*Global and Stochastic Analysis*
*Vol. 13 No. 2 (2026)*

**DOI: 10.64837/GSA.13.2.3**

*Received: 12th December 2025* *Revised: 18th February 2026* *Accepted: 20th March 2026*

# DATA-PROCESSING INEQUALITIES FOR NON-SYMMETRIC FUZZY DIVERGENCE MEASURES

ASOKAN VASUDEVAN[1], YOGEESH N[2]*, MOHAMMED EL KHIDER[3], S. VAIRACHILAI[4], ZETTY PAKIR MASTAN[5], PARAMESHWARAN CHANDRA SEGARAN[6]

**ABSTRACT**: In this study, we introduce a class of mass-adjusted directed fuzzy $f$-divergences for discrete fuzzy states and establish data-processing inequalities under stochastic channels. Each expected fuzzy object is decomposed into total mass and a normalized profile, and the resulting divergence is the sum of a Csiszar

profile term and a convex mass-ratio penalty. We proved nonnegativity, strict separation, channel monotonicity, equality under sufficiency, coarse-graining inequalities, and also a strong data-processing estimate driven by a contraction coefficient. Here under two-sided curvature bounds on the generator, the profile term is quantitatively comparable with the fuzzy Kullback-Leibler divergence, yielding Pinsker-type lower bounds of this study. Binary and four-state examples, together with illustrate the contraction of directed fuzzy divergences under aggregation and noisy processing. The framework of this study covers fuzzy Kullback-Leibler, Pearson, Hellinger-type, and power divergences as special cases.

**Keywords**: fuzzy divergence, non-symmetric divergence, data-processing inequality, fuzzy channel, Csiszár $f$-divergence, coarse graining, process innovation.

## 1. Introduction

Since in the beginning Zadeh's introduction of fuzzy sets, information-theoretic quantities have played a central role in the measurement of uncertainty, discrimination, and decision quality in fuzzy environments [39]. The concept entropy viewpoint was axiomatized by De Luca and Termini [7, 8], while fuzzy divergence measures were developed by Bhandari, Pal, and Majumder [2] and placed on a structural footing by Montes, Couso, Gil, and Bertoluzza [24]. More recent and updated work extends directed divergence ideas to generalized, intuitionistic, picture, and Pythagorean fuzzy settings, often with various decision-making applications [1,3,14,17,16].

On the probabilistic side, directed divergences and their monotonicity properties are very known and classical. Foundational milestones include Kullback-Leibler divergence [20], Shannon's information theory [30, 31], Csiszár's $f$-divergence and informativity functionals [5], Sibson's information radius [33], and modern refinements such as sharp $f$-divergence inequalities, asymptotic comparison, and also for strong data-processing theory [9,10,11,12,13]. The data-processing principle is one of the most robust structural facts in the subject, which was after stochastic post-processing, information cannot increase.

This study develops an analogous theorem-driven module for nonsymmetric fuzzy divergence measures. The key most observation is that a discrete fuzzy object naturally carries two kinds of information that are a shape component, encoded by a normalized membership profile, and an amplitude component, encoded by the total fuzzy mass. This leads to a mass-adjusted directed fuzzy divergence

$$\mathcal{D}^{\lambda}_{f,g}(a\|b) = D_f(p(a)\|p(b)) + \lambda g\left(\frac{m(a)}{m(b)}\right),$$

where $D_f$ is a classical Csiszár divergence applied to the normalized profiles and $g$ is a convex penalty for mass mismatch. The first term is genuinely channel-sensitive; the second term is invariant under mass-preserving channels. This separation allows us to prove clean fuzzy data-processing inequalities with essentially no loss of analytical sharpness.

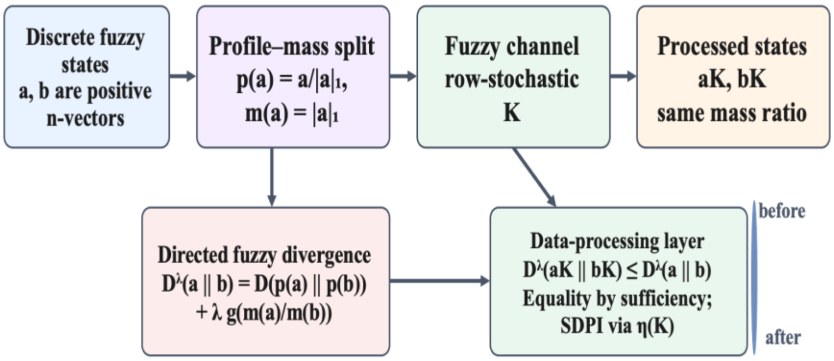

**Figure 1**. The profile-mass formalism used in the paper.

A discrete fuzzy state is split into its normalized profile and expected total mass of this study. Channel action affects the profile term through a stochastic matrix and leaves the mass ratio invariant. This makes the data-processing argument transparent.

Our emphasis is deliberately mathematical. In the finite-state setting we establish:

(i) nonnegativity, strict separation, and an exact characterization of asymmetry;
(ii) a data-processing inequality for arbitrary row-stochastic fuzzy channels;
(iii) equality under a sufficiency condition formulated by a reverse stochastic kernel;
(iv) a strong data-processing estimate via the contraction coefficient of the profile term;
(v) comparison theorems with fuzzy Kullback-Leibler divergence under curvature bounds on the generator;
(vi) explicit binary and multistate examples with figures and numerical tables.

The resulting theory is flexible enough to recover fuzzy Kullback-Leibler, Pearson, Hellinger-type, and power divergences [32, 23].

## 2. Preliminaries and basic properties

We work on the positive cone

$$\mathcal{M}_n^{\circ} = \{a = (a_1, \dots, a_n) \in (0, \infty)^n\}.$$

A finite fuzzy set with strictly positive membership vector is a particular element of $\mathcal{M}_n^{\circ}$. The positivity hypothesis simplifies the formulas and the proof of the data-processing inequality; the nonnegative case with the usual absolute-continuity convention follows by approximation.

**Definition 2.1.** For $a \in \mathcal{M}_n^\circ$ we define its total mass and normalized profile by

$$m(a) = \sum_{i=1}^{n} a_i, p(a) = \frac{a}{m(a)} \in \Delta_n^\circ$$

where

$$\Delta_n^\circ = \left\{p = (p_1, \dots, p_n) \in (0,1)^n : \sum_{i=1}^{n} p_i = 1\right\}$$

**Remark 2.2.** The profile $p(a)$ captures the relative distribution of membership intensity across the universe, while $m(a)$ measures its overall amplitude. The proposed divergence treats these two sources of asymmetry separately.

**Definition 2.3.** A fuzzy channel from $n$ states to $m$ states is a row-stochastic matrix $K = (K_{ij}) \in [0,1]^{n \times m}$ satisfying

$$\sum_{j=1}^{m} K_{ij} = 1 \text{ for every } i = 1, \dots, n$$

We always discard zero output columns, so that each column of $K$ receives positive mass from at least one input state.

For a row vector $a \in \mathcal{M}_n^\circ$, the processed fuzzy state is the row vector $aK \in \mathcal{M}_m^\circ$. Since $K$ is row-stochastic,

$$m(aK) = m(a) \text{ and } p(aK) = p(a)K$$

The channel therefore preserves the total mass and acts linearly on the normalized profile.

**Definition 2.4.** Let $f: (0, \infty) \to \mathbb{R}$ be convex with $f(1) = 0$, let $g: (0, \infty) \to [0, \infty)$ be convex with $g(1) = 0$, and let $\lambda \geq 0$. The associated mass-adjusted directed fuzzy divergence is defined by

$$\mathcal{D}_{f,g}^{\lambda}(a \| b) = \sum_{i=1}^{n} p(b)_i f\left(\frac{p(a)_i}{p(b)_i}\right) + \lambda g\left(\frac{m(a)}{m(b)}\right), a, b \in \mathcal{M}_n^\circ$$

The first term is the Csiszár $f$-divergence of the profiles; the second term penalizes mass asymmetry.

The profile part will be denoted by

$$D_f(p \| q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right), p, q \in \Delta_n^\circ$$

Hence

$$\mathcal{D}_{f,g}^{\lambda}(a \| b) = D_f(p(a) \| p(b)) + \lambda g(m(a)/m(b))$$

**Proposition 2.5.** Let $f$ and $g$ be convex with $f(1) = g(1) = 0$.

(a) For all $a, b \in \mathcal{M}_n^\circ$ one has $\mathcal{D}_{f,g}^{\lambda}(a \| b) \geq 0$.

(b) If $f$ is strictly convex at 1, then

$$\mathcal{D}^{0}_{f,g}(a\|b) = 0 \iff p(a) = p(b).$$

**Table 1**. Representative members of the mass-adjusted directed fuzzy $f$-divergence family. Here $\rho = m(a)/m(b)$ and $\alpha \neq 0,1$.

| **Generator $f(t)$** | **Mass penalty $g(\rho)$** | **Resulting measure** | **Symmetry status** | **DPI** |
|---|---|---|---|---|
| $t\log t - t + 1$ | $\rho\log\rho - \rho + 1$ | fuzzy Kullback-Leibler divergence | directed | yes |
| $(t-1)^2$ | $(\rho-1)^2$ | fuzzy Pearson $\chi^2$ divergence | directed | yes |
| $(\sqrt{t}-1)^2$ | $(\sqrt{\rho}-1)^2$ | Hellinger-type fuzzy divergence | profile term symmetric | yes |
| $\dfrac{t^\alpha - \alpha t + \alpha - 1}{\alpha(\alpha-1)}$ | $\dfrac{\rho^\alpha - \alpha\rho + \alpha - 1}{\alpha(\alpha-1)}$ | power divergence family | generically directed | yes |

(c) If $f$ and $g$ are strictly convex at 1 and $\lambda > 0$, then

$$\mathcal{D}^{\lambda}_{f,g}(a\|b) = 0 \iff a = b.$$

**Proof**. By Jensen's inequality,

$$D_f(p(a)\|p(b)) = \sum_{i=1}^{n} p(b)_i f\left(\frac{p(a)_i}{p(b)_i}\right) \geq f\left(\sum_{i=1}^{n} p(a)_i\right) = f(1) = 0.$$

Since $g \geq 0$ and $\lambda \geq 0$, part (a) follows.

If $\lambda = 0$ and $\mathcal{D}^{0}_{f,g}(a\|b) = 0$, then equality holds in Jensen's inequality above. Strict convexity of $f$ implies that the ratios $p(a)_i/p(b)_i$ are constant in $i$. Summing against $p(b)_i$ shows that the constant must be 1, hence $p(a) = p(b)$. The converse is immediate. This proves (b).

For (c), the vanishing of $\mathcal{D}^{\lambda}_{f,g}(a\|b)$ implies both $D_f(p(a)\|p(b)) = 0$ and $g(m(a)/m(b)) = 0$. By (b) the profiles coincide, and strict convexity of $g$ at 1 gives $m(a) = m(b)$. Therefore $a = m(a)p(a) = m(b)p(b) = b$.

The next proposition makes the source of non-symmetry explicit.

**Proposition 2.6**. Define the reciprocal generators

$$f^{\diamond}(t) = tf(1/t), g^{\diamond}(t) = g(1/t), t > 0$$

Then, for all $a, b \in \mathcal{M}_n^{\circ}$,

$$\mathcal{D}^{\lambda}_{f,g}(b\|a) = \mathcal{D}^{\lambda}_{f^{\diamond},g^{\diamond}}(a\|b).$$

In particular, $\mathcal{D}^{\lambda}_{f,g}$ is symmetric if and only if $f = f^{\diamond}$ and $g(t) = g(1/t)$ for all $t > 0$.

**Proof.** Write $p = p(a)$ and $q = p(b)$. Then

$$D_f(q\|p) = \sum_{i=1}^{n} p_i f\left(\frac{q_i}{p_i}\right) = \sum_{i=1}^{n} q_i \frac{p_i}{q_i} f\left(\frac{q_i}{p_i}\right) = \sum_{i=1}^{n} q_i f^{\circ}\left(\frac{p_i}{q_i}\right) = D_{f^{\circ}}(p\|q).$$

The identity for the mass term is immediate.

Table 1 lists several important choices of generator and mass penalty covered by the theory.

### 3. Data-processing inequalities

We now prove the main structural result. The argument is the classical Jensen proof of the data-processing inequality, but the mass term requires separate bookkeeping.

**Theorem 3.1** (*Data-processing inequality*). Let $f: (0,\infty) \to \mathbb{R}$ be convex with $f(1) = 0$, let $g: (0,\infty) \to [0,\infty)$ be convex with $g(1) = 0$, and let $K$ be a fuzzy channel. Then, for all $a, b \in \mathcal{M}_n^{\circ}$,

$$\mathcal{D}_{f,g}^{\lambda}(aK\|bK) \le \mathcal{D}_{f,g}^{\lambda}(a\|b)$$

**Proof**. Set $p = p(a)$ and $q = p(b)$. Since $m(aK) = m(a)$ and $m(bK) = m(b)$, the mass penalty is invariant:

$$g\left(\frac{m(aK)}{m(bK)}\right) = g\left(\frac{m(a)}{m(b)}\right)$$

Thus, it remains to prove

$$D_f(pK\|qK) \le D_f(p\|q)$$

Fix an output index $j$. Because $K$ is row-stochastic and the $j$-th output column is nonzero,

$$(qK)_j = \sum_{i=1}^{n} q_i K_{ij} > 0$$

Define weights

$$\omega_{ij} = \frac{q_i K_{ij}}{(qK)_j}, i = 1, \dots, n$$

Then $\omega_{ij} \ge 0$ and $\sum_i \omega_{ij} = 1$. Moreover,

$$\frac{(pK)_j}{(qK)_j} = \frac{\sum_i p_i K_{ij}}{\sum_i q_i K_{ij}} = \sum_{i=1}^{n} \omega_{ij} \frac{p_i}{q_i}$$

By convexity of $f$,

$$f\left(\frac{(pK)_j}{(qK)_j}\right) \le \sum_{i=1}^{n} \omega_{ij} f\left(\frac{p_i}{q_i}\right)$$

Multiplying by $(qK)_j$ and summing over $j$ gives

$$
\begin{aligned}
D_f(pK\|qK) &= \sum_{j=1}^{m} (qK)_j f\left(\frac{(pK)_j}{(qK)_j}\right) \\
&\leq \sum_{j=1}^{m} \sum_{i=1}^{n} q_i K_{ij} f\left(\frac{p_i}{q_i}\right) \\
&= \sum_{i=1}^{n} q_i \left(\sum_{j=1}^{m} K_{ij}\right) f\left(\frac{p_i}{q_i}\right) \\
&= \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right) = D_f(p\|q)
\end{aligned}
$$

which proves the claim.

A first consequence is the monotonicity of fuzzy divergence under aggregation.

**Corollary 3.2** (***Coarse graining***). Let $\{B_1, \dots, B_m\}$ be a partition of the underlying universe and let $K_{\mathcal{P}}$ be the deterministic channel defined by

$$
(K_{\mathcal{P}})_{ij} = \begin{cases} 1, & x_i \in B_j \\ 0, & x_i \notin B_j \end{cases}
$$

Then

$$
\mathcal{D}_{f,g}^{\lambda}(aK_{\mathcal{P}}\|bK_{\mathcal{P}}) \leq \mathcal{D}_{f,g}^{\lambda}(a\|b)
$$

**Definition 3.3**. A fuzzy channel $K$ is said to be sufficient for the pair ( $a, b$ ) if there exists a row-stochastic matrix $R$ such that

$$
p(a) = (p(a)K)R, p(b) = (p(b)K)R.
$$

**Theorem 3.4** (***Equality under sufficiency***). If $K$ is sufficient for ( $a, b$ ), then

$$
\mathcal{D}_{f,g}^{\lambda}(aK\|bK) = \mathcal{D}_{f,g}^{\lambda}(a\|b)
$$

**Proof**. By Theorem 3.1,

$$
\mathcal{D}_{f,g}^{\lambda}(aK\|bK) \leq \mathcal{D}_{f,g}^{\lambda}(a\|b)
$$

Applying Theorem 3.1 again to the channel $R$ and the processed pair ( $aK, bK$ ) yields

$$
\mathcal{D}_{f,g}^{\lambda}(a\|b) = \mathcal{D}_{f,g}^{\lambda}((aK)R\|(bK)R) \leq \mathcal{D}_{f,g}^{\lambda}(aK\|bK)
$$

Combining the two inequalities proves equality.

The strong data-processing coefficient of the profile term is defined by

$$
\eta_f(K) = \sup_{\substack{p,q \in \Delta_n^{\circ} \\ p \neq q}} \frac{D_f(pK\|qK)}{D_f(p\|q)}
$$

By Theorem 3.1, one has $0 \leq \eta_f(K) \leq 1$.

**Proposition 3.5** (***Strong data-processing estimate***). For every $a, b \in \mathcal{M}_n^{\circ}$,

$$
\mathcal{D}_{f,g}^{\lambda}(aK\|bK) - \lambda g\left(\frac{m(a)}{m(b)}\right) \leq \eta_f(K)\left[\mathcal{D}_{f,g}^{\lambda}(a\|b) - \lambda g\left(\frac{m(a)}{m(b)}\right)\right].
$$

In particular, if $m(a) = m(b)$, then

$$\mathcal{D}^{\lambda}_{f,g}(aK\|bK) \leq \eta_f(K)\mathcal{D}^{\lambda}_{f,g}(a\|b).$$

**Proof**. With $p = p(a)$ and $q = p(b)$ one has

$$\mathcal{D}^{\lambda}_{f,g}(a\|b) - \lambda g\left(\frac{m(a)}{m(b)}\right) = D_f(p\|q)$$

and similarly for the processed pair. Hence

$$\begin{aligned}\mathcal{D}^{\lambda}_{f,g}(aK\|bK) - \lambda g\left(\frac{m(a)}{m(b)}\right) &= D_f(pK\|qK)\\ &\leq \eta_f(K)D_f(p\|q)\\ &= \eta_f(K)\left[\mathcal{D}^{\lambda}_{f,g}(a\|b) - \lambda g\left(\frac{m(a)}{m(b)}\right)\right]\end{aligned}$$

If $m(a) = m(b)$, then $g(1) = 0$ and the second statement follows.

**Proposition 3.6**. Let $K$ be a fuzzy channel.

(a) If $K$ is a permutation matrix, then $\eta_f(K) = 1$.
(b) If all rows of $K$ are identical, then $\eta_f(K) = 0$.

**Proof.** A permutation merely relabels coordinates, so $D_f(pK\|qK) = D_f(p\|q)$ for all $p, q$, which gives (a). If all rows are identical, then $pK = qK$ for every pair of profiles $p, q$, hence $D_f(pK\|qK) = 0$ and (b) follows.

## 4. Comparison with fuzzy Kullback-Leibler divergence

Let

$$\Phi(t) = t\log t - t + 1, t > 0$$

Then $D_\Phi(p\|q) = \mathrm{KL}(p\|q)$, the usual Kullback-Leibler divergence. The fuzzy Kullback-Leibler divergence with matching mass penalty is therefore

$$\mathcal{D}^{\lambda}_{\Phi,\Phi}(a\|b) = \mathrm{KL}(p(a)\|p(b)) + \lambda\Phi\left(\frac{m(a)}{m(b)}\right)$$

The next theorem gives a generator-comparison principle under a curvature bound.
**Theorem 4.1**. Let $0 < r \leq R < \infty$, and let $f \in C^2([r, R])$ satisfy $f(1) = 0$. Assume that

$$0 < c_1 \leq tf''(t) \leq c_2 \text{ for all } t \in [r, R].$$

If $p, q \in \Delta_n^\circ$ and

$$r \leq \frac{p_i}{q_i} \leq R \text{ for } i = 1, \dots, n$$

then

$$c_1\mathrm{KL}(p\|q) \leq D_f(p\|q) \leq c_2\mathrm{KL}(p\|q)$$

**Proof.** Since $\Phi''(t) = 1/t$, the assumptions imply

$$(f - c_1\Phi)''(t) = f''(t) - \frac{c_1}{t} \geq 0, (c_2\Phi - f)''(t) = \frac{c_2}{t} - f''(t) \geq 0$$

on $[r,R]$. Hence both $f-c_1\Phi$ and $c_2\Phi-f$ are convex on $[r,R]$, and both vanish at 1. Applying their induced Csiszár divergences to the ratios $p_i/q_i\in[r,R]$ yields

$$D_{f-c_1\Phi}(p\|q)\geq 0, D_{c_2\Phi-f}(p\|q)\geq 0$$

which is exactly the desired pair of inequalities.

**Corollary 4.2** (***Pinsker-type lower bound***). Under the assumptions of Theorem 4.1,

$$D_f(p\|q)\geq\frac{c_1}{2}\|p-q\|_1^2$$

Consequently, if $m(a)=m(b)$, then

$$\mathcal{D}^{\lambda}_{f,g}(a\|b)\geq\frac{c_1}{2}\|p(a)-p(b)\|_1^2.$$

**Proof.** Combine Theorem 4.1 with the classical Pinsker inequality

$$\mathrm{KL}(p\|q)\geq\frac{1}{2}\|p-q\|_1^2$$

If $m(a)=m(b)$, then the mass penalty vanishes.

**Remark 4.3.** For the Pearson generator $f(t)=(t-1)^2$ one has $tf''(t)=2t$. Hence, on a ratio interval $[r,R]$,

$$2r\mathrm{KL}(p\|q)\leq\chi^2(p\|q)\leq 2R\mathrm{KL}(p\|q).$$

This provides a clean bridge between fuzzy Pearson divergence and fuzzy KullbackLeibler divergence, complementing bounds of the type studied in [21, 29, 32].

## 5. Binary channels and explicit contraction formulas

For $\varepsilon\in[0,1/2]$, let

$$B_\varepsilon=\begin{pmatrix}1-\varepsilon & \varepsilon\\ \varepsilon & 1-\varepsilon\end{pmatrix}$$

be the binary symmetric channel. If

$$a_{m,p}=m(p,1-p), b_{n,q}=n(q,1-q)$$

then

$$a_{m,p}B_\varepsilon=m(p_\varepsilon,1-p_\varepsilon), b_{n,q}B_\varepsilon=n(q_\varepsilon,1-q_\varepsilon)$$

where

$$p_\varepsilon=(1-2\varepsilon)p+\varepsilon, q_\varepsilon=(1-2\varepsilon)q+\varepsilon$$

Hence the fuzzy Kullback-Leibler divergence takes the explicit form

$$\mathcal{D}^{\lambda}_{\Phi,\Phi}\left(a_{m,p}B_\varepsilon\|b_{n,q}B_\varepsilon\right)=p_\varepsilon\log\frac{p_\varepsilon}{q_\varepsilon}+(1-p_\varepsilon)\log\frac{1-p_\varepsilon}{1-q_\varepsilon}+\lambda\Phi(m/n)$$

**Proposition 5.1.** If $0\leq\varepsilon_1\leq\varepsilon_2<1/2$, then for every convex normalized generator $f$ and every convex normalized mass penalty $g$,

$$\mathcal{D}^{\lambda}_{f,g}\left(aB_{\varepsilon_2}\|bB_{\varepsilon_2}\right)\leq\mathcal{D}^{\lambda}_{f,g}\left(aB_{\varepsilon_1}\|bB_{\varepsilon_1}\right)$$

**Proof.** Set

$$\delta = \frac{\varepsilon_2 - \varepsilon_1}{1 - 2\varepsilon_1}$$

Since $0 \leq \varepsilon_1 \leq \varepsilon_2 < 1/2$, one has $0 \leq \delta \leq 1/2$, and a direct computation shows

$$B_{\varepsilon_2} = B_{\varepsilon_1} B_\delta$$

Theorem 3.1 applied to the channel $B_\delta$ proves the claim.

Figure 2 visualizes Proposition 5.1. The mass ratio is fixed, so the mass penalty contributes the same vertical shift for every $\varepsilon$; in the picture we display the profile term only. Figure 3 shows grid-based estimates of the strong data-processing coefficient for binary channels under the Kullback-Leibler and Pearson generators.

## 6. A four-state numerical illustration

We now evaluate the theory on a simple four-state fuzzy system. Let

$$a = (0.9, 0.4, 0.7, 0.2), b = (0.6, 0.5, 0.3, 0.4), \frac{m(a)}{m(b)} = \frac{2.2}{1.8} = 1.222 \ldots$$

We compare four channels:

- (i) the identity channel;
- (ii) a local diffusion channel preserving the four-state resolution;
- (iii) a deterministic two-block aggregation channel merging states 1,2 and 3,4;
- (iv) a rank-one channel whose rows are all equal.

We evaluate the fuzzy Kullback-Leibler divergence with $\lambda = 0.25$ and the Pearsontype fuzzy divergence with the same weight and a quadratic mass penalty. The resulting values are shown in Table 2.

The numerical behavior is consistent with the theory. First, every processed value is smaller than the identity value, as predicted by Theorem 3.1. Second, the rank-one channel reduces all profile information to zero, in accordance with Proposition 3.6; what remains is exactly the mass penalty. Third, the aggregation

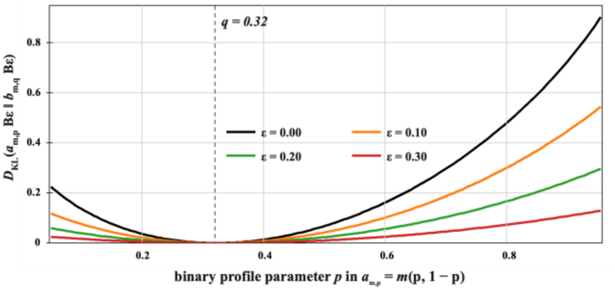

**Figure 2**. Profile-term contraction for the binary symmetric channel.

The curves show $\mathrm{KL}(pB_\varepsilon \| qB_\varepsilon)$ as a function of the input profile parameter $p$, with $q = 0.32$. Increasing channel noise shrinks the directed divergence uniformly, in agreement with Proposition 5.1.

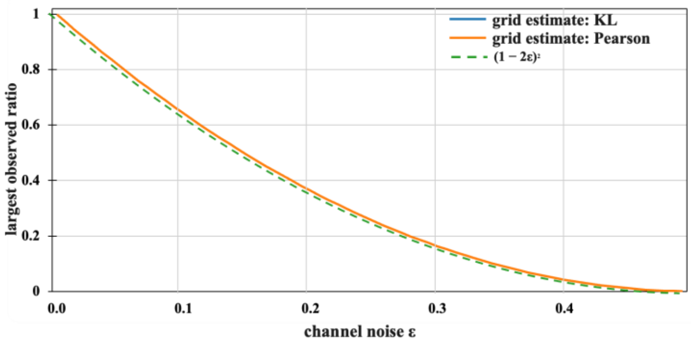

**Figure 3**. Grid-based estimates of the strong data-processing coefficient for binary channels.

The two observed contraction curves almost coincide and closely follow $(1-2\varepsilon)^2$. The figure gives a numerical picture of Proposition 3.5.

channel contracts more strongly than the local diffusion channel, reflecting the loss of resolution under coarse graining.

**Remark 6.1**. Because the total mass ratio is preserved by every row-stochastic channel, all contraction effects in Table 2 are attributable to the profile term. This is precisely the advantage of the profile-mass decomposition: it separates intrinsic information loss from the channel-invariant amplitude discrepancy.

**Table 2**. Numerical contraction of two non-symmetric fuzzy divergences under representative channels. The identity channel gives the largest value, while the rank-one channel annihilates the profile term and leaves only the mass penalty.

| Channel | Output dimension | $\mathcal{D}^{0.25}_{\Phi,\Phi}(aK\Vert bK)$ | Pearson-type fuzzy divergence | Mass ratio |
|---|---|---|---|---|
| Identity | 4 | 0.136972 | 0.278048 | 1.222222 |
| Local diffusion | 4 | 0.054390 | 0.103969 | 1.222222 |
| Two-block aggregation | 2 | 0.006614 | 0.014063 | 1.222222 |
| Rank-one collapse | 4 | 0.005760 | 0.012346 | 1.222222 |

## 7. Conclusion

We proposed a mathematically transparent family of non-symmetric fuzzy divergence measures based on a profile-mass decomposition and proved a general data-processing inequality for fuzzy channels. The theory combines the flexibility of Csiszár $f$-divergences with a mass-ratio penalty that preserves sensitivity to amplitude mismatch. Besides the basic monotonicity theorem, we obtained equality under sufficiency, strong data-processing estimates, and generator-comparison bounds with the fuzzy Kullback-Leibler divergence. The finite-state examples confirm that both aggregation and noise suppress directed fuzzy divergence in a quantitatively predictable way.

This study suggests several directions for further study that are optimal contraction constants for wider classes of channels, continuous-state fuzzy kernels, and information geometry for directed fuzzy divergences on manifolds or statistical models. These concept that fit naturally into nonlinear and stochastic/global analysis themes while remaining will address in rigorous fuzzy information theory.

## References

[1] H. D. Arora and A. Dhiman: On some generalised information measure of fuzzy directed divergence and decision making, Int. J. Comput. Sci. Math. 7 (2016), 263-273. DOI: 10.1504/IJCSM.2016.077856.

[2] D. Bhandari, N. R. Pal and D. D. Majumder: Fuzzy divergence, probability measure of fuzzy events and image thresholding, Pattern Recognit. Lett. 13 (1992), 857-867. DOI: 10.1016/0167-8655(92)90085-E.

[3] P. K. Bhatia and S. Singh: A New Measure of Fuzzy Directed Divergence and Its Application in Image Segmentation, Int. J. Intell. Syst. Appl. 5 (2013), 81-89. DOI: 10.5815/IJISA.2013.04.08.

[4] A. Cichocki, S. Cruces and S.-I. Amari: Generalized Alpha-Beta divergences and their application to robust nonnegative matrix factorization, Entropy 13 (2011), 134-170. DOI: 10.3390/e13010134.

[5] I. Csiszár: A class of measures of informativity of observation channels, Period. Math. Hung. 2 (1972), 191-213. DOI: 10.1007/BF02018661.

[6] I. Csiszár and P. C. Shields: Information theory and statistics: A tutorial, Found. Trends Commun. Inf. Theory 1 (2004), 417-528. DOI: 10.1561/0100000004.

[7] A. De Luca and S. Termini: A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory, Inform. Control 20 (1972), 301-312. DOI: 10.1016/S0019-9958(72)90199-4.

[8] A. De Luca and S. Termini: Entropy of $L$-fuzzy sets, Inform. Control 24 (1974), 55-73. DOI: 10.1016/S0019-9958(74)80023-9.

[9] D. M. Endres and J. E. Schindelin: A new metric for probability distributions, IEEE Trans. Inform. Theory 49 (2003), 1858-1860. DOI: 10.1109/TIT.2003.813506.

[10] A. L. Gibbs and F. E. Su: On choosing and bounding probability metrics, Internat. Statist. Rev. 70 (2002), 419-435. DOI: 10.1111/j.1751-5823.2002.tb00178.x.

[11] G. L. Gilardoni: On Pinsker's and Vajda's type inequalities for Csiszár's $f$-divergences, IEEE Trans. Inform. Theory 56 (2010), 5377-5386. DOI: 10.1109/TIT.2010.2068710.

[12] A. Guntuboyina, S. Saha and G. Schiebinger: Sharp inequalities for $f$-divergences, IEEE Trans. Inform. Theory 60 (2014), 104-121. DOI: 10.1109/TIT.2013.2288674.

[13] P. Harremoës and I. Vajda: On pairs of $f$-divergences and their joint range, IEEE Trans. Inform. Theory 57 (2011), 3230-3235. DOI: 10.1109/TIT.2011.2137353.

[14] R. Kadian and S. Kumar: A new picture fuzzy divergence measure based on Jensen-Tsallis information measure and its application to multicriteria decision making, Granul. Comput. 7 (2022), 113-126. DOI: 10.1007/s41066-021-00254-6.

[15] T. Kailath: The divergence and Bhattacharyya distance measures in signal selection, IEEE Trans. Commun. Technol. 15 (1967), 52-60. DOI: 10.1109/TCOM.1967.1089532.

[16] R. Kaushik, R. K. Bajaj and T. Kumar: On intuitionistic fuzzy divergence measure with application to edge detection, Procedia Comput. Sci. 70 (2015), 2-8. DOI: 10.1016/j.procs.2015.10.017.

[17] V. Kobza: Divergence measures on intuitionistic fuzzy sets, Notes Intuitionistic Fuzzy Sets 28 (2022), 413-427. DOI: 10.7546/nifs.2022.28.4.413-427.

[18] V. Kobza, V. Janiš and S. Montes: Generalizated local divergence measures, J. Intell. Fuzzy Systems 33 (2017), 337-350. DOI: 10.3233/JIFS-161647.

[19] S. Kullback: A lower bound for discrimination information in terms of variation, IEEE Trans. Inform. Theory 13 (1967), 126-127. DOI: 10.1109/TIT.1967.1053968.

[20] S. Kullback and R. A. Leibler: On information and sufficiency, Ann. Math. Statist. 22 (1951), 79-86. DOI: 10.1214/aoms/1177729694.

[21] P. Kumar and S. Chhina: A symmetric information divergence measure of the Csiszár's $f$-divergence class and its bounds, Comput. Math. Appl. 49 (2005), 575-588. DOI: 10.1016/j.camwa.2004.07.017.

[22] J. Lin: Divergence measures based on the Shannon entropy, IEEE Trans. Inform. Theory 37 (1991), 145-151. DOI: 10.1109/18.61115.

[23] S. Maheshwari, S. Thakur and M. A. Siddiqui: A Novel Generalized Fuzzy Divergence Measure and Its Application in Decision Making, Global Stoch. Anal. 12 (2025), no. 6. DOI: 10.64837/GSA.12.6.2.

[24] S. Montes, I. Couso, P. Gil and C. Bertoluzza: Divergence measure between fuzzy sets, Int. J. Approx. Reason. 30 (2002), 91-105. DOI: 10.1016/S0888-613X(02)00063-4.

[25] M. C. Pardo and I. Vajda: On asymptotic properties of information-theoretic divergences, IEEE Trans. Inform. Theory 49 (2003), 1860-1868. DOI: 10.1109/TIT.2003.813509.

[26] Y. Polyanskiy and Y. Wu: Dissipation of information in channels with input constraints, IEEE Trans. Inform. Theory 62 (2016), 35-55. DOI: 10.1109/TIT.2015.2482978.

[27] M. S. Saidin, L. S. Lee, H.-V. Seow and S. Pickl: Fuzzy Divergence Measure Based on Technique for Order of Preference by Similarity to Ideal Solution Method for Staff Performance Appraisal, Mathematics 12 (2024), 714. DOI: 10.3390/math12050714.

[28] I. Sason: Tight bounds on symmetric divergence measures and a refined bound for lossless source coding, IEEE Trans. Inform. Theory 61 (2015), 701-707. DOI: 10.1109/TIT.2014.2387065.

[29] I. Sason and S. Verdú: f-divergence inequalities, IEEE Trans. Inform. Theory 62 (2016), 5973-6006. DOI: 10.1109/TIT.2016.2603151.

[30] C. E. Shannon: A mathematical theory of communication, Bell System Tech. J. 27 (1948), 379-423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

[31] C. E. Shannon: A mathematical theory of communication, Bell System Tech. J. 27 (1948), 623-656. DOI: 10.1002/j.1538-7305.1948.tb00917.x.

[32] A. Sharma and R. N. Saraswat: Bounds of Non-Symmetric Fuzzy Divergence Measure via Fuzzy Kullback-Leibler Divergence Measure, Global Stoch. Anal. 11 (2024), no. 2. DOI: 10.64837/GSA.11.2.6.

[33] R. Sibson: Information radius, Z. Wahrscheinlichkeitstheorie verw. Gebiete 14 (1969), 149-160. DOI: 10.1007/BF00537520.

[34] Y. Song, Q. Fu, Y.-F. Wang and X. Wang: Divergence-based cross entropy and uncertainty measures of Atanassov's intuitionistic fuzzy sets with their application in decision making, Appl. Soft Comput. 84 (2019), 105703. DOI: 10.1016/j.asoc.2019.105703.

[35] I. J. Taneja: Seven means, generalized triangular discrimination, and generating divergence measures, Information 4 (2013), 198-239. DOI: 10.3390/info4020198.

[36] T. van Erven and P. Harremoës: Rényi divergence and Kullback-Leibler divergence, *IEEE* Trans. Inform. Theory 60 (2014), 3797-3820. DOI: 10.1109/TIT.2014.2320500.

[37] I. Vajda: Note on discrimination information and variation, IEEE Trans. Inform. Theory 16 (1970), 771-773. DOI: 10.1109/TIT.1970.1054557.

[38] F. Xiao and W. Ding: Divergence measure of Pythagorean fuzzy sets and its application in medical diagnosis, Appl. Soft Comput. 79 (2019), 254-267. DOI: 10.1016/j.asoc.2019.03.043.

[39] L. A. Zadeh: Fuzzy sets, Inform. Control 8 (1965), 338-353. DOI: 10.1016/S0019-9958(65)90241-X.

[40] Q. Zhou, H. Mo and Y. Deng: A New Divergence Measure of Pythagorean Fuzzy Sets Based on Belief Function and Its Application in Medical Diagnosis, Mathematics 8 (2020), 142. DOI: 10.3390/ math8010142.

[1] **ASOKAN VASUDEVAN** : FACULTY OF BUSINESS AND COMMUNICATIONS, INTI INTERNATIONAL UNIVERSITY, NILAI 71800, MALAYSIA.
EMAIL: asokan.vasudevan@newinti.edu.my, ORCID: https://orcid.org/0000-0002-9866-4045

[2] **YOGEESH N** : RESEARCH FELLOW, INTI INTERNATIONAL UNIVERSITY, 71800 NILAI, NEGERI SEMBILAN, MALAYSIA, DEPARTMENT OF MATHEMATICS, GOVERNMENT FIRST GRADE COLLEGE, TUMKUR-572102, INDIA. EMAIL: yogeesh.r@gmail.com ORCID ID: orcid.org/0000-0001-8080-7821

[3] **MOHAMMED EL KHIDER** : DEPARTMENT OF GENERAL UNDERGRADUATE CURRICULUM REQUIREMENTS, UNIVERSITY OF DUBAI, UNITED ARAB EMIRATES, P. O BOX 14143. EMAIL: mohammeduk79@gmail.com, ORCID ID: https://orcid.org/0009-0007-1006-0786

[4] **S. VAIRACHILAI** : SCHOOL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE, SR UNIVERSITY, WARANGAL, TELANGANA 506371, INDIA. EMAIL: vairachilai2676@gmail.com, ORCID ID: https://orcid.org/0000-0002-8577-6195

[5] **ZETTY PAKIR MASTAN** : FACULTY OF ENGINEERING AND QUANTITY SURVEYING, INTI INTERNATIONAL UNIVERSITY, MALAYSIA. EMAIL: srzetty.pakirmastan@newinti.edu.my, ORCID: 0009-0003-2395-7120

[6] **PARAMESHWARAN CHANDRA SEGARAN** : FACULTY OF BUSINESS AND COMMUNICATIONS, INTI INTERNATIONAL UNIVERSITY, PERSIARAN PERDANA BBN PUTRA NILAI, 71800 NILAI, NEGERI SEMBILAN, MALAYSIA. EMAIL: parameshwaran0702@gmail.com; ORCID: https://orcid.org/0009-0002-5200-7055

*CORRESPONDING AUTHOR: YOGEESH N, yogeesh.r@gmail.com