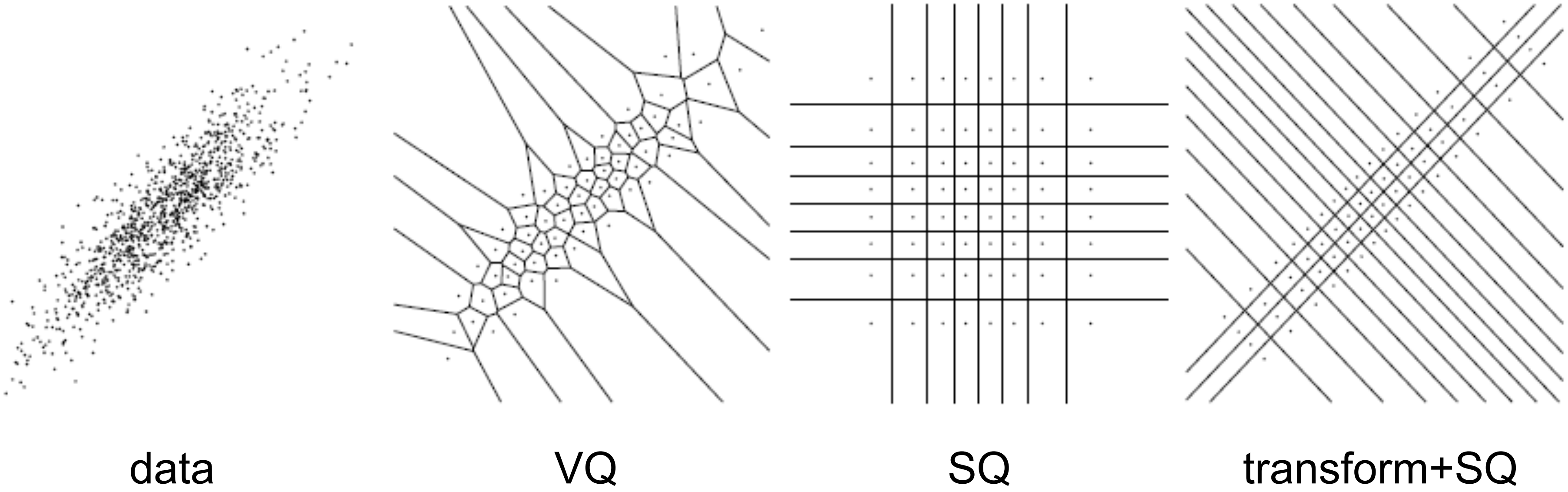
Transform Coding

- Coding sources with memory
- Orthonormal transforms
- Rate distortion function for transform coding
- Transform coding gain
- Maximizing the transform coding gain
- Karhunen-Loeve Transform (KLT)
- Energy concentration by KLT

Coding Sources with Memory

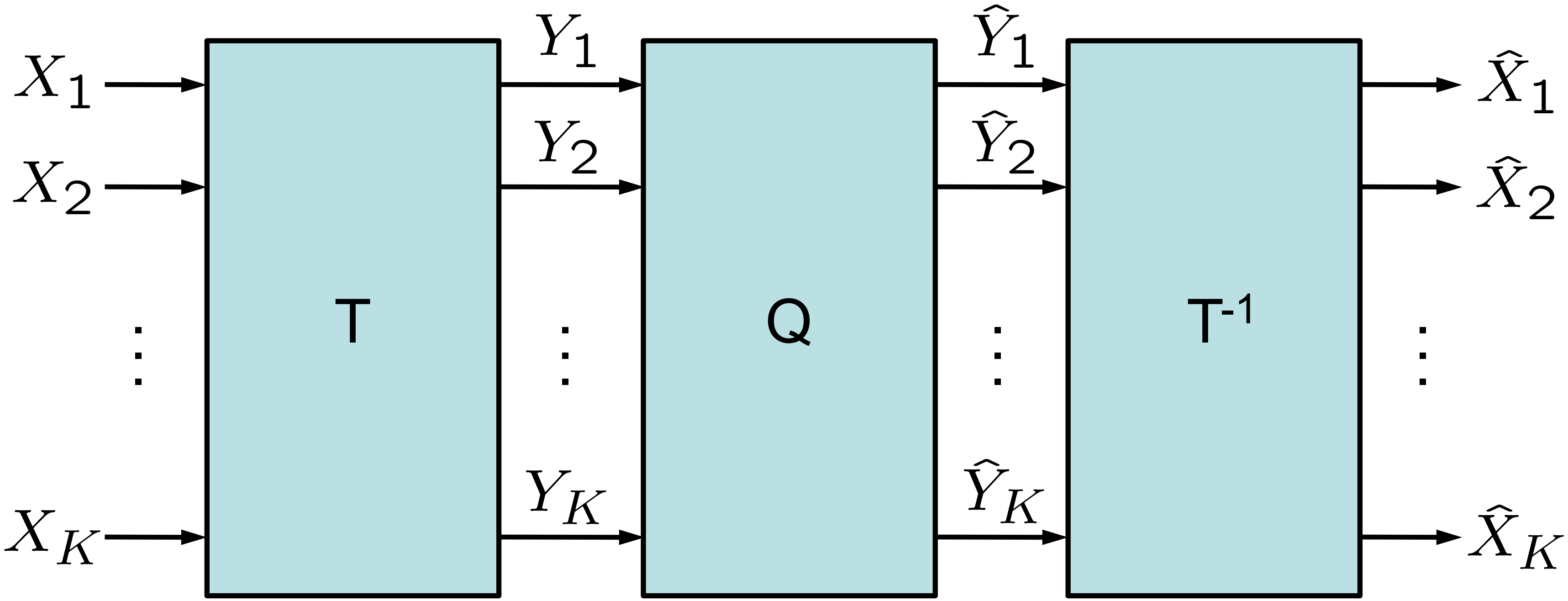
- Example: Two-dimensional source



- Exploit memory of the source by linear transform

Using Transforms for Coding Vector Sources

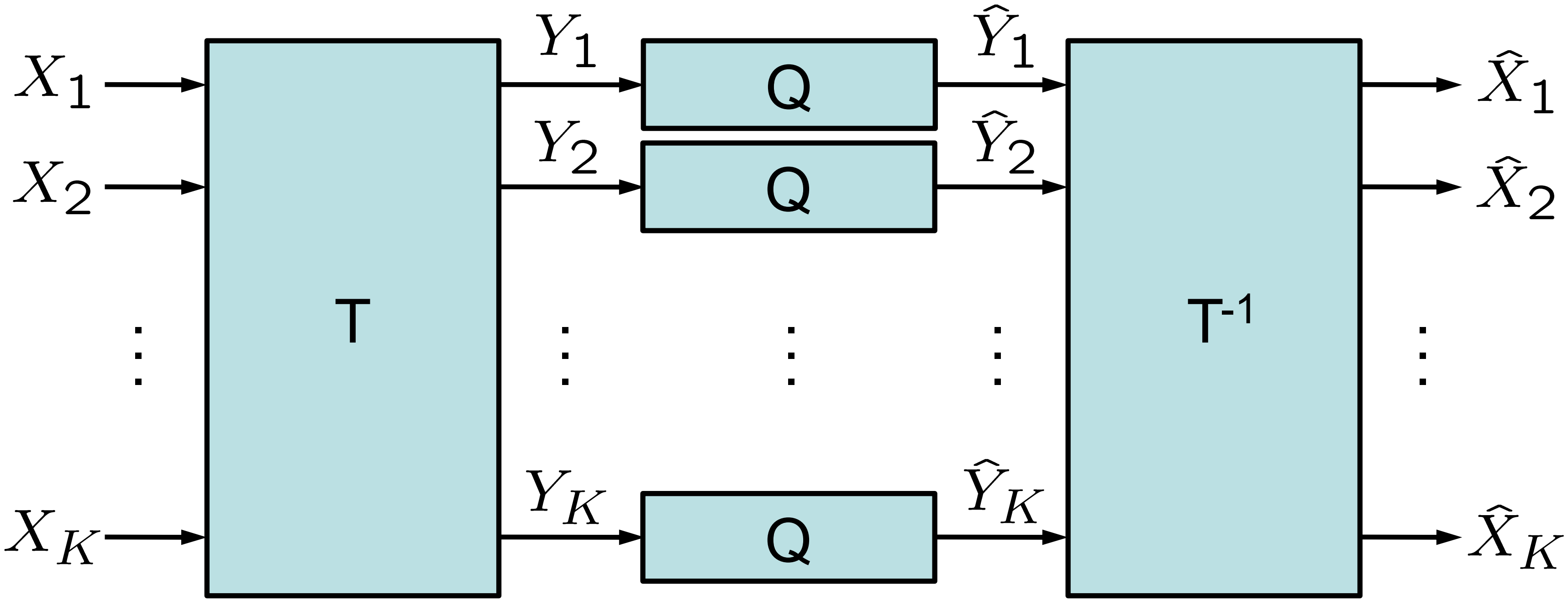
- K-dimensional vector source



- How to choose the transform *T*?

Independent Transform Coefficients

- Let the transform generate independent coefficients



- Further, use independent quantizers (no sphere packing)

Orthonormal Transforms

- Consider a transform T that produces independent transform coefficients
- The synthesis transform exists and is given by T^{-1}
- The component quantizers reconstruct to their respective centroids
- **Theorem:** To minimize the mean square error distortion, it is sufficient to consider transforms that satisfy $TT^T = I$. *[Goyal, 2001]*

Energy Conservation

- For any orthonormal transform *y*= *Tx* we obtain

$$||y||_2^2 = y^T y = x^T T^T T x = ||x||_2^2$$

- Vector lengths („energies“) are conserved.
- Interpretation: every orthonormal transform is simply a rotation of the coordinate system.

Energy Distribution for Orthonormal Transform

- Energy is conserved, but typically will be unevenly distributed among coefficients.
- Autocorrelation matrix

$$C_{yy} = E\{yy^T\} = E\{Txx^T T^T\} = TC_{xx}T^T$$

- Mean squared values („average energies") of the coefficients y_i are on the diagonal of C_{yy}

Distortion Rate Function for Transform Coding

- Component distortion at high rates

$$D_n(R_n) = \epsilon^2 \sigma_{y_n}^2 2^{-2R_n} \quad \text{for} \quad D_n \leq \epsilon^2 \sigma_{y_n}^2$$

- Optimum rate allocation

$$\min \frac{1}{N} \sum_{n=1}^{N} D_n(R_n) \quad \text{s.t.} \quad R = \frac{1}{N} \sum_{n=1}^{N} R_n$$

- Optimum rate allocation at high rates

$$D_n \leq \epsilon^2 \sigma_{y_n}^2 \quad : \quad \frac{\partial (D + \lambda R)}{\partial R_i} = 0 \quad \Rightarrow \quad \lambda = 2 \ln(2) D_i$$

$$D_i = D_j \quad \forall \quad i, j$$

Distortion Rate Function for Transform Coding

- Component rate at high rates

$$R_n = \frac{1}{2}\log_2\left(\frac{\epsilon^2\sigma_{y_n}^2}{D_n}\right) \quad \text{for} \quad D_n \le \epsilon^2\sigma_{y_n}^2$$

- Total rate at high rates

$$R = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{2}\log_2\left(\frac{\epsilon^2\sigma_{y_n}^2}{D}\right) = \frac{1}{2}\log_2\left(\frac{\epsilon^2\left[\prod_{n=1}^{N}\sigma_{y_n}^2\right]^{\frac{1}{N}}}{D}\right)$$

- Distortion rate function for transform coding

$$D(R) = \epsilon^2\left[\prod_{n=1}^{N}\sigma_{y_n}^2\right]^{\frac{1}{N}} 2^{-2R} \quad \text{for} \quad D \le \epsilon^2\sigma_{y_n}^2$$

Transform Coding Gain

- Distortion gain of transform coding with block size N over scalar quantization of source samples.
- Transform coding gain

$$G_{TC} = \frac{D_{SQ}(R)}{D_{TC}(R)}$$

$$G_{TC} = \frac{\sigma_x^2}{\left[\prod_{n=1}^{N} \sigma_{y_n}^2\right]^{\frac{1}{N}}} = \frac{\frac{1}{N}\sum_{n=1}^{N} \sigma_{y_n}^2}{\left[\prod_{n=1}^{N} \sigma_{y_n}^2\right]^{\frac{1}{N}}}$$

Remark: Hadamard Inequality

- Consider a symmetric positive-definite matrix A
- Choose B, b, and x such that $Bx = -b$

$$
\begin{aligned}
\det \begin{bmatrix} B & b \\ b^H & c \end{bmatrix} &= \det \begin{bmatrix} I & 0 \\ x^H & 1 \end{bmatrix} \begin{bmatrix} B & b \\ b^H & c \end{bmatrix} \begin{bmatrix} I & x \\ 0 & 1 \end{bmatrix} \\
&= \det \begin{bmatrix} B & 0 \\ 0 & b^H x + c \end{bmatrix} \\
&= \det(B)(b^H x + c) \\
&\leq \det(B) c
\end{aligned}
$$

- Note that $b^H x = -x^H B x$ is negative

$$
\det(A) \leq \prod_i A_{ii}
$$

Maximizing Transform Coding Gain

- Lower bound for product of coefficient variances

$$\det(C_{yy}) \leq \prod_{n=1}^{N} \sigma_{y_n}^2$$

- Maximize transform coding gain

$$\max G_{TC} \quad \Leftrightarrow \quad \min \prod_{n=1}^{N} \sigma_{y_n}^2$$

- Transform with maximum transform coding gain satisfies

$$\det(C_{yy}) = \prod_{n=1}^{N} \sigma_{y_n}^2$$

- The transform that decorrelates the signal maximizes the transform coding gain

Eigenmatrix of the Autocorrelation Matrix

Definition: Eigenmatrix Φ of the autocorrelation matrix C_{xx}

- Φ is unitary
- The columns of Φ form a set of eigenvectors of C_{xx}, i.e.,

$$C_{xx}\Phi = \Phi\Lambda$$

Λ is a diagonal matrix of eigenvalues λ_i

$$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_K \end{bmatrix}$$

- C_{xx} is symmetric nonnegative definite, hence $\lambda_i \geq 0$ for all i
- C_{xx} is normal matrix, i.e., $C_{xx}^H C_{xx} = C_{xx} C_{xx}^H$, hence unitary eigenmatrix exists

Karhunen-Loeve Transform

- Unitary transform with matrix

$$T = \Phi^H$$

  where the columns of Φ are ordered according to decreasing eigenvalues.
- Transform coefficients are pairwise uncorrelated

$$C_{yy} = T C_{xx} T^H = \Phi^H C_{xx} \Phi = \Phi^H \Phi \Lambda = \Lambda$$

- The KLT maximizes the transform coding gain.

Energy Concentration Property of KLT

- No other unitary transform packs as much energy into the first J coefficients, where J is arbitrary
- Mean squared approximation error by choosing only first J coefficients is minimized.

Optimum Energy Concentration by KLT

- To show optimum energy concentration property, consider the truncated coefficient vector

$$b = I_J y$$

- Where I_J contains ones on the first J diagonal positions, else zeros.
- Energy in first J coefficients for arbitrary transform A

$$E = tr(C_{bb}) = tr(I_J C_{yy} I_J) = tr(I_J T C_{xx} T^H I_J) = \sum_{k=1}^{J} t_k^T C_{xx} t_k^*$$

where t_k^T is the k-th row of T.

- Minimize Lagrangian cost function to enforce unit-length basis vectors

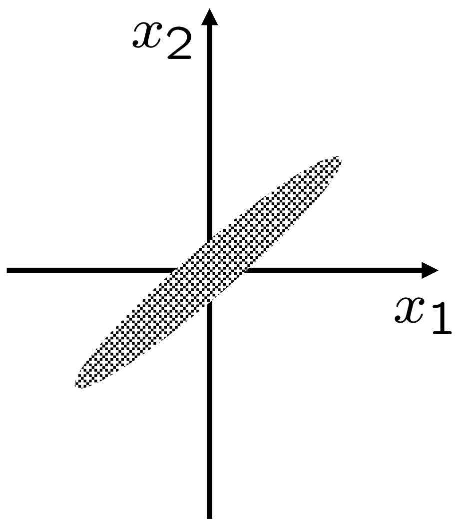
$$L = -E + \sum_{k=1}^{J} \lambda_k (t_k^T t_k^* - 1) = -\sum_{k=1}^{J} t_k^T C_{xx} t_k^* + \sum_{k=1}^{J} \lambda_k (t_k^T t_k^* - 1)$$

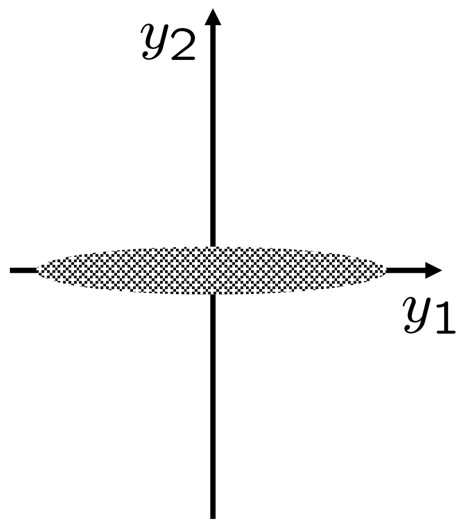
- Differentiating L with respect to t_j yields necessary condition

$$C_{xx} t_j^* = \lambda_j t_j^* \quad \forall \quad j \le J$$

Illustration of Energy Concentration

$$T = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$





Before KLT:
Strongly correlated samples, equal energies

After KLT:
Uncorrelated samples, most of the energy in first coefficient