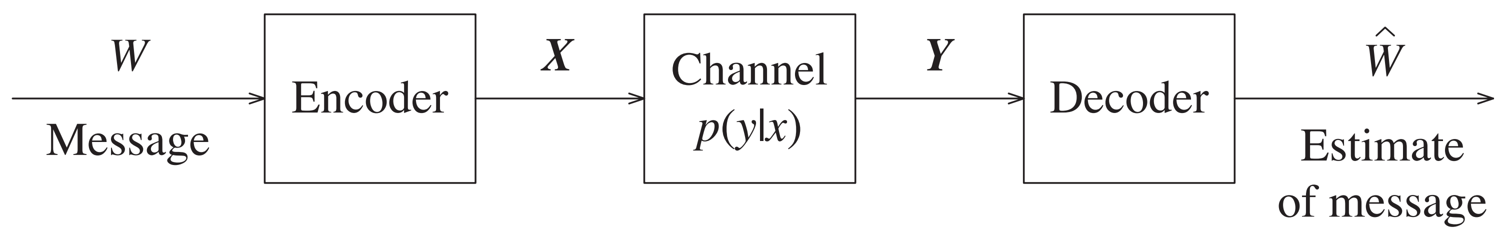
$$-E \log q(\mathbf{Y}|\mathbf{X}) = -E \log \prod_{i=1}^{n} p(Y_i|X_i) = \sum_{i=1}^{n} \left[-E \log p(Y_i|X_i) \right],$$

or

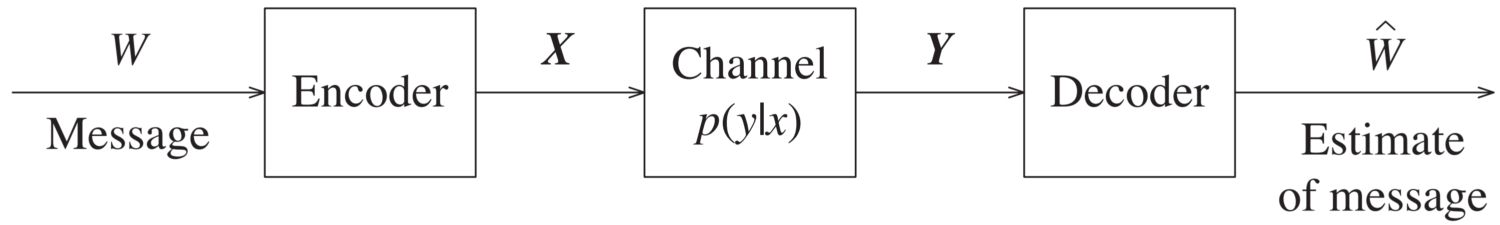
$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i).$$

Why C is related to I(X;Y)?

Why C is related to I(X;Y)?



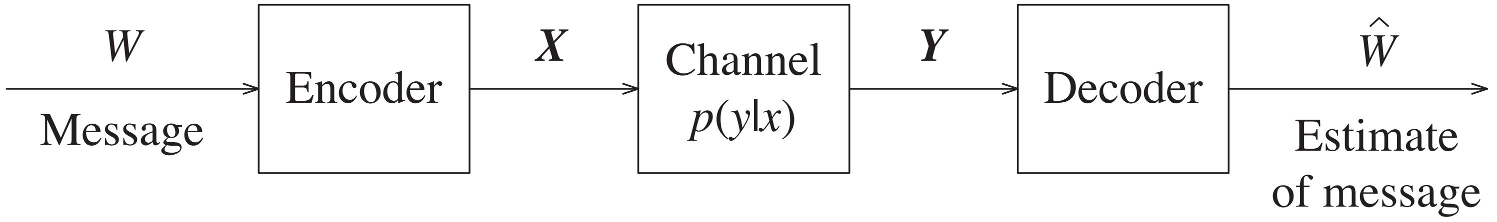
Why C is related to I(X;Y)?



- Consider the information diagram for

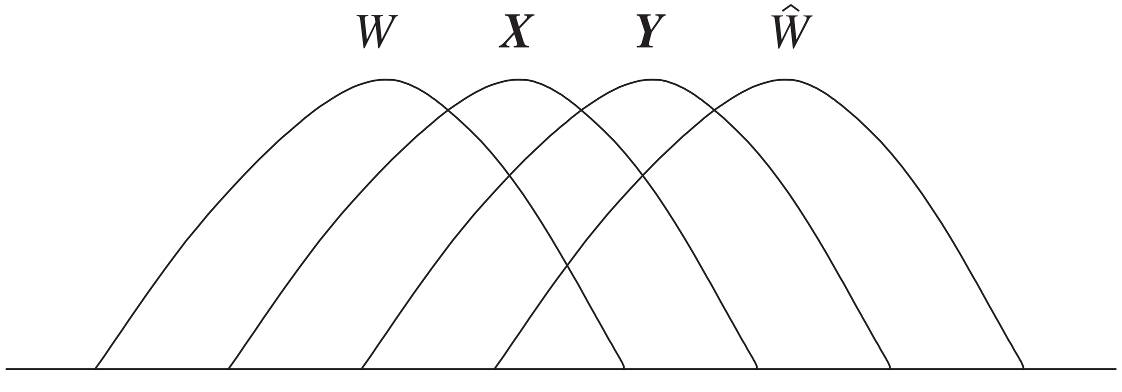
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

Why C is related to I(X;Y)?

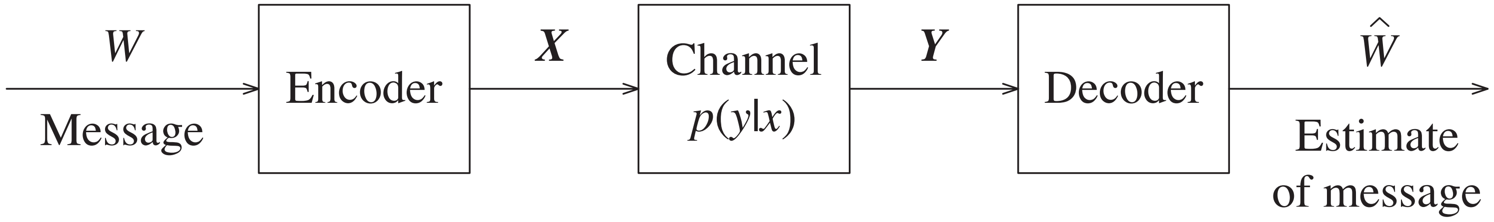


- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

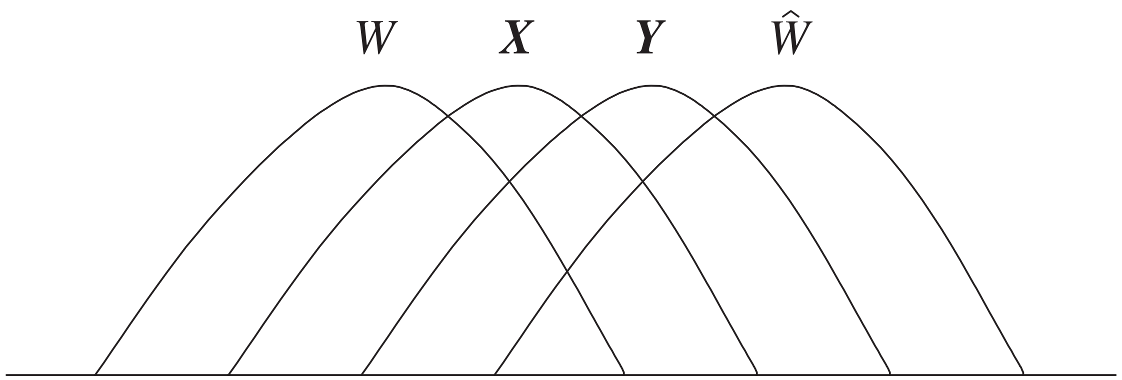


Why C is related to I(X;Y)?

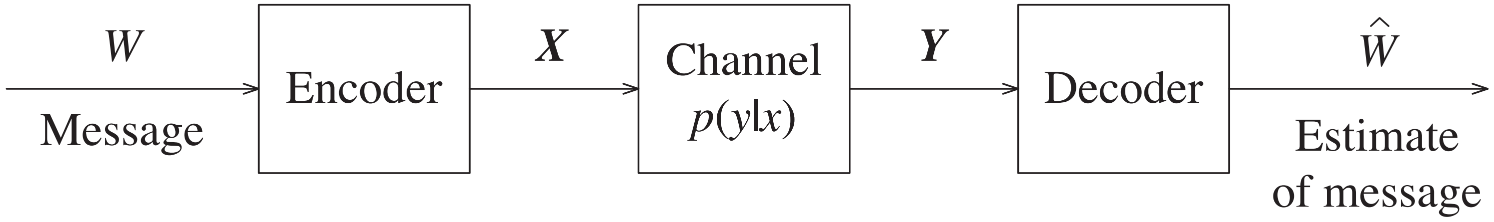


- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$



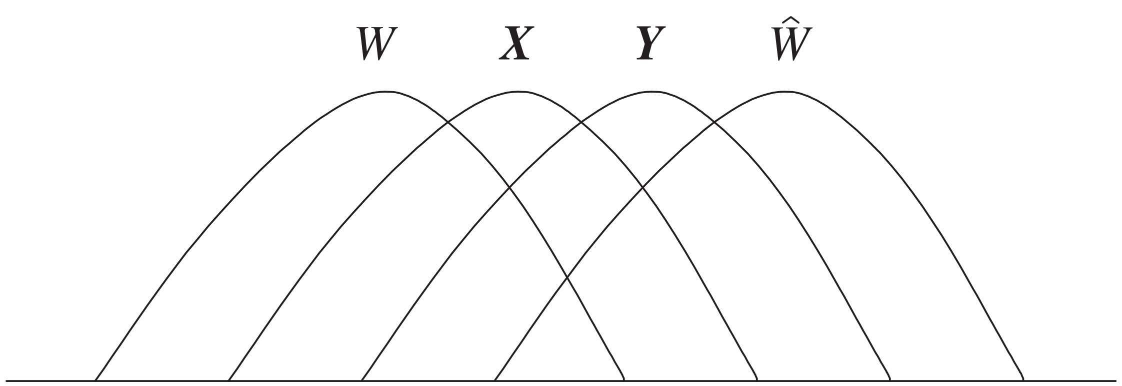
Why C is related to I(X;Y)?



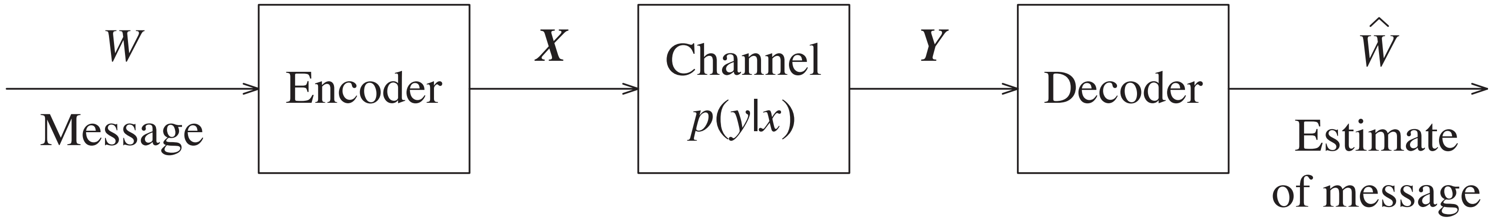
- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$



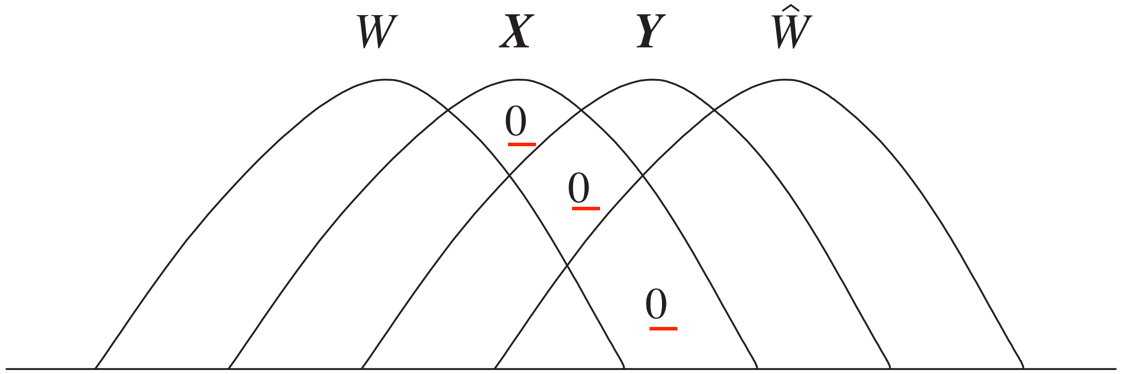
Why C is related to I(X;Y)?



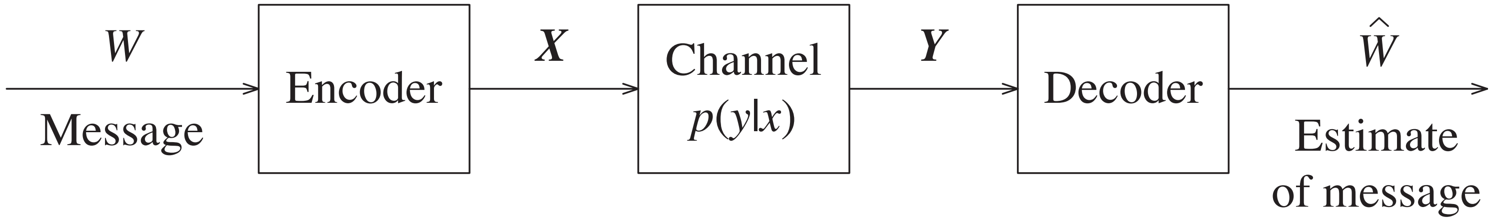
- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$



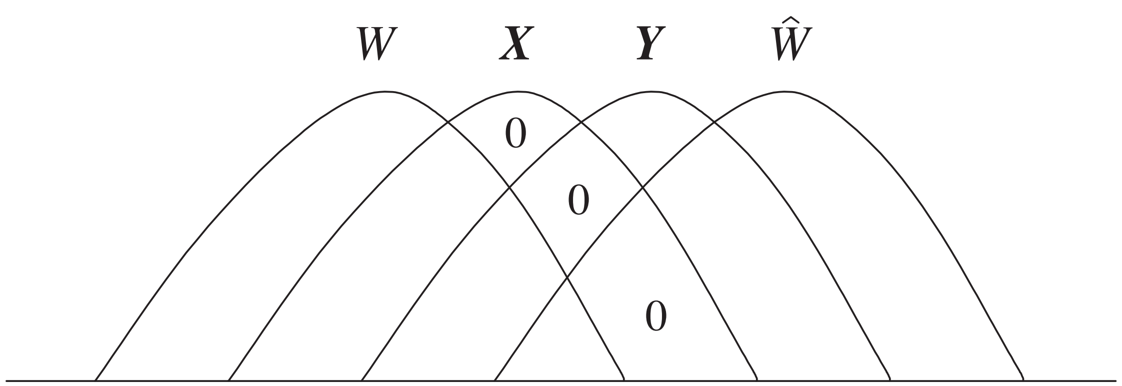
Why C is related to I(X;Y)?



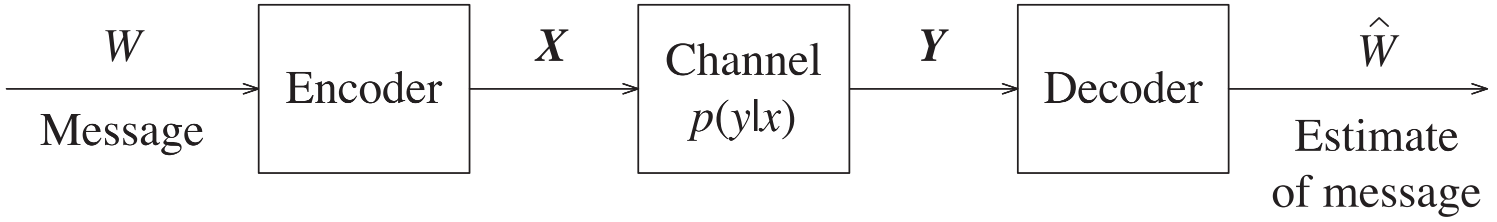
- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$



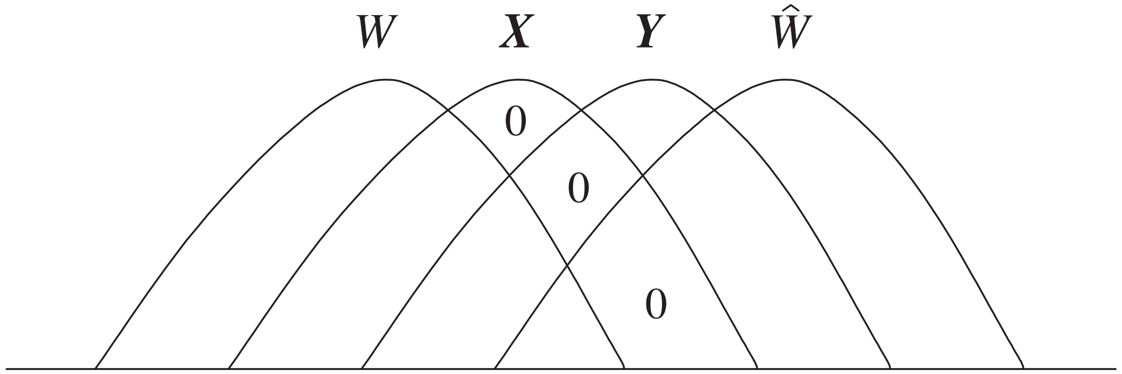
Why C is related to I(X;Y)?



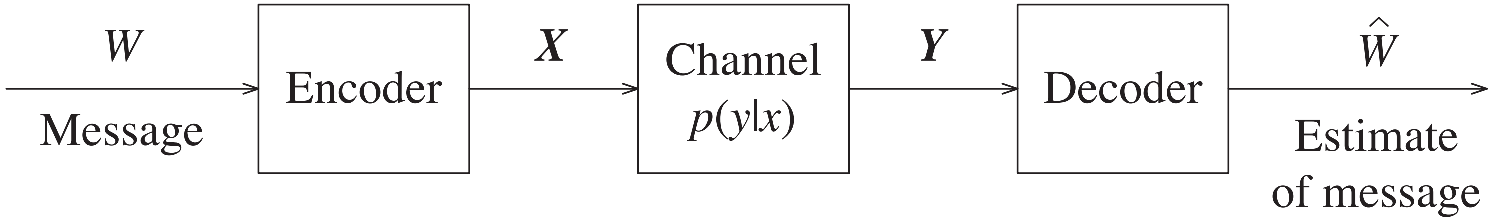
- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$



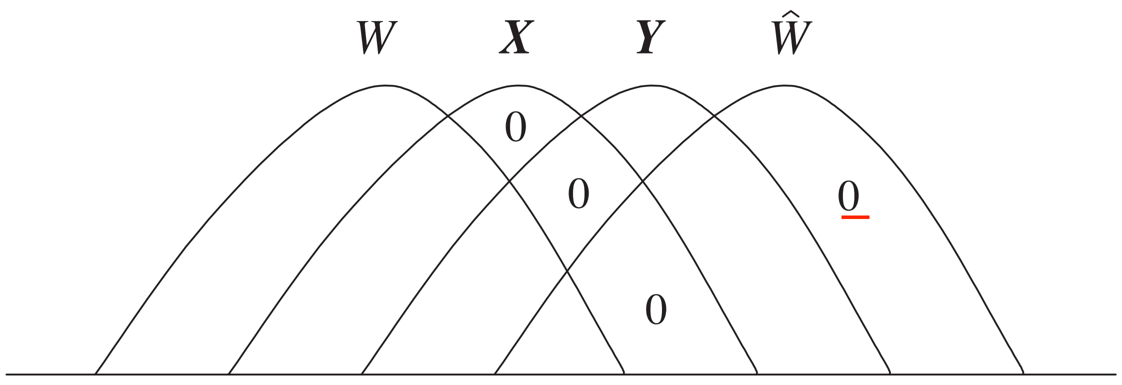
Why C is related to I(X;Y)?



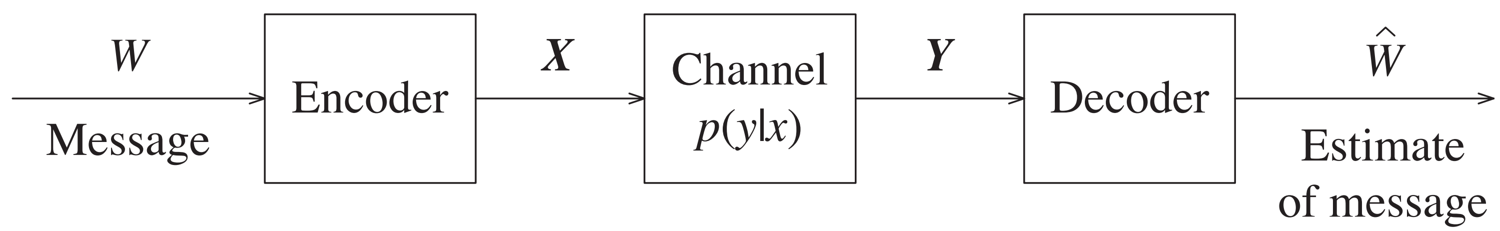
- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$



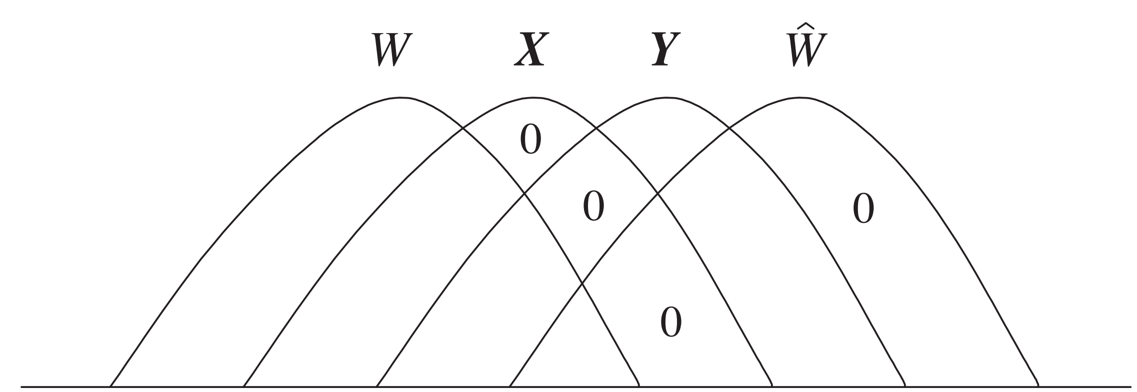
Why C is related to I(X;Y)?



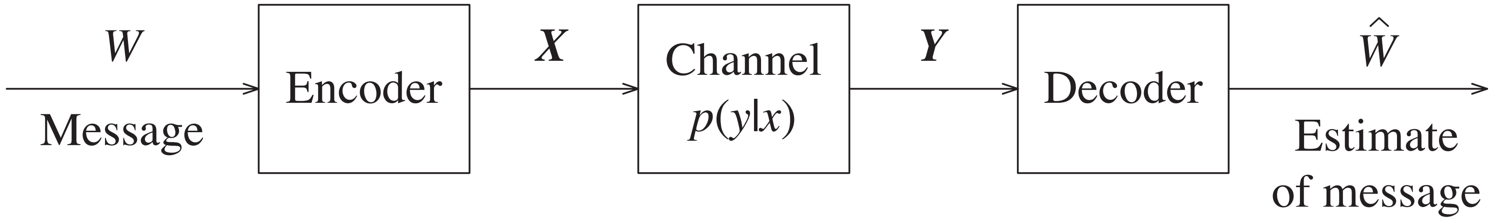
- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$



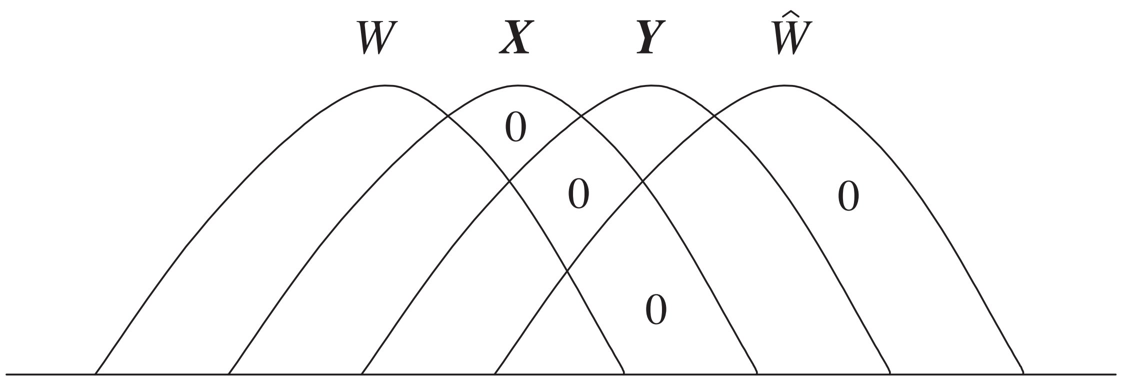
Why C is related to I(X;Y)?



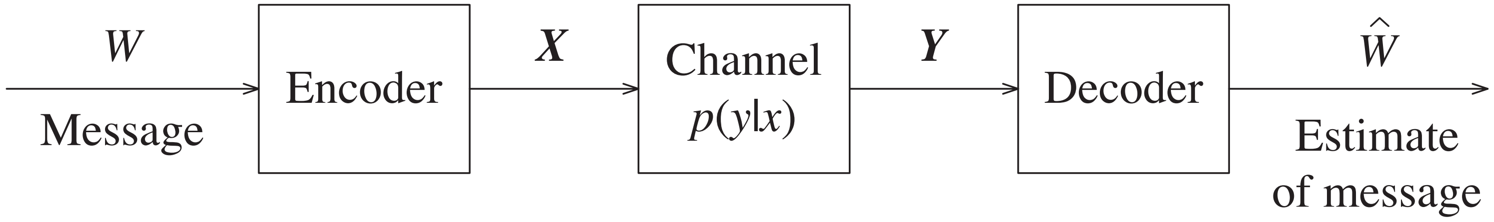
- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume



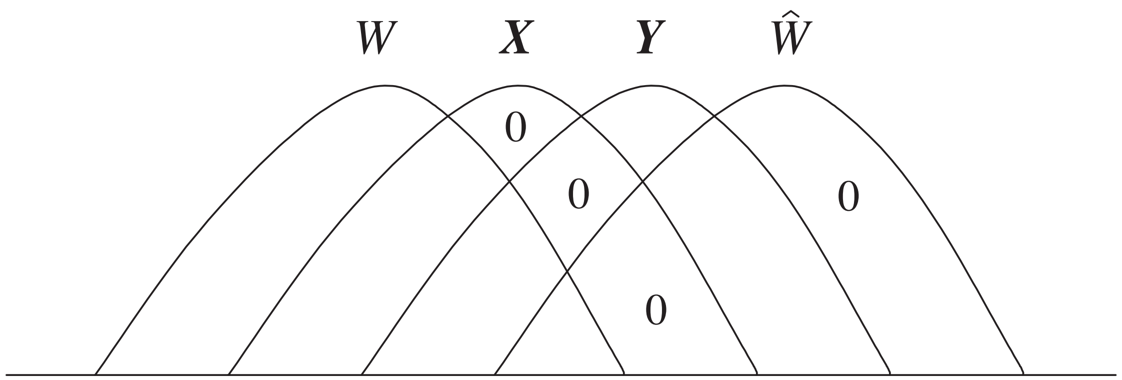
Why C is related to I(X;Y)?



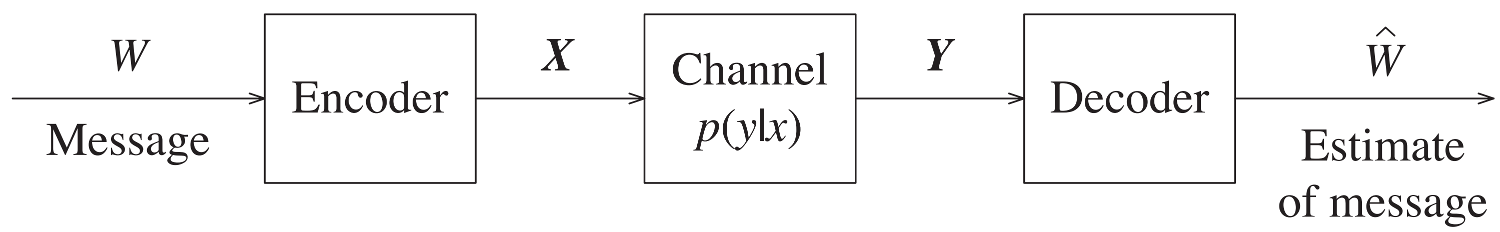
- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume



Why C is related to I(X;Y)?

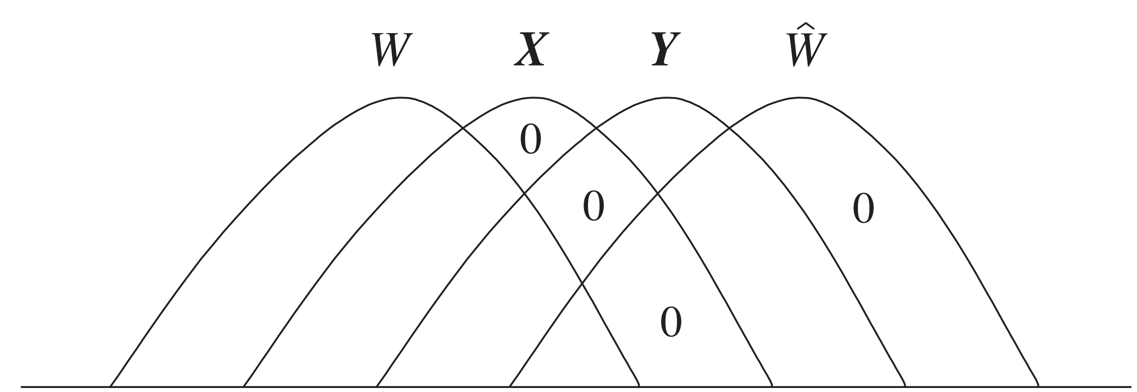


- Consider the information diagram for

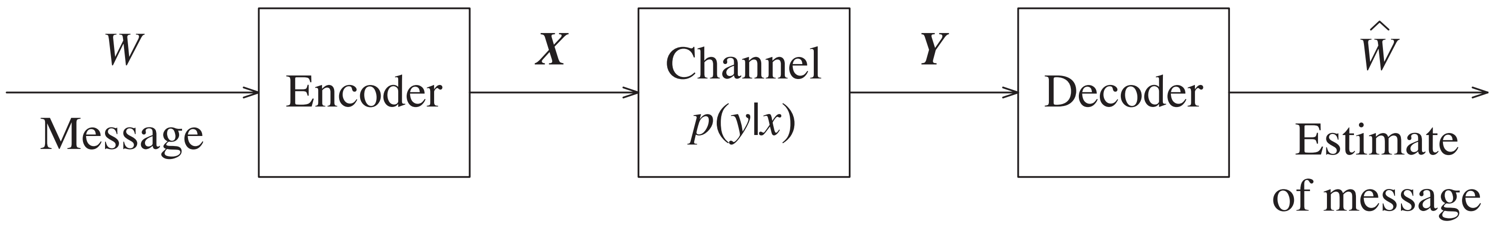
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$



Why C is related to I(X;Y)?

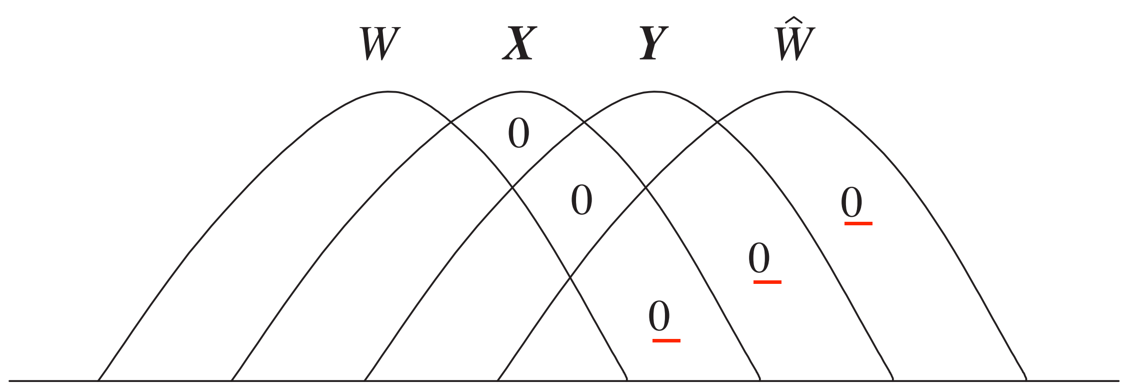


- Consider the information diagram for

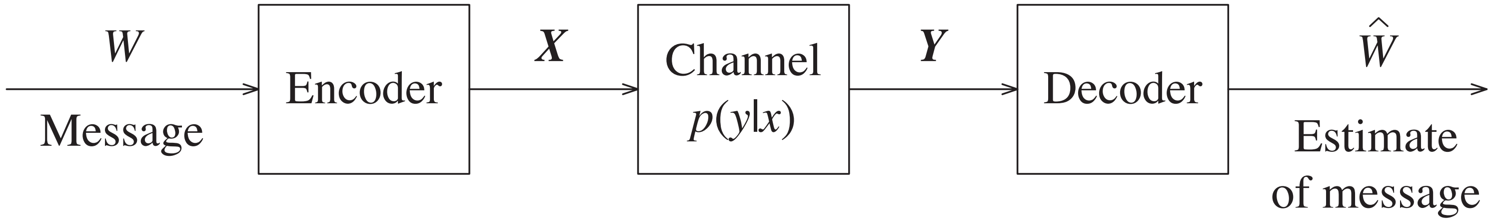
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$



Why C is related to I(X;Y)?

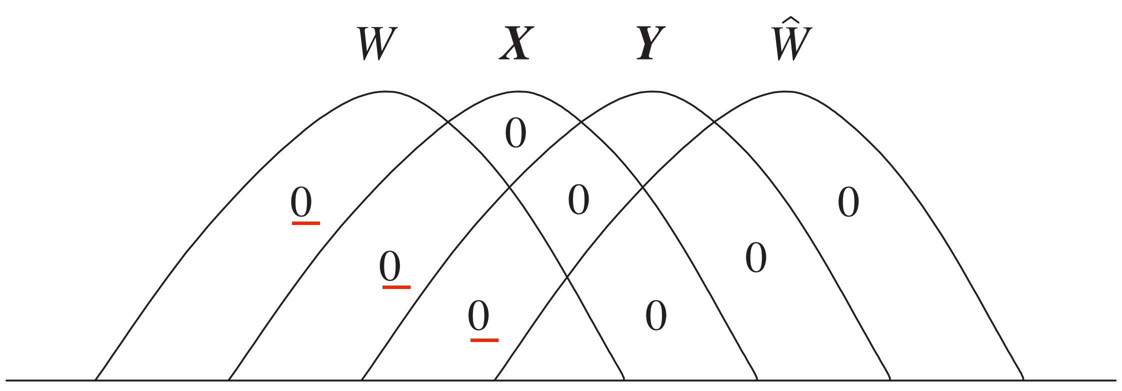


- Consider the information diagram for

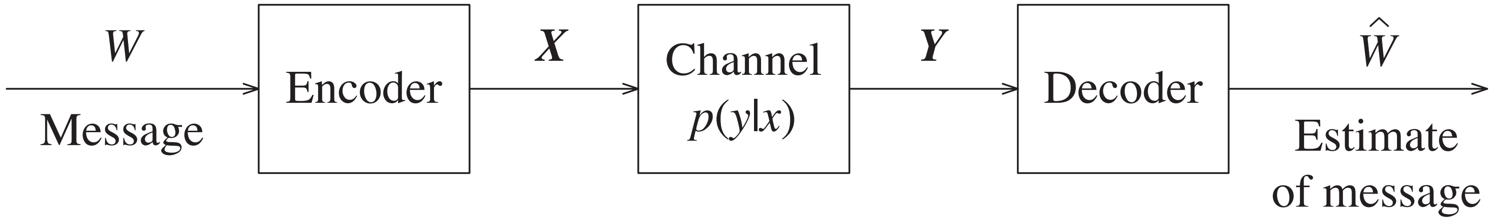
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$



Why C is related to I(X;Y)?

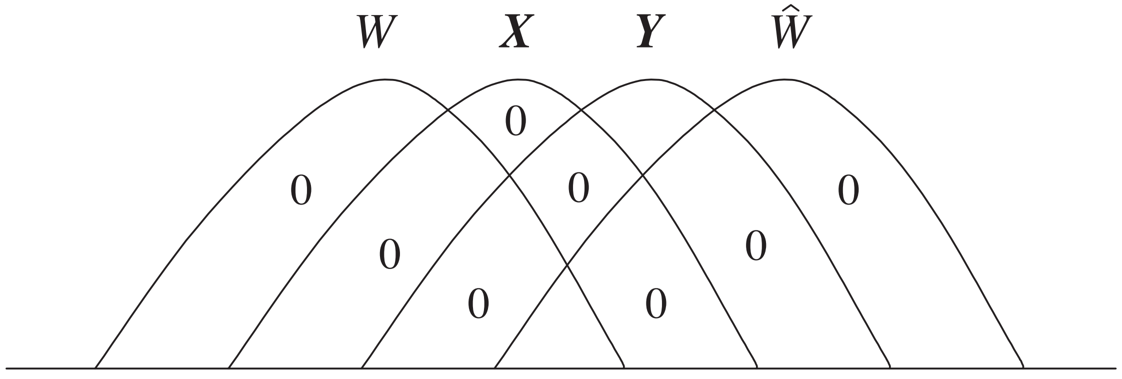


- Consider the information diagram for

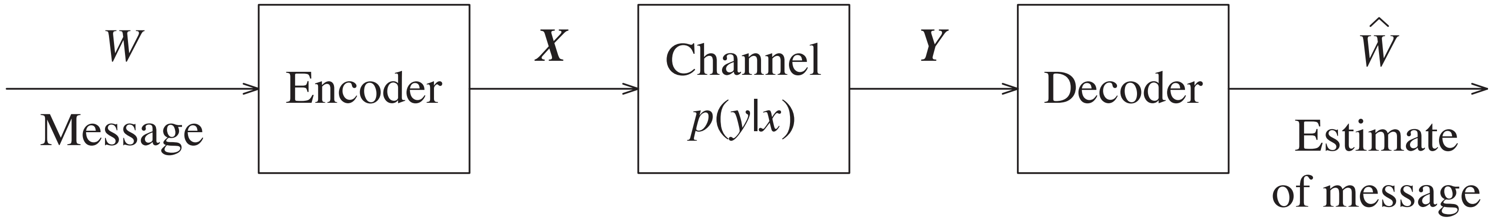
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$



Why C is related to I(X;Y)?



- Consider the information diagram for

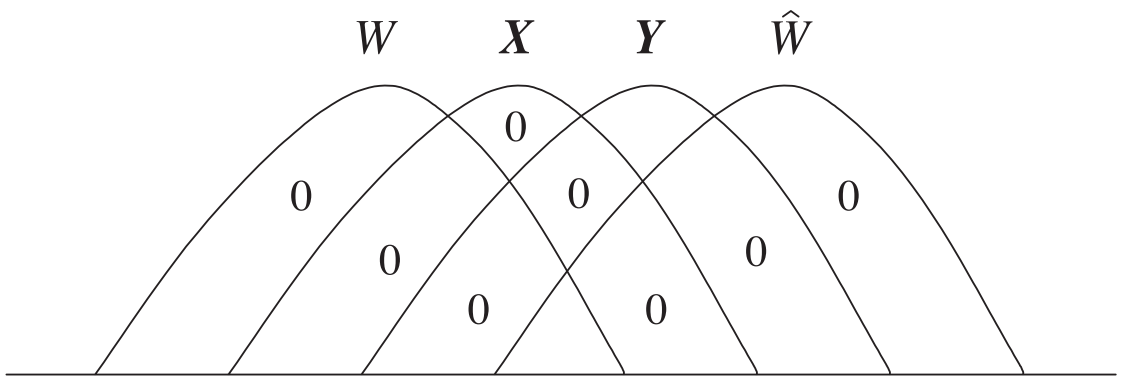
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

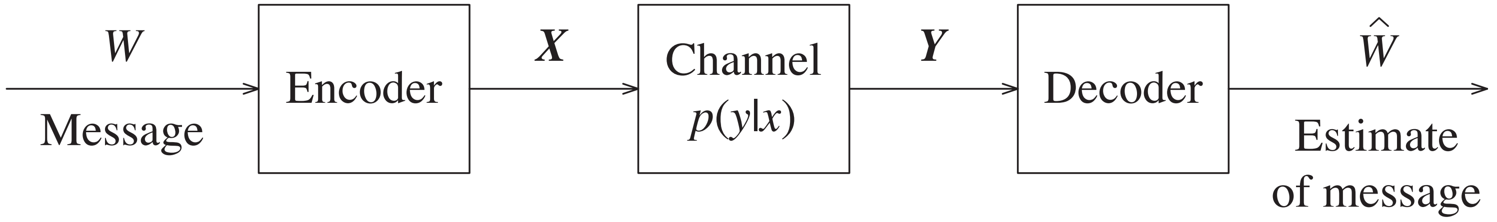
$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$

- Then we see that

$$H(W) = I(\mathbf{X};\mathbf{Y}).$$



Why C is related to I(X;Y)?



- Consider the information diagram for

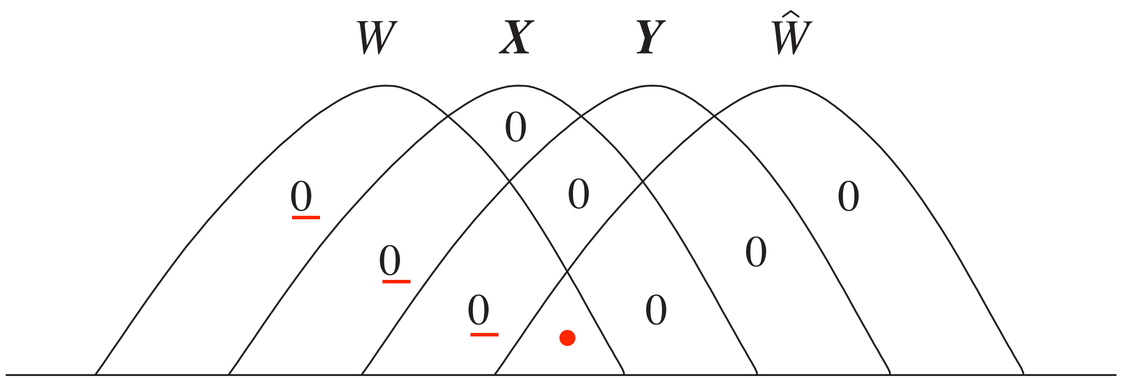
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

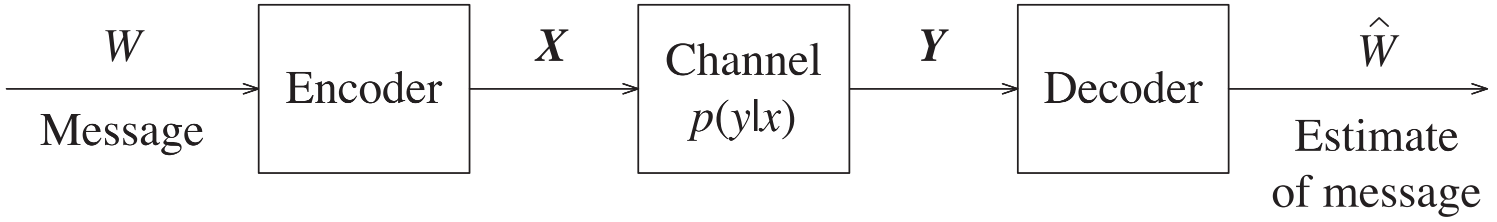
$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$

- Then we see that

$$H(W) = I(\mathbf{X};\mathbf{Y}).$$



Why C is related to I(X;Y)?



- Consider the information diagram for

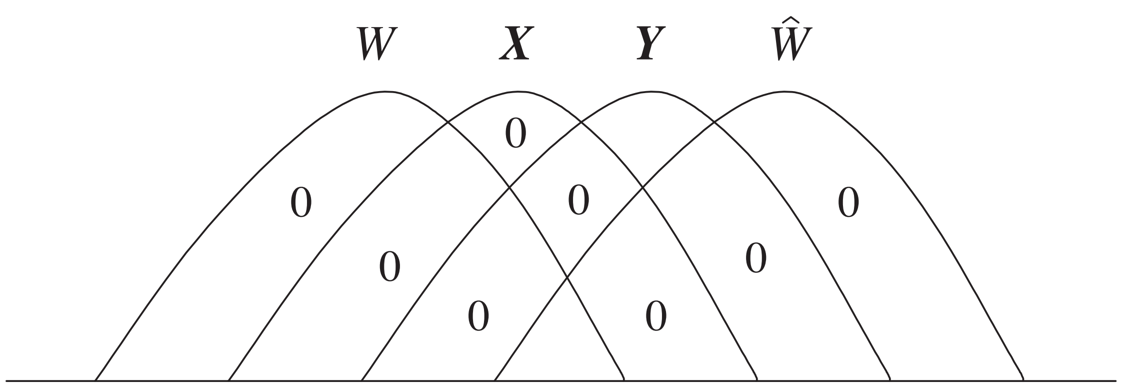
$$W \rightarrow \mathbf{X} \rightarrow \mathbf{Y} \rightarrow \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

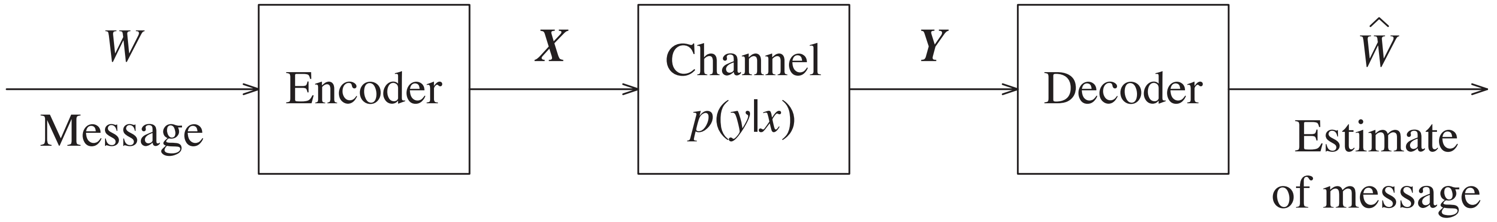
$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$

- Then we see that

$$H(W) = I(\mathbf{X};\mathbf{Y}).$$



Why C is related to I(X;Y)?



- Consider the information diagram for

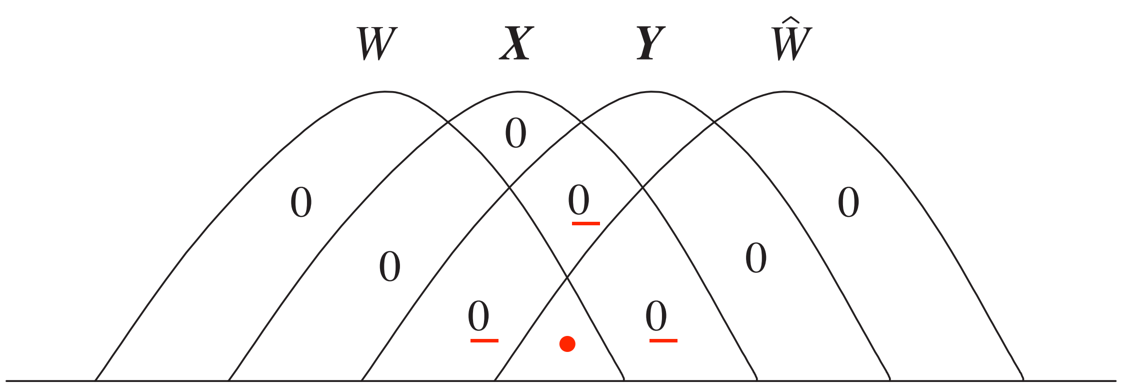
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

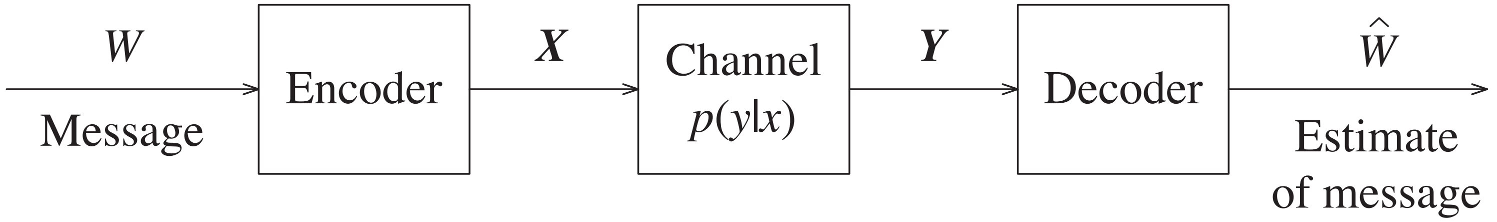
$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$

- Then we see that

$$H(W) = I(\mathbf{X};\mathbf{Y}).$$



Why C is related to I(X;Y)?

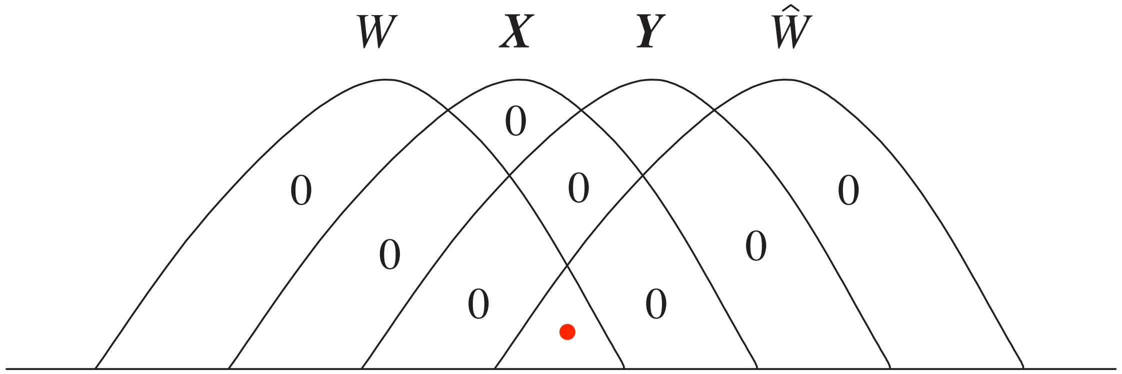


- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

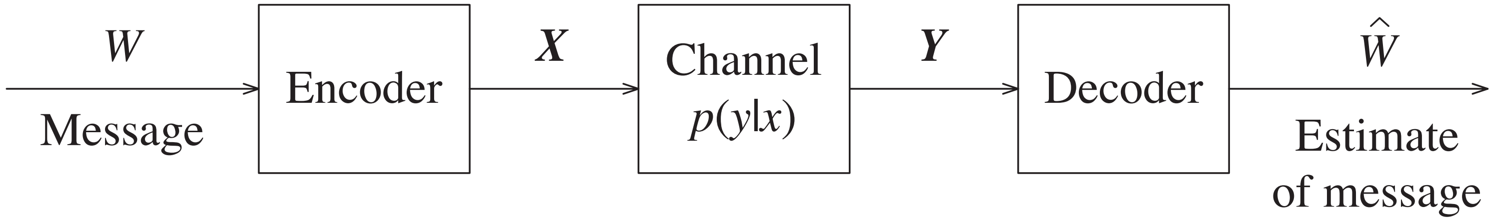
$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$



- Then we see that

$$H(W) = I(\mathbf{X};\mathbf{Y}).$$

Why C is related to I(X;Y)?



- Consider the information diagram for

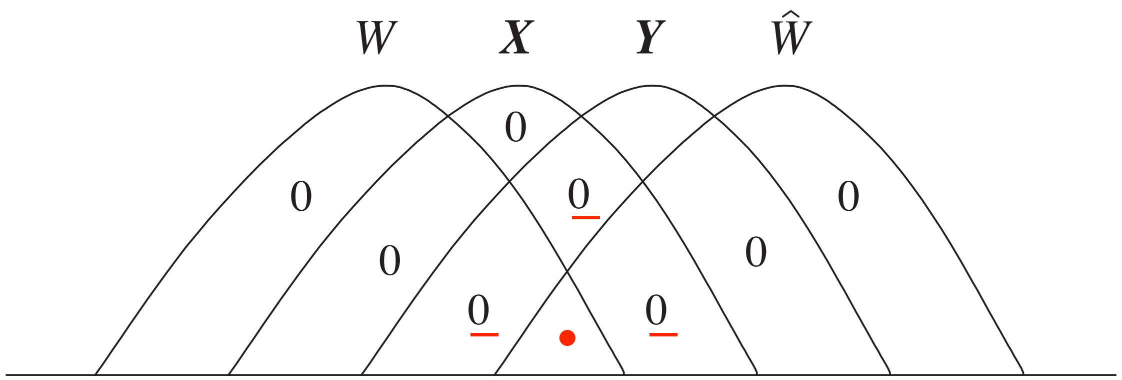
$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$

- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

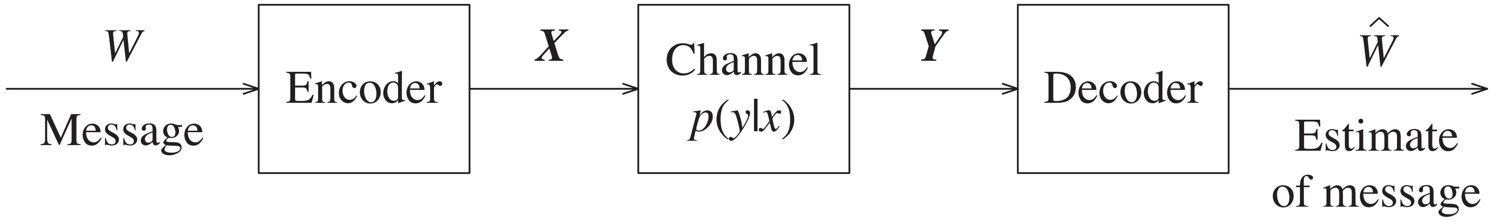
$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$

- Then we see that

$$H(W) = I(\mathbf{X};\mathbf{Y}).$$

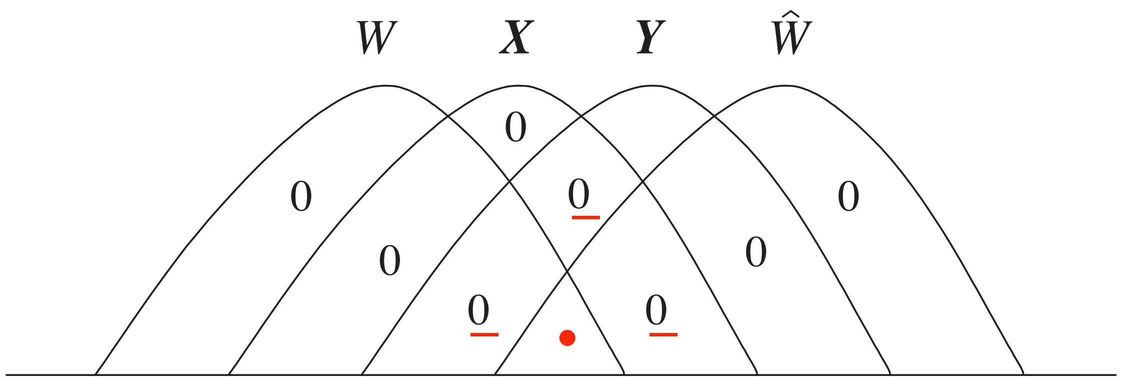


Why C is related to I(X;Y)?



- Consider the information diagram for

$$W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}.$$



- $H(\mathbf{X}|W) = 0$
- $H(\hat{W}|\mathbf{Y}) = 0$
- Since W and $\hat{W}$ are essentially identical for reliable communication, assume

$$H(\hat{W}|W) = H(W|\hat{W}) = 0.$$

- Then we see that

$$H(W) = I(\mathbf{X};\mathbf{Y}).$$

- This suggests that the channel capacity is obtained by maximizing $I(X;Y)$.

Building Blocks of the Converse

Building Blocks of the Converse

- For all $1 \leq i \leq n$,
$$I(X_i; Y_i) \leq C = \max_{p(x)} I(X; Y)$$

Building Blocks of the Converse

- For all $1 \leq i \leq n$,

$$I(X_i; Y_i) \leq C = \max_{p(x)} I(X; Y)$$

- Then

$$\sum_{i=1}^{n} I(X_i; Y_i) \leq nC$$

Building Blocks of the Converse

- For all $1 \leq i \leq n$,

$$I(X_i; Y_i) \leq C = \max_{p(x)} I(X; Y)$$

- Then

$$\sum_{i=1}^{n} I(X_i; Y_i) \leq nC$$

- To be established in Lemma 7.16,

$$I(\mathbf{X}; \mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i; Y_i).$$

Building Blocks of the Converse

- For all $1 \leq i \leq n$,
$$I(X_i;Y_i) \leq C = \max_{p(x)} I(X;Y)$$

- Then
$$\sum_{i=1}^{n} I(X_i;Y_i) \leq nC$$

- To be established in Lemma 7.16,
$$I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i).$$

- Therefore,
$$\frac{1}{n}\log M = \frac{1}{n}\log|\mathcal{W}| = \frac{1}{n}H(W) \approx \frac{1}{n}I(\mathbf{X};\mathbf{Y}) \leq \frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq C.$$

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Proof

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Proof

1. From the previous proposition, we have

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Proof

1. From the previous proposition, we have

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$$

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Proof

1. From the previous proposition, we have

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$$

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Proof

1. From the previous proposition, we have

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$$

2. Then

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Proof

1. From the previous proposition, we have

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$$

2. Then

$$I(\mathbf{X};\mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X})$$

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Proof

1. From the previous proposition, we have

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$$

2. Then

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X}) \\ &\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \end{aligned}$$

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Proof

1. From the previous proposition, we have

$$H(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} H(Y_i|X_i)$$

2. Then

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X}) \\ &\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\ &= \sum_{i=1}^{n} I(X_i;Y_i). \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\log M = H(W)$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \end{aligned}$$

Data Processing Theorem If $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$, then

$$I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y}).$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \le C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \end{aligned}$$

Data Processing Theorem If $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$, then

$$I(W;\hat{W}) \le I(\mathbf{X};\mathbf{Y}).$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \end{aligned}$$

Data Processing Theorem If $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$, then

$$I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y}).$$

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \le C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i)
\end{aligned}$$

Data Processing Theorem If $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$, then

$$I(W;\hat{W}) \le I(\mathbf{X};\mathbf{Y}).$$

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \le \sum_{i=1}^{n} I(X_i;Y_i)$.

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

Data Processing Theorem If $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$, then

$$I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y}).$$

Lemma 7.16 $I(\mathbf{X};\mathbf{Y}) \leq \sum_{i=1}^{n} I(X_i;Y_i)$.

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

Data Processing Theorem If $W \to \mathbf{X} \to \mathbf{Y} \to \hat{W}$, then

$$I(W;\hat{W}) \leq I(\mathbf{X};\mathbf{Y}).$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$H(W|\hat{W}) < 1 + P_e \log |\mathcal{W}|$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned}
H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\
&= 1 + P_e \log M.
\end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \le C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\le H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned}
H(W|\hat{W}) &< 1 + P_e \log|\mathcal{W}| \\
&= 1 + P_e \log M.
\end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned}
H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\
&= 1 + P_e \log M.
\end{aligned}$$

4. Then,

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\log M \leq H(W|\hat{W}) + nC$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\log M \leq H(W|\hat{W}) + nC$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \le C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\le H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\le H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\le H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + \underline{P_e} \log M + nC \\ &\leq 1 + \underline{\lambda_{max}} \log M + nC \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned}
H(W|\hat{W}) &< 1 + P_e \log|\mathcal{W}| \\
&= 1 + P_e \log M.
\end{aligned}$$

4. Then,

$$\begin{aligned}
\log M &\leq H(W|\hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\leq 1 + \lambda_{max} \log M + nC
\end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log|\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned}
H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\
&= 1 + P_e \log M.
\end{aligned}$$

4. Then,

$$\begin{aligned}
\log M &\leq H(W|\hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\leq 1 + \lambda_{max} \log M + nC \\
&< 1 + \underline{\epsilon \log M} + nC,
\end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log|\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned}
H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\
&= 1 + P_e \log M.
\end{aligned}$$

4. Then,

$$\begin{aligned}
\log M &\leq H(W|\hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\leq 1 + \lambda_{max} \log M + nC \\
&< 1 + \epsilon \log M + nC,
\end{aligned}$$

or

$$(1-\epsilon) \log M < 1 + nC$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon) \log M &< 1 + nC \\ \log M &< \frac{1 + nC}{1 - \epsilon} \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log|\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max}\log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1 + nC}{1-\epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}. \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \le C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\le H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\le H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log|\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\le H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\le 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1 + nC}{1 - \epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1 - \epsilon}. \end{aligned}$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1+nC}{1-\epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}. \end{aligned}$$

5. Therefore,

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \le C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\le H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\le H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log|\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\le H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\le 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1+nC}{1-\epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}. \end{aligned}$$

5. Therefore,

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \le C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n}\log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\le H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\le H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned}
H(W|\hat{W}) &< 1 + P_e \log|\mathcal{W}| \\
&= 1 + P_e \log M.
\end{aligned}$$

4. Then,

$$\begin{aligned}
\log M &\le H(W|\hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\le 1 + \lambda_{max} \log M + nC \\
&< 1 + \epsilon \log M + nC,
\end{aligned}$$

or

$$\begin{aligned}
(1-\epsilon)\log M &< 1 + nC \\
\log M &< \frac{1+nC}{1-\epsilon} \\
\frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}.
\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon) \log M &< 1 + nC \\ \log M &< \frac{1 + nC}{1 - \epsilon} \\ \frac{1}{n} \log M &< \frac{\frac{1}{n} + C}{1 - \epsilon}. \end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n} \log M$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \le C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\le H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\le H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\le H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\le H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\le 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon) \log M &< 1 + nC \\ \log M &< \frac{1 + nC}{1 - \epsilon} \\ \frac{1}{n} \log M &< \frac{\frac{1}{n} + C}{1 - \epsilon}. \end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n} \log M < \frac{\frac{1}{n} + C}{1 - \epsilon}.$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1 + nC}{1 - \epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1 - \epsilon}. \end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{\frac{1}{n} + C}{1 - \epsilon}.$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned} \log M &= H(W) \\ &= H(W|\hat{W}) + I(W;\hat{W}) \\ &\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\ &\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\ &\leq H(W|\hat{W}) + nC. \end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned} H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\ &= 1 + P_e \log M. \end{aligned}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon) \log M &< 1 + nC \\ \log M &< \frac{1 + nC}{1 - \epsilon} \\ \frac{1}{n} \log M &< \frac{\frac{1}{n} + C}{1 - \epsilon}. \end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{\frac{1}{n} + C}{1 - \epsilon}.$$

Theorem 7.15 (Channel Coding Theorem) A rate R is achievable for a discrete memoryless channel if and only if $R \leq C$, the capacity of the channel.

Proof of Converse

1. Let R be an achievable rate, i.e., for any $\epsilon > 0$, there exists for sufficiently large n an (n, M) code such that

$$\frac{1}{n} \log M > R - \epsilon \quad \text{and} \quad \lambda_{max} < \epsilon.$$

2. Consider

$$\begin{aligned}
\log M &= H(W) \\
&= H(W|\hat{W}) + I(W;\hat{W}) \\
&\leq H(W|\hat{W}) + I(\mathbf{X};\mathbf{Y}) \\
&\leq H(W|\hat{W}) + \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq H(W|\hat{W}) + nC.
\end{aligned}$$

3. By Fano's inequality,

$$\begin{aligned}
H(W|\hat{W}) &< 1 + P_e \log |\mathcal{W}| \\
&= 1 + P_e \log M.
\end{aligned}$$

4. Then,

$$\begin{aligned}
\log M &\leq H(W|\hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\leq 1 + \lambda_{max} \log M + nC \\
&< 1 + \epsilon \log M + nC,
\end{aligned}$$

or

$$\begin{aligned}
(1-\epsilon)\log M &< 1 + nC \\
\log M &< \frac{1+nC}{1-\epsilon} \\
\frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}.
\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{\frac{1}{n} + C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

4. Then,

$$\begin{aligned}
\log M &\leq H(W \mid \hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\leq 1 + \lambda_{max} \log M + nC \\
&< 1 + \epsilon \log M + nC,
\end{aligned}$$

or

$$\begin{aligned}
(1 - \epsilon) \log M &< 1 + nC \\
\log M &< \frac{1 + nC}{1 - \epsilon} \\
\frac{1}{n} \log M &< \frac{\frac{1}{n} + C}{1 - \epsilon}.
\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n} \log M < \frac{\frac{1}{n} + C}{1 - \epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

4. Then,

$$\begin{aligned}
\log M &\leq H(W|\hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\leq 1 + \lambda_{max} \log M + nC \\
&< 1 + \epsilon \log M + nC,
\end{aligned}$$

or

$$\begin{aligned}
(1-\epsilon)\log M &< 1 + nC \\
\log M &< \frac{1+nC}{1-\epsilon} \\
\frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}.
\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M < \frac{\frac{1}{n} + C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

4. Then,

$$\begin{aligned}
\log M &\leq H(W|\hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\leq 1 + \lambda_{max} \log M + nC \\
&< 1 + \epsilon \log M + nC,
\end{aligned}$$

or

$$\begin{aligned}
(1-\epsilon)\log M &< 1 + nC \\
\log M &< \frac{1+nC}{1-\epsilon} \\
\frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}.
\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M < \frac{\frac{1}{n} + C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

Corollary For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

Proof

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1+nC}{1-\epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}. \end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M < \frac{\frac{1}{n} + C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

Corollary For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

Proof

4. Then,

$$\begin{aligned}
\log M &\le H(W|\hat{W}) + nC \\
&< 1 + P_e \log M + nC \\
&\le 1 + \lambda_{max} \log M + nC \\
&< 1 + \epsilon \log M + nC,
\end{aligned}$$

or

$$\begin{aligned}
(1-\epsilon)\log M &< 1 + nC \\
\log M &< \frac{1+nC}{1-\epsilon} \\
\frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}.
\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M < \frac{\frac{1}{n} + C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \le C$.

Corollary For large n,

$$P_e \ge 1 - \frac{C}{\frac{1}{n}\log M}.$$

Proof

1. Consider $\log M < 1 + P_e \log M + nC$.

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1+nC}{1-\epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n} + C}{1-\epsilon}. \end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M < \frac{\frac{1}{n} + C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

Corollary For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

Proof

1. Consider $\log M < 1 + P_e \log M + nC$.

4. Then,

$$\begin{aligned}\log M &\le H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\le 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC,\end{aligned}$$

or

$$\begin{aligned}(1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1+nC}{1-\epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n}+C}{1-\epsilon}.\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M < \frac{\frac{1}{n}+C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \le C$.

Corollary For large n,

$$P_e \ge 1 - \frac{C}{\frac{1}{n}\log M}.$$

Proof

1. Consider $\log M < 1 + P_e \log M + nC$.

4. Then,

$$\begin{aligned}\log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC,\end{aligned}$$

or

$$\begin{aligned}(1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1+nC}{1-\epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n}+C}{1-\epsilon}.\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M < \frac{\frac{1}{n}+C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

Corollary For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

Proof

1. Consider $\log M < 1 + P_e \log M + nC$.
2. Then

$$P_e \geq 1 - \frac{1+nC}{\log M}$$

4. Then,

$$\begin{aligned} \log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC, \end{aligned}$$

or

$$\begin{aligned} (1-\epsilon) \log M &< 1 + nC \\ \log M &< \frac{1 + nC}{1 - \epsilon} \\ \frac{1}{n} \log M &< \frac{\frac{1}{n} + C}{1 - \epsilon}. \end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n} \log M < \frac{\frac{1}{n} + C}{1 - \epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

Corollary For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n} \log M}.$$

Proof

1. Consider $\log M < 1 + P_e \log M + nC$.
2. Then

$$P_e \geq 1 - \frac{1 + nC}{\log M} = 1 - \frac{\frac{1}{n} + C}{\frac{1}{n} \log M}$$

4. Then,

$$\begin{aligned}\log M &\leq H(W|\hat{W}) + nC \\ &< 1 + P_e \log M + nC \\ &\leq 1 + \lambda_{max} \log M + nC \\ &< 1 + \epsilon \log M + nC,\end{aligned}$$

or

$$\begin{aligned}(1-\epsilon)\log M &< 1 + nC \\ \log M &< \frac{1+nC}{1-\epsilon} \\ \frac{1}{n}\log M &< \frac{\frac{1}{n}+C}{1-\epsilon}.\end{aligned}$$

5. Therefore,

$$R - \epsilon < \frac{1}{n}\log M < \frac{\frac{1}{n}+C}{1-\epsilon}.$$

6. Letting $n \to \infty$ and then $\epsilon \to 0$ to conclude that $R \leq C$.

Corollary For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

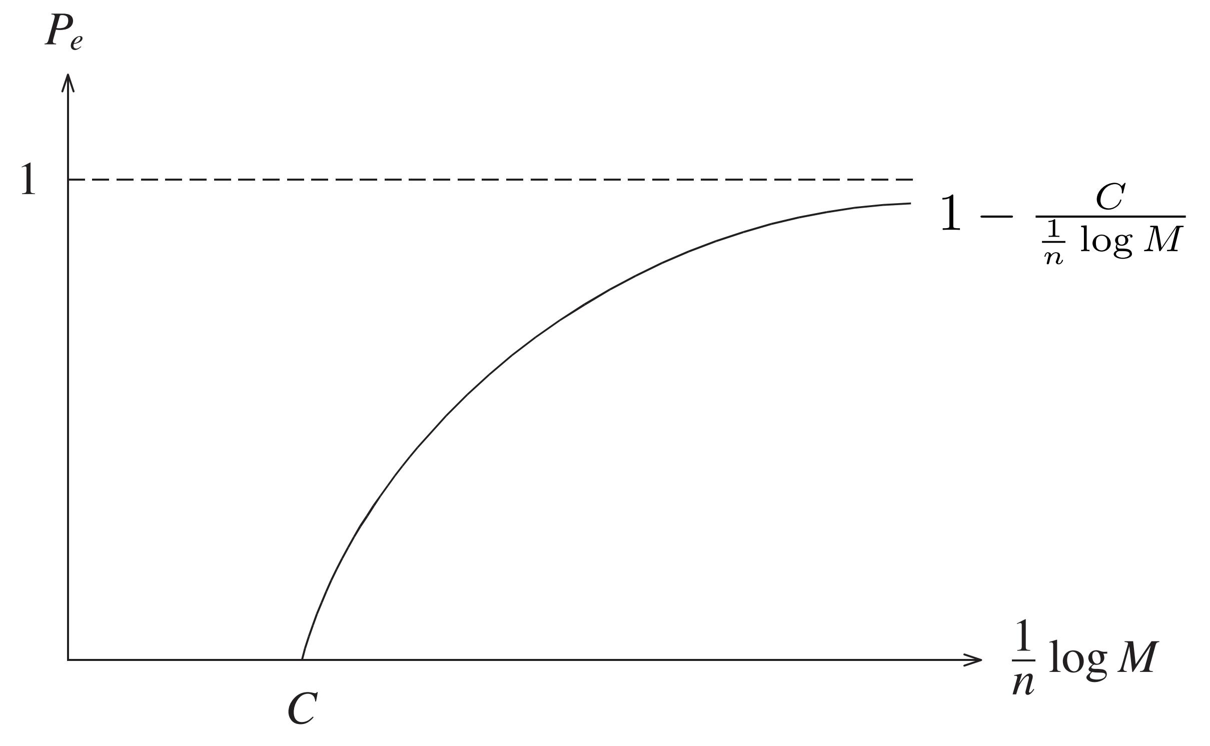
Proof

1. Consider $\log M < 1 + P_e \log M + nC$.
2. Then

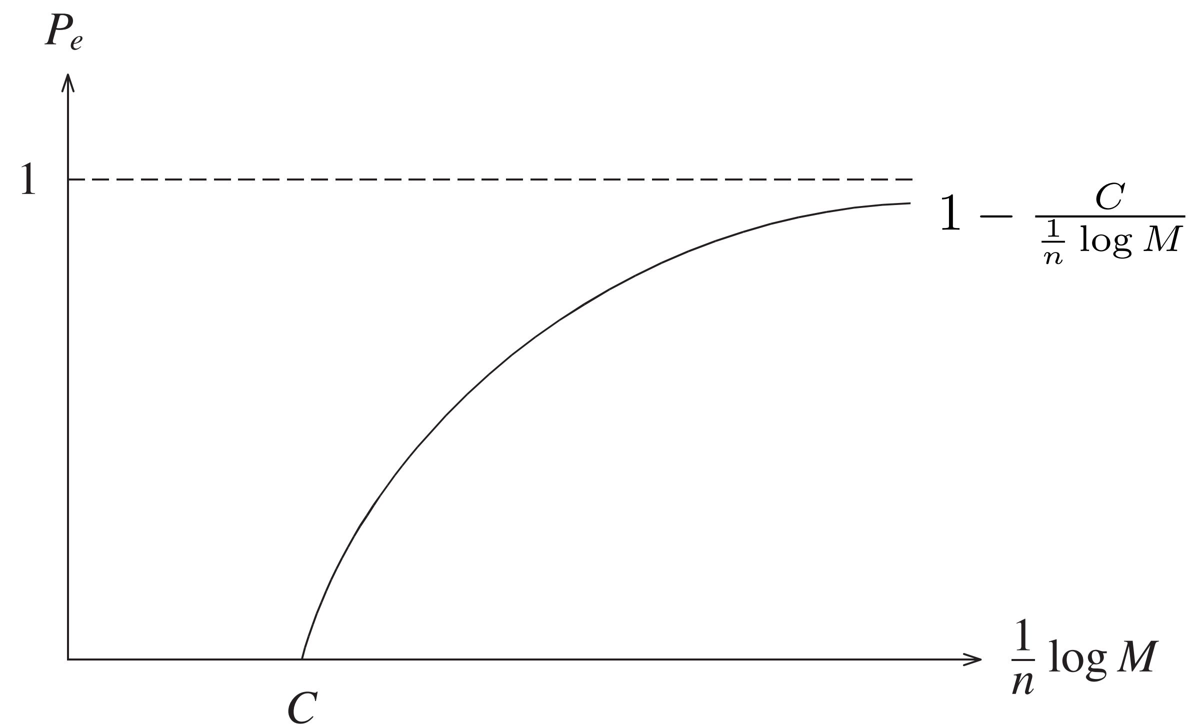
$$P_e \geq 1 - \frac{1+nC}{\log M} = 1 - \frac{\frac{1}{n}+C}{\frac{1}{n}\log M} \approx 1 - \frac{C}{\frac{1}{n}\log M}$$

for large n.

Asymptotic Analysis of P_e

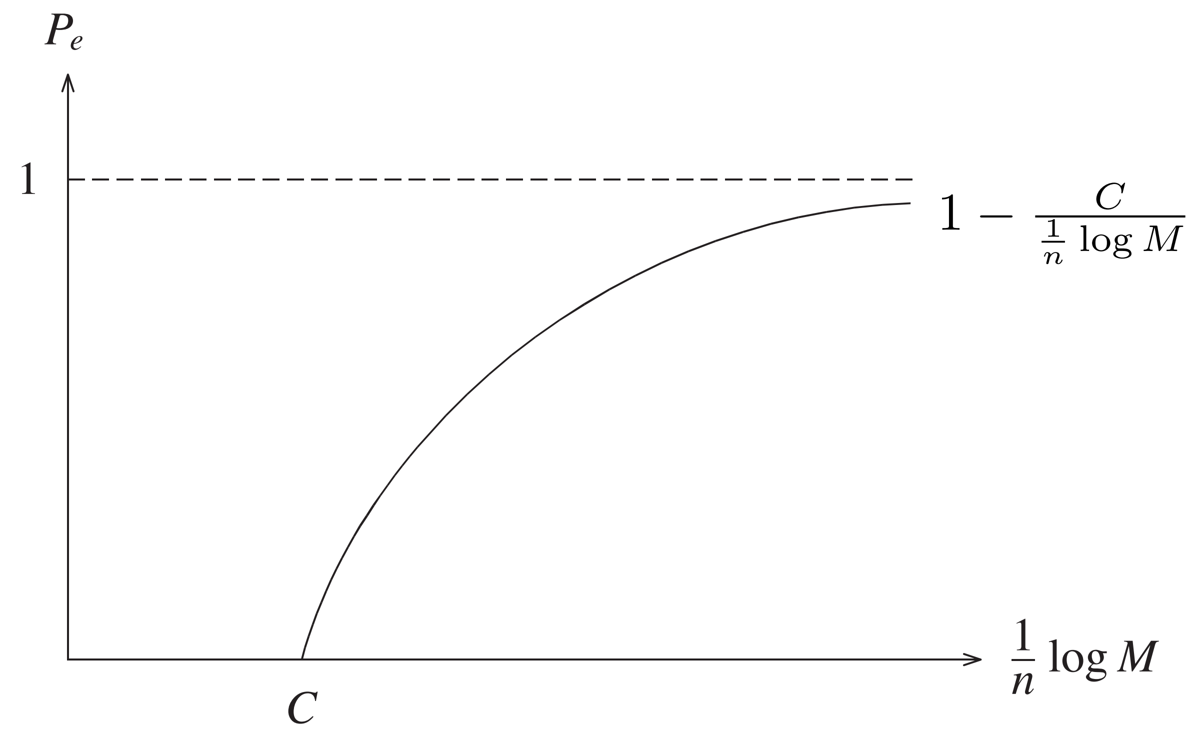


Asymptotic Analysis of P_e



- For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

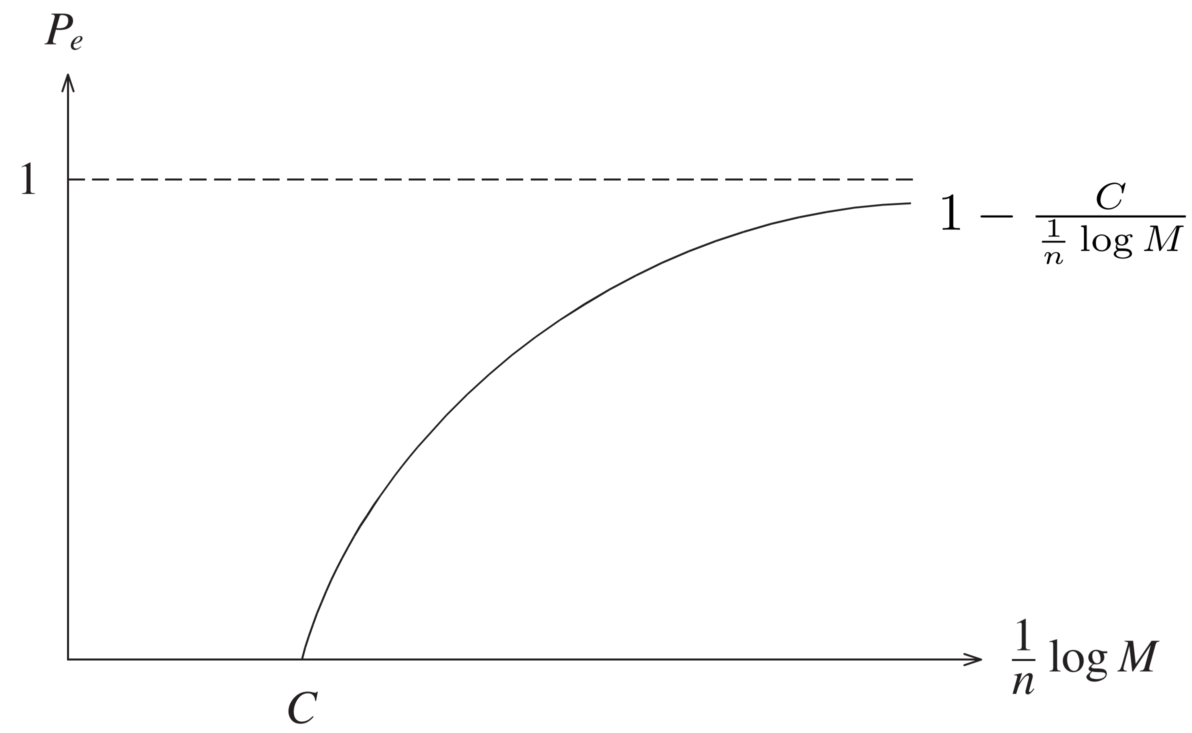


Asymptotic Analysis of P_e

- For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

- $\frac{1}{n}\log M$ is the rate of the channel code.

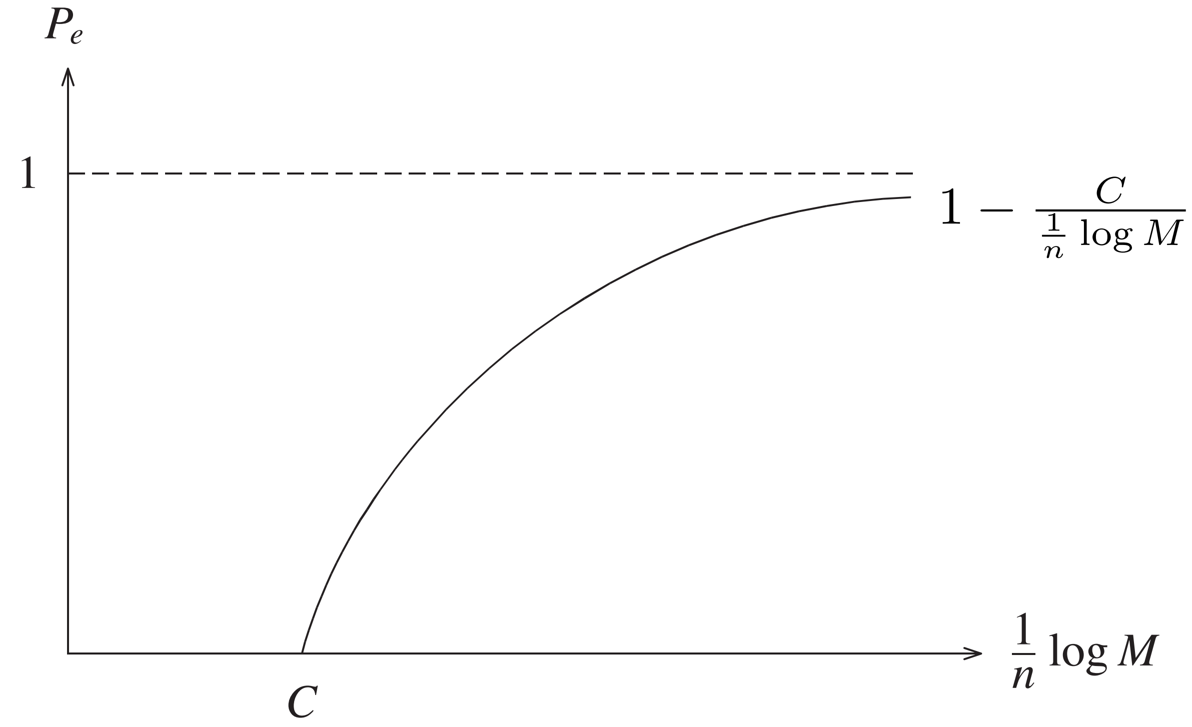


Asymptotic Analysis of P_e

- For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

- $\frac{1}{n}\log M$ is the rate of the channel code.



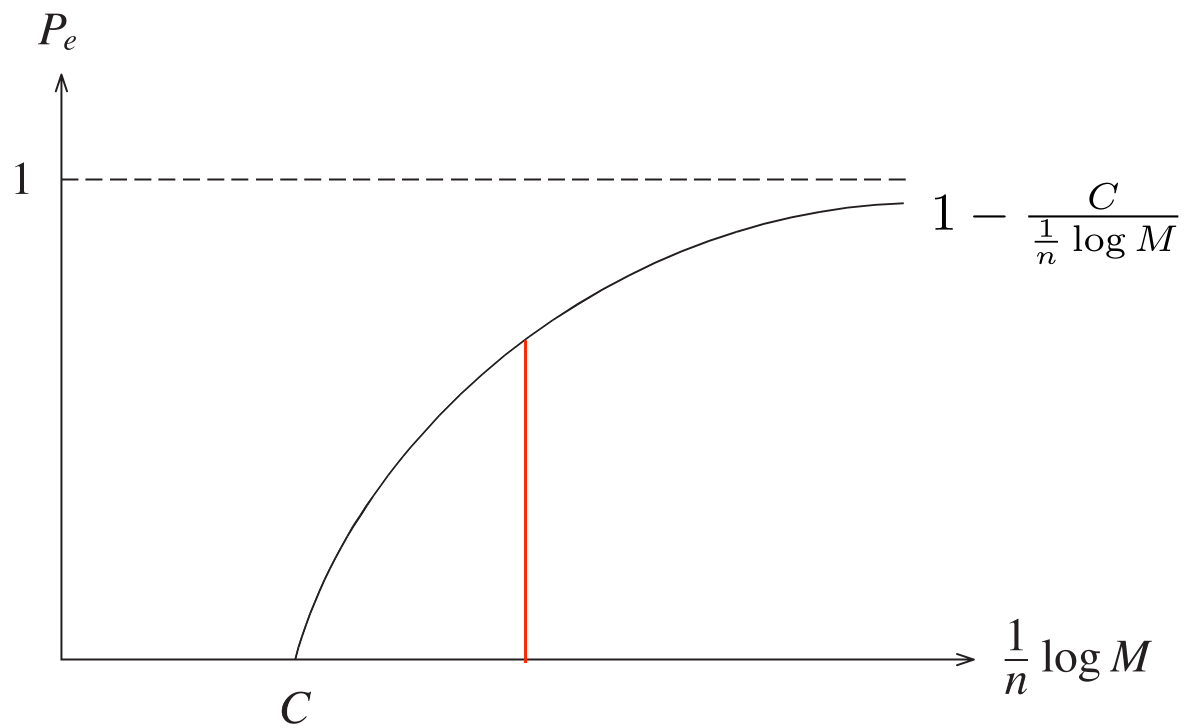
Asymptotic Analysis of P_e

- For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

- $\frac{1}{n}\log M$ is the rate of the channel code.

- If $\frac{1}{n}\log M > C$, then P_e is bounded away from 0 for large n.



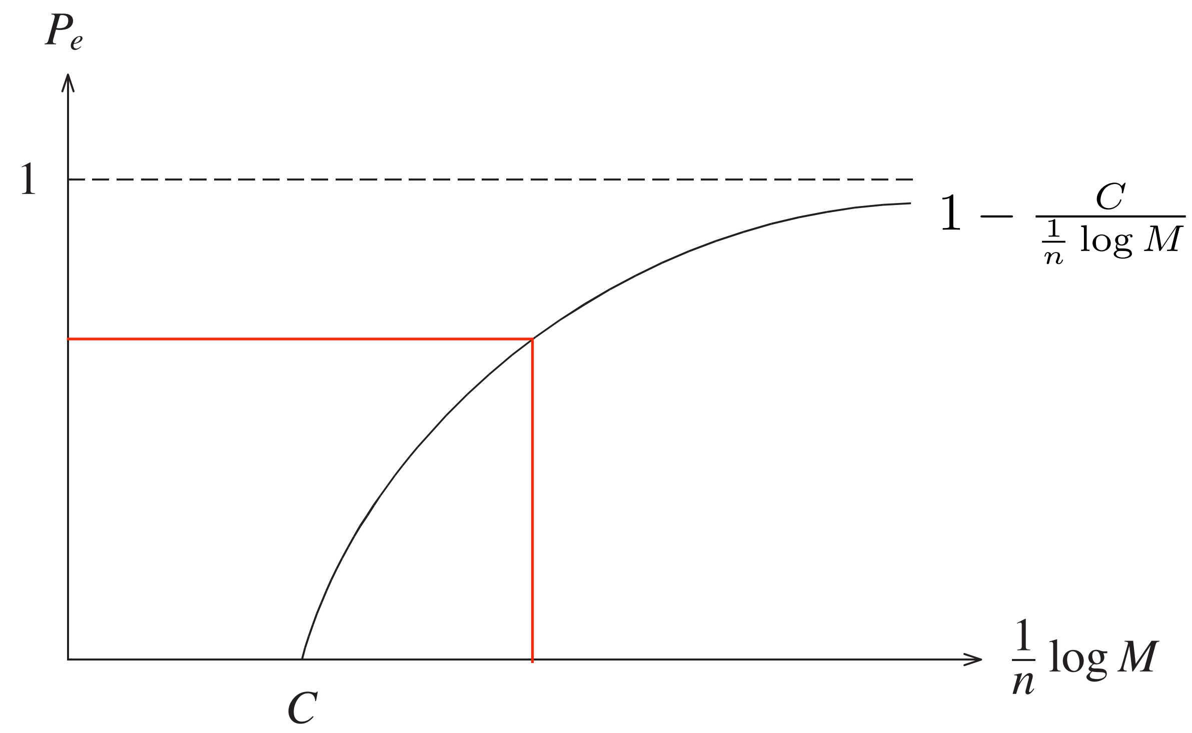
Asymptotic Analysis of P_e

- For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

- $\frac{1}{n}\log M$ is the rate of the channel code.

- If $\frac{1}{n}\log M > C$, then P_e is bounded away from 0 for large n.



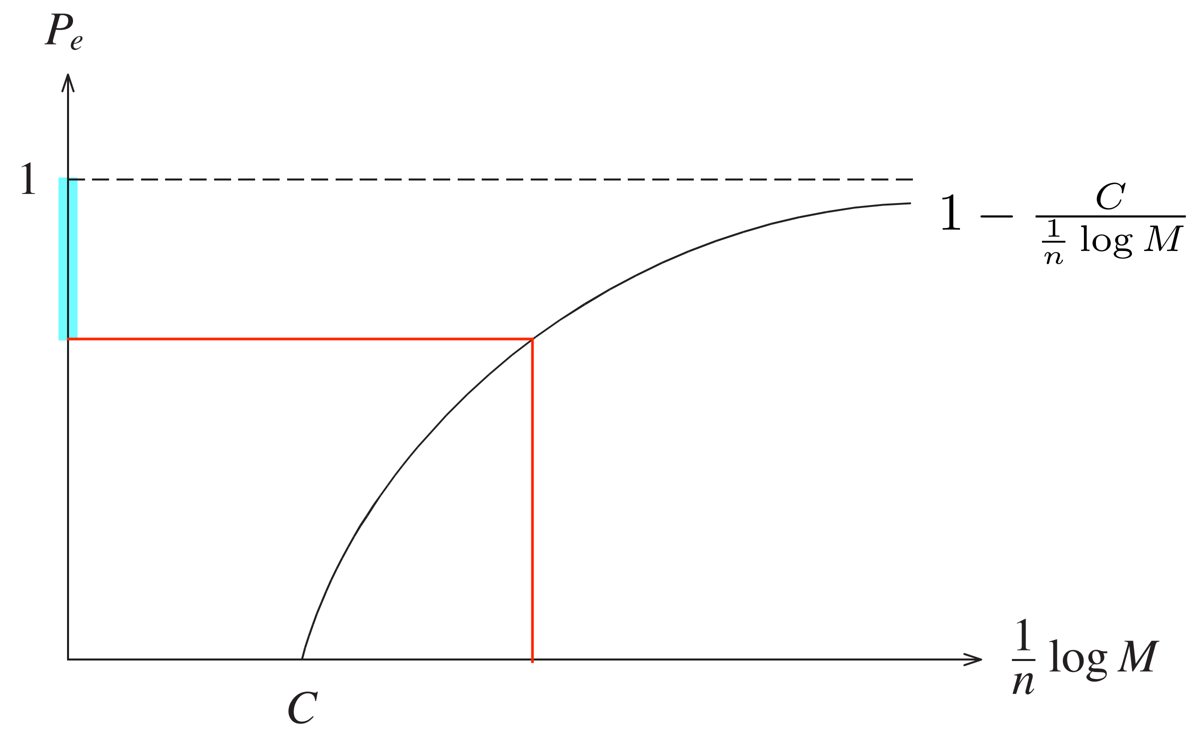
Asymptotic Analysis of P_e

- For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

- $\frac{1}{n}\log M$ is the rate of the channel code.

- If $\frac{1}{n}\log M > C$, then P_e is bounded away from 0 for large n.



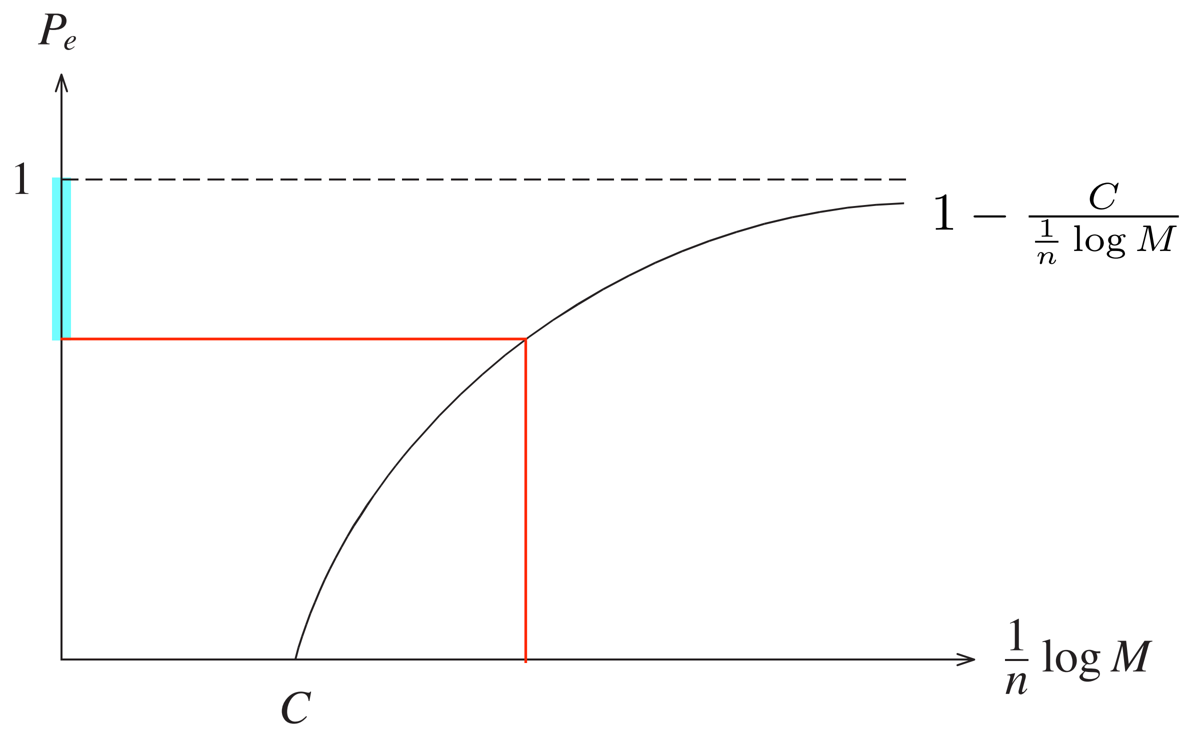
Asymptotic Analysis of P_e

- For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

- $\frac{1}{n}\log M$ is the rate of the channel code.

- If $\frac{1}{n}\log M > C$, then P_e is bounded away from 0 for large n.



Asymptotic Analysis of P_e

- For large n,

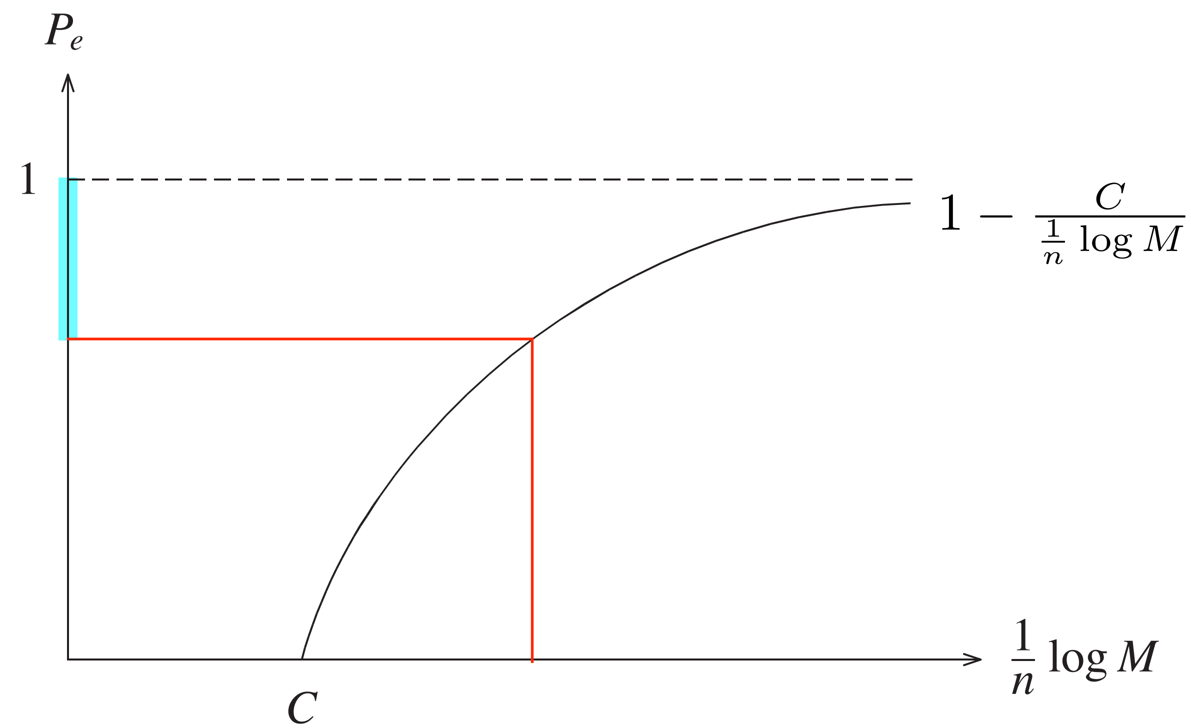
$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

- $\frac{1}{n}\log M$ is the rate of the channel code.

- If $\frac{1}{n}\log M > C$, then P_e is bounded away from 0 for large n.

- This implies that if $\frac{1}{n}\log M > C$, then

$$P_e > 0 \text{ for all } n.$$



Asymptotic Analysis of P_e

- For large n,

$$P_e \geq 1 - \frac{C}{\frac{1}{n}\log M}.$$

- $\frac{1}{n}\log M$ is the rate of the channel code.
- If $\frac{1}{n}\log M > C$, then P_e is bounded away from 0 for large n.
- This implies that if $\frac{1}{n}\log M > C$, then

$$P_e > 0 \text{ for all } n.$$

- Also note that this lower bound on P_e tends to 1 as $\frac{1}{n}\log M \to \infty$.