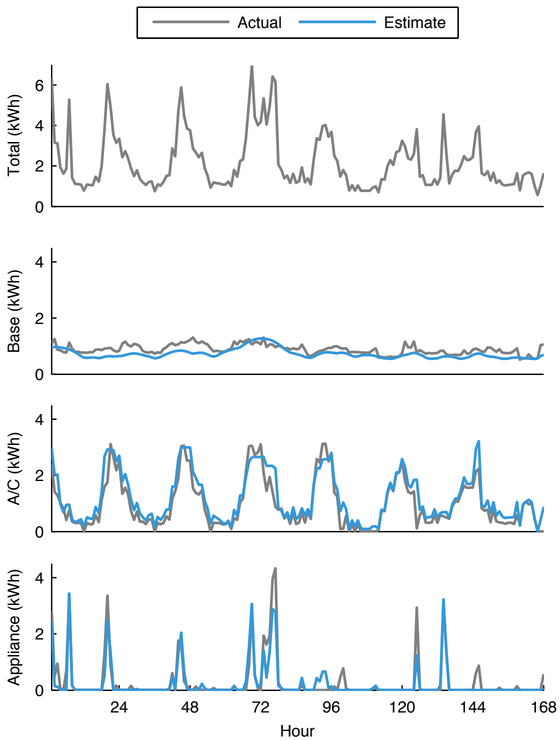

Figure 2: Energy disaggregation results over one week and a single home from the Pecan Street dataset.

Category	Mean	NNSC	Contextual
Base	0.2534	0.2793	0.1849
A/C	0.2849	0.2894	0.1919
Appliance	0.2262	0.2416	0.1900
Average	0.2548	0.2701	**0.1889**

Table 2: Comparison of performance on Pecan Street dataset, measured in mean absolute error (MAE).

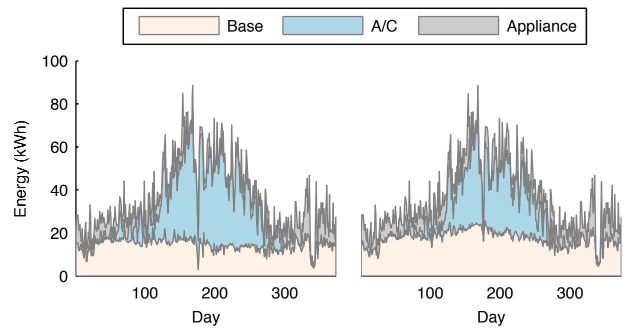

Figure 3: Energy disaggregation results over entire time period for a single home from the Pecan Street dataset with estimated (left) and actual (right).

noise. As can be seen in Table 1, by using ℓ_1 loss for y_2 and adding $g_i(y_i)$ terms penalizing $\|Dy_1\|_2^2$ and $\|Dy_2\|_1$, error decreases by 37% over just ℓ_2 loss alone; in Figure 1, we observe that our estimation recovers the true source signals closely with the g_i terms helping to capture the dynamics of the noise model for w_2.

As a baseline for this result, we compare to the mean prediction heuristic (predicting at each time point a breakdown proportional to the overall probability of each signal) and to a state-of-the-art unsupervised method, nonnegative sparse coding (Hoyer 2002). We apply sparse coding by segmenting the input signal into 1000 examples of 50 time points (1/4 the period of the sine wave, $X_1(t)$) and fit a sparse model of 200 basis functions. We report the best possible source separation by assigning each basis function according to an oracle measuring correlation with the true source signal and using the best value over a grid of hyperparameters. As can be seen in Table 1, the mean prediction heuristic is nearly 5 times worse and sparse coding is nearly 4 times worse than our best contextually supervised model.

Energy disaggregation with ground truth. Next we consider the ability of contextual supervision to recover the sources of energy consumption on a real dataset from Pecan Street consisting of 84 homes each with at least 1 year worth of energy usage data. As contextual information we construct a temperature time series using data from Weather Underground (http://www.wunderground.com/) measuring the temperature at the nearby airport in Austin, Texas. The Pecan Street dataset includes fine-grained energy usage information at the minute level for the entire home with an energy breakdown labeled according to each electrical circuit in the home. We group the circuits into categories representing air conditioning, large appliances and base load and aggregate the data on an hourly basis to mimic the scenario presented by smart meter data.

The specification of our energy disaggregation model is given in Table 3—we capture the non-linear dependence on temperature with radial-basis functions (RBFs), include a "Base" category which models energy used as a function of time of day, and featureless "Appliance" category representing large spikes of energy which do not correspond to any available context. For simplicity, we penalize each category's deviations from the model using ℓ_1 loss; but for heating and cooling we first multiply by a smoothing matrix S_n (1's on the diagonal and n super diagonals) capturing the thermal mass inherent in heating and cooling: we expect energy usage to correlate with temperature over a window of time, not immediately. We use $g_i(y_i)$ and the difference operator to encode our intuition of how energy consumption in each category evolves over time. The "Base" category represents an aggregation of many sources which we expect to evolve smoothly over time, while the on/off behavior in other categories is best represented by the ℓ_1 penalty. Finally we note that in the Pecan Street data, there is no labeled circuit corresponding exclusively to electric heating ("Heating"), and thus we exclude this category for this dataset.

In Table 2, we compare the results of contextual supervision with the mean prediction heuristic and see that contextual supervision improves by 26% over this baseline which is already better than nonnegative sparse coding. Qualitatively we consider the disaggregated energy results for a sin-

Category	Features	ℓ_i	g_i
Base	Hour of day	$\alpha_1\|y_1 - X_1\theta_1\|_1$	$\beta_1\|Dy_1\|_2^2$
Heating	RBFs over temperatures $< 50°$F	$\alpha_2\|S_2(y_3 - X_3\theta_3)\|_1$	$\beta_2\|Dy_3\|_1$
A/C	RBFs over temperatures $> 70°$F	$\alpha_3\|S_2(y_2 - X_2\theta_2)\|_1$	$\beta_3\|Dy_2\|_1$
Appliance	None	$\alpha_4\|y_4\|_1$	$\beta_4\|Dy_4\|_1$

Table 3: Model specification for contextually supervised energy disaggregation.

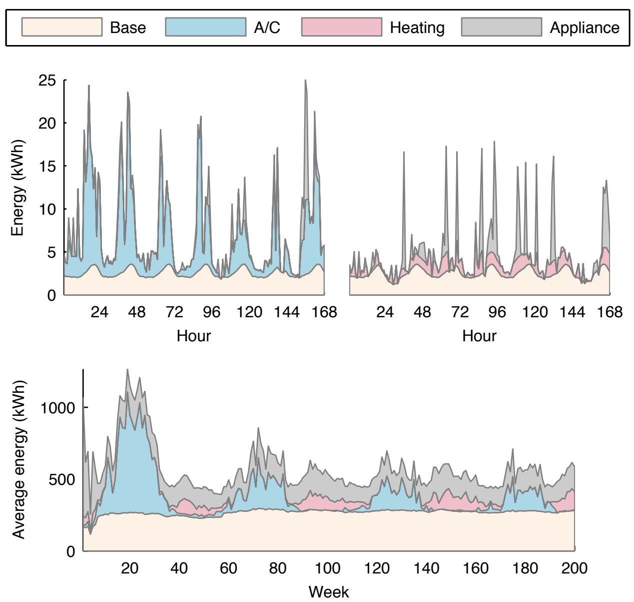

Figure 4: Disaggregated energy usage for a single home near Fresno, California over a summer week (top left) and a winter week (top right); aggregated over 4000+ homes over nearly four years (bottom)

gle home over one week in Figure 2 and see that contextual supervision correctly assigns energy usage to categories—a large amount of energy is assigned to A/C which cycles on and off roughly corresponding to external temperature, large spikes from appliances happen at seemingly random times and the smoothly varying base load is captured correctly. In Figure 3, we consider the disaggregation results for the same home across the entire time period and see that the contextually supervised estimates correspond very closely to the actual sources of energy consumption.

Large-scale energy disaggregation. Next, we turn to the motivating problem for our model: disaggregating large-scale low-resolution smart meter data into its component sources of consumption. Our dataset consists of over 4000 homes and was collected by PG&E from customers in Northern California who had smart meters between 1/2/2008 and 12/31/2011. According to estimations based on survey data, heating and cooling (air conditioning and refrigerators) comprise over 39% of total consumer electricity usage (U.S. Energy Information Administration 2009) and thus are dominant uses for consumers. Clearly, we expect temperature to have a strong correlation with these uses and thus we provide contextual supervision in the form of temperature information. The PG&E data is anonymized, but the location of individual customers is identified at the census block level and we use this information to construct a parallel temperature dataset as in the previous example.

We present the result of our model at two time scales, starting with Figure 4 (top), where we show disaggregated energy usage for a single home over a typical summer and winter week. Here we see that in summer, the dominant source of energy consumption is estimated to be air conditioning due to the context provided by high temperature. In winter, this disappears and is replaced to a smaller extent by heating. In Figure 4 (bottom), itemized energy consumption aggregated across all 4000+ homes demonstrates these basic trends in energy usage. Quantitatively, our model assigns 15.6% of energy consumption to air conditioning and 7.7% to heating, reasonably close to estimations based on survey data (U.S. Energy Information Administration 2009) (10.4% for air conditioning and 5.4% for space heating). We speculate that our higher estimation may be due to the model conflating other temperature-related energy usages (e.g. refrigerators and water heating) or to differences in populations between the survey and smart meter customers.

Conclusion and discussion

The disaggregation of smart meter data into itemized energy uses creates large opportunities for increases in efficiency; as smart meters are already widely deployed and have been collecting data for the past several years, millions of homes stand to benefit. However, disaggregating smart meter data is a challenging task due to its low-resolution sampling and lack of supervised information. We believe that with the development of contextual supervision described in this paper, we have made a significant advancement in this area that has been previously dominated by methods that rely on either high-resolution or supervised data that, unlike the smart meter data, is not readily available.

An interesting direction for future work is the explicit connection of our large-scale low-resolution methods with the more sophisticated appliance models developed on smaller supervised datasets with high-frequency measurements. However, there are clear limitations as to what can be observed in a whole home power trace that is only sampled once an hour. The development of refined statistical models that produce confidence intervals around their estimations is one avenue for dealing with this uncertainty. Still, the largest gains are likely to come from an increase in sampling frequency, perhaps in a hybrid approach that varies the sampling rate in order to capture more accurate high-frequency snapshots during periods of high activity.

References

Bell, A. J., and Sejnowski, T. J. 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural computation* 7(6):1129–1159.

Blumensath, T., and Davies, M. 2005. Shift-invariant sparse coding for single channel blind source separation. *SPARS* 5:75–78.

Comon, P. 1994. Independent component analysis, a new concept? *Signal processing* 36(3):287–314.

Darby, S. 2006. The effectiveness of feedback on energy consumption. Technical report, Environmental Change Institute, University of Oxford.

Davies, M., and James, C. 2007. Source separation using single channel ica. *Signal Processing* 87(8):1819–1832.

Ehrhardt-Martinez, K.; Donnelly, K. A.; and Laitner, S. 2010. Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities. In *American Council for an Energy-Efficient Economy*.

Ghahramani, Z., and Jordan, M. I. 1997. Factorial hidden markov models. *Machine learning* 29(2-3):245–273.

Hart, G. W. 1992. Nonintrusive appliance load monitoring. *Proceedings of the IEEE* 80(12):1870–1891.

Hoyer, P. O. 2002. Non-negative sparse coding. In *Neural Networks for Signal Processing, 2002. Proceedings of the 2002 12th IEEE Workshop on*, 557–565. IEEE.

Institute for Electric Efficiency. 2012. Utility-scale smart meter deployments, plans, and proposals. Technical report, Institute for Electric Efficiency.

Kim, H.; Marwah, M.; Arlitt, M. F.; Lyon, G.; and Han, J. 2011. Unsupervised disaggregation of low frequency power measurements. In *SDM*, volume 11, 747–758. SIAM.

Kolter, J. Z., and Jaakkola, T. 2012. Approximate inference in additive factorial hmms with application to energy disaggregation. In *International Conference on Artificial Intelligence and Statistics*, 1472–1482.

Kolter, J. Z.; Batra, S.; and Ng, A. Y. 2010. Energy disaggregation via discriminative sparse coding. In *Neural Information Processing Systems*, 1153–1161.

Lewicki, M. S., and Sejnowski, T. J. 2000. Learning overcomplete representations. *Neural computation* 12(2):337–365.

Neenan, B., and Robinson, J. 2009. Residential electricity use feedback: A research synthesis and economic framework. Technical report, Electric Power Research Institute.

Olshausen, B. A., and Field, D. J. 1997. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision research* 37(23):3311–3326.

Parson, O.; Ghosh, S.; Weal, M.; and Rogers, A. 2012. Nonintrusive load monitoring using prior models of general appliance types. In *26th AAAI Conference on Artificial Intelligence*.

Roweis, S. T. 2001. One microphone source separation. *Advances in neural information processing systems* 793–799.

Schmidt, M. N., and Mørup, M. 2006. Nonnegative matrix factor 2-d deconvolution for blind single channel source separation. In *Independent Component Analysis and Blind Signal Separation*. Springer. 700–707.

Schmidt, M., and Olsson, R. 2006. Single-channel speech separation using sparse non-negative matrix factorization.

Smaragdis, P.; Raj, B.; and Shashanka, M. 2006. A probabilistic latent variable model for acoustic modeling. *Advances in models for acoustic processing, NIPS* 148.

U.S. Energy Information Administration. 2009. 2009 RECS survey data. Available at http://www.eia.gov/consumption/residential/data/2009/.

Ziefman, M., and Roth, K. 2011. Nonintrusive appliance load monitoring: Review and outlook. *IEEE Transactions on Consumer Electronics* 57(1):76–84.