There is increasing interest in the potential advantages of using quantum computing technologies as sampling engines to speedup machine learning and probabilistic programming tasks. However, some pressing challenges in state-of-the-art quantum annealers have to be overcome before we can assess their actual performance. Most notably, the effective temperature at which samples are generated is instance-dependent and unknown, the interaction graph is sparse, the parameters are noisy, and the dynamic range of the parameters is finite. Of all these limitations, the sparse connectivity resulting from the local interaction between quantum bits in physical hardware implementations, is considered the most severe limitation to the quality of constructing powerful machine learning models. Here we show how to overcome or mitigate these limitations and illustrate our findings by training probabilistic generative models with arbitrary pairwise connectivity. Our model can be trained in quantum hardware without full knowledge of the effective parameters specifying the corresponding quantum Boltzmann-like distribution. Therefore, inference of the effective temperature is avoided and the effect of noise in the parameters is mitigated. We illustrate our findings by successfully training hardware-embedded models with all-to-all connectivity on a real dataset of handwritten digits and two synthetic datasets. In each of these datasets we show the generative capabilities of the models learned with the assistance of the quantum annealer in experiments with up to 940 quantum bits. Additionally, we show a visual Turing test with handwritten digit data, where the machine generating the digits is a quantum processor. Such digits, with a remarkable similarity to those generated by humans, are extracted from the experiments with 940 quantum bits.