

The low dimensional world of echolocating bats

Abstract

This template helps you to create a properly formatted L^AT_EX manuscript.

1. introduction

Echolocating bats are long-lived, mobile animals. They use the landscape intensively for foraging, roosting, and commuting. They have excellent spatial memory (Barchi et al., 2013; Von Helversen & Winter, 2005) and can rely on sonar to navigate to notable locations like foraging grounds, drinking places (see Vanderelst et al., 2016; Vanderelst & Peremans, 2017, for references). Furthermore, bats have been reported to use sonar-based recognition of landmarks, both under experimental (Jensen et al., 2005; Yu et al., 2019), and natural (Verboom et al., 1999) conditions.

Research on animal navigation has revealed a range of potential strategies that can be used to reach a distant goal (Reviewed by Franz & Mallot, 2000). Irrespective of the strategy an animal uses, a mechanism for recognizing previously visited places is required (Vanderelst et al., 2016; Vanderelst & Peremans, 2017). This implies that echolocating bats can identify previously visited sites (or landmarks defining those places) using sonar. Currently, it is still an open question of how bats use sonar to recognize locations in the environment.

One mechanism for place recognition would be for bats to parse the stream of echoes returning from the environment into localized and identified objects (Lee et al., 2017; Barchi et al., 2013; Moss & Surlykke, 2001; Schnitzler et al., 2003; Simmons, 2012; Ulanovsky & Moss, 2008; Clare & Holderied, 2015; Surlykke et al., 2016). However, as we argued before, it is highly questionable whether bats can interpret echoes in terms of a 3D environment model or acoustic image (e.g., Vanderelst, 2015; Vanderelst et al., 2016; Steckel & Peremans, 2013). Sonar maps the three-dimensional structure of the environment on two one-dimensional waveforms arriving at the ears. While the time of arrival of echoes encodes object distance, azimuth and elevation directions are not directly available from the acoustic signal. In principle, azimuth and elevation can be estimated using spectral cues imposed by the head-related transfer function [REF]. However, complex environments, and even single plants [REF], return many overlapping and interfering echoes, making reconstructing the direction of origin of echoes much harder, and perhaps impossible. Extracting spatial information is further hampered by the limited temporal resolution of the sonar system (Simmons

et al., 1989; Wiegrebe & Schmidt, 1996; Surlykke & Bojesen, 1996).

We have suggested before that limitations of the bat's sonar system result in bats having access to a sparse, low dimensional description of the outside 3D world only (Vanderelst et al., 2015, 2016). Some recent evidence seems to confirm this. First, Geberl et al. (2019) confirmed that bio-sonar has a low spatial resolution, both for distance and angular dimensions. Also, neurophysiological data suggest that closely spaced objects are not perceived separately by echolocating bats. In contrast, neural responses to echo cascades from densely spaced objects seem to represent a single extended stimulus (Warnecke et al., 2018).

Look at recent Simmons paper where bats refuse to fly through hoop

Low dimensional descriptions potentially contain a lot of information (Kuc, 1997b,a). We previously demonstrated that bats might use sonar template-based scene recognition (Vanderelst et al., 2016; Vanderelst & Peremans, 2017). In this approach, bats are assumed to compare the acoustic signature of the environment to a set of previously stored signatures. We operationalized these acoustic signatures as time-intensity profiles, which could be derived directly from the cochlear output.

In this paper, we directly evaluate the hypothesis that sonar provides bats inherently with a low dimensional, yet sufficiently detailed, description of the world. We test this hypothesis in two steps. In a first step, we generate an extensive collection of synthetic echo sequences. We assess the compressibility of these echo signals using Principal Component Analysis (PCA). Next, we filter the echo sequences using a model of the bat's auditory periphery (Wiegrebe, 2008). The synthetic echo sequences differ greatly in complexity. Nevertheless, the output of the cochlear model can be mapped onto a low-dimensional space. In a second step, we show that echoes collected in real bat habitats can be mapped into the same low-dimensional space, without loss of information. We do this by demonstrating that the low dimensional feature space allows recognizing locations in real habitats.

Together these results suggest that echolocating bats live in a low dimensional perceptual world. The physics of the sonar system and the auditory periphery map the

complexity of the 3D environment onto a small number of independent features, irrespective of whether we use artificial or real echo sequences. Nevertheless, this limited set of features is sufficiently rich for bats to recognize locations, and as we discuss, support flight control.

2. Methods

2.1. Synthetic echoes

In this section of the paper, we assess how the filtering in the auditory periphery reduces the complexity of input signals. We do this by presenting a simple generative model for echo sequences received by bats. Next, we assess whether the cochlear representation derived from these synthetic echo sequences can be projected into a low dimensional space.

We created artificial echo sequences by randomly spacing ideal point reflectors. The distances of the point reflectors were selected using a two-step process. First we selected n_s seed points s , uniformly distributed over a range from 0.5 to 7 meters. Next, for every seed point s , we selected n_c cloud points normally distributed around s , with standard deviation 0.25 m. This two step process to modeling natural reflectors was inspired by the results of Yovel et al. (2009), who found that echoes returning from vegetation could be better modeled using a similar two-step process than simply using uniformly spaced points. We generated xx sets of points for each combination of $n_s \in [xx, xx, xx, xx]$ and $n_c \in [xx, xx, xx, xx]$. Synthetic echoes consisting of few reflectors are likely to consist of only weak echoes. To ensure that all sequences included at least one strong echo, we added a single reflector at a distance between xx and xx to each collection of points.

Each set of points was converted to an impulse response with a sampling frequency f_s of 219 kHz. The impulse response length was set to 7499 samples. These parameters correspond to the recording length and the sampling frequency used in the acoustic data collection (see below). In converting the point to an impulse response, we modeled attenuation of the echoes due to spherical spreading (both ways). We varied the amplitude of the impulses according to a normal distribution with a standard deviation of 10 dB.

The impulse responses were converted to echo sequences by convolving them with a 1 millisecond long linear frequency-swept cosine starting at 100 kHz and ending at 40 KHz. Again, these parameters correspond to those used in the acoustic data collection described below. Examples of the resulting echo sequences are shown in figures 1 and 2.

Artificial echo sequences were filtered using a model of the bat’s auditory periphery, as proposed by Wiegbe (2008). This model simulates the temporal and spectral resolution of the bat’s cochlea. The model consists of a gammatone filterbank followed by half-wave rectification and non-linear compression. Finally, the model assumes

a 1 kHz low-pass filter modeling the response characteristics of the inner hair cells. In the current implementation, we used a gammatone filterbank with xx filters, centered around xx kHz. The filters were spaced to correspond equivalent rectangular bandwidths (Glasberg & Moore, 1990). The output of the channels were filtered with a time (distance) weighted factor corresponding to the atmospheric attenuation at the channel’s center frequency. This filtering approximated frequency-dependent atmospheric attenuation. The first 6 ms of data were removed. This was done to ensure the data corresponded with the acoustic data. As indicated below, we removed the first 6 ms of those data to avoid including the pickup of the emitted signal.

Figure 1 depicts a single example of a synthetic echo sequence. It also illustrates the resulting response for each of the xx channels of the Wiegbe (2008) model. From these responses, it can be inferred that these are highly correlated. This is confirmed when analyzing the average cross-channel correlation for all xx synthetic echo sequences (Fig. 1xx). From these high cross-channel correlations, we conclude that the output of the cochlear model can be summarized by averaging across channels without extensive loss of information. An example of such a cross-channel sum is plotted in figure 1xx. In correspondence to our previous work (Vanderelst et al., 2016), we refer to this average across frequency channels as a template.

We used PCA to assess the dimensionality of the templates generated from the synthetic echo sequences. We found that almost all of the variance in the templates could be captured by 25 components (see figure 1xx). Some examples of templates reconstructed from the principal components are shown in 2xx).

In this section, we used a generative model for naturalistic echoes. We assumed that echoes sequences received by bats can be considered as generated by anisotropically distributed point reflectors. Using this model, we have derived a low dimensional space ($D = 25$), which can be used to describe the templates resulting from the artificial echo sequences. This implies that, after filtering through (a model of) the bat’s auditory system, these artificial echoes can be converted to low dimensional descriptions. The original echo sequences, generated by xx to xx point reflectors effectively map to 25 dimensional vector. In the remainder of the paper, we will assess whether this low dimensional space allows describing the echoes from real bat habitats as collected using an ensonification device.

2.2. Natural Templates

2.2.1. Acoustic data

We reused the large library of echoes gathered in real bat habitats by Vanderelst et al. (2016). Below, we briefly discuss the construction of templates based on these echoes, and we refer to our previous work for more details. Vanderelst et al. (2016) constructed an ensonification device featuring 31 Knowles FG series microphones connected to

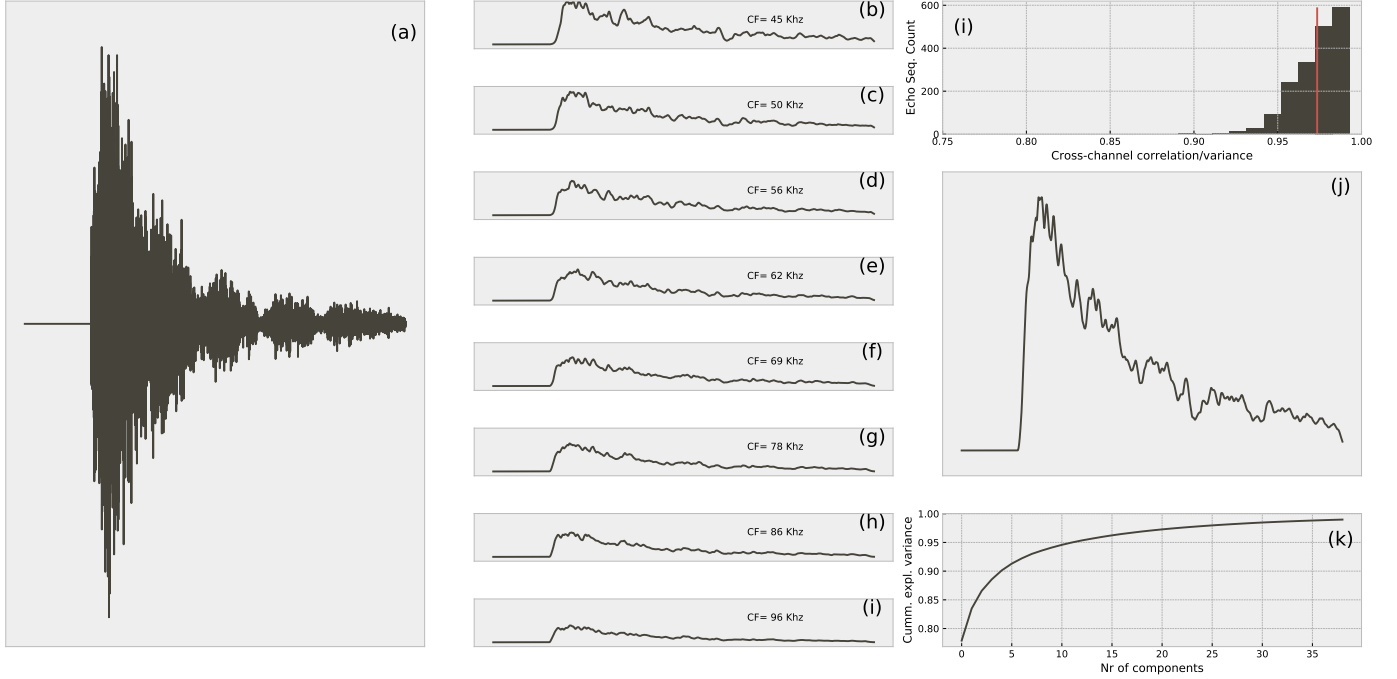


Figure 1: Example of synthetic echo sequence and its processing using the Wiegrobe (2008) model. (xx) Example of a single synthetic echo sequence. (xx-xx) Response of the xx frequency channels in the Wiegrobe (2008) model. From these plots, it can be seen that channels highly correlate. This is confirmed in panel xx, depicting the average correlation across channels for all xx synthetic echo sequences analyzed. (xx) The output of the Wiegrobe (2008) model summed across the frequency channels, is referred to as a template in this paper. The template resulting from the echo sequence in panel xx is shown in panel xx. Panel xx shows the cumulative proportion of explained variance as a function of number of PC across all templates for all synthetic echo sequences.

a custom-built sonar data acquisition board (Steckel et al., 2013). The emitter was a Sensecomp 7000 Series ultrasonic broadband speaker. The ensonification device was mounted on a pan-tilt system allowing to rotate it in azimuth and elevation. The device and pan-tilt system was mounted on a tripod.

In our previous work, the ensonification device was used to ensonify bat habitats at three different sites. However, here we only use data collected along two linear pathways through natural habitats. At the first site, a public park in Sde Boker, Israel (Fig. 3). The environment consisted of a walkway lined with rocks and vegetation. Unpublished bat monitoring data at the site confirmed that bat used this corridor to commute. The ensonification device was placed at 50 positions along a straight line spanning a total of 10 meters. Locations were spaced 20 cm apart. At each location, echoes were collected for 31 azimuth directions (-150 to +150 degrees) and 7 elevation directions (-40 to +25 degrees). Therefore, at each location, echoes were collected from $31 \times 7 = 217$ directions. At each azimuth and elevation direction, three repeat measurements were recorded. In this way, a total of 32,550 echo trains (50 positions \times 217 directions \times 3 repeats) for each of the 31 microphones in the device were collected.

The second site, at the University of Bristol Royal Fort Gardens (UK), data were collected at 40 different positions along a line (Fig. 3). Locations were spaced 25 cm apart.

As for the Israel site, we collected data for 217 directions, and measurements were repeated three times. This resulted in 26,040 echo trains (40 locations \times 217 directions \times 3 repeats) for each of 31 microphones. The recording of the echoes started with the emission of a hyperbolic frequency modulated sweep and was recorded for 34 ms (~ 5.7 m) at a sampling rate of 219 kHz.

2.2.2. Template construction

For every azimuth and elevation direction at each of the 40 or 50 locations, a biologically plausible acoustic signature of the environment was derived from the echoes, i.e., a template. Below, we briefly outline the steps used to derive templates from the acoustic data.

As we did for the synthetic echoes sequences, the echoes trains collected in bat habitats were filtered using a model of the bat's auditory periphery as proposed by Wiegrobe (2008). We concluded from analyzing the synthetic echoes that the response in different frequency channels is highly correlated. Therefore, the output of the cochlear model was averaged across frequencies to obtain a time-intensity profile, or template. The beam of the emitter used was more narrow than the typically combined hearing and emission directionality of bats (Vanderelst et al., 2010; Reijniers et al., 2010; Jakobsen et al., 2012). Therefore, templates were averaged across three neighboring directions. The first 6 ms of data were removed to avoid using emitted signals for analysis. The processing steps outlined above

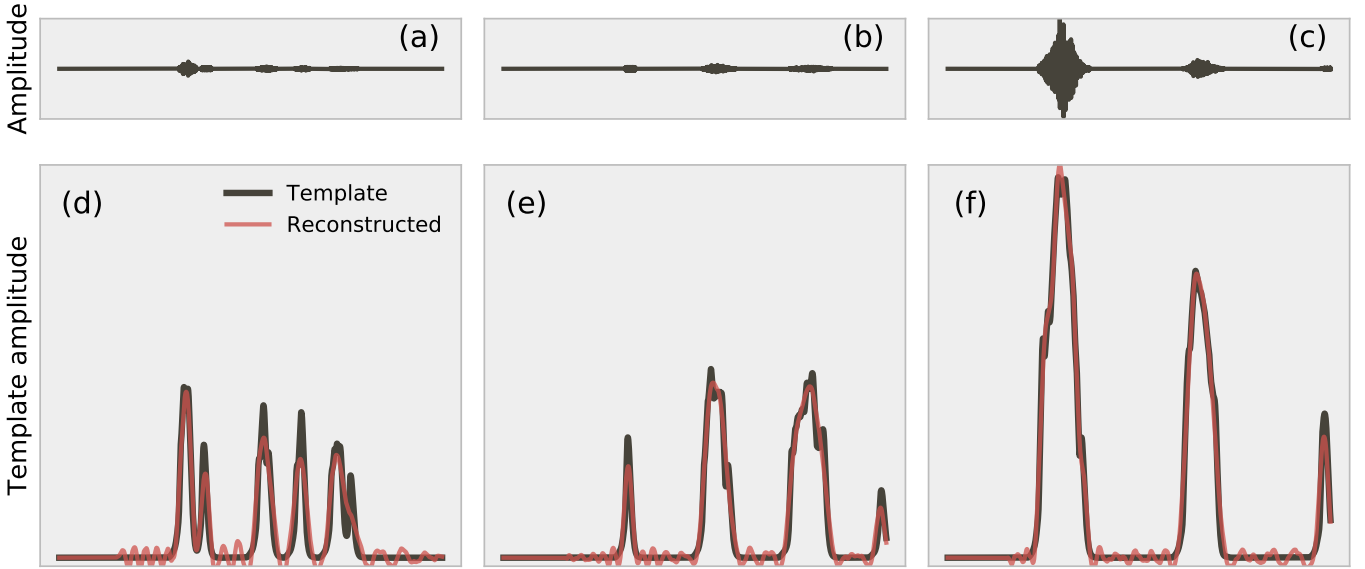


Figure 2: Example of synthetic echo sequences, their corresponding templates and the templates reconstructed from the 25 PCs. (xx-xx) Three examples of synthetic echo sequence waveforms. Panels xx-xx depict the resulting templates (i.e., summation across the output channels of the Wiegube (2008) model). These panels also show the result of projecting the templates into a 25D space and transforming the result back to the original coordinates.

resulted in 3×217 templates, for each location at each site, corresponding to 217 directions and three repeats.

Among the sets of templates, some had very little energy in them, notably those collected for directions without reflectors. For example, aiming the ensonification device upwards often resulted in no echoes returning. These templates could not be classified (Vanderelst et al., 2016). Therefore, we removed them from the current analyses by omitting all templates with very low energy. The resulting number of templates included in the present analyses are given in table 4.2.

2.2.3. Noise floor

Internal (and, potentially, external) noise results in template values greater than zero, even in the absences of reflectors returning echoes. To avoid training the neural network on this noise, we estimated the noise floor in our data by collecting templates while pointing the device to empty space. The maximum template value obtained from these template measurements was used as the noise floor \mathcal{N}_f . The value of the noise floor \mathcal{N}_f used had a numerical value of 0.15. Any template values below this value were set to this value. As we argued before (Vanderelst et al., 2016), using the empirical measurement noise to estimate \mathcal{N}_f is biologically warranted as bats have a better signal to noise ratio in their auditory systems compared to our measuring device. Indeed, their dynamic range is at least 30 dB larger than our measurement device.

2.2.4. Mapping templates into the 25-dimensional space

The templates from both sites were mapped into the 25-dimensional space derived from the synthetic echoes,

described above. We assessed how well the templates collected in real habitats can be represented by the 25 components by mapping the templates into this space and transforming the results back to the original space. Next, we calculated the correlation (and explained variance) between the original templates and the reconstructed templates.

2.3. Artificial Neural Networks

We wished to assess whether the templates projected into the 25-dimensional space still contained sufficient information to be recognized. Previously, we assessed whether templates can be used to recognize locations using, what Baddeley et al. (2012) called, a perfect memory approach. Templates were stored in a lookup table. The Euclidean distance between two echo templates was the metric used to assess how well bats should be able to distinguish between templates. This table lookup and distance calculation require a comprehensive search of memory. Hence, in our previous work, we established that templates could be classified, but this required a significant amount of memory and computational resources.

Here, we used a more biologically plausible memory system: an artificial neural network. These are a neurally inspired way of distributed information storage in artificial neural networks (McLeod et al., 1998). However, we acknowledge that current training methods most likely have no biological counterpart.

Simple artificial feedforward neural networks were implemented using Tensorflow 2 for Python. The first layer of the neural network added Gaussian noise to each input sample ($\sigma(\mathcal{N}_s) = 0.01$). The following layers were



Figure 3: Pictures of the sonar data acquisition process. Top: the ensonification device used at the Israeli site at Sde Boker. Bottom: The device used at the Bristol University Royal Fort Gardens. The tape measure indicated the transect along which the device was positioned in 40 steps.

completely connected and used the rectified linear activation function. The output layer used a softmax activation function (or the normalized exponential function), normalizing the output into a probability distribution consisting of probabilities proportional to the exponentials of the output layer’s activation. The loss was calculated as the cross-entropy between the target output and the actual output of the neural network. Weights were updated after each epoch using the Adam algorithm (Kingma & Ba, 2014).

For each site, we trained three neural networks, each taking the 25-sample templates as input. The first neural network was trained to output the azimuth direction associated with the template (31 targets). The second network was trained to return the elevation (7 different targets). The final network returned the location (50 or 40 different targets). The targets for the network were encoded using the one-hot encoding scheme.

Training a different neural network for each of these features ensured that the training algorithm paid equal attention to each feature. When training a single network for all features, the optimizer could preferably reduce errors in the location predictions as this feature has the largest variance. The full architecture can be considered as a sin-

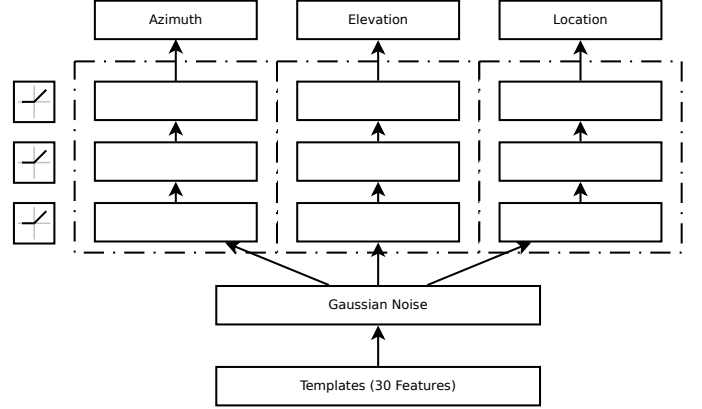


Figure 4: Depiction of the neural network architecture trained for each site.

gle network with different streams for each feature (fig. 4).

We should point out that we did not take any steps to avoid so-called overfitting by the neural networks. Overfitting occurs when a neural network (Ghotra & Dua, 2017), or other mathematical models (Hawkins, 2004), extract complicated relationships between inputs and outputs that do not hold beyond the training examples. Overfitting typically reduces the ability of the model to generalize its predictive power to new data. However, in this case, the model is not required to generalize. Instead, we are only interested in whether the neural networks can learn and classify the set of training examples.

2.3.1. Training examples

We wished to ensure that the neural networks encoded the templates and were able to recognize them in the presence of realistic noise. Indeed, it can be assumed that neural noise interferes with encoding or memorizing templates. Therefore, we trained the neural networks using templates to which noise was added.

In our previous work, we describe a procedure, based on Dau et al. (1996), to derive a realistic model for stochastic noise. In brief, we empirically determined a noise level that resulted in a 75% correct discrimination of two artificial bat calls differing by 2 dB in amplitude. This criterion is a conservative estimate of the just noticeable difference in loudness in bats. Based on this procedure we assumed internal (independent) Gaussian stochastic noise \mathcal{N}_s for each sample with $\sigma(\mathcal{N}_s) = 0.01$ and $\mu = 0$.

2.4. Perfect Memory

To be able to directly compare the results of the neural networks with the perfect memory approach (Baddeley et al., 2012), we implemented a version of the classification scheme employed in our earlier work (Vanderelst et al., 2016). We generated noisy templates by adding Gaussian noise to them ($\sigma(\mathcal{N}_s) = 0.01$). Next, we performed the classification of these noisy templates by finding the closest template (according to the Euclidean distance) in the

original set of templates. The location, azimuth, and elevation of the closest template were considered as the noisy template’s classification. This was repeated ten times, and we report on the average performance across these repeats.

3. Results

We found that the templates derived from acoustic data collected in real habitats could be well represented by the 25 components derived from synthetic data. This 25-dimensional space could represent xx% of the variance in the templates. Figure 5 shows examples of templates and their reconstructed counterparts. Therefore, although the model used to generate echo sequences was very simple, the resulting underlying 25D feature space could adequately represent templates resulting from real acoustic data.

In our previous work [Vanderelst et al. \(2016\)](#), we showed that the templates could be used to recognize locations in bat habitats. As we have shown that the templates can be projected into a lower 25D feature space without loss, it can be expected that these 25 features can be used to locations as well. We found this to be indeed the case.

Figure A.9 depicts the training history of all networks. This figure shows that, in all cases, training almost reached an asymptote around or before 250 epochs. However, some additional training can be seen up to xx epochs.

Figure 6 summarizes the performance of the three neural networks trained on the Israel templates. The performance was obtained by assuming that the most active output neuron represented the classification of the network. Inspecting the cumulative distribution of error reveals that the neural network encoding outperformed the perfect memory approach. The number of small errors was larger for the neural network than for the perfect memory. The neural network was able to classify about 70% of templates with an error of fewer than 30 degrees in azimuth. In elevation, almost all templates were classified with an error below 15 degrees. Most (70%) errors in location attribution were smaller than 0.5 meters. The confusion matrices confirm that most classification errors were small.

Interestingly, for azimuth and elevation, classification errors were not uniform across directions. This can be seen by the classification entropy plotted in figure 6. In particular, these curves show interesting peaks around zero azimuth and medium elevations. The geometry of the environment can explain these peaks. The corridor (fig. 3) does not return a lot of echoes from azimuths around zero. Indeed, this direction is mostly open space. The same is true for elevations above 0. Elevations around 25 return very few echoes and this serves as a reliable indicator for classifying these templates, explaining the drop in entropy for elevation = 25 degrees.

Figure 7 shows the results for the Royal Fort Park data. These results mirror those in figure 6. Again the neural network outperformed the perfect memory approach by

having more smaller errors. Overall, the performance for this data set was somewhat better than for the Israel data.

4. Discussion

4.1. The low dimensional world of bats

[Reijniers & Peremans \(2010\)](#) quantified the transmission of auditory information in the cochlea of bats. In that study, the authors adapted the Meddis inner hair cell model ([Meddis, 2006](#)) to assess the amount of information in the acoustic signal preserved by the cochlea. These authors concluded that spectral notches introduced by overlapping echoes are faithfully encoded by auditory nerve spike patterns. However, quantifying the information transfer rate also found that about half of the information (bits) are lost in the transduction from acoustic energy to spike trains. In the current paper, we expand on this work and quantify the dimension reduction due to cochlear encoding. Moreover, we show that this reduction in complexity still supports an ecologically valid task: sonar-based scene recognition.

We assumed echoes received by bats can be modeled as originating from a large collection of point reflectors. Passing these synthetic echoes through a model of the auditory periphery, we obtained a time-intensity profile (template) for each echo sequence. These templates could be well described by a low dimensional space of 25 principal components (figure 2). We found that the same 25D space could be used to describe templates originating from echoes collected in real bat habitats (figure 5). This suggests the filtering of the auditory periphery provides the bat with a greatly reduced acoustic description of its environment. To identify the locus of the compression, we performed a number of additional analyses. In particular, we evaluated whether the low-dimensional description is a property of the echo sequences, the gammatone filters, or the 1 kHz low-pass filter in the [Wiegbe \(2008\)](#) model.

We generated templates from the synthetic echo sequences using two versions of the [Wiegbe \(2008\)](#) model; with and without the 1 KHz low-pass filter. Per default, the model includes this filter to model the response characteristics of the inner hair cells ([Meddis, 2006](#); [Reijniers & Peremans, 2010](#)). By removing this lowpass filter, we can assess to what extent it is responsible for the observed compressibility of the templates. We also generated templates using echo sequences of which the samples were randomly shuffled. This allows inferring whether the low-dimensionality of the templates is a property of the echo signals (before being filtered through the cochlear model).

Figure 8 shows the number of PCs needed to capture 99% of the templates variance. From this plot, it can be seen that the number of PCs increases when shuffling the echo sequences increases from xx to xx. Omitting the 1 kHz filter from the [Wiegbe \(2008\)](#) model increases the number of PCs to xx. Inspecting the average spectrum of the templates (fig. 8, inset) shows that most energy

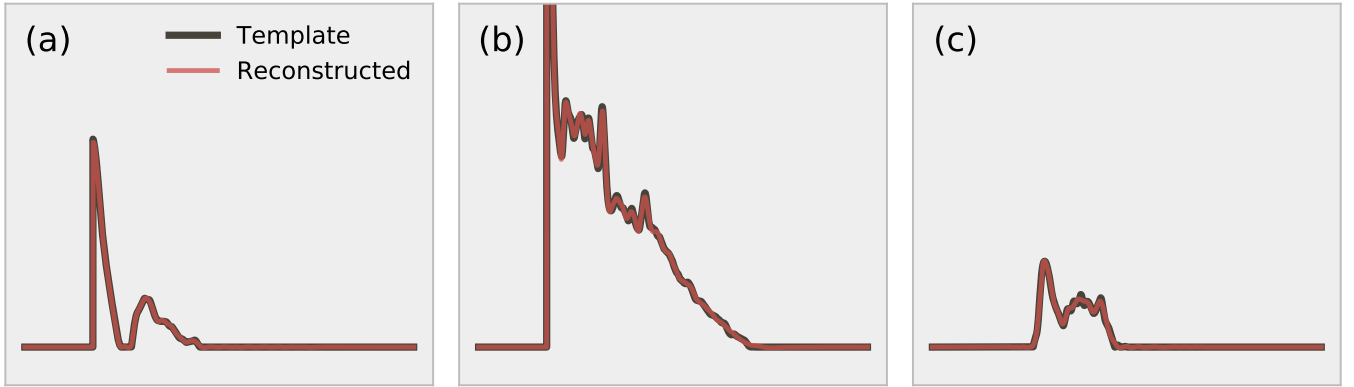


Figure 5: Examples of three templates generated from the echoes collected at the Israel and Royal Fort Gardens site. The panels also depict the result of projecting the templates in the 25-dimensional space and transforming the result back to the original coordinates.

in the templates is contained in frequencies below 1 kHz. Therefore, it ought to be possible to completely represent templates by sampling them at 2 kHz, or about 70 samples per template, as templates consist of 7499 samples at 219 kHz sample rate. This corresponds to the number of PCs required to represent templates derived from echo sequences that were shuffled (fig. 8). The number of PCs required to represent unshuffled templates is about half of this, i.e., 25. Therefore, we conclude that the 1 kHz filter is the main, but not exclusive, locus of the dimensionality reduction. The structure of the echo signals also increases the dimensionality of the templates.

Despite the low dimensional character of the templates, they are sufficient to recognize environments. Indeed, this was demonstrated in our previous work (Vanderelst et al., 2016). Here, we directly demonstrated that the 25 principal components are sufficient to discriminate their associated locations using neural networks.

Here, we show that a low dimensional signal is sufficient for recognizing environments. This is a critical component of navigation. However, bats do more than navigating. They also use sonar to control flight, for example, avoid obstacles or following guiding structures. The description the bat has access to for flight control has the same low-level character than that used for place recognition. Therefore, our work suggests that this low-level description is sufficient for flight control as well. We have shown before that a low-level description could be sufficient for obstacle avoidance and corridor following (Vanderelst et al., 2015; Mansour et al., 2019). However, in that work, we converted both the echoes at the left and at the right ear to template-like representation. Subsequent work will have to demonstrate that the parameters of the 25D space presented here can be used directly for controlling flight. In other words, we conjecture it should be possible to extract 25 features from an echo signal and convert this vector into a motor command. One can presume a neural network that converts the 25 features to flight control parameters.

Site	T	n	W	W (nb)	R
Israel	206661	5299	53322	52425	25.8%
Royal	302991	7769	51945	51075	17.1%

Table 1: Table indicating the size of both template data sets and number of weights in the networks. The template size T is obtained as $25 \times n$, with n the number of templates in the set. The number of weights W includes all connections between neurons as well as the bias weights for each of the three networks trained. $W(nb)$ omits the bias weights. The ratio R gives the ratio W/T .

4.2. Template encoding

We were unable to find numbers on the number of neurons in the cortex of bat species. However, the mouse cortex has been estimated to contain about $9 \cdot 10^4$ neurons or $7 \cdot 10^8$ synapses per cubic millimeter (Braitenberg & Schüz, 2013). The neural networks presented in this paper consisted of about 200 neurons and 37 thousand connections. Therefore, in as much as these neural networks can be considered as biologically plausible models of memory, they are orders of magnitude smaller than a single cubic millimeter of mouse cortex. Let alone the complete cortex, which has been estimated to contain 75 million neurons (Williams, 2000).

Despite their small size, the networks in this paper were able to encode (memorize) the templates as least as well as the perfect memory approach. The neural networks presented here are tiny compared to the density of cortical matter. However, the networks are also small compared to the data they encode. Table 4.2 lists the number of weights for the artificial neural networks and compares this number with the size of the 25-dimensional templates. The number of floats making up the templates is larger than the number of weights by a factor of 2-3. Therefore, the networks have a smaller memory footprint than the data they encode. It should be noted that we have not made an attempt to try and to optimize (minimize) the size of the networks.

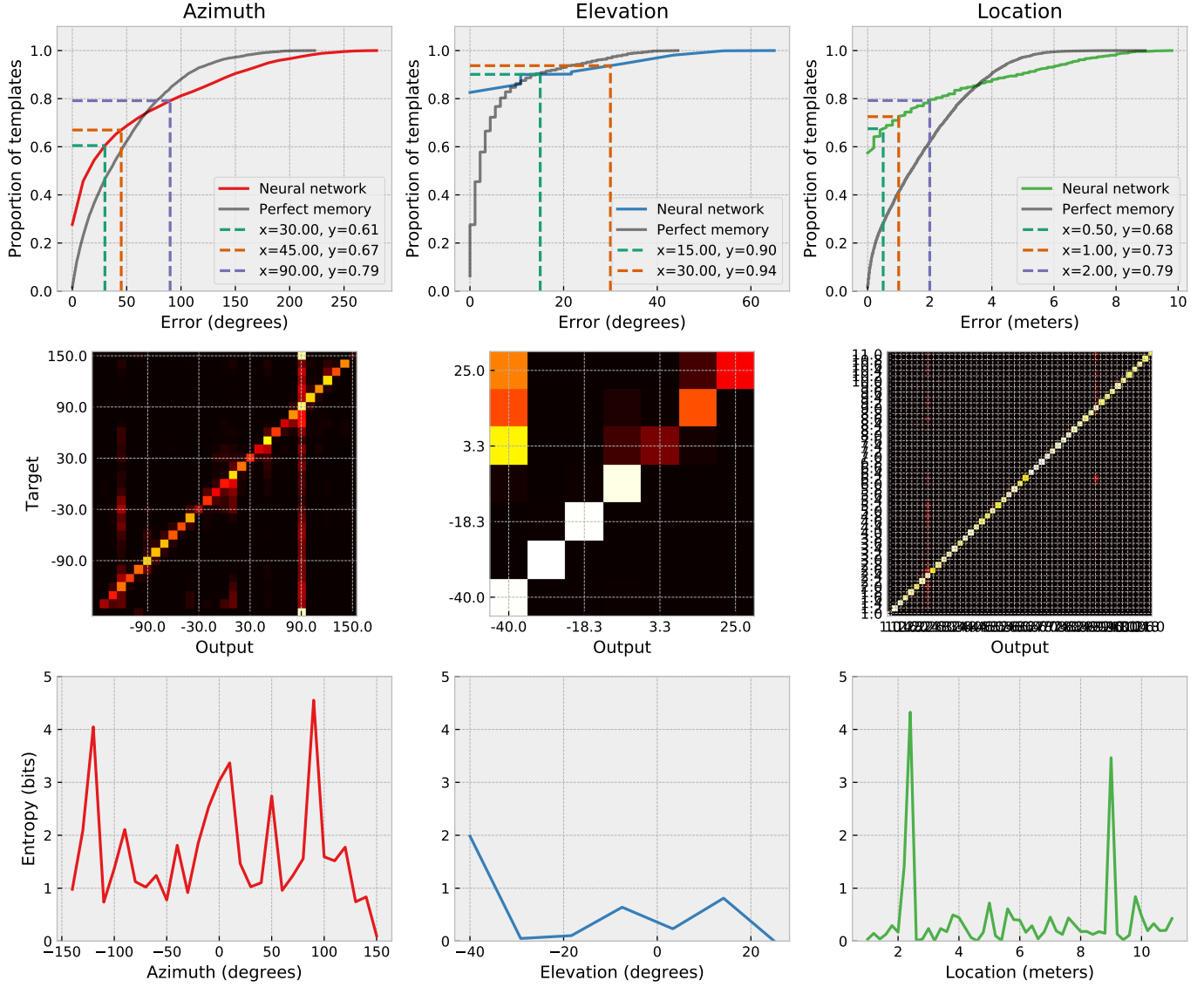


Figure 6: Classification results for the Israel site. Top row: cumulative error distributions for the azimuth, elevation and location classification. Middle row: Confusion matrices for the classifications. Bottom row: remaining classification entropy for each target (derived from the confusion matrices).

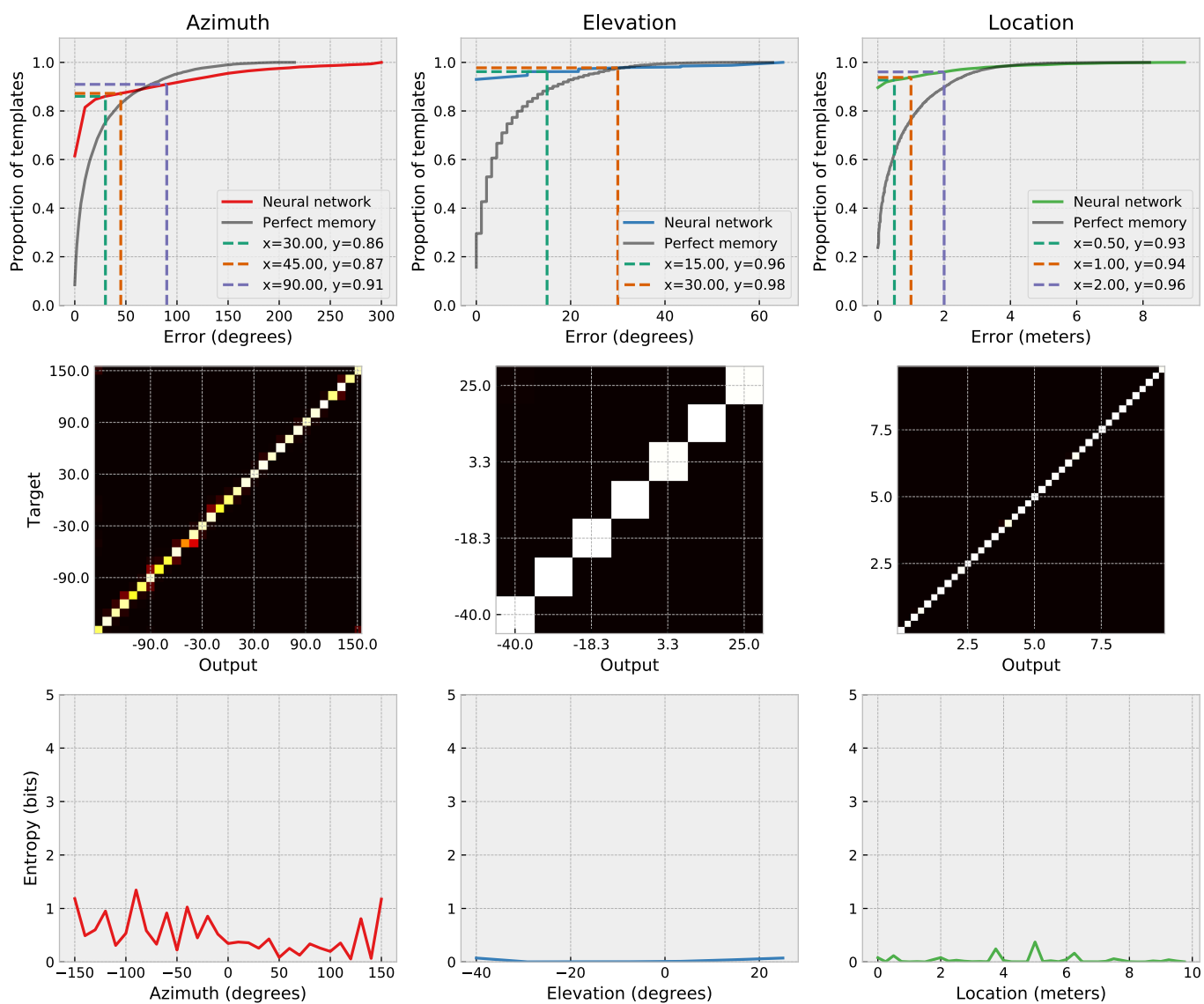


Figure 7: Similar as fig 6, but for the Royal Fort Park site.

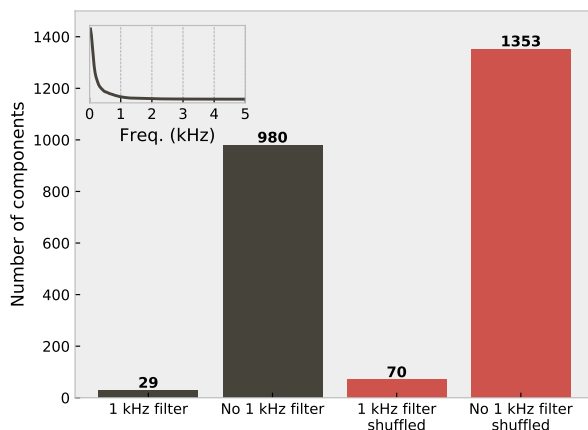


Figure 8:

4.3. Conclusion

From our current work, in combination with previous results (Vanderelst et al., 2016, 2015; Mansour et al., 2019), we conclude that the auditory periphery of bats maps complex echoes to a limited number of features. These features can be efficiently stored to recognize scenes. Moreover, we conjecture that a (binaural) version of the features should allow for sonar-based flight control. As such, a view of echolocating bats living in a low dimensional sensory space emerges.

References

- Baddeley, B., Graham, P., Husbands, P., & Philippides, A. (2012). A model of ant route navigation driven by scene familiarity. *PLoS computational biology*, 8, e1002336.
- Barchi, J. R., Knowles, J. M., & Simmons, J. A. (2013). Spatial memory and stereotypy of flight paths by big brown bats in cluttered surroundings. *Journal of Experimental Biology*, 216, 1053–1063.
- Braitenberg, V., & Schüz, A. (2013). *Cortex: statistics and geometry of neuronal connectivity*. Springer Science & Business Media.
- Clare, E. L., & Holderied, M. W. (2015). Acoustic shadows help gleanings bats find prey, but may be defeated by prey acoustic camouflage on rough surfaces. *Elife*, 4, e07404.
- Dau, T., Püschel, D., & Kohlrausch, A. (1996). A quantitative model of the “effective” signal processing in the auditory system. i. model structure. *The Journal of the Acoustical Society of America*, 99, 3615–3622.
- Franz, M. O., & Mallot, H. A. (2000). Biomimetic robot navigation. *Robotics and autonomous Systems*, 30, 133–153.
- Geberl, C., Kugler, K., & Wiegrebe, L. (2019). The Spatial Resolution of Bat Biosonar Quantified with a Visual-Resolution Paradigm. *Current Biology*, (p. S0960982219304750).
- Ghotra, M. S., & Dua, R. (2017). *Neural Network Programming with TensorFlow: Unleash the power of TensorFlow to train efficient neural networks*. Packt Publishing Ltd.
- Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing research*, 47, 103–138.
- Grunwald, J.-E., Schörnich, S., & Wiegrebe, L. (2004). Classification of natural textures in echolocation. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 5670–5674.
- Hawkins, D. M. (2004). The Problem of Overfitting. *Journal of Chemical Information and Computer Sciences*, 44, 1–12.
- Jakobsen, L., Ratcliffe, J. M., & Surlykke, A. (2012). Convergent acoustic field of view in echolocating bats. *Nature*, 493, 93–96.
- Jensen, M. E., Moss, C. F., & Surlykke, A. (2005). Echolocating bats can use acoustic landmarks for spatial orientation. *Journal of experimental biology*, 208, 4399–4410.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, .
- Kuc, R. (1997a). Biomimetic sonar recognizes objects using binaural information. *The Journal of the Acoustical Society of America*, 102, 689–696.
- Kuc, R. B. (1997b). Biomimetic sonar differentiates coin head from tail. *The Journal of the Acoustical Society of America*, 101, 3198.
- Lee, W.-J., Falk, B., Chiu, C., Krishnan, A., Arbour, J. H., & Moss, C. F. (2017). Tongue-driven sonar beam steering by a lingual-echolocating fruit bat. *PLoS biology*, 15, e2003148.
- Mansour, C. B., Koreman, E., Laurijssen, D., Steckel, J., Peremans, H., & Vanderelst, D. (2019). Robotic models of obstacle avoidance in bats. In *The 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)* (pp. 463–464). MIT Press.
- McLeod, P., Plunkett, K., & Rolls, E. T. (1998). *Introduction to connectionist modelling of cognitive processes*. Oxford University Press.
- Meddis, R. (2006). Auditory-nerve first-spike latency and auditory absolute threshold: a computer model. *The Journal of the Acoustical Society of America*, 119, 406–417.
- Moss, C. F., & Surlykke, a. (2001). Auditory scene analysis by echolocation in bats. *The Journal of the Acoustical Society of America*, 110, 2207–2226.
- Reijnen, J., & Peremans, H. (2010). On population encoding and decoding of auditory information for bat echolocation. *Biological Cybernetics*, 102, 311–326.
- Reijnen, J., Vanderelst, D., & Peremans, H. (2010). Morphology-induced information transfer in bat sonar. *Physical review letters*, 105, 148701.
- Schnitzler, H.-U., Moss, C. F., & Denzinger, A. (2003). From spatial orientation to food acquisition in echolocating bats. *Trends in Ecology & Evolution*, 18, 386–394.
- Simmons, J. A. (2012). Bats use a neuronally implemented computational acoustic model to form sonar images. *Current Opinion in Neurobiology*, 22, 311–319.
- Simmons, J. A., Freedman, E. G., Stevenson, S. B., Chen, L., & Wohlgenant, T. J. (1989). Clutter interference and the integration time of echoes in the echolocating bat, *Eptesicus fuscus*. *The Journal of the Acoustical Society of America*, 86, 1318–1332.
- Steckel, J., Boen, A., & Peremans, H. (2013). Broadband 3-D sonar system using a sparse array for indoor navigation. *IEEE Transactions on Robotics*, 29, 161–171.
- Steckel, J., & Peremans, H. (2013). Batslam: Simultaneous localization and mapping using biomimetic sonar. *PLoS one*, 8, e54076.
- Surlykke, A., & Bojesen, O. (1996). Integration time for short broadband clicks in echolocating fm-bats (*Eptesicus fuscus*). *Journal of Comparative Physiology A*, 178, 235–241.
- Surlykke, A., Simmons, J. A., & Moss, C. F. (2016). Perceiving the world through echolocation and vision. In *Bat Bioacoustics* (pp. 265–288). Springer.
- Ulanovsky, N., & Moss, C. F. (2008). What the bat’s voice tells the bat’s brain. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 8491–8498.
- Vanderelst, D. (2015). Social network analysis as a tool for research policy. *Plos Neglected Tropical Diseases*, . In press.
- Vanderelst, D., De Mey, F., Peremans, H., Geipel, I., Kalko, E., & Firzlaff, U. (2010). What noseleaves do for fm bats depends on their degree of sensorial specialization. *PLOS ONE*, 5, e11893.
- Vanderelst, D., & Peremans, H. (2017). A computational model of mapping in echolocating bats. *Animal Behaviour*, 131, 73–88.
- Vanderelst, D., Peremans, H., & Holderied, M. (2015). Sensorimotor model of obstacle avoidance in echolocating bats. *PLoS Computational Biology*, . Paper accepted pending minor revisions.
- Vanderelst, D., Steckel, J., Boen, A., Peremans, H., & Holderied,

- M. W. (2016). Place recognition using batlike sonar. *Elife*, 5, e14188.
- Verboom, B., Boonman, A. M., & Limpens, H. J. (1999). Acoustic perception of landscape elements by the pond bat (*myotis dasycneme*). *Journal of Zoology*, 248, 59–66.
- Von Helversen, O., & Winter, Y. (2005). Bat ecology. chapter Glossophagine Bats and Their Flowers: Costs and Benefits for Plants and Pollinators. (p. 346). University of Chicago Press.
- Warnecke, M., Macías, S., Falk, B., & Moss, C. F. (2018). Echo interval and not echo intensity drives bat flight behavior in structured corridors. *The Journal of Experimental Biology*, 221, jeb191155.
- Wiegrefe, L. (2008). An autocorrelation model of bat sonar. *Biological cybernetics*, 98, 587–595.
- Wiegrefe, L., & Schmidt, S. (1996). Temporal integration in the echolocating bat, *megaderma lyra*. *Hearing research*, 102, 35–42.
- Williams, R. W. (2000). Mapping genes that modulate mouse brain development: a quantitative genetic approach. In *Mouse brain development* (pp. 21–49). Springer.
- Yovel, Y., Stilz, P., Franz, M. O., Boonman, A., & Schnitzler, H. U. (2009). What a plant sounds like: The statistics of vegetation echoes as received by echolocating bats. *PLoS Computational Biology*, 5.
- Yu, C., Luo, J., Wohlgemuth, M., & Moss, C. F. (2019). Echolocating bats inspect and discriminate landmark features to guide navigation. *The Journal of Experimental Biology*, (p. jeb.191965).

Appendix A. Supplementary material

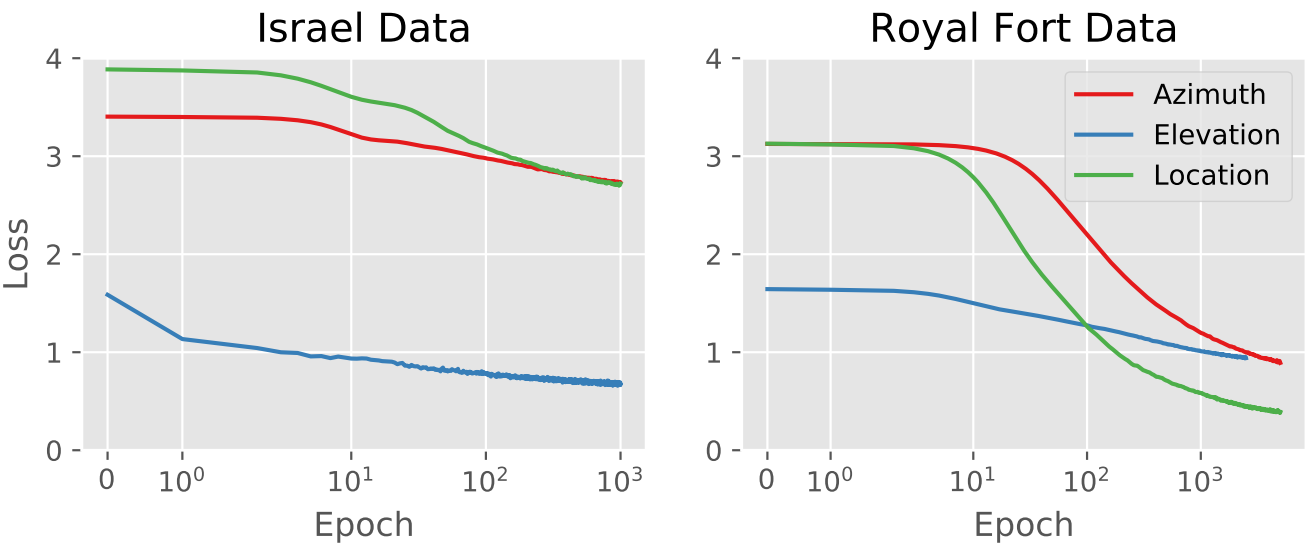


Figure A.9: Training history of the neural networks. The plots give the value of the loss function as a function of training epoch. Notice the logarithmic x-axis.

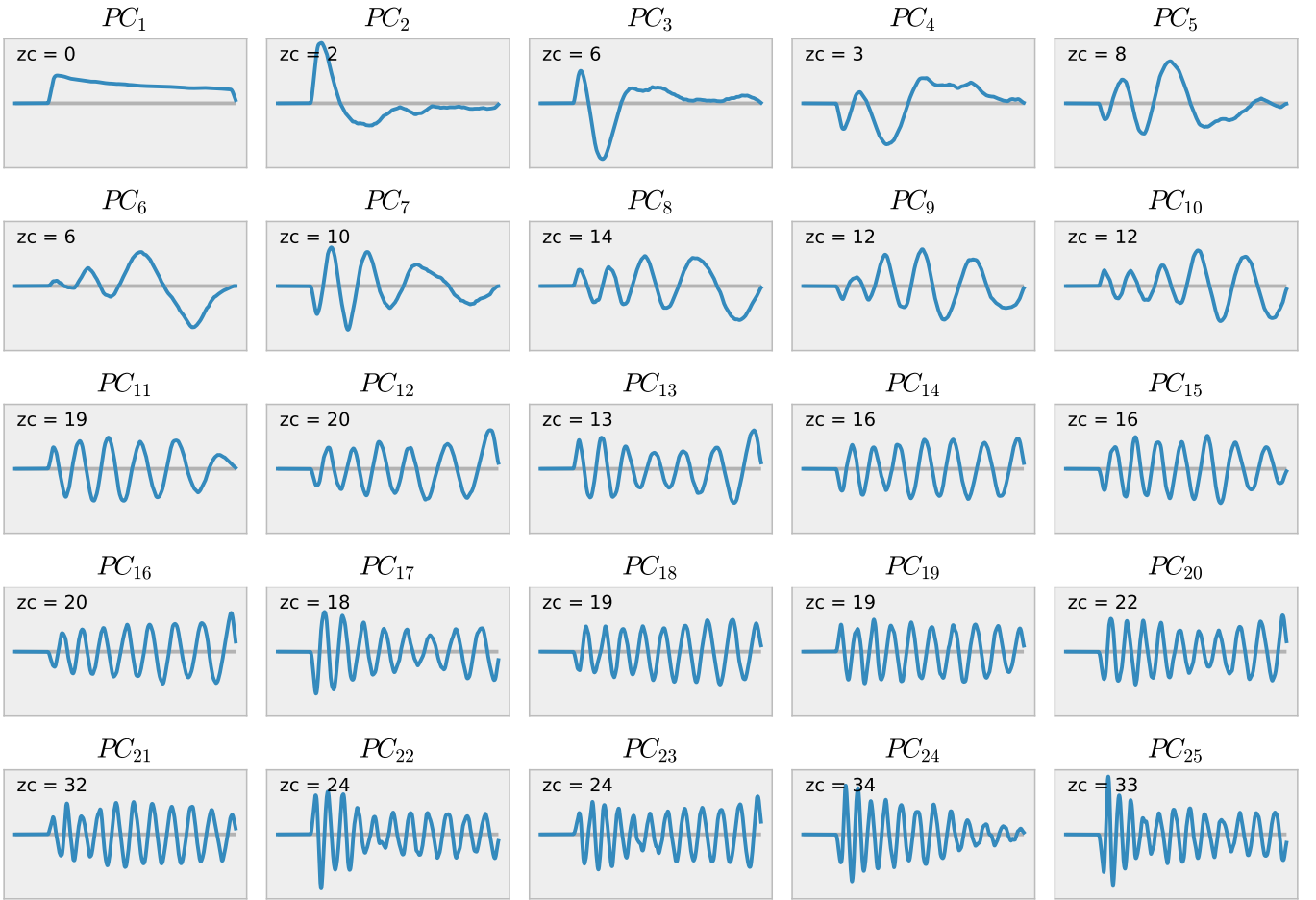


Figure A.10: