

Discrete Universal Denoiser

In [information theory](#) and [signal processing](#), the **Discrete Universal Denoiser** (DUDE) is a [denoising](#) scheme for recovering sequences over a finite alphabet, which have been corrupted by a [discrete memoryless channel](#). The DUDE was proposed in 2005 by Tsachy Weissman, Erik Ordentlich, Gadiel Seroussi, Sergio Verdú and Marcelo J. Weinberger.^[1]

Contents

- 1 **Overview**
- 2 **The discrete denoising problem**
- 3 **The DUDE scheme**
 - 3.1 Step 1: Calculating the empirical distribution in each context
 - 3.2 Step 2: Calculating the Bayes response to each context
 - 3.3 Step 3: Estimating each symbol by the Bayes response to its context
- 4 **Asymptotic optimality properties**
 - 4.1 For a stationary source
 - 4.2 For an individual sequence
 - 4.3 Non-asymptotic performance
- 5 **Background**
- 6 **Extensions**
- 7 **Applications**
 - 7.1 Application to image denoising
 - 7.2 Application to channel decoding of uncompressed sources
- 8 **References**

Overview

The Discrete Universal Denoiser^[1] (DUDE) is a [denoising](#) scheme that estimates an unknown signal $\mathbf{x}^n = (x_1 \dots x_n)$ over a finite alphabet from a noisy version $\mathbf{z}^n = (z_1 \dots z_n)$. While most [denoising](#) schemes in the signal processing and statistics literature deal with [signals](#) over an infinite alphabet (notably, real-valued signals), the DUDE addresses the finite alphabet case. The noisy version \mathbf{z}^n is assumed to be generated by transmitting \mathbf{x}^n through a known [discrete memoryless channel](#).

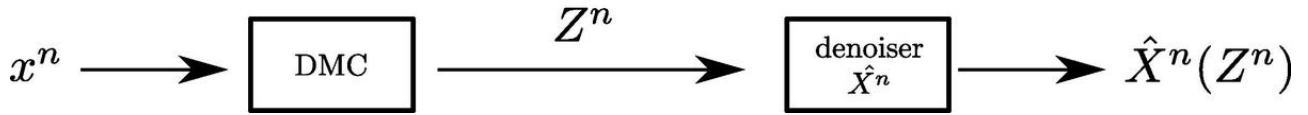
For a fixed *context length* parameter k , the DUDE counts the occurrences of all the strings of length $2k + 1$ appearing in \mathbf{z}^n . The estimated value \hat{x}_i is determined based the two-sided length- k context $(z_{i-k}, \dots, z_{i-1}, z_{i+1}, \dots, z_{i+k})$ of z_i , taking into account all the other tokens in \mathbf{z}^n with the same context, as well as the known channel matrix and the loss function being used.

The idea underlying the DUDE is best illustrated when \mathbf{x}^n is a realization of a random vector \mathbf{X}^n . If the conditional distribution $\mathbf{X}_i | \mathbf{Z}_{i-k}, \dots, \mathbf{Z}_{i-1}, \mathbf{Z}_{i+1}, \dots, \mathbf{Z}_{i+k}$, namely the distribution of the noiseless symbol \mathbf{X}_i conditional on its noisy context $(\mathbf{Z}_{i-k}, \dots, \mathbf{Z}_{i-1}, \mathbf{Z}_{i+1}, \dots, \mathbf{Z}_{i+k})$ was available, the optimal estimator $\hat{\mathbf{X}}_i$ would be the Bayes Response to $\mathbf{X}_i | \mathbf{Z}_{i-k}, \dots, \mathbf{Z}_{i-1}, \mathbf{Z}_{i+1}, \dots, \mathbf{Z}_{i+k}$. Fortunately, when the channel matrix is known and non-degenerate, this conditional distribution can be expressed in terms of the conditional distribution $\mathbf{Z}_i | \mathbf{Z}_{i-k}, \dots, \mathbf{Z}_{i-1}, \mathbf{Z}_{i+1}, \dots, \mathbf{Z}_{i+k}$, namely the distribution of the noisy symbol \mathbf{Z}_i conditional on its noisy context. This conditional distribution, in turn, can be estimated from an individual observed noisy signal \mathbf{Z}^n by virtue of the Law of Large Numbers, provided n is “large enough”.

Applying the DUDE scheme with a context length k to a sequence of length n over a finite alphabet \mathcal{Z} requires $O(n)$ operations and space $O(\min(n, |\mathcal{Z}|^{2k}))$.

Under certain assumptions, the DUDE is a universal scheme in the sense of asymptotically performing as well as an optimal denoiser, which has oracle access to the unknown sequence. More specifically, assume that the denoising performance is measured using a given single-character fidelity criterion, and consider the regime where the sequence length n tends to infinity and the context length $k = k_n$ tends to infinity “not too fast”. In the stochastic setting, where a doubly infinite sequence noiseless sequence \mathbf{x} is a realization of a stationary process \mathbf{X} , the DUDE asymptotically performs, in expectation, as well as the best denoiser, which has oracle access to the source distribution \mathbf{X} . In the single-sequence, or “semi-stochastic” setting with a *fixed* doubly infinite sequence \mathbf{x} , the DUDE asymptotically performs as well as the best “sliding window” denoiser, namely any denoiser that determines $\hat{\mathbf{x}}_i$ from the window $(z_{i-k}, \dots, z_{i+k})$, which has oracle access to \mathbf{x} .

The discrete denoising problem



Block diagram description of the discrete denoising problem

Let \mathcal{X} be the finite alphabet of a fixed but unknown original “noiseless” sequence $\mathbf{x}^n = (x_1, \dots, x_n) \in \mathcal{X}^n$. The sequence is fed into a discrete memoryless channel (DMC). The DMC operates on each symbol \mathbf{x}_i independently, producing a corresponding random symbol \mathbf{Z}_i in a finite alphabet \mathcal{Z} . The DMC is known and given as a \mathcal{X} -by- \mathcal{Z} Markov matrix $\mathbf{\Pi}$, whose entries are $\pi(\mathbf{x}, z) = \mathbb{P}(Z = z | X = \mathbf{x})$. It is convenient to write π_z for the z -column of $\mathbf{\Pi}$. The DMC produces a random noisy sequence $\mathbf{Z}^n = (z_1, \dots, z_n) \in \mathcal{Z}^n$. A specific realization of this random vector will be denoted by \mathbf{z}^n . A denoiser is a function $\hat{\mathbf{X}}^n : \mathcal{Z}^n \rightarrow \mathcal{X}^n$ that attempts to recover the noiseless sequence \mathbf{x}^n from a distorted version \mathbf{z}^n . A specific denoised sequence is denoted by $\hat{\mathbf{x}}^n = \hat{\mathbf{X}}^n(\mathbf{z}^n) = (\hat{x}_1(\mathbf{z}^n), \dots, \hat{x}_n(\mathbf{z}^n))$. The problem of choosing the denoiser $\hat{\mathbf{X}}^n$ is known as signal estimation, filtering or smoothing. To compare candidate denoisers, we choose a single-symbol fidelity criterion $\Lambda : \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty)$ (for example, the Hamming loss) and define the per-symbol loss of the denoiser $\hat{\mathbf{X}}^n$ at $(\mathbf{x}^n, \mathbf{z}^n)$ by

$$L_{\hat{\mathbf{X}}^n}(\mathbf{x}^n, \mathbf{z}^n) = \frac{1}{n} \sum_{i=1}^n \Lambda(x_i, \hat{x}_i(\mathbf{z}^n)).$$

Ordering the elements of the alphabet \mathcal{X} by $\mathcal{X} = (a_1, \dots, a_{|\mathcal{X}|})$, the fidelity criterion can be given by a $|\mathcal{X}|$ -by- $|\mathcal{X}|$ matrix, with columns of the form

$$\lambda_{\hat{x}} = \begin{pmatrix} \Lambda(a_1, \hat{x}) \\ \vdots \\ \Lambda(a_{|\mathcal{X}|}, \hat{x}) \end{pmatrix}.$$

The DUDE scheme

Step 1: Calculating the empirical distribution in each context

The DUDE corrects symbols according to their context. The context length k used is a tuning parameter of the scheme. For $k+1 \leq i \leq n-k$, define the left context of the i -th symbol in z^n by $l^k(z^n, i) = (z_{i-k}, \dots, z_{i-1})$ and the corresponding right context as $r^k(z^n, i) = (z_{i+1}, \dots, z_{i+k})$. A two-sided context is a combination (l^k, r^k) of a left and a right context.

The first step of the DUDE scheme is to calculate the empirical distribution of symbols in each possible two-sided context along the noisy sequence z^n . Formally, a given two-sided context $(l^k, r^k) \in \mathcal{Z}^k \times \mathcal{Z}^k$ that appears once or more along z^n determines an empirical probability distribution over \mathcal{Z} , whose value at the symbol z is

$$\mu(z^n, l^k, r^k)[z] = \frac{|\{k+1 \leq i \leq n-k \mid (z_{i-k}, \dots, z_{i+k}) = l^k z r^k\}|}{|\{k+1 \leq i \leq n-k \mid l^k(z^n, i) = l^k \text{ and } r^k(z^n, i) = r^k\}|}.$$

Thus, the first step of the DUDE scheme with context length k is to scan the input noisy sequence z^n once, and store the length- $|\mathcal{Z}|$ empirical distribution vector $\mu(z^n, l^k, r^k)$ (or its non-normalized version, the count vector) for each two-sided context found along z^n . Since there are at most $N_{n,k} = \min(n, |\mathcal{Z}|^{2k})$ possible two-sided contexts along z^n , this step requires $O(n)$ operations and storage $O(N_{n,k})$.

Step 2: Calculating the Bayes response to each context

Denote the column of single-symbol fidelity criterion Λ , corresponding to the symbol $\hat{x} \in \mathcal{X}$, by $\lambda_{\hat{x}}$. We define the *Bayes Response* to any vector \mathbf{v} of length $|\mathcal{X}|$ with non-negative entries as

$$\hat{X}_{\text{Bayes}}(\mathbf{v}) = \operatorname{argmin}_{\hat{x} \in \mathcal{X}} \lambda_{\hat{x}}^\top \mathbf{v}.$$

This definition is motivated in the [background](#) below.

The second step of the DUDE scheme is to calculate, for each two-sided context (l^k, r^k) observed in the previous step along z^n , and for each symbol $z \in \mathcal{Z}$ observed in each context (namely, any z such that $l^k z r^k$ is a substring of z^n) the Bayes response to the vector $\Pi^{-\top} \mu(z^n, l^k, r^k) \odot \pi_z$, namely

$$g(l^k, z, r^k) := \hat{X}_{\text{Bayes}}(\Pi^{-\top} \mu(z^n, l^k, r^k) \odot \pi_z).$$

Note that the sequence \mathbf{z}^n and the context length k are implicit. Here, π_z is the z -column of $\mathbf{\Pi}$ and for vectors \mathbf{a} and \mathbf{b} , $\mathbf{a} \odot \mathbf{b}$ denotes their Schur (entrywise) product, defined by $(\mathbf{a} \odot \mathbf{b})_i = a_i b_i$. Matrix multiplication is evaluated before the Schur product, so that $\mathbf{\Pi}^{-\top} \mu \odot \pi_z$ stands for $(\mathbf{\Pi}^{-\top} \mu) \odot \pi_z$.

This formula assumed that the channel matrix $\mathbf{\Pi}$ is square ($|\mathcal{X}| = |\mathcal{Z}|$) and invertible. When $|\mathcal{X}| \leq |\mathcal{Z}|$ and $\mathbf{\Pi}$ is not invertible, under the reasonable assumption that it has full row rank, we replace $(\mathbf{\Pi}^\top)^{-1}$ above with its Moore-Penrose pseudo-inverse $(\mathbf{\Pi}\mathbf{\Pi}^\top)^{-1}\mathbf{\Pi}$ and calculate instead

$$g(l^k, z, r^k) := \hat{X}_{\text{Bayes}}((\mathbf{\Pi}\mathbf{\Pi}^\top)^{-1}\mathbf{\Pi}\mu(z^n, l^k, r^k) \odot \pi_z) .$$

By caching the inverse or pseudo-inverse $\mathbf{\Pi}^{-\top}$, and the values $\lambda_{\hat{\mathbf{x}}} \odot \pi_z$ for the relevant pairs $(\hat{\mathbf{x}}, z) \in \mathcal{X} \times \mathcal{Z}$, this step requires $O(N_{k,n})$ operations and $O(N_{k,n})$ storage.

Step 3: Estimating each symbol by the Bayes response to its context

The third and final step of the DUDE scheme is to scan \mathbf{z}^n again and compute the actual denoised sequence $\hat{\mathbf{X}}^n(z^n) = (\hat{X}_1(z^n), \dots, \hat{X}_n(z^n))$. The denoised symbol chosen to replace z_i is the Bayes response to the two-sided context of the symbol, namely

$$\hat{X}_i(z^n) := g(l^k(z^n, i), z_i, r^k(z^n, i)) .$$

This step requires $O(n)$ operations and used the data structure constructed in the previous step.

In summary, the entire DUDE requires $O(n)$ operations and $O(N_{k,n})$ storage.

Asymptotic optimality properties

The DUDE is designed to be universally optimal, namely optimal (in some sense, under some assumptions) regardless of the original sequence \mathbf{x}^n .

Let $\hat{X}_{DUDE}^n : \mathcal{Z}^n \rightarrow \mathcal{X}^n$ denote a sequence of DUDE schemes, as described above, where \hat{X}_{DUDE}^n uses a context length k_n that is implicit in the notation. We only require that $\lim_{n \rightarrow \infty} k_n = \infty$ and that $k_n |\mathcal{Z}|^{2K_n} = o\left(\frac{n}{\log n}\right)$.

For a stationary source

Denote by \mathcal{D}_n the set of all n -block denoisers, namely all maps $\hat{X}^n : \mathcal{Z}^n \rightarrow \mathcal{X}^n$.

Let \mathbf{X} be an unknown stationary source and \mathbf{Z} be the distribution of the corresponding noisy sequence. Then

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[L_{\hat{X}_{DUDE}^n}(\mathbf{X}^n, \mathbf{Z}^n) \right] = \lim_{n \rightarrow \infty} \min_{\hat{X}^n \in \mathcal{D}_n} \mathbf{E} \left[L_{\hat{X}^n}(\mathbf{X}^n, \mathbf{Z}^n) \right] ,$$

and both limits exist. If, in addition the source \mathbf{X} is ergodic, then

$$\limsup_{n \rightarrow \infty} L_{\hat{X}_{DUDE}^n}(X^n, Z^n) = \lim_{n \rightarrow \infty} \min_{\hat{X}^n \in \mathcal{D}_n} \mathbf{E} \left[L_{\hat{X}^n}(X^n, Z^n) \right], \text{ almost surely.}$$

For an individual sequence

Denote by $\mathcal{D}_{n,k}$ the set of all n -block k -th order sliding window denoisers, namely all maps $\hat{X}^n : \mathcal{Z} \rightarrow \mathcal{X}$ of the form $\hat{X}_i(z^n) = f(z_{i-k}, \dots, z_{i+k})$ with $f : \mathcal{Z}^{2k+1} \rightarrow \mathcal{X}$ arbitrary.

Let $\mathbf{x} \in \mathcal{X}^\infty$ be an unknown noiseless sequence stationary source and \mathbf{Z} be the distribution of the corresponding noisy sequence. Then

$$\lim_{n \rightarrow \infty} \left[L_{\hat{X}_{DUDE}^n}(x^n, Z^n) - \min_{\hat{X}^n \in \mathcal{D}_{n,k}} L_{\hat{X}^n}(x^n, Z^n) \right] = 0, \text{ almost surely.}$$

Non-asymptotic performance

Let \hat{X}_k^n denote the DUDE on with context length k defined on n -blocks. Then there exist explicit constants $A, C > 0$ and $B > 1$ that depend on (Π, Λ) alone, such that for any n, k and any $x^n \in \mathcal{X}^n$ we have

$$\frac{A}{\sqrt{n}} B^k \leq \mathbf{E} \left[L_{\hat{X}_k^n}(x^n, Z^n) - \min_{\hat{X}^n \in \mathcal{D}_{n,k}} L_{\hat{X}^n}(x^n, Z^n) \right] \leq \sqrt{k} \frac{C}{\sqrt{n}} |Z|^k,$$

where Z^n is the noisy sequence corresponding to x^n (whose randomness is due to the channel alone) [2].

In fact holds with the same constants A, B as above for *any* n -block denoiser $\hat{X}^n \in \mathcal{D}^n$.^[1] The lower bound proof requires that the channel matrix Π be square and the pair (Π, Λ) satisfies a certain technical condition.

Background

To motivate the particular definition of the DUDE using the Bayes response to a particular vector, we now find the optimal denoiser in the non-universal case, where the unknown sequence x^n is a realization of a random vector X^n , whose distribution is known.

Consider first the case $n = 1$. Since the joint distribution of (X, Z) is known, given the observed noisy symbol z , the unknown symbol $X \in \mathcal{X}$ is distributed according to the known distribution $\mathbb{P}(X = x | Z = z)$. By ordering the elements of \mathcal{X} , we can describe this conditional distribution on \mathcal{X} using a probability vector $\mathbf{P}_{X|z}$, indexed by \mathcal{X} , whose x -entry is $\mathbb{P}(X = x | Z = z)$. Clearly the expected loss for the choice of estimated symbol \hat{x} is $\lambda_{\hat{x}}^\top \mathbf{P}_{X|z}$.

Define the *Bayes Envelope* of a probability vector \mathbf{v} , describing a probability distribution on \mathcal{X} , as the minimal expected loss $U(\mathbf{v}) = \min_{\hat{x} \in \mathcal{X}} \mathbf{v}^\top \lambda_{\hat{x}}$, and the *Bayes Response* to \mathbf{v} as the prediction that achieves this minimum,

$\hat{X}_{\text{Bayes}}(\mathbf{v}) = \operatorname{argmin}_{\hat{x} \in \mathcal{X}} \mathbf{v}^\top \lambda_{\hat{x}}$. Observe that the Bayes response is scale invariant in the sense that $\hat{X}_{\text{Bayes}}(\mathbf{v}) = \hat{X}_{\text{Bayes}}(\alpha \mathbf{v})$ for $\alpha > 0$.

For the case $n = 1$, then, the optimal denoiser is $\hat{X}(z) = \hat{X}_{Bayes}(\mathbf{P}_{X|z})$. This optimal denoiser can be expressed using the marginal distribution of \mathbf{Z} alone, as follows. When the channel matrix $\mathbf{\Pi}$ is invertible, we have $\mathbf{P}_{X|z} \propto \mathbf{\Pi}^{-\top} \mathbf{P}_Z \odot \pi_z$ where π_z is the z -th column of $\mathbf{\Pi}$. This implies that the optimal denoiser is given equivalently by $\hat{X}(z) = \hat{X}_{Bayes}(\mathbf{\Pi}^{-\top} \mathbf{P}_Z \odot \pi_z)$. When $|\mathcal{X}| \leq |\mathcal{Z}|$ and $\mathbf{\Pi}$ is not invertible, under the reasonable assumption that it has full row rank, we can replace $\mathbf{\Pi}^{-1}$ with its Moore-Penrose pseudo-inverse and obtain

$$\hat{X}(z) = \hat{X}_{Bayes}((\mathbf{\Pi}\mathbf{\Pi}^\top)^{-1}\mathbf{\Pi}\mathbf{P}_Z \odot \pi_z) .$$

Turning now to arbitrary n , the optimal denoiser $\hat{X}^{opt}(z^n)$ (with minimal expected loss) is therefore given by the Bayes response to $\mathbf{P}_{X_i|z^n}$

$$\hat{X}_i^{opt}(z^n) = \hat{X}_{Bayes} \mathbf{P}_{X_i|z^n} = \operatorname{argmin}_{\hat{x} \in \mathcal{X}} \lambda_{\hat{x}}^\top \mathbf{P}_{X_i|z^n} ,$$

where $\mathbf{P}_{X_i|z^n}$ is a vector indexed by \mathcal{X} , whose x -entry is $\mathbb{P}(X_i = x | Z^n = z^n)$. The conditional probability vector $\mathbf{P}_{X_i|z^n}$ is hard to compute. A derivation analogous to the case $n = 1$ above shows that the optimal denoiser admits an alternative representation, namely $\hat{X}_i^{opt}(z^n) = \hat{X}_{Bayes}(\mathbf{\Pi}^{-\top} \mathbf{P}_{Z_i, z^{n \setminus i}} \odot \pi_{z_i})$, where $z^{n \setminus i} = (z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n) \in \mathcal{Z}^{n-1}$ is a given vector and $\mathbf{P}_{Z_i, z^{n \setminus i}}$ is the probability vector indexed by \mathcal{Z} whose z -entry is $\mathbb{P}((Z_1, \dots, Z_n) = (z_1, \dots, z_{i-1}, z, z_{i+1}, \dots, z_n))$. Again, $\mathbf{\Pi}^{-\top}$ is replaced by a pseudo-inverse if $\mathbf{\Pi}$ is not square or not invertible.

When the distribution of \mathbf{X} (and therefore, of \mathbf{Z}) is not available, the DUDE replaces the unknown vector $\mathbf{P}_{Z_i, z^{n \setminus i}}$ with an empirical estimate obtained along the noisy sequence z^n itself, namely with $\mu(Z_i, l^k(Z^n, i), r^k(Z^n, i))$. This leads to the above definition of the DUDE.

While the convergence arguments behind the optimality properties above are more subtle, we note that the above, combined with the Birkhoff Ergodic Theorem, is enough to prove that for a stationary ergodic source, the DUDE with context-length k is asymptotically optimal all k -th order sliding window denoisers.

Extensions

The basic DUDE as described here assumes a signal with a one-dimensional index set over a finite alphabet, a known memoryless channel and a context length that is fixed in advance. Relaxations of each of these assumptions have been considered in turn.^[3] Specifically:

- Infinite alphabets^{[4][5][6][7]}
- Channels with memory^{[8][9]}
- Unknown channel matrix^{[10][11]}
- Variable context and adaptive choice of context length^{[12][13][14][15]}
- Two-dimensional signals^[16]

Applications

Application to image denoising

A DUDE-based framework for grayscale image denoising^[6] achieves state-of-the-art denoising for impulse-type noise channels (e.g., "salt and pepper" or "M-ary symmetric" noise), and good performance on the Gaussian channel (comparable to the Non-local means image denoising scheme on this channel). A different DUDE variant applicable to grayscale images is presented in.^[7]

Application to channel decoding of uncompressed sources

The DUDE has led to universal algorithms for channel decoding of uncompressed sources.^[17]

References

1. T. Weissman, E. Ordentlich, G. Seroussi, S. Verdú, and M.J. Weinberger. Universal discrete denoising: Known channel. *IEEE Transactions on Information Theory*, 51(1):5–28, 2005.
2. K. Viswanathan and E. Ordentlich. Lower limits of discrete universal denoising. *IEEE Transactions on Information Theory*, 55(3):1374–1386, 2009.
3. Ordentlich, E.; Seroussi, G.; Verdú; Weinberger, M. J.; Weissman, T. "Reflections on the DUDE" (http://www.stanford.edu/~tsachy/pdf_files/Reflections%20on%20the%20DUDE.pdf) (pdf).
4. A. Dembo and T. Weissman. Universal denoising for the finite-input-general-output channel. *IEEE Trans. Inform. Theory*, 51(4):1507–1517, April 2005.
5. K. Sivaramakrishnan and T. Weissman. Universal denoising of discrete-time continuous amplitude signals. In *Proc. of the 2006 IEEE Intl. Symp. on Inform. Theory, (ISIT'06)*, Seattle, WA, USA, July 2006.
6. G. Motta, E. Ordentlich, I. Ramírez, G. Seroussi, and M. Weinberger, "The DUDE framework for continuous tone image denoising," *IEEE Transactions on Image Processing*, 20, No. 1, January 2011.
7. K. Sivaramakrishnan and T. Weissman. Universal denoising of continuous amplitude signals with applications to images. In *Proc. of IEEE International Conference on Image Processing, Atlanta, GA, USA, October 2006*, pp. 2609–2612
8. C. D. Giurcaneanu and B. Yu. Efficient algorithms for discrete universal denoising for channels with memory. In *Proc. of the 2005 IEEE Intl. Symp. on Inform. Theory, (ISIT'05)*, Adelaide, Australia, Sept. 2005.
9. R. Zhang and T. Weissman. Discrete denoising for channels with memory. *Communications in Information and Systems (CIS)*, 5(2):257–288, 2005.
10. G. M. Gemelos, S. Sigurjonsson, T. Weissman. Universal minimax discrete denoising under channel uncertainty. *IEEE Trans. Inform. Theory*, 52:3476–3497, 2006.
11. G. M. Gemelos, S. Sigurjonsson and T. Weissman. Algorithms for discrete denoising under channel uncertainty. *IEEE Trans. Signal Processing*, 54(6):2263–2276, June 2006.
12. E. Ordentlich, M.J. Weinberger, and T. Weissman. Multi-directional context sets with applications to universal denoising and compression. In *Proc. of the 2005 IEEE Intl. Symp. on Inform. Theory, (ISIT'05)*, Adelaide, Australia, Sept. 2005.
13. J. Yu and S. Verdú. Schemes for bidirectional modeling of discrete stationary sources. *IEEE Trans. Inform. Theory*, 52(11):4789–4807, 2006.
14. S. Chen, S. N. Diggavi, S. Dusad and S. Muthukrishnan. Efficient string matching algorithms for combinatorial universal denoising. In *Proc. of IEEE Data Compression Conference (DCC)*, Snowbird, Utah, March 2005.
15. G. Gimel'farb. Adaptive context for a discrete universal denoiser. In *Proc. Structural, Syntactic, and Statistical Pattern Recognition, Joint IAPR International Workshops, SSPR 2004 and SPR 2004, Lisbon, Portugal, August 18–20, pp. 477–485*

16. E. Ordentlich, G. Seroussi, S. Verdú, M.J. Weinberger, and T. Weissman. A universal discrete image denoiser and its application to binary images. In Proc. IEEE International Conference on Image Processing, Barcelona, Catalonia, Spain, September 2003.
 17. E. Ordentlich, G. Seroussi, S. Verdú, and K. Viswanathan, "Universal Algorithms for Channel Decoding of Uncompressed Sources," IEEE Trans. Information Theory, vol. 54, no. 5, pp. 2243–2262, May 2008
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Discrete_Universal_Denoiser&oldid=744364913"

This page was last edited on 14 October 2016, at 19:04.

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.