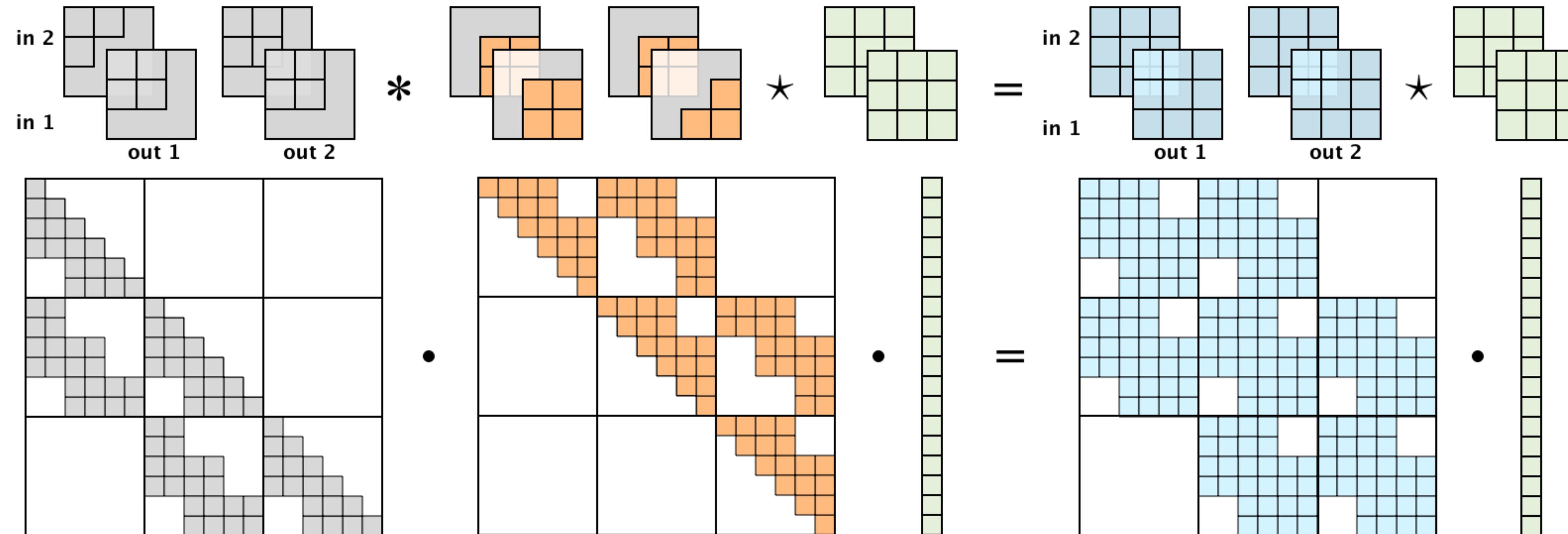


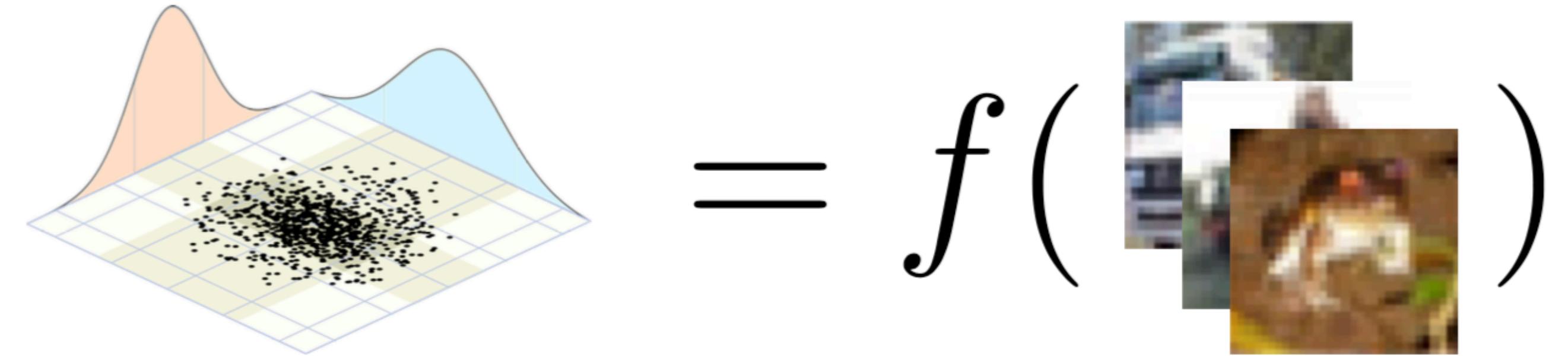
# Emerging Convolutions for Generative Normalizing Flows

by Emiel Hoogeboom, Rianne van den Verg, Max Welling



# Invertible functions

$$p_X(x) = p_Z(z) \left| \frac{dz}{dx} \right| ; z = f(x)$$



- The change of variable formula *holds*
- Admits exact log-likelihood optimization (opposed to VAEs, GANs)
- Fast sampling (opposed to PixelCNNs)

# Background

## Convolutions

# Background: Convolution as matrix multiplication

The diagram illustrates two ways to represent convolution:

Left side (Convolution): A 3x3 kernel  $w$  (labeled  $\star$ ) and a 3x3 feature map  $x$ . The result is a 3x3 output where each element is the sum of products of corresponding elements from  $w$  and  $x$ .

a	b	c
d	e	f
g	h	i

$\star$

0	1	2
3	4	5
6	7	8

$w$        $x$

Right side (Matrix Multiplication): A 3x3 kernel  $w$  and a 3x8 feature map  $\vec{x}$  (labeled  $\bullet$ ). The result is a 3x8 output where each row is the product of the kernel  $w$  and a column of the feature map  $\vec{x}$ .

e	f	h	i				
d	e	f	g	h	i		
	d	e		g	h		
b	c		e	f		h	i
a	b	c	d	e	f	g	h
	b	c		d	e		g
	b	c		e	f		
a	b	c	d	e	f		

$\bullet$

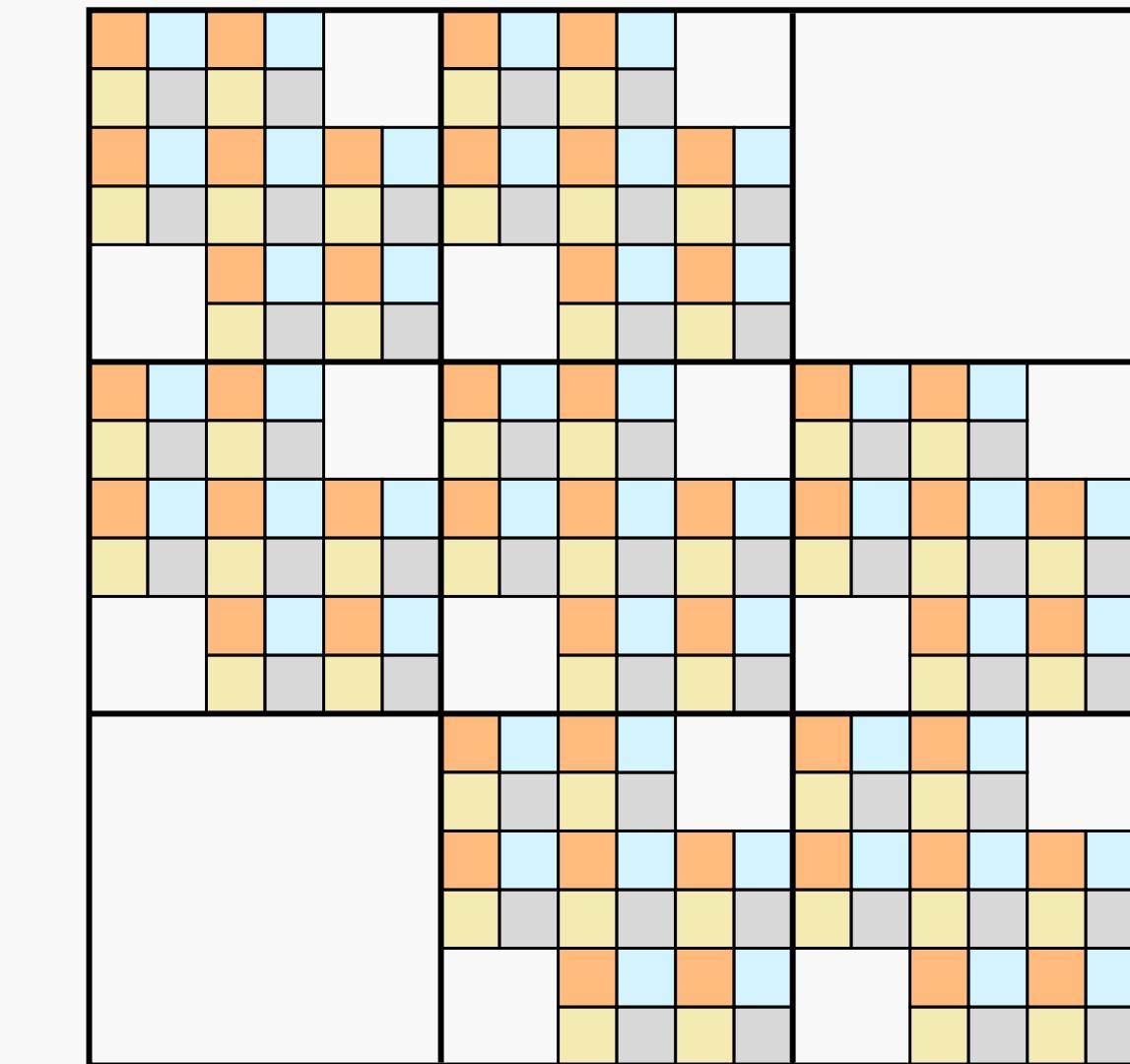
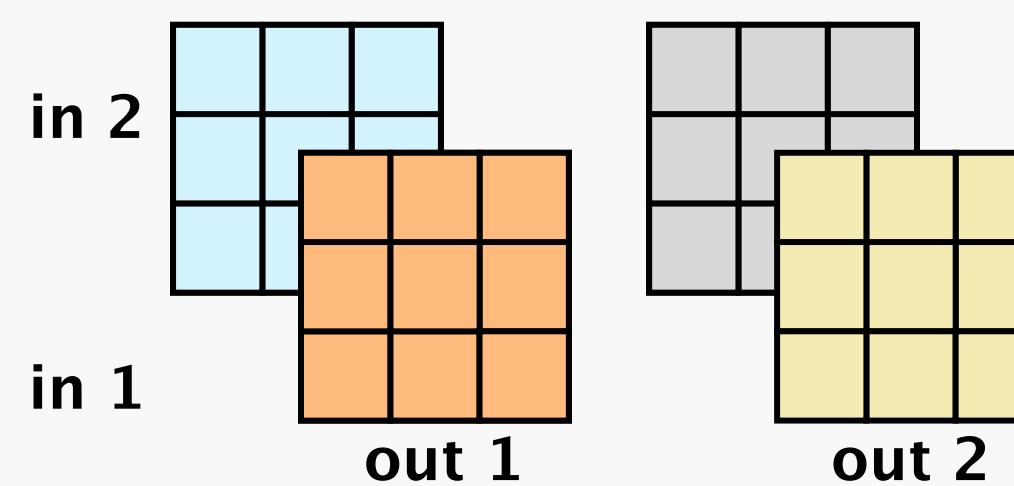
0
1
2
3
4
5
6
7
8

$w$        $\vec{x}$

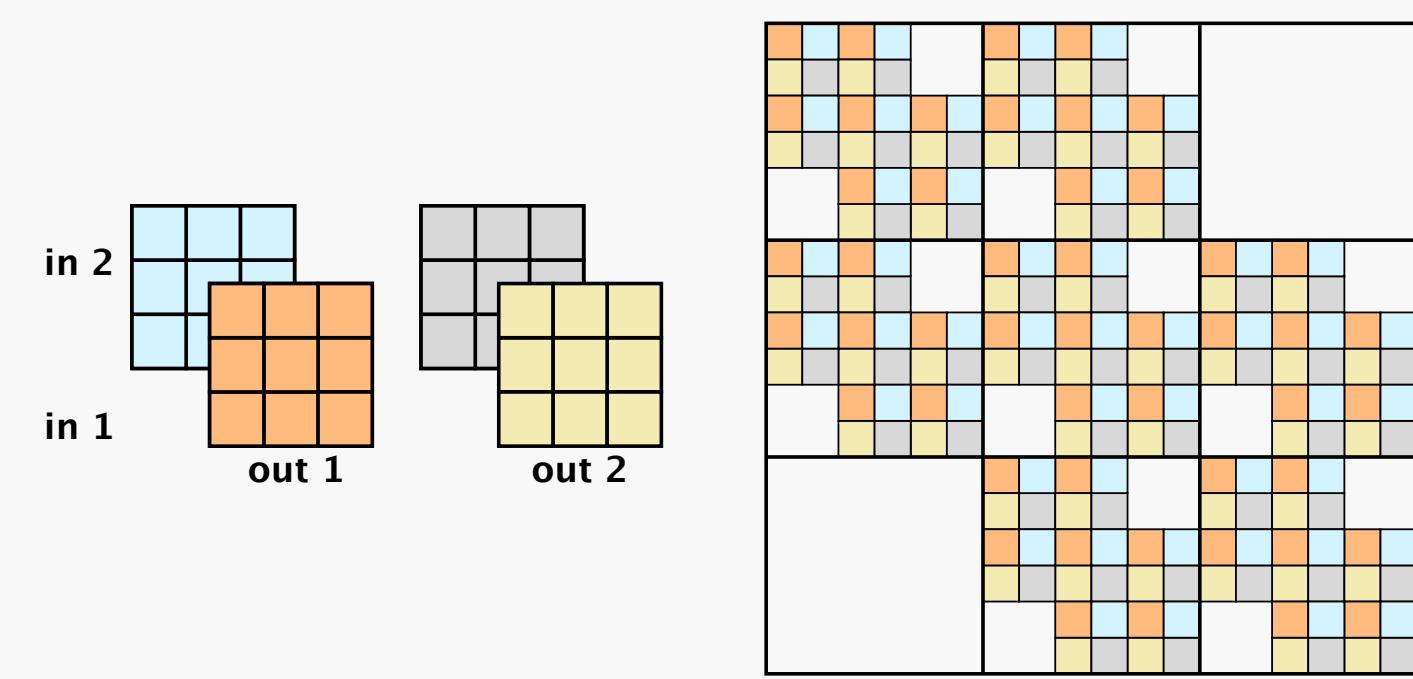
- Let  $w$  be a kernel, and  $x$  a feature map
- A convolution is equivalent to a matrix multiplication

# Convolution as matrix multiplication

$$\begin{array}{c}
 \begin{array}{cc}
 \begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix} \star \begin{matrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{matrix} & | \\
 \text{w} & \text{x} & \text{W} & \vec{\text{x}}
 \end{array} \\
 \begin{array}{c}
 \begin{matrix} e & f & h & i & & \\ d & e & f & g & h & i \\ & d & e & & g & h \\ b & c & & e & f & h & i \\ a & b & c & d & e & f & g & h & i \\ & b & c & & d & e & & g & h \\ & & b & c & & e & f & \\ a & b & c & d & e & f & g & h & i \\ & b & c & & d & e & & g & h \\ & & b & c & & e & f & \\ & & & a & b & c & d & e & f \\ & & & b & c & & d & e & \end{matrix} \\
 \bullet \quad \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix}
 \end{array}
 \end{array}$$

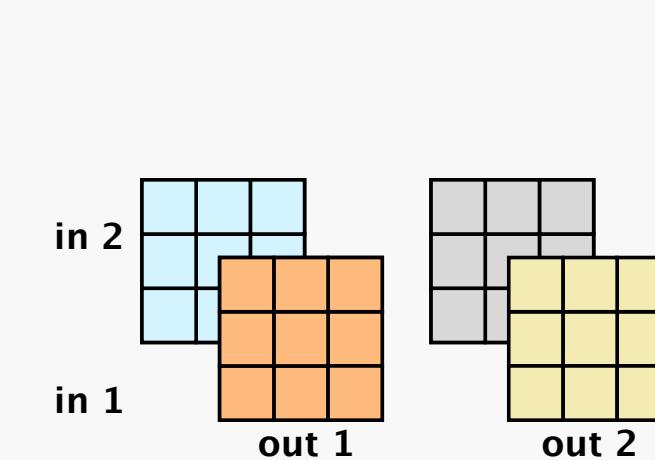


# Autoregressive Convolutions

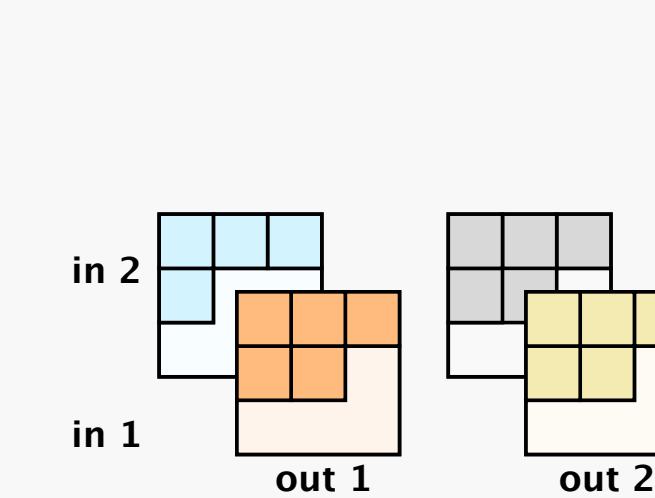
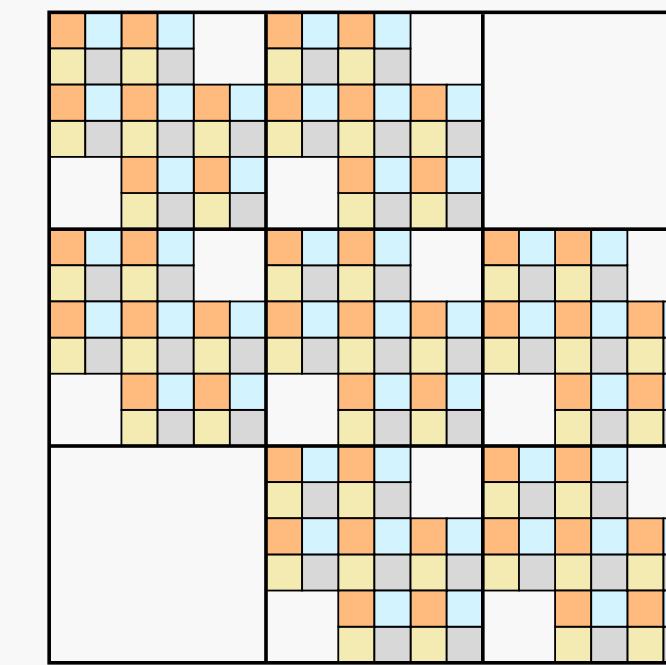


Standard convolution

# Autoregressive Convolutions



Standard convolution



Autoregressive convolution

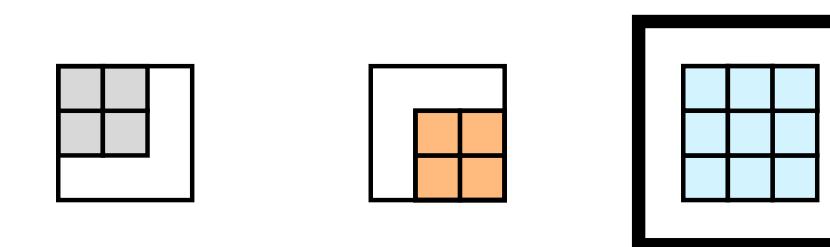
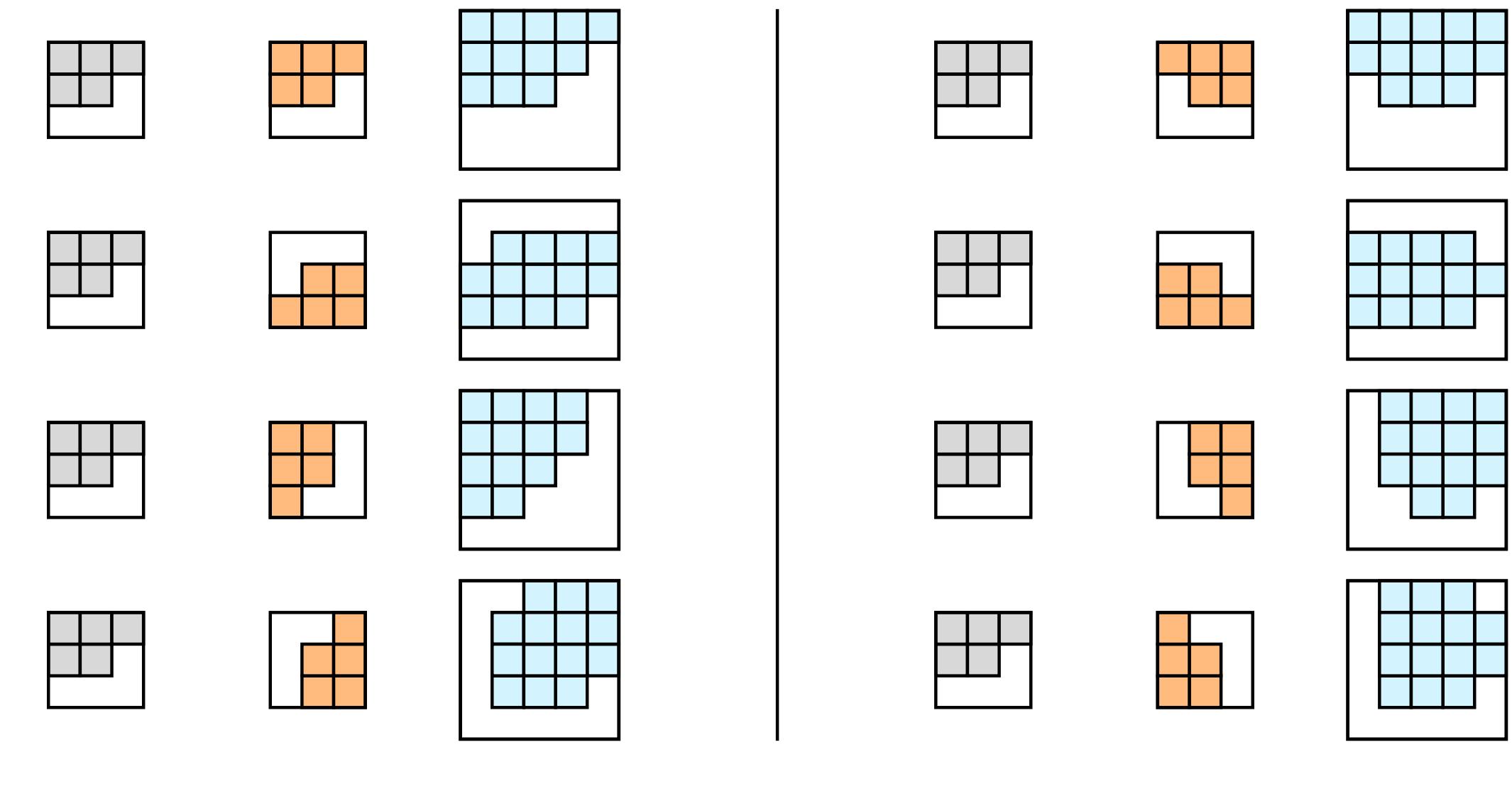
- Tractable Jacobian determinant
- Straightforward to invert

# Method

Emerging convolutions

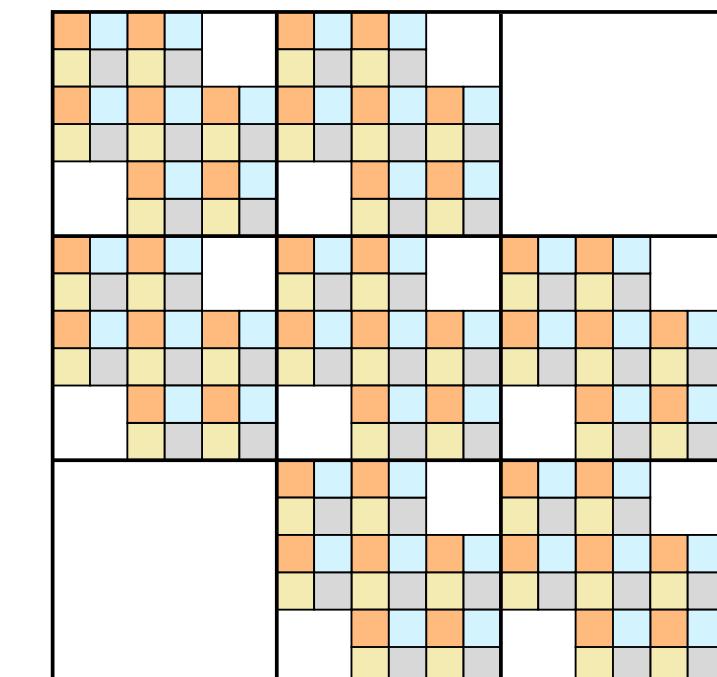
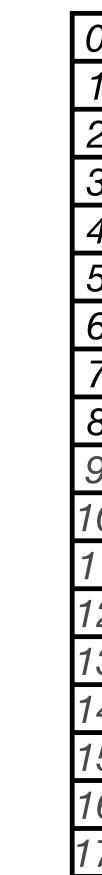
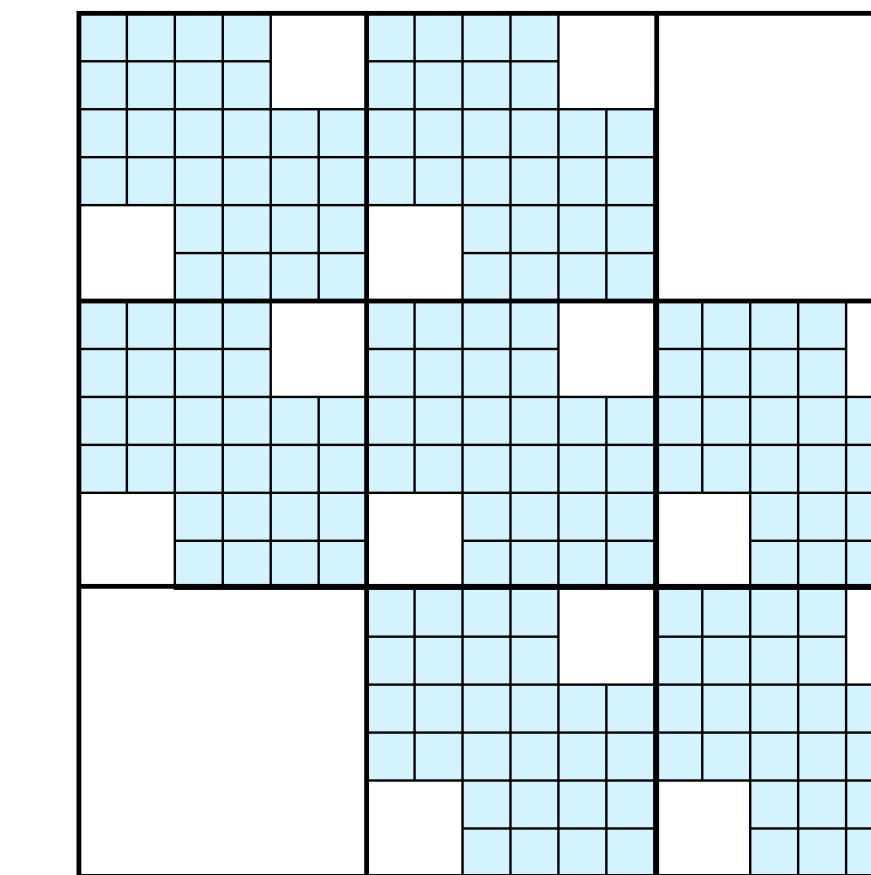
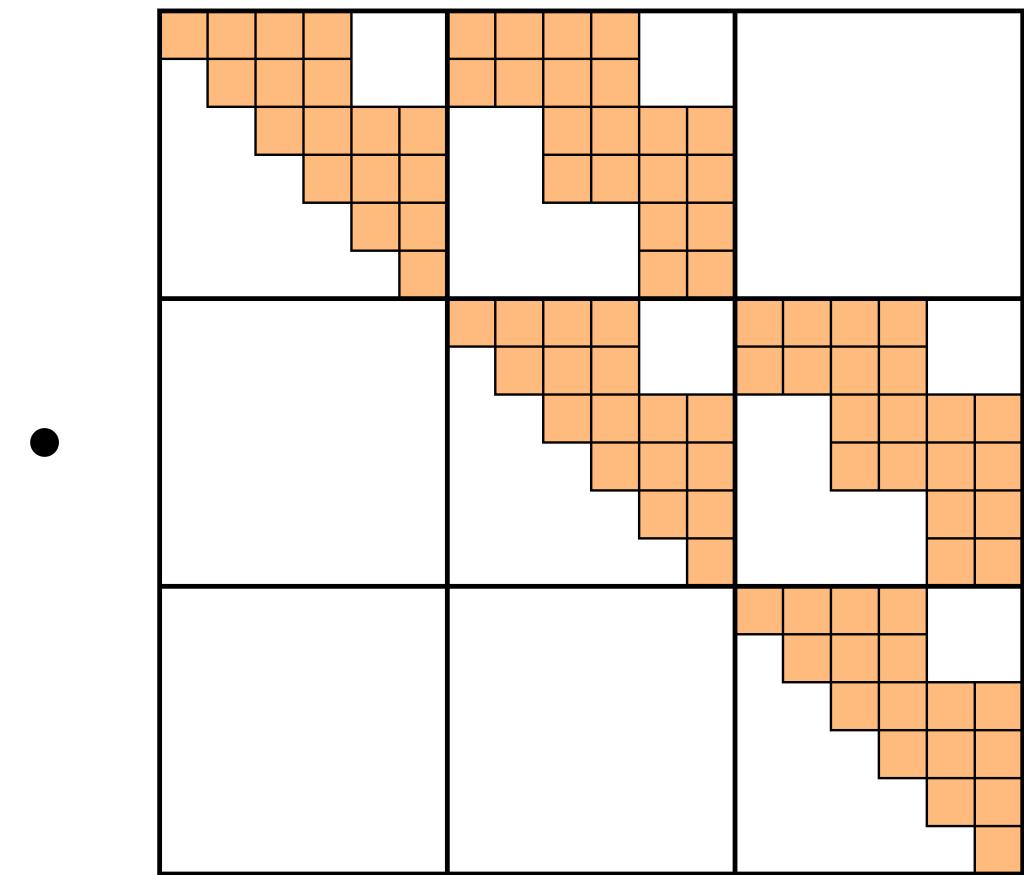
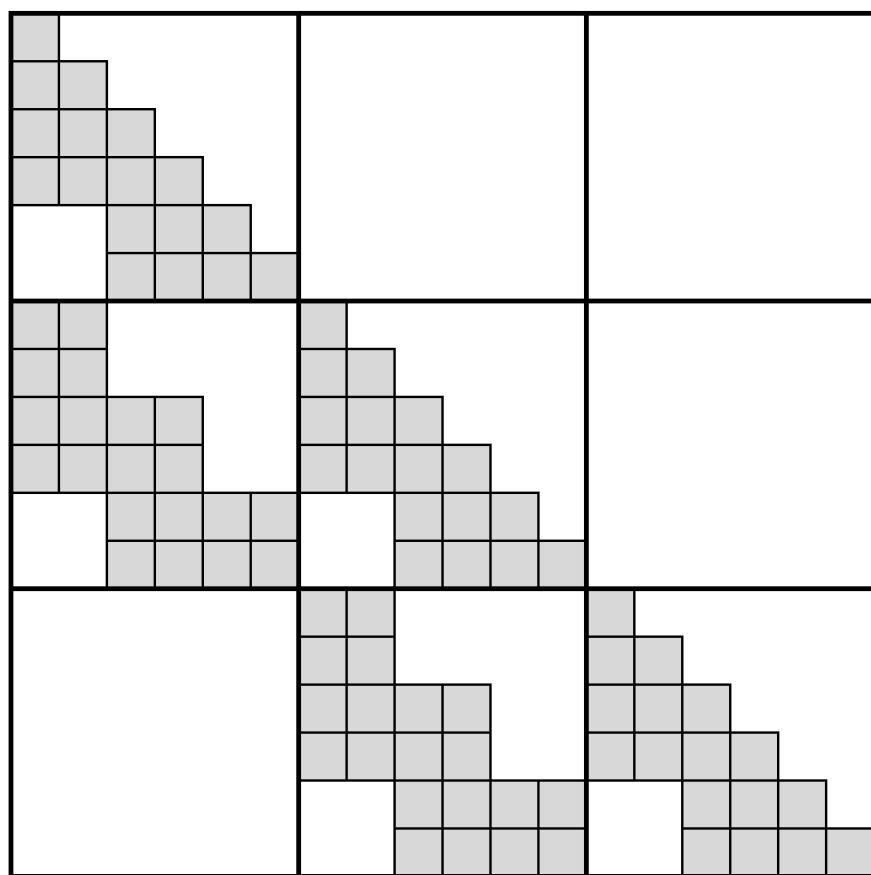
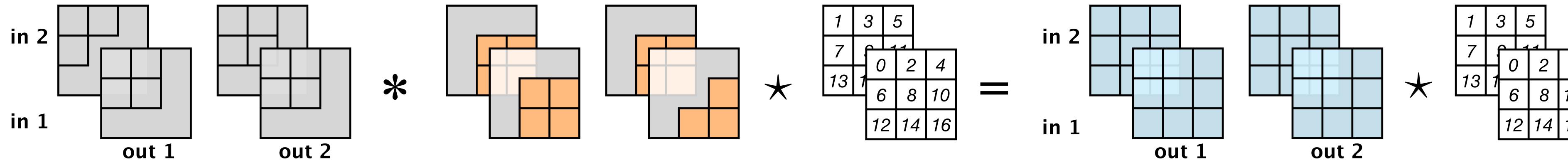
# Emerging Convolutions

- Combine autoregressive convolutions
- Special case: receptive field identical to standard convolutions



Receptive fields of emerging convolutions

# Emerging Convolutions



*Emerging convolution*

$$k_1 \star (k_2 \star f) = (k_1 * k_2) \star f$$

*Equivalent filter*

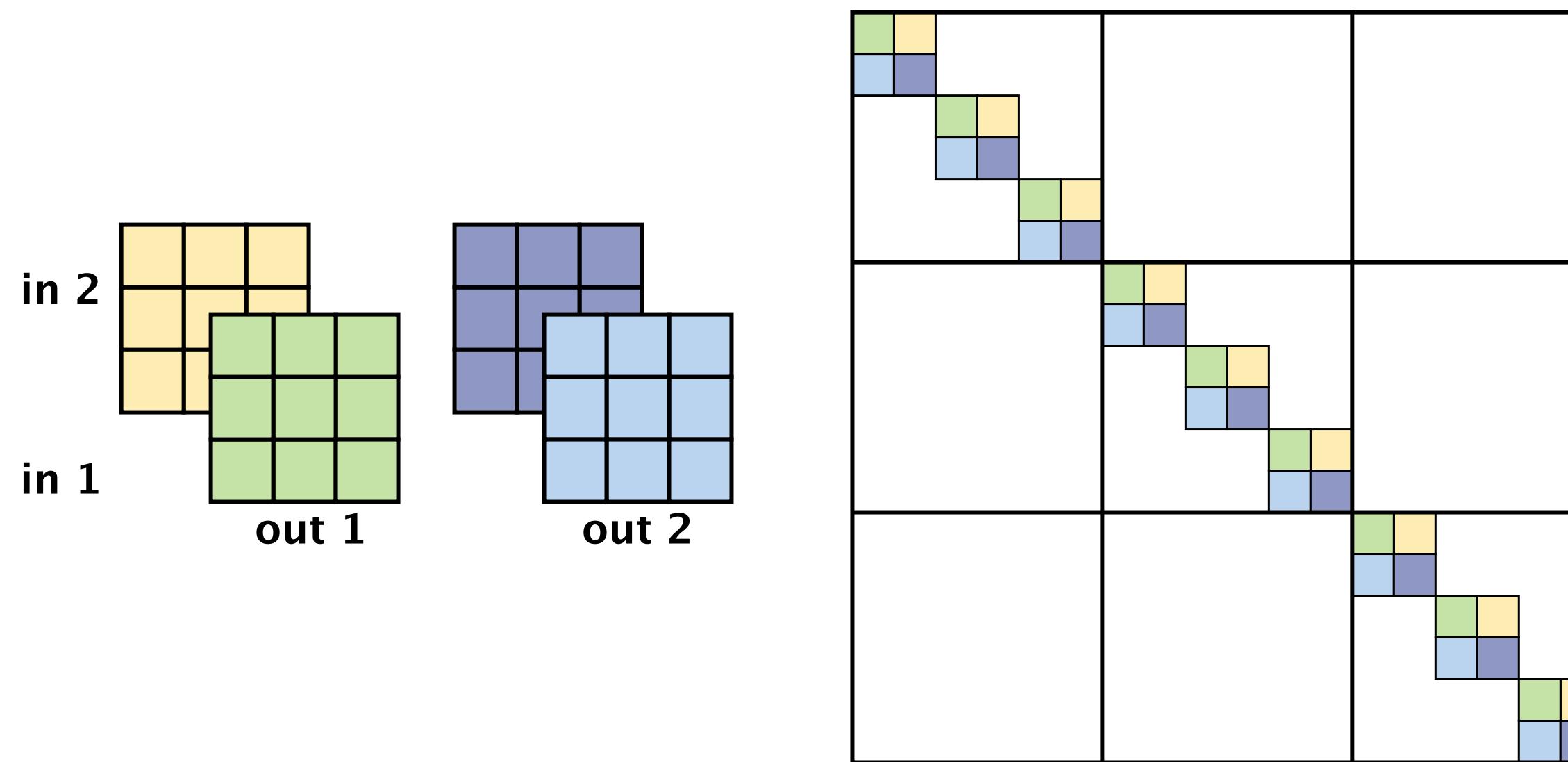
*Standard convolution*

# Method

Periodic convolutions

# Invertible Periodic Convolutions

- Leverages the *convolution theorem*
- The determinant and inverse are computed in *frequency domain*



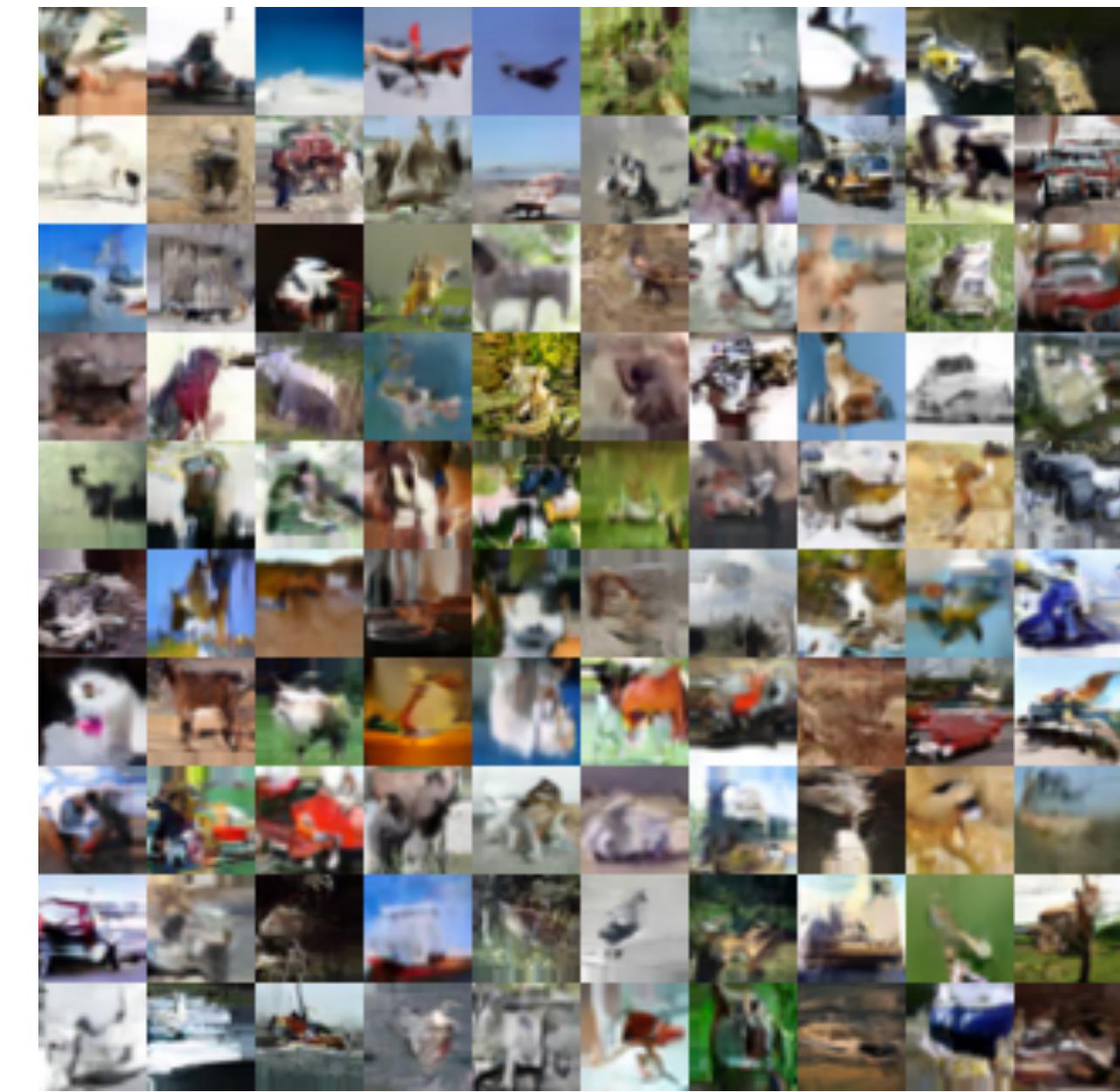
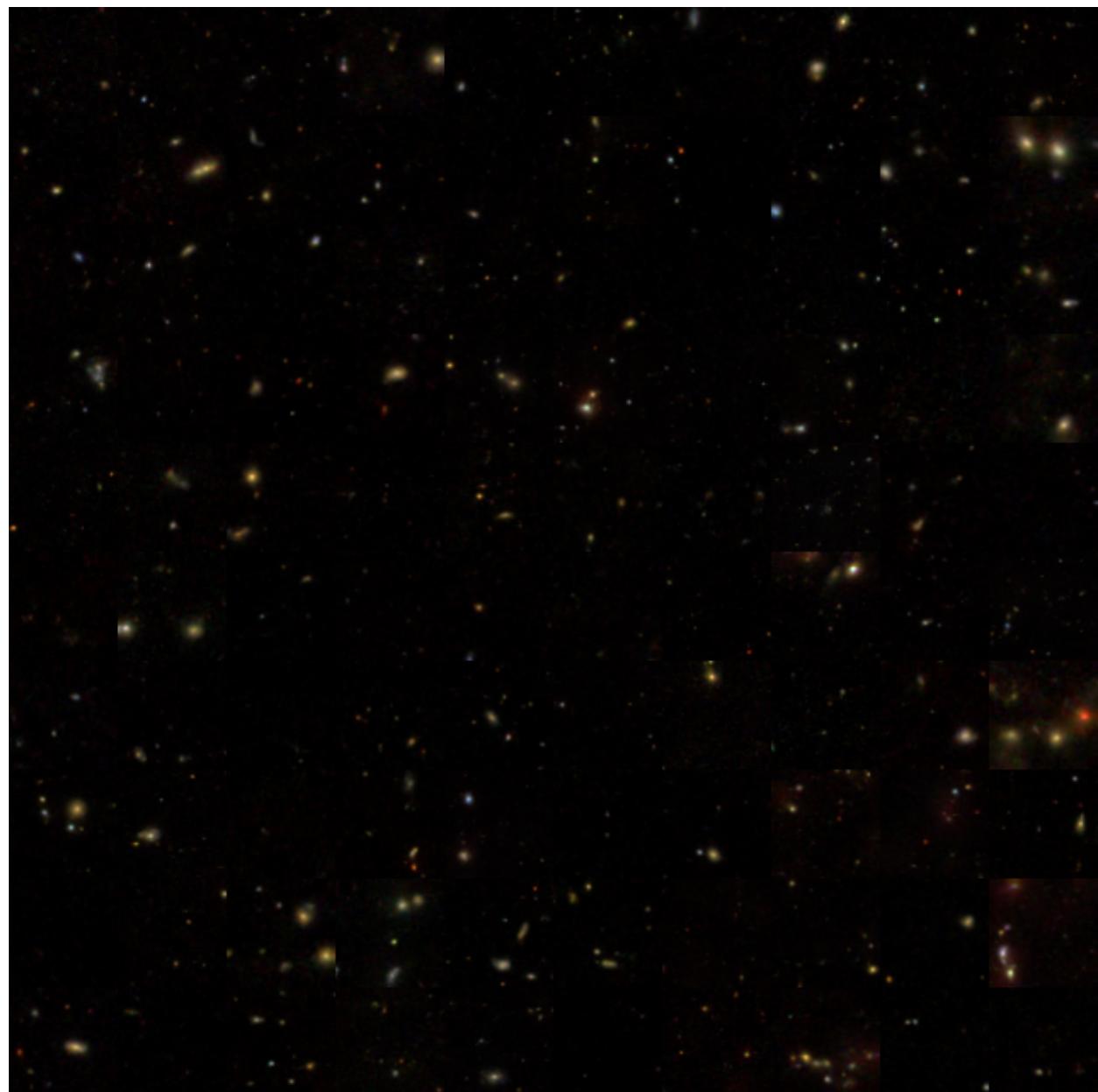
$$\hat{\mathbf{z}}_{:,uv} = \hat{\mathbf{W}}_{uv} \hat{\mathbf{x}}_{:,uv}$$

$$\hat{\mathbf{x}}_{:,uv} = \hat{\mathbf{W}}_{uv}^{-1} \hat{\mathbf{z}}_{:,uv}$$

$$\sum_{u,v} \log |\det \hat{\mathbf{W}}_{uv}|$$

# Conclusion

- Emerging convolutions
- Invertible periodic convolutions
- Stable, flexible  $1 \times 1$  QR convolutions
- Poster #8



Galaxy	
$1 \times 1$ (Glow)	$2.03 \pm 0.026$
Periodic $3 \times 3$	<b><math>1.98 \pm 0.003</math></b>
Emerging $3 \times 3$	<b><math>1.98 \pm 0.007</math></b>

	CIFAR10	ImageNet 32x32	D
$1 \times 1$ (Glow)	$3.46 \pm 0.005$	$4.18 \pm 0.003$	8
Emerging	<b><math>3.43 \pm 0.004</math></b>	<b><math>4.16 \pm 0.004</math></b>	8
$1 \times 1$ (Glow)	$3.56 \pm 0.008$	$4.28 \pm 0.008$	4
Emerging	<b><math>3.51 \pm 0.001</math></b>	<b><math>4.25 \pm 0.002</math></b>	4