



HAL
open science

COMPUTATION OF RATE-DISTORTION-PERCEPTION FUNCTION UNDER f -DIVERGENCE PERCEPTION CONSTRAINTS

Giuseppe Serra, Photios A Stavrou, Marios Kountouris

► **To cite this version:**

Giuseppe Serra, Photios A Stavrou, Marios Kountouris. COMPUTATION OF RATE-DISTORTION-PERCEPTION FUNCTION UNDER f -DIVERGENCE PERCEPTION CONSTRAINTS. 6G-WFF 2023, 6G Wireless Foundations Forum, Jul 2023, Sophia Antipolis (06), France. hal-04445972

HAL Id: hal-04445972

<https://hal.science/hal-04445972>

Submitted on 8 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

COMPUTATION OF RATE-DISTORTION-PERCEPTION FUNCTION UNDER f -DIVERGENCE PERCEPTION CONSTRAINTS



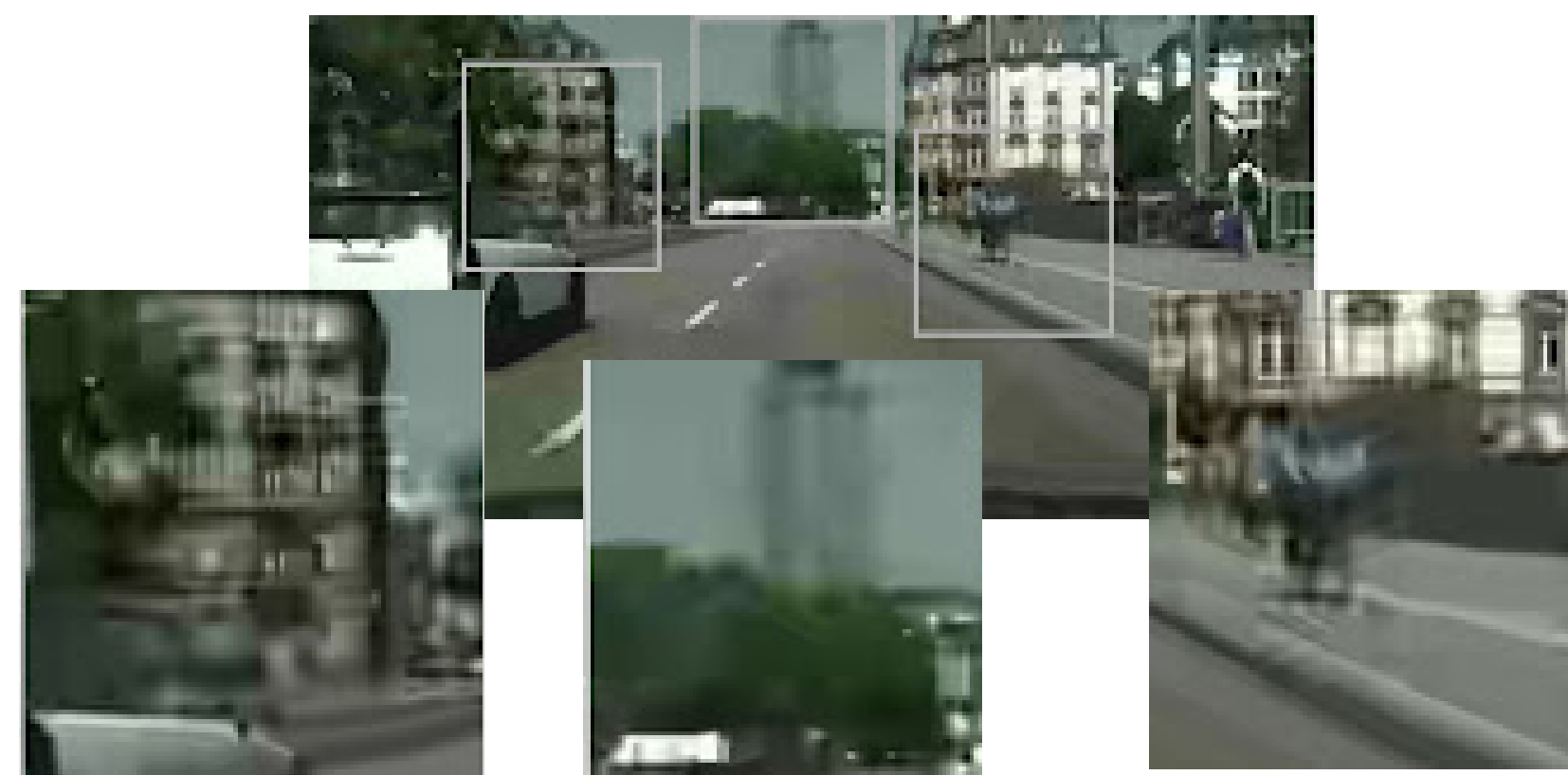
Giuseppe Serra, Photios A. Stavrou, Marios Kountouris

EURECOM, Communication Systems Department

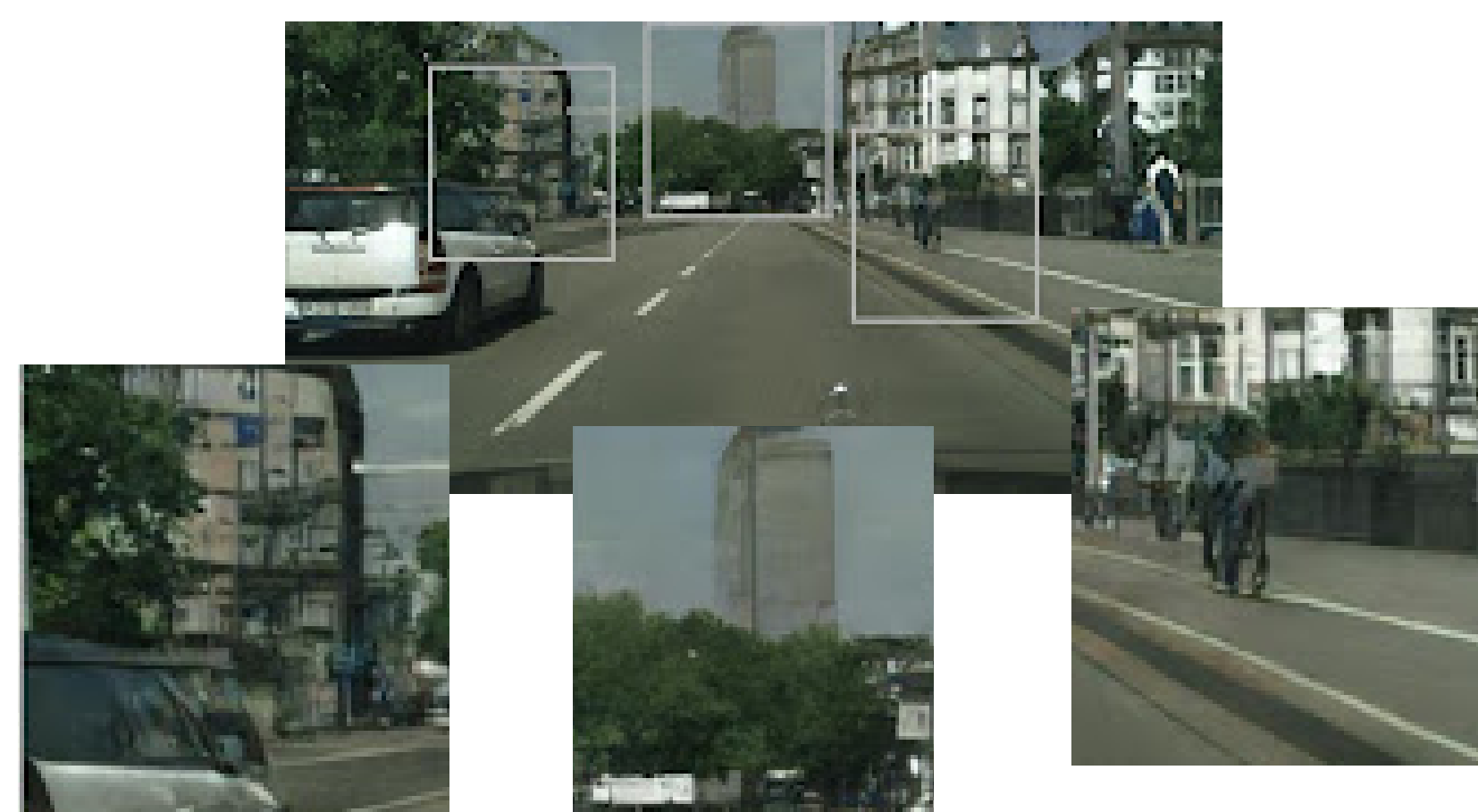
European Research Council
Established by the European Commission

MOTIVATIONS

In recent years, the application of machine learning techniques in the domains of computer vision and multimedia compression has shown the limits of the traditional source coding framework of rate-distortion (RD) theory. The key insight lies in observing that focusing exclusively on distortion minimization does not imply good perceptual quality, whereas perceptual quality refers to the property of a sample to appear visually pleasing from the human perspective. The introduction of the rate-distortion-perception (RDP) framework generalizes the notion of RD, complementing the classical distortion constraint between a source message and its reconstruction with a divergence constraint between the source distribution and the induced distribution of the reconstructed messages. This additional constraint acts as a proxy of human perception, measuring the degree of satisfaction in the consumption of data from a human perspective.



Bad Perceptual Quality - Low Distortion



Good Perceptual Quality - High Distortion

A different interpretation of this constraint links the divergence measure to a semantic quality metric, which measures the degree of relevance of the reconstructed source from the perspective of the observer.

THE RATE-DISTORTION-PERCEPTION (RDP) FUNCTION

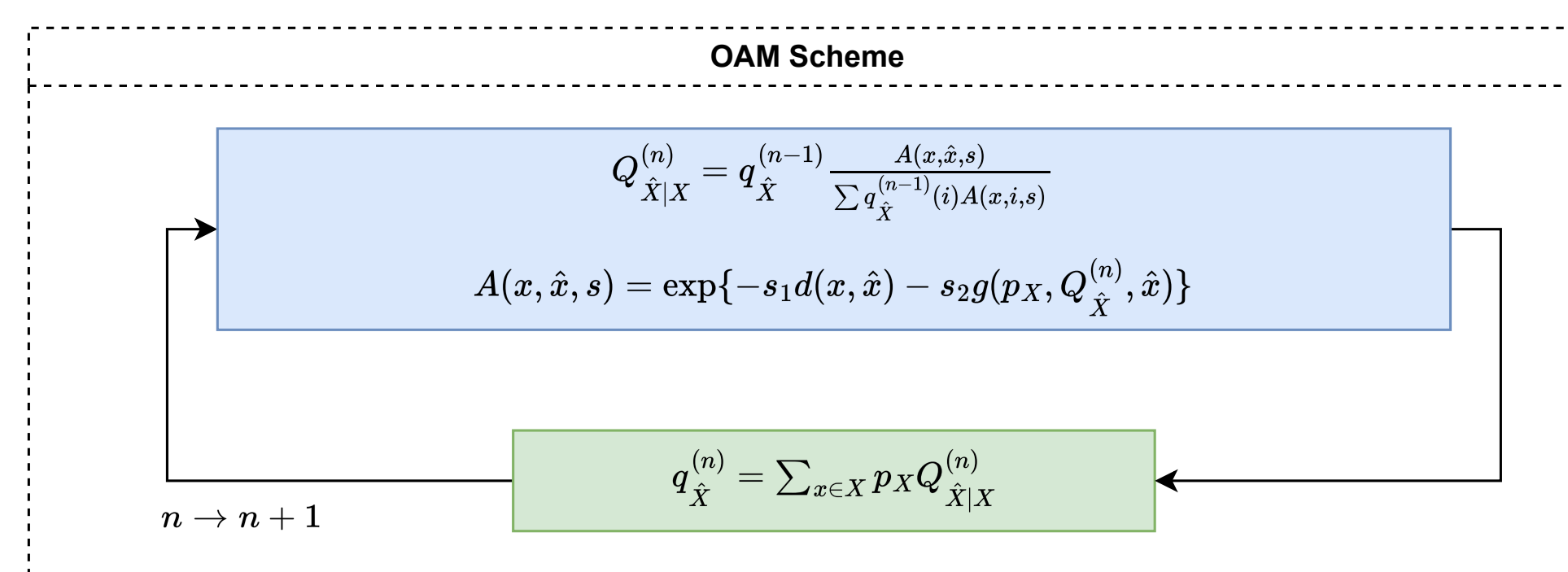
For a given finite alphabet source distribution p_x , a single-letter distortion measure $d(\cdot, \cdot)$ and a f -divergence measure $D_f(\cdot || \cdot)$, the RDPF is characterized as follows:

$$R(D, P) = \min_{Q_{\hat{x}|x}} I(X, \hat{X}) \quad \text{s.t.} \quad E[d(x, \hat{x})] \leq D \quad D_f(p_x || Q_{\hat{x}}) \leq P$$

where $D \in [D_{\min}, D_{\max}] \subset (0, \infty)$, $P \in [P_{\min}, P_{\max}] \subset (0, \infty)$

OPTIMAL ALTERNATING MINIMIZATION (OAM) SCHEME

The optimal solutions derived imposing optimality conditions to the RDP problem allow to define an iterative minimization scheme, for which we provide convergence guarantees. However, due to a recursive definition in the iteration structure, the scheme is not easily implementable.



RDP Problem

$$R(D, P) = \min_{Q_{\hat{x}|x}} I(X, \hat{X}) \quad \text{s.t.} \quad E[d(X, \hat{X})] \leq D, \quad D_f(p_x, Q_{\hat{x}}) \leq P$$

Property of Mutual Information I

Double minimization form

$$R(D, P) = \min_{Q_{\hat{x}|x}} \min_{q_{\hat{x}}} D_{KL}(p_x Q_{\hat{x}|x}, p_x q_{\hat{x}}) \quad \text{s.t.} \quad E[d(X, \hat{X})] \leq D, \quad D_f(p_x, Q_{\hat{x}}) \leq P$$

Lagrangian Formulation

$$R(D, P) = \min_{Q_{\hat{x}|x}} \min_{q_{\hat{x}}} D_{KL}(p_x Q_{\hat{x}|x}, p_x q_{\hat{x}}) + s_1 (E[d(X, \hat{X})] - D) + s_2 (D_f(p_x, Q_{\hat{x}}) - P)$$

Using KKT optimality conditions

Fixing $q_{\hat{x}}$

$$Q_{\hat{x}|x} = q_{\hat{x}} \frac{A(x, \hat{x}, s)}{\sum_{\hat{x}} q_{\hat{x}}(i) A(x, i, s)}$$

$$A(x, \hat{x}, s) = \exp\{-s_1 d(x, \hat{x}) - s_2 g(p_x, Q_{\hat{x}}, \hat{x})\}$$

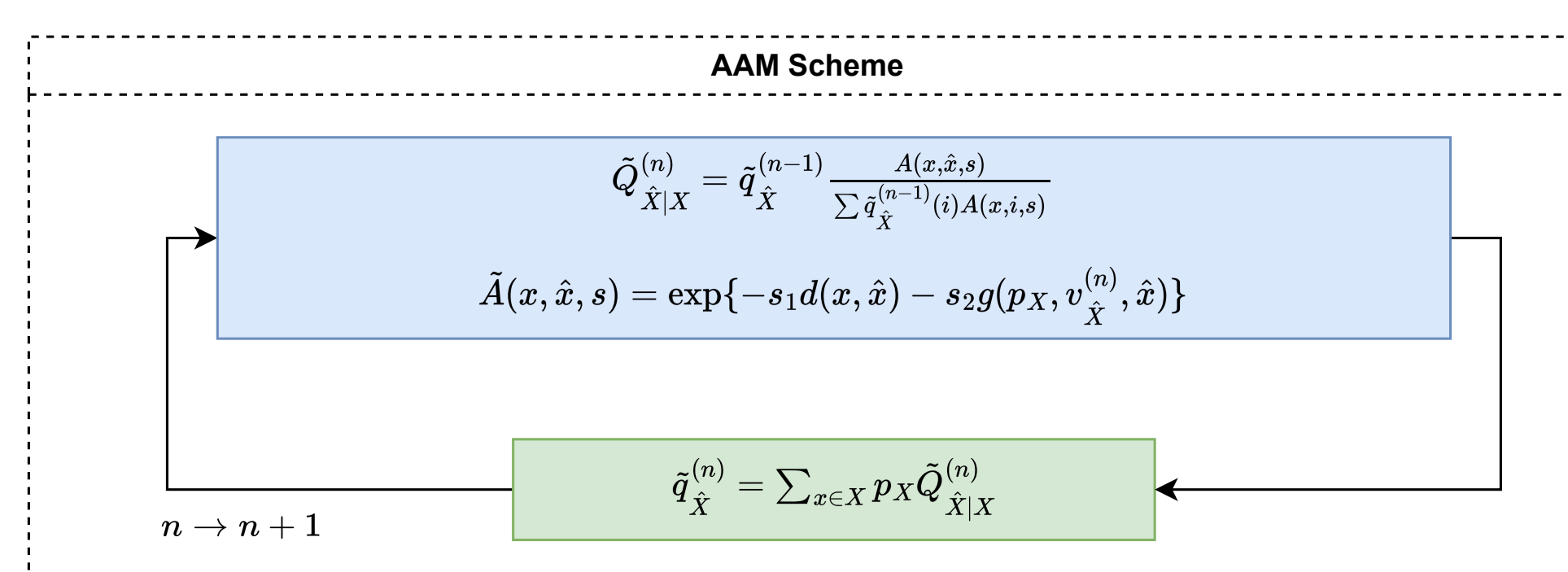
$$g(p_x, Q_{\hat{x}}, \hat{x}) = f\left(\frac{p_x(\hat{x})}{Q_{\hat{x}}(\hat{x})}\right) - \frac{p_x(\hat{x})}{Q_{\hat{x}}(\hat{x})} \partial f\left(\frac{p_x(\hat{x})}{Q_{\hat{x}}(\hat{x})}\right)$$

Fixing $Q_{\hat{x}|x}$

$$q_{\hat{x}} = \sum_{x \in \mathcal{X}} p_x Q_{\hat{x}|x}$$

APPROXIMATE ALTERNATING MINIMIZATION (AAM) SCHEME

To circumvent the implementation problem of the OAM scheme, we introduce a relaxation in its iteration structure. Under the condition that the introduced approximation $v^{(n)}$ ensures $\lim_{n \rightarrow \infty} (\hat{q}_{\hat{x}}^{(n+1)} - v_{\hat{x}}^{(n)}) \rightarrow 0$ with linear rate of convergence, then the associated AAM scheme converges to the optimal point of the RDP curve.



Optimal Solution

Fixing $q_{\hat{x}}$

$$Q_{\hat{x}|x} = q_{\hat{x}} \frac{A(x, \hat{x}, s)}{\sum_{\hat{x}} q_{\hat{x}}(i) A(x, i, s)}$$

$$A(x, \hat{x}, s) = \exp\{-s_1 d(x, \hat{x}) - s_2 g(p_x, Q_{\hat{x}}, \hat{x})\}$$

Sub-optimal solution

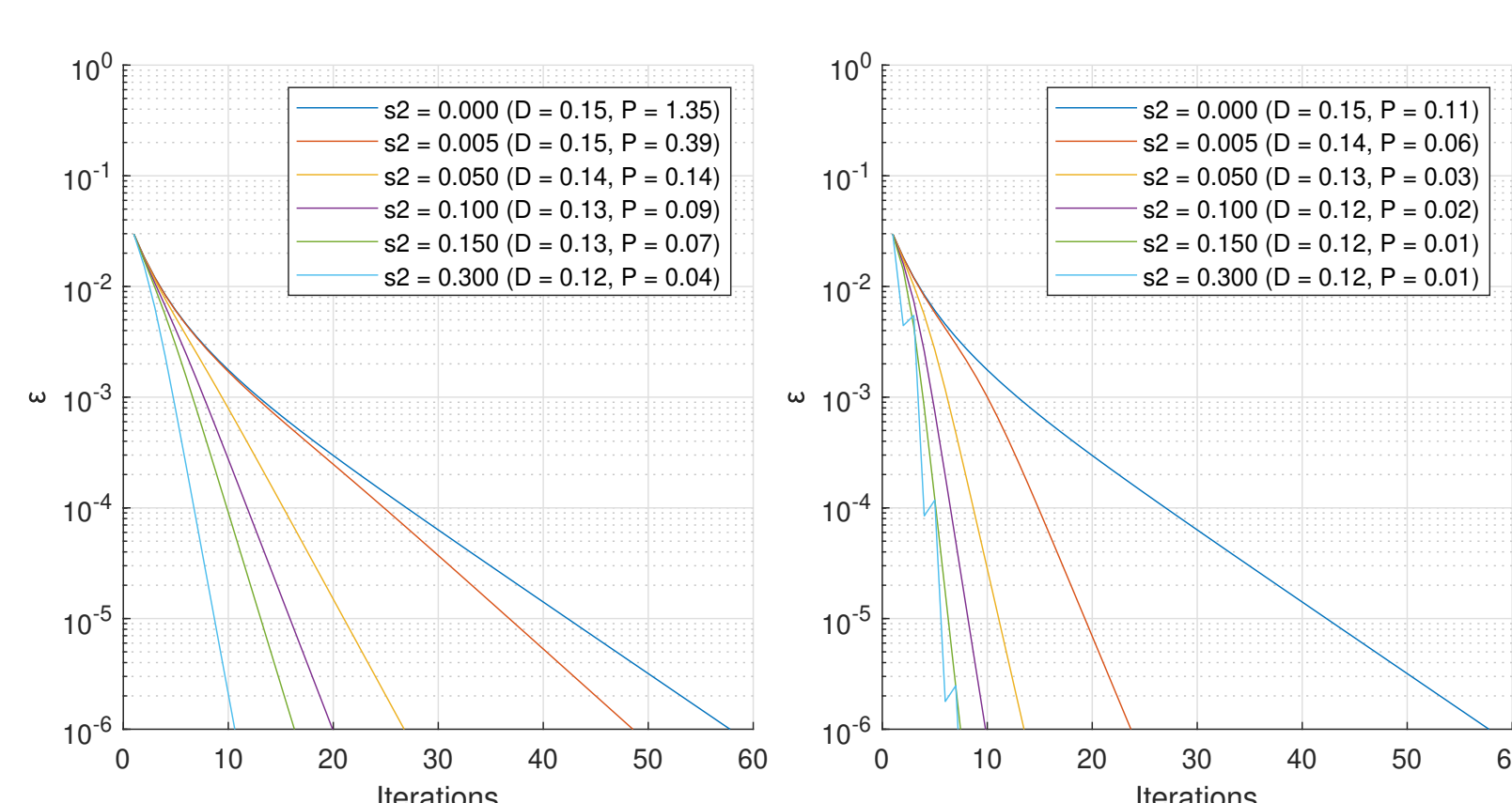
Fixing $q_{\hat{x}}$

$$\tilde{Q}_{\hat{x}|x} = \tilde{q}_{\hat{x}} \frac{A(x, \hat{x}, s)}{\sum_{\hat{x}} \tilde{q}_{\hat{x}}(i) A(x, i, s)}$$

$$\tilde{A}(x, \hat{x}, s) = \exp\{-s_1 d(x, \hat{x}) - s_2 g(p_x, v_{\hat{x}}, \hat{x})\}$$

Introducing an approximant $v_{\hat{x}}$

ASYMPTOTIC CONVERGENCE RATE ANALYSIS



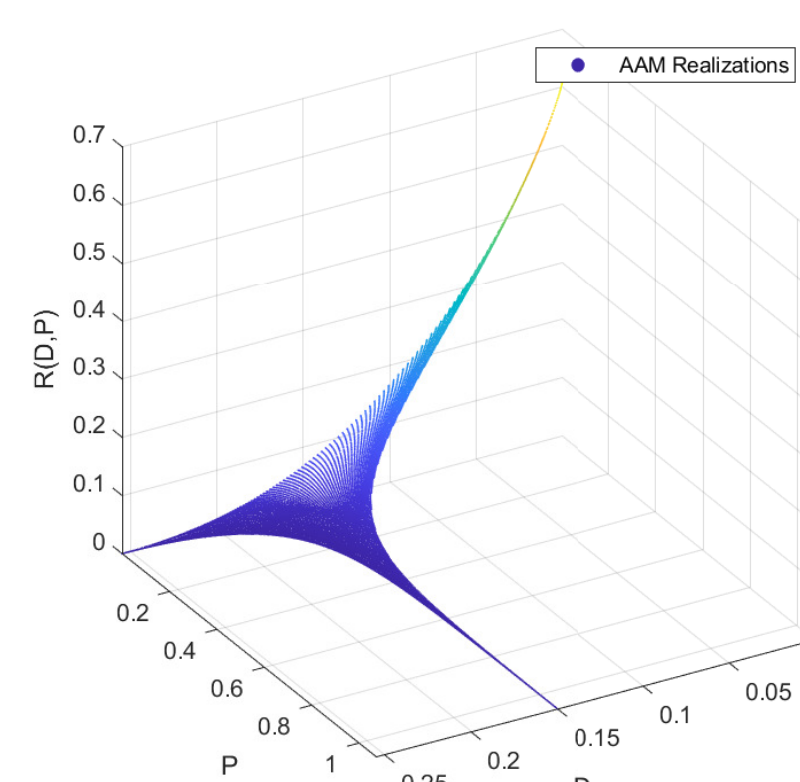
There exist sufficient conditions to ensure the linear convergence of the proposed schemes:

- for the OAM scheme, it suffices that the distortion metric $d: \mathcal{X} \times \hat{\mathcal{X}} \rightarrow \mathcal{R}_0^+$ induces a full-rank matrix $K = [\exp\{-s_1 d(i, j)\}]_{(i, j) \in \mathcal{X} \times \hat{\mathcal{X}}}$.
- for the AAM scheme, in addition to the previous condition, we need to constraint the set of allowable Lagrangian multipliers (s_1, s_2) .

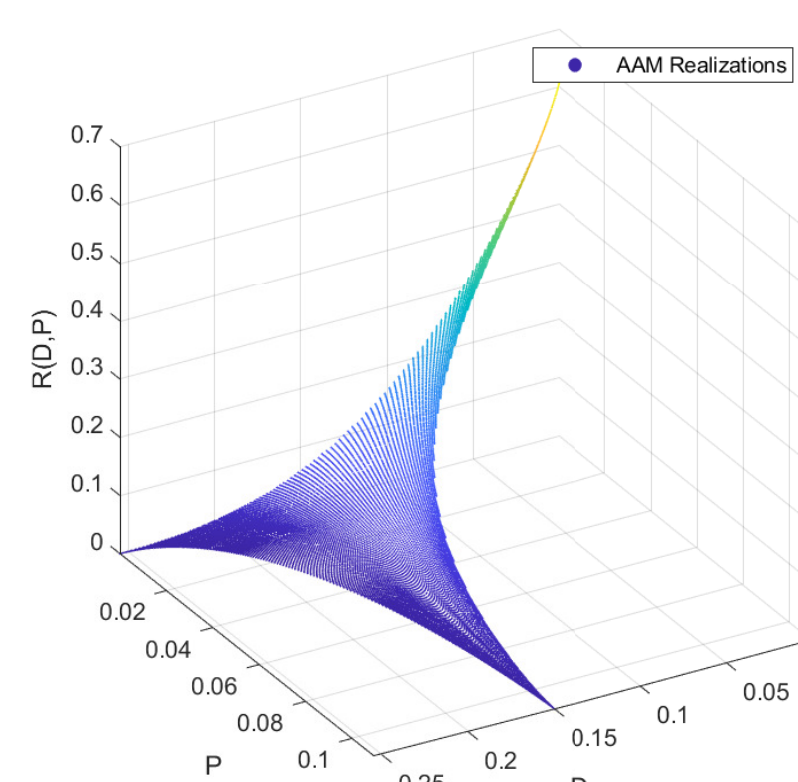
In the plots we show the convergence behavior of the AAM algorithm for Kullback-Leibler divergence (left) and Jensen-Shannon divergence (right).

NUMERICAL EXAMPLES

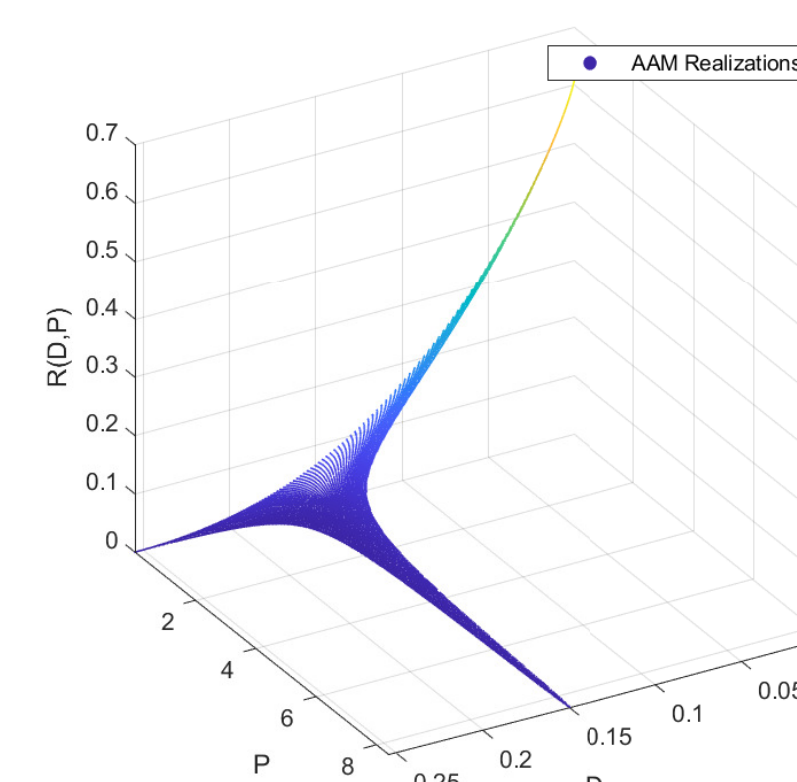
We consider a binary source alphabet $\{0, 1\}$ and a source $X \sim \text{Bern}(0.15)$. The distortion measure is the Hamming distortion $d_H(i, j) = \delta_{i \neq j}$ and the perception measure is selected among commonly used f -divergences.



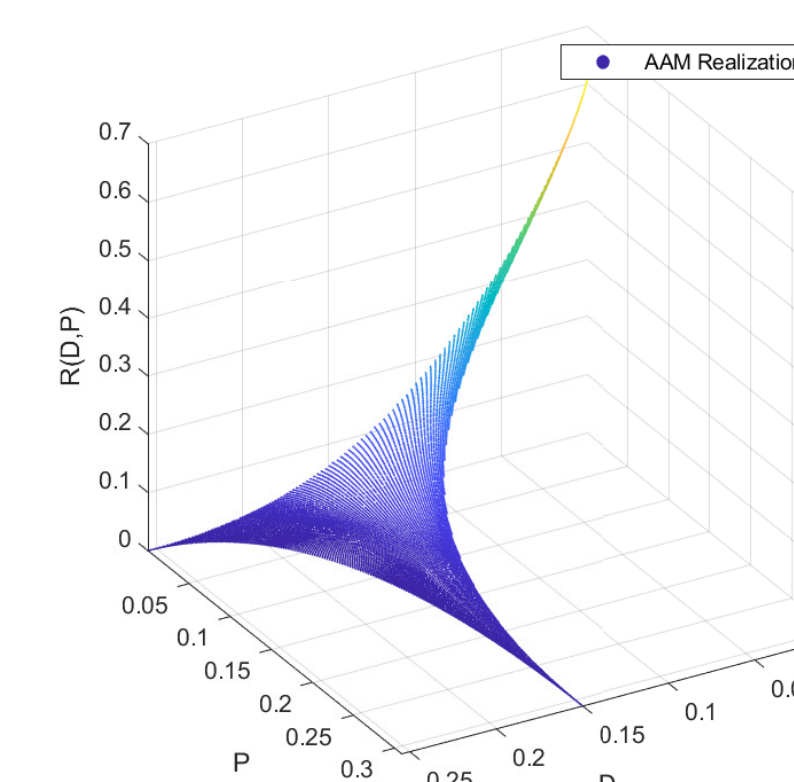
KL divergence



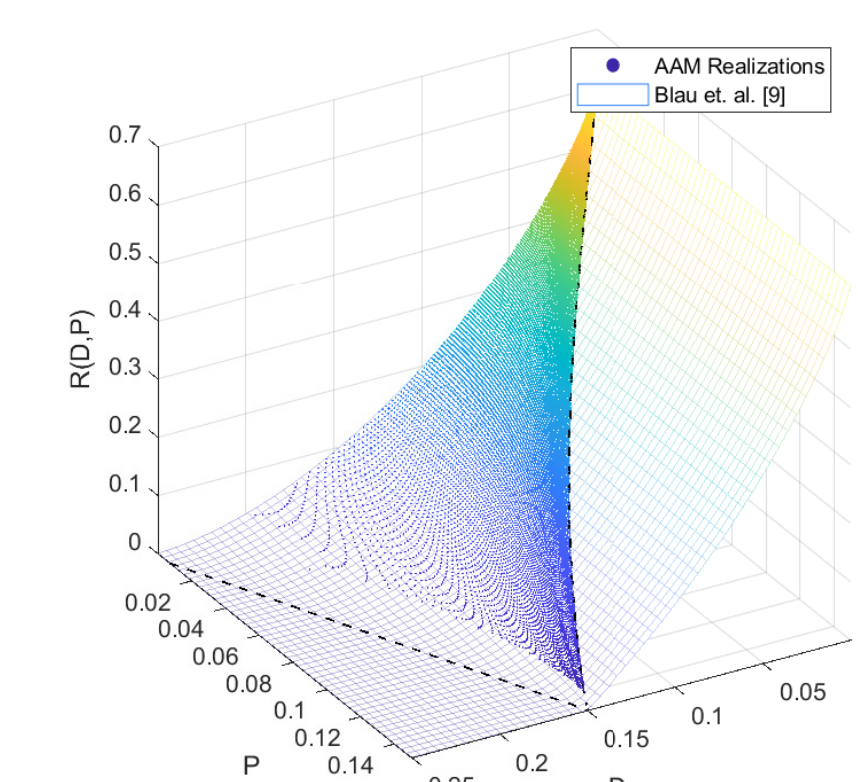
JS divergence



χ^2 divergence



α divergence ($\alpha = \frac{1}{2}$)



TV divergence