

Increasing Signal-to-Noise Ratio in Gas Chromatography - Electroantennography Using a Deans Switch Effluent Chopper

Andrew J. Myrick & Thomas C. Baker

Journal of Chemical Ecology

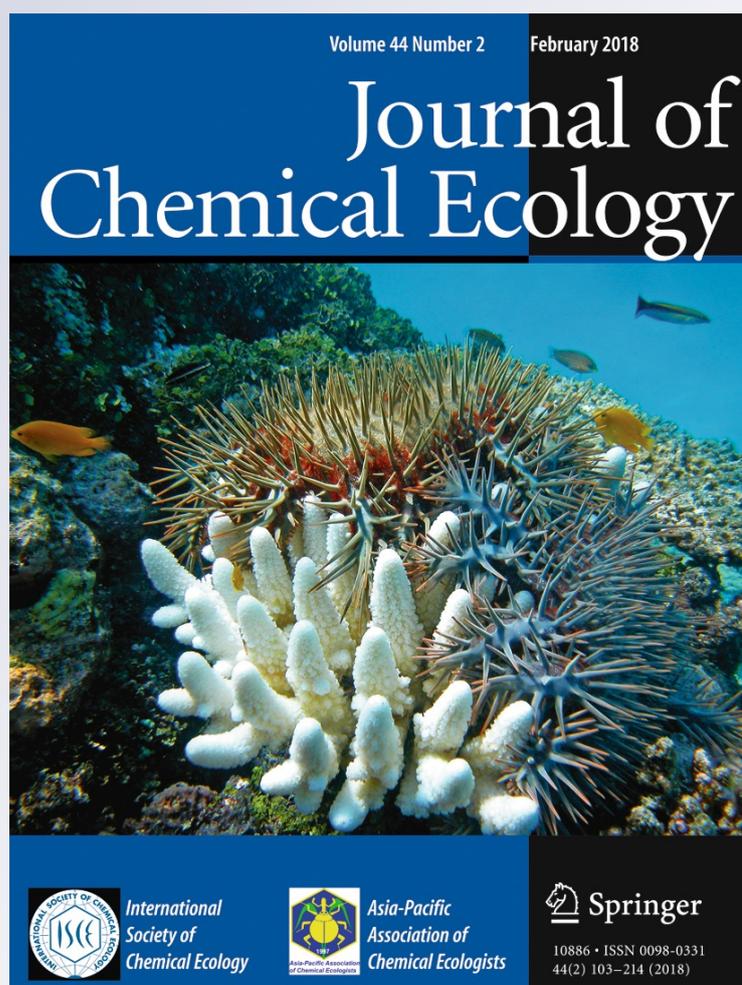
ISSN 0098-0331

Volume 44

Number 2

J Chem Ecol (2018) 44:111-126

DOI 10.1007/s10886-017-0916-y



Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".



Increasing Signal-to-Noise Ratio in Gas Chromatography - Electroantennography Using a Deans Switch Effluent Chopper

Andrew J. Myrick¹ · Thomas C. Baker¹

Received: 28 August 2017 / Revised: 16 November 2017 / Accepted: 15 December 2017 / Published online: 6 January 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Gas-chromatography-electroantennographic detection (GC-EAD) is a technique used in the identification of volatile organic compounds (VOCs), such as pheromones and plant host odors, which are physiologically relevant to insects. Although pheromones often elicit large EAD responses, other behaviorally relevant odors may elicit responses that are difficult to discern from noise. Lock-in amplification has long been used to reduce noise in a wide range of applications. Its utility when incorporated with GC-EAD was demonstrated previously by chopping (or pulsing) effluent-laden air that flowed over an insect antenna. This method had the disadvantage that it stimulated noise-inducing mechanoreceptors and, in some cases, disturbed the electrochemical interfaces in a preparation, limiting its performance. Here, the chopping function necessary for lock-in amplification was implemented directly on the GC effluent using a simple Deans switch. The technique was applied to excised antennae from female *Heliothis virescens* responding to phenethyl alcohol, a common VOC emitted by plants. Phenethyl alcohol was always visible and quantifiable on the flame ionization detector (FID) chromatogram, allowing the timing and amount of stimulus delivered to the antennal preparation to be measured. In our new chopper EAG configuration, the antennal preparation was shielded from air currents in the room, further reducing noise. A dose-response model in combination with a Markov-chain monte-carlo (MCMC) method for Bayesian inference was used to estimate and compare performance in terms of error rates involved in the detection of insect responses to GC peaks visible on an FID detector. Our experiments showed that the predicted single-trial phenethyl alcohol detection limit on female *H. virescens* antennae (at a 5.0% expected error rate) was 140,330 pg using traditional EAG recording methods, compared to 2.6–6.3 pg (5th to the 95th percentile) using Deans switch-enabled lock-in amplification, corresponding to a 10.4–12.7 dB increase in signal-to-noise ratio.

Keywords GC-EAD · GC-EAG · Electroantennogram · Deans switch · Signal processing · Lock-in amplification · Signal-to-noise ratio

Introduction

Gas chromatography-electroantennographic-detection (GC-EAD) (Moorhouse et al. 1969) is used in the identification of odorant compounds that insect antennae can detect (Schneider 1957). In many cases, the technique has been utilized to identify pheromones, (Andreadis et al. 2015; Bouwer et al. 2015; Sillam-Dussès et al. 2009) but has also found applications in other studies, such as unraveling plant host-finding behavior in insects (Pan et al. 2015) or other chemical

communication between plants and insects (Schiestl and Marion-Poll 2002). Whereas EAG depolarizations tend to be very large when pheromonal components are encountered, other types of compounds often result in relatively small EAG responses. Efforts to identify EAG-active non-pheromonal compounds benefit from increased EAG sensitivity. Such efforts have, for instance, utilized preparations incorporating multiple antennae (Moore 1981; Park and Baker 2002), high-pass combined with low-pass filters, or other automated detection methods designed to recognize the shapes of EAG depolarizations resulting from GC peaks (Slone and Sullivan 2007).

Signal chopping for the general purpose of increasing signal-to-noise ratio (SNR) is a well known technique (Michels and Curtis 1941; Scofield 1994) and was, in fact, (manually) used in the original GC-EAD experiments performed by Moorhouse et al. (1969), in which GC peak widths

✉ Andrew J. Myrick
ajm25@psu.edu

¹ Department of Entomology, Chemical Ecology Laboratory, Penn State University, University Park, PA 16802, USA

were impractically long to be used directly on antennae. Marion-Poll and Thiéry (1996) suggested some form of pulse shaping to increase sensitivity of their GC-EAD peaks, which were approximately 10 s in duration. Later, Gouinguéné et al. (1998) recognized the advantage of chopping to sidestep low frequency noise and thermally modulated the GC effluent directed over the antennal preparation, resulting in systematically larger EAG depolarizations when compared to traditional methods.

Adaptation has also been cited as a possible source of reduced EAG signal amplitude in traditional recordings compared to chopped recordings. Today's standard capillary GC peaks are long enough that adaptation can occur in an antenna used as a detector. For instance, in French et al. (2014), single unit recordings of CO₂ receptors adapted to stimuli below frequencies of about 1 Hz, whereas Fourier transforms of GC peak waveforms contain (spectral energy at) frequencies mostly limited to below 0.2 to 0.3 Hz.

More recently, two publications (Myrick and Baker 2011, 2012) reported on the use of chopper modulation of a stimulus delivered to an antenna to increase SNR and substantially increase the chances that antennal responses to GC peaks are detected. The method introduced in those works improved upon previous chopping efforts by increasing the achievable chop frequency, applying optimal waveform processing and quantifying the performance advantage. Matched filtering in colored noise (Helstrom 1960; Robey et al. 1992), which ideally maximizes SNR, was applied to EAG waveforms. EAG waveforms tend to have baselines that slowly and randomly drift, but also have small perturbations that occur with greater frequency. This is a characteristic of “red” (colored) noise, in which slow random fluctuations are larger than faster ones. Matched filtering can be thought of as optimally searching through the red-noise in the EAG waveform for the FID peak elution waveform.

Presence/absence testing was then accomplished using a simple threshold test (Trees 2001). In our model, measurements below the threshold are categorized as noise, and values above it as EAGs. However, detection errors inevitably occur. Possible error types include misses, which occur with a probability P_M , and false alarms, with an associated probability of P_{FA} . Misses occur when a below-threshold EAG occurs and is, therefore, missed, while false alarms occur when the noise is higher than the threshold and no EAG has occurred.

Here, we report on the modulation of the effluent from the oven of a gas chromatograph. Because mechanical movement of the odor delivery tube (Myrick and Baker 2011; 2012) was eliminated, techniques commonly used when making high quality traditional GC-EAD recordings could be implemented. Often the antenna recording site is shielded in some way from air currents that may be present in the laboratory, and the odor delivery tube is cooled. These increase the time antennae

are useful for recording by increasing relative humidity of the air and minimize the interfering effects of mechanoreceptor-stimulating air currents.

Chopping of effluent inside a GC oven was accomplished using a Deans switch (Deans 1968). A Deans switch can be used to quickly direct effluent from a common port to either of 2 ports. It is proposed that the effluent alternate between the FID and the EAG. A description of the operation of a Deans switch may be found in Seeley et al. (2007) and Sharif et al. (2016).

In Myrick and Baker (2012), the air carrying the effluent to the antennal preparations was modulated at 8 Hz and this increased the detection sensitivity for (*Z*)-11-hexadecenal (Z11–16:Ald), the major pheromone component of *Heliothis subflexa*, using antennae of male *H. subflexa*. In the experiments described here, a plant-derived compound, phenethyl alcohol (PEA), was used as the stimulus and antennae from female *Heliothis virescens* as detectors. The detection sensitivity for PEA by female *H. virescens* is much lower (i.e., detection limit is much higher) relative to that for the major pheromone component, Z11–16:Ald, by male *H. subflexa*. Thus, PEA represents a general odorant compound from plants that, if it were detectable by GC-EAD in volatile samples, could be more easily identified from such samples. An additional benefit of using PEA was that it is detectable by a GC-flame ionization detector (FID) when an EAG response is elicited. This enables the measurement of the dose and elution time of the PEA from the FID signal which, in turn, allows the selection of the known timing (KT) method (Myrick and Baker 2012) that reduces error rates in the detection of GC-EAD olfactory responses. In Myrick and Baker (2012), an envelope detector (Trees 2001, p. 341) was employed. However, precise knowledge of the timing between the FID and EAG further allowed coherent detection to be applied, which selects the ‘in-phase’ component of the amplitude measurement. The in-phase component is consistently aligned in time with the chopping waveform. Selecting the in-phase component eliminated the ‘quadrature’ component, thus eliminating half of the noise and increasing the SNR by about 3 dB.

Methods and Materials

Overview After acquiring GC-EAD waveforms, amplitudes of EAG depolarizations were estimated after maximizing SNR using techniques detailed in Myrick and Baker (2011). Briefly, chopped waveforms were first demodulated, which can be thought of as measuring the amplitude due to the on-off action of the chopping and replacing it with a smooth waveform. Next, matched filtering in colored noise was used to maximize SNR and to obtain estimates of the amplitude of EAG (and FID) responses. Data obtained via amplitude estimates as a

function of dose were then used to infer parameters of a dose-response model for the purpose of estimating error rates (probabilities) involved in the detection of an EAG response. The dose-response model was assumed to be linear on a log-log scale or equivalently followed a power-law. Evidence for such a relationship may also be found in several publications (Guerin et al. 1983; Mayer et al. 1987; Park and Hardie 1998; Schott et al. 2013, 1999; Weissbecker et al. 2004). Because the EAG noise was found to be near Gaussian (via quantile-quantile plots, not shown here) when the antennae were cooled and shielded from air currents, no empirical distribution functions for the noise were necessary, making the analysis amenable to available software that performs Bayesian inference using Markov-chain monte-carlo (MCMC) methods and analytic probability density functions. JAGS, a free software package written for the purpose of using MCMC sampling to infer model parameters was employed. More information about the methods used by JAGS may be found in Gelman et al. (2004) and Ntzoufras (2009). The dose-response model was used to optimize the detection threshold, by minimizing modeled detection errors, as a function of dose. Performance was then evaluated by comparing predicted error rates in detecting the presence of EAG responses as a function of dose.

Insects Moths of *H. virescens* used in this experiment were maintained over many generations on general purpose Lepidoptera diet (F9772, Frontier Agricultural Science, Newark, DE, USA). Single larvae were placed into individual diet-filled cups, removed at pupation, and sexed under a dissecting scope. Sexed pupae were placed into separate screen cages lined with paper toweling and provided with a 10% sucrose solution for adults to feed on. Moths were kept in a Percival environmental chamber (Percival Scientific, Boone, IA USA) at 20 °C and 50% or higher relative humidity, under a standard light cycle, 16 L:8D.

Gas Chromatograph Settings A 6190 N (Agilent, Santa Clara, CA, USA) GC was operated in split mode with an 11:1 split ratio and was fitted with an Alltech EC-5 (Grace, Columbia, MD, USA) column, 30 m in length, 0.32 mm ID, with a 1.0 μm internal coating. The GC was operated with a helium carrier gas head pressure of 171 kPa (24.8 PSI). The oven temperature was held at 125 °C for all injections. The injector and FID detector were held at 250 °C. The FID electrometer was operated in “fast peaks” mode with an analog output voltage range of 1 V, range value of 0 and attenuation value of 0.

GC Oven Setup A Silflow Deans switch kit (123,703, SGE Analytical Science by Trajan, Ringwood, Victoria, Australia) was utilized for these experiments and connected as shown in Fig. 1a. The chromatographic column was attached to the center port of the Deans switch. Only one output of the

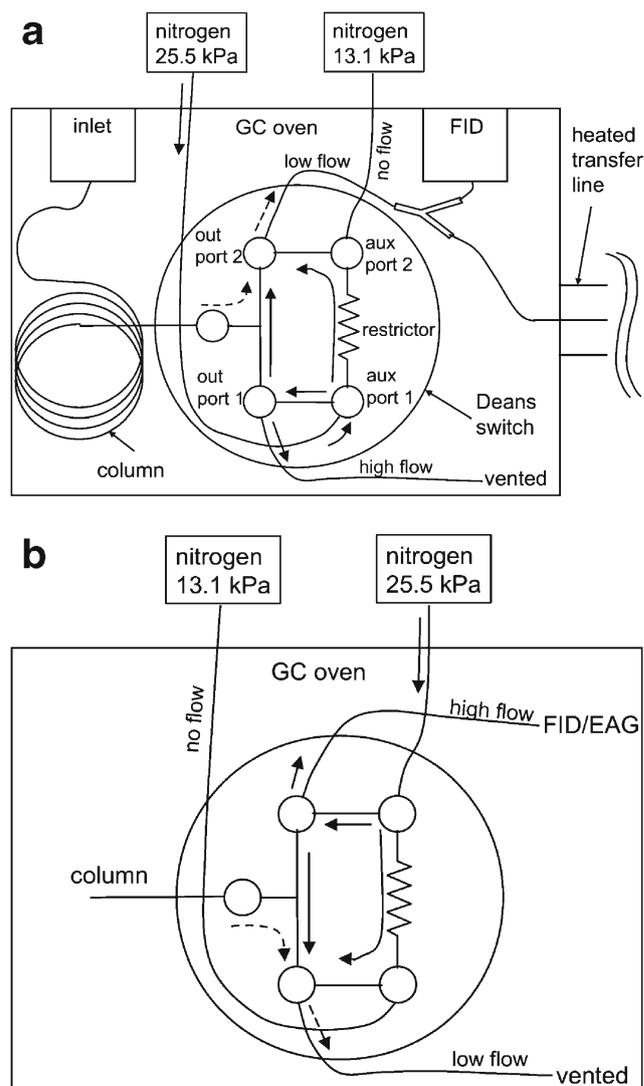


Fig. 1 Alternating valve configurations for chopping at 1.5 Hz. Dashed arrows indicate effluent flow. Solid arrows indicate steady state flow of nitrogen makeup gas. **a** Effluent is directed toward the flame ionization detector (FID) and electroantennogram detector (EAD). Nitrogen flow from auxiliary port 1 (A1) of the Deans switch sweeps past the input port, where it is combined with the effluent from the column and directed through out port 2, toward the FID and EAD, through a splitter. No flow is occurring through auxiliary port 2 (A2), preventing effluent from splitting at out port 2 and exiting into A2. Flow into the splitter (low flow) is lower than that which is vented. **b** Effluent is vented. Nitrogen flow now flows past the column from A2 to output port 2. Note the flow rate to the FID and EAD is increased in this state. GC = gas chromatograph

Deans switch was directed to both the FID and the EAG via a fused silica splitter (23,631, Supelco, Bellefonte, PA, USA) so that the dose could be quantified on the FID for both chopped and traditional measurements with the same setup. Deactivated 0.32 mm ID silica tubing connected the Deans switch, the splitter, the FID and the EAG. A piece of deactivated silica tubing was connected to the second Deans switch output, so that the Deans switch output flows were

symmetrical. The auxiliary ports of the Deans switch were attached to the kit-provided stainless-steel tubing. Note, that it is possible to route fully half of the effluent to the FID and the EAG during both chopped and traditional measurements by removing the splitter and connecting one Deans switch output to the FID and the other to the EAG. While traditional measurements are being made, the Deans switch may be used as a splitter by applying the same pressure to both auxiliary lines (Boeker et al. 2013).

Deans Switch External Components A pressure regulator (Model 4000, Parker, Cleveland, IN, USA) connected to a nitrogen tank regulator output was combined with three low-cost solenoid valves (S070C-6CG-32, SMC, Tokyo, Japan) so that pressures of 13.1 kPa and 25.5 kPa were applied to the two auxiliary inputs and alternated at 1.5 Hz. A digital timer on an Arduino's (Arduino Uno Rev. 3, BMCI labs, Boston, MA, USA) microcontroller was used to generate a 1.5 Hz, 50% duty cycle digital signal to drive the solenoid valves. For further information about the switching scheme, see Appendix 3.

Switching the direction of the Deans switch flow had the effect of abruptly changing the flow rate through the two exit lines. This is illustrated in Fig. 1 where the low flow and high flow designations alternate as the pressures on the auxiliary lines are exchanged. Under low flow conditions, the FID/EAG tubing connected to the Deans switch outlet is nearly filled with effluent during the half-period. After switching, clean nitrogen under high flow conditions swept out the effluent more quickly than it was filled, resulting in a pulse that had higher flux than that flowing into the Deans switch, but shorter in time than half of the switching period.

Custom Odor Delivery Tube Humidified air flowed at $500 \text{ ml} \cdot \text{min}^{-1}$ into a custom tube, constructed using 6.35 mm (0.25 in.) OD, 3.18 mm (0.125 in.) ID aluminum tubing with a length of 140 mm (5.5 in.), for mixing with the GC effluent (Fig. 2). A hole was drilled 15.9 mm (0.625 in.) from the exit-end of the aluminum tube for the GC effluent to enter. The relatively narrow ID was used so that GC effluent would be swept away quickly. Although unnecessary for the slow chop rate of 1.5 Hz, higher chopping frequencies aided.

The temperature of the odor delivery tube was maintained at 20 °C using a Peltier cooler. The cooled odor delivery tube assembly consisted of a Peltier element sandwiched between the odor delivery tube on its cool side and a CPU heatsink with a fan on its warm side. A proportional-integral-derivative (PID) controller was implemented using the Arduino microcontroller board for the purpose of maintaining constant temperature.

The odor delivery tube flow exited into a custom polystyrene antenna holder whose airflow-path was cylindrical with an ID of 6.35 mm (0.25 in.) and had a length of 15.9 mm (0.625 in.). The antenna was located 6.35 mm (0.25 in.) from the end of the aluminum tubing.

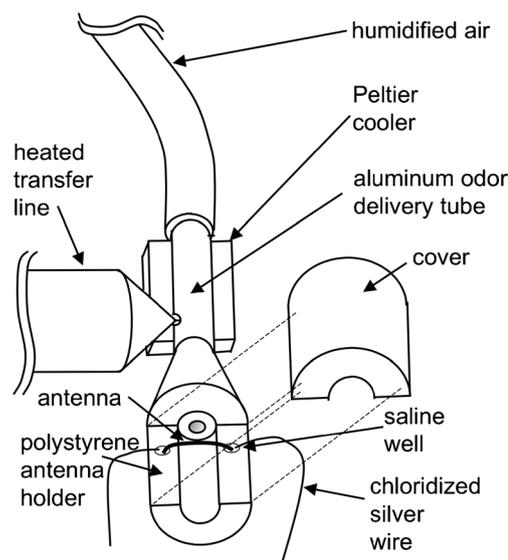


Fig. 2 ‘Odor delivery tube’ and antenna holder. Heated gas chromatograph effluent is introduced into a humidified airstream in the temperature-controlled odor delivery tube. The mixture is then passed over the antenna in the antenna holder. The ends of the antenna bathe in saline, which is in turn connected to chloridized silver wires. A cover over the antennal preparation prevents interference-inducing mechanical disturbances due to external air currents

Antennal Preparations Electrical connections to excised antennae were made through saline solution. Borosilicate capillary tubes (1B150F-4, World Precision Instruments, Sarasota, FL, USA) were drawn to sharp tips, which were scored and broken off so that the antennae would fit inside for mechanical stability. The tubes were shortened and glued inside wells created in the antenna holder. The wells were filled with saline, and connections to the saline were made using chloridized silver wire. Opposite ends of excised antennae, with the first few terminal segments removed, were then inserted into the borosilicate tubes. With the axis of an antenna perpendicular to the flow and horizontal, the half of the antenna with sensilla was arranged so that the sensilla pointed “up”, perpendicular to the flow. Following this, the preparation was covered to prevent laboratory air currents from stimulating mechanoreceptors on the antenna.

Data Collection A total of 8 measurement sets were obtained from 16 antennae, in which 2 antennae were used for each measurement set. Measurements were made so that traditional and chopped measurements could be paired on the same antenna. To ‘balance’ possible aging effects, chopped measurements preceded traditional measurements on one antenna, while the order was reversed for the other antenna. Each measurement set was used to test three half-decade-spaced doses and consisted of 8 recordings, 4 from each antenna, in which each recording consisted of a noise-only segment and two sets of

three half-decade-spaced doses (Fig. 3). This ensured that only one order of magnitude in dosage was tested on each antenna. This was done to reduce possible effects of very large relative doses on subsequent measurements made using small doses. A recording was first started with the chopper enabled. After the last peak eluted, a second recording was made in the same manner, resulting in 12 total measurements. The chopper was then disabled for traditional recordings. Two recordings were made using the same injection procedure with the same amounts injected into the GC inlet. A second antenna was then subjected to the same treatment, although the traditional method preceded the chopped method. This procedure was conducted using various dose ranges resulting in 384 total measurements.

GC-EAD Recording System The GC-EAD system was custom built. Amplification of the EAG signal was accomplished by a CMOS-type, high impedance DC-coupled amplifier, band-limited to 50 Hz. A second, AC-coupled (auto-zeroed) EAG signal was generated by applying a single pole high pass filter with a time constant of 1.8 s to the EAG signal. The FID analog output was amplified using a bipolar-type DC-coupled amplifier. The two EAG signals and the FID analog output from the GC were simultaneously digitized at 100 Sa/s (samples per second) and stored for later analysis. The EAG gain was sufficient to cause the EAG to be the dominant source of noise at the output of the digitizer. The AC-coupled EAG signal was used for all measurements unless otherwise specified. Ideally, auto-zeroing has no effect on the matched-filtered output.

Waveform Demodulation Demodulated waveforms for both the chopped FID and EAG were obtained using direct digital down conversion incorporating two product detectors (Ziemer and Tranter 1995). Product detectors followed by finite impulse response (FIR) low-pass filters, with 3 dB cutoff frequencies of 0.6 Hz, established the baseband in-phase (I) and quadrature (Q) signals. The FIR low-pass filter was

produced using the product of a sine function and a Gaussian function. The filter coefficients are described by the following equation,

$$h[nT] = 2fT \frac{\sin(2\pi f(nT-d))}{2\pi f(nT-d)} e^{-\frac{(nT-d)^2}{2\sigma^2}}, \quad (1)$$

where $T = 0.01$ s, $f = 0.75$ Hz, $d = 3$ s and $\sigma = 0.6$ s. The filter was 6 s in duration. Following this, the baseband signals were decimated to 10 Sa/s. The scaling constant of $2fT$ is approximate but accurate to within 1%.

Signal and Noise Models Signal models for the chopped EAG were constructed using the methods described in Myrick and Baker (2011). Signals were created in 20 s windows, from a Gaussian shaped pulse centered at 10 s. The Gaussian pulse input was modeled with a full width half maximum (FWHM) value of 1.6 s and a peak tailing time constant of 1 s. For traditional recordings, the high-pass time constant parameter was 1.8 s.

Signal models were also necessary for estimation of peak amplitudes and phases on FID waveforms and were identical to the signal model used for the chopped EAG.

The noise segment of each recording from 20 to 180 s was used to estimate the 'unbiased' autocovariance. The autocovariance was used to estimate the 200×200 covariance matrix, \mathbf{M} , assuming the noise was stationary and ergodic.

Waveform Measurements

FID Measurements The FID output allowed the dose, location (in time) and, in the case of chopped measurements, relative phase (which can be considered part of the location measurement) between the FID and EAG at 1.5 Hz to be estimated. Traditional FID amplitude measurements were produced using the methods applied to EAG peaks described as having uncertain timing (UT) in Myrick and Baker (2012). Briefly, FID amplitude measurements were made using peak searches

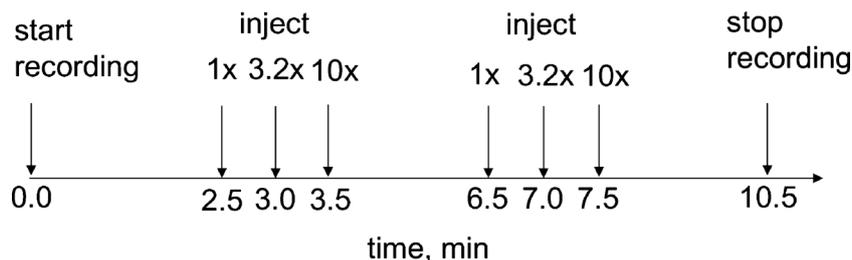


Fig. 3 Illustration of sequence used to make a gas chromatograph-electroantennogram detection recording. After a recording was started, 2.5 min were allowed to pass to create a segment to be used for autocovariance estimation. Three manual injections spaced by

approximately 30 s were then made using the lowest dose first (represented by x) and ending with the highest. After a total of 6.5 min were allowed to pass, 3 more identical 30-s-spaced injections were made

(sometimes by hand, in the case of low doses) of colored noise-matched filtered FID data. The elution time of each FID peak was retained for the purpose of measuring corresponding EAG amplitudes. The FID voltage was scaled to dose in nanograms, \hat{d}_{FID} , using a scaling parameter determined by measuring the amplitude of several 1 ng splitless injections, assuming a linear relationship between voltage and dose. The caret symbol (^) is often used to denote a variable that represents an estimated quantity.

Chopped FID voltage and location measurements were produced in the same manner as traditional measurements. In addition to the elution time, the phase (proportional to delay) relative to the demodulation signal of each FID peak was retained for the purpose of measuring corresponding EAG amplitudes. Modulated FID waveforms had a duty cycle of about 34%, and, therefore, required a different scaling factor than traditional measurements.

EAG Measurements Denoted by m_t , where t refers to measurement type (traditional is represented by 1 and chopped by 2), EAG measurements are estimates of amplitudes of depolarizations that have been normalized with respect to noise. Amplitude estimates, \hat{a} , which, in general are represented by complex numbers, of traditional EAG signals were made using a matched filter in colored noise with known timing (KT) described in Myrick and Baker (2012). Amplitude estimates were then normalized by dividing the estimate by the root mean square (RMS) power estimate of matched-filtered noise alone, $\hat{\delta}_{vm}$. An expression for m_1 can be written,

$$m_1 = \frac{\hat{a}_1}{\hat{\delta}_{vm}}, \quad (2)$$

where the 1 subscript denotes traditional measurements.

Chopped measurements were treated differently, in order to obtain the “in-phase” component of the demodulated waveform, which requires knowledge of the timing (measured using phase differences) between the FID and EAG. This timing was determined by measuring the phase differences of high SNR FID-EAG peak pairs. The information was especially important for predicting the phase differences in low-dose, low-SNR data. Phase difference estimates of individual measurements, $\hat{\phi}$, embedded in Euler’s formula, between FID-EAG peak pairs were computed using the following equation,

$$e^{j\hat{\phi}} = \frac{\hat{a}_2 \hat{d}_{FID}^*}{|\hat{a}_2| |\hat{d}_{FID}^*|}, \quad (3)$$

where \hat{d}_{FID} is the complex amplitude measurement made on the FID waveform. Equation (3) says that the phase difference

was obtained by subtracting the phase of \hat{d}_{FID} from that of \hat{a}_2 . The asterisk notation indicates the complex conjugate is taken (or equivalently, the phase should be negated). The vertical lines, $|\cdot|$, indicate that the magnitude of the complex number between the lines should be taken.

Under normal linear conditions, the phase differences are expected to be the same across measurements and the average phase difference should be predictive of all phase measurements. However, due to non-linearity of the EAG response, the phase differences were found to be slightly dose-dependent. A straight-line, log-dose dependent fit to high SNR (doses above 10 pg) phase estimates obtained from data obtained from (3) was used to approximate the expected phase for each measurement. The in-phase component, \hat{a}_ϕ , of the complex amplitude, \hat{a}_2 , was obtained by rotating it into the direction of the real axis and subsequently taking the real component.

$$\hat{a}_\phi = \text{Re} \left\{ \hat{a}_2 \frac{\hat{d}_{FID}^*}{|\hat{d}_{FID}|} e^{-j\hat{\phi}(d)} \right\} = |\hat{a}_2| \cos(\hat{\phi} - \phi(d)), \quad (4)$$

where $\phi(d)$ is the log-dose-dependent straight-line phase estimate. $\text{Re}(\cdot)$ denotes the real component of a complex number. The in-phase EAG amplitude estimate, \hat{a}_ϕ , was normalized using the average noise energy by dividing it by the root mean square (RMS) power estimate of in-phase noise alone, $\hat{\delta}_{vm\phi}$

$$m_2 = \frac{\hat{a}_\phi}{\hat{\delta}_{vm\phi}} \quad (5)$$

Dose-Response Model The dose-response model employed assumed a linear relationship between the log of the median of the probability density of noise-free normalized amplitudes, a_N , and the log of the dose measured on the FID. At a given dose, a_N was considered to be log-normally distributed. Model fitting to the data employed Bayesian MCMC methods that were used to draw values of parameters (e.g., variance and mean of normally distributed data) that describe the model, based on Bayesian inference from the data.

Remaining model assumptions may be found in Appendix 1.

Error Rate Computations

To compare performance, error rates, $(P_{FA} + P_M)/2$, as a function of dose and threshold were estimated based on parameter values for modeled densities drawn from the MCMC simulation. Optimal thresholds as a function of dose were then chosen to minimize the (average) modeled error rate.

To optimize threshold values, 10^4 Monte Carlo simulations, thinned by a factor of 10 to obtain 10^3 uncorrelated samples, were used to estimate expected error rates from the

model parameter distribution through a range of doses and thresholds, choosing and storing the thresholds that minimized the expected error rate for each dose and measurement type. After optimal thresholds for each dose were found, 10^5 Monte Carlo iterations were thinned by a factor of 10 to obtain 10^4 uncorrelated samples of parameters describing possible models. The samples were then used to estimate error rates at optimal thresholds for each dose. Expected values of error rates were obtained by averaging, while 5th and 95th percentiles were obtained by sorting the Monte Carlo error rates. This was accomplished using the methods described in Appendix 2.

Dose Ratio Computations

After determining optimal thresholds for each dose and measurement type, performance comparison of two methods involved sampling the probability density function (PDF) of the ratio between the traditional dosage and chopped dosage, given they had the same error rate. This was accomplished using the methods described in Appendix 2.

Results

Sample Waveforms Sample waveforms, shown in Fig. 4, illustrate the typical EAG traces generated by our experimental protocol employing either chopped or traditional methods illustrated in short (Fig. 4a) and long (Fig. 4b) time scales. Figure 4a includes a 14 s segment of both the recording obtained from the FID and the EAG simultaneously at the relatively high dose of 1.9 ng. It can be seen in the FID waveform that the flow of PEA is eliminated between pulses. Also notable is that the ‘on’ time is less than the ‘off’ time. The duty cycle is 34%, rather than 50%, because flow increases when effluent is exiting the capillary tubing into the odor delivery tube and FID. During each pulse, the rate of material entering both the FID and odor delivery tube increased by about 47% over the rate it entered the Deans switch, but for a shorter period of time. Given equal amounts of compound eluting, a chopped FID elution recording will have an instantaneous maximum peak height about 3 times higher than that of the traditional peak, because it is only eluting about 1/3 of the time. Another effect that is apparent over the time course of the chopping is adaptation. Evidence for adaptation is apparent because of the greater size of the EAG response to the first visible FID pulse compared to the last.

Figure 4b includes two 620 s (10.33 min) recordings under both traditional and chopped conditions. Here, the depolarizations resulting from the experimental sequence described in the methods may be seen. The baselines for both types of recordings are similar and quite steady with rare ‘extraneous’

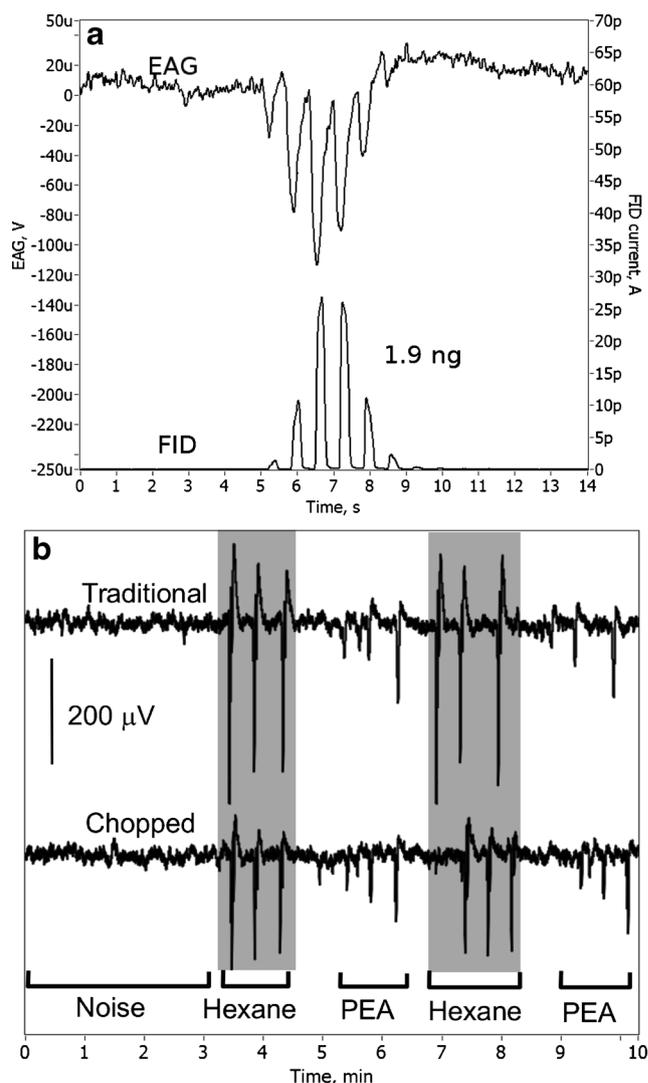


Fig. 4 Sample waveforms. **a** Chopped response to a 1.9 ng dose. **b** Full-length recordings from the same antenna illustrate both the injection sequence and the typical appearance of electroantennogram (EAG) recordings. Note similar baselines. Sets of three depolarizations due to phenethyl alcohol (PEA) are preceded by those due to large amounts of eluting hexane solvent, shaded gray. FID = flame ionization detection

activity. The doses applied to the chopped recording are roughly half of those applied to the traditional recording, due to the experimental setup. It is notable that the amplitudes of the responses under traditional and chopped conditions are similar. However the noise power in the vicinity of the chop frequency is lower, leading to increased SNR that becomes apparent after demodulation and matched filtering (Fig. 6).

Reduction of Chopping-Related Mechanoreceptor Stimulation In Myrick and Baker (2012), the chopping mechanism involved mechanical movement that occurred close to the antenna. Moreover, the antenna was also not ‘shielded’ from external air currents. Mechanoreceptors likely responded to both types of stimulation at that time. The Deans switch was

intended to eliminate both sources of noise introduced by the previous chopping arrangement. Figure 5 is included to compare the baseline noise between the older experiments in Myrick and Baker (2012) and the experiments described here. The amplitude spectral density (ASD) is the square root of the average chopped noise power spectra, obtained by averaging the power spectra from many noise segments in different recordings. In the previous, older experiments, a large amount of power is visible at the chop frequency of 8 Hz. Moreover, “sidebands” that interfere with the EAG signal about the 8 Hz chop frequency are also visible. The power at the chop frequency of 1.5 Hz in the experiments described here is much lower and, most importantly, does not include any sidebands. Harmonics of the chop frequency are visible at 3, 4.5 and 7.5 Hz. In addition, the noise power at lower frequencies is significantly lower, which we attribute to shielding from room air currents rather than differences in species or sex.

Typical Dose-Response Matched-Filtered Waveforms

Equivalent comparisons between traditional and chopped responses may be made using matched filtering. Figure 6 compares the traditional and chopped matched filtered waveforms. The amounts of PEA eluting are not repeatable functions of the amount injected. We attribute this to the injections taking place manually under split conditions in the GC inlet. Presumably, the depth of the needle in the inlet and the speed of the plunger affect how much PEA enters the column. However, the amount of PEA could reliably be measured on the FID. Also visible, between the PEA peaks, are responses to an unknown impurity present in the PEA source. Note, that as a result of the experimental setup, the amount of PEA passing over the antenna during traditional runs is about double the amount during chopped runs. The normalized responses, despite having smaller amounts under chopped

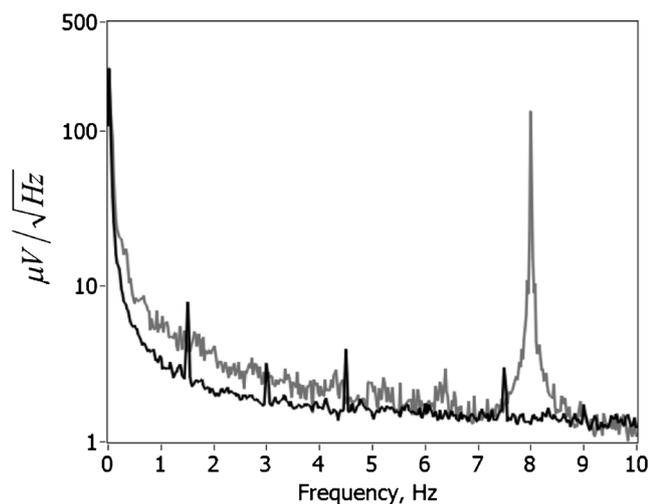


Fig. 5 Noise voltage spectra. The gray line is for experiments for 2012 and the black line for experiments described here. The units of $\mu\text{V}/\sqrt{\text{Hz}}$ arise because power is proportional to the voltage squared

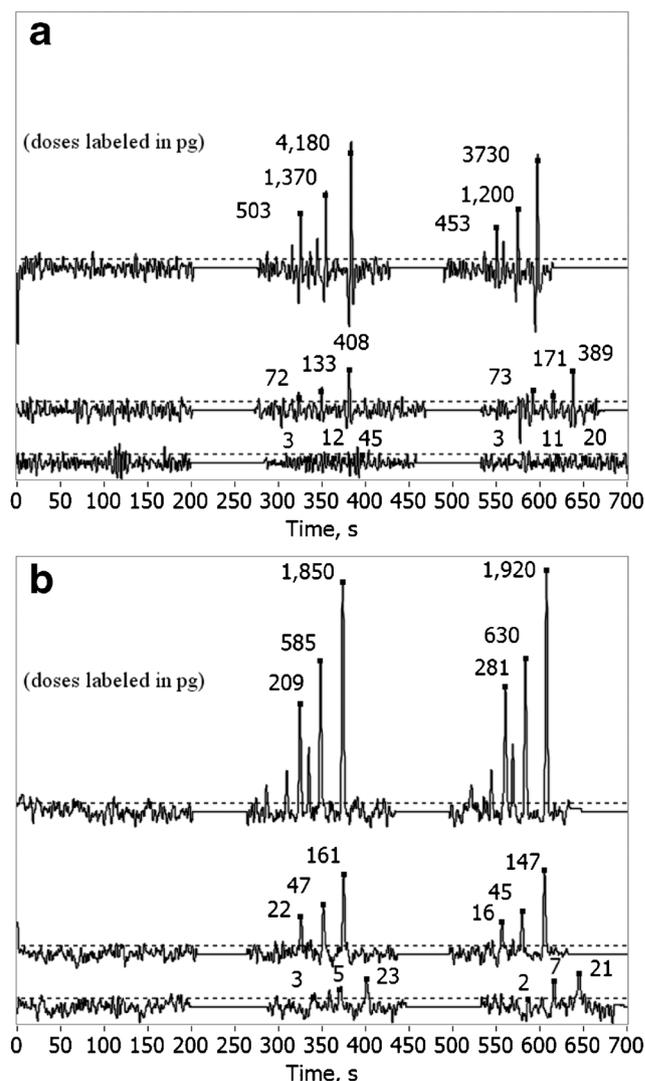


Fig. 6 Representative matched filtered and normalized electroantennogram (EAG) waveforms over a range of about 3 orders of magnitude in dose for both traditional (a) and chopped (b) recordings. Dose range is approximately 4 pg to 4 ng in (a) and 2 pg to 2 ng in (b). Doses applied to traditional recordings are approximately double those applied to chopped recordings, due to the use of only a single output of the Deans switch for both the flame ionization detector (FID) and EAG. Black squares label the PEA elution times obtained from the FID, above which measured doses are labeled in picograms. Hexane solvent responses have been removed for clarity. The horizontal dotted lines show the location of the detection threshold with an associated 5% false alarm rate

conditions, are significantly higher than those obtained under traditional conditions. At the lowest of doses, responses are not readily apparent for the traditional recordings. However the labeled locations on the waveforms are always positive (greater than zero). On the other hand, at the lowest dose range under chopped conditions, the peaks are almost always detectable in the figure. One other notable difference between the two types of measurements is that the chopped response maxima can be more reliably predicted from the FID peak.

Phase Measurements Phase measurements [see Eq. (3)] changed slightly as a function of dose, which had a decreasing lag of about 0.13 rad (7.4°) per decade dose, highlighting the non-linearity of the response. A straight-line, log-dose dependent fit to the phase, $\phi(d)$, used in (4), was accurate to within about 0.2 rad for doses above 0.01 ng. Below about 0.01 ng, individual phase estimates were increasingly corrupted by noise.

Monte Carlo Results The mean log-dose, which appears in Eq. 3A (Appendix 1), had a value of $\ln(0.28 \text{ ng})$ for traditional measurements and $\ln(0.14 \text{ ng})$ for chopped measurements. The SNR improvement is approximated here by the difference between the chopped and traditional methods' SNRs at 1 ng. The difference has an expected value of 11.6 dB (a voltage ratio of about 3.8) and a 90% confidence interval from the 5th percentile value of 10.4 dB to the 95th percentile value of 12.7 dB. A responsive antenna tended to be more responsive in both chopped and traditional types of measurements and *vice versa*, however, the correlation was not very high. The correlation coefficient had an expected value of 0.64 and a 90% confidence interval from the 5th percentile value of 0.32 to a 95th percentile value of 0.86. Other parameters obtained for the model from the Monte Carlo results are summarized in Table 1. Same-measurement type, same-antenna measurement variation about the dose-response curve, σ_A , was found to be roughly half the amount of variation accounted for by different antennal preparations, σ_{AA} . The slopes, in dB per decade of concentration, were also significantly different between measurement types as summarized in Table 1.

Normalized amplitude data (filled circles) vs. dose for both traditional and chopped methods can be seen in Fig. 7. Also shown are the expected value and 90% confidence interval as a function of dose for the “expected” model obtained from Monte-Carlo results. Responses appear to fit well to the power-law model with normalized amplitudes of each type increasing linearly with dose on the log-log plot. Chopped responses have an average normalized amplitude about 4 times higher than traditional responses given the same dose. Chopped responses also have an expected dose about 50 times lower, given the same normalized amplitude.

Table 1 Traditional and chopped model parameter ranges

| Traditional | Mean | 5th | 95th | Chopped | Mean | 5th | 95th |
|--------------------|------|------|------|--------------------|------|------|------|
| slope, dB/dec | 8.71 | 8.15 | 9.28 | slope, dB/dec | 7.41 | 6.99 | 7.83 |
| 1 ng response, dB | 18.5 | 17.2 | 19.8 | 1 ng response, dB | 30.1 | 28.9 | 31.2 |
| σ_A , dB | 1.35 | 1.14 | 1.56 | σ_A , dB | 1.20 | 1.08 | 1.35 |
| σ_{AA} , dB | 2.78 | 1.89 | 4.00 | σ_{AA} , dB | 2.60 | 1.87 | 3.58 |

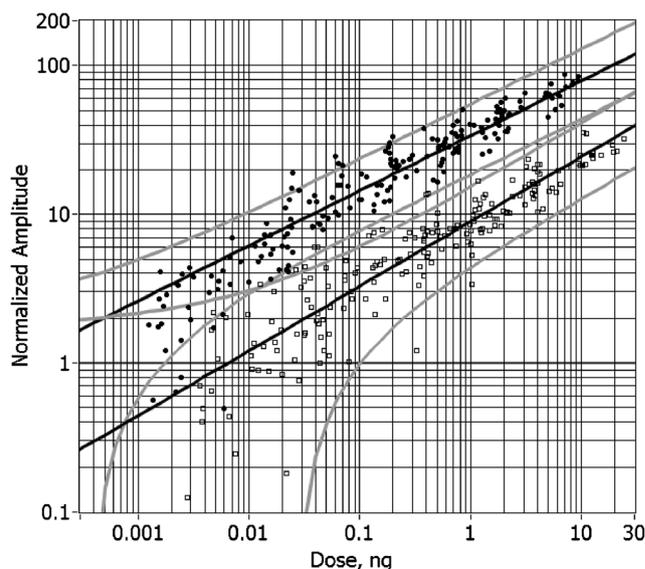


Fig. 7 Electroantennogram dose-response data and expected model. Traditional and chopped responses are represented by open square and filled black circles, respectively. Straight, dark lines indicate the expected normalized amplitude as a function of dose. The lighter curves represent the 5–95% confidence interval about the expected normalized amplitude

In Fig. 8a, it is apparent that predicted error rates associated with chopped measurements are well below those associated with traditional measurements. Assigning a single-depolarization detection limit as that which has an error rate of 5%, the expected detection limit of traditional measurements is 200 pg, while the expected detection limit of chopped measurements is 3.9 pg. Thus, the expected detection limit is reduced by about a factor, or “dose ratio”, of 50.6. However, there is a large amount of uncertainty in the dose ratio, as seen in Fig. 8b. There, at a traditional dose of 200 pg, the ratio from the 5th to 95th percentile can be from 31.2–82.6. The dose ratio at different error rates may also be read off the graph. For instance, at a 1% error rate, corresponding to the traditional dose of 800 pg, the dose ratio ranges from approximately 20–80.

Discussion

Our results indicate that chopping in the circumstances described here results in a reduction of the detection limit of GC-induced electroantennograms by a factor ranging from

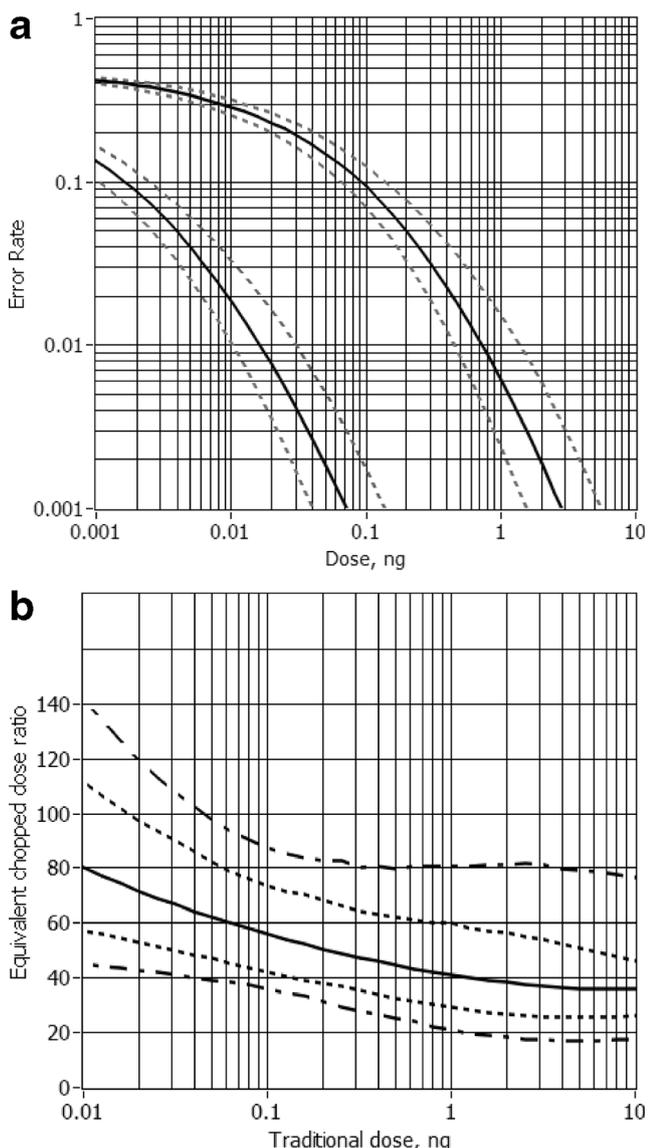


Fig. 8 Predicted optimal error rate and dose ratio as a function of dose. Panel A plots expected error rate vs. dose for traditional and chopped methods. Solid line: expected value of the error rate. Dotted lines are the 5th and 90th percentiles of the expected error rate. Panel B illustrates the 'dose ratio' vs. traditional dose. See text for details. Solid line: expected value. Inner dotted lines are the 16th and 84th percentiles, while outer dotted lines are the 5th and 95th percentiles

31.2–82.6, with 90% confidence. The reduction in the detection limit is accompanied by an approximately 11.6 dB improvement in SNR (at 1 ng).

Although it remains to be determined, the detection limit of the instrument is perhaps approaching that of the insects themselves. Mayer et al. (1987) investigated behavior thresholds in insects for various compounds, and reasoned based on behavioral thresholds for pheromones that the behavioral threshold of non-pheromonal compounds was at most about 100 times lower than the (traditional) EAG threshold. With a dose-ratio of about 50, the instrument here has improved detection by a

similar factor. Mayer et al. (1987) estimated the non-pheromonal behavioral detection limit for insects in general to be in the range of 10^{-8} $\mu\text{mol}/\text{cm}^3$. For comparison, 10^{-8} $\mu\text{mol}/\text{cm}^3$ corresponds to a dose in this system of about 15 pg of PEA, which is close to the 3.9 pg computed detection limit of GC-EAD here.

In addition to the SNR increase, an advantage of chopping is that it slightly increases the resolution of GC-EAD. In traditional waveforms, information about the FID waveform shape at low frequencies (long time scales) is obscured by noise. When optimally filtered, low frequencies are removed where the SNR is low. As a result, the width of the band of frequencies where the signal power was significant (bandwidth), for the resulting depolarizations is reduced. As a rule of thumb, the duration of a signal and its bandwidth are inversely related. The duration of the filtered waveform increases when bandwidth is reduced. In Fig. 6a, the overshoot and ringing are evidence of the increased duration of the waveform. As a result, neighboring filtered responses may tend to overlap more, reducing the time resolution of the EAG.

The matched-filtered traditional and chopped EAG waveforms shown in Fig. 6 do not compare the same dose range. Note, that under chopped conditions, only half of the effluent enters the splitter, whereas all of the effluent enters the splitter under traditional conditions in our setup (Fig. 2). It is for this reason that we recommend modification of the setup to use the Deans switch itself as a splitter for traditional recordings and simply alternate the effluent between the EAG and FID for chopped recordings.

Previously (Gouinguéné et al. 1998; Myrick and Baker 2011), peak tailing on EAGs has been observed during chopping (pulsing) experiments. This was possibly due to reversible condensation of the effluent on the column exit or on the walls of the odor delivery tube. Odor re-released from the internal surfaces of the odor delivery tube may have continued to flow over the antenna after the compound had eluted from the column. Here, on the FID, no peak tailing is evident on the timescale of an FID peak in the EAG (Fig. 4a). This may be due to differences in the design of the odor delivery tube, having a short capillary column exiting the heated transfer line, or the makeup gas maintaining a high temperature there. Because of the short time scale of the pulsing, careful attention to odor delivery may be important to the implementation of chopped GC-EAD.

The SNR is given by the normalized amplitude squared. As a result, normalized amplitudes plotted in Fig. 7 are illustrative of the SNR increase obtainable by chopping and demodulating. It is worth mentioning that the setup was disassembled for different experiments, and then re-assembled for doses below about 20 pg (chopped) and 40 pg (traditional). Effects of changes to the setup are possibly evident, because responses for both measurement types appear to drop slightly at those doses. Models to the data may have had

slightly increased slopes, wider σ_{AA} estimates, and slightly lowered sensitivities. It appears that the effect, if present, is stronger on the chopped responses because they were associated with higher SNR.

Because the SNR increases slowly as dose is increased, small changes in SNR result in large changes to the dose associated with a similar SNR (Fig. 7). SNR is strongly associated with the error rate here; comparison of the expected error rates for both measurement types showed a similar dose-ratio of about 50 (Fig. 8). However the error rate is more strongly affected by SNR, where the error rates associated with the same dose have a ratio of up to 150 or even more, in comparison to the SNR, which is about 15, corresponding to an amplitude ratio of only about 4.

Detection of very low sensitivity (though not resolution) compounds in chopped GC-EAD analyses would likely be aided by large-capacity, long elution time columns, when large samples are available, whereas for traditional GC-EAD, long elution times are a disadvantage. The reason for the disadvantage in traditional recordings is that longer elution time peaks occupy low frequencies, where the EAG noise is most abundant. In contrast to long EAG waveforms being obscured by noise in traditional EAG waveforms, chopping longer waveforms increases the amount of time that a chopped signal is present, without increasing noise, thus allowing more noise to be averaged out. For a given dose, much of the advantage obtained by increasing the averaging time would be countered by the reduction in the amplitude of the EAG, making the main advantage the increased capacity of large ID or packed columns. A hindrance to detecting long peaks, however, is the adaptation that occurs even when the effluent is chopped (Fig. 4a).

More improvement to the method may possibly be achieved by increasing the high flow to low flow ratio through the transfer line capillary, so that the resulting pulses are narrower and more concentrated. This would be expected to benefit most antennal preparations that are able to respond quickly to stimulus pulses.

Another area of possible improvement to GC-EAD is in the signal processing and methods for EAG depolarization detection and/or quantization, e.g., p -value estimation. It should be noted that the matched-filtering signal processing here was used to aid evaluation of detection performance and is, despite its usefulness, not yet general purpose enough to be used to fully automate GC-EAD waveform analysis. Simple matched filtering lacks the capacity to distinguish electrophysiological signal resulting from stimulus delivered via the GC capillary from other types of disturbance to the EAD baseline. Moreover it does not handle overlapping responses. Nevertheless, matched filtering is a useful tool for the analysis of GC-EAD waveforms. In this context, the use of matched filtering as implemented here may be hindered by the

requirement to estimate the autocorrelation, or equivalently, the noise spectrum. Because the autocorrelation estimate itself is uncertain, and because effects, such as the decline of the antenna, are not accounted for, it is possible the number of parameters to estimate could be reduced without sacrificing accuracy. For instance, the frequency spectrum of EAG waveforms can be assumed to be Brownian, or some other simple spectral distribution, thus removing the need to measure it. Making this assumption, for traditional recordings, it would be possible to construct a matched filter using only a single parameter, the EAG peak width. Signal processing that processes multiple waveforms would be a natural next step for GC-EAD. A similar problem has been treated in radar applications (Heidbreder and Mitchell 1967).

It is not uncommon for advances in instrumentation to have unforeseen applications. For instance, it is possible to detect the presence of pheromonal compounds in environmental samples at reduced trace levels, perhaps for population monitoring or early detection of known pests.

Acknowledgements This work was supported by the National Science Foundation under Grant Number DBI-1353870.

Appendix 1

Noise-corrupted measurements, m_t , where t refers to measurement type (traditional is represented by 1 and chopped by 2) were assumed to be normally distributed with a standard deviation of 1 about the true normalized signal amplitude, a_{Nt} .

$$f_{Mt|A}(m_t|a_{Nt}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(m_t - a_{Nt})^2}{2}} \quad (1A)$$

The logs of a pair of the true normalized amplitudes, $\mathbf{a}_{Nt} = [\ln(a_{N1}), \ln(a_{N2})]$ on an individual antenna were assumed to be independently normally distributed about the antenna's dose-dependent medians, $\boldsymbol{\mu} = [\mu_1, \mu_2]$. The order of presentation of measurement types and the age of the preparation were not accounted for.

$$f_{AL}(\mathbf{a}_{Nt}|\boldsymbol{\mu}, \boldsymbol{\Sigma}_A) = \frac{1}{2\pi \sqrt{|\boldsymbol{\Sigma}_A|}} e^{-\frac{1}{2}(\mathbf{a}_{Nt} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_A^{-1} (\mathbf{a}_{Nt} - \boldsymbol{\mu})} \quad (2A)$$

where $\boldsymbol{\Sigma}_A = \begin{bmatrix} \sigma_{A1}^2 & 0 \\ 0 & \sigma_{A2}^2 \end{bmatrix}$. σ_{A1} describes the standard deviation of the true normalized amplitudes about the antenna's dose-dependent median, whereas σ_{A2} describes the standard deviation associated with chopped measurements. Note that the same variance was assumed for all antennae.

The equation describing the dose-dependent medians places the y intercept at the mean experimental dose. Thus, for an antenna, the medians are given by

$$\boldsymbol{\mu} = \boldsymbol{\beta}_1 (\mathbf{d}_l - \mathbf{d}_\mu) + \boldsymbol{\beta}_0, \tag{3A}$$

where $\boldsymbol{\beta}_1 = \begin{bmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{bmatrix}$ describes the assumed independent slope parameters for each measurement type, $\mathbf{d}_l = [\ln(d_1), \ln(d_2)]$ is the log-dose in ng and $\boldsymbol{\beta}_0$ describes the responsiveness of the antenna to the PEA stimulant for each measurement type at the mean log-dose, \mathbf{d}_μ . Because measurement types are paired on each antenna and covary, $\boldsymbol{\beta}_0$ is modeled as a joint bivariate normal variable that varies about the mean of the response at the mean experimental log-dose, $\bar{\mathbf{b}}$.

$$f_{\boldsymbol{\beta}_0}(\boldsymbol{\beta}_0 | \bar{\mathbf{b}}, \Sigma_{AA}) = \frac{1}{2\pi\sqrt{|\Sigma_{AA}|}} e^{-\frac{1}{2}(\boldsymbol{\beta}_0 - \bar{\mathbf{b}})^T \Sigma_{AA}^{-1} (\boldsymbol{\beta}_0 - \bar{\mathbf{b}})} \tag{4A}$$

In the model, all parameters were assigned non-informative prior distributions, with the exception of the Wishart prior density used for the covariance matrix $\Sigma_{AA} = \begin{bmatrix} \sigma_{AA1}^2 & \sigma_{AA12} \\ \sigma_{AA21} & \sigma_{AA2}^2 \end{bmatrix}$. Parameters to be inferred include $\mathbf{P} = \{\boldsymbol{\beta}_1, \Sigma_A, \Sigma_{AA}, \bar{\mathbf{b}}\}$. The complete model is summarized below.

$$\left. \begin{aligned} & a_{N1\dots2,i,j} \sim N\left(\mu_{1\dots2,i,j}, \begin{bmatrix} \sigma_{A11} & 0 \\ 0 & \sigma_{A22} \end{bmatrix}\right) \\ & \mu_{1\dots2,i,j} = \beta_{1,1\dots2,1\dots2} d_{1\dots2,i,j} + \beta_{0,1\dots2,j} \\ & m_{1\dots2,i,j} \sim N(a_{N1\dots2,i,j}, 1) \\ & a_{N1\dots2,i,j} = \exp(a_{M1\dots2,i,j}) \\ & \beta_{0,j,1\dots2} \sim N(\bar{b}_{1\dots2}, \Sigma_{AA1\dots2,1\dots2}) \\ & \bar{b}_{1\dots2} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10^6 & 0 \\ 0 & 10^6 \end{bmatrix}\right) \\ & \beta_{1,1\dots2} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 10^6 & 0 \\ 0 & 10^6 \end{bmatrix}\right) \\ & \sigma_{A1} \sim \text{Gamma}(10^{-3}, 10^{-3}) \\ & \sigma_{A2} \sim \text{Gamma}(10^{-3}, 10^{-3}) \\ & \Sigma_{AA,1\dots2,1\dots2} \sim \text{Wishart}\left(\begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, 2\right) \end{aligned} \right\} \begin{array}{l} 1 \leq i \leq 12 \\ 1 \leq j \leq 16 \end{array}$$

Appendix 2

The analysis here closely resembles the one in Myrick and Baker (2012), however uses analytical PDFs. In the detection problem (Trees 2001), given a measurement m , (of traditional

or chopped types), it is desired to distinguish between two hypotheses H_0 and H_1 ,

$$\begin{aligned} H_1 &: m = a_N + n \\ H_0 &: m = n \end{aligned}, \tag{1B}$$

where H_1 represents the hypothesis that a noise-corrupted EAG response is present and H_0 that only noise, n , is present. A tool for making a decision between H_0 and H_1 is known as the likelihood ratio test (LRT). The likelihood ratio, $\Lambda(m)$, can be written as a ratio of probability densities associated with each hypothesis.

$$\Lambda(m) = \frac{f_M(m|a_N > 0)}{f_M(m|a_N = 0)} \underset{< H_0}{\overset{> H_1}{>}} \eta \tag{2B}$$

A threshold, η , can be applied to the likelihood ratio to decide whether a measurement can be attributed to H_0 or H_1 . Here the likelihood ratio increases monotonically with m , so it is possible to simply apply a threshold to the measurement itself.

$$m \underset{< H_0}{\overset{> H_1}{>}} T \tag{3B}$$

The optimal threshold chosen here is that which minimizes the error rate associated with the posterior predictive distribution, assuming H_0 and H_1 are equally probable. We evaluate distributions of possible error rates as a function of the distribution of model parameters, \mathbf{P} . As a result, error rate, P_E , is computed for each sample drawn, which is half the sum of the false alarm rate, P_{FA} and the miss rate, P_M . Error rate is also dependent on log-dose, d_l , and threshold, T .

$$P_E(T, d_l, \mathbf{P}) = (P_M(T, d_l, \mathbf{P}) + P_{FA}(T))/2 \tag{4B}$$

The false alarm rate is independent of dose and threshold because it describes the chances that noise is above the threshold. We seek the threshold that minimizes the error rate associated with the posterior predictive distribution.

$$P_E(T, d_l) = E_{\mathbf{P}}[P_E(T, d_l, \mathbf{P})] = \int_{\mathbf{P}} f_{\mathbf{P}}(\mathbf{P}) P_E(T, d_l, \mathbf{P}) d\mathbf{P} \tag{5B}$$

The optimal threshold, T^* , is found as a function of the log-dose by minimizing the error rate obtained from (5B).

$$T^*(d_l) = \underset{T}{\operatorname{argmin}} (P_E(T, d_l)) \tag{6B}$$

The optimal error rate can be computed by substituting the optimal threshold into (4B), removing dependence on threshold and resulting in the following expression for the optimized error rate as a function of log-dose.

$$P_E^*(d_l, \mathbf{P}) = (P_M^*(d_l, \mathbf{P}) + P_{FA}^*(d_l))/2 \tag{7B}$$

It is the error rate drawn from (7B) that is used to compare performance between the measurement types.

P_{FA} . The probability of error due to a false-alarm, given there is no EAG response present, is the probability that normally distributed noise will be above the threshold, T . The false alarm rate is given by,

$$\Pr(m_t > T) = 1 - F_M(T|a_N = 0) = \frac{1}{2} \left[1 + \operatorname{erfc} \left(\frac{T}{\sqrt{2}} \right) \right], \quad (8B)$$

where the cumulative distribution function (CDF) is indicated by the capital F . The CDF in (8B) is obtained by integrating the normal density describing noise in (1A).

P_M . The probability of error due to a miss, P_M , given there is an EAG present, is the probability that its associated measurement is less than the threshold, T . It is most convenient computationally to compare the performance of antennae independently selected for each measurement type, rather than many single randomly selected antennae whose performance

is compared for each measurement type. Comparing individual antennae is further complicated by the Bayesian model, which did not take into account the order of presentation of measurement types or the age to the measurements into account. Anecdotally, the order of presentation had little effect. Fortunately, the correlation of the performance of an antenna between the two measurement types is low, which means that limited information can be obtained by comparing the performance of single antennae. In the case of drawing different antennae (rather than a different run on the same antenna), it is necessary to “integrate over” the distribution of antennal responsiveness, β_0 .

$$f_{AL}(\mathbf{a}_{NI}|\mathbf{P}, \mathbf{d}_t) = \iint_{\mathbb{R}^2} f_{AL}(\mathbf{a}_{NI}|\beta_0, \beta_1, \Sigma_A, \mathbf{d}_t) f_{\beta_0}(\beta_0|\bar{\mathbf{b}}, \Sigma_{AA}) d^2\beta_0 \quad (9B)$$

After the substitution of (2A), (3A), and (4A) into (9B), the integral can be fully written out as

$$= \frac{1}{2\pi} \iint_{\mathbb{R}^2} \frac{1}{\sqrt{|\Sigma_A|}} e^{-\frac{1}{2}(\mathbf{a}_{NI} - (\beta_1(\mathbf{d}_t - \mathbf{d}_\mu) + \beta_0))^T \Sigma_A^{-1} (\mathbf{a}_{NI} - (\beta_1(\mathbf{d}_t - \mathbf{d}_\mu) + \beta_0))} \frac{1}{\sqrt{|\Sigma_{AA}|}} e^{-\frac{1}{2}(\beta_0 - \bar{\mathbf{b}})^T \Sigma_{AA}^{-1} (\beta_0 - \bar{\mathbf{b}})} d^2\beta_0. \quad (10B)$$

Solving the integral in (10B) has the following result.

$$f_{AL}(\mathbf{a}_{NI}|\mathbf{P}, \mathbf{d}_t) = \frac{1}{2\pi\sqrt{|\Sigma_A + \Sigma_{AA}|}} e^{-\frac{1}{2}(\mathbf{a}_{NI} - (\beta_1(\mathbf{d}_t - \mathbf{d}_\mu) + \bar{\mathbf{b}}))^T (\Sigma_A + \Sigma_{AA})^{-1} (\mathbf{a}_{NI} - (\beta_1(\mathbf{d}_t - \mathbf{d}_\mu) + \bar{\mathbf{b}}))} \quad (11B)$$

Only the two marginal densities in the bivariate distribution (11B) are needed to construct thresholds for each measurement type on two randomly selected antennae and test them independently for a given sample. The marginal densities for the log-normalized amplitudes corresponding to the two measurement types in (11B) are normally distributed and can be written

$$f_{ALt}(a_{NI,t}|\theta_t, d_t) = \int_{-\infty}^{\infty} f_{AL}(\mathbf{a}_{NI}|\mathbf{P}, \mathbf{d}_t) da_{NI(3-t)} \quad (12B)$$

$$= \frac{1}{\sigma_{at}\sqrt{2\pi}} e^{-\frac{(a_{NI} - \mu_t(d_t))}{2\sigma_{at}^2}},$$

where the variance is the corresponding variance in $\Sigma_A + \Sigma_{AA}$,

$$\sigma_{at}^2 = (\sigma_{At}^2 + \sigma_{AA_t}^2), \quad (13B)$$

and the mean is linearly related to the log of the dose,

$$\mu_t = \beta_{1t}(d_t - d_{\mu}) + \bar{b}_t. \quad (14B)$$

$\theta_t = (\beta_{1t}, \bar{b}_t, \sigma_{at})$ is the parameter vector derived from the sample, \mathbf{P} , drawn from the Monte Carlo simulation for the measurement type, t . (12B) describes the log of a_N , a_{NI} , as normally distributed, thus a_N is log-normally distributed.

$$f_A(a_N|\theta, d_t) = \frac{1}{a_N\sigma_a\sqrt{2\pi}} e^{-\frac{(\ln(a_N) - \beta_1(d_t - d_\mu) - \bar{b}))^2}{2\sigma_a^2}} \quad (15B)$$

In (15B), the t subscript denoting the measurement type has been removed, but it is implied in the following text. To obtain

the density of the noise-corrupted measurements, integration over a_N is required.

$$f_M(m|\theta, d_l) = \int_{-\infty}^{\infty} f_A(a_N|\theta, d_l) f_M(m|a_N) da_N \quad (16B)$$

Integrating again, over m , and then exchanging the order of integration yields the CDF of unknown signal plus noise,

$$F_M(T|\theta, d_l) = \int_{-\infty}^{\infty} \int_{-\infty}^T f_M(m|a_N) f_A(a_N|\theta, d_l) dm da_N. \quad (17B)$$

The CDF of known signal plus noise appears in (17B), and may be evaluated first.

$$F_M(T|a_N) = \int_{-\infty}^T f_M(m|a_N) dm, \quad (18B)$$

where the integrand refers to (1A). The computed results of (18B) remain the same for all draws from the simulation and may be stored in lookup tables. Substituting (18B) into (17B) has the result

$$\begin{aligned} P_M(T, d_l, \mathbf{P}) &= F_M(T|\theta, d_l) \\ &= \int_{-\infty}^{\infty} F_M(T|a_N) f_A(a_N|\theta, d_l) da_N. \end{aligned} \quad (19B)$$

The equation above may be substituted into (4B) or (5B). (18B), which can be written

$$F_M(T|a_N) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{T - a_N}{\sqrt{2}} \right) \right), \quad (20B)$$

is a function of $T - a_N$ and (19B) can be considered a convolution integral. Denoting the distribution of the parameters based on the observed data as simply $f_{\theta}(\theta)$, after substituting (19B) into (5B), the following expression is obtained.

$$P_E(T, d_l) = \frac{\int_{-\infty}^{\infty} F_M(T|a_N) \int_{\theta} f_{\theta}(\theta) f_A(a_N|\theta, d_l) d\theta da_N + P_{FA}(T)}{2} \quad (21B)$$

The equation above can be simplified. The inner integral marginalizes the lognormal density, $f_A(a_N|\theta, d_l)$, over θ .

$$f_A(a_N|d_l) = \int_{\theta} f_{\theta}(\theta) f_A(a_N|\theta, d_l) d\theta \quad (22B)$$

Substitution of (22B) into (21B) results in the following equation.

$$P_E(T, d_l) = \frac{\int_{-\infty}^{\infty} F_M(T|a_N) f_A(a_N|d_l) da_N + P_{FA}(T)}{2} \quad (23B)$$

(23B) is substituted into (6B) to optimize the threshold as a function of dose. The optimal error rate can then be written as a function of dose and distribution parameters, \mathbf{P} , by substituting the optimal threshold (6B) back into (19B).

$$P_E(d_l, \mathbf{P}) = \frac{\int_{-\infty}^{\infty} F_M(T^*(d_l)|a_N) f_A(a_N|\theta, d_l) da_N + P_{FA}(T^*(d_l))}{2} \quad (24B)$$

P_{FA} computations

In the implementation, Eq. (8B) was computed and stored in increments of $\Delta T = 0.1$.

P_M computations

$F_M(T|a_N)$ given by Eq. (20B) was computed and stored in increments of $\Delta(T - a_N) = 0.1$. The log-normal density of normalized amplitudes, $f_A(a_N|\theta, d_l)$, (15B), was marginalized over θ by averaging samples indexed by n , creating a series of average normalized amplitude densities at each log-dose, indexed by k .

$$\%f_A(a_{Ni}|d_{lk}) = E_{\theta}[f_A(a_{Ni}|\theta, d_{lk})] \cong \frac{1}{1000} \sum_{n=0}^{999} f_A(a_{Ni}|\theta_n, d_{lk}) \quad (25B)$$

where $a_{Ni} = i\Delta a_N$ and $\Delta a_N = 0.1$ for $a_{Ni} = -5$ to 20. These log-doses, (corresponding to doses in ng), used for the computations were given by

$$d_{1,k} = \ln \left(10^{(0.1k-3.0)} \right) = \ln(10)(0.1k-3.0) \text{ (1 pg to 79 ng)}$$

$$d_{2,k} = \ln \left(10^{(0.1k-4.5)} \right) = \ln(10)(0.1k-4.5) \text{ (0.032 pg to 2.5 ng)}$$

for k from 0 to 49. Data from tables were convolved to compute log-dose dependent missed detection rate functions as a function of threshold, indexed by j . The discrete convolution operation in (23B) is given by

$$P_M(T_j, d_{lk}) = F_M(T_j|d_{lk}) \cong \sum_{i=-50}^{199} f_A(a_{Ni}|d_{lk}) F_M(T_j|a_{Ni}) \Delta a_N \quad (26B)$$

Thresholds were tested in increments of 0.1, and the minimum, near-optimal value was stored for each dose,

$$T^*(d_{lk}) = \operatorname{argmin}_{T_j} (F_M(T_j|d_{lk}) + P_{FA}(T_j)), \quad (27B)$$

The discrete convolution in (24B) was computed in the following manner.

$$\begin{aligned} P_M^*(d_{lk}, \mathbf{P}_n) &= F_M(T_j^*(d_{lk})|\theta_n, d_{lk}) \\ &\cong \sum_{i=-50}^{199} f_A(a_{Ni}|\theta_n, d_{lk}) F_M(T_j^*(d_{lk})|a_{Ni}) \Delta a_N \end{aligned} \quad (28B)$$

yielding samples of the optimized missed detection probability as a function of dose. These, added to P_{FA} resulted in samples of P_E (7B).

Dose Ratio Computations Performance comparison of two methods involved sampling the PDF of the ratio between the traditional dosage and chopped dosage given the same optimized error rate and a uniform prior density for $\ln(d_2)$, where d_1 and d_2 refer to the doses measured on the FID using traditional and chopped methods, respectively. That PDF may be written

$$f_R(d_1 d_2^{-1} | d_1). \tag{29B}$$

For every sample drawn from the Monte Carlo simulation, enumerated by n , the chopped dose with the same optimally thresholded error rate as that from each traditional dose was found. Thus, for the k^{th} log-dose, and the n^{th} sample,

$$P_{E2}^*(d_{l2,k,n}, \theta_{2n}) = P_{E1}^*(d_{l1,k}, \theta_{1n}). \tag{30B}$$

After defining

$$P_{E2n}^*(d_l) = P_{E2}^*(d_l, \theta_{2n})$$

and

$$P_{E1n}^*(d_l) = P_{E1}^*(d_l, \theta_{1n}),$$

an expression for the chopped dose corresponding to the k^{th} traditional log-dose and n^{th} sample can be obtained by rearrangement.

$$d_{l2,k,n} = P_{E2n}^{*-1}(P_{E1n}^*(d_{l1,k})). \tag{31B}$$

The optimized error rate was always a monotonically decreasing function of dose, making its inverse (the log-dose) single valued as a function of error rate. In the implementation, the inverse error rate (corresponding chopped log-dose) was found by linearly interpolating the inverse chopped error rate, P_{E2n}^{*-1} , for all k and n . After finding d_{l2} , the "dose ratio" can be sampled by solving for d_1/d_2 .

$$\left(\frac{d_1}{d_2}\right)_{k,n} = \exp(d_{l1,k} - d_{l2,k,n}) \tag{32B}$$

10^4 draws of the dose ratio were made for the each traditional dose enumerated by k (25B)). The samples were sorted and the 5th, 16th, 84th and 95th percentiles were found.. The average value was also computed at each dose.

Because the measurement types were both made on the same antennae, the mean SNR estimates, \bar{b}_l , of the measurement types are also correlated. Thus, the measurement types are not completely separate when threshold-optimized error rates are computed for comparison. This simply reflects that uncertainty in the sensitivity of the average antenna is correlated due to the way measurements were made.

Appendix 3

Three valves were connected as illustrated in Fig. 9. Two 91.4 cm (3 ft.) pieces of flexible rubber tubing with an ID of 1.59 mm (1/16 in.) and outer diameter (OD) of 3.18 mm (1/8 in.) were used as a reservoir to provide back-pressure (13.1 kPa), designed to limit effluent flow out of the low-pressure auxiliary line. The back pressure could be adjusted by plugging the reservoirs at different locations. Interconnects were made using nylon tubing with an OD of 3.18 mm (1/8 in.). While traditional measurements were being made (steady state flow), the relay positions remained in the configuration labeled C (Fig. 9) for the duration of the GC run.

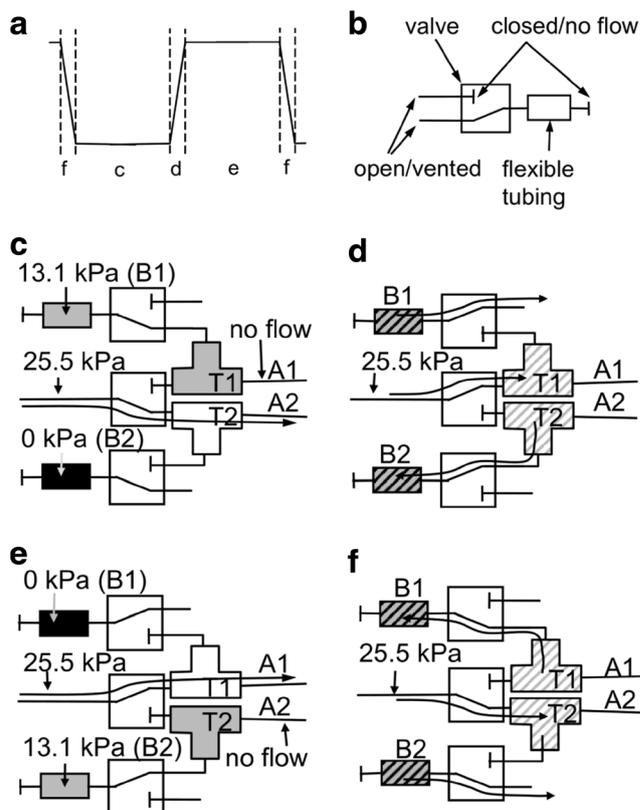


Fig. 9 Operation of valve system used to control pressure to Deans switch auxiliary lines. **a** Illustration of pressure at auxiliary line 1 (A1) over time. Letters correspond to figure items that depict the labeled time period. **b** Symbol key. Closed empty shapes indicate tubing containing significant volume. **c** Steady pressures are maintained at auxiliary line 1 (A1) and line 2 (A2) leading to the Deans switch. Gray color indicates backpressure tubing 1 (B1) that is pressurized at 13.1 kPa, preventing backflow from the Deans switch through A1. White T2 is held at a regulated pressure of 25.5 kPa, driving nitrogen into the Deans switch through A2. Backpressure line 2 (B2) is vented and therefore at 0 kPa. **d** Transient flows (only) immediately after valve positions are switched. Flow from 25.5 kPa input increases the pressure of T1. T2, initially at 25.5 kPa, escapes into B2, eventually reaching 13.1 kPa. B1 vents to the atmosphere. **e** Alternate steady pressure state which is symmetrical with that illustrated in panel C. **f** Transient flows occurring as pressure is brought back to state in panel C

References

- Andreadis SS, Cloonan KK, Myrick AJ, Chen H, Baker TC (2015) Isolation of a female-emitted sex pheromone component of the fungus gnat, *Lycoriella Ingenua*, attractive to males. *J Chem Ecol* 41: 1127–1136
- Boeker P, Leppert J, Mysliweitz B, Lammers PS (2013) Comprehensive theory of the deans' switch as a variable flow splitter: fluid mechanics, mass balance, and system behavior. *Anal Chem* 85:9021–9030
- Bowyer MC, Slippers B, Degefu D, Wingfield MJ, Lawson S, Rohwer ER (2015) Identification of the sex pheromone of the tree infesting Cossid moth *Coryphodema tristis* (Lepidoptera: Cossidae). *PLoS One* 10:e0118575
- Deans DR (1968) A new technique for heart cutting in gas chromatography. *Chromatographia* 1:18–22
- French AS, Meisner S, CY S, Torkkeli PH (2014) Carbon dioxide and fruit odor transduction in drosophila olfactory neurons. What controls their dynamic properties? *PLoS One* 9:e86347. <https://doi.org/10.1371/journal.pone.0086347>
- Gelman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis, 2nd edn. Chapman & Hall/CRC, Boca Raton
- Gouinguené S, Id C, Jvd P, Wadhams L, Marion-Poll F (1998) A new method to improve olfactory responses to GC effluents. *Chem Senses* 23:647–652
- Guerin PM, Stadler E, Buser HR (1983) Identification of host plant attractants for the carrot fly, *Psila rosae*. *J Chem Ecol* 9:843–861
- Heidbreder GR, Mitchell RL (1967) Detection probabilities of log-normally distributed signals. *IEEE T Aero Elec Sys* AES-3:5–13
- Helstrom CW (1960) Statistical theory of signal detection vol 9. International series of monographs on electronics and instrumentation. Pergamon Press, New York
- Marion-Poll F, Thiéry D (1996) Dynamics of EAG responses to host-plant volatiles delivered by a gas chromatograph. *Entomol Exp Appl* 80:120–123
- Mayer MS, Mankin RW, Grant AJ (1987) Quantitative comparison of behavioral and neurophysiological responses of insects to odorants: inferences about central nervous system processes. *J Chem Ecol* 13: 509–531
- Michels WC, Curtis NL (1941) A pentode lock-in amplifier of high frequency selectivity. *Rev Sci Instrum* 12:444–447
- Moore I (1981) Biological amplification for increasing electroantennogram discrimination between two female sex pheromones of *Spodoptera littoralis* (Lepidoptera: Noctuidae). *J Chem Ecol* 7:791–798
- Moorhouse JE, Yeadon R, Beavor PS, Nesbitt BF (1969) Method for use in studies of insect chemical communication. *Nature* 223:1174–1175
- Myrick AJ, Baker TC (2011) Chopper stabilized gas chromatography - electroantennography part I: background, signal processing and example. *Biosens Bioelectron* 31:197–204
- Myrick AJ, Baker TC (2012) Chopper-modulated locked in amplified gas chromatography - electroantennography part II: signal processing and performance comparisons. *IEEE Sensors J* 12:2974–2983. <https://doi.org/10.1109/JSEN.2012.2206380>
- Ntzoufras I (2009) Bayesian modeling using WinBUGS. John Wiley & Sons, Hoboken
- Pan H et al (2015) Volatile fragrances associated with flowers mediate host plant alternation of a polyphagous mirid bug. *Sci Rep* 5:14805
- Park KC, Baker TC (2002) Improvement of signal-to-noise ratio in electroantennogram responses using multiple insect antennae. *J Insect Physiol* 48:1139–1145
- Park KC, Hardie J (1998) An improved aphid electroantennogram. *J Insect Physiol* 44:919–928
- Robey FC, Fuhrmann DR, Kelly EJ, Nitzberg R (1992) A CFAR adaptive matched filter detector. *IEEE T Aero Elec Sys* 28:208–216
- Schiestl FP, Marion-Poll F (2002) Detection of physiologically active flower volatiles using gas chromatography coupled with electroantennography. In: Jackson JF, Linskens HF (eds) Analysis of taste and aroma, vol 21. Molecular methods of plant analysis. Springer, Berlin
- Schneider D (1957) Electrophysiologische Untersuchungen von Chemo- und Mechanorezeptoren der Antennae des Seidenspinner *Bombix mori* L. *Z Vergle Physiol* 40:8–41
- Schott M, Wehrenfennig C, Gasch T, Düring R-A, Vilcinskas A (2013) A portable gas chromatograph with simultaneous detection by mass spectrometry and electroantennography for the highly sensitive *in situ* measurement of volatiles. *Anal Bioanal Chem* 405:7457–7467
- Schroth P, Schöning MJ, Kordoš P, Schütz S, Weißbecker B, Hummel HE (1999) Insect-based BioFets with improved signal characteristics. *Biosens Bioelectron* 14:303–308
- Scofield JH (1994) A frequency-domain description of a lock-in amplifier. *Am J Phys* 62:129–133
- Seeley JV, Micyus NJ, Bandurski SV, Seeley SK, McCurry JD (2007) Microfluidic deans switch for comprehensive two-dimensional gas chromatography. *Anal Chem* 79:1840–1847
- Sharif KM, Chin S-T, Kulsing C, Marriott PJ (2016) The microfluidic deans switch: 50 years of progress, innovation and application. *TrAC Trends Anal Chem* 82:35–54. <https://doi.org/10.1016/j.trac.2016.05.005>
- Sillam-Dussès D et al (2009) Identification by GC-EAD of the two-component trail-following pheromone of *Prorhinotermes simplex* (Isoptera, Rhinotermitidae, Prorhinotermitinae). *J Insect Physiol* 55:751–757
- Slone DH, Sullivan BT (2007) An automated approach to detecting signals in electroantennogram data. *J Chem Ecol* 33:1748–1762
- Trees HLV (2001) Detection, estimation, and modulation theory, part I: detection, estimation and linear modulation theory. John Wiley & Sons, Inc., New York
- Weissbecker B, Holighaus G, Schütz S (2004) Gas chromatography with mass spectrometric and electroantennographic detection: analysis of wood odorants by direct coupling of insect olfaction and mass spectrometry. *J Chromatogr A* 1056:209–216
- Ziemer RE, Tranter WH (1995) Principles of communications systems, modulation, and noise, 4th edn. John Wiley & Sons, Inc., New York