Metric entropy in linear inverse scattering

Maria Antonia Maisto, Raffaele Solimene, Rocco Pierri

Dipartimento di Ingegneria Industriale e dell’Informazione,
Seconda Università degli studi di Napoli, Aversa (Ce), Italy
*corresponding author, E-mail: mariaantonia.maisto@unina2.it

Abstract

The role of multiple views and/or multiple frequencies on the achievable performance in linear inverse scattering problems is addressed. To this end, the impact of views and frequencies on the Kolmogorov entropy measure is studied. This way the metric information that can be conveyed back from data to the unknown can be estimated. For the sake of simplicity, the study deals with strip scatterers and the cases of discrete angles of incidence and/or frequencies.

1. Introduction

Linear inverse electromagnetic scattering problems amount to inverting an integral equation of the first kind for an object function which is related to the scatterer’s features like the shape, the support, the dielectric contrast, etc. [1]. As is well known, this problem is ill-posed in the sense of Hadamard [2]. Indeed, since the kernel function is a square integrable function, the integral operator is of Hilbert-Schimdt class and hence compact [3]. This entails that the inverse operator is not continuous and in order to obtain meaningful solutions some regularisation method must be employed [4]. Therefore, while in principle an object can be perfectly reconstructed when the operator is injective, actually due to the noise and uncertainties one must content of approximate solutions resulting from a trade-off between accuracy and stability imposed by the adopted regularisation scheme [5].

As known, the achievable performance in the reconstruction can be improved by probing the scattering scene at different angles of incidence (multi-view configuration) or at different frequencies (multi-frequency configuration). When only view or frequency diversity is exploited the cor-

mation that can be conveyed back from data to the unknown [9].

Of course, a criterion for identifying the significant singular values must be given. It is clear that this question belongs to the more general realm of regularisation theory. Here, we implicitly adopt a regularisation method that is based on a spectral cut-off strategy [10]. However, unless the singular values exhibit a step-like behaviour (and this is not the case herein) the choice of the cut-off (truncation) threshold is not so obvious. Indeed, some a priori information is usually required. Here, we follow the same philosophy as in [11], therefore it is assumed that bounds about the norm of the noise as well as of the solution are known. Accordingly, we have that

\[ \| A_i \chi - \tilde{E}_S \| \leq \epsilon, \| \chi \| \leq E \]  

where \( i \in (v, f) \) and \( \tilde{E}_S \) is in general the noisy scattered field. Of course, it is furthermore assumed that the set of object functions \( \chi \) for which both constraints hold true is not empty. By recalling that \( A_i \) are compact operators the choice of the truncation threshold can be pursued within the framework of the topological information theory in terms of the \( \epsilon \)-entropy \( H\epsilon \) and the maximum number of \( \epsilon \)- distinguishable messages \( M\epsilon \) that can be conveyed back on the unknown space. In particular, the ensemble composed by the maximum number of distinguishable messages that can be sent back to recover an approximation of the unknown solution, constitutes the backward information flow. In [11], it has been shown that the cut-off threshold must be chosen as \( N\epsilon = \max\{ n > 0 : \sigma_n(A_i) \geq \epsilon/E \}, \sigma_n(A_i) \) being the singular values. By further assuming \( E = 1 \), the following estimates have been derived

\[ H_{\epsilon} \geq \sum_{n=0}^{N\epsilon} \log_2 \sigma_n(A_i)/\epsilon \quad \text{and} \quad M_{\epsilon} \geq 2^{H_{\epsilon}} \]  

Equation (2) makes it evident that not only the number of singular values but also their magnitudes are important. In [12], the focus was addressed on how the diversities change both these attributes of the singular value behaviour. In particular, mathematical arguments that allow to analytically estimate the singular values were developed. Here, the aim is to exploit the results obtained there in order to derive an analytical expression of the right side of equation (2) in terms of the parameters of the configurations. That
is, how the $\epsilon$-entropy changes by varying of the scattering parameters is addressed.

In order to tackle the problem under the simplest conditions, a canonical configuration is considered. In more detail, the scattering domain is assumed to be a strip illuminated by plane waves whereas the scattered field is assumed to be collected in the far zone. Invariance is assumed along the plane wave polarisation direction which in turn is orthogonal to the strip. Accordingly, the study is developed for the two-dimensional scalar configuration sketched in Figure 1.

2. View diversity

Consider the geometry sketched in figure 1. Let $I = [-a; a]$ be the scattering domain which is assumed to lay along the $x$-axis and $\Omega_o = [-u_{max}; u_{max}]$ the observation angular corner with $u_o = \sin \theta_o \in \Omega_o$. Moreover, thanks to view diversity the scattered field is collected for different incidence directions $u_v = \sin \theta_v \in \Omega_v$ with $\Omega_v \subseteq \Omega_o$. Let us assume that $\Omega$, be a discrete subset of $\Omega_o$. In particular, say $M$ the number of views which are taken by uniformly sampling $\Omega_v$.

The pertinent scattering operator then writes as

$$ A_v : \chi(x) \in L^2 I \rightarrow E_S(u_o, u_v) = A_v \chi \in L^2 (\Omega_o \times \Omega_v) \tag{3} $$

with

$$ A_v \chi = \int_{-a}^{a} \exp \left[ j k_0 (u_o - u_v) x \right] \chi(x) dx $$

$k_0$ being the wavenumber at single frequency adopted in the scattering experiment.

In order to estimate the singular system of equation (3), it is convenient to study the eigenspectrum of the associated operator $A_v^* A_v$, with $A_v^*$ being the adjoint operator of $A_v$.

In this case it results that

$$ A_v^* A_v = \sum_{m=1}^{M} \frac{2 \pi (M - m)}{k_0} P_l B_{\Omega_m} P_I \tag{4} $$

where $P_I$ and $B_{\Omega}$ denote the spatial limiting and the band limiting projectors over the spatial $I$ and frequency $\Omega_m$ intervals, respectively and $\Omega_m = [-k_0 u_{max} + k_0 u_{im}, k_0 u_{max} + k_0 u_{im}]$. The operator $A_v^* A_v$ can be conveniently rearranged as (see Figure 2 for a graphical explanation)

$$ A_v^* A_v = \sum_{m=1}^{M-1} \frac{2 \pi (M - m)}{k_0} \left[ P_I B_{\Omega_m} P_I + P_I B_{\Omega_m} P_I \right] \tag{5} $$

Figure 2: The case of three angles of incidence $\Omega_i = \{-u_{max}, 0, u_{max}\}$. The top panel gives a pictorial view of the frequency bands which now overlap. The bottom panel shows how the spectrum of the kernel can be still given in terms of disjoint bands.

Let be $\lambda_n[A_v^* A_v]$ the eigenvalues of $A_v^* A_v$. They are not known in closed form. However, as shown in [13], if the spatial-bandwidth products of each single operator in equation (5) ($c_M = a \Delta / 2$) are greater than 4, the eigenvalues of $A_v^* A_v$ can be very well approximated by the union of those associated to each single operator, the latter being extensively studied in literature [14]-[15]. Therefore, the eigenvalues of $\lambda_n[A_v^* A_v]$ can be given in terms of those associated to the prolate spheroidal wave-functions ($\lambda_l(c_M)$) as

$$ \{ \lambda_n[A_v^* A_v] \}_{n=0}^{\infty} \simeq \bigcup_{m=1}^{M-1} \frac{2 \pi (M - m)}{k_0} \left[ \{ \lambda_l(c_M) \}_{l=0}^{\infty} \cup \{ \lambda_h(c_M) \}_{h=0}^{\infty} \right] \tag{6} $$

Figure 1: A pictorial view of the scattering configuration considered in this paper is sketched.
The same holds also true for the singular values of the scattering operator, denoted by $\sigma_n[A_\nu]$, that are the square root of such eigenvalues. In particular, as the behaviour of $\{\lambda_l(c_M)\}_{l=0}^\infty \cup \{\lambda_h(c_M)\}_{l=0}^\infty$ is a step-like one with the knee at the index $\lfloor 4c_M/\pi \rfloor$, the singular values of $A_v$ exhibit $M - 1$ steps where the first knee occurs at $N_1 = \lfloor 4c_M/\pi \rfloor$. The successive $m$-th step is comprised between the indexes $N_m = N_1 + \sum_{x=0}^{m-2} \lfloor 4c_M/\pi \rfloor$ and $N_m + \lfloor 4c_M/\pi \rfloor$, for $m \in \{2, 3, \cdots, M - 1 \}$. Moreover, on each step the singular values are equal to $\sqrt{(M - m)/\pi}k_0$ with $m \in \{1, 2, \cdots, M - 1 \}$. An example confirming this result is reported in figure 3.

By exploiting the results obtained for the singular values of $A_\nu$, an upper bound of the $\epsilon$-entropy in equation (2) can be analytically obtained. Assume that the bound about the norm of the noise $\epsilon$ is known. Hence, the cut-off threshold must be chosen as $N_\epsilon = \max\{n > 0 : \sigma_n[A_\nu] \geq \epsilon \}$. As the $\sigma_n(A_\nu)$ exhibit a $M - 1$ steps behaviour, estimating $N_\epsilon$ entails evaluating how many steps are above $\epsilon$. Accordingly let $\tilde{m}_\epsilon = \max\{m : (M - m)2\pi/k_0 \geq \epsilon^2 \}$ be the number of steps above $\epsilon$, then

$$H_{\epsilon v} \geq \frac{4c_M}{\pi} \sum_{m=1}^{\tilde{m}_\epsilon} \log_2 \frac{\sqrt{(M - m)(2\pi/k_0)}}{\epsilon} \quad M \geq 2$$

(7)

By substituting $c_M = ak_0u_{\text{max}}/(M-1)$, it can be obtained an estimate of the $\epsilon$-entropy in terms of the parameter of scattering

$$H_{\epsilon v} \geq \frac{4ak_0u_{\text{max}}}{(M-1)\pi} \sum_{m=1}^{\tilde{m}_\epsilon} \log_2 \frac{\sqrt{(M - m)(2\pi/k_0)}}{\epsilon} \quad M \geq 2$$

(8)

In figure 4, a comparison between the metric entropy evaluated from the exact singular values and the one estimated by the equation (8). In particular, it can be appreciated how the estimation given by the equation (8) works very well. Let $\mathcal{H}_\epsilon$ be the $\epsilon$-entropy for the single view / single frequency configuration. Its expression in terms of the parameter of the configuration is

$$\mathcal{H}_\epsilon \geq \frac{2ak_0u_{\text{max}}}{\pi} \log_2 \frac{2\pi/k_0}{\epsilon} \quad (9)$$

As expected, to collect the field for different angles of illumination entails an increasing on "the information" that can be conveyed back from data to the unknown with respect to the single view case and this improvement, if $\epsilon^2 < 2\pi/k_0$, can be quantified as

$$\frac{H_{\epsilon v}}{\mathcal{H}_\epsilon} \geq \left\{ \begin{array}{ll} \frac{2}{M=2} & \frac{2}{2[1 + \frac{1}{(M-1)} \log_2 \frac{(M-1)!}{\log_2(2\pi/k_0)}]} \quad M > 2 \end{array} \right.$$

(10)

In [12], it has been shown that in order to obtain the maximum number of measures ‘linearly independent ’ two extremal views (i.e., having angles corresponding to the edges of $\Omega_v$) are sufficient. Accordingly, multiple views can be considered redundant. From the equation (10), it can be concluded that also the redundant views yield an increasing on “the information” that can be conveyed back from data to the unknown. In fact, the redundancies "amplify" some spectral components of the data against the noise. This suggests that the optimal strategy of data collection would be to illuminate the object by two extremal views and probing the scene more times for each view. In this way a maximisation of the NDF and of the $\epsilon$-entropy is obtained, contemporaneously. When $\epsilon^2 > 2\pi/k_0$, $\mathcal{H}_\epsilon$ is zero and only collecting the data for $M > 2$ views allows to reconstruct the unknown.

3. Frequency diversity

In this section, the role of frequency diversity is studied. Therefore, let us assume to probe the field by varying the
frequency of the incident waves within the discrete interval \( \Omega_k = \{k_{min}, k_{max}\} \). In this case the scattering operator particularises as

\[
A_f : \chi(x) \in \mathcal{L}_d^a \rightarrow E_S(u_o, k) \in \mathcal{L}_d^{\Omega_k \times \Omega_k}
\]

where now \( u_o = \sin \theta_o \) and \( u_i \) has been assumed equal to zero (normal incidence).

Let \( \Omega_k \) consists of \( M \) frequencies \( k_m \) taken uniformly between \( \{k_{min}, k_{max}\} \) at a step of \( \Delta k = (k_{max} - k_{min})/(M - 1) \). Accordingly, it results that

\[
A_f^* A_f = \sum_{m=1}^{M} \frac{2\pi}{k_m} P_I B_{\Omega_m} P_I + \sum_{m=1}^{M} \sum_{l=1}^{M} \frac{2\pi}{k_1} \left( P_I B_{\Omega_m} P_I + P_I B_{\Omega_l} P_I \right)
\]

where now \( \Omega_m = [-k_m u_{max}, k_m u_{max}] \). For the sake of simplicity, from now on, we assume that \( u_{max} = 1 \) so that \( \Omega_m = [-k_m, k_m] \).

The operator \( A_f^* A_f \) can be conveniently rearranged as (see figure 5 for a graphical explanation)

\[
A_f^* A_f = \sum_{m=1}^{M} \frac{2\pi}{k_m} P_I B_{\Omega_m} P_I + \sum_{m=1}^{M} \sum_{l=1}^{M} \frac{2\pi}{k_1} \left( P_I B_{\Omega_m} P_I + P_I B_{\Omega_l} P_I \right)
\]

where \( \hat{\Omega}_m = \{k_{min} + (m - 1)\Delta k, k_{min} + m\Delta k\} \) and \( \hat{\Omega}_m = \{-k_m - m\Delta k, -k_m - (m - 1)\Delta k\} \).

Now, \( A_f^* A_f \) is in a suitable form to apply the results reported in [12]. Hence, it has been shown that as long as \( \epsilon M = a\Delta k/2 \gg 1 \) and \( ak_{min} \gg 1 \) the \( \lambda_{nf} \) \( A_f^* A_f \) exhibit \( M \) steps with knees occurring at \( N_m = \lfloor 2k_{min} a/\pi \rfloor + (m - 1)\lfloor 2a\Delta k/\pi \rfloor \). Moreover, on the \( m \)-th step the numerical value is \( \sum_{l=m}^{M} 2\pi/k_l \), with \( k_l \) represents the \( l \)-th frequency in \( \Omega_{k_l} \). An example confirming this result is reported in figure 6.

The following analytical estimation for the \( \epsilon \)-entropy \( \mathcal{H}_{nf} \) follows

\[
\mathcal{H}_{nf} \geq \frac{2k_{min} a}{\pi} \log_2 \left( \frac{\sum_{l=1}^{M} (2\pi/k_l)}{\epsilon} \right) + \frac{2(k_{max} - k_{min}) a}{\pi(M - 1)} \sum_{m=2}^{M} \log_2 \left( \frac{\sum_{l=m}^{M} (2\pi/k_l)}{\epsilon} \right)
\]

where \( \hat{m}_c = \max\{m : \sum_{l=m}^{M} (2\pi/k_l) \geq \epsilon^2\} \). Also here, this estimation works very well (see figure 7).

As expected, collecting data for different frequencies below \( k_{max} \) entails an increasing of the information as compared to the single frequency \( k_0 = k_{max} \) case. This improvement, if \( \epsilon^2 < 2\pi/k_{max} \), is estimated as

\[
\mathcal{H}_{nf}/\mathcal{H}_{e} \geq 1 + \frac{1}{(M - 1)k_{max}} \left( \log_2 \sqrt{2\pi/k_{max}} + \frac{k_{min} \log_2 \sqrt{\sum_{l=1}^{M} (2\pi/k_l)}}{\epsilon} \right)
\]

In [12], it has been concluded that while the maximum number of measurements ‘linearly independent’ depends on the highest adopted frequency \( k_{max} \), by using more frequencies (lower than the highest one) shapes the singular
value behaviour so that it resemble a staircase. From equation (14), it can be observed that the presence of multiple frequencies yield an increasing of the $\epsilon$-entropy and this improvement depends on the parameters of the configuration. Also here, an optimal collecting strategy of the field can be reached by illuminating the object at the frequency $k_{max}$ and probing the scene more time at the same frequency. In this way a maximisation of the NDF and of the $\epsilon$-entropy is obtained, contemporaneously.

4. Comparison between the diversities

The expressions obtained in the previous sections allow to obtain an estimation of the $\epsilon$-entropy in terms of the scattering parameters. In this section, such results are used to compare the different configurations considered herein.

Assume that the bound about the norm of the noise $\epsilon$ is known and equal to $\epsilon = 0.21$ and that the observation domain is $\Omega_o = [-1, 1]$. For the single frequency case (which means for both the single-view/single-frequency and multi-view/single-frequency configurations) the object is illuminated at the spatial frequency $k_0 = k_{max}$.

Equation (9) normalised with respect to $\alpha$ becomes

$$\hat{H}_\epsilon \geq \frac{2k_{max}}{\pi} \log_2 \sqrt{\frac{2\pi/M}{\epsilon}}$$

where $\hat{H}_\epsilon$ denotes the normalised $\epsilon$-entropy.

When the multi-view configuration is considered, the scattering scene is probed at $M$ different angles of incidence. Accordingly, equation (8) normalised to $\alpha$ becomes

$$\hat{H}_{\alpha} \geq \frac{4k_{max}}{(M-1)\pi} \sum_{m=1}^{M-1} \log_2 \sqrt{\frac{(M-m)(2\pi/k_{max})}{\epsilon}}$$

$$M \geq 2$$

(17)

From equation (17) it can be observed that in order to obtain better performance in terms of $\epsilon$-entropy only $k_{max}$ and $M$ play a role. For multi-frequency configuration only normal incidence is considered. In this case, also $k_{min}$ enters and affects the ‘information’. Equation (14) normalised is

$$\hat{H}_{fe} \geq \frac{2k_{min}}{\pi} \log_2 \left\{ \sum_{i=1}^{M} \left( \frac{2\pi/k_i}{\epsilon} \right) \right\} +$$

$$\frac{2(k_{max} - k_{min})}{\pi(M-1)} \sum_{m=2}^{M} \log_2 \left\{ \sum_{l=m}^{M} \left( \frac{2\pi/k_l}{\epsilon} \right) \right\}$$

(18)

In order to perform the comparison let us assume $k_{min} = 2\pi m^{-1}$. This way, the different configurations can be compared by varying $k_{max}$ and $M$ only.

In figures 8 and 9, the normalised upper bound of the $\epsilon$-entropy is shown as a function of $k_{max}$ for each configuration, when $M = 2$ and $M = 3$, respectively. In particular, the blue line is associated to single view/single frequency configuration (that is when no diversities are employed), while the red and the green lines are associated to the multi-view and multi-frequency configurations, respectively.

As expected, there is a gain in terms of $\epsilon$-entropy obtained by exploiting the diversities with respect to the single-frequency/single-view configuration. This is always true as the scattering parameters vary.

Furthermore, as can be seen, the multi-view configuration allows to obtain better results than the multi-frequency one and this improvement increases as more views or frequencies are introduced.

However, this cannot allows to state that the multi-view configuration is the optimal one because in general the $\epsilon$-entropy depends on the chosen $k_{min}$ as well. Indeed, for
some parameters the multi-view configuration can be better than the multi-frequency one (as in this case) and for other parameters the opposite is true.

So, the presented analytical analysis allows to get the answer to what is the optimal configuration when the parameters have been fixed according to some constraints that one may have.

5. Conclusion

The impact of view and frequency diversities on the Kolmogorov entropy measure has been studied in the framework of linear inverse scattering problems.

Canonical strip-like scatterers have been considered. This has allowed to determine and estimate the $\epsilon$-entropy for all the considered configurations. Also, these expressions clearly make it evident the role of the scattering parameters (i.e., observation corner, number of views, frequencies, etc.). Numerical checks of such estimations showed that they are in very good agreement with the actual entropy.

As expected, the use of diversities increases the information content. Also, some numerical examples have been run to compare the view and the frequency diversities. It is shown that, for the considered parameters, view diversity outperforms the frequency one. However, in general this cannot be the case.

The main results of this paper is to provide a tool to foresee the information content once the configuration parameters have been fixed. Conversely, the reported estimations can be used to set the configuration parameters in order to maximise the achievable information content.

Acknowledgement

This work was supported by the Italian Ministry of University and Research through the FIRB initiative under the project MICENEA (RBFR12A7CD).

References