

---

# MAGAN: Aligning Biological Manifolds

---

Matthew Amodio<sup>1</sup> Smita Krishnaswamy<sup>2 1</sup>

## Abstract

It is increasingly common in many types of natural and physical systems (especially biological systems) to have different types of measurements performed on the same underlying system. In such settings, it is important to align the manifolds arising from each measurement in order to integrate such data and gain an improved picture of the system; we tackle this problem using generative adversarial networks (GANs). Recent attempts to use GANs to find correspondences between sets of samples do not explicitly perform proper alignment of manifolds. We present the new Manifold Aligning GAN (MAGAN) that aligns two manifolds such that related points in each measurement space are aligned. We demonstrate applications of MAGAN in single-cell biology in integrating two different measurement types together: cells from the same tissue are measured with both genomic (single-cell RNA-sequencing) and proteomic (mass cytometry) technologies. We show that MAGAN successfully aligns manifolds such that known correlations between measured markers are improved compared to other recently proposed models.

## 1. Introduction

We commonly have samples from a pair of related domains and want to ask the natural question of how samples from one relate to samples from the other. Our motivational system for this is two types of measurements on cells sampled from the same population in a biological system. It is important for the discovery of new biology to integrate these datasets, which are often generated at great cost and expense. However, a fundamental challenge is that there are exponentially many possible relationships that could exist

<sup>1</sup>Department of Computer Science, Yale University

<sup>2</sup>Department of Genetics, Yale University. Correspondence to: Matthew Amodio <matthew.amodio@yale.edu>.

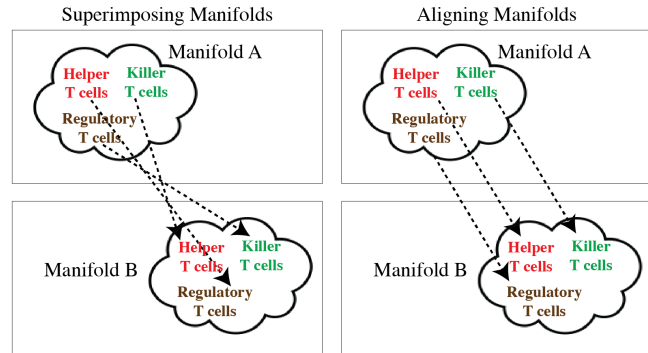


Figure 1. There are exponentially many mappings that superimpose the two manifolds, fooling a GAN’s discriminator. By aligning the manifolds, we maintain pointwise correspondences.

between the two domains of measurement and the system must learn a logical way to map between them.

One way to approach this task to use dual GANs mapping between each domain. The first of these approaches required supervised paired examples from each domain, an impractical demand for many applications (Isola et al., 2016). Recently, there have been attempts at performing the same task without the supervision of paired data (Zhu et al., 2017; Yi et al., 2017; Kim et al., 2017; Li et al., 2017). Like these previous models, MAGAN learns to map between distinct domains from unsupervised, unpaired data without pretraining. However, unlike them, MAGAN aligns rather than superimposes the manifolds of the two domains.

We draw the connection between unsupervised domain mapping with dual GANs and the alignment of the domain manifolds (Domingos, 2012). Much work has framed the generation problem of GANs as sampling points from this manifold (Park et al., 2017; Zhu et al., 2016). In the dual GAN framework, each domain’s GAN learns a mapping that fools a discriminator by generating a point on the other domain’s manifold when given a point on its manifold. There are exponentially many mappings that produce the same distribution of outputs from the same distribution of inputs but do so by mapping different individual input points to different individual output points. To the GAN’s discriminator, these mappings are identical. But when we interpret the

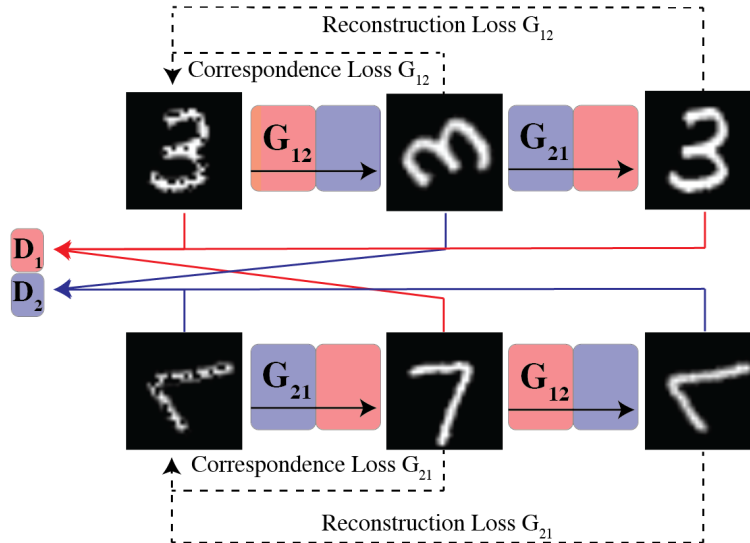


Figure 2. MAGAN’s architecture with two generators, two discriminators, reconstruction loss, and correspondence loss. Domain 1 comprises upright images of 3’s and 7’s, Domain 2 comprises rotated images of 3’s and 7’s.

generator as finding correspondences between domains, we instead have preferences amongst them.

We motivate our preference for some mappings between domain manifolds by differentiating between *aligning* manifolds and *superimposing* manifolds. A mapping that superimposes two manifolds would make the two indistinguishable by making their supports and densities identical. However, in our setting, we have produced each domain by measuring cells from the same underlying tissue twice. The two manifolds could be superimposed without aligning the two observations of each latent cell. To consider the two manifolds aligned rather than just superimposed, we require that a cell’s representation in one manifold be aligned with that same cell’s representation on the other manifold.

In this paper we propose the novel concept of using adversarial neural networks for alignment of manifolds arising from different biological experimental data measurement types. Single-cell biological experiments create many situations where manifold alignment problems are of interest. New technologies allow for measurements to be made at the granularity of each cell, rather than older technologies which could only acquire aggregate summary statistics for whole populations of cells. While these instruments allow us to discover biological phenomena that were not apparent before, it is a challenge to integrate and analyze this information in a unified fashion for biological discovery. Further, even for the same technology, experiments run on different days or in different batches can show variations even on the same populations, possibly due to calibration differences. In such cases even replicate experiments need alignment

before comparison. Two such technologies that we examine are single-cell RNA sequencing which measures cells in thousands of gene (mRNA) dimensions and mass cytometry which measures protein abundances in several dozen dimensions (Bendall et al., 2012; Klein et al., 2015).

In all of these examples we have two data manifolds with a latent physical cell being measured analogously in each manifold. In some applications it might be adequate to simply superimpose these manifolds in any way. In many applications though, including the ones demonstrated here, we would like to be able to align them such that the two representations of each latent cell are aligned. MAGAN improves upon neural models for manifold alignment by finding the mapping between the manifolds (*correspondence*) that models these latent points by penalizing differences in each point’s representation in the two manifolds.

We summarize the contributions of this paper as follows:

1. The introduction of a novel GAN architecture that aligns rather than superimposes manifolds to find relationships between points in two distinct domains
2. The demonstration of novel applications made possible by the new architecture in the analysis of single-cell biological data

The rest of this paper is organized as follows. First, there is a detailed description of the MAGAN architecture. Next, there is a validation of its performance on artificial data and the standard MNIST dataset. Then, there are demonstrations

on three real-world biological applications: mapping between two replicate cytometry domains, mapping between two different cytometry domains, and mapping between one cytometry domain and a single-cell RNA sequencing domain.

## 2. Model

### 2.1. Architecture

MAGAN (Figure 2) is composed of two GANs, each with a generator network  $G$  that takes as input  $X$  and outputs a target dataset  $X'$ . We refer to each generator as a *mapping* from the input domain to the output domain. Each generator attempts to make its output  $G(X)$  indistinguishable by  $D$  from  $X'$ . Denote the two datasets  $X_1$  and  $X_2$ . Let the generator mapping from  $X_1$  to  $X_2$  be  $G_{12}$  and the generator mapping from  $X_2$  to  $X_1$  be  $G_{21}$ . The discriminator that tries to separate true points from mapped ones for the first domain is  $D_1$  and the discriminator doing so for the second domain is  $D_2$ .

The loss for  $G_1$  on minibatches  $x_1$  and  $x_2$  is:

$$\begin{aligned} x_{12} &= \mathbf{G}_{12}(x_1) \\ x_{21} &= \mathbf{G}_{21}(x_2) \\ L_r &= L_{reconstruction} = L(x_1, x_{12}) \\ L_d &= L_{discriminator} = -\mathbb{E}_{x_1 \sim P_{X_1}} [\log \mathbf{D}_2(x_{12})] \\ L_c &= L_{correspondence} = L(x_1, x_{12}) \\ \mathbf{L}_{G_1} &= L_r + L_d + L_c \end{aligned}$$

where  $L$  is any loss function, here mean-squared error (MSE).

Similarly, the loss for  $G_2$  is:

$$\begin{aligned} x_{21} &= \mathbf{G}_{21}(x_2) \\ x_{12} &= \mathbf{G}_{12}(x_1) \\ L_r &= L(x_2, x_{21}) \\ L_d &= -\mathbb{E}_{x_2 \sim P_{X_2}} [\log \mathbf{D}_1(x_{21})] \\ L_c &= L(x_2, x_{21}) \\ \mathbf{L}_{G_2} &= L_r + L_d + L_c \end{aligned}$$

The losses for  $D_1$  and  $D_2$  are:

$$\begin{aligned} \mathbf{L}_{D_1} &= -\mathbb{E}_{x_1 \sim P_{X_1}} [\log \mathbf{D}_1(x_1) + \log \mathbf{D}_1(x_{12})] \\ &\quad - \mathbb{E}_{x_2 \sim P_{X_2}} [\log(1 - \mathbf{D}_1(x_{21}))] \\ \mathbf{L}_{D_2} &= -\mathbb{E}_{x_2 \sim P_{X_2}} [\log \mathbf{D}_2(x_2) + \log \mathbf{D}_2(x_{212})] \\ &\quad - \mathbb{E}_{x_1 \sim P_{X_1}} [\log(1 - \mathbf{D}_2(x_{12}))] \end{aligned}$$

### 2.2. Correspondence Loss

Previous models included only two restrictions: (1) that the two generators be able to reconstruct a point after it moves

to the other domain and back, and (2) that the discriminators not be able to distinguish batches of true and mapped points. To do this, the generators could learn arbitrarily complex mappings as long as they superimpose the two manifolds.

To instead enforce the manifolds be fully aligned, MAGAN includes a correspondence loss between a point in its original domain and that point's representation after being mapped to the other domain. This correspondence loss needs to be chosen appropriately for the manifolds in any particular problem. We propose two such formulations: one unsupervised and one supervised.

#### 2.2.1. UNSUPERVISED CORRESPONDENCE

In the biological domains considered here, we measure the same physical system in two different experiments where a subset of the dimensions in each experiment are shared. For example, in Section 3.4, we measure the amount of 35 proteins in a physical tissue in the first experiment and then the amount of 31 proteins in the same physical tissue in a second experiment. Sixteen of the proteins (CD4 for example) are measured in both experiments. Thus, the pre-mapping amount of CD4 in the first domain should equal the post-mapping amount of CD4 in the second domain. This leverages the information from the domains partially overlapping while generating a point in the full space (i.e. the generator maps a first domain point to the full 31-dimensional second domain by preserving the values of the 16 shared dimensions and then filling in the values of the 15 unique dimensions to make a plausible second domain point).

Formally, for each shared dimension pair  $(i, j)$ , the correspondence loss is:

$$L_c = MSE(\mathbf{G}_{12}(x_1)_j, (x_1)_i) + MSE(\mathbf{G}_{21}(x_2)_i, (x_2)_j)$$

We note that there are many types of relationships between a dimension in the first domain and a dimension in the second domain that could define the unsupervised correspondence loss in other experimental settings. For example, rather than knowing that two dimensions of the experiment are identical, we might know that two markers are negatively correlated. A cell in one domain high in a T cell identifier like CD8 should not be mapped to a cell in the other domain high in a B cell identifier like CD19. The unsupervised correspondence loss can be formulated to enforce this by penalizing deviations from this known relationship.

#### 2.2.2. SEMI-SUPERVISED CORRESPONDENCE

In the case where no known relationship between the domains is known, the correspondence loss could alternatively be formulated as a semi-supervised learning setting. Of course, if each point in  $X_1$  already had a known correspon-

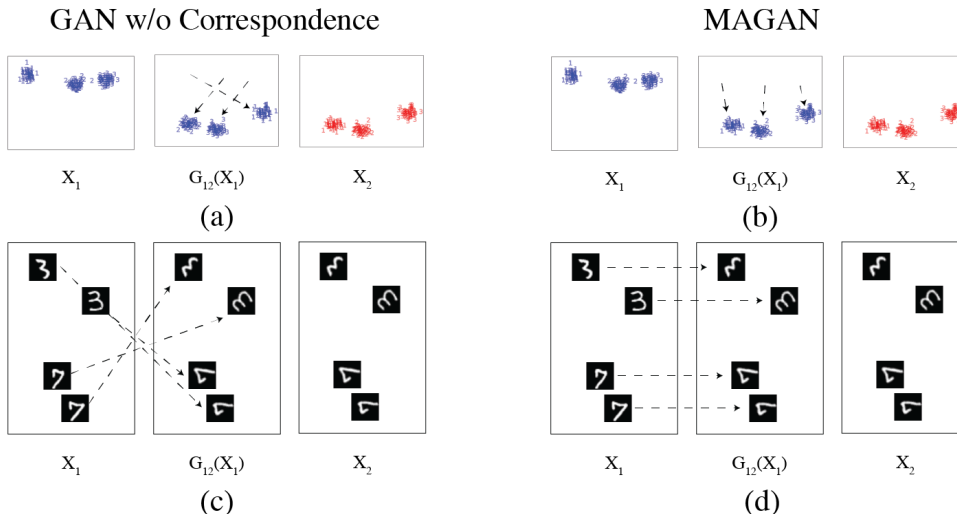


Figure 3. Both models superimpose the manifolds, meaning the first domain ( $X_1$ ) is mapped to the second domain ( $X_2$ ) such that the dataset of the first domain after mapping ( $G_{12}(X_1)$ ) matches the second domain. Without the correspondence loss, though, this mapping is arbitrary and thus the relationships found vary. With the correspondence loss, the relationships found are coherent. This is confirmed with (a) a GAN without correspondence loss on artificial data (b) MAGAN on artificial data (c) a GAN without correspondence loss on MNIST and (d) MAGAN on MNIST.

dence with a point in  $X_2$ , no framework of dual GANs would be necessary to discover relationships. In some domains, though, it is easy to acquire a very small number of labeled pairs. We would like a model that learns from unsupervised data but can improve with any small number of labels that can be acquired. In those situations, we want to leverage both (1) the information that the unsupervised model has learned on all of the data and (2) incorporate the information the labels provide where they exist.

Here we choose the loss function to be nonzero only at the paired points in each domain. Its value is then the sum of the losses on each labeled pair, where the loss for a particular labeled pair  $(x_{1i}, x_{2j})$ ,  $x_{1i} \in X_1, x_{2j} \in X_2$  is:

$$L_c = MSE(\mathbf{G}_{12}(x_{1i}), x_{2j}) + MSE(\mathbf{G}_{21}(x_{2j}), x_{1i})$$

### 2.3. Manifold Data Augmentation

MAGAN also utilizes a novel technique for data augmentation, leveraging the imperfect reconstructions each generator produces within its domain. It has been well established that autoencoders model and reconstruct from the data manifold (Hinton et al., 1997; Vincent et al., 2008). We note that the dual GANs within each domain function as an autoencoder, meaning their reconstruction  $x'_i$  of a sample  $x_i$  is another point near the underlying manifold, but importantly  $x'_i \neq x_i$ . By letting each discriminator see the reconstructions as true samples from the real domain, we both (1) augment the original data with new samples from the manifold and (2) prevent the discriminators from learning to separate real

from generated examples by modeling the noise around the manifold, which differs between  $X_1$  and  $G_{21}(X_2)$  and between  $X_2$  and  $G_{12}(X_1)$ . This is especially important in biological settings, where the number of measurements per cell dwarfs the number of cells measured and dropout in the measuring process produces sparsity.

## 3. Experiments

All experiments were performed with the MAGAN framework with discriminators of five layers each and generators of three layers each. Layer sizes depended on the dataset, while Leaky ReLU activations were used on all layers except the output layers of the discriminators (which were sigmoid) and the generators (which were linear). Dropout of 0.9 was applied during training and for images convolutional layers were used. Optimization was performed on 100,000 iterations of batches of size 256 by the ADAM optimizer with learning rate 0.001.

As with other GANs, the generators and discriminators are trained alternatively, so they each must get progressively better as their adversaries make their tasks harder and harder. One known difficulty in the adversarial training process is preventing a collapse of the generator into mapping all inputs to one point, chasing the minimum probability region of the discriminator as it moves. To combat this, MAGAN includes the approach outlined in (Salimans et al., 2016). This involves giving the discriminator access to minibatch information by having a subset of the network process a

rotation of the original data matrix.

### 3.1. Artificial Data

We first test MAGAN on a generated example of points sampled from Gaussian distributions with varying means. Figure 3a shows the three subpopulations in the first domain  $X_1$  in blue and the three in the second domain  $X_2$  in red with an example mapping where, without the correspondence loss, each subpopulation in  $X_1$  is mapped to a subpopulation in  $X_2$ , but not to the closest one. Even though the distribution of  $G_{12}(X_1)$  matches the distribution of  $X_2$ , for an individual point  $x_{1i} \in X_1$ ,  $G_{12}(x_{1i})$  is not the member of  $X_2$  that is most closely analogous to it. MAGAN finds a mapping that fools the discriminator, too: the one that least alters the original input (Figure 3b).

Without the correspondence loss, not only is a less-preferred manifold superimposition chosen, but the one chosen varies from run to run of the model. We compare the variability of the learned mappings across multiple runs of each model with 100 independent trials. In each trial we evaluate the relationships by calculating  $G_{12}(x_{1i})$  for each  $x_{1i} \in X_1$  and calculating its nearest neighbor  $x_{2j}$  in the real  $X_2$ . Then, this is repeated for the other domain. Figure 4a confirms that for the GAN without the correspondence loss, the learned manifold superimposition (and thus the correspondences) varies with repeated training the model. Figure 4b confirms MAGAN instead aligns the manifolds and finds the same correspondence every time.

### 3.2. MNIST

Next we test a subset of the MNIST handwritten digit data by taking only 3's and 7's as the first domain  $X_1$ , and a 120 degree rotation of each image as the second domain  $X_2$ . Without the correspondence loss (Figure 3c), each subpopulation in  $X_1$  maps to one of the subpopulations in  $X_2$ , but the original 3's go to the rotated 7's and vice versa. There is no term in the objective function to create a preference for the mapping that sends original 3's to rotated 3's. It would be difficult to define a distance measure that captures the notion of alignment with these manifolds, but it is a natural place where a small number of labeled pairs could be easily acquired. The semi-supervised correspondence loss with just a single labeled pair of points finds the desired manifold alignment and gets the correct correspondences for all of the other points that are unlabeled (Figure 3d).

Using the same simulation design as in the previous section, we can test the robustness of the models in finding these particular mappings. The GAN without the correspondence loss discovers either relationship with roughly even probability (Figure 4c). Remarkably, MAGAN is able to use the single labeled example to learn that (except for a few sloppily written 3's that in fact look more like 7's) the original

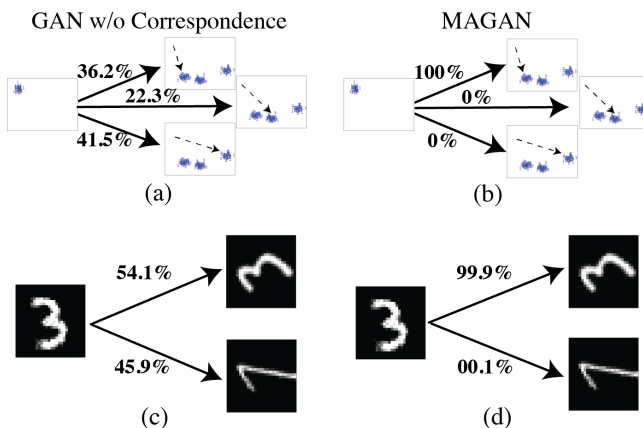


Figure 4. In simulations of 100 complete training runs of each model, without correspondence loss the resulting relationships learned varied randomly in both the (a) toy and (c) MNIST datasets. With correspondence loss, the most coherent relationship was found repeatedly for both (b) toy and (d) MNIST datasets.

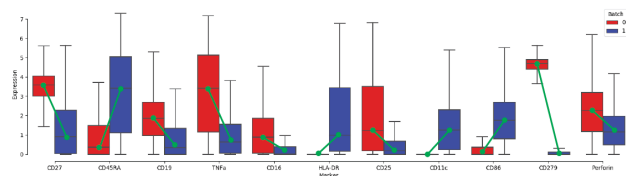


Figure 5. Selected markers illustrating large batch effects that separate the two data manifolds.

3's correspond to the rotated 3's and that the original 7's correspond to the rotated 7's every time (Figure 4d).

### 3.3. Correspondence: CyTOF Replicates

We now test MAGAN on real biological data from single-cell time-of-flight mass cytometry (CyTOF) measurements of protein abundance. Each protein, also referred to as a *marker*, is measured individually for each cell, allowing for more granular analysis than processes that only measure population totals for the cells in a given sample. Here the same sample was run twice in different batches (*replicates*), but due to machine calibration and other experimental details that are impossible to reproduce precisely each time, there are distortions between the batches. Thus, even though the same physical blood sample is being measured, the data manifold of each batch is different. The type of noise introduced by these distortions is not known *a priori*, need not fit any parametric assumption, and is likely to be highly nonlinear.

To analyze these two batches together, we need to know

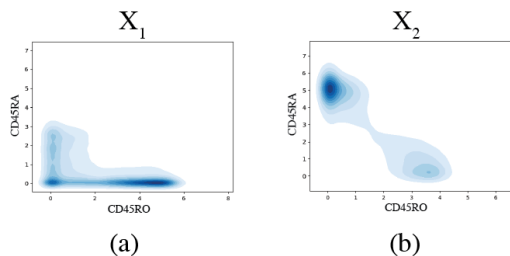


Figure 6. Two distinct populations of T-cells (CD45RA+CD45RO- and CD45RA-CD45RO+) with severe dropout in the CD45RA marker that causes a difference between that between the (a) first batch and (b) second batch.

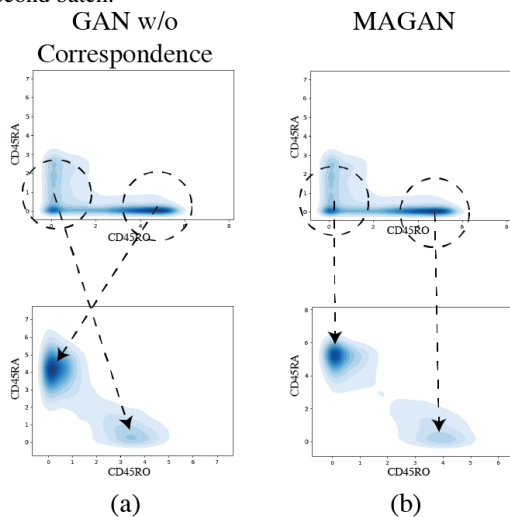


Figure 7. (a) Without correspondence loss, the GAN corrects the batch effect but subpopulations are reversed. (b) MAGAN still corrects the batch effect and subpopulations are preserved.

which cells in the first batch correspond to which cells in the second batch. To do this, we learn a mapping with MAGAN between the batches, each of which contains 75,000 cells with 34 individual markers measured. Figure 5 shows that the two batches indeed contain distinct differences in both the values of each marker and their distribution. For example, the mean value of HLA-DR in the second batch is higher than the *maximum* value in the first batch.

We demonstrate that MAGAN with its correspondence loss preserves crucial information that is lost with the mapping from the GAN without the correspondence loss. Often, analysis starts by identifying subpopulations of interest. For example, naive T-cells and central memory T-cells serve distinct functions and can be identified by looking at two isoforms of the CD45 marker, CD45RA and CD45RO

(Capra et al., 1999). In naive T-cells, CD45RA is present while CD45RO is not (CD45RA+CD45RO-), and in central memory T-cells CD45RA is not present while CD45RO is (CD45RA-CD45RO+). Figure 6a shows that very few cells had any CD45RA readings in the first batch, a typical case of instrument-induced dropout. Figure 6b shows proper readings for CD45RA in the second batch, where the two distinct subpopulations are clearly seen.

Both models learn a mapping for the first batch of cells  $x_1$  such that  $G_{12}(x_1)$  fools their discriminators by looking like the second batch of cells  $x_2$ . However, in the GAN without the correspondence loss (Figure 7a), naive T-cells in the first batch are mapped to central memory T-cells in the second batch and vice versa. If we went through the manual process of gating (selecting cells by manually looking at relative marker expression) central memory T-cells in the first batch and wanted to know whether their expression was similar in the second batch, we would be led to believe incorrectly that either there are none of these cells in the second batch or their expression profile is radically different.

MAGAN learns a different mapping (Figure 6b), the one in which subpopulation correspondences are preserved. Notably, the resulting mapped dataset  $G_{12}(x_1)$  is not negatively affected by the correspondence loss. Instead, out of the two mappings that have similar results at the aggregate level, the one that maintains pointwise correspondences is learned. With the cell correspondences from other manifold superimpositions, the wrong biological conclusions could be made. This application necessitates MAGAN’s manifold alignment.

### 3.4. Correspondence: Different CyTOF Panels

Next we demonstrate MAGAN’s ability to align two manifolds in domains whose dimensionality only partly overlap. Despite the other advantages of CyTOF instruments, one disadvantage is that CyTOF experiments can only measure the expression of 30-40 markers per cell. Each experiment chooses which 30-40 markers to measure and refers to this set as the *panel*. Even though each panel has a limited capacity, different panels can be run on different samples from the same physical blood or tissue. MAGAN provides the opportunity to combine the results from these multiple panels and effectively increase the number of expression measurements acquired for each cell.

To test this, we use the datasets from two experiments published in (Setty et al., 2016) where each experiment had a different panel that was run on samples from the same population of cells. The first panel measured 35 markers, the second panel measured 31 markers, and 16 of those were measured in both. Without any advanced methods, all we would be able to do across experiments is compare population summary statistics — and lose all of the information

at a single-cell resolution that motivated these experiments being done in the first place.

If we can identify points in each panel that measure the same cell, we can combine the measurements and have an augmented 50-dimensional dataset. To accomplish this, we take the first experiment’s panel as one domain and the second experiment’s panel as the other domain and use MAGAN to learn a mapping between the two. We then combine the original 35 dimensions of a cell in the first experiment  $x_{1i}$  with the 15 dimensions unique to the second experiment from that cell after mapping  $G_{12}(x_{1i})$ .

For combining the measurements from each experiment to be meaningful, the mapped point  $G_{12}(x_{1i})$  must correspond accurately to the true point  $x_{2i}$ . This notion can be captured by taking the correspondence loss function to be the MSE across the 16 dimensions that are shared between the experiments. In other words, MAGAN should use the shared measurements to match cells between experiments, and then learn the required mapping for all of the measurements that are not shared. Without incorporating this correspondence measure into the model,  $x_{1i}$  need not be analogous to  $G_{12}(x_{1i})$  in any way, and their information could not be combined.

We evaluate the accuracy of each model’s learned correspondence by removing one of the markers measured in both experiments, CD3, from the first experiment. Then, we map points from the first experiment to the second experiment and evaluate how well the discovered CD3 values correspond with the true, held-out CD3 values for each cell from the first experiment.

Figure 8a shows that the GAN without the correspondence loss finds a manifold superimposition that does not preserve the values of CD3 for each cell accurately. Quantitatively, we can evaluate this with the correlation coefficient between the real, held-out CD3 values and the CD3 values predicted after mapping each point to the other domain. For the GAN without the correspondence loss (Figure 8a), the correlation is  $-.275$ , while for MAGAN (Figure 8b) it is  $.801$ . The negative correlation means that without the correspondence loss, the GAN will systematically map cells in one panel to different cells in the other panel.

We perform cross-validation by repeating this test with each of the 16 shared markers in turn for the GAN without correspondence loss (Figure 8c) and MAGAN (Figure 8d). While some of the markers have more shared information than others and are recovered more accurately, in all cases the correlation is better with the MAGAN.

If we had not measured one of these in the first experiment, we would have been able to use the learned value from the mapping in its place with remarkable accuracy. MAGAN can powerfully increase the impact of CyTOF experiments

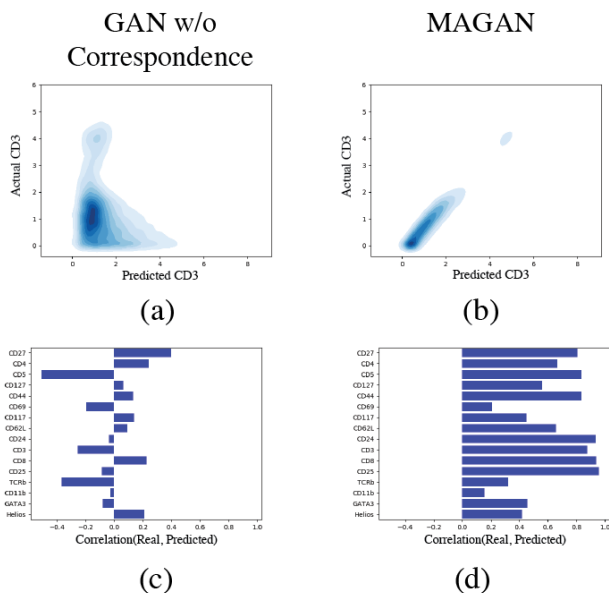


Figure 8. Using MAGAN’s correspondence loss, measurements from each experiment can be combined. Their true values are known because they are measured in both experiments. Performing cross-validation by holding each out from the first experiment, we can measure the correlation between the predicted value and the real, correct value.

by expanding their limited capacity of markers that can be measured at any one time.

### 3.5. Correspondence: Cytometry and scRNA-seq

To demonstrate MAGAN aligning manifolds of domains with radically different dimensionality and underlying structure, we use it to find correspondences between flow cytometry (FACS-sorted) and scRNA-seq measurements made on the same set of cells. These two types of measurements have advantages and disadvantages, including the throughput, quality, and amount of information acquired from each. Being able to combine their information offers the possibility of getting the best from each and finding insights that might not otherwise be obtainable. In order to do this, though, it is crucial for pointwise correspondences to be accurate, or else features of a data point in the scRNA-seq domain will be ascribed to the incorrect point in the cytometry domain and the relationships will be meaningless.

To test MAGAN in this setting, we use a dataset consisting of 2830 measurements, where the dimensionality of each domain is 12 and 12496 for cytometry and scRNA-seq, respectively (Velten et al., 2017). The scRNA-seq data was normalized with the inverse hyperbolic sine transform and preprocessed with MAGIC (van Dijk et al., 2017). Here we know the true correspondences of which points in the

Table 1. With MAGAN’s correspondence loss, the accuracy of the learned mapping is dramatically improved, as measured by the MSE between the known real point  $x$  and the predicted point  $G(x)$  after mapping.

Paired Cytometry & scRNA-seq	GAN	MAGAN	Modified LLE
$\text{MSE}(x_1, G_{21}(x_2))$	99.3	22.0	10.5
$\text{MSE}(x_2, G_{12}(x_1))$	33.7	7.1	2.4

two domains are the same cell. In this setting we use the semi-supervised correspondence loss and show the impact of providing the pairing of just 10 cells, which can easily be acquired via inspection.

We evaluate the quality of the correspondences learned by calculating the *correspondence error*, or MSE between the true known value  $x_{1i} \in X_1$  and the predicted correspondence  $G_{21}(x_{2i})$ . In this semi-supervised setting where we have some known labels, we compare MAGAN to both the GAN without correspondence loss and the method of (Ham et al., 2003) which uses a modification of locally linear embeddings for manifold alignment. We first find equal-sized manifolds for each domain separately. Then, for two observations we know to be corresponding, we constrain the point on each manifold to be identical. This aligns the two manifolds as anchored by the ground truth corresponding points. Then, in order to obtain correspondences in the original space rather than the manifold space, we interpolate between points in the original space based on a point’s nearest neighbors in the latent manifold space.

We note that not only is this method only defined in the semi-supervised setting, it also requires finding the eigenvectors of a square matrix of size  $n = 2830$ , so cannot scale to the dataset sizes that neural network approaches like MAGAN can. Table 1 shows the correspondence error for the correspondences mapping to and from each domain. The LLE approach achieves the best error, showing that for small datasets with some ground truth labels, this method is preferable. We see that the correspondence loss is necessary for the GAN framework to perform manifold alignment, as MAGAN comes much closer to the performance of the baseline while the GAN is off by an order of magnitude.

#### 4. Discussion

The use of GANs for manifold alignment is well motivated. By virtue of being parallelizable deep learning models trained with minibatch gradient descent, GANs can work with massive datasets. In contrast, other methods are graph-based and involve eigen decomposition of matrices that scale with the number of data points. Additionally, manifold

alignment methods themselves are not inherently generative. They align points in the latent space and need to be augmented with original space interpolation (meaningful interpolation may not be achievable in all domains), while GANs explicitly generate points in the original space.

Furthermore, GANs learn a general non-linear mapping between manifolds that is not limited to any specific assumptions such as a linear latent dimension of correspondence (Butler et al., 2018) or that correlation distance is preserved across manifolds (Haghverdi et al., 2018).

Outside of GANs, much previous work has been devoted to the field of manifold alignment. In (Ham et al., 2003; 2005), semi-supervised approaches are used, but the learned alignments are only defined on the training points, unlike MAGAN which learns a universal mapping guided by its semi-supervised loss. In (Wang & Mahadevan, 2008), Procrustes alignment is used to provide extension beyond the training points. This is still a two-step alignment process, though, which first finds a manifold for each domain and then separately aligns them. As such, information from the original space lost in the reduction to the estimated manifold is lost and not used in alignment. MAGAN jointly learns the manifolds and aligns them, allowing each to inform the other. In (Wang & Mahadevan, 2009), they learn the manifolds and perform alignment jointly and do so without semi-supervised correspondences. They do this by matching the local geometry (thus requiring a meaningful distance metric in the original space), as opposed to MAGAN’s unsupervised correspondence loss which can be an arbitrary function defined over the original spaces.

#### 5. Conclusion

MAGAN discovers relationships between domains by aligning their manifolds rather than just superimposing them. Crucially, this can be used when one system is measured in two different ways and thus forms two different manifolds. In this case, the point in each manifold for one object in the underlying system are linked. This preserves information at a pointwise (rather than just population aggregate) level.

MAGAN facilitates integration of datasets from multiple biological modalities. As each type of experiment captures different information with different strengths and weaknesses, combining them makes possible discoveries that could not be found otherwise.

#### References

- Bendall, S. C., Nolan, G. P., Roederer, M., and Chattopadhyay, P. K. A deep profiler’s guide to cytometry. *Trends in immunology*, 33(7):323–332, 2012.
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija,



- R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature biotechnology*, 36(5):411, 2018.
- Capra, J. D., Janeway, C. A., Travers, P., and Walport, M. *Immunobiology: the immune system in health and disease*. Garland Publishing., 1999.
- Domingos, P. A few useful things to know about machine learning. *Communications of the ACM*, 55(10):78–87, 2012.
- Haghverdi, L., Lun, A. T., Morgan, M. D., and Marioni, J. C. Batch effects in single-cell rna-sequencing data are corrected by matching mutual nearest neighbors. *Nature biotechnology*, 2018.
- Ham, J., Lee, D. D., and Saul, L. K. Semisupervised alignment of manifolds. In *AISTATS*, pp. 120–127, 2005.
- Ham, J. H., Lee, D. D., and Saul, L. K. Learning high dimensional correspondences from low dimensional manifolds. 2003.
- Hinton, G. E., Dayan, P., and Revow, M. Modeling the manifolds of images of handwritten digits. *IEEE transactions on Neural Networks*, 8(1):65–74, 1997.
- Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. Image-to-image translation with conditional adversarial networks. *arXiv preprint arXiv:1611.07004*, 2016.
- Kim, T., Cha, M., Kim, H., Lee, J., and Kim, J. Learning to discover cross-domain relations with generative adversarial networks. *arXiv preprint arXiv:1703.05192*, 2017.
- Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., and Kirschner, M. W. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, 161(5):1187–1201, 2015.
- Li, C., Liu, H., Chen, C., Pu, Y., Chen, L., Henao, R., and Carin, L. Alice: Towards understanding adversarial learning for joint distribution matching. In *Advances in Neural Information Processing Systems*, pp. 5495–5503, 2017.
- Park, N., Anand, A., Moniz, J. R. A., Lee, K., Chakraborty, T., Choo, J., Park, H., and Kim, Y. Mmgan: Manifold matching generative adversarial network for generating images. *arXiv preprint arXiv:1707.08273*, 2017.
- Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. In *Advances in Neural Information Processing Systems*, pp. 2234–2242, 2016.
- Setty, M., Tadmor, M. D., Reich-Zeliger, S., Angel, O., Salame, T. M., Kathail, P., Choi, K., Bendall, S., Friedman, N., and Pe’er, D. Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nature biotechnology*, 34(6):637, 2016.
- van Dijk, D., Nainys, J., Sharma, R., Kathail, P., Carr, A. J., Moon, K. R., Mazutis, L., Wolf, G., Krishnaswamy, S., and Pe’er, D. Magic: A diffusion-based imputation method reveals gene-gene interactions in single-cell rna-sequencing data. *BioRxiv*, pp. 111591, 2017.
- Velten, L., Haas, S. F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B. P., Hirche, C., Lutz, C., Buss, E. C., Nowak, D., et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nature cell biology*, 19(4):271, 2017.
- Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P.-A. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, pp. 1096–1103. ACM, 2008.
- Wang, C. and Mahadevan, S. Manifold alignment using procrustes analysis. In *Proceedings of the 25th international conference on Machine learning*, pp. 1120–1127. ACM, 2008.
- Wang, C. and Mahadevan, S. Manifold alignment without correspondence. In *IJCAI*, volume 2, pp. 3, 2009.
- Yi, Z., Zhang, H., Tan, P., and Gong, M. Dualgan: Unsupervised dual learning for image-to-image translation. *arXiv preprint*, 2017.
- Zhu, J.-Y., Krähenbühl, P., Shechtman, E., and Efros, A. A. Generative visual manipulation on the natural image manifold. In *European Conference on Computer Vision*, pp. 597–613. Springer, 2016.
- Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint arXiv:1703.10593*, 2017.