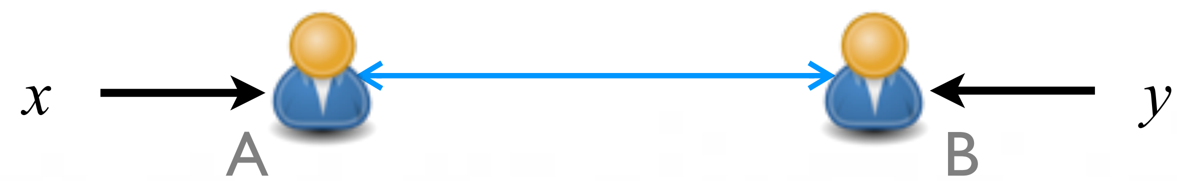
Rényi Information Complexity

Manoj Prabhakaran

Vinod Prabhakaran

Communication Complexity



How many bits do they need to exchange to compute $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$?

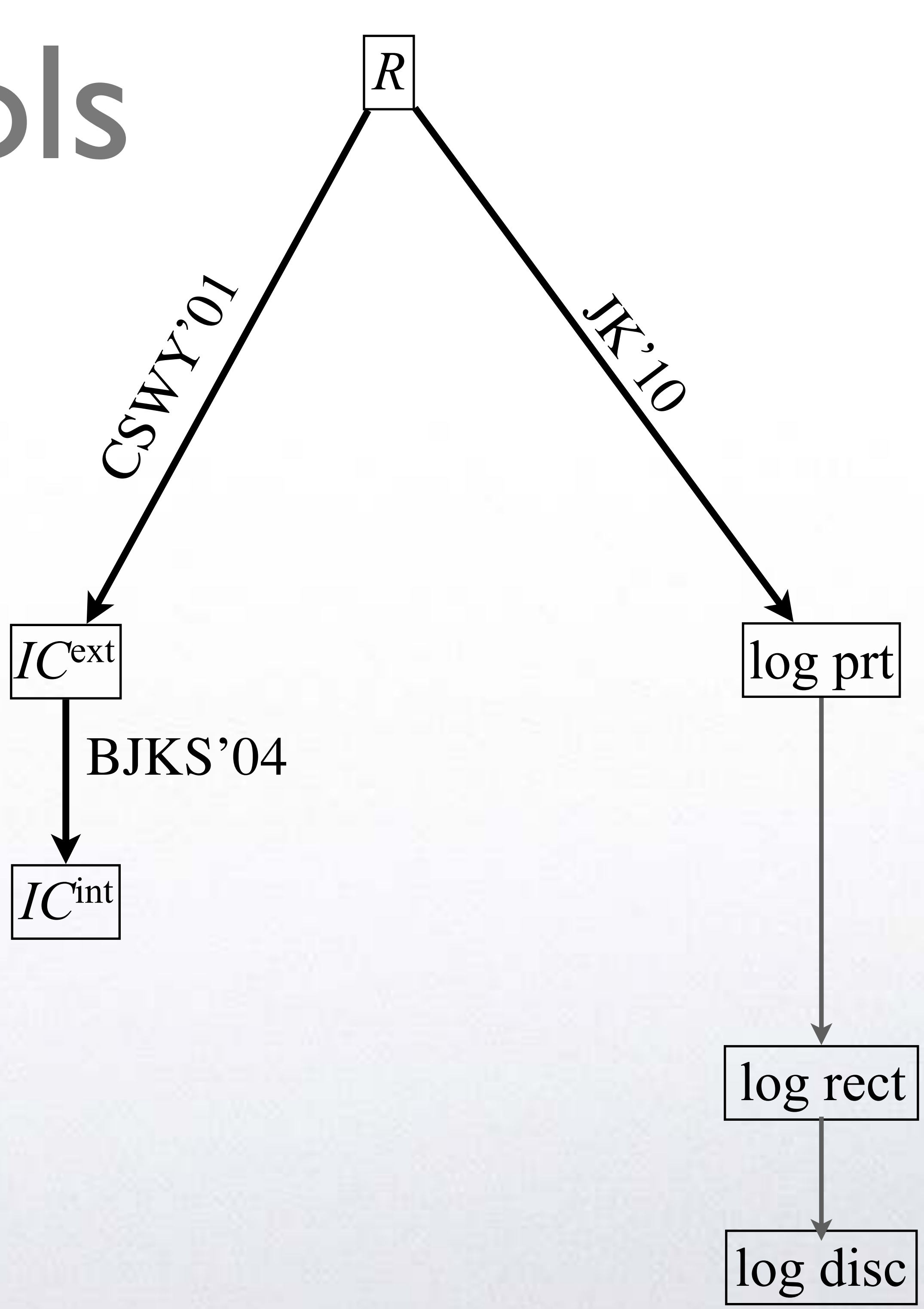
$$R(f, \varepsilon) = \min_{\substack{\text{rand protocol } \pi : \\ \text{err}(\pi, f) \leq \varepsilon}} \; \max_{(x,y)} \; \#\text{bits}(\pi(x,y))$$

$\forall (x,y) \;\; \Pr[\pi(x,y) \neq f(x,y)] \leq \varepsilon$

Connections to circuit complexity, data structures, streaming algorithms, property testing, game theory, …

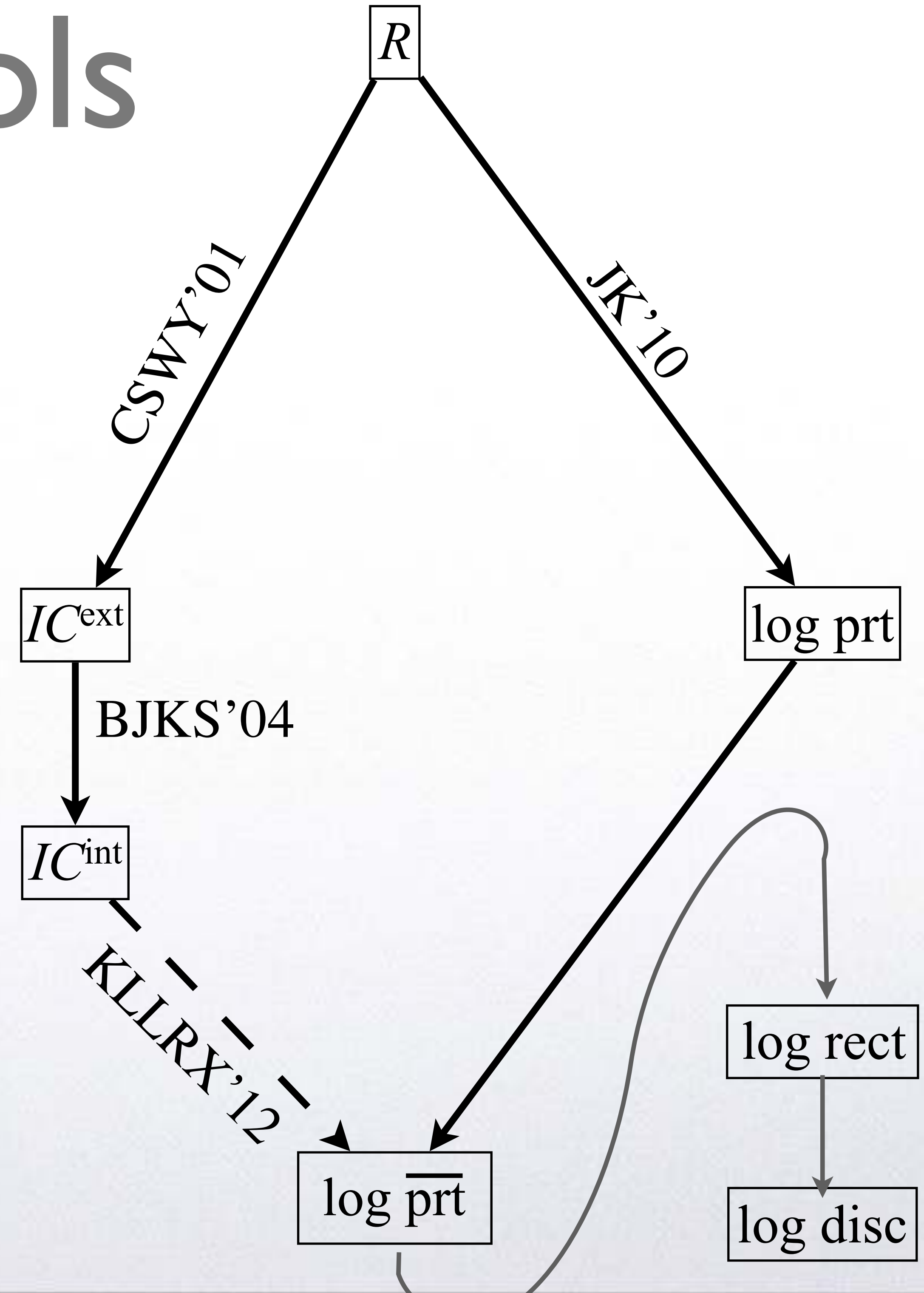
Lower Bound Tools

- Combinatorial bounds
  - Dominated by partition bounds [Jain-Klauck'10]
- Information theoretic bounds
  - Dominated by (external) information complexity



Lower Bound Tools

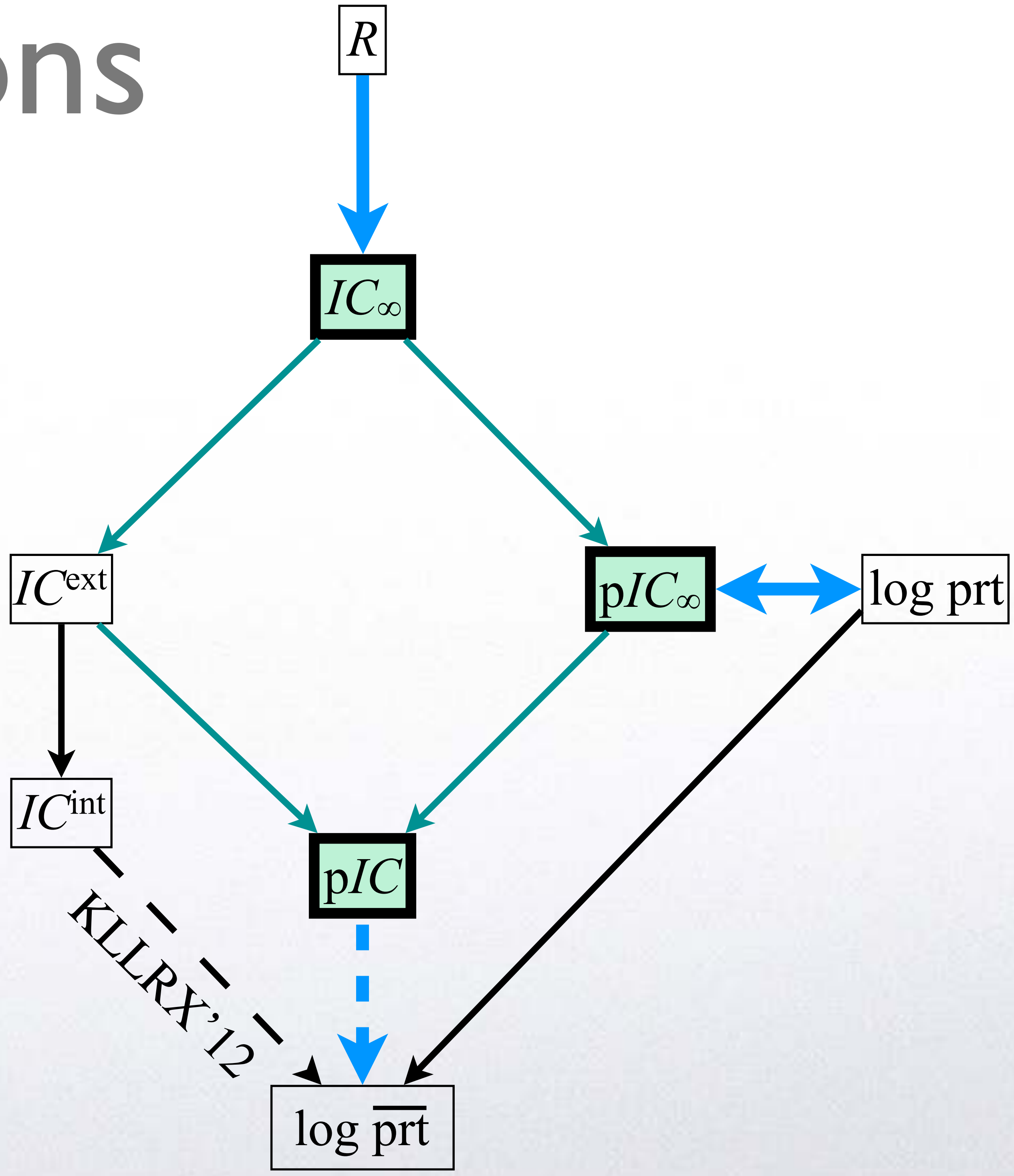
- Combinatorial bounds
  - Dominated by partition bounds [Jain-Klauck'10]
- Information theoretic bounds
  - Dominated by (external) information complexity
- Best connection between the two by Kerenidis et al. [KLLRX'12]



New Connections

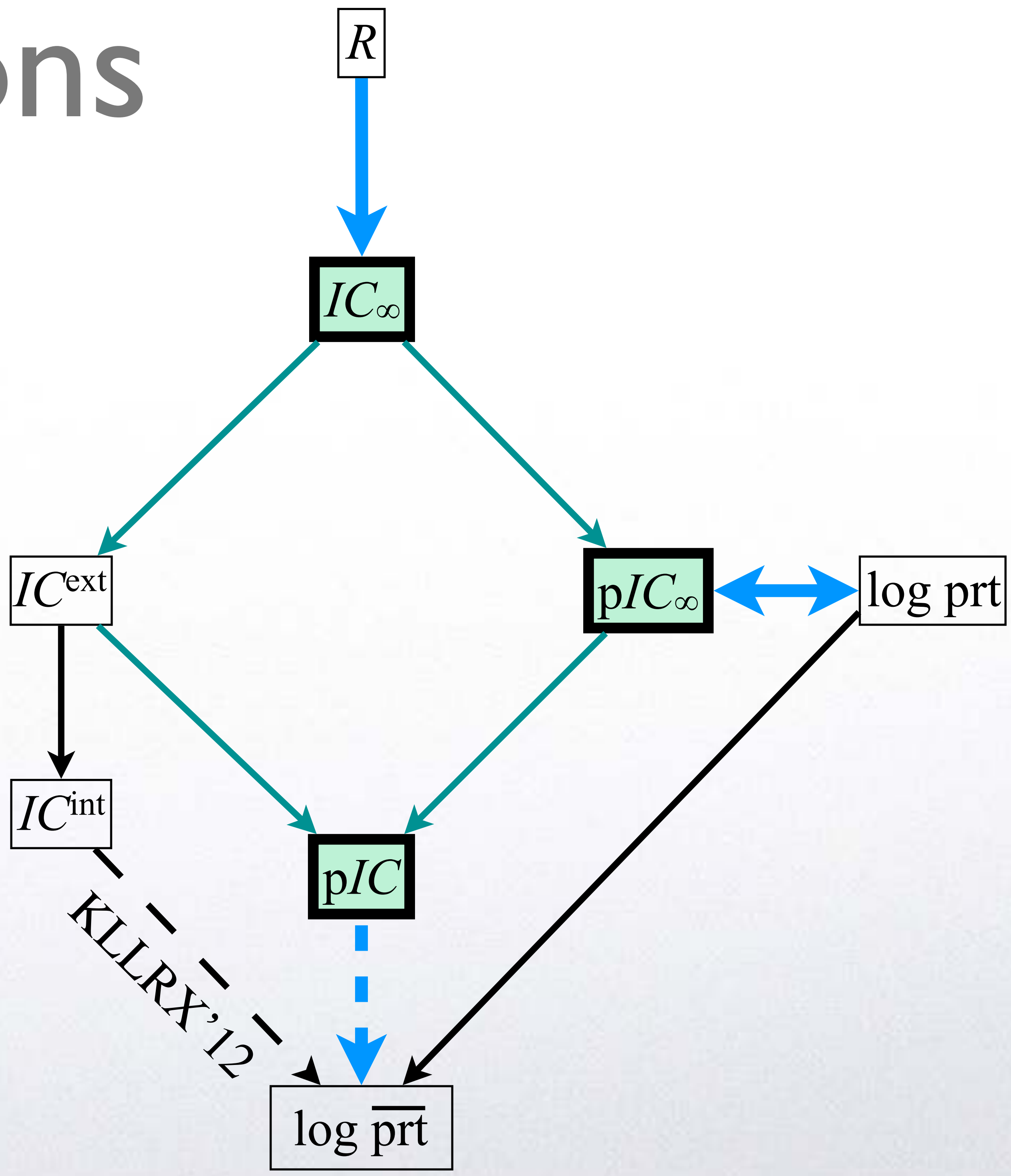
- A new lower bound IC_∞

- *Natural* relaxations yield IC^{ext} and $\log \text{prt}$



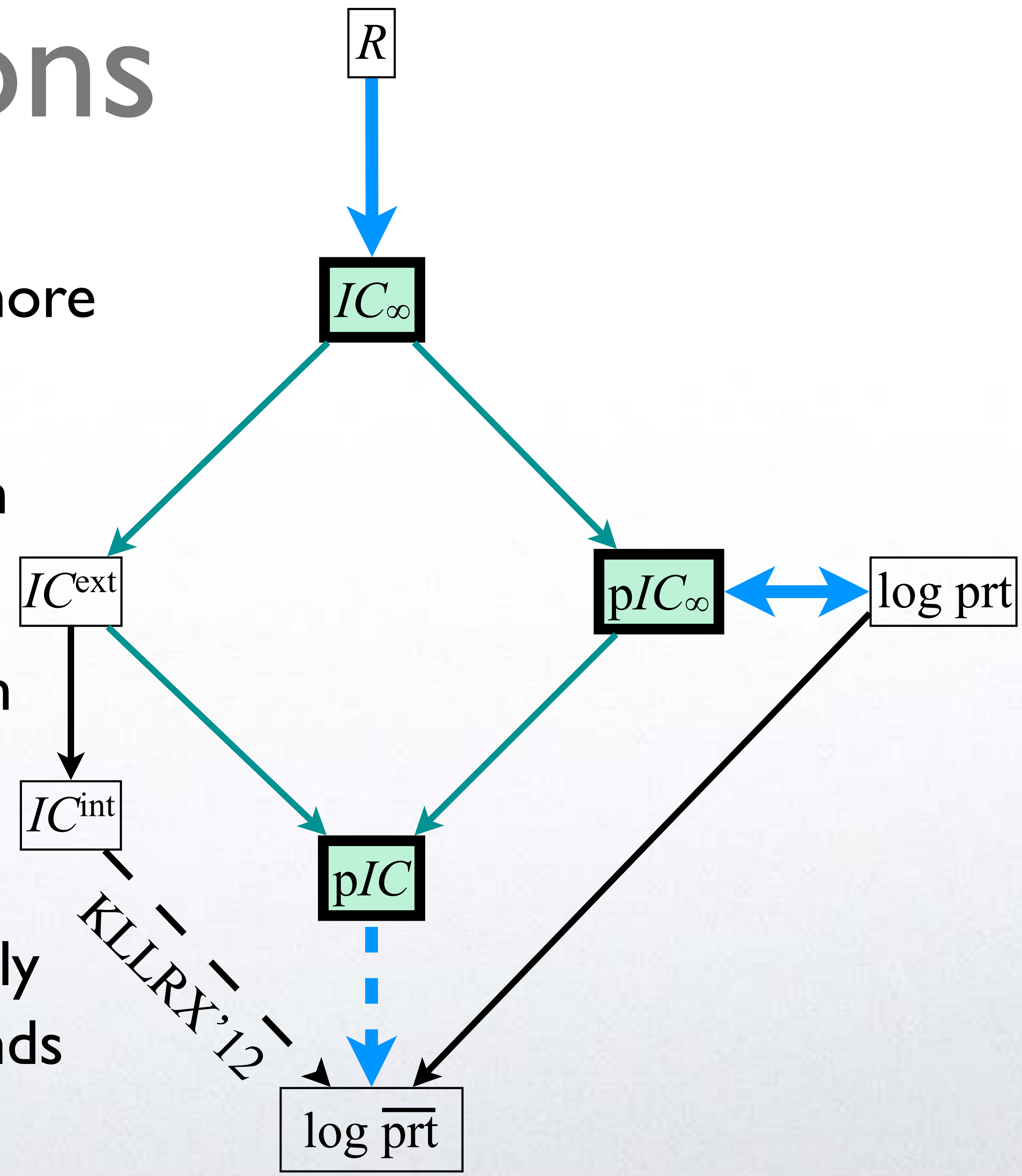
New Connections

- A new lower bound IC_∞
- *Natural* relaxations yield IC^{ext} and $\log \text{prt}$
- Applying both relaxations together yields $\mathrm{p}IC$, which dominates $\log \overline{\text{prt}}$



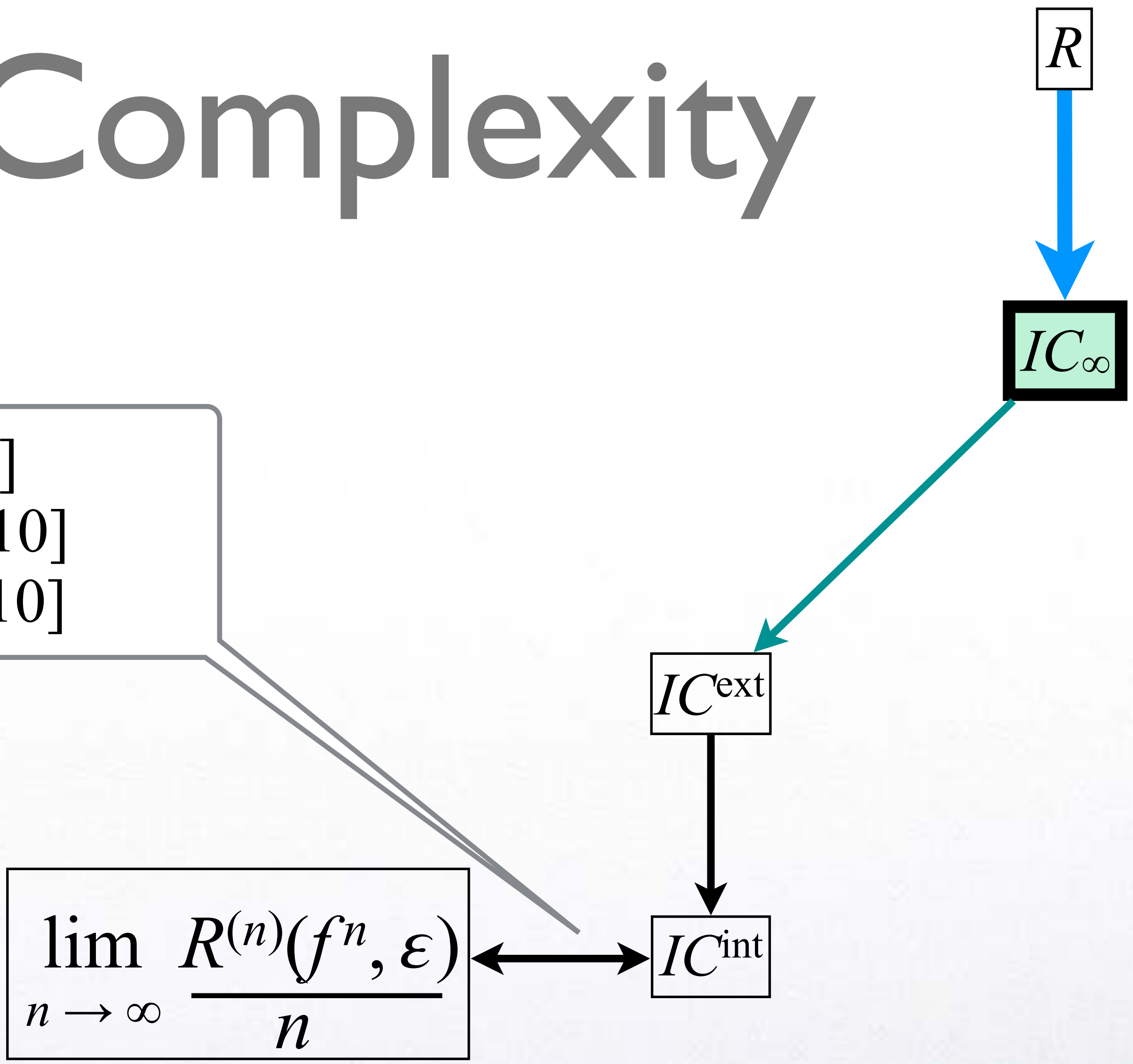
New Connections

- IC_∞ a *potentially stronger lower bound*, and potentially can separate R and IC^{ext} for more parameter ranges than currently known
- $\mathrm{p}IC_\infty$ gives a *new definition* of the partition bound
- $\mathrm{p}IC$ vs $\log \overline{\mathrm{prt}}$ implies a similar result as in [KLLRX'12], for IC^{ext}, but with better parameters
- Some lower bounds derived for IC^{ext} apply to $\mathrm{p}IC$ as well (e.g., [BJKS'04]). Such bounds *cannot beat the partition bound*.



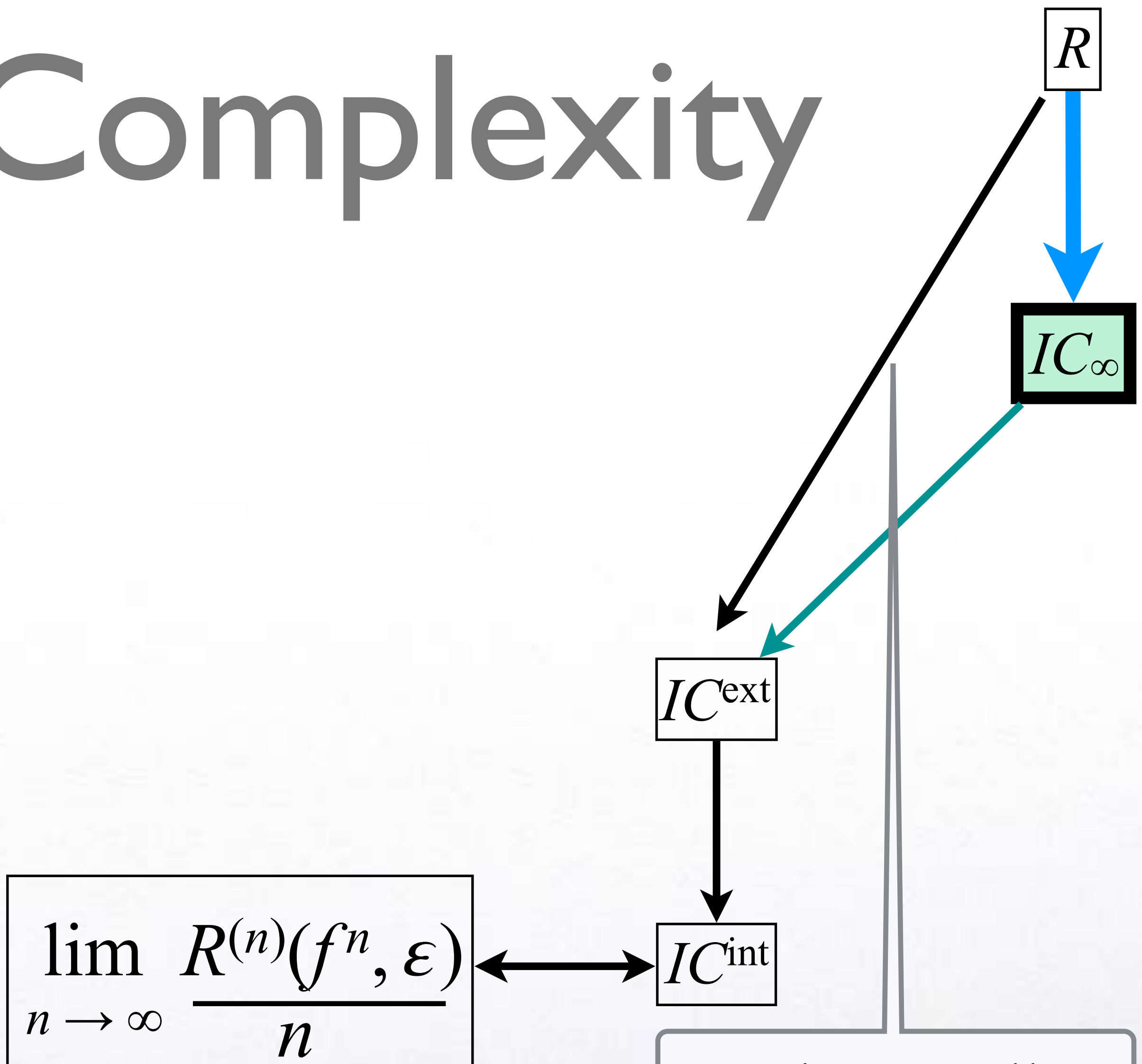
Information Complexity

[Braverman '12]
[Braverman,Rao'10]
[Ma,Ishwar '08-'10]



$$IC^{\text{int}}(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I(X;\Pi \mid Y) + I(Y;\Pi \mid X)}$$

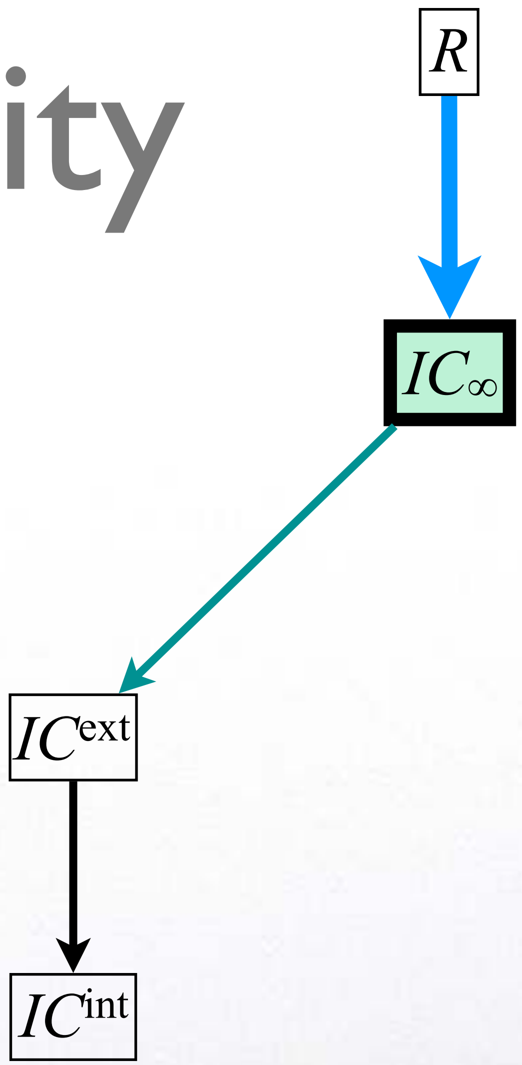
Information Complexity



When small, can have an exponential gap! [Ganor,Kol,Raz '14,'15]

$$IC^{\text{int}}(f, \varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \max_{\text{input distr } \mu} I(X; \Pi \mid Y) + I(Y; \Pi \mid X)$$

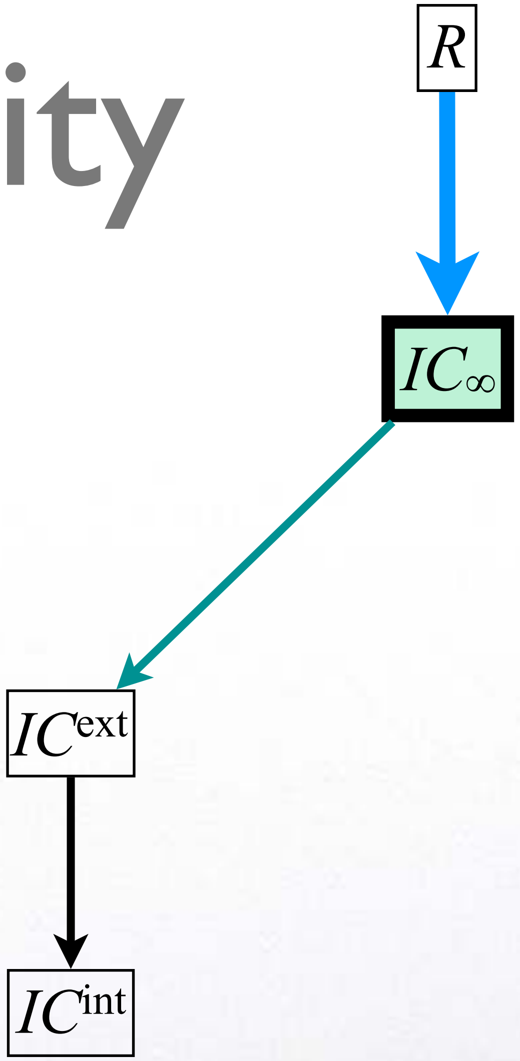
Information Complexity



$$IC^{\text{ext}}(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I(X,Y;\Pi)}$$

$$IC^{\text{int}}(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I(X;\Pi \mid Y) + I(Y;\Pi \mid X)}$$

Rényi Information Complexity



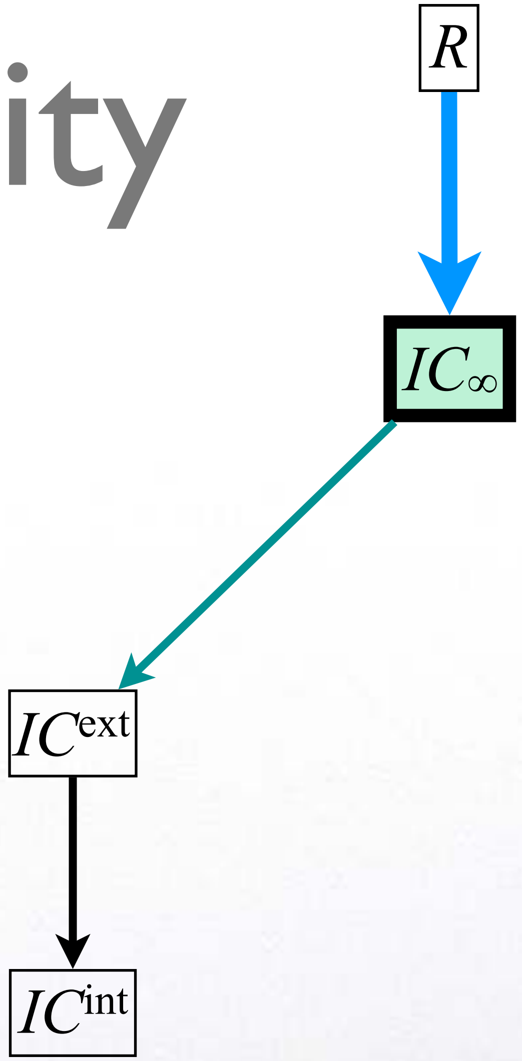
$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \mathrm{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I_\alpha(X,Y;\Pi)}$$

$$IC^{\mathrm{ext}}(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \mathrm{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I(X,Y;\Pi)}$$

$$IC^{\mathrm{int}}(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \mathrm{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I(X;\Pi \mid Y) + I(Y;\Pi \mid X)}$$

Rényi Information Complexity

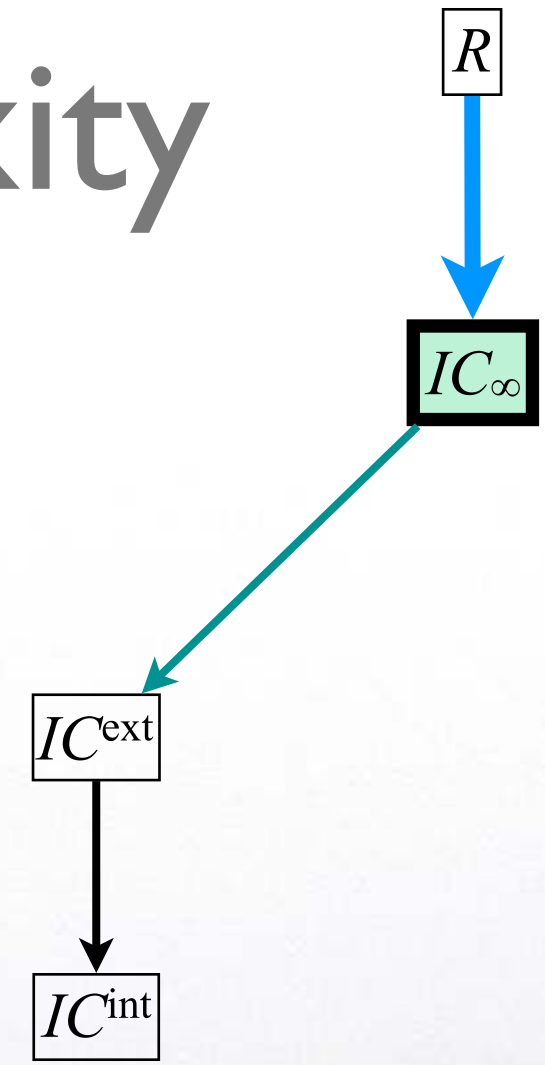
$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi,f)\le\varepsilon}} \boxed{\max_{\text{input distr } \mu} I_\alpha(X,Y;\Pi)}$$



$$I_\alpha(U;V) = \frac{\alpha}{\alpha-1}\log\sum_{v\in\mathcal{V}}\left(\sum_{u\in\mathcal{U}} p_U(u)p^\alpha_{V|U}(v|u)\right)^{\frac{1}{\alpha}}$$

[Sibson'69,Verdú'15]

Rényi Information Complexity



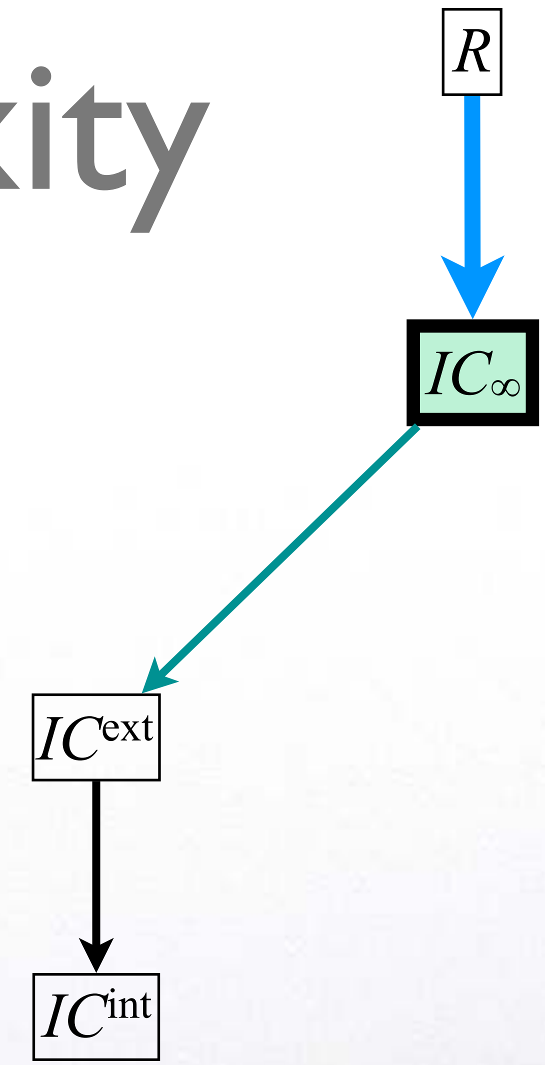
$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \leq \varepsilon}} \boxed{\max_{\text{input distr } \mu} I_\alpha(X,Y;\Pi)}$$

$$I_\alpha(U;V) = \frac{\alpha}{\alpha-1} \log \sum_{v\in\mathcal{V}} \left(\sum_{u\in\mathcal{U}} p_U(u) p^\alpha_{V|U}(v|u) \right)^{\frac{1}{\alpha}}$$

$$I_\infty(U;V) = \log \sum_{v\in\mathcal{V}} \max_{u:p_U(u)>0} p_{V|U}(v|u)$$

[Sibson'69,Verdú'15]

Rényi Information Complexity



$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi,f)\le\varepsilon}} \boxed{\max_{\text{input distr } \mu} I_\alpha(X,Y;\Pi)}$$

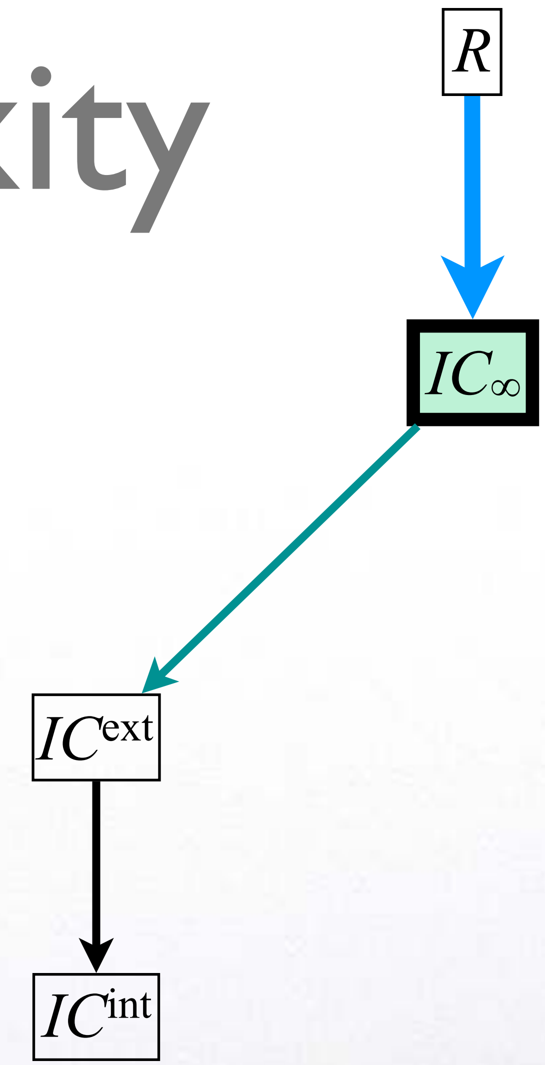
$$IC_\infty(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi,f)\le\varepsilon}} \boxed{\log \sum_q \max_{x,y} p_{\Pi|X,Y}(q\,|\,x,y)}$$

$$I_\alpha(U;V) = \frac{\alpha}{\alpha-1}\log\sum_{v\in\mathcal{V}}\left(\sum_{u\in\mathcal{U}} p_U(u)\,p^\alpha_{V|U}(v|u)\right)^{\frac{1}{\alpha}}$$

$$I_\infty(U;V) = \log\sum_{v\in\mathcal{V}}\max_{u:p_U(u)>0} p_{V|U}(v|u)$$

[Sibson'69, Verdú'15]

Rényi Information Complexity



$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I_\alpha(X,Y;\Pi)}$$

$$IC_\infty(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \boxed{\log \sum_q \max_{x,y} p_{\Pi|X,Y}(q \mid x,y)}$$

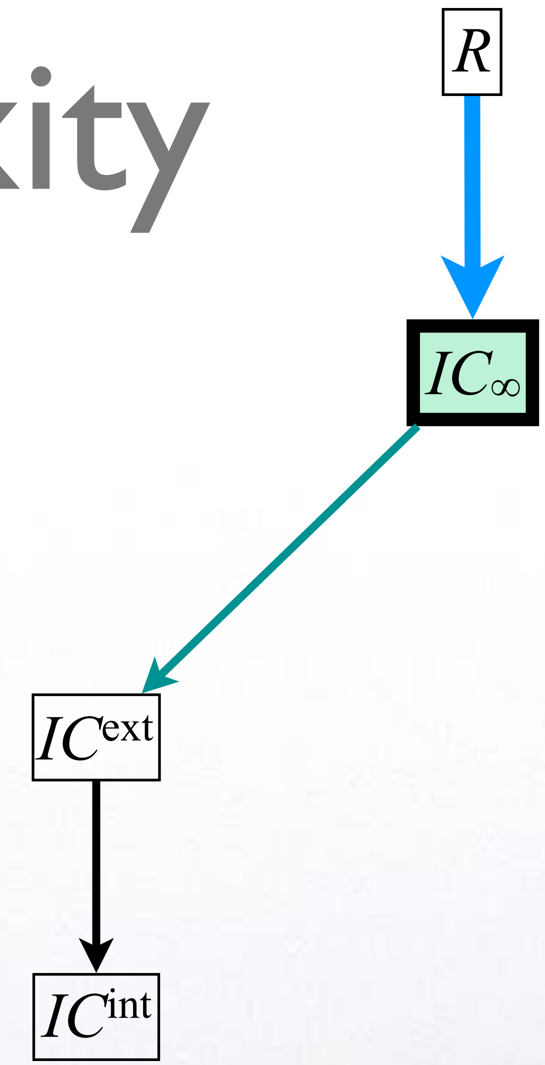
Previously used in one-shot communication problems [Ziv-Zakai'73]

$$I_\alpha(U;V) = \frac{\alpha}{\alpha - 1} \log \sum_{v \in \mathcal{V}} \left(\sum_{u \in \mathcal{U}} p_U(u) p^\alpha_{V|U}(v|u) \right)^{\frac{1}{\alpha}}$$

$$I_\infty(U;V) = \log \sum_{v \in \mathcal{V}} \max_{u : p_U(u) > 0} p_{V|U}(v|u)$$

[Sibson'69,Verdú'15]

Rényi Information Complexity

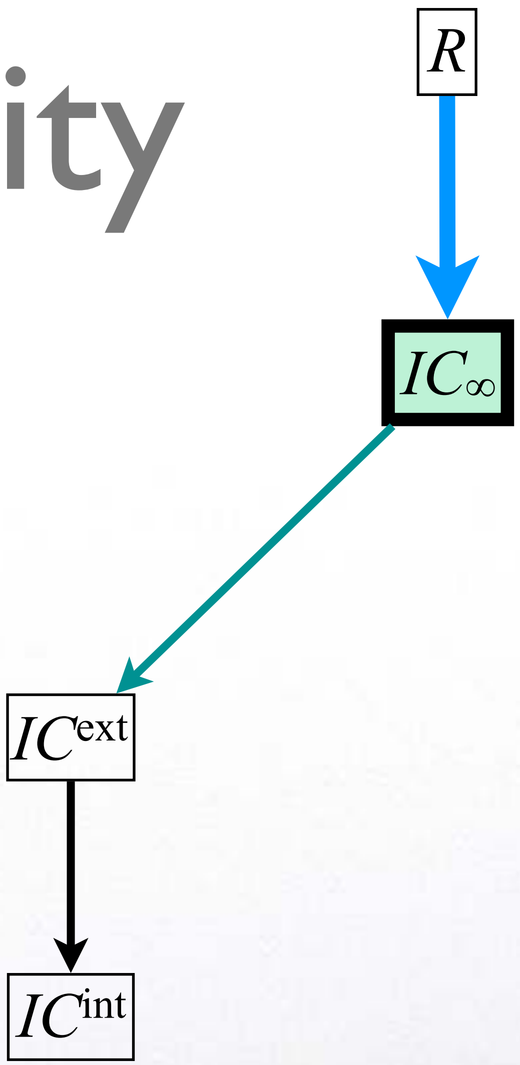


$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \max_{\text{input distr } \mu} I_\alpha(X, Y; \Pi)$$

$$IC_\infty(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \log \sum_q \max_{x,y} p_{\Pi|X,Y}(q \mid x,y)$$

No input distribution!
Depends only on
the conditional distribution
of pseudo-transcripts for each (x,y)

Rényi Information Complexity



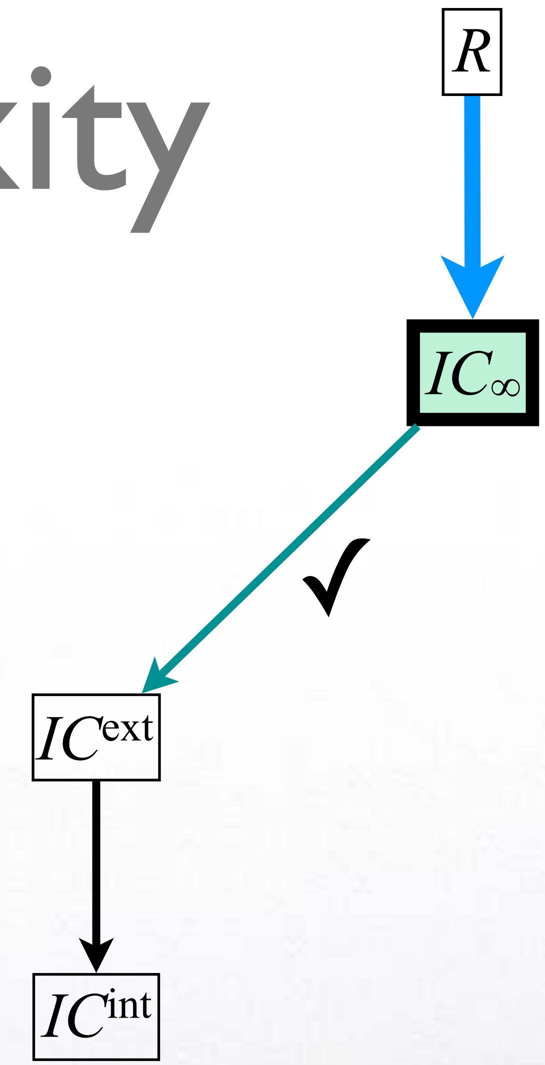
$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} \; I_\alpha(X,Y;\Pi)}$$

$$IC_\infty(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} \boxed{\log \sum_q \max_{x,y} p_{\Pi|X,Y}(q\,|\,x,y)}$$

$$IC^{\text{ext}}(f,\varepsilon) = \lim_{\alpha \to 1} IC_\alpha(f,\varepsilon)$$

$$\lim_{\alpha \to 1} I_\alpha(U;V) = I(U;V)$$

Rényi Information Complexity



$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \leq \varepsilon}} \boxed{\max_{\text{input distr } \mu} I_\alpha(X, Y; \Pi)}$$

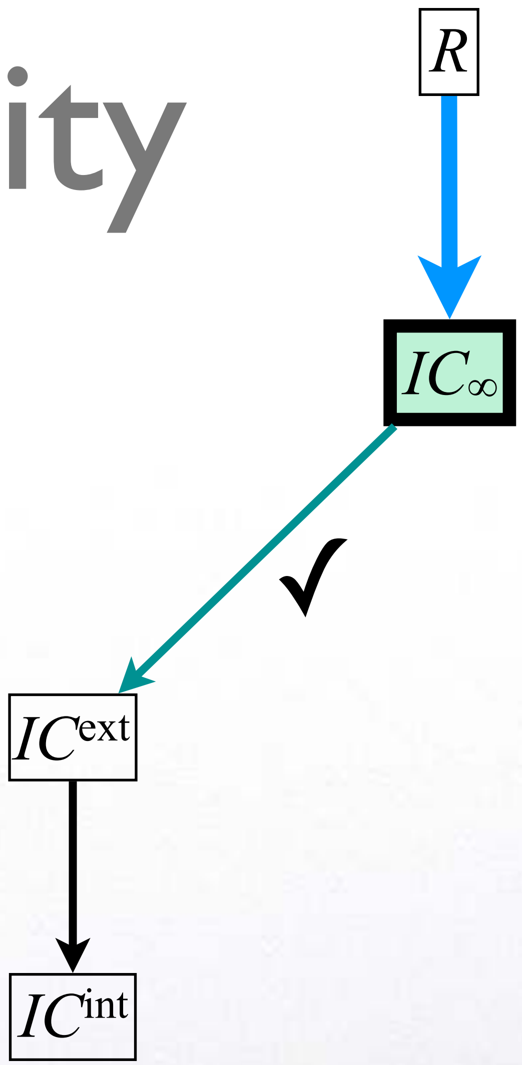
$$IC_\infty(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \leq \varepsilon}} \boxed{\log \sum_q \max_{x,y} p_{\Pi|X,Y}(q \mid x,y)}$$

$$IC^{\text{ext}}(f,\varepsilon) = \lim_{\alpha \to 1} IC_\alpha(f,\varepsilon)$$

$$\lim_{\alpha \to 1} I_\alpha(U; V) = I(U; V)$$

$I_\alpha(U; V)$ monotonically increases with α

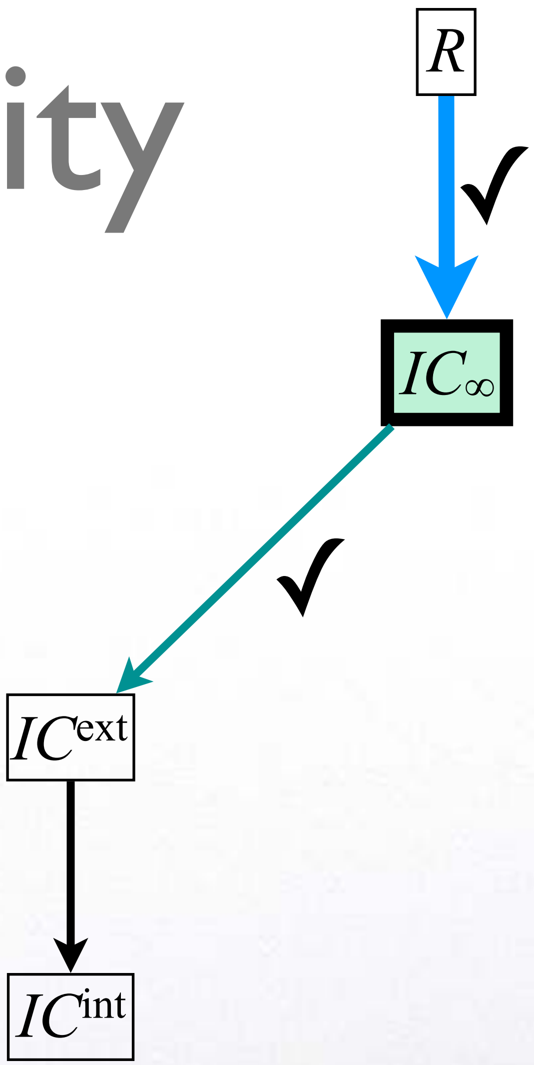
Rényi Information Complexity



$$IC_\infty(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} IC_\infty(\pi)$$

$$\begin{aligned} IC_\infty(\pi) &= \log \sum_q \max_{x,y} p_{\Pi|X,Y}(q \mid x,y) \\ &\le \log \sum_q 1 \\ &\le \#\text{bits}(\pi) \end{aligned}$$

Rényi Information Complexity



$$IC_\infty(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \text{err}(\pi, f) \le \varepsilon}} IC_\infty(\pi)$$

$$IC_\infty(\pi) = \log \sum_q \max_{x,y} p_{\Pi|X,Y}(q \mid x,y)$$

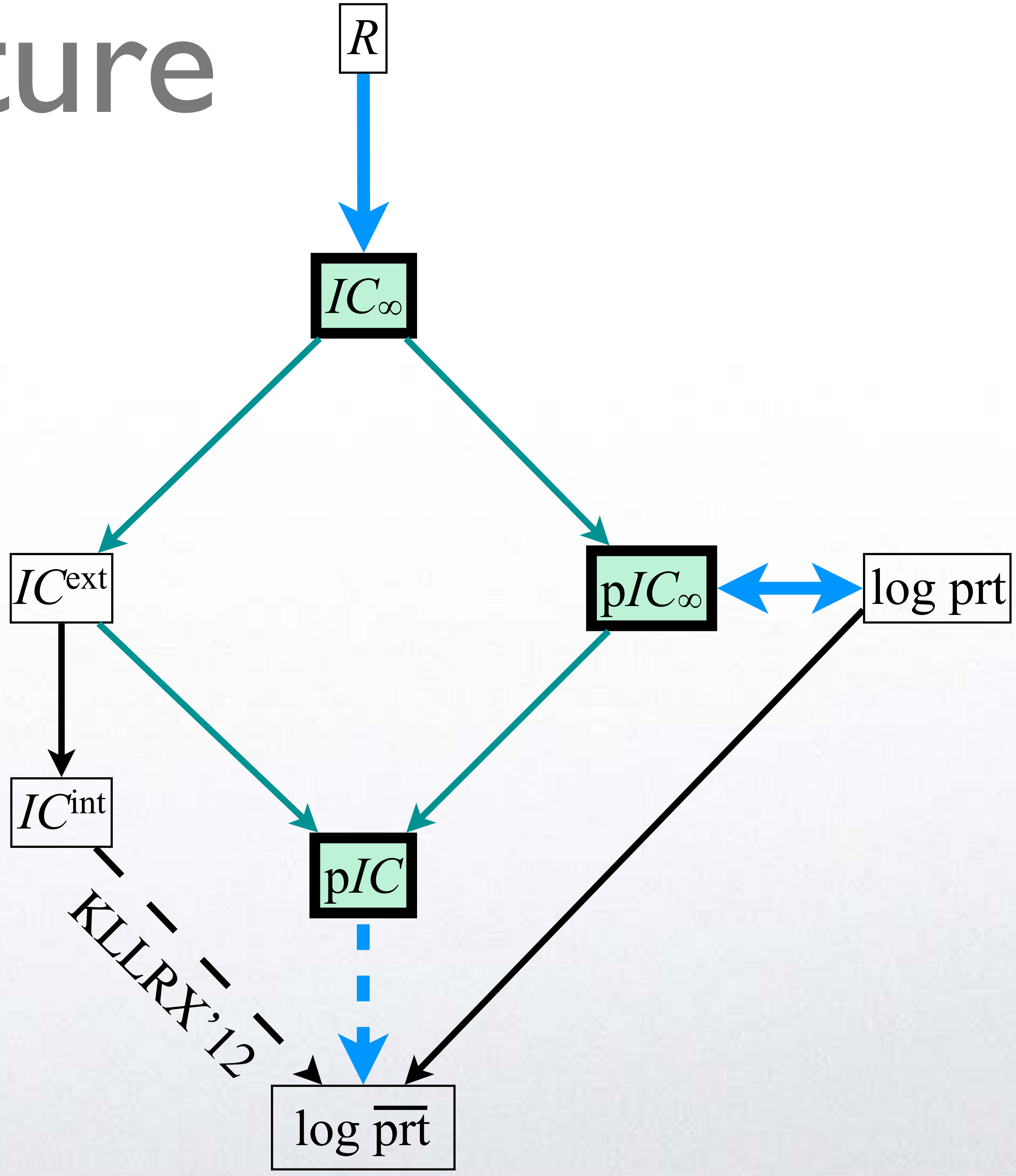
$$\le \log \sum_q 1$$

$$\le \#\text{bits}(\pi)$$

Generalizes to public-coin protocols too

Recall The Big Picture

- How to relax IC_∞ to pIC_∞?

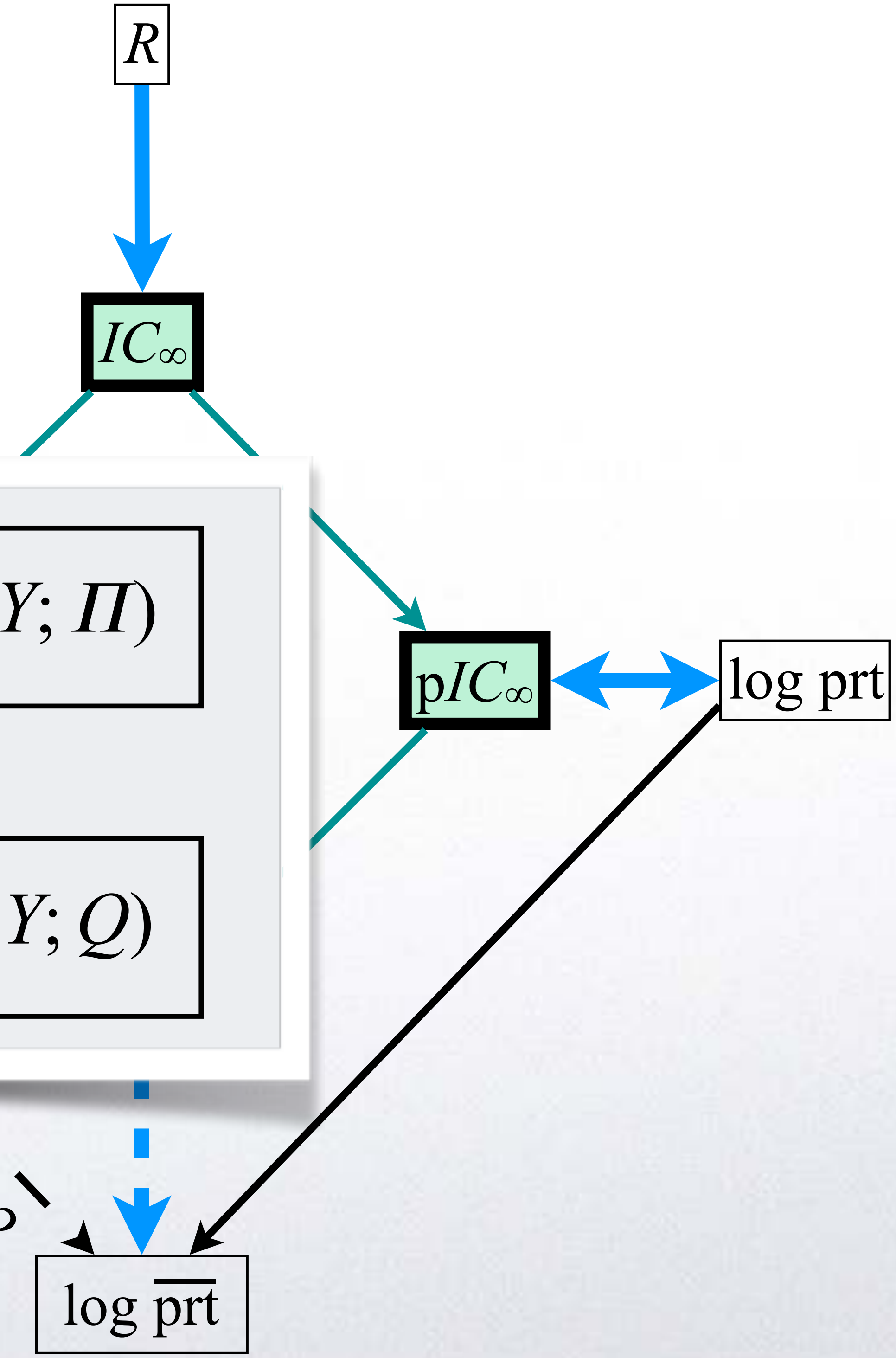


Recall The Big Picture

- How to relax IC_∞ to pIC_∞?

$$IC_\alpha(f,\varepsilon) = \inf_{\substack{\text{rand protocol } \pi: \\ \mathrm{err}(\pi, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I_\alpha(X, Y; \Pi)}$$

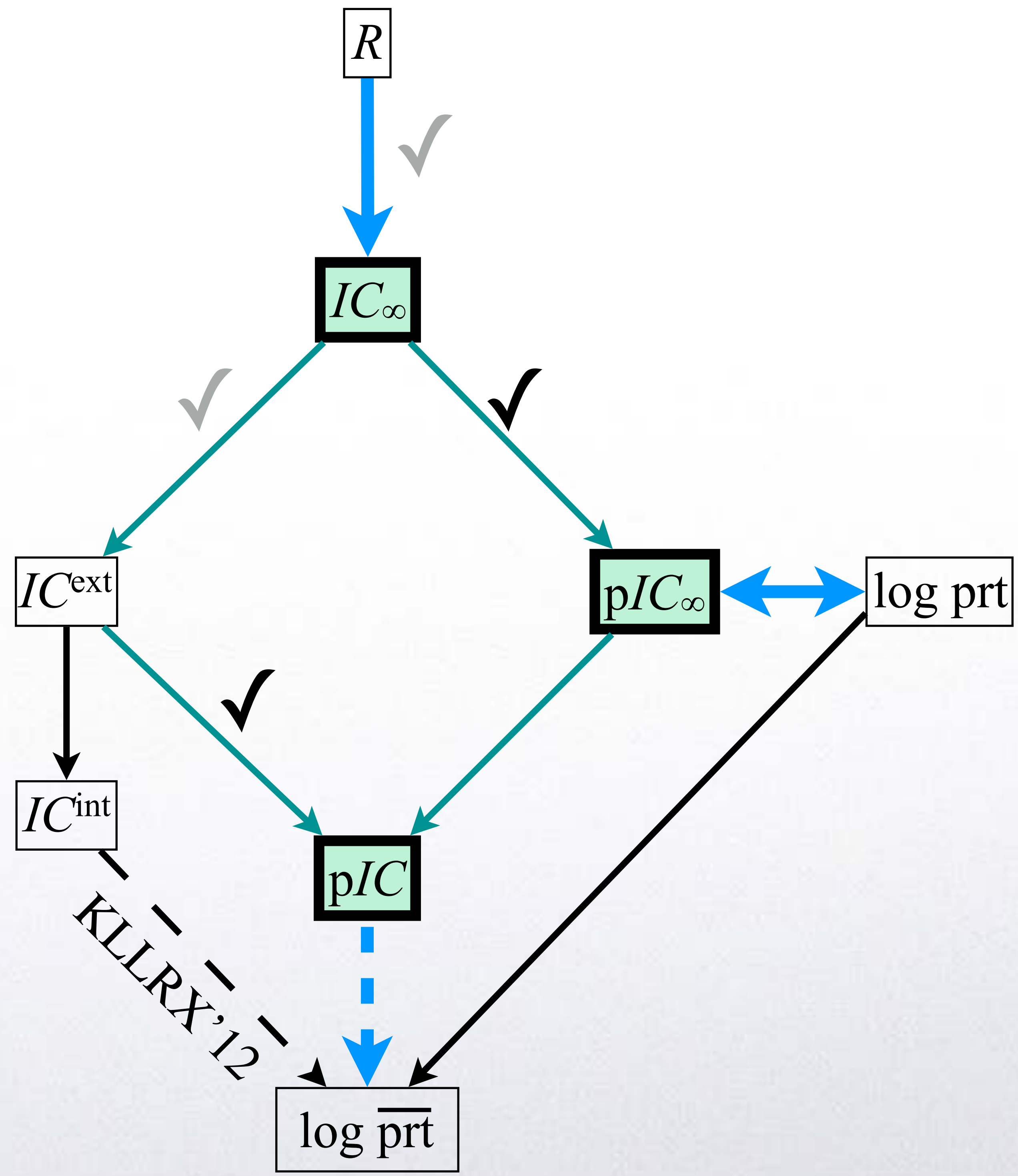
$$\mathrm{p}IC_\alpha(f,\varepsilon) = \inf_{\substack{\textbf{pseudo trans. } Q: \\ \mathrm{err}(Q, f) \le \varepsilon}} \boxed{\max_{\text{input distr } \mu} I_\alpha(X, Y; Q)}$$



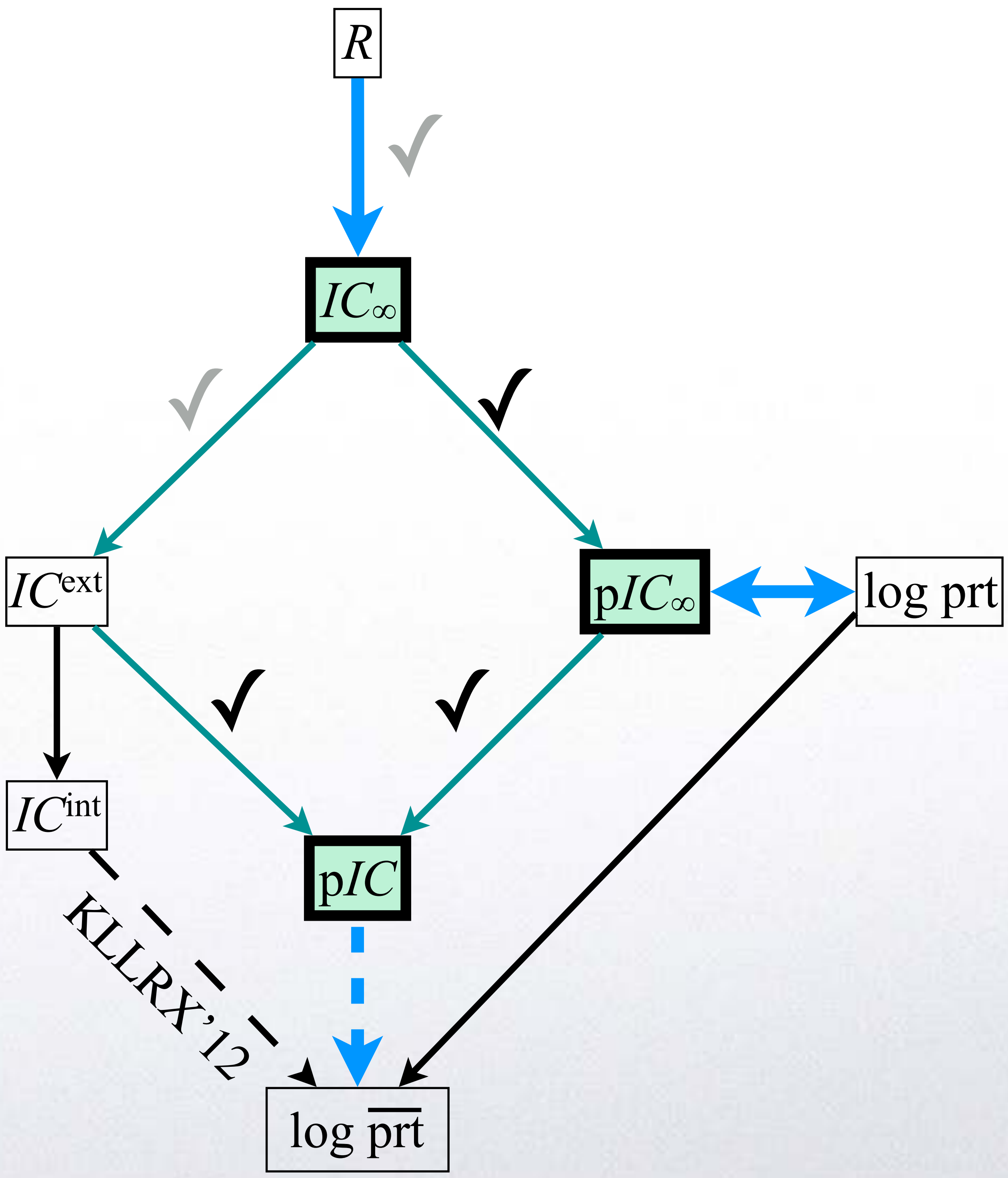
Pseudo Transcripts

- In a protocol, transcripts satisfy the "factorization property," $p(q \mid x, y) = \alpha(q, x) \times \beta(q, y)$
  - We require (~ w.l.o.g.) that outputs by both parties are the same and it is part of the transcript : z_q
- **Pseudo transcript**: any random variable Q such that $p_{Q|XY}$ has the factorization property (along with a function mapping $q \mapsto z_q$)
- Error of Q w.r.t. f : $\mathrm{err}(Q, f) = \max_{x,y} p[\, z_q \neq f(x,y) \mid x, y \,]$

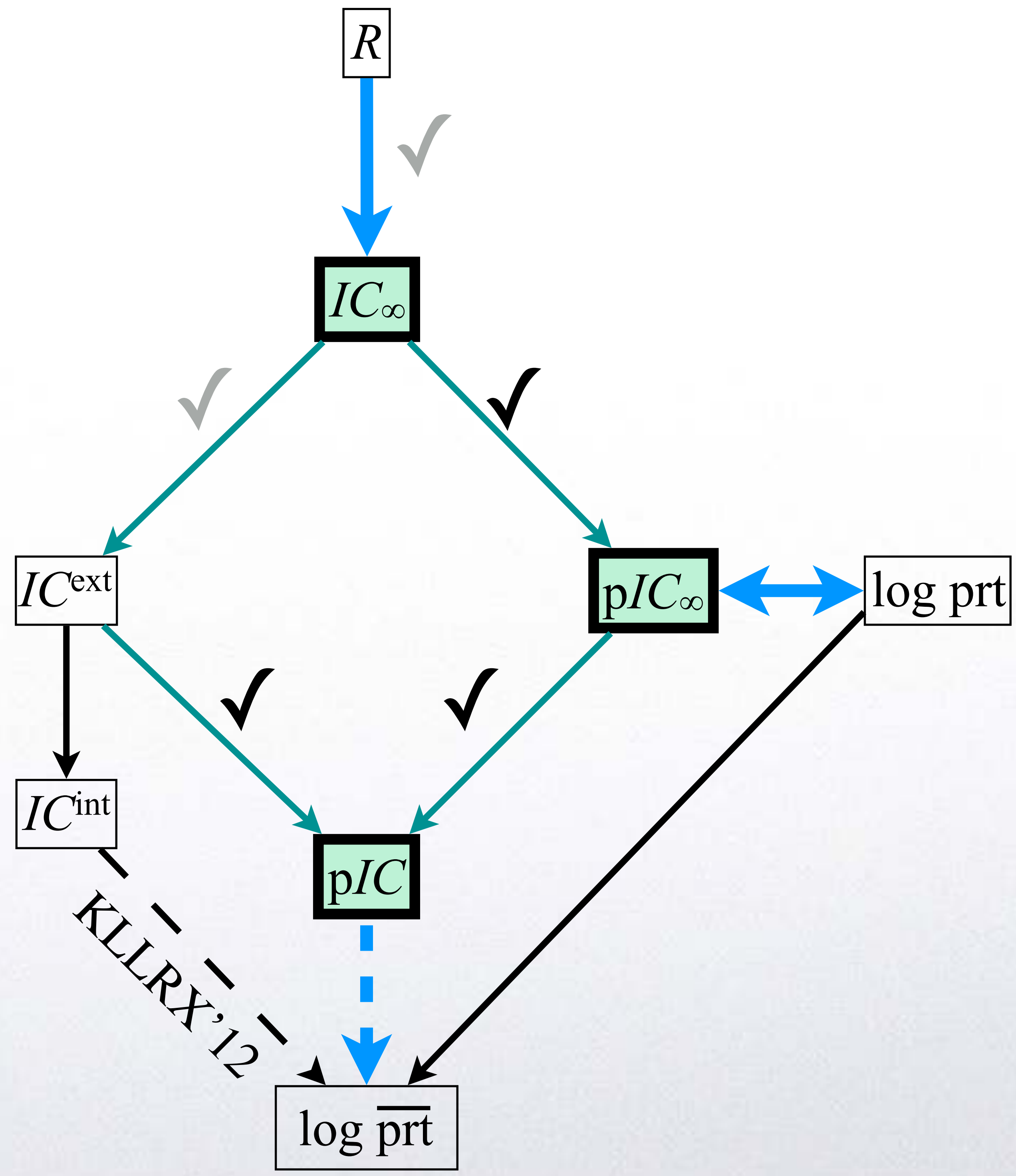
- A transcript is a pseudo-transcript (with the same error)



- A transcript is a pseudo-transcript (with the same error)
- And as before, monotonicity of I_α implies monotonicity of $\mathrm{p}IC_\alpha$



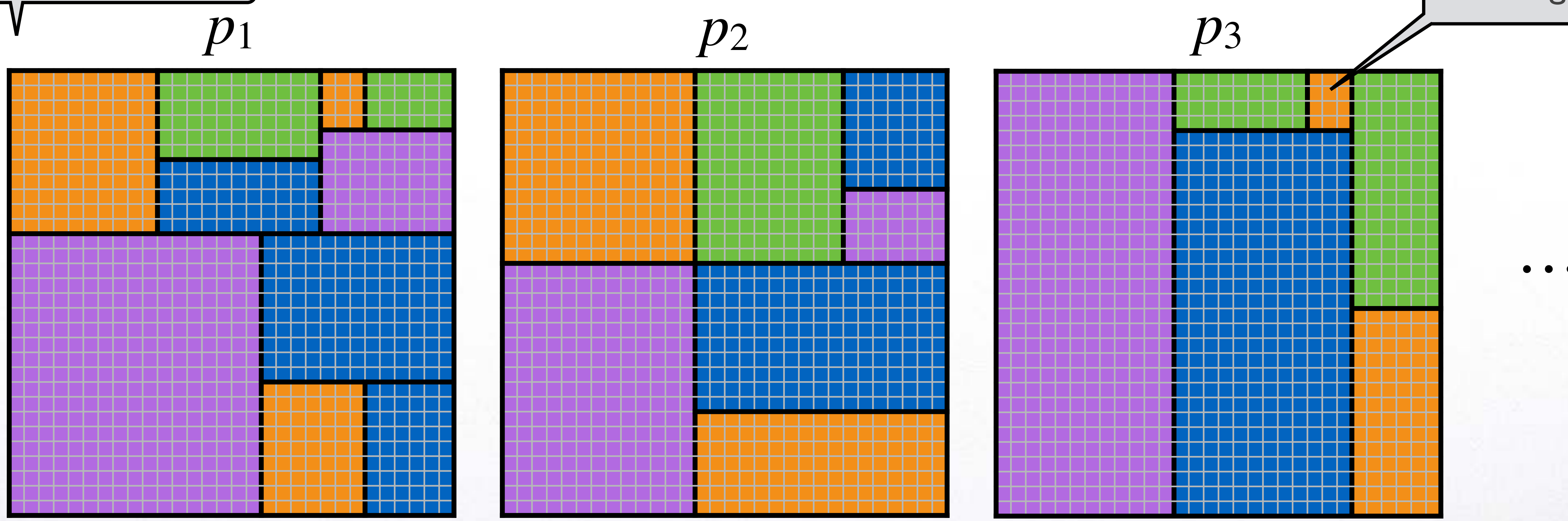
- A transcript is a pseudo-transcript (with the same error)
- And as before, monotonicity of I_α implies monotonicity of $\mathrm{p}IC_\alpha$
- Next: $\mathrm{p}IC_\infty(f,\varepsilon) = \log \mathrm{prt}\ (f,\varepsilon)$
- <u>**A consequence:**</u> If an IC-bound is derived only using the pseudo-transcript property of the protocol, *then it cannot beat the partition bound*



Partition Bound

Distribution over deterministic protocols

Tile: colored "rectangle"

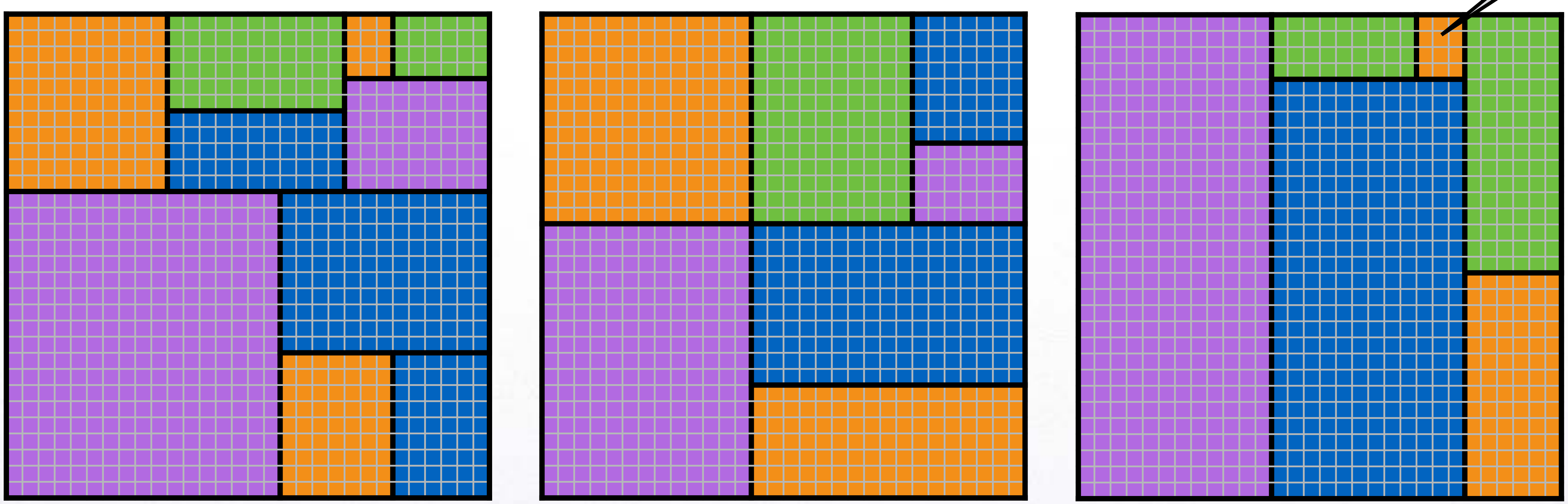


Partition Bound

Distribution over deterministic protocols

Tile: colored "rectangle"

p_1 p_2 p_3



. . .

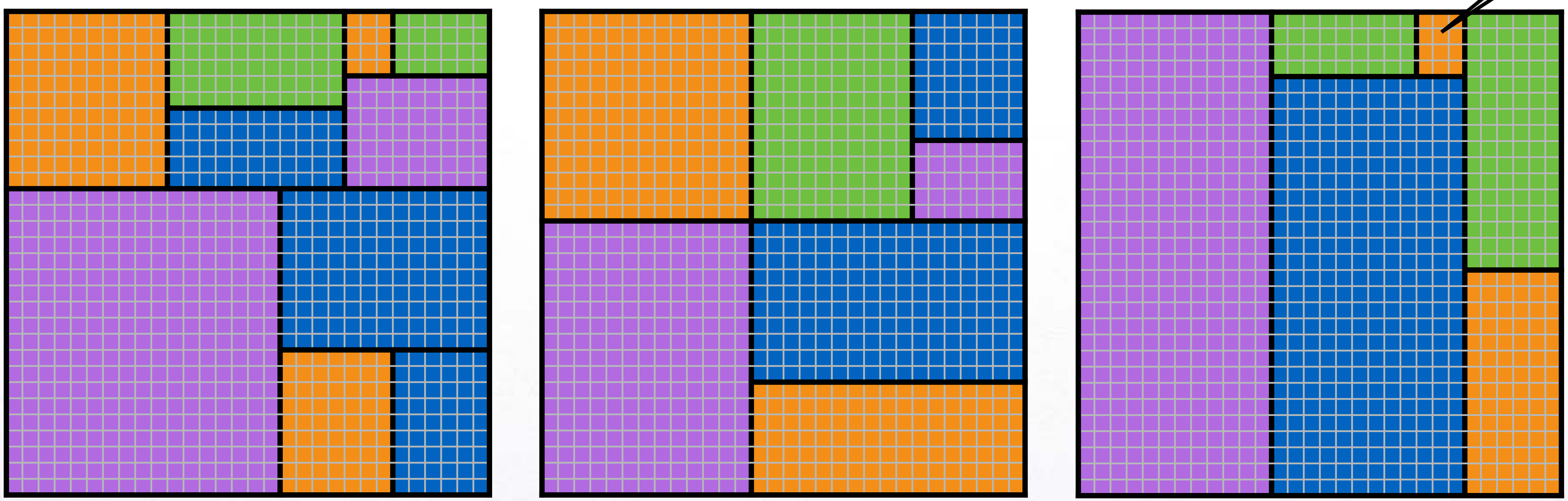
$$w(T) = \mathrm{p}(T \mid x,y)\ \forall (x, y) \in T$$

Partition Bound

Distribution over deterministic protocols

Tile: colored "rectangle"

p_1 p_2 p_3



$|\text{Tiling}_1| = 10$ $|\text{Tiling}_2| = 7$ $|\text{Tiling}_3| = 6$ …

$w(T) = \mathrm{p}(T \mid x,y)\ \forall (x, y) \in T$

$\Sigma_i\, p_i \cdot |\text{Tiling}_i| = \Sigma_{\text{tiles } T}\ w(T)$

Partition Bound

$$\text{prt}(f,\varepsilon) = \min \sum_{\text{tiles } T} w(T)$$

$$\forall\, x,y \qquad \sum_{T:(x,y)\in T} w(T) = 1$$

$$\forall\, x,y \in f^{-1} \qquad \sum_{T:(x,y)\in T,\ \text{color}(T)=f(x,y)} w(T) \geq 1-\varepsilon$$

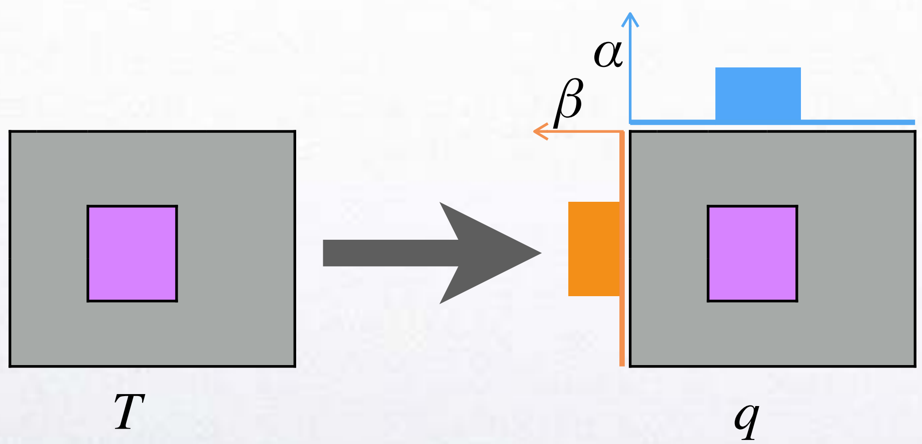
$$\forall\, T \qquad w(T) \geq 0$$

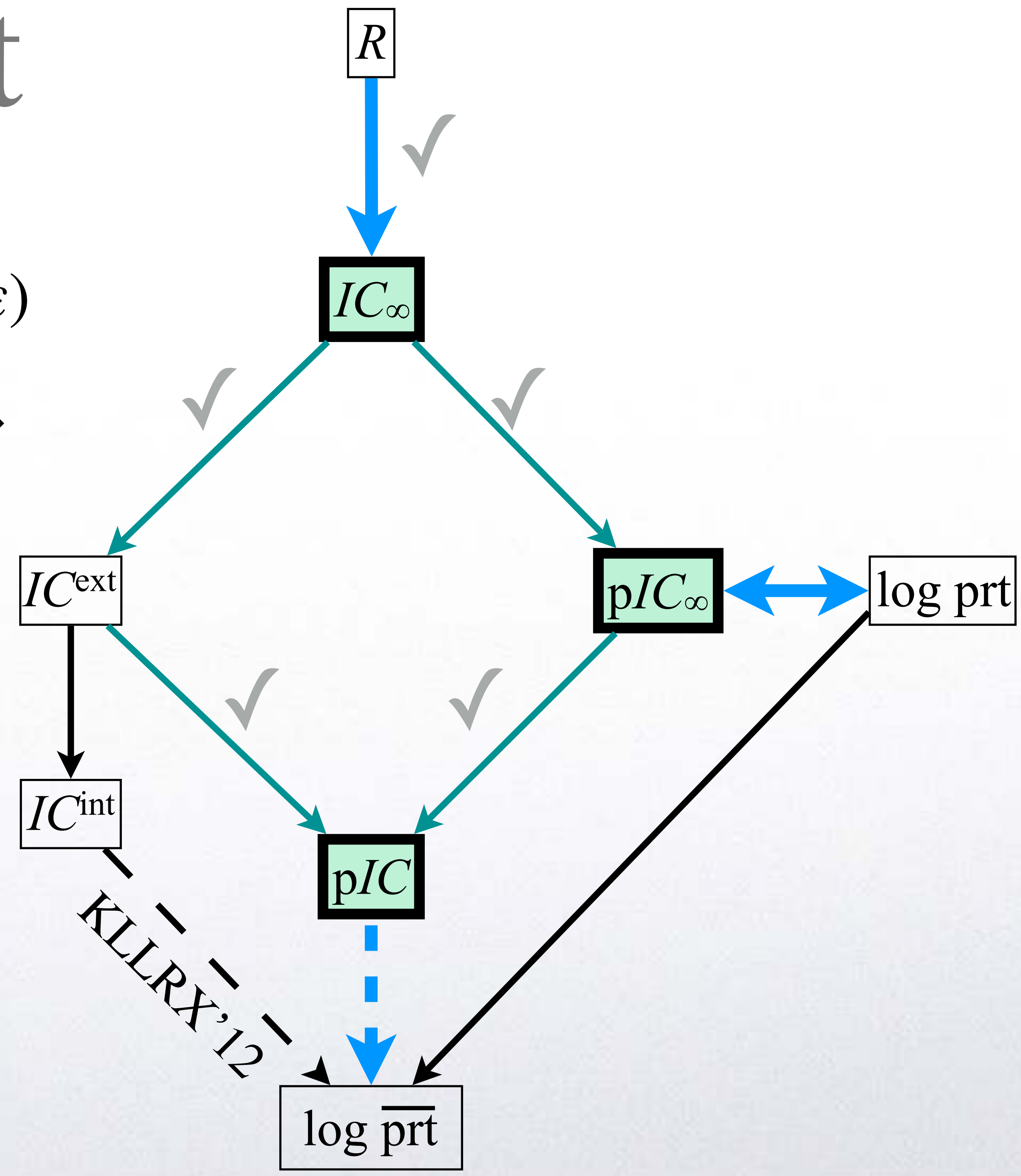
$w(T) = \text{p}(T \mid x,y)\ \forall (x,y) \in T$

$\sum_i p_i \cdot |\text{Tiling}_i| = \sum_{\text{tiles } T} w(T)$

pIC_∞ ≤ log prt

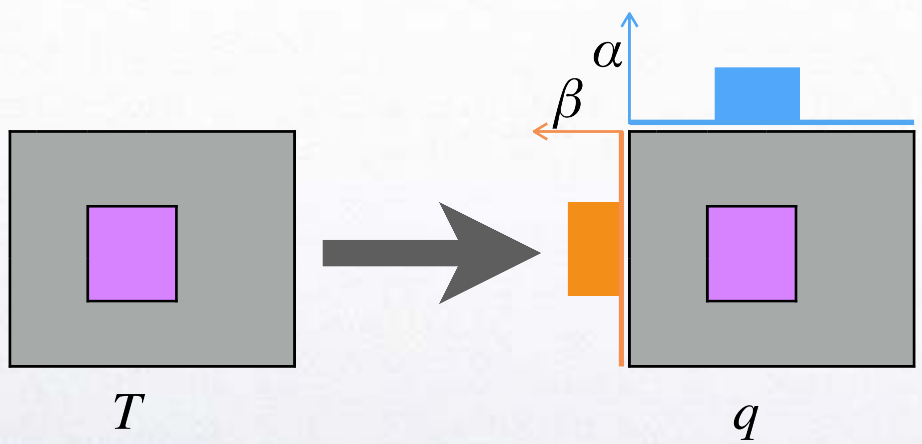
- Easy direction: p$IC_\infty(f,\varepsilon)$ ≤ log prt (f,ε)
- Partition (set of tiles and weights) → Pseudo-transcript distribution with p(q | x, y) = $\alpha(q, x) \times \beta(q,y)$



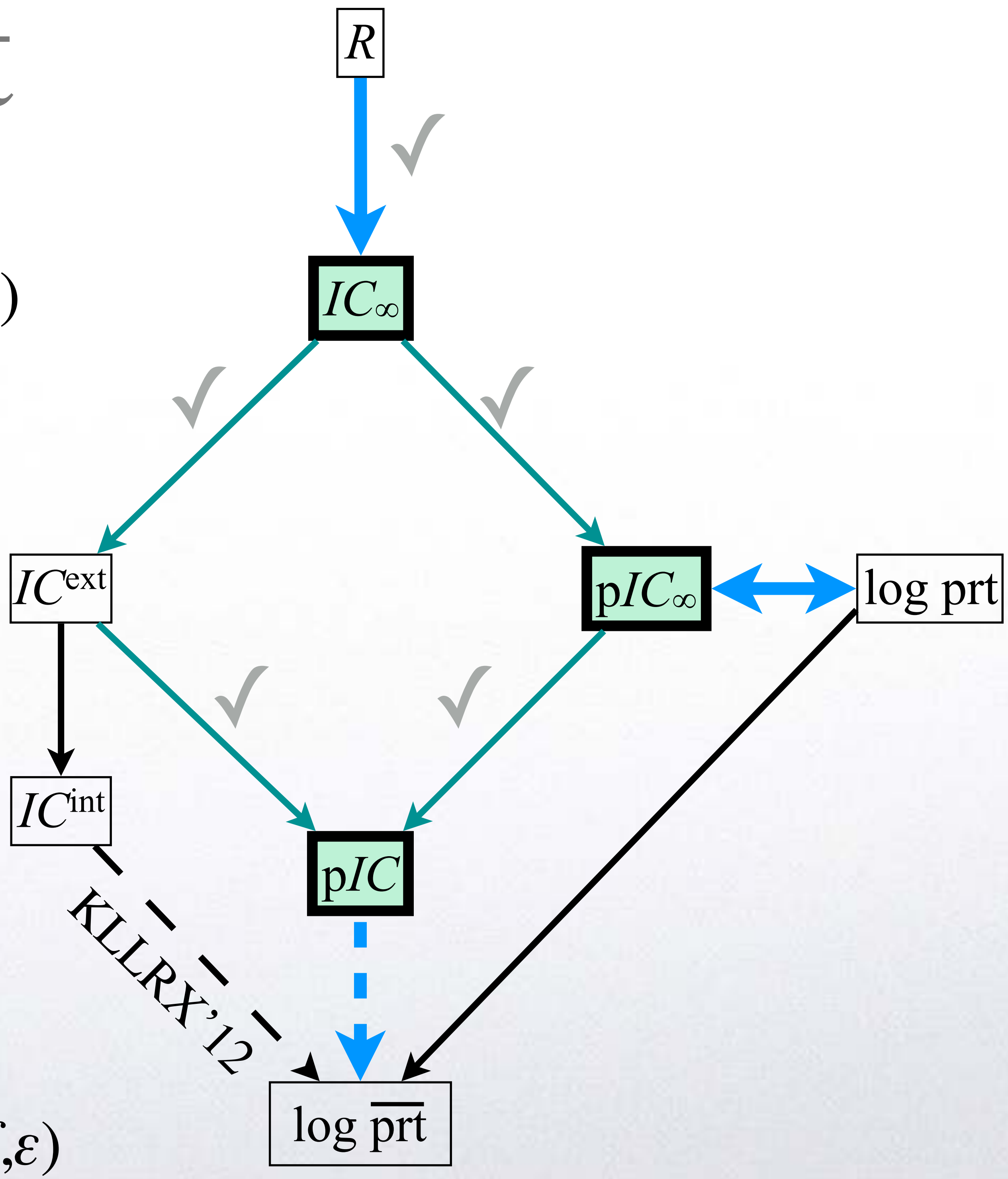


pIC_∞ ≤ log prt

- Easy direction: p$IC_\infty(f,\varepsilon)$ ≤ log prt (f,ε)
- Partition (set of tiles and weights) → Pseudo-transcript distribution with p$(q \mid x, y) = \alpha(q, x) \times \beta(q,y)$

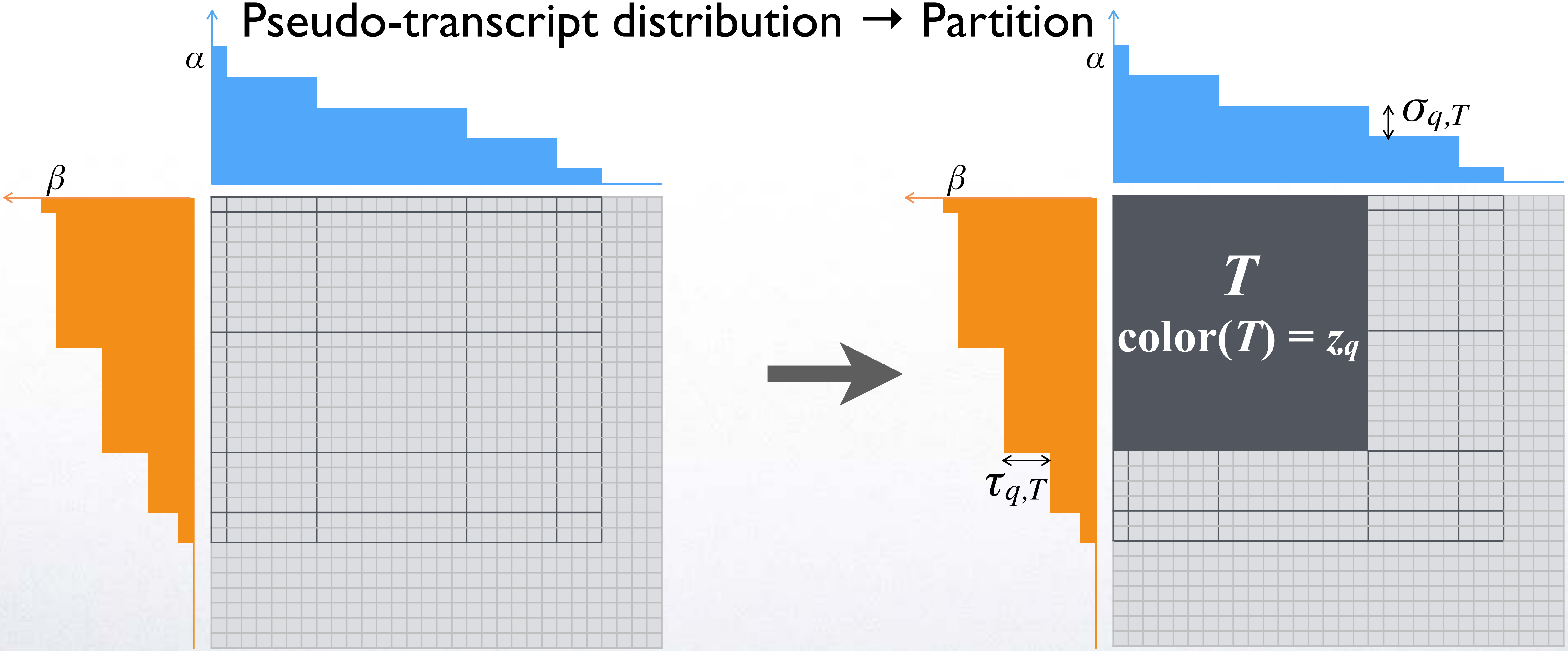


- $I_\infty(X,Y; Q)$ = log Σ_q max$_{x,y}$ p$(q|x,y)$
  = log Σ_T $w(T)$ = log prt (f,ε)



$\mathrm{p}IC_{\infty} \geq \log \mathrm{prt}$

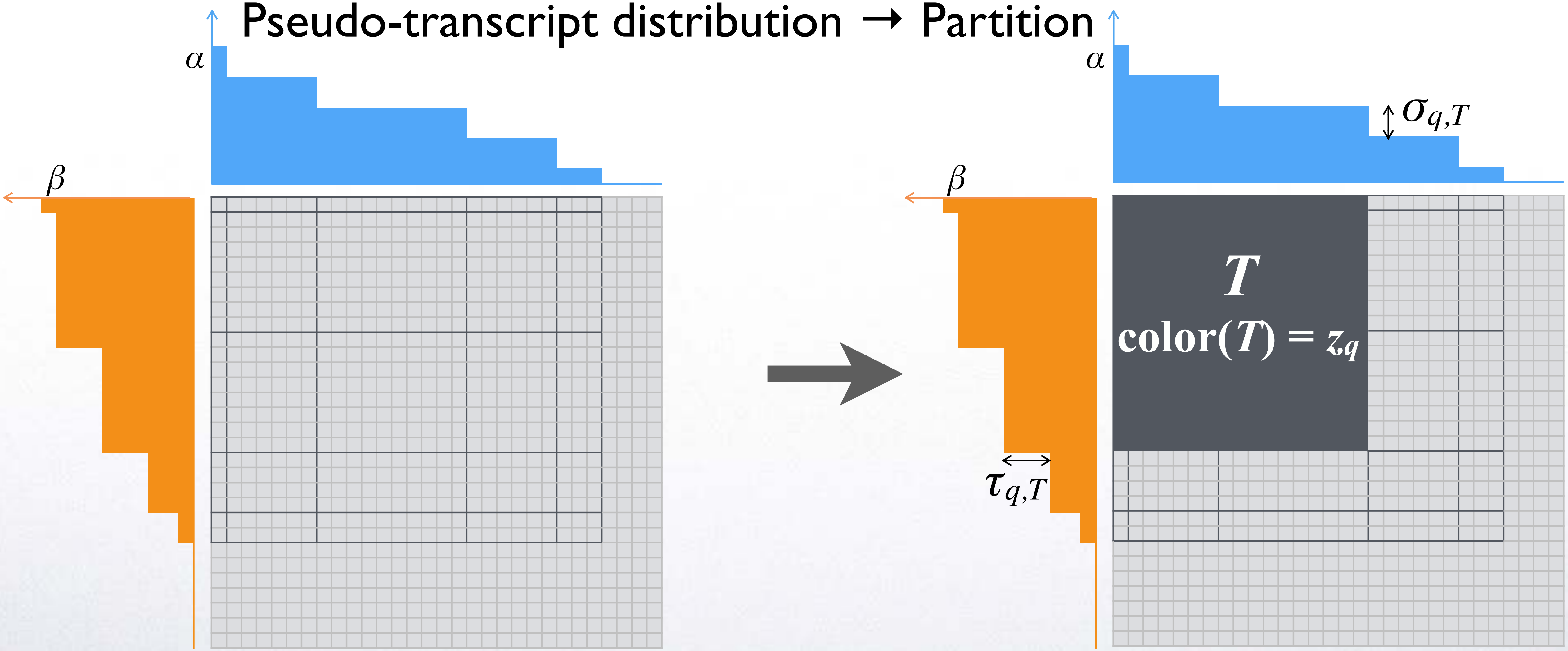
Pseudo-transcript distribution → Partition



$$w(T) = \Sigma_q \, \sigma_{q,T} \times \tau_{q,T}$$

p$IC_\infty \geq$ log prt

Pseudo-transcript distribution → Partition

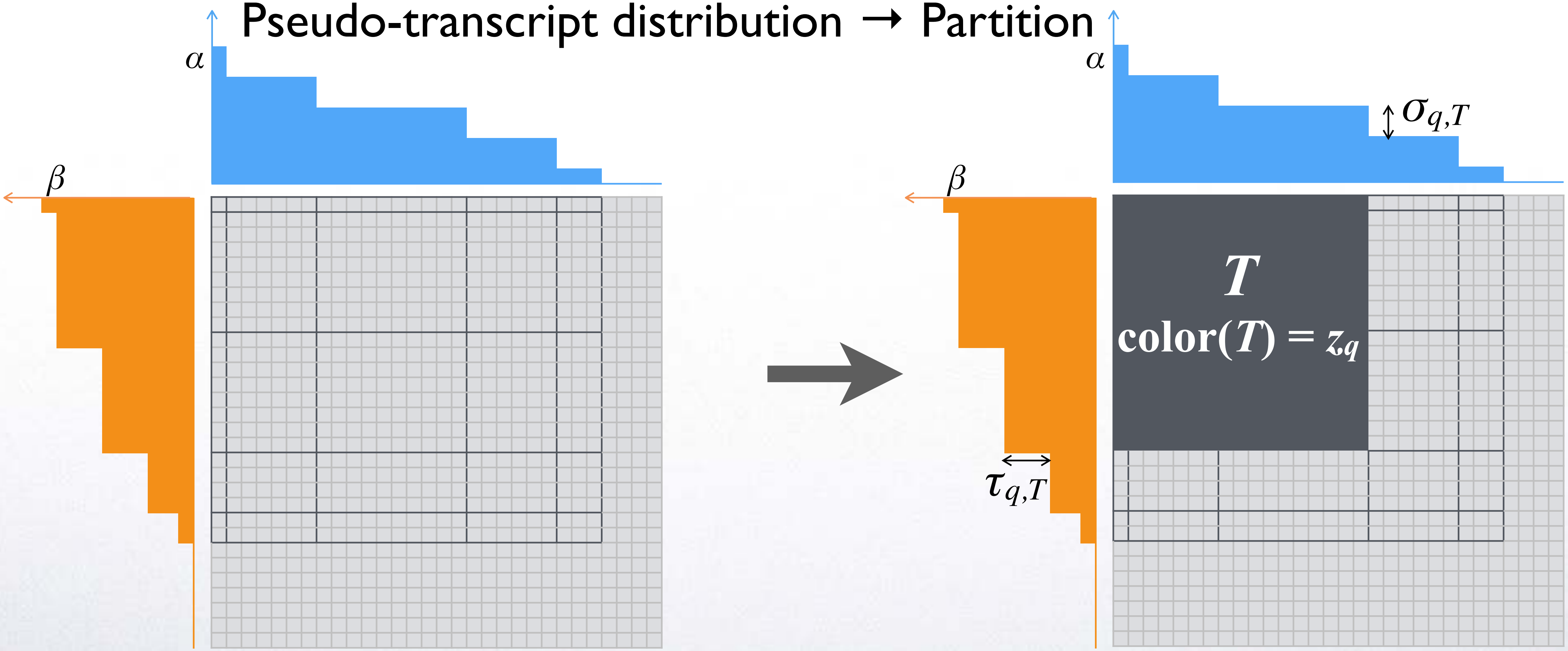


$w(T) = \Sigma_q \, \sigma_{q,T} \times \tau_{q,T}$

$\Sigma_T \, w(T) =$ … $= \exp(\mathrm{p}IC_\infty(Q))$

$\mathrm{p}IC_\infty \geq \log \mathrm{prt}$

Pseudo-transcript distribution → Partition

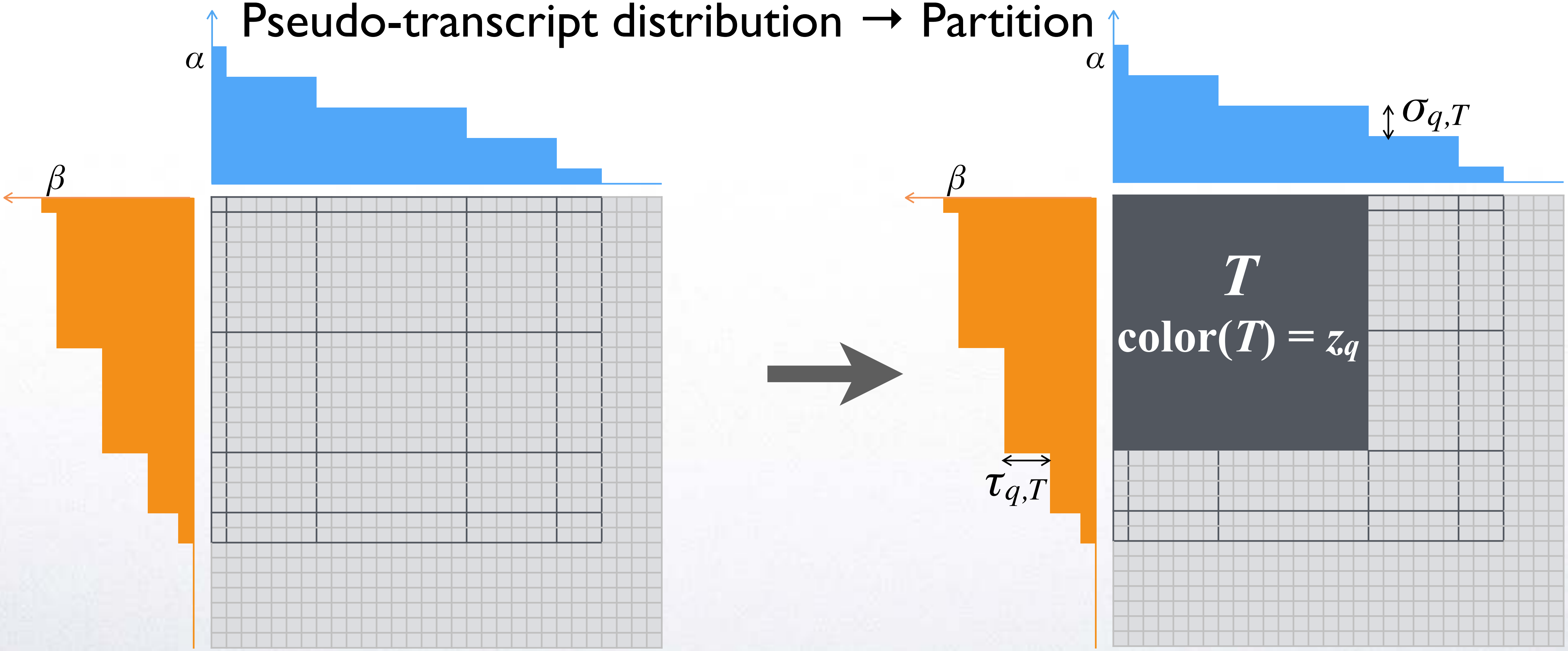


$$w(T) = \Sigma_q\ \sigma_{q,T} \times \tau_{q,T}$$

$$\Sigma_T\ w(T) = \Sigma_q\ (\Sigma_T\ \sigma_{q,T}) \cdot (\Sigma_T\ \tau_{q,T}) = \quad \ldots \quad = \exp(\mathrm{p}IC_\infty(Q))$$

p$IC_\infty \geq$ log prt

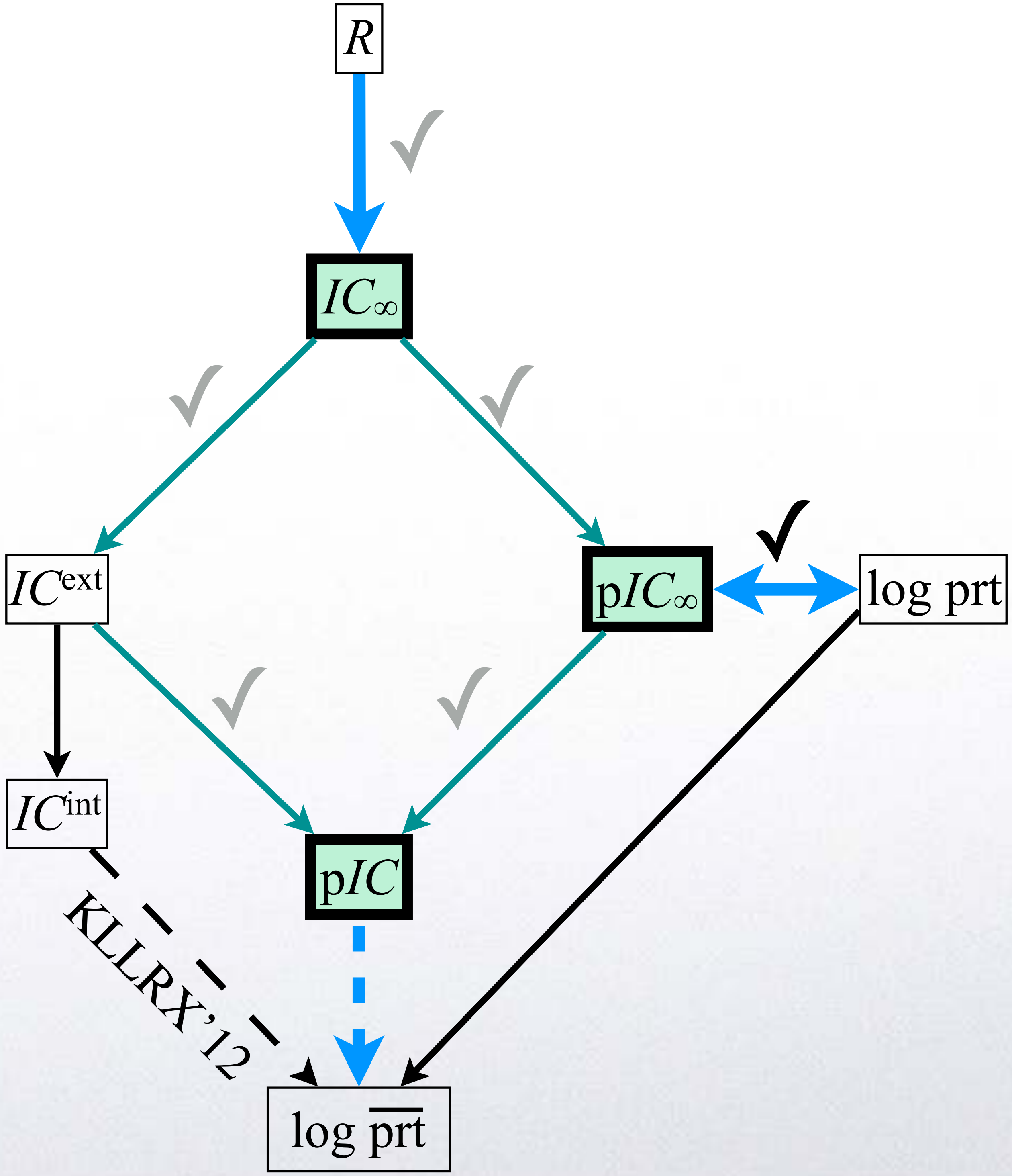
Pseudo-transcript distribution → Partition



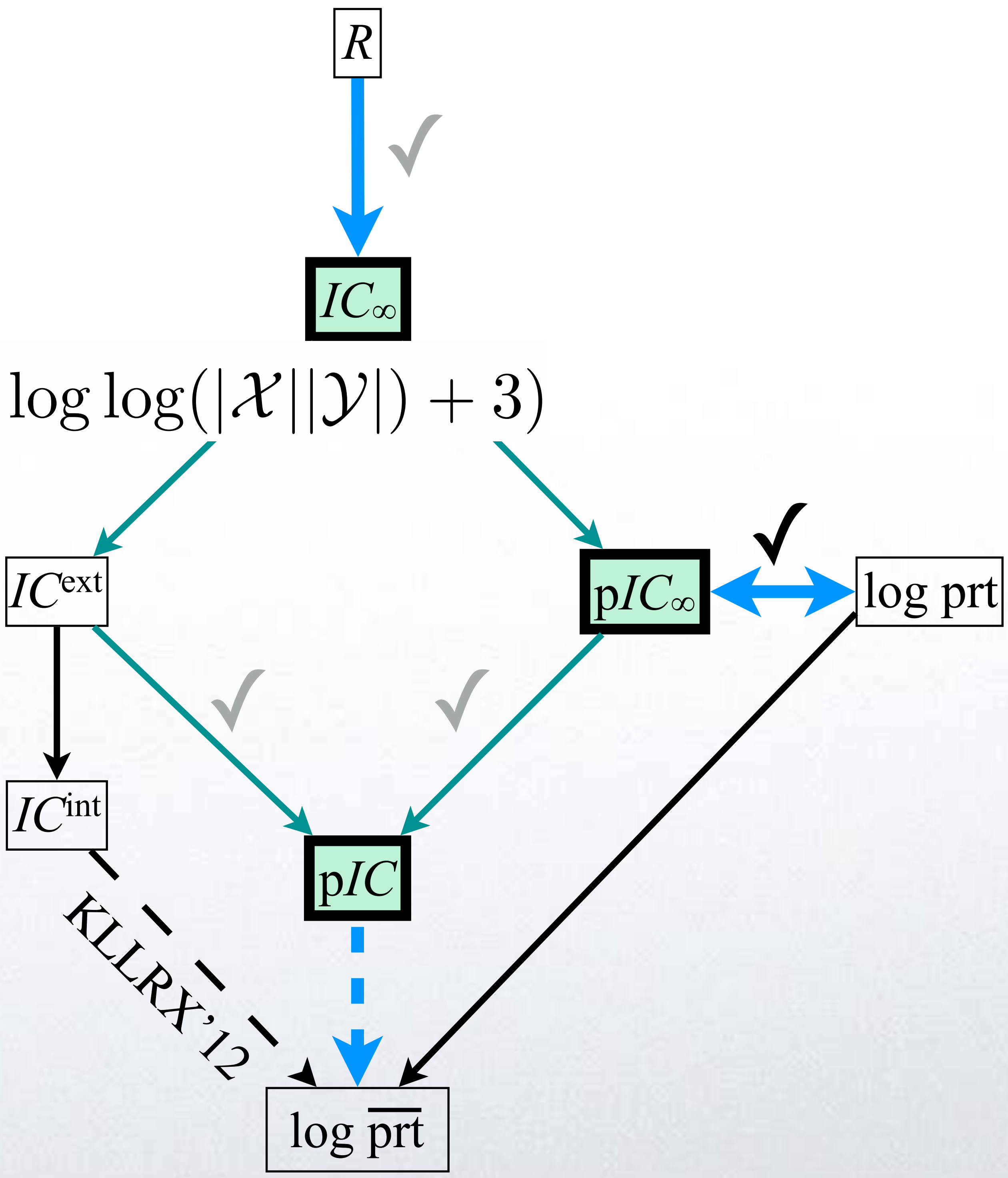
$$w(T) = \Sigma_q\, \sigma_{q,T} \times \tau_{q,T}$$

$$\Sigma_T\, w(T) = \Sigma_q\, (\Sigma_T\, \sigma_{q,T}) \cdot (\Sigma_T\, \tau_{q,T}) = \Sigma_q\, \max_{(x,y)} \mathrm{p}(q \mid x,y) = \exp(\mathrm{p}IC_\infty(Q))$$

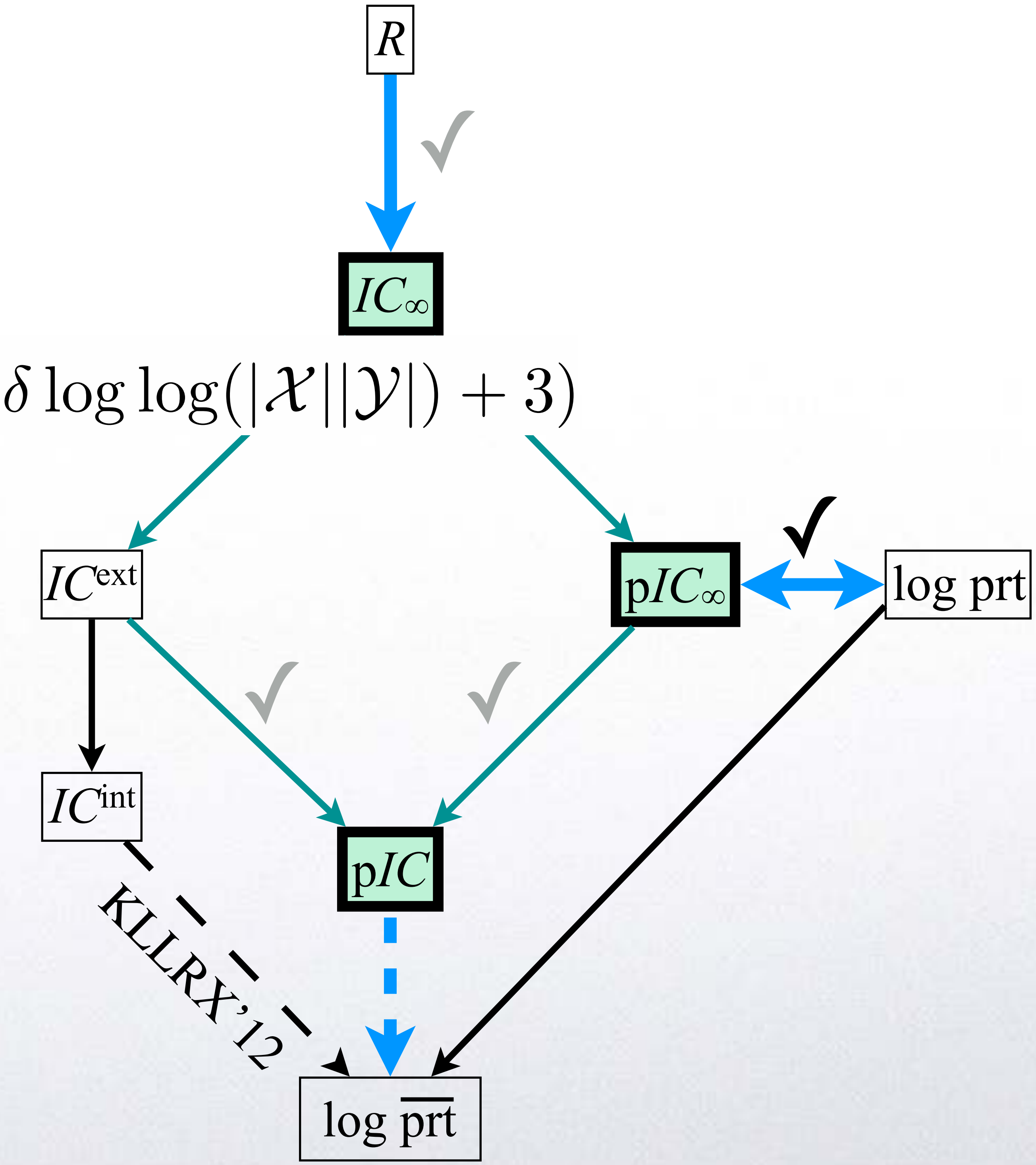
- **Theorem**: $\mathrm{p}IC_\infty(f,\varepsilon) = \log \mathrm{prt}\,(f,\varepsilon)$



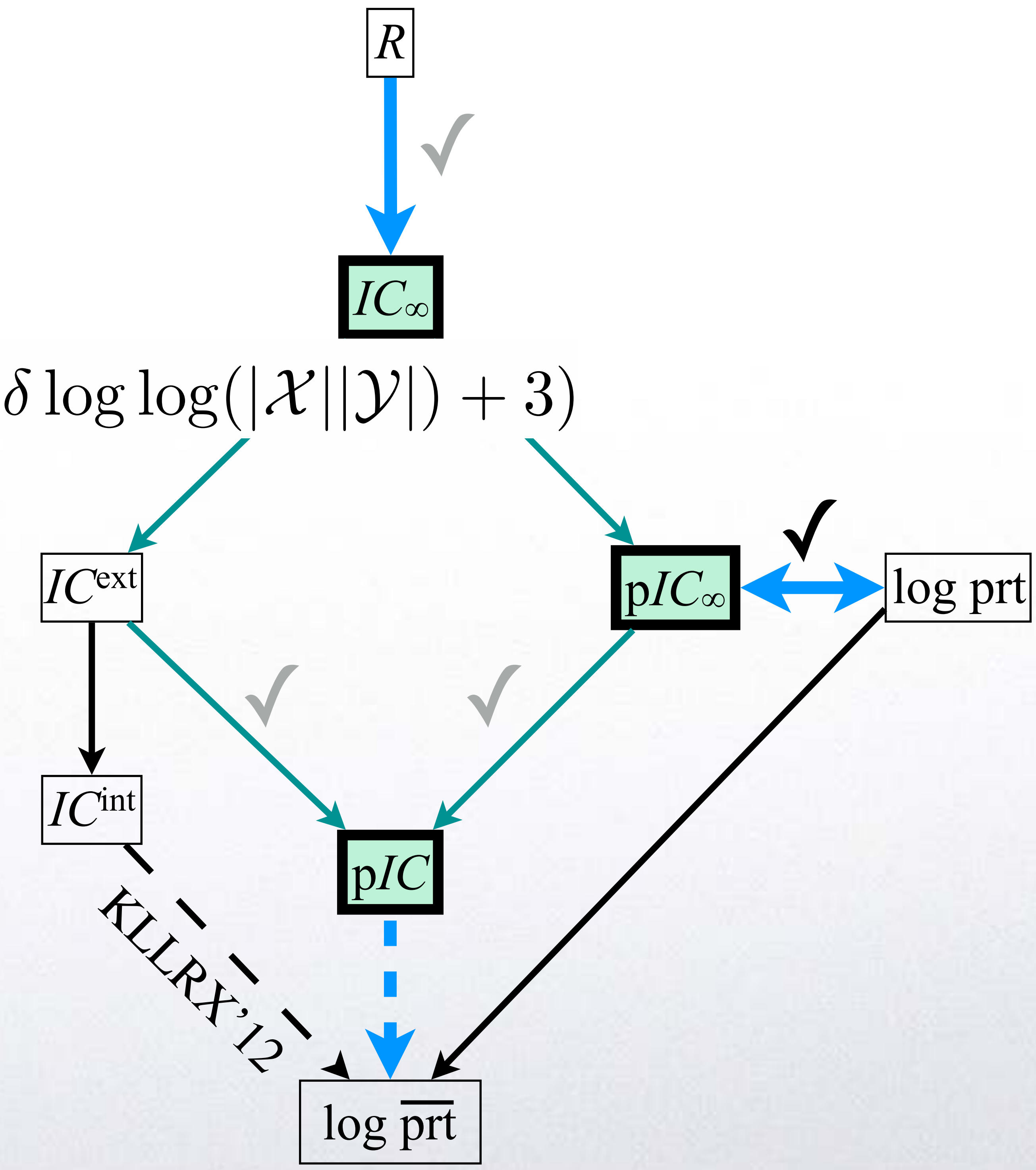
- **Theorem**: $\mathrm{p}IC_\infty(f,\varepsilon) = \log \mathrm{prt}\,(f,\varepsilon)$
- **Theorem**:
$$\mathrm{p}IC(f,\varepsilon) \geq \delta \log \overline{\mathrm{prt}}(f,\varepsilon+\delta) - (\delta \log\log(|\mathcal{X}||\mathcal{Y}|) + 3)$$



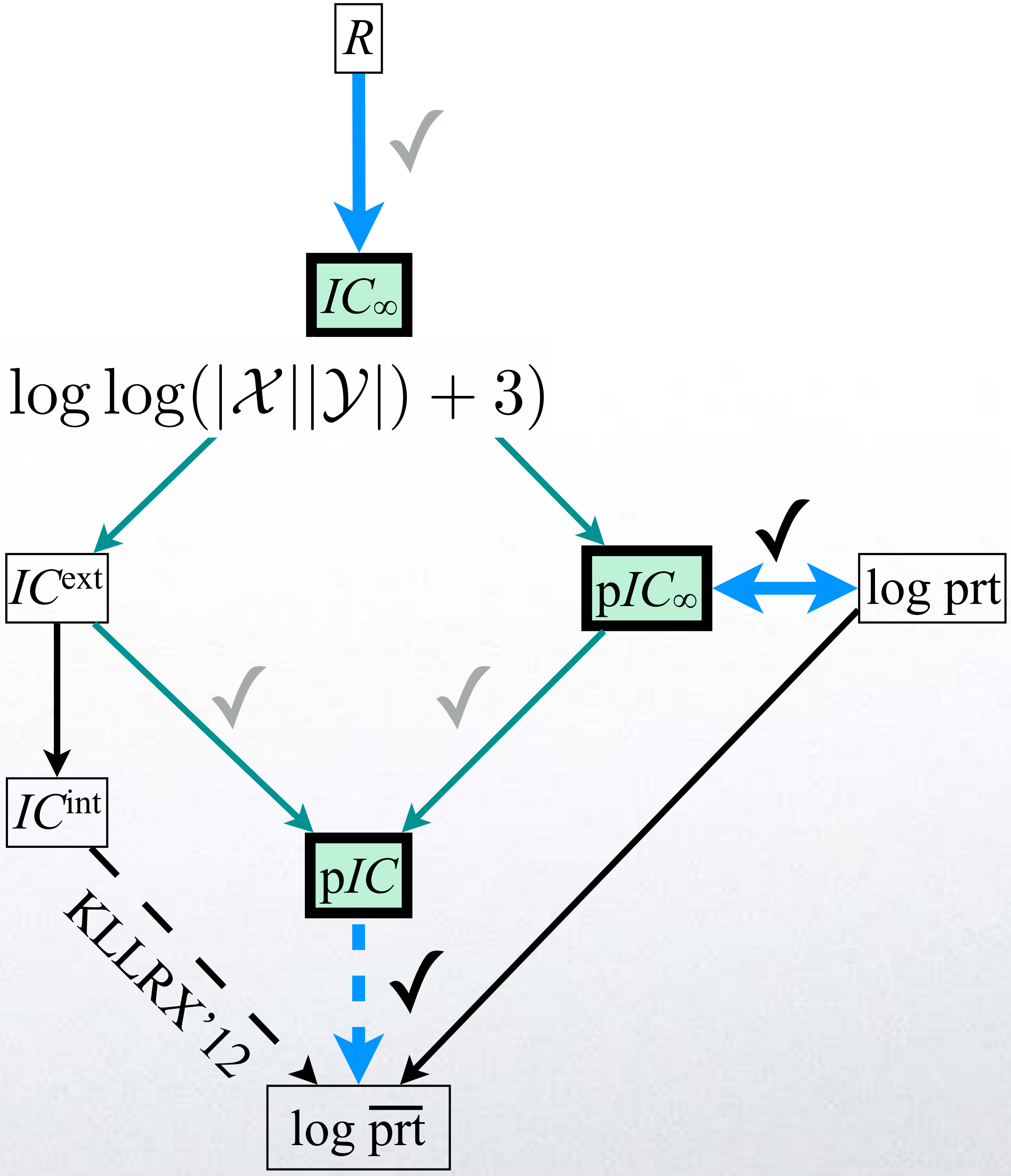
- **Theorem**: $\mathrm{p}IC_\infty(f,\varepsilon) = \log \mathrm{prt}\,(f,\varepsilon)$
- **Theorem**:
  $$\mathrm{p}IC(f,\varepsilon) \geq \delta \log \overline{\mathrm{prt}}(f,\varepsilon+\delta) - (\delta \log\log(|\mathcal{X}||\mathcal{Y}|) + 3)$$
- cf. [KLLRX'12] gives a factor of δ^2 instead of δ in the leading term, but gives a comparison with IC^{int}



- **Theorem**: $\mathrm{p}IC_\infty(f,\varepsilon) = \log \mathrm{prt}\,(f,\varepsilon)$
- **Theorem**:
  $\mathrm{p}IC(f,\varepsilon) \geq \delta \log \overline{\mathrm{prt}}(f,\varepsilon+\delta) - (\delta \log\log(|\mathcal{X}||\mathcal{Y}|) + 3)$
- cf. [KLLRX'12] gives a factor of δ^2 instead of δ in the leading term, but gives a comparison with IC^{int}
- **Idea**: $w(T) = \Sigma_{q:q\in G(T)}\, \sigma_{q,T} \cdot \tau_{q,T}$, s.t. $q \in G(T)$ iff T is a "large" tile for q
  - Gives a lower cost $\Sigma_T\, w(T) \approx I(XY;Q)/\delta$. Bound the extra error incurred.



- **Theorem**: $\mathrm{p}IC_\infty(f,\varepsilon) = \log \mathrm{prt}\,(f,\varepsilon)$
- **Theorem**:
  $$\mathrm{p}IC(f,\varepsilon) \geq \delta \log \overline{\mathrm{prt}}(f,\varepsilon+\delta) - (\delta \log\log(|\mathcal{X}||\mathcal{Y}|)+3)$$
- cf. [KLLRX'12] gives a factor of δ^2 instead of δ in the leading term, but gives a comparison with IC^{int}
- **Idea**: $w(T) = \Sigma_{q:q\in G(T)}\, \sigma_{q,T} \cdot \tau_{q,T}$, s.t. $q \in G(T)$ iff T is a "large" tile for q
  - Gives a lower cost $\Sigma_T\, w(T) \approx I(XY;Q)/\delta$. Bound the extra error incurred.



Summary

- New lower bounds IC_α, $\mathrm{p}IC_\alpha$
- Clarifies the connection between partition bound and information complexity
- Questions:
  - Techniques to lower bound IC_∞ that don't apply to prt or IC?
  - Separate R and IC^{ext} via IC_∞?
  - New techniques to lower bound IC^{ext} that don't apply to $\mathrm{p}IC$?
  - Consider IC_∞ (and $\mathrm{p}IC_\infty$) corresponding to IC^{int}? What is the analog of prt then?