A TUTORIAL ON OPTIMAL MULTI-SENSORY DECISION-MAKING

JAN DRUGOWITSCH

Contents

1. Introduction

The aim of this tutorial is to re-investigate some of the observations made by Drugowitsch et al. (2014) about decision making in the presence of evidence from multiple cues. This includes testing cue combination in a reaction time task using standard Bayesian cue combination tests, deriving optimal evidence accumulation for multiple cues present, and building diffusion models that aim at re-capturing the observed psychometric and chronometric curves.

1.1. **The task.** Drugowitsch et al. (2014) had human decision makers sit on a motion platform in front of a screen, and decide as quickly and as accurately as possible if they were moving to the left or right from straight-ahead. The decision makers received evidence about self-motion either from motion of the platform (vestibular condition), from a random dot flow field presented on the screen (visual condition), or from both cues simultaneously (and congruently; combined condition). Self-motion had a Gaussian velocity profile, with motion velocity initially increasing, peaking after 1s, and decreasing until stimulus offset (see `va_profile.m` for details; Fig. 2). The maximum stimulus duration was 2s, and the maximum distance travelled was 30cm. Decision makers could respond at any time after stimulus onset. In the vestibular and the combined condition, platform motion was stopped at response, at which point the platform returned to its initial location. In the visual condition, the visual stimulus was turned off at response. Therefore, the task was a 2-alternative forced choice (2AFC) reaction time task.

1.2. **Factors influencing evidence reliability.** The reliability about self-motion information was modulated in four ways. First, the visual flow field consisted of a set of dots in virtual 3D space that the decision maker moved through. Reliability of the flow field information was reduced by, in each frame, replacing a subset of these dots randomly. The fraction of dots moving coherently, called the *coherence*, remained constant within trials, but varied across trials, taking values out of $\{0.25, 0.37, 0.7\}$[1]. Second, the heading direction did not modulate the reliability of self-motion information itself, but the task-dependent information about if self-motion was towards the left or the right of straight-ahead. Choosing $h = 0°$ as straight-ahead, large heading magnitudes $|h|$ implied more information towards the requested decision. Third, physical limitations of the experimental setup required time-varying velocity and acceleration profiles. This additionally modulated reliability of self-motion information, as low velocity implies little information from the visual modality, and high velocity implies lots of information. Forth, the presence of multiple cues provides more information about heading direction than if only a single cue were present. All of these factors will come into play when formulating a model for how to optimally accumulate evidence over time and across cues.

[1] Additional coherences were used in a second experiment, but they were excluded from the dataset.

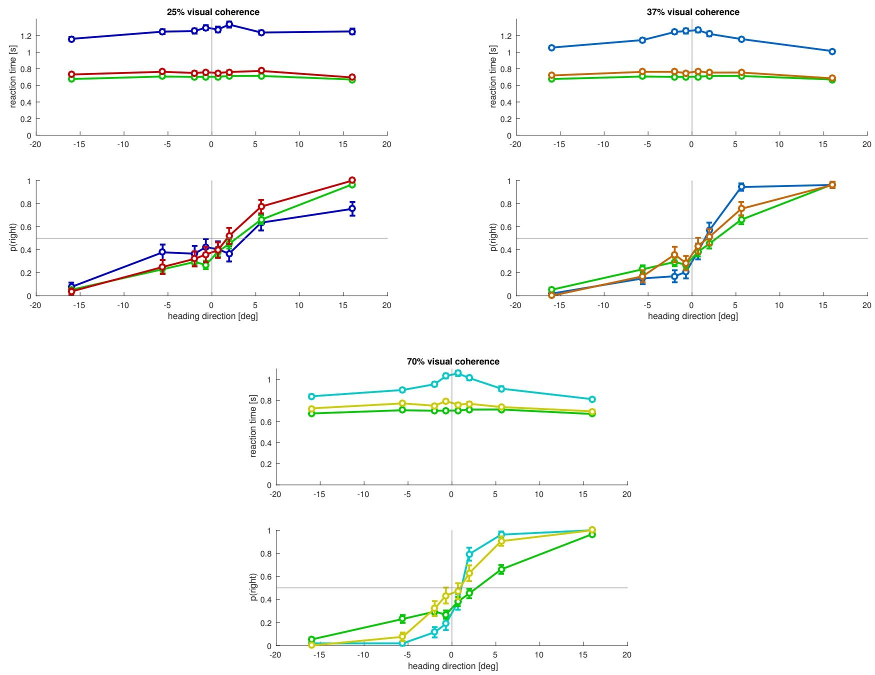

FIGURE 1. Example output of `plot_psych_chron.m`. The data is for subject A.

1.3. **The data.** The dataset consists of the trial-by-trial data of seven human decision makers performing the task. Each file, one per subject, contains a set of vectors, all of the same length, that characterize the condition in each trial and the associated response. The vectors are

- `oris` the heading direction in each trial, in degrees. Negative values are motion to the left, positive values are motion to the right from straight-ahead.
- `mods` modalities present. 1 = visual only, 2 = vestibular only, 3 = both visual/vestibular information present.
- `cohs` coherence of the visual modality, one of $\{0.25, 0.37, 0.7\}$. Set to zero for vestibular-only trials.
- `choice` decision made in each trial. 0 = left, 1 = right.
- `rt` reaction time, giving time from stimulus onset to response in seconds.

See `plot_psych_chron.m` for how to use these data to plot psychometric and chronometric curves for each condition, and Fig. 1 for an example output.

2. What you should do

What follows is a list of things that you should do or attempt to do. Even though the derivations are optional, and results are provided in the appendix, you should at least attempt going through them by yourself. Code templates are provided for most of the steps, except for the bonus questions.

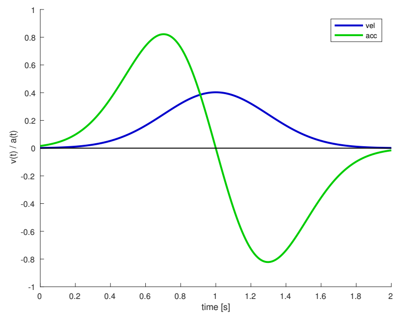

FIGURE 2. Output of `plot_va_profile.m`, showing the velocity and acceleration profile.

2.1. **Get familiar with the data and behavior.** Load the datasets and get familiar with the behavior in multiple different ways. The script `plot_psych_chron.m` provides one way of doing so, by plotting the psychometric and chronometric curves (see Fig. 1). Look through the script and make sure that you understand what it does, and how it achieves it. Feel free to also go beyond the analysis that the script provides, by, for example, looking at reaction time distributions through quantile plots or other means.

2.2. **Perform standard Bayesian cue combination analysis.** Bayesian cue combination analysis predicts the performance in the combined condition from the performance observed in the two unimodal conditions. It does so by fitting the psychometric curves in the unimodal conditions to estimate the heading direction estimation reliability of each of the two cues individually. These reliabilities are then used to predict the reliability when both cues are present. This prediction is in turn compared to the observed reliability in the combined condition, to assess if cue combination is Bayes-optimal.

Section 3 provides the theoretical rational for the analysis. `fit_cumul_gauss.m` allows fitting the reliability of individual conditions, and `test_cumul_gauss.m` provides an example for how to use this script. `test_standard_cue_comb.m` provides a template for performing the cue combination analysis, with critical sections to be completed.

Questions you should think of are: what do the results of this analysis tell us? What does this analysis ignore about the experimental setup? How could it be fixed?

2.3. **Bayes-optimal evidence accumulation for time-varying evidence reliability.** Optimal performance in the task requires accumulating evidence across time and – in the combined condition – across cues. We will first focus on evidence accumulation across time. The novel aspect here is that the reliability of the evidence varies across time. For the visual modality it is affected by the flow field's velocity, and for the vestibular modality it is affected by the platform's acceleration. This causes the optimal evidence accumulation strategy to differ from that arising from a time-invariant reliability.

In the first step, you should look at the setup discussed in Section 4, and use it to derive the Bayes-optimal strategy to accumulate evidence across time with time-varying evidence reliability, as described in Section 4.2. A step-by-step derivation

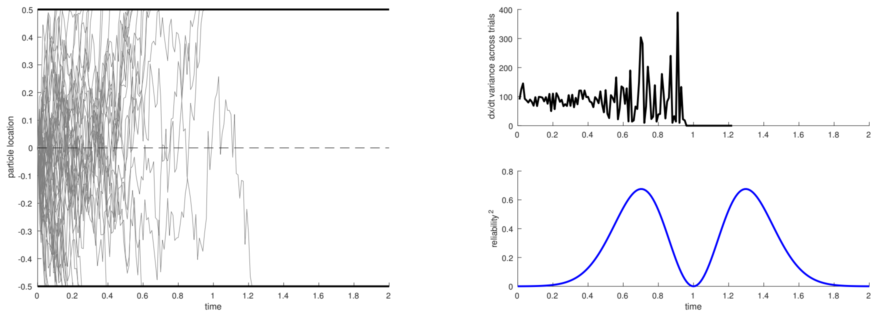

FIGURE 3. Default output of `sim_weighted_diffusion.m`. The left panel shows example traces of the diffusion model. The left panel shows the variance (across trials) of changes of the particle location over time, and the reliability time-course of the momentary evidence. By default, this script does not perform any reliability-weighting, as can be seen by the constant particle increment variance in the top right panel. This variance is noisy due to the limited number of trials that the script uses to estimate this variance.

is provided in Appendix A, in case you get stuck. Nonetheless, please try it for yourself, and consider the impact of a changing reliability on evidence accumulation. What happens if the information is highly unreliable? What happens if it is very reliable? What are the assumptions made about the reliability time-course? Is it assumed known?

In the second step, you should implement a diffusion model that optimally accumulates evidence across time with time-varying evidence reliability. Use the script `sim_weighted_diffusion.m` as a template. By default, this script does not perform any reliability-weighting, as can be seen by the constant diffusion variance in Fig. 3. Can you illustrate the impact of time-varying evidence reliability on evidence accumulation? What would you expect to happen to reaction time distributions?

2.4. **Bayes-optimal evidence accumulation across multiple cues.** The next step is to extend the model to accumulate evidence across multiple cues. We will first focus on assuming that the reliability does not change over time. This setup is discussed in Section 4.3. Try to first go through the derivation yourself, and then compare it to the results provided in Appendix A. The aim is to show that a single diffusion model is sufficient for evidence accumulation, even if the evidence is provided from multiple modalities. What happens if the reliability of the visual modality drops/increases?

The next step is to extend the model to the case that the evidence reliability of both modalities changes over time. This derivation is a bit more involved, and you can either attempt it by going through Section 4.4, or go straight to Appendix A to check the result. Please look at the interaction between reliability time-course and reliability, and how they impact evidence accumulation. Try to predict what happens if the overal reliabilities of the visual or vestibular modality changes.

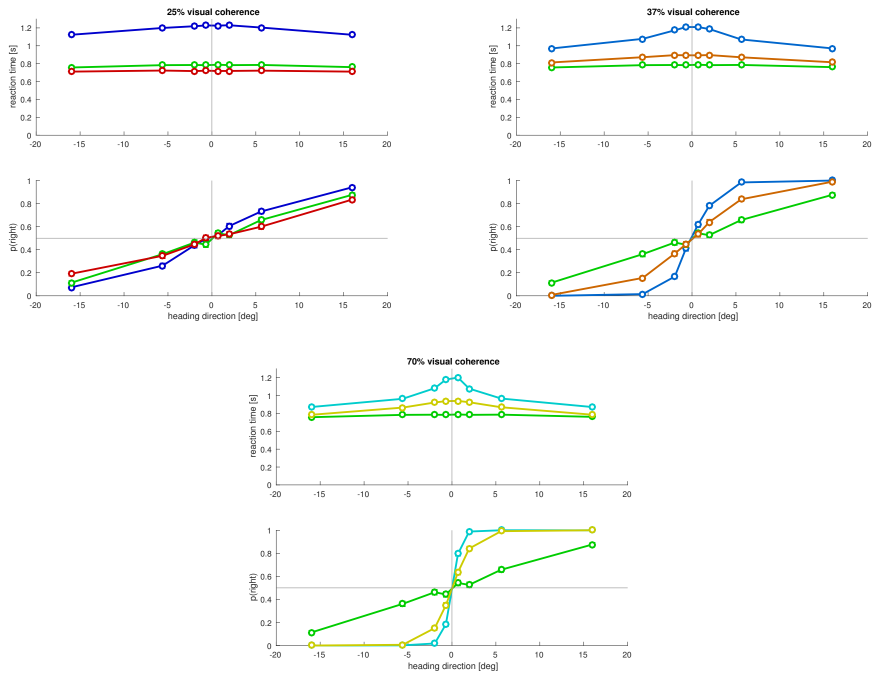

FIGURE 4. Example output of `sim_behavior.m`, once completed and with tuned parameters.

At last, try to modify `sim_weighted_diffusion.m` to combine information from both the visual and the vestibular modality, within a single diffusion model. This will require you to modify the expressions for `k` and `dn` in that script. Change the reliabilities `kvis` and `kvest` of the visual and vestibular modalities to see if your intuitions of a change in these reliabilities was correct.

2.5. **Simulating behavior in a virtual experiment.** Let's put all of the above together by simulating behavior for different modalities and coherences and looking at the resulting psychometric and chronometric curves. To do so, inspect `sim_behavior.m` that provides a framework to simulate choices and reaction times for the same conditions as used in the human subject experiment. The script is incomplete, and you should complete the missing sections, as outlined in the script's comments.

By default, the script's parameters don't replicate the qualitative features of human behavior. In particular, the psychometric curve in the combined condition is usually steeper than that in the visual condition, contrary to what has been observed in experiments. Can you adjust the parameters to more closely match human behavior? To do so, consider the different effects of drift and boundary change on the reaction times and choice probabilities in diffusion models. See Fig. 4 for an example simulation with tuned parameters.

2.6. **Bonus: refining the simulated behavior.** Currently, `sim_behavior.m` cannot capture all details of human behavior. One of the most significant deviations is that the predicted reaction time for straight-ahead motion does not depend on visual coherence, contrary to what is seen in the data. This prediction arises from the fact that the mean of the momentary evidence is proportional to $\sin(h)$, where h is heading direction, which is zero for straight-ahead motion, irrespective of the coherence of the visual modality (which comes in as another multiplicative factor $k_{vis}(c)$).

One way to achieve a modulation of reaction time by visual coherence even for straight-ahead motion is to have this coherence affect the variance of the momentary evidence. This is a reasonable assumption, given that lowering the coherence causes random dot replacement in the visual stimulus without lowering the saliency of the visual display itself. Two steps are required to introduce such a coherence-modulated variance into `sim_behavior.m`. First, the diffusion model simulation function within that script needs to take an additional argument that specifies the diffusion model variance. Second, an additional parameter needs to specify this variance for different coherences, and the diffusion model simulation needs to be called with the appropriate parameter, dependent on coherence.

With these modifications, try to adjust parameters to achieve a qualitative match to human behavior, where, in the visual-only condition, reaction times are seen to generally decrease with increasing coherence.

2.7. **Bonus: optimal decision boundaries for time-varying evidence reliability.** So far we have focused on optimal evidence accumulation, but modeled the speed/accuracy trade-off with a simple constant boundary on the accumulated evidence. This ignores that such a time-variant boundary might not be optimal in the sense that it might not maximize the expected reward per decision. Finding the optimal decision boundary requires extending the dynamic programming approach to finding optimal decision boundaries to the case of time-varying evidence reliability. Section 5 provides an overview over a possible way for how to approach this.

What are the features of the optimal decision boundary? What kind of behavior would one expect following these boundaries? Does this match what is seen in human behavior?

3. Standard cue combination analysis

A standard assumption in the Bayesian analysis of cue combination is that both cues provide independent evidence with a Gaussian likelihood. Based on this assumption, the expected performance in the combined condition, in which both cues are simultaneously present, follows from combining both likelihoods to form a posterior upon which the decision is based (e.g., Ernst and Banks, 2002; Körding and Wolpert, 2004).

3.1. **Estimating evidence reliability from the psychometric curve.** Let us focus on the vestibular-only condition. If we denote the true heading direction to be h, then we assume that our vestibular organ provides information x_{vest} about h with some additive Gaussian noise, such that

$$p(x_{vest}|h) = \mathcal{N}\left(x_{vest}|h, \sigma^2_{vest}\right), \tag{1}$$

which is a Gaussian likelihood with mean h and variance σ^2_{vest}. Here, σ^2_{vest} controls how much x_{vest} tells us about h, which lower variance σ^2_{vest} meaning that x_{vest} is more informative. Let us now assume that, a priori, we have a uniform prior belief about the true h, that is $p(h) \propto 1$. In this case, the posterior over $\hat{h}$ (adding a $\hat{\cdot}$ to indicate that this is the decision maker's belief) given the observed x_{vest} becomes

$$p(\hat{h}|x_{vest}) \propto p(x_{vest}|\hat{h}) \propto \mathcal{N}\left(\hat{h}|x_{vest}, \sigma_{vest}^2\right). \tag{2}$$

The above would be the estimate within a single trial. However, the experimenter does not have access to x_{vest} within individual trials. Instead, we can average across multiple trials with the same h, to find the marginal posterior

$$\begin{aligned} p(\hat{h}|h) &= \int p(\hat{h}|x_{vest})p(x_{vest}|h)\mathrm{d}x_{vest} \\ &= \int \mathcal{N}\left(\hat{h}|x_{vest}, \sigma_{vest}^2\right)\mathcal{N}\left(x_{vest}|h, \sigma_{vest}^2\right)\mathrm{d}x_{vest} \\ &= \mathcal{N}\left(\hat{h}|h, 2\sigma_{vest}^2\right). \end{aligned} \tag{3}$$

Here, the variance is doubled, as it reflects on one hand the variability of x_{vest} across trials, and on the other hand the decision maker's uncertainty about $\hat{h}$ upon observing x_{vest}.

The task is not to identify the heading direction h, but to decide if it is left or right from straight-ahead, that is, if $h < 0$ or $h \geq 0$. Assuming that the decision maker chooses "right" if $\hat{h} \geq 0$, the probability of choosing "right" for a given heading direction h is thus given by

$$p(\hat{h} \geq 0|h) = \int_0^\infty p(\hat{h} = x|h)\mathrm{d}x = \Phi\left(\frac{h}{\sqrt{2}\sigma_{vest}}\right), \tag{4}$$

where $\Phi(a) = \int_{-\infty}^a \mathcal{N}(b|0,1)\mathrm{d}b$ is the cumulative distribution function of a standard Gaussian. Therefore, this model predicts that increasing h should increase the probability of choosing "right" in a way that follows the Gaussian cumulative function. In other words, the psychometric curves that plots this probability over different heading directions should have this functional form. Furthermore, the only free parameters (ignoring choice biases, etc.) in this function is σ_{vest}. Therefore, we can estimate σ_{vest} by fitting this function to the psychometric curve. See `fit_cumul_gauss.m` and `test_fit_cumul_gauss.m` for scripts that perform this fit.

3.2. **Bayes-optimal cue combination.** In the presence of both cues, we assume that the decision maker in each trial observes both x_{vest} and x_{vist} that relate to the true heading direction h through their likelihoods

$$p(x_{vest}|h) = \mathcal{N}\left(x_{vest}|h, \sigma_{vest}^2\right), \qquad p(x_{vis}|h, c) = \mathcal{N}\left(x_{vis}|h, \sigma_{vis}^2(c)\right), \tag{5}$$

where c is the coherence of the visual stimulus, and the reliability $\sigma_{vis}^2(c)$ is modulated by this coherence. In particular, the likelihood variance $\sigma_{vis}^2(c)$ is expected to increase with decreasing coherence, indicating that the visual stimulus provides less reliable information for lower coherences.

Assuming as before an uninformative prior $p(c) \propto 1$, the posterior heading direction combines the information in both cues, and is given by

$$\begin{aligned} p(\hat{h}|x_{vis}, x_{vest}, c) &\propto p(x_{vis}|\hat{h}, c) p(x_{vest}|\hat{h}) \\ &= \mathcal{N}\left(x_{vis}|\hat{h}, \sigma^2_{vis}(c)\right) \mathcal{N}\left(x_{vest}|\hat{h}, \sigma^2_{vest}\right) \\ &\propto \mathcal{N}\left(\hat{h}|\frac{\sigma^2_{vest}}{\sigma^2_{vis}(c)+\sigma^2_{vest}} x_{vis} + \frac{\sigma^2_{vis}(c)}{\sigma^2_{vis}(c)+\sigma^2_{vest}} x_{vest}, \frac{\sigma^2_{vis}(c)\sigma^2_{vest}}{\sigma^2_{vis}(c)+\sigma^2_{vest}}\right). \end{aligned} \tag{6}$$

The mean of the posterior is a weighted combination of x_{vis} and x_{vest} with weights proportional to their respective precisions (i.e., inverse variances) $1/\sigma^2_{vis}(c)$ and $1/\sigma^2_{vest}$. This means that information is combined by weighting it according to its reliability. The posterior variance can be shown to be smaller than (or at most the same as) the individual variances, indicating that combining information across multiple cues increases the decision maker's certainty.

As before, the experimenter doesn't directly observe the x_{vis}'s and x_{vest}'s in individual trials. However, we can again average across trials to find

$$\begin{aligned} p(\hat{h}|h, c) &= \iint p(\hat{h}|x_{vis}, x_{vest}, c) p(x_{vis}|h, c) p(x_{vest}|h) \mathrm{d}x_{vis} \mathrm{d}x_{vest} \\ &= \mathcal{N}\left(\hat{h}|h, 2\frac{\sigma^2_{vis}(c)\sigma^2_{vest}}{\sigma^2_{vis}(c)+\sigma_{vest}}\right), \end{aligned} \tag{7}$$

resulting, as before, in twice the variance due to a combination of stochastic averaging and the decision-maker's uncertainty. From the above, we find the predicted probability of choosing "right" to be given by

$$p\left(\hat{h} \geq 0|h, c\right) = \Phi\left(\frac{h}{\sqrt{2}\sigma_{comb,pred}(c)}\right), \tag{8}$$

with

$$\sigma^2_{comb,pred}(c) = \frac{\sigma^2_{vis}(c)\sigma^2_{vest}}{\sigma^2_{vis}(c)+\sigma_{vest}}. \tag{9}$$

This provides the prediction for the psychometric curve in the combined condition, based on the fitted psychometric curves in the unimodal conditions.

4. Accumulating evidence across time from multiple cues

The previous section has assumed that all heading direction evidence observed in a single trial is for each cue encapsulated in a single observation x, while ignoring the temporal dynamics of this evidence. Here, we re-formulate the problem to take into account this temporal dynamics. Let us first focus on evidence from a single modality. We will late expand the framework to multiple cues.

4.1. A single cue with time-invariant evidence reliability.

Let us start by assuming an evidence reliability that is constant over time. In this case, the informativeness of the momentary evidence is modulated by heading direction h and coherence c (for the visual modality). As the task is to identify if the heading is towards the left or right of straight-ahead, we only care about the projection of the heading direction orthogonal to straight-ahead, given by $z(h) = \sin(h)$.

Recall that for diffusion models applied to the random dot motion task, the momentary evidence δx in some small time δt can be assumed to have likelihood

$p(\delta x|\mu) = \mathcal{N}(\delta x|\mu\delta t, \delta t)$. In this case, the magnitude of μ modulates the evidence reliability. Translated to the heading direction discrimination task, and including all the components that modulate reliability, the likelihood will be assumed to be given by

$$p(\delta x|h, c) = \mathcal{N}(\delta x|z(h)k(c)\delta t, \delta t), \tag{10}$$

where $k(c)$ is a free parameter that described how reliability changes with a change in coherence.

With the above likelihood, we can ask how to optimally accumulate evidence over time. Our aim is to find the posterior $p(z(h) \geq 0|\delta x_{1:n}, c)$, which is the belief that the heading direction is right from straight ahead, given momentary evidence $\delta x_1, \dots, \delta x_n$. To do so, we will assume a uniform prior over $z(h)$, that is $p(z(h)) \propto 1$, and find $p(z(h)|\delta x_{1:n}, c)$ by Bayes' rule,

$$\begin{aligned} p(z(h)|\delta x_{1:n}, c) &\propto \prod_{j=1}^{n} \mathcal{N}(\delta x_j|z(h)k(c)\delta t, \delta t) \\ &\propto e^{-\sum_{j=1}^{n} \frac{(\delta x_j - z(h)k(c)\delta t)^2}{2\delta t}} \\ &\propto e^{-\frac{z(h)^2 k(c)^2 \sum_{j=1}^{n} \delta t}{2} + z(h)k(c)\sum_{j=1}^{n}\delta x_j} \\ &= e^{-\frac{z(h)^2 k(c)^2 t}{2} + z(h)k(c)x(t)} \\ &\propto \mathcal{N}\left(z(h)|\frac{x(t)}{k(c)t}, \frac{1}{k(c)^2 t}\right), \end{aligned} \tag{11}$$

where we have used $\sum_{j=1}^{n} \delta t = t$ and $\sum_{j=1}^{n} \delta x_j = x(t)$. To move from the fourth to the last line, we have used

$$\mathcal{N}(x|\mu, \sigma^2) \propto e^{-\frac{1}{2\sigma^2}x^2 + \frac{\mu}{\sigma^2}x}, \tag{12}$$

which allows us to directly read off the inverse variance $1/\sigma^2$ and mean-to-variance ratio μ/σ^2, from which we can find the mean and variance of the posterior. The expression for the posterior shows that it is thus sufficient to track $x(t)$ and t rather than the whole "particle" trajectory $\delta x_{1:n}$. From the above we find $p(z(h) \geq 0|x(t), t)$ by

$$p(z(h) \geq 0|x(t), t) = \int_0^\infty p(z(h)|x(t), t, c)\mathrm{d}z(h) \approx \Phi\left(\frac{x(t)\sqrt{k(c)^2 t}}{k(c)t}\right) = \Phi\left(\frac{x(t)}{\sqrt{t}}\right), \tag{13}$$

where the integral is approximate as $z(h)$ (true value, not noisy estimate) is bounded by $z(h) \in [-1, 1]$. To solve the integral, we have used $\int_0^\infty \mathcal{N}(x|\mu, \sigma^2)\,\mathrm{d}x = \Phi(\mu/\sigma)$. This demonstrates a posterior belief about "right" being correct that increases with $x(t)$, but for the same $x(t)$ drops over time. At any point in time, the decision-maker would choose "right" if $p(z(h) \geq 0|x(t), t) \geq 1/2$, and "left" otherwise.

4.2. **A single cue with time-dependent evidence reliability.** Let us continue by assuming that evidence reliability is modulated by some time-varying signal. Focusing for now on the visual modality, this signal is assumed to be velocity $v(t)$. To include this time-varying reliability, we change the momentary evidence likelihood to be given by

$$p(\delta x_n|z(h), c) = \mathcal{N}(\delta x_n|z(h)v_n k(c)\delta t, \delta t), \tag{14}$$

where $v_n = v(n\delta t)$ is the velocity in the nth time step.

Given the momentary evidence likelihood, the posterior over the heading direction being "right" of straight-ahead is found in two steps, as before:

(1) We need to find the posterior $p(z(h)|\delta x_{1:n}, c)$ given n momentary evidence time-steps. Proceed as in the previous section to find this posterior, where you should use $V(t) = \sum_{j=1}^{n} v_j^2 \delta t$ and $x_v(t) = \sum_{j=1}^{n} v_j \delta x_j$. Make sure that, if $v_n = 1$ for all n, you recover the posterior from the previous section.

(2) In the second step, we need to find $p\left(z(h) \geq 0 | x_v(t), V(t), c\right)$ from the the posterior $p\left(z(h)|x_v(t), V(t), c\right)$. This is again done by integrating this posterior from 0 to ∞, using $\int_0^\infty \mathcal{N}\left(x|\mu, \sigma^2\right) \mathrm{d}x = \Phi(\mu/\sigma)$. This will lead to an expression for the belief that the heading direction was "right", given $x_v(t)$, $V(t)$, and c.

You should attempt the full derivation by yourself, before checking the result provided in Appendix A.

4.3. **Multiple cues with time-invariant evidence reliability.** As a next step, let us focus on combining information from multiple cues, whose reliability is assumed to remain constant within a trial. Specifically, we will assume visual evidence $\delta x_{vis,1}, \delta x_{vis,2}, \ldots$ and vestibular evidence $\delta x_{vest,1}, \delta x_{vest,2}, \ldots$, with likelihoods

$$\begin{aligned} p(\delta x_{vis,n}|z(h), c) &= \mathcal{N}\left(\delta x_{vis,n}|z(h)k_{vis}(c)\delta t, \delta t\right), \\ p(\delta x_{vest,n}|z(h)) &= \mathcal{N}\left(\delta x_{vest,n}|z(h)k_{vest}\delta t, \delta t\right), \end{aligned} \tag{15}$$

where only the reliability of the visual modality, $k_{vis}(c)$, depends on coherence c.

Given the momentary evidence likelihoods, the posterior over the heading direction being "right" of straight-ahead can be found using the same steps as before:

(1) We fist need to find the posterior $p(z(h)|\delta x_{vis,1:n}, \delta x_{vest,1:n}, c)$ given n momentary evidence time-steps. This posterior can again be found by the use of Bayes' rule,

$$p(z(h)|\delta x_{vis,1:n}, \delta x_{vest,1:n}, c) \propto \prod_{j=1}^{n} p(\delta x_{vis,j}|z(h), c) p(\delta x_{vest,j}|z(h)), \tag{16}$$

this time combining both likelihoods, which we assume to be independent, conditional on $z(h)$. Use $\sum_{j=1}^{n} \delta t = t$, $\sum_{j=1}^{n} \delta x_{vis,j} = x_{vis}(t)$ and $\sum_{j=1}^{n} \delta x_{vest,j} = x_{vest}(t)$. You will find this posterior computation to be implementable by a weighted combination of two diffusion models. What are the weights? Try using $k_{comb}(c)^2 = k_{vis}(c)^2 + k_{vest}^2$ to find them. Can the weights be related to standard Bayesian cue combination principles, where $1/\sigma_{comb}(c)^2 = 1/\sigma_{vis}(c)^2 + 1/\sigma_{vest}^2$?

(2) Two diffusion models imply the use of two evidence accumulators to perform this task. Is it possible to combined the momentary evidence $\delta x_{vis,n}$ and $\delta x_{vest,n}$ in some way to use a single diffusion model for this accumulation? Would using the same combination weights to combine the diffusion models also work to combine the momentary evidence? You can check this by computing the posterior with the likelihood of the combined momentary evidence, to see if this posterior matches that arising from the weighted combination of diffusion models.

(3) The last step involved finding the posterior belief of the heading direction being "right" from straight-ahead, $p(z(h) \geq 0|x_{vis}(t), x_{vest}(t), c)$ from the

previously found posterior, using again $\int_0^\infty \mathcal{N}\left(x|\mu,\sigma^2\right) \mathrm{d}x = \Phi(\mu/\sigma)$. Find this posterior belief, and express it again as a combination of two diffusion models, or as a single diffusion model.

You should attempt the full derivation by yourself, before checking the result provided in Appendix A.

4.4. **Bonus: multiple cues with time-dependent evidence reliability.** Let us at last consider the full setup, in which the visual reliability is modulated by heading velocity, and the vestibular reliability is modulated by acceleration. In this case, the two momentary likelihoods are given by

$$\begin{aligned} p(\delta x_{vis,n}|z(h),c) &= \mathcal{N}\left(\delta x_{vis,n}|z(h)v_n k_{vis}(c)\delta t, \delta t\right), \\ p(\delta x_{vest,n}|z(h)) &= \mathcal{N}\left(\delta x_{vest,n}|z(h)a_n k_{vest}\delta t, \delta t\right), \end{aligned} \tag{17}$$

where $v_n = v(n\delta t)$ and $a_n = a(n\delta t)$ are velocity and acceleration at the time the momentary evidence is provided.

Finding the posterior(s) follows the same scheme as in the previous section:

(1) The posterior $p(z(h)|\delta x_{vis,1:n}, \delta x_{vest,1:n}, c)$ can again be found by Bayes' rule, where you should use $\sum_{j=1}^n v_j^2 \delta t = V(t)$, $\sum_{j=1}^n a_j^2 \delta t = A(t)$, and the accumulated evidence $\sum_{j=1}^n v_j \delta x_{vis,j} = x_{vis}(t)$ and $\sum_{j=1}^n a_j \delta x_{vest,j} = x_{vest}(t)$ with accumulation weighted by reliability over time. This leads to an expression that is a combination of the expressions found for time-varying evidence reliability, and that for multiple cues.

(2) Try to combine the two diffusing particle into a single particle / momentary evidence that results in the same posterior. This time, adequate re-weighting needs to be performed across both time and cues.

(3) Find the posterior belief for right-ward motion from the above posterior $z(h)$.

The results of the full derivation are provided in Appendix A .

5. The optimal decision boundary for time-varying evidence reliability

To find the optimal decision boundaries, we will assume that the decision maker knows about the modalities present and the coherence of the visual modality (if present), and is only ignorant about the heading direction. More specifically, we assume the momentary evidence to have likelihood

$$p(\delta x_n|z(h),k) = \mathcal{N}\left(\delta x_n|z(h)kd_n\delta t, \delta t\right), \tag{18}$$

where reliability k and reliability time-course $d_n = d(n\delta t)$ are known, and $z(h)$ needs to be inferred. As we have previously derived, the posterior $z(h)$ assuming a uniform prior and some momentary evidence $\delta x_1, \dots, \delta x_n$ is given by

$$p\left(z(h)|x_d(t),t,k\right) = \mathcal{N}\left(\frac{x_d(t)}{kD(t)}, \frac{1}{k^2 D(t)}\right), \tag{19}$$

where $x_d(t)$ is the d-weighted accumulated evidence, $x_d(t) = \sum_{j=1}^n d_j \delta x_j$, and $D(t)$ is the "power" of the evidence time-course, $D(t) = \sum_{j=1}^n d_j^2$. The posterior belief of right-ward motion is thus given by

$$g(x_d,t) \equiv p\left(z(h) \geq 0|x_d(t),t\right) = \Phi\left(\frac{x_d(t)}{\sqrt{D(t)}}\right). \tag{20}$$

For a fixed time (and assuming $D(t) > 0$), the mapping between g and x_d is bijective. That is, for some fixed t, we can uniquely recover x_d from g, and vice versa.

If we assume a reward of one/zero for correct/incorrect choices, then the expected reward for a "right" choice is $g(x_d, t)$, and that for a "left" choice is $1 - g(x_d, t)$. We will assume a cost of c per second for accumulating evidence.

5.1. **Finding the optimal boundaries by dynamic programming.** We can formulate the problem of finding the optimal decision boundaries as a dynamic programming problem. At any point in time after stimulus onset, we can choose among three possible options. Either can we choose "right" or "left", yielding rewards $g(x_d, t)$ or $1 - g(x_d, t)$, or we can decide to accumulate more evidence for some time δt and choose later, which comes at cost $c\delta t$. This yields Bellman's equation

$$V(g, t) = \max \left\{ g, 1 - g, \langle V(\tilde{g}, t + \delta t) \rangle_{p(\tilde{g}|g,t)} - c\delta t \right\}, \tag{21}$$

where the expectation $\langle \cdot \rangle$ is over future beliefs $\tilde{g}$ when accumulating evidence for another small time-step δt. We chose to formulate Bellman's equation in g rather than x_d, as g is bounded by $g \in [0, 1]$, whereas x_d isn't.

In the above, we know all components other than $p(\tilde{g}|g, t + \delta t)$. The next section outlines how to derive this density.

With this density, we can discretize the belief space $g \in [0, 1]$ in small steps δg and solve Bellman's equation backwards in time in steps of δt. To do so, we can assume a final time T in which it is always beneficial to choose immediately, thus that the last term in Bellman's equation disappears. At this time, the value function $V(g, T)$ can thus be computed for all g. Based on $V(g, T)$ we can then compute $\langle V(\tilde{g}, T) \rangle_{p(\tilde{g}|g,T-\delta t)}$, with which we find $V(g, T - \delta t)$ for all g. Continuing this procedure, we find $V(g, T - 2\delta t), V(g, T - 3\delta t), \dots$ up to $V(g, 0)$. Having computed these value functions, the optimal decision boundary is at the belief at which the expected reward for deciding immediately equals that for accumulating more evidence. Using the relationship between g and x_d, we can map this optimal decision boundary into the diffusion model space.

5.2. **Finding the belief transition density.** Here we discuss how to find the belief transition density $p(\tilde{g}|g, t)$ where $\tilde{g}$ is the belief when holding belief g at time t and accumulating more evidence for some time δt. This belief transition is best found by relating it to the particle location,

$$p(\tilde{g}|g, t) \left| \frac{\mathrm{d}\tilde{g}}{\mathrm{d}\tilde{x}_d} \right| = p(\tilde{x}_d | x_d, t). \tag{22}$$

Thus, once we know the derivative $\mathrm{d}\tilde{g}/\mathrm{d}\tilde{x}_d$ and the particle transition density $p(\tilde{x}_d|x_d, t)$, we can find the belief transition density with the above equality.

The particle transition density can be found by using the momentary evidence likelihood and the posterior $p(z(h)|x_d(t), t)$, which is specified further above. In particular, $\tilde{x}_d(t) = x_d(t) + \delta x(t)$, such that

$$p(\tilde{x}_d|x_d, t, z(h)) = \mathcal{N}\left(\tilde{x}_d | x_d + z(h)kd(t)\delta t, \delta t\right), \tag{23}$$

where we have used the expression for the likelihood of the momentary evidence, $\delta x(t)$. The issue with the above is that we don't know $z(h)$, but can marginalize it out, using the current posterior,

$$p(\tilde{x}_d|x_d,t) = \int p(\tilde{x}_d|x_d,t,z(h))p(z(h)|x_d(t),t)\mathrm{d}z(h). \tag{24}$$

Both densities are Gaussian, such that the particle transition density is Gaussian as well. What remains as an exercise to the reader is to perform the marginalization, and use the result to derive the belief transition density. See Appendix C for a self-contained derivation of this belief transition density.

References

Drugowitsch, J., DeAngelis, G. C., Klier, E. M., Angelaki, D. E., and Pouget, A. (2014). Optimal multisensory decision-making in a reaction-time task. *eLife*, 3:e03005.

Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415:429–433.

Körding, K. P. and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427:244–247.

Appendix A. Bayes-optimal evidence accumulation across time

A.1. **Time-dependent evidence reliability.** The derivation uses the momentary evidence likelihood

$$p(\delta x_n|z(h), c) = \mathcal{N}\left(\delta x_n|z(h)v_n k(c)\delta t, \delta t\right), \tag{25}$$

where $v_n = v(\delta n)$ is the velocity in the nth timestep. This velocity controls the time-dependence of the evidence reliability. We first find the posterior $p\left(z(h)|\delta x_{1:n}, c\right)$ by Bayes' rule, resulting in

$$\begin{aligned} p\left(z(h)|\delta x_{1:n}, c\right) &\propto \prod_{j=1}^{n} \mathcal{N}\left(\delta x_j|z(h)v_j k(c)\delta t, \delta t\right) \\ &\propto e^{-\sum_{j=1}^{n} \frac{\left(\delta x_j - z(h)v_j k(c)\delta t\right)^2}{2\delta t}} \\ &\propto e^{-\frac{z(h)^2 k(c)^2 \sum_{j=1}^{n} v_j^2 \delta t}{2} + z(h)k(c)\sum_{j=1}^{n} v_j \delta x_j} \\ &= e^{-\frac{z(h)^2 k(c)^2 V(t)}{2} + z(h)k(c)x_v(t)} \\ &\propto \mathcal{N}\left(z(h)|\frac{x_v(t)}{k(c)V(t)}, \frac{1}{k(c)^2 V(t)}\right), \end{aligned} \tag{26}$$

where we have used $V(t) = \sum_{j=1}^{n} v_j^2 \delta t$ and $x_v(t) = \sum_{j=1}^{n} v_j \delta x_j$. This $V(t)$ in some sense indicates the total evidence reliability that has been available up to time t, by summing up the squared velocity up to that point in time. It will increase in proportion to the evidence available at any point in time. Thus, if the velocity is close to zero, $V(t)$ will remain mostly constant. $x_v(t)$ represents a similar sum, but now weighting the momentary evidence δx_j by its current reliability v_j. If this reliability is small, the momentary evidence is down-weighted, and thus ignored. If it is large, then the contribution of the momentary evidence to $x_v(t)$ is boosted.

We find the posterior over "right" by

$$p\left(z(h) \geq 0|x_v(t), V(t), c\right) = \int_0^1 p\left(z(h)|\delta x_{1:n}, c\right) \approx \Phi\left(\frac{x_v(t)}{\sqrt{V(t)}}\right). \tag{27}$$

Again, if $v(t) \approx 0$, neither $x_v(t)$ nor $V(t)$ increase, leaving the posterior unchanged in the absence of evidence. This would not happen if we were to simply sum up the momentary evidence without weighting it by reliability.

A.2. **Multiple cues with time-invariant evidence reliability.** The derivation uses the momentary evidence likelihoods

$$\begin{aligned} p(\delta x_{vis,n}|z(h), c) &= \mathcal{N}\left(\delta x_{vis,n}|z(h)k_{vis}(c)\delta t, \delta t\right), \\ p(\delta x_{vest,n}|z(h)) &= \mathcal{N}\left(\delta x_{vest,n}|z(h)k_{vest}\delta t, \delta t\right), \end{aligned} \tag{28}$$

In this case, the posterior $z(h)$ given some visual evidence $\delta x_{vis,1:n}$ and vestibular evidence $\delta x_{vest,1:n}$ results in

$$
\begin{aligned}
&p(z(h)|\delta x_{vis,1:n}, \delta x_{vest,1:n}, c) \\
&\propto \prod_{j=1}^{n} \mathcal{N}\left(\delta x_{vis,j}|z(h)k_{vis}(c)\delta t, \delta t\right) \mathcal{N}\left(\delta x_{vest,j}|z(h)k_{vest}\delta t, \delta t\right) \\
&\propto e^{-\sum_{j=1}^{n} \frac{(\delta x_{vis,j} - z(h)k_{vis}(c)\delta t)^2 + (\delta x_{vest,j} - z(h)k_{vest}\delta t)^2}{2\delta t}} \\
&\propto e^{-\frac{z(h)^2}{2}\left(k_{vis}(c)^2 + k_{vest}^2\right)\sum_{j=1}^{n}\delta t + z(h)\begin{pmatrix} k_{vis}(c)\sum_{j=1}^{n}\delta x_{vis,j} \\ + k_{vest}\sum_{j=1}^{n}\delta x_{vest,j}\end{pmatrix}} \\
&= e^{-\frac{z(h)^2}{2}\left(k_{vis}(c)^2 + k_{vest}^2\right)t + z(h)(k_{vis}(c)x_{vis}(t) + k_{vest}x_{vest}(t))} \\
&\propto \mathcal{N}\left(z(h)|\frac{k_{vis}(c)x_{vis}(t) + k_{vest}x_{vest}(t)}{\left(k_{vis}(c)^2 + k_{vest}^2\right)t}, \frac{1}{\left(k_{vis}(c)^2 + k_{vest}^2\right)t}\right).
\end{aligned}
\tag{29}
$$

If we now use $k_{comb}(c)^2 = k_{vis}(c)^2 + k_{vest}^2$, we can re-write this posterior as

$$
P(z(h)|x_{vis}(t), x_{vest}(t), c) = \mathcal{N}\left(z(h)|\frac{w_{vis}(c)x_{vis}(t) + w_{vest}(c)x_{vest}(t)}{k_{comb}(c)t}, \frac{1}{k_{comb}(c)^2 t}\right), \tag{30}
$$

with combination weights

$$
w_{vis}(c) = \frac{k_{vis}(c)}{k_{comb}(c)}, \qquad w_{vest}(c) = \frac{k_{vest}}{k_{comb}(c)}. \tag{31}
$$

Note that, these weights satisfy $w_{vis}(c)^2 + w_{vest}(c)^2 = 1$, but only $w_{vis}(c) + w_{vest}(c) = 1$ if one of $k_{vis}(c)$ or k_{vest} is zero. As for standard Bayesian cue combination, the diffusing particles, $x_{vis}(t)$ and $x_{vest}(t)$, are combined by weighting them according to their reliabilities. The difference to such standard cue combination is that these weights don't sum to one, which results from having reliabilities affecting the momentary evidence means rather than variances.

The use of these combination weights suggest that the momentary evidence $\delta x_{vis,n}$ and $\delta x_{vest,n}$ can be combined similarly by

$$
\delta x_{comb,n} = w_{vis}(c)\delta x_{vis,n} + w_{vest}(c)\delta x_{vest,n}, \tag{32}
$$

yielding the likelihood

$$
p(\delta x_{comb,n}|z(h), c) = \mathcal{N}\left(\delta x_{comb,n}|z(h)k_{comb}(c)\delta t, \delta t\right). \tag{33}
$$

This likelihoods results in the posterior $z(h)$,

$$
p(z(h)|x_{comb}(t), c) = \mathcal{N}\left(z(h)|\frac{x_{comb}(t)}{k_{comb}(c)t}, \frac{1}{k_{comb}(c)^2 t}\right), \tag{34}
$$

which corresponds to the previously derived posterior. This demonstrates that, by combining the momentary evidence of each modality, weighted by its reliability, evidence can again be accumulated in a single diffusion model that tracks $x_{comb}(t)$ and time.

Using this posterior, we find the posterior belief about right-ward motion to be given by

$$
p(z(h) \geq 0|x_{comb}(t), c) = \Phi\left(\frac{x_{comb}(t)}{\sqrt{t}}\right), \tag{35}
$$

which is the same expression as for evidence accumulation with a single cue and time-invariant evidence reliability, only with $x(t)$ replaced by $x_{comb}(t)$.

A.3. **Multiple cues with time-varying evidence reliability.** We assume momentary evidence likelihoods

$$\begin{aligned} p(\delta x_{vis,n}|z(h),c) &= \mathcal{N}\left(\delta x_{vis,n}|z(h)v_n k_{vis}(c)\delta t, \delta t\right), \\ p(\delta x_{vest,n}|z(h)) &= \mathcal{N}\left(\delta x_{vest,n}|z(h)a_n k_{vest}\delta t, \delta t\right), \end{aligned} \tag{36}$$

where the reliability of visual information $\delta x_{vis,n}$ is modulated by velocity $v_n = v(n\delta t)$, and the reliability of vestibular information $\delta x_{vest,n}$ by acceleration $a_n = a(n\delta t)$. As in the previous section, we find the posterior $z(h)$ by Bayes' rule, resulting in

$$\begin{aligned} &p(z(h)|\delta x_{vis,1:n}, \delta x_{vest,1:n}, c) \\ &\propto e^{-\frac{z(h)^2}{2}\left(k_{vis}(c)^2 V(t) + k_{vest}^2 A(t)\right) + z(h)(k_{vis}(c)x_{vis}(t) + k_{vest}x_{vest}(t))} \\ &\propto \mathcal{N}\left(z(h)|\frac{k_{vis}(c)x_{vis}(t) + k_{vest}x_{vest}(t)}{k_{vis}(c)^2 V(t) + k_{vest}^2 A(t)}, \frac{1}{k_{vis}(c)^2 V(t) + k_{vest}^2 A(t)}\right), \end{aligned} \tag{37}$$

where we have used $\sum_{j=1}^n v_j^2 \delta t = V(t)$, $\sum_{j=1}^n a_j^2 \delta t = A(t)$, and the accumulated evidence $\sum_{j=1}^n v_j \delta x_{vis,j} = x_{vis}(t)$ and $\sum_{j=1}^n a_j \delta x_{vest,j} = x_{vest}(t)$. The time-varying reliability impacts the posterior mean through a reliability-weighted temporal accumulation of the momentary evidence in $x_{vis}(t)$ and $x_{vest}(t)$. The difference in reliability across cues affects the posterior mean by a weighted combination of $x_{vis}(t)$ and $x_{vest}(t)$, whose weight is proportional to the overall reliability of the respective cue.

We can combine the two cue-dependent diffusing particles into a single particle as follows. First, let us defined a combined reliability, as before, by $k_{comb}(c)^2 = k_{vis}(c)^2 + k_{vest}^2$. Second, we want to find a time-dependent reliability component $D(t)$ that satisfies

$$k_{comb}(c)^2 D(t) = k_{vis}(c)^2 V(t) + k_{vest}^2 A(t). \tag{38}$$

Solving for $D(t)$ results in the weighted combination of $V(t)$ and $A(t)$,

$$D(t) = \frac{k_{vis}(c)^2}{k_{comb}(c)^2} V(t) + \frac{k_{vest}^2}{k_{comb}(c)^2} A(t). \tag{39}$$

If we define $D(t) = \sum_{j=1}^n d(n\delta t)^2 \delta t$, analogous to $V(t)$ and $A(t)$, we find that the reliability time-course for the combined particle is given by

$$d(t) = \sqrt{\frac{k_{vis}(c)^2}{k_{comb}(c)^2} v(t)^2 + \frac{k_{vest}^2}{k_{comb}(c)^2} a(t)^2}. \tag{40}$$

Note that this time-course now depends on a combination of both the global reliabilities, $k.$, for each cue, and their reliability time-courses.

This leads to the combined momentary evidence likelihood, based on

$$\delta x_{comb,n} = \frac{k_{vis}(c)v_n}{k_{comb}(c)d_n} \delta x_{vis,n} + \frac{k_{vest}a_n}{k_{comb}(c)d_n} x_{vest,n}, \tag{41}$$

to be given by

$$p(\delta x_{comb,n}|z(h),c) = \mathcal{N}\left(\delta x_{comb,n}|z(h)d_n k_{comb}(c)\delta t, \delta t\right), \tag{42}$$

where $d_n = d(n\delta t)$. Using previous results for the posterior with time-varying evidence reliability, the posterior $z(h)$ based on this momentary evidence is given by

$$p(z(h)|x_{comb}(t), c) = \mathcal{N}\left(\frac{x_{comb}(t)}{k_{comb}(c)D(t)}, \frac{1}{k_{comb}(c)^2 D(t)}\right). \tag{43}$$

It is easy to show that this posterior matches the previously derived one based on two sources of evidence. This again demonstrates that evidence accumulation can be performed by a single diffusion model even in the presence of multiple cues whose reliability varies over time.

With the above, the posterior belief of right-ward motion is given by

$$p\left(z(h) \geq 0 | x_{comb}(t), c\right) = \Phi\left(\frac{x_{comb}(t)}{\sqrt{D(t)}}\right). \tag{44}$$

Appendix B. Suggestions for code completion

This appendix provides suggestions for how to complete particular scripts. Please try to complete the scripts yourself before consulting this section!

B.1. `test_standard_cue_comb.m`. The discrimination thresholds per coherence can be found by

```
vis_sigs = NaN(1, ncohs);
comb_sigs = NaN(1, ncohs);
for icoh = 1:ncohs
coh = unique_cohs(icoh);
    [ioris, ichoice] = extract_cond(oris, mods, cohs, choice, NaN(size(choice)), 1, coh);
    vis_sigs(icoh) = fit_cumul_gauss(ioris, ichoice);

    [ioris, ichoice] = extract_cond(oris, mods, cohs, choice, NaN(size(choice)), 3, coh);
    comb_sigs(icoh) = fit_cumul_gauss(ioris, ichoice);
end

[ioris, ichoice] = extract_cond(oris, mods, cohs, choice, NaN(size(choice)), 2, coh);
vest_sig = fit_cumul_gauss(ioris, ichoice);
```

Based on these thresholds, the discimination threshold in the combined condition is predicted by

```
pred_sigs = 1 ./ sqrt(1 ./ vis_sigs.^2 + 1 / vest_sig^2);
```

B.2. `sim_weighted_diffusion.m`. Reliability-weighted evidence accumulation in a diffusion model is achieved by

```
xs = NaN(trials, nt);
sqrtdt = sqrt(dt);
for trial = 1:trials
    xs(trial, 1) = 0.0;
    for i = 2:nt
        dxi = zh * dn(i) * k * dt + sqrtdt * randn();
        xs(trial, i) = xs(trial, i-1) + dn(i) * dxi;

        if abs(xs(trial,i)) > theta
            xs(trial,i) = theta * sign(xs(trial,i));
            break
        end
    end
end
```

The reliability and reliability time-course in the combined condition can be computed by

```
k = sqrt(kvis^2 + kvest^2);
dn = sqrt((kvis/k * vn).^2 + (kvest/k * an).^2);
```

B.3. `sim_behavior.m`. The diffusion model in `sim_dmm(.)` can simulate reliability-weighted evidence accumulation by

```
sqrtdt = sqrt(dt);
x = 0;
n = length(dn);
for i = 2:n
    dx = dt * ksinh * dn(i) + sqrtdt * randn();
    x = x + dn(i) * dx;
    if abs(x) > theta
        if x > 0
            choice = 1;
        else
            choice = 0;
        end
        rt = i * dt + tnd;
        return
    end
end

if x > 0
    choice = 1;
else
    choice = 0;
end
rt = n * dt + tnd;
```

Optimal evidence accumulation across multiple cues is achieved by

```
kcomb = sqrt(kvest^2 + kvis(icoh)^2);
dn = sqrt((kvis(icoh)/kcomb * vn).^2 + (kvest/kcomb * an).^2);
```

Parameters that generate behavior similar to that seen in human subjects are

```
kvest = 20;                 % vestibular  reliability
kvis = [30 120 350];        % visual  reliability  for  coh [0.25  0.37  0.7]
thetas = [0.15 0.15  0.06]; % decision bounds for [vis  vest  comb] mod.
tnd = 0.3;                  % non-decision time
```

Appendix C. Belief transition density for time-varying evidence reliability

Here, we provide a self-contained derivation of the belief transition density if the evidence reliability varies across time.

C.1. **Model assumptions.** We assume a latent state $z \in \{-1, 1\}$ and a non-negative nuisance parameter y that together form $\mu = z \times y$. The sign of μ reveals z and its magnitude is given by y. We assume, a-priori, that μ is distributed according to a zero-mean Gaussian with variance σ_μ^2, making both $z = 1$ and $z = -1$ a-priori equally likely. In any small time step of size δt, the decision maker makes a noisy observation of μ whose reliability is additionally modulated by $d(t)$ at any time t. Here $d(t) = 1$ implies no modulation, $d(t) > 1$ implies more evidence per unit time, and $d(t) < 1$ less evidence per unit time. Overall, this yields the generative model

$$p(\mu) = \mathcal{N}\left(\mu|0, \sigma_\mu^2\right), \tag{45}$$

$$p(\delta x_n|\mu) = \mathcal{N}\left(\delta x_n|\mu d_n \delta t, \delta t\right), \tag{46}$$

where $d_n = d(d\delta t)$ is the time-dependent reliability modulation in the nth time step. This model assumes that the time-course $d(t)$ is known and remains the same across trials. What is unknown and needs to be inferred is the sign of μ. Its magnitude is also unknown and is a nuisance parameters that the decision-maker is not interested in but needs to infer nonetheless (as it impacts the evidence).

C.2. **Bayes-optimal evidence accumulation.** We derive optimal evidence accumulation, assuming that we observe n pieces of momentary evidence $\delta x_{1:n}$ and are interested in inferring the sign of μ. We do this in two steps. First we find the full posterior over μ and as a second step we find the probability that it is positive.

The full posterior $p(\mu|\delta x_{1:n})$ is found by Bayes' rule and results in

$$\begin{aligned}
p(\mu|\delta x_{1:n}) &\propto p(\mu) \prod_{j=1}^{m} p(\delta x_j|\mu), \\
&= \mathcal{N}\left(\mu|0, \sigma_\mu^2\right) \prod_{j=1}^{n} \mathcal{N}\left(\delta x_j|\mu d_n \delta t, \delta t\right), \\
&\propto e^{-\frac{\mu^2}{2}\left(\frac{1}{\sigma_\mu^2} + D(t)\right) + \mu x_d(t)}
\end{aligned} \tag{47}$$

where we have defined

$$D(t) \approx \sum_{j=1}^{n} d_n^2 \delta t, \qquad x_d(t) = \sum_{j=1}^{n} d_n \delta x_n. \tag{48}$$

The above posterior is quadratic in μ in the exponential, such that it is a Gaussian posterior. Adding the normalization constant yields the mean and variance

$$p(\mu|x_d(t), t) = \mathcal{N}\left(\mu|\frac{x_d(t)}{\sigma_\mu^{-2} + D(t)}, \frac{1}{\sigma_\mu^{-2} + D(t)}\right). \tag{49}$$

Note that if we were to assume a uniform prior over μ, we have $\sigma_\mu^2 \to \infty$ or $\sigma_\mu^{-2} \to 0$, such that only $D(t)$ remains in the above denominators.

To find the posterior belief that $z = 1$, we find the probability that $\mu \geq 0$, which is given by

$$g(t) \equiv p(\mu \geq 0|x_d(t), t) = \int_0^\infty p(\mu|x_d(t))\mathrm{d}\mu = \Phi\left(\frac{x_d(t)}{\sqrt{\sigma_\mu^{-2} + D(t)}}\right). \tag{50}$$

We will call this posterior simply the "belief" $g(t)$. Note that, for some fixed time t, the mapping between the "particle location" $x_d(t)$ and $g(t)$ is bijective, such that we can find the particle location $x_d(t)$ that corresponds to a given belief $g(t)$ and vice versa.

C.3. **Belief transition density.** In order to find the optimal policy by dynamic programming, we need to have an estimate for how the belief evolves when accumulating evidence for another time δt. Let us denote the current belief after n time steps of momentary evidence as $g = g(t) = g(n\delta t)$, and the belief after one more time step as $\tilde{g} = g(t + \delta t) = g((n + 1)\delta t)$. Then, the probability we are interested in finding is given by $p(\tilde{g}|g, t)$. This density depends on g as well as t, as how this belief evolves over time is time-dependent. This is a consequence of the belief itself depending on both $x_d(t)$ and time t.

We will find the belief transition density by using its relation to the particle location transition density,

$$p(\tilde{g}|g, t)\left|\frac{\mathrm{d}\tilde{g}}{\mathrm{d}\tilde{x}}\right| = p(\tilde{x}|x, t), \tag{51}$$

where $x = x_d(t) = x_d(n\delta t)$ and $\tilde{x} = x_d(t + \delta t) = x_d((n + 1)\delta t)$. Thus, we will first find $p(\tilde{x}|x, t)$ and then map it to $p(\tilde{g}|g, t)$ using the above relationship.

To find the particle location density, observe that we known the posterior $p(\mu|x_d(t), t)$. Furthermore, we know that $\tilde{x} = x + \delta x_{n+1}$ and $\delta x_{n+1}|mu \sim \mathcal{N}(\mu d_{n+1}\delta t, \delta t)$, such that

$$p(\tilde{x}|x, \mu, t) = \mathcal{N}(\tilde{x}|x + \mu d_{n+1}\delta t, \delta t). \tag{52}$$

Thus, we can handle not actually "knowing" μ by marginalizing it out, resulting in

$$\begin{aligned} p(\tilde{x}|x, t) &= \int p(\tilde{x}|x, \mu, t)p(\mu|x, t)\mathrm{d}\mu \\ &= \int \mathcal{N}(\tilde{x}|x + \mu d_{n+1}\delta t, \delta t)\,\mathcal{N}\left(\mu|\frac{x}{\sigma_\mu^{-2} + D(t)}, \frac{1}{\sigma_\mu^{-2} + D(t)}\right) \\ &= \mathcal{N}\left(\tilde{x}|x\left(1 + \frac{d_{n+1}\delta t}{\sigma_\mu^{-2} + D(t)}\right), \delta t\left(1 + \frac{d_{n+1}^2\delta t}{\sigma_\mu^{-2} + D(t)}\right)\right). \end{aligned} \tag{53}$$

This yields the particle location transition density.

We use the above mapping to find the belief transition density. This mapping requires the derivative

$$\frac{\mathrm{d}\tilde{g}}{\mathrm{d}\tilde{x}} = \frac{\mathrm{d}}{\mathrm{d}\tilde{x}}\Phi\left(\frac{\tilde{x}}{\sqrt{\sigma_\mu^{-2} + D(t + \delta t)}}\right) = \mathcal{N}\left(\tilde{x}|0, \sigma_\mu^{-2} + D(t + \delta t)\right), \tag{54}$$

for which we have used

$$\frac{\mathrm{d}}{\mathrm{d}x}\Phi\left(\frac{x - \mu}{\sigma}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\int_0^x \mathcal{N}\left(a|\mu, \sigma^2\right)\mathrm{d}a = \mathcal{N}\left(x|\mu, \sigma^2\right). \tag{55}$$

Putting all the terms together results in

$$p(\tilde{g}|g,t) = \frac{\mathcal{N}\left(\tilde{x}|x\left(1+\frac{d_{n+1}\delta t}{\sigma_\mu^{-2}+D(t)}\right), \delta t\left(1+\frac{d_{n+1}^2\delta t}{\sigma_\mu^{-2}+D(t)}\right)\right)}{\mathcal{N}\left(\tilde{x}|0, \sigma_\mu^{-2}+D(t+\delta t)\right)}, \tag{56}$$

where the x and $\tilde{x}$ are given by

$$x = \Phi^{-1}(g)\sqrt{\sigma_\mu^{-2}+D(t)}, \qquad \tilde{x} = \Phi^{-1}(\tilde{g})\sqrt{\sigma_\mu^{-2}+D(t+\delta t)}. \tag{57}$$

To evaluate the above density, we are only interested in terms proportional to $\tilde{g}$. Expanding the normal distributions, these are given by

$$p(\tilde{g}|g,t) \propto_{\tilde{g}} e^{-\frac{\tilde{x}^2}{2\delta t\left(1+\frac{d_{n+1}^2\delta t}{\sigma_\mu^{-2}+D(t)}\right)}+\frac{\tilde{x}^2}{2\left(\sigma_\mu^{-2}+D(t+\delta t)\right)}+\frac{x\tilde{x}\left(1+\frac{d_{n+1}\delta t}{\sigma_\mu^{-2}D(t)}\right)}{\delta t\left(1+\frac{d_{n+1}^2\delta t}{\sigma_\mu^{-2}+D(t)}\right)}}. \tag{58}$$

Substituting the expressions for x and $\tilde{x}$ and using

$$\delta t_{eff} = \frac{\delta t}{\sigma_\mu^{-2}+D(t)} \approx \frac{\delta t}{\sigma_\mu^{-2}+D(t+\delta t)} \tag{59}$$

we find

$$p(\tilde{g}|g,t) \propto e^{\frac{\left(\Phi^{-1}(\tilde{g})\right)^2}{2}-\frac{\left(\Phi^{-1}(\tilde{g})-(1+d_{n+1}\delta t_{eff})\Phi^{-1}(g)\right)^2}{2\delta t_{eff}(1+d_{n+1}^2\delta t_{eff})}}. \tag{60}$$

Department of Neurobiology, Harvard Medical School, Boston, MA
E-mail address: jan_drugowitsch@hms.harvard.edu